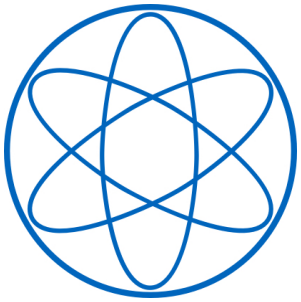
Lehrstuhl E15
Physik - Department



Data Analysis in Sterile Neutrino Searches

Dissertation
von

Birgit Marina Monika Neumair

Technische Universität München

Technische Universität München

Fakultät für Physik
Lehrstuhl für Experimentelle Astroteilchenphysik, E15
Prof. Dr. Lothar Oberauer

Data Analysis in Sterile Neutrino Searches

Birgit Marina Monika Neumair

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende/-r:	Prof. Dr. Alejandro Ibarra
Prüfende/-r der Dissertation:	1. Prof. Dr. Lothar Oberauer
	2. Prof. Dr. Susanne Mertens

Die Dissertation wurde am 08.02.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 16.03.2021 angenommen.

Abstract

In the last decades, several neutrino oscillation experiments reported results which are not compatible with the 3-neutrino model and hint at the existence of light sterile neutrinos. To test this hypothesis, a large number of short-baseline experiments is currently searching for oscillations from active to sterile neutrinos. The present work encompasses two aspects of the data analysis in the sterile neutrino search.

The first aspect is the data analysis of the SOX (Short distance neutrino Oscillations with BoreXino) experiment. SOX planned to place an artificial ^{144}Ce-$\bar{\nu}_e$ source underneath the liquid scintillator detector Borexino, which is located at the LNGS underground laboratory. The sterile neutrino signature would be the reduction of the detected antineutrino interaction rate with an oscillatory pattern as a function of the neutrino energy and traveled distance. The data selection of the antineutrinos was optimized to increase the expected sensitivity, taking into account the expected signal, background, and systematic effects. A detection efficiency of $(93.0 \pm 0.5)\,\%$ was achieved for antineutrino signals, which is an improvement of $10.8\,\%$ in comparison to the previously-performed antineutrino analysis in Borexino. Analyzing Borexino data, the expected background was examined to $28.2^{+2.7}_{-4.2}$ events, resulting to a signal-to-background ratio of ~ 200 for a source activity of 100 kCi. SOX is sensitive to oscillation amplitudes $\sin^2(2\theta) \gtrsim 0.06$ for mass splittings $\Delta m^2 = 1\,\mathrm{eV}^2$ ($\sin^2(2\theta) \gtrsim 0.09$ for $\Delta m^2 > 10\,\mathrm{eV}^2$) at the 95 % CL, and can hence probe a large fraction of the preferred parameter space of light sterile neutrinos. To further improve the sensitivity, a hardware modification of the trigger system of Borexino was tested.

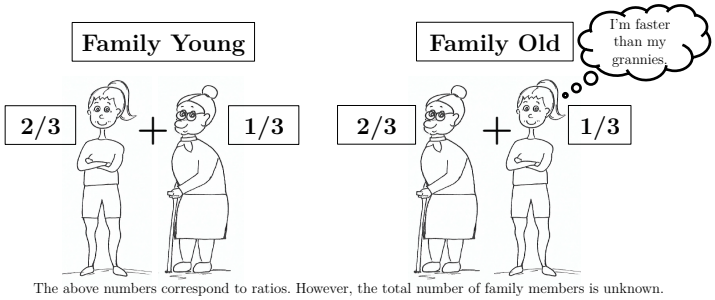
The second part analyzes the non-trivial statistical issues and the different applied analysis methods of the current short-baseline experiments. Experiments in which a neutrino oscillation would enhance or reduce the neutrino interaction rate in the detector are both studied using toy models. The dependency on the interaction rate and the neutrino energy spectrum for exclusion and discovery scenarios are tested. The most-likely hypothesis for a specific data set is typically given by a sterile neutrino signal, meaning that the maximum-likelihood estimators are biased. The sensitivity and reconstructed confidence regions for a specific data set change significantly for the different methods, where the discrepancy is especially pronounced for a signal discovery. The differences could be traced back to the choice of the alternative hypothesis in the test statistic and the assumptions on the asymptotic probability distributions. In particular, Wilks' theorem was found not to be valid and the probability distributions of the test statistic need to be constructed via Monte Carlo techniques. As a consequence, the significance of a signal is overestimated when Wilks' theorem is applied. The standardized analysis based on the most general alternative hypothesis proposed in the frame of this work has found a large consensus in the sterile neutrino field and has been endorsed by many experiments.

Zusammenfassung

Mehrere Neutrinooszillationsexperimente erzielten in den letzten Jahrzehnten Ergebnisse, welche von dem 3-Neutrino-Modell abweichen und auf die Existenz leichter steriler Neutrinos hinweisen. Um diese Hypothese zu testen, suchen zur Zeit eine Vielzahl an Experimenten nach Oszillationen von aktiven zu sterilen Neutrinos auf kurzen Entfernungen. Die vorliegende Arbeit behandelt zwei Themen zur Datenanalyse für die Suche nach sterilen Neutrinos.
Der erste Teil behandelt die Datenanalyse des SOX (Short distance neutrino Oscillations with BoreXino) Experiments. SOX plante eine künstliche ^{144}Ce-$\bar{\nu}_e$ Quelle unterhalb des Borexino Detektors, welcher sich im LNGS Untergrundlabor befindet, zu platzieren. Die Signatur eines sterilen Neutrinos wäre die Reduktion der detektierten Antineutrinointeraktionsrate mit einem Oszillationsmuster in Abhängigkeit von Energie und Flugdistanz der Neutrinos. Die Selektion der Antineutrinoereignisse wurde optimiert und somit die Sensitivität des Experiments erhöht. Hierbei wurde das erwartete Neutrinosignal, Untergrundereignisse und systematische Effekte miteinbezogen. Eine Detektionseffizienz von $(93.0 \pm 0.5)\,\%$ wurde für Antineutrinosignale erreicht, welches einer Verbesserung um $10.8\,\%$ im Vergleich zur vorherigen durchgeführten Antineutrinoanalyse in Borexino entspricht. Durch eine Analyse von Borexino-Daten wurde eine Untergrunderwartung von $28.2^{+2.7}_{-4.2}$ Ereignissen berechnet, welches bei einer Aktivität von 100 kCi zu einem Signal-zu-Untergrund Verhältnis von ~ 200 führt. SOX is sensitiv auf Oszillationsamplituden $\sin^2(2\theta) \gtrsim 0.06$ für Massenunterschiede $\Delta m^2 = 1\,\mathrm{eV}^2$ ($\sin^2(2\theta) \gtrsim 0.09$ für $\Delta m^2 > 10\,\mathrm{eV}^2$) bei einem $95\,\%$ Konfidenzniveau und kann damit einen Großteil des bevorzugten Parameterbereichs steriler Neutrinos überprüfen. Um die Sensitivität weiter zu verbessern, wurde eine Modifikation des Triggersystems von Borexino getestet.
Im zweiten Teil werden die nicht-trivialen statistischen Aspekte und die unterschiedlichen Analysemethoden der aktuellen Experimente zur Suche nach sterilen Neutrinos analysiert. Experimente, in denen Neutrinooszillationen die Interaktionsrate im Detektor erhöhen oder reduzieren würden, werden beide mit Hilfe von Monte Carlo-Modellen studiert. Die Abhängigkeit der Interaktionsrate und des Neutrinoenergiespektrums für Ausschluss- und Entdeckungsszenarien werden getestet. Die wahrscheinlichste Hypothese für einen spezifischen Datensatz ist üblicherweise durch ein Oszillationssignal gegeben. Dies bedeutet, dass die Maximum-Likelihood-Schätzwerte verzerrt sind. Die Sensitivität und die rekonstruierten Konfidenzregionen für einen speziellen Datensatz verändern sich signifikant für die unterschiedlichen Analysemethoden. Hierbei ist die Diskrepanz besonders ausgeprägt, falls ein Signal entdeckt wird. Die Unterschiede konnten auf die Wahl der alternativen Hypothese in der Teststatistik und der Annahme der asymptotischen Wahrscheinlichkeitsverteilungen zurückgeführt werden. Da das Wilks-Theorem seine Gültigkeit verliert und dessen Annahme die Signifikanz von Signalen überschätzen würde, müssen die Wahrscheinlichkeitsverteilungen der Teststatistiken mit Monte Carlo-Techniken erzeugt werden. Die Standardanalyse, welche auf der allgemeinsten alternativen Hypothese basiert und im Rahmen dieser Arbeit vorgeschlagen wurde, wurde von vielen Experimenten übernommen und stellt den derzeitigen Konsens in der sterilen Neutrinosuche dar.

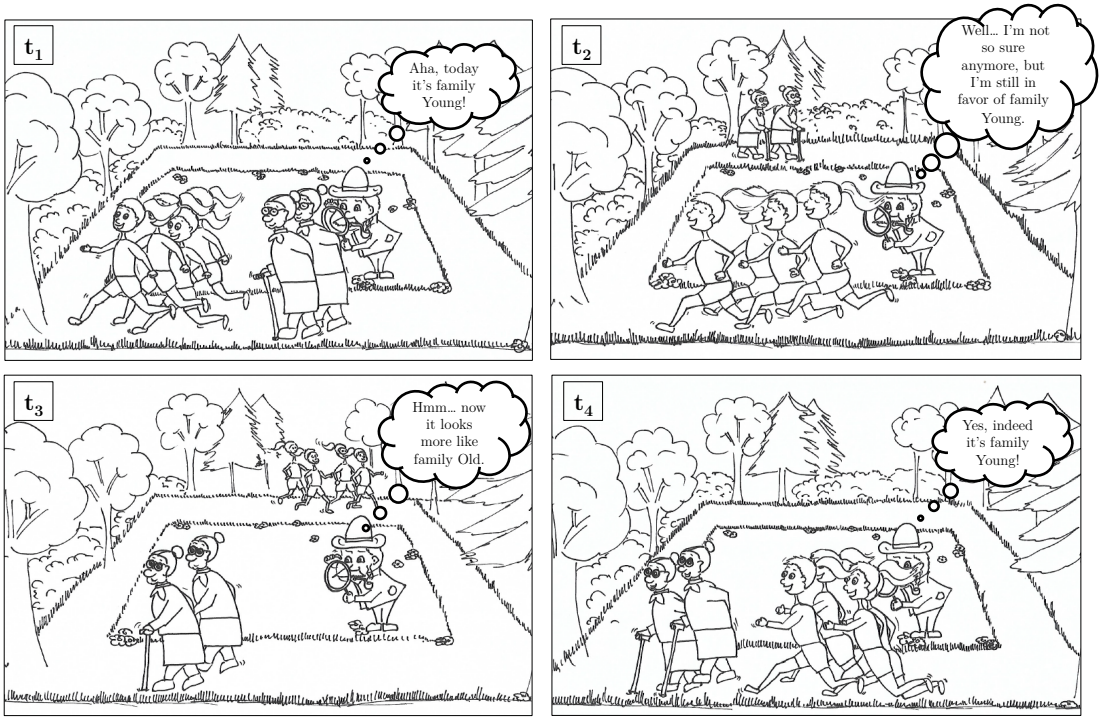
Contents

Family Oscillations

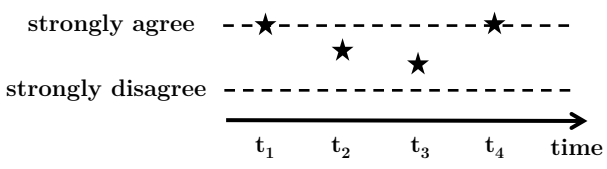


The above numbers correspond to ratios. However, the total number of family members is unknown.

Both families like to go for a walk in the park. Today family Young is in the park, but the observer does not know this. Nevertheless, he likes to guess the family. However, as the observer's field of view is limited, his guess depends on time.



Q: Observer, do you think it's Family Young?



Similarly, the three known neutrino families consist of three members with different phase velocities. There exists therefore a chance, which depends on the traveled time, that the initial neutrino family is observed as a different neutrino family. This phenomenon is called **neutrino oscillations**.

1 Neutrinos in the Standard Model and beyond

Neutrinos are elementary, almost massless, and only weakly interacting particles. This enables neutrinos to fly straight and, thus, to point back to their sources. Hence, with the measurement of neutrinos the properties of their sources can be studied. In addition, given a well-known neutrino source, the properties of the neutrinos themselves can be investigated. The latter approach is the one used for the search of light sterile neutrinos which is the subject of this thesis. Sterile neutrinos are hypothetical particles beyond the standard model of particle physics (Section 1.1) and are introduced in Section 1.5. The search for light sterile neutrinos (Section 1.5.5) is especially motivated by several experimental hints which are described in Section 1.5.3. Before that, an overview of the neutrino properties including neutrino oscillations, their sources, and the current status in neutrino physics is given in Sections 1.1-1.4.

1.1 Neutrinos in the Standard Model

The standard model (SM) of particle physics describes the theory of strong, electromagnetic, and weak interactions [1]. The discovery of the Higgs boson [2, 3] completed the list of predicted elementary particles and confirmed the well-established SM yet another time.
Among these elementary particles neutrinos are the only fermions that can only interact weakly. This implies they carry only the weak charge (weak isospin) and are colorless (not coupling to the strong force), chargeless (not coupling to the electromagnetic force), and massless (not coupling to the gravitational force). However, the latter has been disproven by the observation of neutrino oscillations which will be discussed in the next section.
The weak interaction is mediated via the massive gauge bosons Z^0 ($m_Z = 92\,\mathrm{GeV}$) and $W^\pm$ ($m_W = 80\,\mathrm{GeV}$) leading to neutral current (NC) and charged current (CC) interactions, respectively [1]. Since the decay width of the Z^0-boson depends on the possible decay channels and hence on the number of weakly interacting neutrinos, the number of so-called active neutrinos could be determined in e^-e^+ collisions at the large electron-positron collider (LEP) [4] to

$$N_\nu = 2.984 \pm 0.008. \tag{1.1}$$

From this measurement follows that every additional neutrino must have either a mass larger than $m_Z/2$ or must not couple to the Z^0-boson, i.e. be sterile.
These three neutrinos and antineutrinos (their antiparticles) appear in three families (flavors) named after their corresponding charged lepton partner: $\bar{\nu}_e$, $\bar{\nu}_\mu$, and $\bar{\nu}_\tau$. Each lepton family carries its own lepton family number that is conserved in the weak interaction. Additionally, the total lepton number is conserved given by the sum of the individual lepton family numbers.

The weak interaction is special in the sense that only left-handed particles and right-handed antiparticles can take part in it. The handiness describes the chirality of the fermions which is connected to the weak isospin I. Left-handed and right-handed fermions carry a weak isospin of $I = 1/2$ and $I = 0$, respectively. This maximal-parity violating theory was introduced after the actual observation of parity violation of the weak interaction by Wu et al. [5]. Since neutrinos were thought to be massless[1] at this time, one assumed that neutrinos have a definite helicity which then equals the chirality. Finally, the helicity of neutrinos was determined by the Goldhaber experiment to be left-handed [6]. Therefore, there are no right-handed neutrinos and no left-handed antineutrinos defined in the SM.
The weak and electromagnetic forces can be unified and described in the electroweak theory with a $\mathrm{SU(2)_L} \times \mathrm{U(1)_R}$ gauge group. The left-handed leptons form lepton doublets of the weak isospin. The right-handed components of the charged leptons form singlets which carry no weak isospin and are hence not taking part in the weak interaction [7]:

$$\begin{array}{lll} \text{lepton doublets} & \begin{pmatrix}\nu_e \\ e\end{pmatrix}_{\mathrm{L}} \begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}_{\mathrm{L}} \begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}_{\mathrm{L}} & I = 1/2 \\ \text{charged lepton singlets} & e_{\mathrm{R}} \quad \mu_{\mathrm{R}} \quad \tau_{\mathrm{R}} & I = 0 \end{array}$$

Although the SM is very successful, there is evidence for physics beyond the SM. The observed neutrino oscillations are direct evidence for at least two neutrino masses being different from zero. In addition, the lepton family number is not conserved during neutrino propagation. The next section will discuss possible extensions of the SM to generate neutrino masses and the phenomenology of neutrino oscillations.

1.2 Massive Neutrinos and Neutrino Oscillations

The masses of all elementary particles in the SM are produced via the Higgs mechanism [8]. However, in order to produce the masses of the neutrino via the Higgs mechanism, right-handed neutrinos are required that are currently not present in the SM. To obtain right-handed neutrinos, the SM needs to be extended by either right-handed neutrinos and, thus, new particles. On the contrary, the SM can be extended by total lepton number violation which allows that neutrinos are Majorana particles, i.e. neutrinos are their own antiparticles. In the latter case, right-handed neutrinos can then be obtained through the charge-conjugation transformation of the left-handed Majorana neutrinos.
The neutrino masses can then be generated via a Dirac or a Majorana mass term in the Lagrangian [9]. In both scenarios, the neutrino masses are produced when the flavor eigenstates couple to the Higgs field. Hence, the mass terms are expressed in the flavor eigenstates that generate a complex symmetric mass matrix M. To obtain an expression of the Lagrangian in the mass eigenstates and the neutrino masses, the mass matrix is diagonalized via a unitary transformation which is sketched in Figure 1.1. Thus, the weakly interacting flavor eigenstates ν_e, ν_μ, ν_τ are a superposition of the mass eigenstates ν_1, ν_2, ν_3 [1]:

$$\begin{pmatrix}\nu_e \\ \nu_\mu \\ \nu_\tau\end{pmatrix} = U \begin{pmatrix}\nu_1 \\ \nu_2 \\ \nu_3\end{pmatrix} \tag{1.2}$$

The mass eigenstates ν_1, ν_2, ν_3 have the respective mass eigenvalues m_1, m_2, m_3 and describe the propagation of the neutrinos in vacuum. The superposition of the mass eigenstates is

[1] No neutrino masses have been measured so far and the limits were about a million times smaller than the mass of the lightest other particle, the electron.

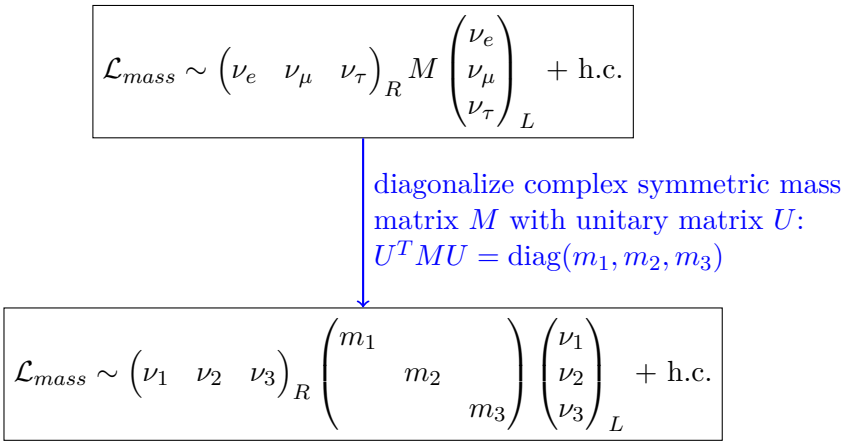


Figure 1.1: Simplified Scheme of the Neutrino Mass Generation for Three Neutrinos The top box shows the Lagrangian mass term $\mathcal{L}_{mass}$ as a function of the flavor neutrino fields ν_α ($\alpha = e, \mu, \tau$). The left(L)- and right(R)-handed flavor neutrino fields couple to the Higgs field and generate the symmetric mass matrix M. The right-handed neutrinos can be either the new (currently not in the SM) right-handed fields of Dirac neutrinos or the charge-conjugated field of Majorana neutrinos. Through a unitary transformation U the mass matrix can be diagonalized and the neutrino mass eigenstates ν_i ($i = 1, 2, 3$) with the respective neutrino masses m_i ($i = 1, 2, 3$) are obtained. The Lagrangian mass term is then expressed in the massive neutrino fields (bottom box). Note that the notation in this illustration is extremely simplified. An exact expression of the Lagrangian mass term can be found for example in Reference [9].

described by the unitary matrix that diagonalizes the mass term in the Lagrangian and is called PMNS matrix [1]

$$\begin{aligned} U &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{1.3}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ with the rotation angle θ_{ij} and $i, j = 1, 2, 3$. δ denotes the Dirac CP-violating phase which is still unknown. In case that the neutrino is a Majorana particle, the PMNS matrix contains two further CP-violating Majorana phases.
Given different neutrino masses m_i, the wave functions of the mass eigenstates with the same momentum propagate with different frequencies. The initial mass composition of one neutrino flavor changes as the neutrino is moving through space and can lead to constructive or destructive interference of the wave functions. Therefore, there is a probability that a neutrino with its initial flavor α can change to another flavor β while propagating ($\alpha, \beta = e, \mu, \tau$). This phenomenon is called neutrino oscillations. The oscillation probability is a function of the energy E and the traveled distance L of the neutrino and can often be approximated by the two-flavor formalism [1]:

$$P(\nu_\alpha \to \nu_\beta) \approx \left| \delta_{\alpha\beta} - \sin^2(2\theta_{ij}) \cdot \sin^2\left(\frac{1.27 \cdot \Delta m^2_{ij}[\text{eV}^2] \cdot L[\text{m}]}{E[\text{MeV}]} \right) \right|. \tag{1.4}$$

Here, the mixing angles θ_{ij} ($i, j = 1, 2, 3$) are connected to the entries of the PMNS matrix

via

$$\sin^2(\theta_{12}) = \frac{|U_{e2}|^2}{1-|U_{e3}|^2}, \quad \sin^2(\theta_{13}) = |U_{e3}|^2, \quad \sin^2(\theta_{23}) = \frac{|U_{\mu 3}|^2}{1-|U_{e3}|^2}, \tag{1.5}$$

and $\Delta m^2_{ij} = m^2_i - m^2_j$ denotes the mass splittings. Neutrino oscillations can only be observed, if the mixing angles and the mass splittings are different from zero. The term $\sin^2(2\theta_{ij})$ in Equation 1.4 represents the amplitude of the oscillation signal and Δm^2_{ij} the frequency of the oscillation. Larger mass splittings result in larger oscillation frequencies and hence in shorter distances between two oscillation maxima. This can be described by the oscillation length [1]

$$(L[\mathrm{m}]/E[\mathrm{MeV}])_{\mathrm{osc}} \approx 2.48/\Delta m^2[\mathrm{eV}^2]. \tag{1.6}$$

It has to be noted that neutrino oscillations can change in matter. This is described by the so-called Mikheyev Smirnov Wolfenstein (MSW) effect [10–12]. Coherent forward scattering off particles in matter results in an effective mass term that adds to the already existing one in the Lagrangian. Given relative differences in the mass terms for different neutrino flavors, the diagonalization matrix changes and can alter the neutrino oscillation probabilities.
The MSW effect is for example visible in the survival probability of solar neutrinos, where the oscillation probability is enhanced above $\sim 3\,\mathrm{MeV}$ compared to the one in vacuum (see Figure 2.5). The above-mentioned relative difference in the mass terms is introduced, as electron neutrinos can interact with electrons in the Sun not only via NC (as muon and tau neutrinos do) but also via CC reactions. From the observation of the MSW effect for solar neutrinos one concludes that m_2 is larger than m_1.

1.3 Neutrino Sources

An overview of the neutrino flux at Earth as a function of the energy for the individual neutrino sources is shown in Figure 1.2. With the recent observation of high-energy neutrinos from the blazar[2] TXS 0506+056 [15, 16], the third astrophysical source of neutrinos (in addition to the Sun and a core-collapse supernova) has been experimentally identified. Besides that, geo neutrinos produced in the Earth and atmospheric neutrinos, which are secondary particles produced from cosmic rays in the atmosphere, have been observed. Small neutrino fluxes and/or too small/high neutrino energies prevent so far from the observation of the residual natural neutrino sources, as for example the cosmic neutrino background (CNB)[3], the diffuse supernova background (DSNB), and cosmogenic neutrinos. A diffuse flux of high-energy neutrinos ($\sim 10^{15}\,\mathrm{eV}$) has been observed by the IceCube detector. However, their origin has not yet been identified [18]. A recent review of the neutrino flux on Earth from different sources can be found in Reference [13]. In the following, the natural and artificial neutrino sources that can be observed with the Borexino detector are discussed in more detail.

1.3.1 Natural Sources

Solar Neutrinos

Solar neutrinos are the most abundant neutrino source on Earth which can also be observed. When solar neutrinos are measured (for the experimental results with Borexino see Chapter 2), the standard solar model (SSM) and the metallicity of the Sun can be tested [19, 20].

[2] A blazar is a class of Active Galactic Nuclei (AGN) with a jet pointing towards the Earth [14].

[3] Cosmic neutrinos are relic particles from the Big Bang and were indirectly observed by cosmological observations as the cosmic microwave background and the big bang nucleosynthesis [17].

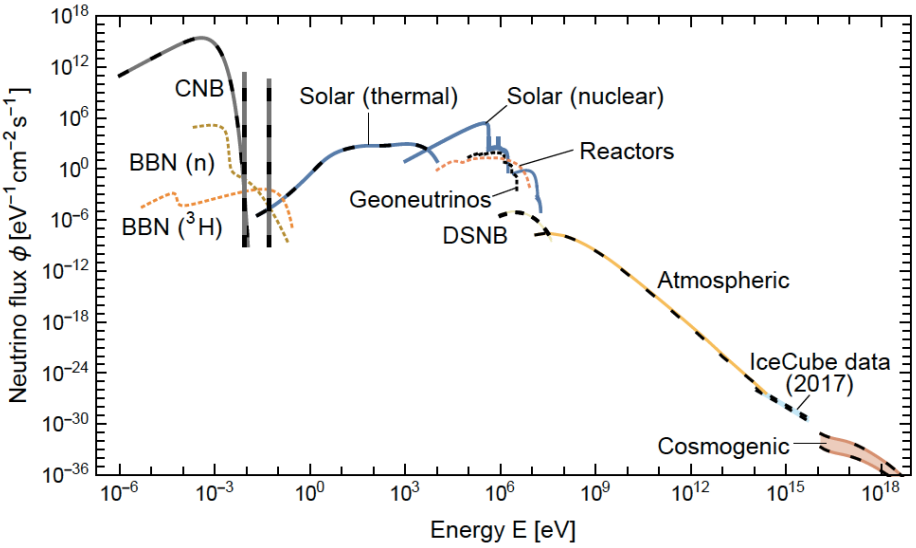


Figure 1.2: Neutrino Flux at Earth The neutrino flux is integrated over directions and summed over flavors. The unit of the flux for mono energetic lines is given in $\mathrm{cm}^{-1}\mathrm{s}^{-1}$. A solid (dashed/dotted) line represents neutrinos (antineutrinos). The cosmic neutrino background (CNB) is shown for the assumed neutrino masses $m_1 = 0$, $m_2 = 8.6\,\mathrm{meV}$, and $m_3 = 50\,\mathrm{meV}$, where the two monochromatic lines correspond to m_2 and m_3. Big bang nucleosynthesis (BBN) neutrinos are shown for neutron and tritium decay. Only the neutrino sources with a sufficiently large flux in the energy range of $(10^6 - 10^{15})\,\mathrm{eV}$ could be directly measured so far. For a detailed discussion of the individual components, see the figure's Reference [13].

In addition, as the Sun is a relatively well-defined neutrino source, one can gain information on the neutrino propagation (neutrino oscillations including matter effects) [21].
Solar (nuclear) neutrinos are electron neutrinos and are produced in fusion reactions in the Sun, where the net reaction is

$$4p \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e. \tag{1.7}$$

In this process an energy of 26.73 MeV is released and the average energy value that is taken away by the two neutrinos is only $\sim 0.6\,\mathrm{MeV}$ [21]. A simple estimation of the neutrino flux can be obtained by the ratio of the solar luminosity over the energy released per neutrino production, corresponding to a solar neutrino flux of $\sim 6 \cdot 10^{10}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ on Earth. Solar neutrinos are produced via two processes, namely the pp-chain and the CNO-cycle, which are both depicted in Figures 1.3 and 1.4. The pp-chain dominates the energy production in the Sun with about 99 %, while the CNO-cycle is expected to dominate in stars heavier than 1.5 solar masses [19]. The calculated neutrino flux as a function of the energy based on the SSM [22, 23] is shown in Figure 1.5. Depending on the specific reaction, the different neutrino species have a monoenergetic or a continuous energy spectrum up to $\sim 10\,\mathrm{MeV}$.
Solar neutrinos have been observed in several experiments using radiochemical, Cherenkov, or liquid scintillator detectors [24–31]. Besides the pep-neutrinos, all solar neutrino species could be measured with Borexino, which will be discussed in Chapter 2 [32, 33]. Moreover, the measurements of solar neutrinos confirm the SSM and neutrino oscillations according to the MSW-LMA[4] solution.
Notice that the Sun produces in thermal processes (mainly plasmon decay, Compton process, and electron bremsstrahlung) also neutrino pairs of all flavors with an energy in the keV-

[4]LMA stands for Large Mixing Angle [12].

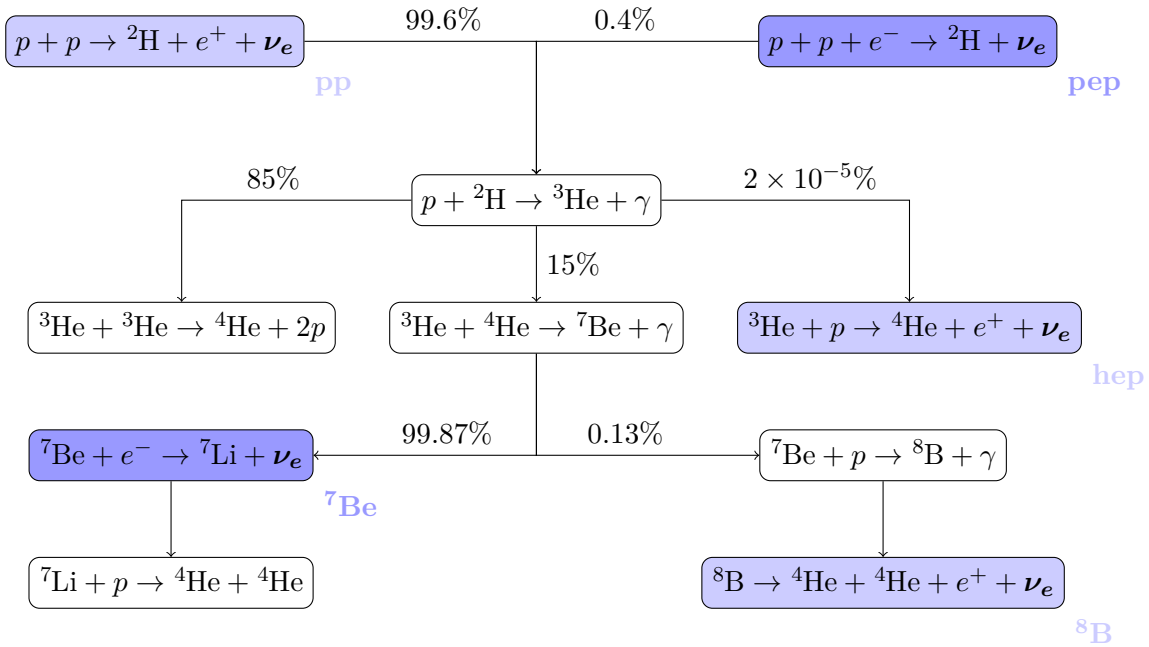


Figure 1.3: Scheme of the pp-chain The pp-chain is the dominant neutrino production mechanism in the Sun. The reactions in which neutrinos with a continuous (monoenergetic) spectrum are produced are indicated in light (dark) blue. The neutrinos are named after the specific production reaction. The branching ratios are taken from [32].

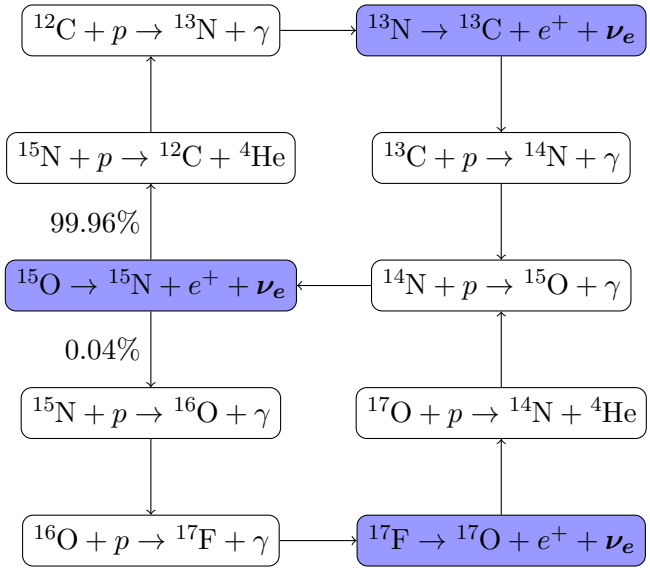


Figure 1.4: Scheme of the CNO-cycle The CNO-cycle consists of two sub-cycles, where the upper cycle dominates. The reactions in which neutrinos are produced are indicated in blue. The branching ratios are taken from [32]. The CNO-cycle plays only a minor role in the Sun, however, it is the main fusion process in heavier stars.

range (see Figure 1.2). However, these thermal solar neutrinos could not yet be observed and no realistic detection potential exists at the moment [13].

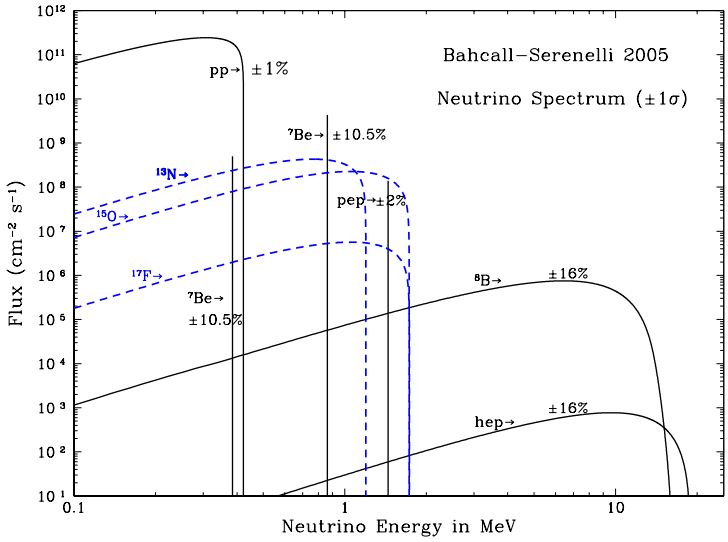


Figure 1.5: Solar Neutrino Spectrum in the SSM The neutrinos produced in the pp-chain (CNO-cycle) are shown in black (blue). For the continuous spectra the flux is given in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{MeV}^{-1}$, for the mono energetic spectra in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The errors given on the flux correspond to the theoretical ones. The figure is taken from Reference [34].

Geo Neutrinos

Geo neutrinos are mainly[5] electron antineutrinos emitted in beta decays of the radioactive isotopes, namely ^{40}K and the decay chains of ^{238}U, ^{235}U, and ^{232}Th, contained in the Earth. The released energy during the radioactive decays is expected to make up more than 50 % of the total terrestrial heat [35]. Geo neutrinos reach energies up to 3.3 MeV and have a flux of the order of $10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Figure 1.6 shows the expected geo neutrino spectrum at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy.
As the exact flux depends on the composition of the Earth, geophysical models can be tested with the measurement of geo neutrinos. Geo neutrinos from the ^{238}U and ^{232}Th decay chain that are above the detection threshold of 1.8 MeV of the inverse beta decay (IBD) have been observed with Kamland [36] and Borexino (see Chapter 2). However, so far statistical uncertainties dominate the observed results and prevent from a differentiation of geophysical models.

Supernova Neutrinos

A massive star explodes at the end of its life in a supernova while its core collapses to a neutron star or a black hole. The density and temperature become so large in the collapsing core that electron neutrinos are produced through electron capture on protons. These electron neutrinos can only escape freely at the beginning of the stellar core collapse. When the density of the stellar plasma becomes too high, the electron neutrinos interact with the infalling matter billions of times and thermalize before they can escape the stellar medium. In these interactions neutrinos and antineutrinos of all flavors are produced, where their mean energy is about 10 MeV. The supernova reaches neutrino luminosities up to $\sim 10^{53}\,\mathrm{erg/s}$. Note that such a supernova neutrino signal lasts only for several seconds [37]. If the supernova is

[5]The isotope ^{40}K undergoes with a probability of 10.7 % electron capture and produces electron neutrinos [35].

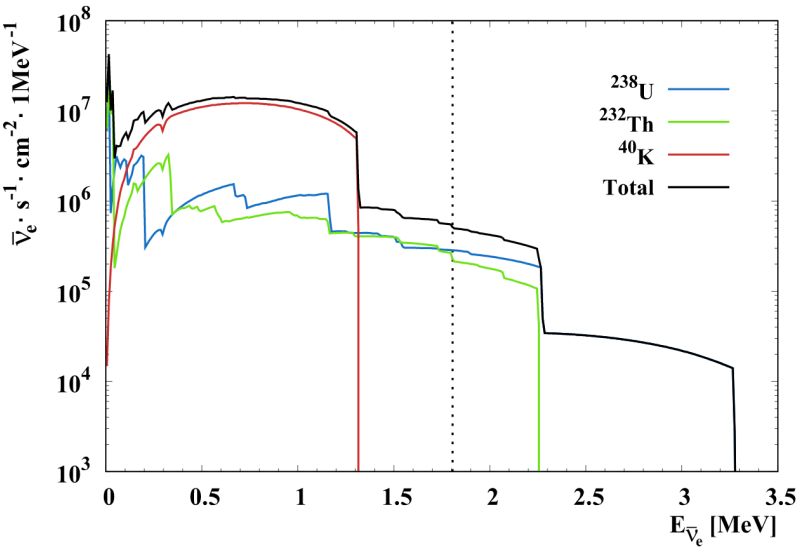


Figure 1.6: Expected Geo Neutrino Flux at LNGS The individual geo neutrino contributions and their sum from ^{40}K and the decay chains of ^{238}U and ^{232}Th are shown. The geo neutrino contributions are modeled for the Borexino site. The dashed line indicates the energy threshold for the inverse beta decay detection reaction. Hence, geoneutrinos from ^{40}K cannot be observed. The figure is taken from [35].

close, the neutrino flux is high enough on Earth and can be detected with neutrino detectors. This was the case for the supernova 1987A in the Large Magellanic Cloud with a distance of $\sim 50\,\mathrm{kpc}$, where the time-integrated neutrino flux of the burst was determined to $\sim 10^{10}$ $\bar{\nu}_e$ per cm^2 [38–40]. The observation confirms the core-collapse supernova model and that the produced neutrinos carry away a large fraction of the gravitational binding energy released in the core collapse. Moreover, the measured energies of the detected ~ 20 events are consistent with the predicted thermal energy spectrum [37].

In addition to a single supernova, one expects the existence of the diffuse supernova background (DSNB), which is the integral neutrino signal of all past core-collapse supernova explosions in the observable Universe. Due to the small flux of $\sim 10^2$ per cm^{2}s [41] and the low energy, it has not yet been observed. However, future large-scale experiments (as JUNO [42,43], Superkamiokande/Hyperkamiokande [44,45], DUNE [46,47], and Theia [41]) might be able to detect such a signal (for a comparison of the discovery potential see Ref [41,48]). The observation of future galactic supernovae as well as the DSNB could provide essential information about the processes during core-collapse supernovae or the formation rates of neutron stars and black holes (for a detailed review see Reference [49]).

Borexino is sensitive to a galactic supernova signal and is a member of the SuperNova Early Warning System (SNEWS) [50]. An upper limit on the DSNB neutrino flux of $< 112.3\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ is reported at the 90 % CL from Borexino in [51].

Atmospheric Neutrinos

Atmospheric neutrinos are secondary particles emitted in pion and kaon decays. These pions and kaons are produced in interactions from primary cosmic rays with nucleons in the Earth's atmosphere. Their broad energy spectrum reaching from 0.1 GeV to several TeV and their wide range of baselines from 10 to about 1000 kilometers allow studying neutrino oscillations. Atmospheric neutrinos are neutrinos as well as antineutrinos and appear in electron and muon

flavor. Atmospheric neutrinos have been detected by a vast list of experiments [52–62], and the first evidence of neutrino oscillations was reported in 1998 from the Super-Kamiokande experiment using atmospheric neutrinos [58]. Atmospheric neutrinos were also observed in Borexino [51,63]. As they can mimic an inverse beta decay signal (the main detection reaction for electron antineutrinos), they contribute to the background for antineutrino searches (see e.g. Reference [51]) and are the most critical background for the detection of the DSNB signal [43].

1.3.2 Artificial Sources

Reactor

Nuclear reactors produce in nuclear fission processes of heavy isotopes, mainly ^{235}U, ^{238}U, ^{239}Pu, and ^{241}Pu, about 600 fission fragment isotopes. These subsequently produce electron antineutrinos of up to $\sim 10\,\mathrm{MeV}$ via beta decays [64]. The flux depends on the fuel composition and the thermal power of the reactor, whereby the flux of a reactor with a thermal power of 1 GW is about $2 \cdot 10^{22}\,\mathrm{s}^{-1}$ [1].
The uncertainty on the flux is rather large with up to several %, as the total flux consists of about 10000 individual beta-decay branches, which complicates the flux prediction [64]. The neutrino flux is modeled in two ways, namely the conversion and ab initio method. The current standard spectrum is derived by Huber und Mueller [65, 66], which exploits the conversion method. Here, the measured integrated beta spectra of ^{235}U, ^{238}U, ^{239}Pu, and ^{241}Pu are converted into the expected neutrino flux. The Huber-Mueller flux differs from the experimental results in the total flux for about 5 % (the so-called reactor antineutrino anomaly [67], see Section 1.5.3) as well as in the spectral shape around $\sim 5\,\mathrm{MeV}$ for about 10 % [68–71]. On the contrary, ab initio calculations directly compute the neutrino flux as the sum of the individual derived beta branches using fission yields and information from nuclear databases [72].
Reactor neutrinos were the first neutrinos that have been detected in 1956 [73] and have been widely used to measure neutrino oscillations [74–77]. Currently, reactor neutrinos are the main neutrino source in experiments searching for sterile neutrinos at the eV-scale (see Section 1.5.5). To counteract the neutrino flux uncertainties, a relative measurement of the energy spectrum at different baselines is performed.
Since there are no nuclear reactors in Italy, the reactor neutrino rate is low at the LNGS. However, reactor neutrinos are measured as the main background component for a geo neutrino analysis in Borexino (see for example [35]).

Accelerator Neutrinos

Accelerator neutrinos are produced in the decays of mesons that are in turn produced when high-energy protons collide onto a target. Accelerator neutrinos are mainly muon (anti-)neutrinos, as pions are the most abundant decay product. A directional neutrino beam with the desired neutrino type can be achieved when the mesons are focused using so-called magnetic horns. The mean energy of accelerator neutrinos lies in the GeV range. The neutrino flux has large uncertainties of 5 %-10 %, mainly because of the uncertainties in the hadron production cross sections [1].
Long- and short-baseline experiments have been looking for neutrino oscillations with a baseline of several $\sim 100\,\mathrm{km}$ and $\sim 1\,\mathrm{km}$, respectively. While the results from the long-baseline experiments confirm the neutrino oscillations within the three-neutrino model [78–82], some of the short-baseline experiments hint at the existence of a sterile neutrino in the eV-scale (see

Table 1.1: Neutrino Oscillation Parameters. The parameters are obtained from a global fit to solar, atmospheric, reactor, and accelerator data from July 2020 assuming three-flavor neutrino oscillations. The values are given for normal ($m_1 < m_2 < m_3$) and inverted ($m_3 < m_1 < m_2$) ordering, where $l = 1$ and $l = 2$, respectively. The values are taken from [85, 89].

parameter	best fit $\pm 1\sigma$	
	normal ordering	inverted ordering
$\sin^2\theta_{12}$	$0.304^{+0.013}_{-0.012}$	
$\sin^2\theta_{23}$	$0.573^{+0.016}_{-0.020}$	$0.575^{+0.016}_{-0.019}$
$\sin^2\theta_{13}$	$0.02219^{+0.000062}_{-0.00063}$	$0.02238^{+0.00063}_{-0.00062}$
$\Delta m^2_{21}[10^{-5}\,\mathrm{eV}^2]$	$7.42^{+0.21}_{-0.20}$	
$\lvert\Delta m^2_{3l}\rvert\,[10^{-3}\,\mathrm{eV}^2]$	$2.517^{+0.026}_{-0.028}$	$-2.498^{+0.028}_{-0.028}$
$\delta_{\mathrm{CP}}[^\circ]$	197^{+27}_{-24}	282^{+26}_{-30}

Section 1.5.3). New experiments searching for sterile neutrinos using accelerator neutrinos are currently in preparation or already taking data (see Section 1.5.5).
Accelerator neutrinos from the CNGS (CERN Neutrinos to Gran Sasso) beam that were sent to the closeby Opera experiment [80] could also be observed in Borexino [83].

1.4 Current Status and Open Questions

Detecting reactor, solar, atmospheric, and accelerator neutrinos with their different energies at different baselines, the neutrino oscillation parameters could be measured (for a detailed review see Reference [1]). The current status of the experimental values obtained from a global fit is listed in Table 1.1. The experiments could measure the three mixing angles and the two mass splittings with $\Delta m^2_{21} \ll |\Delta m^2_{31}| \simeq |\Delta m^2_{32}|$. However, a small discrepancy of 1.4σ between solar and reactor experiments for $\Delta \mathrm{m}^2_{21}$ exists [84], and the octant of δ_{23} is not yet determined.
Neutrino oscillation experiments for which the oscillation probability can be approximated with the two-flavor formula (see Equation 1.4) are not sensitive to the sign of Δm^2. In contrast, three-flavor oscillations and matter effects can be used to determine the neutrino mass ordering. While $m_2 > m_1$ is known from the observation of the MSW effect using solar neutrinos, the neutrino mass ordering (normal: $m_1 < m_2 < m_3$ or inverted: $m_3 < m_1 < m_2$) is still unknown. Current data slightly prefer the normal mass ordering [85]. The neutrino mass ordering will be measured by experiments that perform either a precision spectral measurement of the three-flavor oscillations in the reactor spectrum (JUNO [42]) or look for the matter effects in the Earth of atmospheric and accelerator neutrinos (IceCube-PINGU [86], KM3Net-ORCA [87], DUNE [46], Hyper-Kamiokande [44]).
Further, the value of the CP-violating phase δ is not determined yet. It can be reconstructed by long-baseline accelerator experiments (T2K [88], NOνA [82], and future DUNE [46] and Hyper-Kamiokande [44]) that look for a difference in the neutrino and antineutrino appearance signal of muon to electron neutrinos. Present results hint at a CP violation at the 3σ level, however, only if the inverted mass ordering is assumed [89].
The nature of the neutrino (Dirac or Majorana particle) is investigated in experiments that look for the neutrinoless double-beta decay [90]. If this decay is observed, it can be deduced that neutrinos are Majorana particles. In addition, the then existing Majorana phases that are not observable in neutrino oscillation experiments could possibly be measured. The

current experiments reach limits on the half-life of the neutrinoless double-beta decay of $\sim 10^{26}$ yr. This limit can be converted into an upper limit on the effective Majorana neutrino mass (e.g. $m_{\beta\beta} < 79 - 180$ meV from Gerda [91], $m_{\beta\beta} < 61 - 165$ meV from KamLAND-Zen [92], $m_{\beta\beta} < 93 - 286$ meV from EXO-200 [93], all at the at 90 % CL) and hence into an absolute value of the neutrino mass. The next-generation experiments start to probe the effective Majorana neutrino mass region that is predicted for the inverted mass ordering [94].
Another way to measure the still unknown absolute neutrino mass is the precision spectroscopy of the beta-decay (e.g. KATRIN [95]) and electron-capture (e.g. ECHO [96]) spectra close to the Q-value of the reactions. This is possible, as the Q-value is exactly reduced by the incoherent sum of the neutrino masses m_β. The KATRIN experiment recently reported an upper limit of $m_\beta < 1.1$ eV at the 90 % CL [95].
A complementary indirect measurement for the absolute neutrino mass comes from cosmology that can determine the sum of neutrino masses of those neutrinos that were relativistic in the early Universe. The relativistic neutrinos add up to the energy density of radiation that drives the expansion and temperature evolution of the early Universe. The number and masses of relativistic neutrinos at that time affect, therefore, the signal that can be observed in the Big Bang Nucleosynthesis (BBN), Cosmic Microwave Background (CMB), and the Large Scale Structure (LSS). The current limit is $\sum m_\nu < 0.54$ eV at the 95 % CL [17]. It has to be noted that this limit depends on the assumptions made in the cosmological model, and a laboratory measurement is still mandatory.
After all, the current data agrees well with the three-neutrino model predicted by the SM. However, a few experiments report results that are not compatible and hint at a sterile neutrino at the eV-scale. Sterile neutrinos, which could also help to explain the unresolved question of the generation of the neutrino masses, will be discussed in the next section.

1.5 Sterile Neutrinos

Sterile neutrinos are hypothetical right-handed neutrinos that are neutral lepton singlets under the electroweak $\mathrm{SU(2)_L} \times \mathrm{U(1)_R}$ gauge group [7]. They hence carry no weak charge ($I = 0$) and cannot take part in the weak interaction. However, as they are right-handed, they can form a mass term in the Lagrangian together with the left-handed active neutrinos and couple thus to the active neutrinos [97].
There are no theoretical limits on the number and on the mass of sterile neutrinos [97]. Therefore, experiments look for their existence at various mass scales: sterile neutrinos at the meV scale could affect the oscillation signature of solar [98] and reactor neutrinos [99]; keV sterile neutrinos that could be a possible dark matter candidate are searched for in beta decays [100]; fixed-target and high-energy collider experiments look for sterile neutrinos at the ∼GeV-scale [101–103]. Besides that, a huge experimental effort is currently ongoing for the search of eV sterile neutrinos which will be discussed in the following in more detail and which are the only ones where experimental hints exist for.

1.5.1 Theoretical Motivation

Neutrinos are the only particles that appear solely with left-handed fields in the SM. Thus, many theoretical models extend the SM by right-handed neutrinos which could then also explain unresolved questions of particle physics and cosmology.
Sterile neutrinos can be involved in the neutrino mass generation and a sterile neutrino with $\sim 10^{15}$ GeV could explain at the same time the smallness of the neutrino mass via the so-called "seesaw mechanism" [104–108]. Such a heavy sterile neutrino is naturally predicted by Grand

Unified Theories (GUT) which are theoretical models that try to merge the electroweak and the strong interactions in one theory. The see-saw mechanism works as sketched in Figure 1.1, however, the neutrino flavor vector is extended by the sterile neutrino fields $(\nu_e, \nu_\mu, \nu_\tau) \rightarrow (\nu_e, \nu_\mu, \nu_\tau, \nu_{s_1}, \nu_{s_2}, ..., \nu_{s_N})$, where s_N is the number of total sterile neutrinos [7, 97]. To ease the discussion, in the following one active ν_a and one sterile ν_s neutrino is assumed, where ν_a is massless and ν_s carries the mass m_M [8,109]. In this scenario, neutrinos are Majorana particles and through the interaction with the Higgs field Majorana (m_M) and Dirac[6] (m_D) mass terms are produced with the mass matrix $M = \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix}$. Through the diagonalization of the mass matrix a light ν_{light} and a heavy ν_{heavy} neutrino mass eigenstate are obtained with the masses $m_{light} \approx \frac{m_D^2}{m_M}$ and $m_{heavy} \approx m_M$, respectively. The resulting mass eigenstates are superpositions of the flavor eigenstates ($\nu_{light} \approx \nu_a + \frac{m_D}{m_M}\nu_s$ and $\nu_{heavy} \approx \nu_s - \frac{m_D}{m_M}\nu_a$) and hence have mixed properties of the flavor states. With $m_D \ll m_M$, the almost pure active ν_{light} carries a small mass and the massive ν_{heavy} can take part in the weak interaction which, however, is strongly suppressed.
In order to explain three neutrino masses, at least three sterile neutrinos are needed, but not all of them need to be very heavy. A popular model is the Neutrino Minimal Standard Model (νMSM [110,111]) that can, in addition to the smallness of the neutrino masses, explain the matter-antimatter asymmetry in the Universe [112] (two sterile neutrinos with mass ∼GeV) and predict a dark matter candidate [113] (one sterile neutrino with mass ∼keV).
There are various variations of the seesaw mechanism that could also accompany a light sterile neutrino at the eV-scale (see Reference [114] for an overview). Morever, light sterile neutrinos can be embedded in other models, as for example the gravitational-anomaly mass model that could at the same time also explain neutrino masses [115].

1.5.2 Sterile Neutrino Mixing

In the following, one additional sterile neutrino ν_s in the so-called (3+1) model [7, 97] is assumed with a mass m_4 at the eV scale and the corresponding mass eigenstate ν_4. In this model, neutrinos are generally Majorana particles. The flavor and mass eigenstates are then correlated via an unitary 4×4 mixing matrix [116]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} & U_{\mu 4} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} & U_{\tau 4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{pmatrix} \tag{1.8}$$

In addition to the PMNS matrix, three mixing angles, two Dirac CP-violating phases, and one Majorana CP-violating phase exist. The sterile neutrino ν_s cannot be observed directly. However, the active neutrinos contain a ν_4 component that can affect neutrino oscillations. The current neutrino oscillation data restricts the mixing with the non-standard massive neutrino that must be mostly sterile, i.e.

$$|U_{e4}|^2 \ll 1, |U_{\mu 4}|^2 \ll 1, |U_{\tau 4}|^2 \ll 1. \tag{1.9}$$

For $m_4 \gg m_1, m_2, m_3$ and thus $\Delta m_{41}^2 \approx \Delta m_{42}^2 \approx \Delta m_{43}^2 \sim 1\,\mathrm{eV}^2$ and $\Delta m_{41}^2 \ll \Delta m_{32}^2$, $\Delta m_{41}^2 \ll \Delta m_{21}^2$, the neutrino oscillation probability can be expressed in the two-flavor approximation

[6] Dirac mass terms do not necessarily require Dirac neutrinos. Instead, only left- and right-handed neutrinos are needed.

$(\alpha, \beta = e, \mu, \tau)$:

$$P(\nu_\alpha \to \nu_\beta) \approx \left| \delta_{\alpha\beta} - \sin^2(2\theta_{\alpha\beta}) \cdot \sin^2 \left(\frac{1.27 \cdot \Delta m^2_{41}[\mathrm{eV}^2] \cdot L[\mathrm{m}]}{E[\mathrm{MeV}]} \right) \right|, \tag{1.10}$$

with

$$\sin^2(2\theta_{\alpha\beta}) = 4 \left|U_{\alpha 4}\right|^2 \left| \delta_{\alpha\beta} - \left|U_{\beta 4}\right|^2 \right|. \tag{1.11}$$

The mixing angles of appearance ($\alpha \neq \beta$) and disappearance ($\alpha = \beta$) experiments are connected via

$$\sin^2(2\theta_{\alpha\beta}) = \frac{1}{4} \sin^2(2\theta_{\alpha\alpha}) \sin^2(2\theta_{\beta\beta}). \tag{1.12}$$

1.5.3 Experimental Hints for Sterile Neutrinos

Several experiments report results not compatible with the three-neutrino model which could possibly be explained by a sterile neutrino with the mass at the eV scale. These so-called anomalies are summarized in the following.

LSND and MiniBooNE Anomaly

The LSND experiment [117] was an accelerator experiment with a $\bar{\nu}_\mu$-beam with a few tens of MeV energy. A liquid scintillator detector was placed at a distance of $\sim 30\,\mathrm{m}$ (L/E $\sim$ $1\,\mathrm{m/MeV}$) and an excess of $\bar{\nu}_e$ events at $\sim 3.8\sigma$ compared to the three-flavor neutrino oscillations was observed. This result could be explained by $\bar{\nu}_\mu$-$\bar{\nu}_e$-oscillations enhanced by a sterile neutrino with the mass splitting $\Delta m^2_{41} \gtrsim 0.1\,\mathrm{eV}^2$ [97]. The similar KARMEN experiment [118] with a baseline of $\sim 18\,\mathrm{m}$, however, could not observe such an excess. Due to the lower statistics, KARMEN could only constrain but not exclude the preferred parameter space by the LSND experiment. Therefore, the goal of the MiniBooNE experiment [119] was to test the LSND anomaly. MiniBooNE is sensitive to both neutrino and antineutrino channels (ν_μ-ν_e/$\bar{\nu}_\mu$-$\bar{\nu}_e$) and covers the same L/E-range as LSND, whereas the energy and baseline are by a factor of ten larger than for the LSND experiment ($E \sim$ GeV, $L \sim 540\,\mathrm{m}$). The MiniBooNE data shows an excess in neutrino and antineutrino mode corresponding to 4.7σ [120]. Furthermore, the data is compatible with the LSND data and the combined analysis reaches a significance of 6.0σ. However, the excess is only present in the lowest energy bins which is not fully compatible with the (3+1) model [97, 120]. With that ambiguous result MiniBooNE can not give conclusive results on the LSND anomaly.

Reactor Antineutrino Anomaly

The reactor antineutrino anomaly (RAA) [67] describes an observed deficit of reactor neutrinos in comparison to the expectation for reactor experiments with baselines in between $6\,\mathrm{m}$ to $100\,\mathrm{m}$. The anomaly arose after the reevaluation of the reactor neutrino flux [65] which lead to an increased expected neutrino flux of $\sim 5\,\%$ [7]. The average ratio between experimental and expected neutrino rate is $\bar{R} = 0.933 \pm 0.021$ which corresponds to a significance of $\sim 3.1\sigma$ [97]. The RAA could be explained by neutrino oscillations of reactor neutrinos into sterile neutrinos with a mass splitting of $\Delta m^2_{41} \gtrsim 0.5\,\mathrm{eV}^2$.[7]

[7] There exist some doubts on the RAA which are discussed in Chapter 13.

Gallium Anomaly

In order to exclude unknown systematic errors for their radiochemical detectors, the solar neutrino experiments GALLEX [26] and SAGE [25] exploited high-radioactive ^{37}Ar and ^{51}Cr sources. Both isotopes decay via electron capture and produce electron neutrinos with an energy of $\sim 800\,\text{keV}$. The electron neutrinos were detected via $^{71}\text{Ga}(\nu_e,e^-)^{71}\text{Ge}$ charged-current reactions. The sources were placed near the center of the detectors with an average baseline of $\sim 1\,\text{m}$. Both experiments observed a smaller number of neutrino events than expected, where the average ratio is $\bar{R} = 0.84 \pm 0.05$. This corresponds to a 2.9σ deficit [97]. This discrepancy was confirmed by an independent measurement of the transition strengths in ^{71}Ge [121]. Similar to the RAA, such a signal could be explained by neutrino oscillations to sterile neutrinos with a mass splitting of $\Delta m^2_{41} \gtrsim 1\,\text{eV}^2$.

1.5.4 Global Picture of Sterile Neutrinos in 2015

The compatibility of the global neutrino oscillation data with sterile neutrino models is analyzed and constantly updated by various groups (e.g. [122–125]). In the following, the results from Reference [122] are discussed which reflect the status before the current sterile neutrino search program. The three-neutrino model and hence the absence of sterile neutrinos is excluded at the $\sim 6\sigma$ level which, however, drops to only 2.6σ when the LSND data is not taken into account. Hence, the strongest hint for sterile neutrinos comes from the LSND anomaly. The goodness of fit in the (3+1) model decreases when the MiniBooNE data is included in the global fit which is consistent with the fact that sterile neutrino oscillations cannot describe the MiniBooNE data. Therefore, global fits are often performed without the MiniBooNE data, as shown in Figure 1.7. The preferred parameter space is shown for the oscillation parameters of muon neutrino to electron neutrino appearance ($\Delta m^2_{41} - \sin^2(2\theta_{e\mu})$), electron ($\Delta m^2_{41} - \sin^2(2\theta_{ee})$) or muon ($\Delta m^2_{41} - \sin^2(2\theta_{\mu\mu})$) neutrino disappearance searches. The global best-fit value for the mass splitting is $\sim 1.5\,\text{eV}^2$ and for the oscillation amplitude $\sin^2(2\theta_{ee})$ is ~ 0.1. In addition, the allowed regions at the 3σ level using only disappearance and appearance experiments are indicated. The overlap of these regions is small which shows that appearance and disappearance data are in tension with each other. Moreover, the best-fit value obtained from the appearance experiments is excluded by the disappearance data. Given the experimental hints and the disagreement between appearance and disappearance data, a large number of experiments have been developed over the last years. The next section describes the experimental strategies and efforts in the light sterile neutrino search.

1.5.5 Experimental Strategies in the Sterile Neutrino Search

As sterile neutrinos do not participate in the weak interaction, they can only be detected indirectly via their coupling to the active neutrinos. A large fraction of the experiments looks for neutrino oscillations in appearance and disappearance experiments described by Equation 1.10. Given a neutrino mass splitting of $\sim 1\,\text{eV}^2$, the oscillation length is about $(L/E)_{\text{osc}} \sim \text{m/MeV}$. Therefore, the experiments are constructed to detect neutrinos with a broad L/E range of $\mathcal{O}(\text{m/MeV})$. The broad range is necessary to observe the oscillation signature as a function of the neutrino energy and baseline which would be a smoking gun signature for sterile neutrinos.
On the one hand, this can be realized with a neutrino source with an energy of $\mathcal{O}(1\,\text{MeV})$ (e.g. neutrinos from nuclear reactors or artificial sources). This energy leads to an oscillation length in the baseline L_{osc} with $\mathcal{O}(1\,\text{m})$ that can be directly observed in a large enough detector. On the other hand, having a broad energy spectrum of $\mathcal{O}(100\,\text{MeV})$ as provided

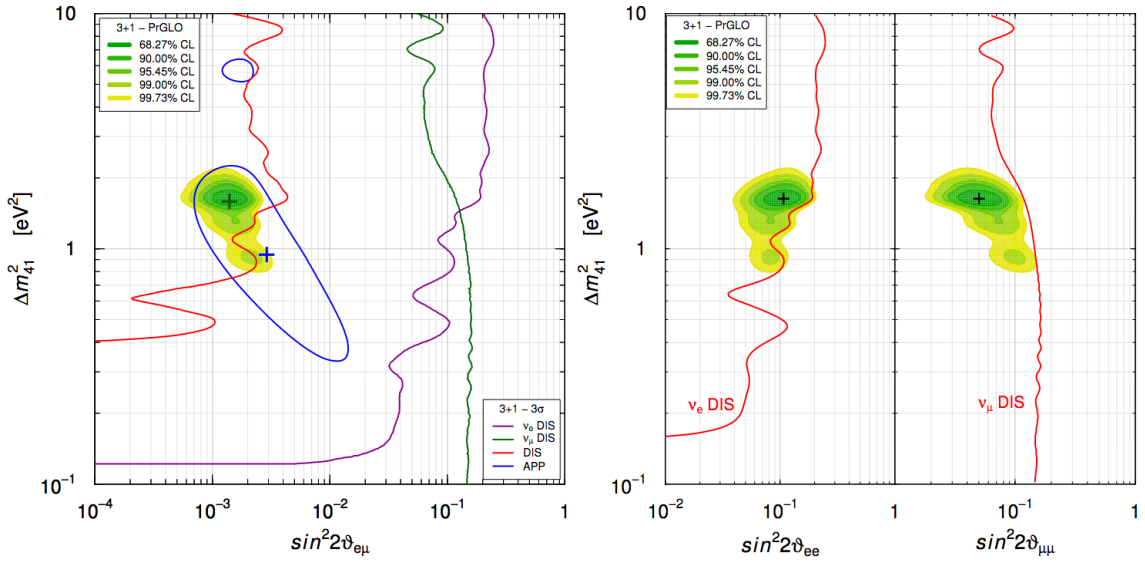


Figure 1.7: Global Picture of Light Sterile Neutrinos The colored area shows the allowed parameter space using global oscillation data in a (3+1) model for the oscillation parameter plane $\Delta m^2_{41} - \sin^2(2\theta_{\mu e})$, $\Delta m^2_{41} - \sin^2(2\theta_{ee})$, and $\Delta m^2_{41} - \sin^2(2\theta_{\mu\mu})$. The lines delimit the 3σ confidence regions when only a subset of the global data is used (violet: ν_e disappearance data, green: ν_μ disappearance data, red: disappearance data, blue: appearance data). The crosses indicate the best-fit values. Comparing the preferred parameter region for appearance (blue) and disappearance (red) data, shows the existing tension between both kind of experiments. The figure is taken from [97].

with accelerator neutrinos, the oscillation signature is also visible in the energy spectrum measured with a detector located at a baseline of $\mathcal{O}(100\,\mathrm{m})$. In order to observe the sought-after oscillation signature, the distance between the detector to the neutrino source needs to be in the same range as the oscillation length. Thus, these experiments place their detectors close to the neutrino source and are called short-baseline (SBL) experiments.

Moreover, to observe the oscillation signature, the neutrino source must be compact and the detector needs to be able to reconstruct the energy and the interaction position of the neutrino. The deposited energy in the detector is often correlated with the initial neutrino energy (e.g. by the detection through the IBD) or can be reconstructed from the energy and direction of the outgoing particles produced in the neutrino interaction. If the energy cannot be reconstructed, monoenergetic neutrino sources are used. To reconstruct the traveled distance, the detector must be either capable of resolving the interaction position or consist of several sub-detectors (or be movable). In appearance experiments using accelerator neutrinos with an energy of $\mathcal{O}(100\,\mathrm{MeV})$, the baseline is $L \sim \mathcal{O}(100\,\mathrm{m})$ and is hence much larger than the source and the detector and can be assumed to be known.

Besides the SBL program, there exist other approaches to look for sterile neutrinos. For example, sterile neutrinos could lead to an enhanced oscillation probability through matter effects in the Earth. The IceCube experiment looks for such a signature using atmospheric neutrinos [136]. Given a light sterile neutrino, the PMNS matrix would not be unitary. Hence, tests of the unitarity of the PMNS matrix can be performed in the search for sterile neutrinos [137]. Moreover, a light sterile neutrino would affect the signature of the experiments that can measure the absolute neutrino mass, i.e. precision spectroscopy of beta decay and electron capture spectra [96, 138], neutrinoless double-beta decay experiments, and cosmol-

Table 1.2: Features and Parameters of Current Short-Baseline Experiments The experiments are grouped according to the source of neutrinos. The kind of neutrinos emitted by the source and those detected by the experiment are shown in the third column. For each experiment the accessible range of distance (L) and energy (E) are given along with the binning used to analyze the data and the expected number of neutrino (N_S) and background (N_B) events. The number of neutrino events is given assuming an oscillation probability of one: $P(\nu_\alpha \to \nu_\alpha) = P(\nu_\alpha \to \nu_\beta) = 1$. Absolute resolutions on the reconstructed baseline and energy (σ_L and σ_E) are quoted for the mean value of L and E. The parameters quoted in this table are sometimes approximated quantities and are intended to give an idea of the signal expected in each experiment.

	Detection	Sought-after	L [m]		E [MeV]		L/E [m/MeV]	binning		
	technology	oscillation	range	σ_L	range	σ_E	range	L×E	N_S	N_B
Experiments with neutrinos from nuclear reactors:										
DANSS [126]	Gd-coated plastic scint.	$\bar{\nu}_e \to \bar{\nu}_e$	11–13	0.2	1–7	0.6	1.5-13	3×24	10^6	10^4
NEOS [71]	Gd-loaded liquid scint.	$\bar{\nu}_e \to \bar{\nu}_e$	24	1	1–7	0.1	3.5–24	1×60	10^5	10^4
NEUTRINO-4 [127]	Gd-loaded liquid scint.	$\bar{\nu}_e \to \bar{\nu}_e$	6–12	0.2	1–6	0.3	1–12	24×9	10^5	10^6
PROSPECT [128]	Li-loaded liquid scint.	$\bar{\nu}_e \to \bar{\nu}_e$	7–9	0.15	1–7	0.1	1–9	6×16	10^4	10^4
SoLid [129]	Li-coated PVT scint.	$\bar{\nu}_e \to \bar{\nu}_e$	6–9	0.05	1–7	0.2	1–9		10^5	10^5
STEREO [130]	Gd-loaded liquid scint.	$\bar{\nu}_e \to \bar{\nu}_e$	9–11	0.3	2–7	0.1	1.3-5.5	6×11	10^4	10^4
Experiments with neutrinos from radioactive sources:										
BEST [131]	Ga radiochemical	$\nu_e \to \nu_e$	0.1–1	0.6	0.4–1.4	–	0.1–2.5		10^4	10^2
SOX [132, 133]	liquid scint.	$\bar{\nu}_e \to \bar{\nu}_e$	4–12	0.15	2–3	0.1	1.3–6.5		10^4	10^2
Experiments with neutrinos from particle accelerators:										
JSNS^2 [134]	Gd-loaded liquid scint.	$\bar{\nu}_\mu \to \bar{\nu}_e$	24	5	10–50	5	0.5–2.5		10^5	10^2
MiniBooNE [120]	mineral oil	$\nu_\mu \to \nu_e$ $\bar{\nu}_\mu \to \bar{\nu}_e$	500	50	200–3000	20	0.2–2.5	1×11	10^5 10^4	10^3 10^2
SBN@FNAL [135]	liquid-Ar TPC	$\nu_\mu \to \nu_e$	110–600	50	200–3000	15	0.1–3		10^5	10^3

ogy [17]. A review of the experimental searches can be found for example in Reference [116]. The current status of these complementary measurements is discussed in Chapter 13.
This thesis focuses on the sterile neutrino searches in SBL experiments. An overview of these experiments and their experimental parameters is given in Table 1.2. It can be seen that a large fraction of the experiments is using reactor and accelerator neutrinos to directly test the reactor and LSND/MiniBooNE anomaly, respectively. In both cases, the neutrino source is intense and high statistical data sets can be achieved. However, the neutrino fluxes suffer from relatively large uncertainties which decrease the sensitivity of these experiments. Therefore, these experiments typically compare the energy spectra measured at different baselines directly to each other and perform in that way a flux-independent measurement. On the contrary, neutrinos from artificial sources can be well characterized and hence are an attractive alternative approach, even though the statistics are limited by the lifetime of the source itself. The SOX [132] experiment was planned to use such an artificial neutrino source, namely $\bar{\nu}_e$ from ^{144}Ce, together with the well-characterized and low-background neutrino detector Borexino. Moreover, using a Gallium radiochemical detector and a ^{51}Cr source, the BEST [131] experiment can directly probe the Gallium anomaly.

In summary, light sterile neutrinos are one of the current hot topics in the neutrino community. Hence, a vast number of SBL experiments are at the moment searching for these new particles beyond the SM. The first part of this thesis discusses the data analysis of one specific experiment, the SOX experiment, which is introduced in more detail in Chapter 2. Further, Chapter 3 reviews statistical tools for the data analysis and gives an overview of the applied statistical methods of the current SBL experiments. Motivated by the incoherent picture of the presentation of their results, the second part of this thesis studies and compares the statistical methods and issues of the SBL experiments.

2 The Solar Neutrino Detector Borexino and the SOX Project

In order to search for a sterile neutrino with a mass of about 1 eV, the distance between the detector and a MeV-neutrino source needs to be in the range of ~ 10 m. One way is to exploit an existing powerful nuclear research reactor with a neutrino flux of $\sim 10^{20}$ antineutrinos per second and to build a small detector with a size of a few meters (for example STEREO [130] or PROSPECT [128]). These reactor-based experiments reach total event numbers of more than 10^5, however, they suffer from the large uncertainties on the neutrino flux. In contrast, the SOX project [132] uses the existing large radio-pure Borexino detector (diameter of active volume is 8.5 m) and produces a well-characterized neutrino source. To achieve a sufficient total event number of $\sim 10^4$ events, the neutrino source needs to have a high activity of more than 10^{15} Bq. Thus, the SOX project is a complementary approach in the sterile neutrino search.
This chapter describes first the Borexino detector with its recently published results (Section 2.1). Second, the SOX project with its planned ^{144}Ce antineutrino source is summarized in Section 2.2.

2.1 The Borexino Detector

The liquid scintillator detector Borexino has been developed to detect solar neutrinos in real time, requiring a low energy threshold, a good energy and vertex resolution, and a low concentration on radioactive impurities. At the moment, Borexino is the only detector with such a low energy threshold of ~ 150 keV which can at the same time reconstruct the deposited energy and the position of neutrino interactions. After a purification campaign in 2011, the concentrations of ^{238}U and ^{232}Th in the detector's center are smaller than 9.4×10^{-20} g/g (95 % CL) and 5.7×10^{-19} g/g (95 % CL), respectively [139]. Borexino is taking data since May 2007 and was able to, thanks to its unprecedented radio-purity, successfully measure the full pp-chain and the CNO cycle of the Sun.
In the following, Sections 2.1.1-2.1.3 describe the Borexino detector, the reconstruction of the inner vessel which contains the active scintillator volume, and the neutrino detection channels. Finally, the backgrounds and recent results of Borexino are summarized in Section 2.1.4 and 2.1.5.

2.1.1 Detector Design and Source Location

The Borexino detector is located in hall C at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy under a rock overburden of 3800 m.w.e. [141]. A schematic drawing of the setup is shown in Figure 2.1. It is composed of two independent light detector systems, namely the inner (ID) and the outer detector (OD), divided by a stainless steel sphere (SSS) with a radius

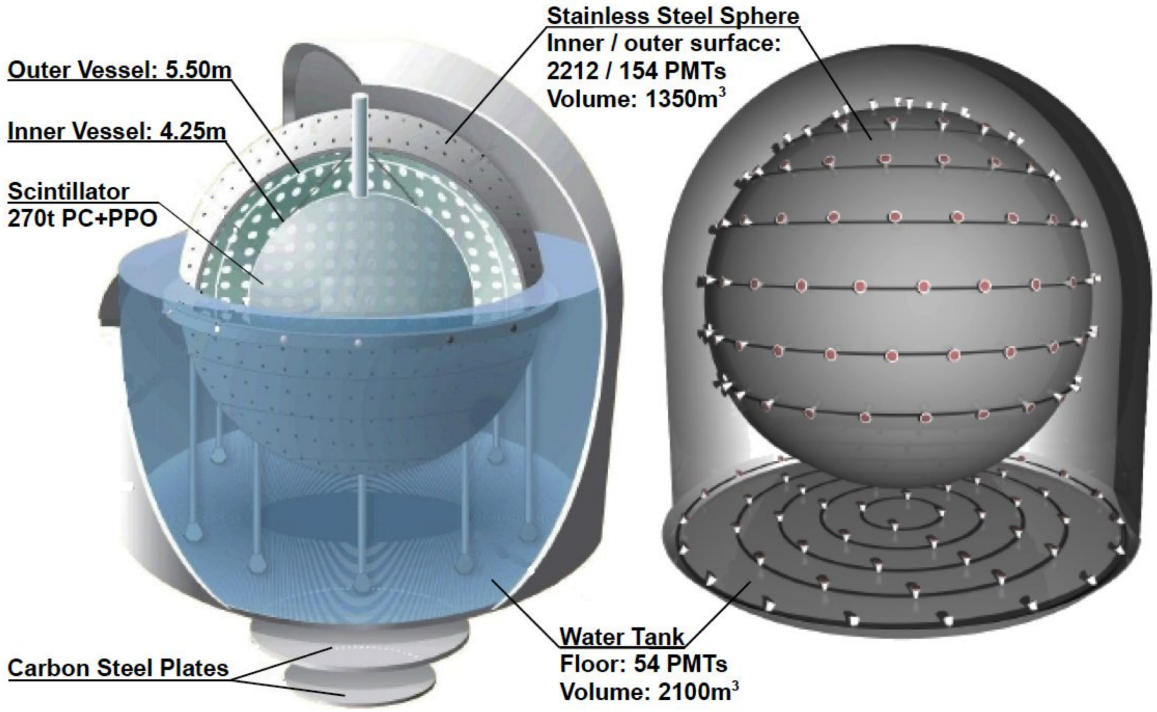


Figure 2.1: Sketch of the Borexino Detector. The Borexino detector consists of two independent light detectors which are divided by a stainless steel sphere. The inner detector contains 270 t of active liquid scintillator hold by a nylon vessel and surrounded by buffer material. For the detection of the scintillation light it is instrumented with 2212 PMTs. The outer detector is a water Cherenkov detector and acts as passive shielding and as active muon veto. The source for SOX is located in a dedicated tunnel (see Figure 2.2) below two carbon steel plates that reduce neutron and gamma background from the source. The figure is taken from [140].

of 6.85 m. The ID is the main detector and observes the scintillation light produced in the active medium. The active medium is a mixture of 270 t pseudocumene (PC,$C_6H_3(CH_3)_3$) and the wavelength shifter PPO ($C_{15}H_{11}NO$) at a concentration of 1.5 g/l. It is contained in a 125 µm nylon vessel at a nominal radius[1] of 4.25 m. The scintillation light with a peak wave length of 360 nm is produced in the fast decay ($\sim$ 3 ns) of excited PPO molecules induced by an incident ionizing particle. Here, the incident ionizing particle excites first PC molecules which in turn excite through a non-radiative energy transfer the PPO molecules. 2212 photomultiplier tubes (PMTs) mounted on the inner surface of the SSS collect the scintillation light with a total optical coverage of $\sim$ 30 %.

The 1024 t buffer mixture of PC and DMP (dimethylphytalate,$C_{10}H_{10}O_{44}$) surrounds the active medium and shields against external background. At the same time, the buffer mixture prevents from the production of scintillation light by quenching it by a factor of $\sim$ 20. A second nylon vessel at 5.5 m separates the buffer into the inner (IB) and outer (OB) buffer. Both vessels are held by nylon strings in the north and south pole of the detector and act in addition as ^{222}Rn barriers.

The ID can reconstruct the energy and position of an event, where the energy is obtained from the total amount of measured scintillation light. The interaction point is deduced using time-of-flight techniques and the arrival times of the scintillation photons. Borexino has a

[1]The exact shape of the so-called inner vessel varies from the designed sphere and is discussed in the next section.

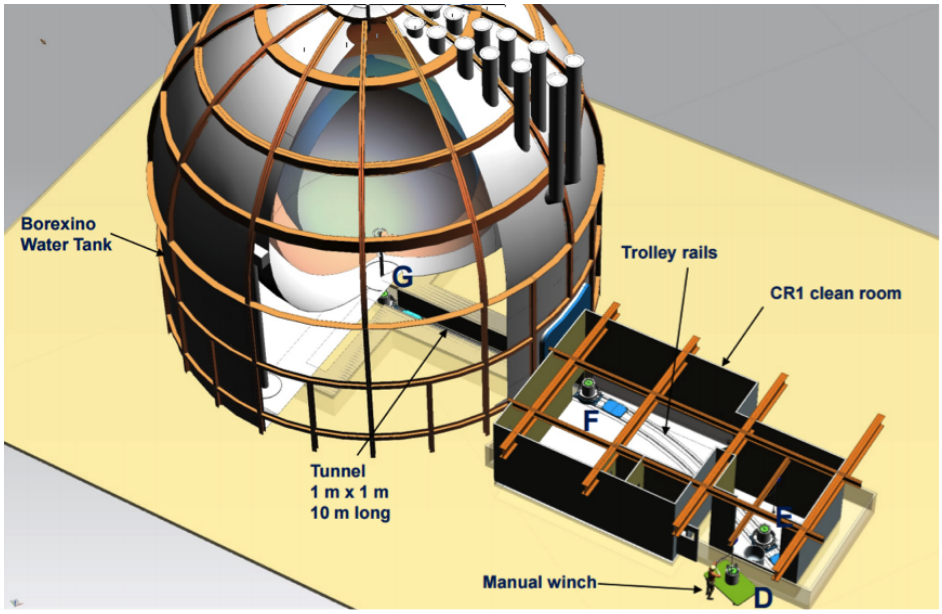


Figure 2.2: Sketch of the Source Location. The source is located in a dedicated pit 8.51 m below the detector center and can be reached via a tunnel and an installed railway system. In addition, the cleanroom, from which the tunnel can be accessed and in which the calorimetric measurements are performed, is shown.

light yield of $\sim 10^4$ photons/MeV and an attenuation length of ∼8 m at 360 nm leading to an energy resolution of 5 % and a spatial resolution of ∼10 cm at 1 MeV. The detector response is unfortunately time dependent. Indeed, the number of active PMTs is decreasing due to aging effects. At the planned start of SOX data taking at the end of 2017, approximately 1400 PMTs were still operational.

Details on the charged particle energy loss affect the light yield and the hit-time distributions. This results into a reduced light yield for alpha particles. The quenching factor of MeV-alpha particles is about a factor of ten higher than for electrons. Moreover, it is possible to distinguish alpha particles from electrons and gammas by means of pulse shape discrimination.

The OD with a diameter of 18 m and a height of 16.9 m is filled with 2100 t of ultra-pure de-ionized water that acts as shielding and muon veto. The emitted Cherenkov light from traversing muons is collected by 208 PMTs mounted on the outer surface of the SSS and the floor of the water tank.

Since 2015 the Borexino detector is thermally insulated with material based on mineral wool. Further, an active temperature control system was installed in January 2016 [33]. This modification of the detector was necessary to reduce convection movements in the active volume due to temperature variations in hall C.

The source position is 8.51 m below the detector's center in a dedicated tunnel of only 1 m height that has been already built at the time of the construction of Borexino. Two carbon steel plates between the source and the detector with dimensions of 8 m x 8 m x 10 cm and 4 m x 4 m x 4 cm are present and shield against neutrons and gamma rays from the source. The tunnel entrance can be accessed via a cleanroom and the source can be moved via a railway system in the tunnel, which is sketched in Figure 2.2. As it is mandatory to avoid any heat transfer from the source to the detector, the source is kept at a constant temperature during data taking using a carefully controlled water cooling system. Moreover, the tunnel

is continuously ventilated.

2.1.2 The Inner Vessel Shape and its Reconstruction

A hole in the inner vessel results in a leak of the scintillator of the inner vessel to the buffer and deforms at the same time the vessel shape. The hole appeared in April 2008 and its location was estimated to 26°< θ <37° and 225°< ϕ <270° [141]. To reduce the leak and the buoyancy forces, the density difference between the buffer and the scintillator has been reduced by lowering the DMP concentration in the buffer from 5 g/l to 2 g/l.

The time-dependent shape of the inner vessel can be reconstructed using the events from radioactive contaminants of the inner vessel, mainly ^{210}Bi, ^{40}K, and ^{208}Tl which is shown in Figure 2.3. The resulting vessel shape is a function of the polar angle, where azimuthal symmetry is assumed (see Figure A.7). To do so, data from three weeks are collected and grouped for the polar angle. The radius for each polar angle is determined from a simultaneous fit of the radial and energy distributions using Monte Carlo (MC) simulations of the three main impurities. The obtained data points for the radius as a function of the polar angle are described with a combined function of a high-order polynomial, a Fourier series, and a Gaussian distribution where the radius at the poles is fixed to 4.25 m (for a detailed description see References [35, 141]). This procedure allows to determine the vessel shape with an uncertainty of 1 % which corresponds to about 5 cm. The uncertainty of the vessel shape is the main systematic source for the SOX experiment and is studied in Chapter 7.

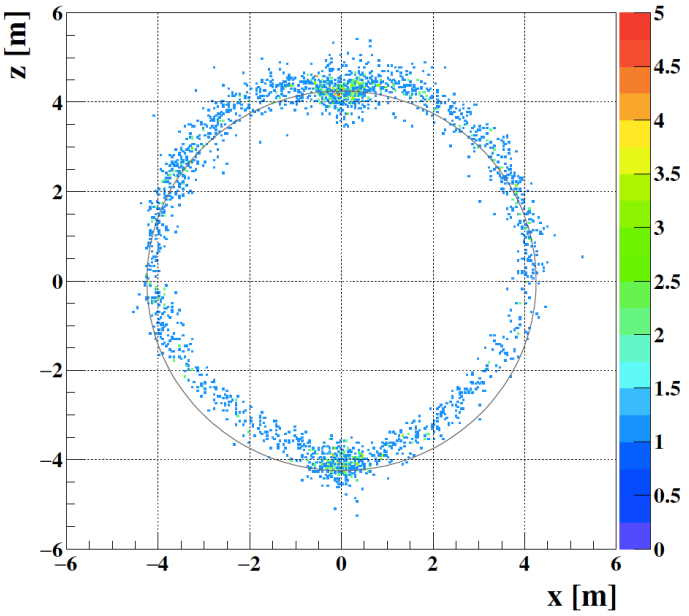


Figure 2.3: Vessel Shape Reconstruction Number of events in a volume of 0.0016 m^3 selected for a time period of three weeks with an energy of (800 - 900) keV and |y|<0.5 m. These events are due to radioactive contaminants of the inner vessel and are used for the vessel shape reconstruction. Due to a hole in the inner vessel and buoyancy forces, the shape of the inner vessel deviates from the nominal vessel shape with a radius of 4.25 m shown in gray. The figure is taken from [35].

2.1.3 Neutrino Detection Channels

Dependent on the neutrino type and its energy, neutrinos can be detected via different detection channels in liquid scintillators (see for example Table 2.1 in Reference [63]). In Borexino, electron neutrinos are detected via elastic scattering and electron antineutrinos mainly via the inverse beta decay (IBD). For both detection channels, the interaction is point-like as the mean free path for the recoil/produced electrons/positrons is only a few centimeters. As the produced scintillation light is in addition emitted isotropically, no directional information can be measured with Borexino.

Elastic Neutrino Electron Scattering

Electron neutrinos scatter off electrons contained in the liquid scintillator

$$\nu_e + e^- \rightarrow \nu_e + e^- \tag{2.1}$$

via charged and neutral current interactions and transfer a part of their energy E_ν to the electrons. The maximum recoil energy T is given for a scattering angle of 180°

$$T \leq \frac{2E_\nu^2}{m_e + 2E_\nu}, \tag{2.2}$$

where m_e is the electron mass. The corresponding cross section is given by [142]:

$$\frac{\mathrm{d}\sigma(E_\nu)}{\mathrm{d}T} = \frac{2G_\mathrm{F}^2 m_e}{\pi E_\nu^2}\left[\mathrm{A}^2 E_\nu^2 + \mathrm{B}^2(E_\nu - T)^2 - \mathrm{AB}m_e T\right] \tag{2.3}$$

where G_F is the Fermi coupling constant. A und B are constants that depend on the weak mixing angle θ_W. For electron neutrinos one obtains $\mathrm{A} = -0.5 - \sin^2\theta_W$ und $\mathrm{B} = -\sin^2\theta_W$. Electron antineutrinos as well as muon/tau (anti-) neutrinos also interact via elastic scattering off electrons, but only via neutral current interactions which leads to reduced cross sections in comparison to electron neutrinos. The main detection channel for electron antineutrinos is hence given by the IBD (cross section is about a factor of ten larger).

The Inverse Beta Decay

Antineutrinos interact mainly via the IBD with protons from the liquid scintillator

$$\bar{\nu}_e + p \rightarrow e^+ + n, \tag{2.4}$$

which is sketched in Figure 2.4. As the neutron is much heavier than the positron, the antineutrino transfers almost all its energy to the positron

$$E_\mathrm{kin}(e^+) \approx E_\nu - Q_\mathrm{IBD}, \tag{2.5}$$

where $Q_\mathrm{IBD} = 1.806\,\mathrm{MeV}$ is the energy threshold of the IBD [143]. The positron annihilates with an electron and two 511 keV gamma rays are emitted. This process happens instantly and provides the so-called prompt signal. The incident neutrino energy can be reconstructed from the measured visible energy E_vis:

$$E_\nu \approx E_\mathrm{vis} + 0.784\,\mathrm{MeV}. \tag{2.6}$$

The neutron thermalizes and after the mean time of $\sim 250\,\mu\mathrm{s}$ it is captured either on Hydrogen ($\sim 99\,\%$) or on Carbon ($\sim 1\,\%$). The exited nuclei de-excite and produce a delayed signal

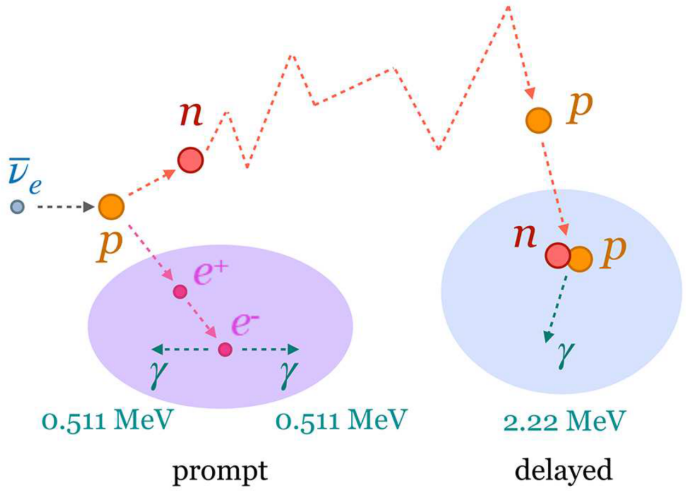


Figure 2.4: Sketch of the Inverse Beta Decay $\bar{\nu}_e$ react via the IBD on protons and produce a positron and a neutron. The positron annihilates instantly with an electron and its kinetic energy is measured together with the one of the two annihilation gammas as a prompt signal. A delayed signal is induced from the neutron which thermalizes and is captured mainly on protons with a mean time of $\sim 250\,\mu s$. The deuteron de-excites and emits a detectable 2.2 MeV gamma. The figure is taken from [35].

by the emission of a 2.22 MeV or 4.95 MeV gamma ray, respectively. This coincidence of the prompt and the delayed event in time and space reduces the background significantly in electron antineutrino searches.
The cross section of the IBD is given in Reference [143]:

$$\sigma_\nu(E_\nu) \approx p_e E_e \left(\frac{E_\nu}{\text{MeV}}\right)^{a(E_\nu)} \cdot 10^{-43} \frac{\text{cm}^2}{\text{MeV}^2}, \tag{2.7}$$

where p_e is the momentum of the electron and

$$a(E_\nu) = -0.07056 + 0.02018 \ln(E_\nu/\text{MeV}) - 0.001953 \ln^3(E_\nu/MeV).$$

2.1.4 Backgrounds in Borexino

From the beginning, the material used in the Borexino detector was selected and handled with special care to reduce radio impurities as much as possible. Moreover, the scintillator was purified several times [144] and an unprecedented radio purity could be achieved. The contamination levels of ^{238}U and ^{232}Th are 9.4×10^{-20} g/g (95 % CL) and 5.7×10^{-19} g/g (95 % CL), respectively which is about ten orders of magnitude lower than the one of natural material on Earth [139].
In the following, an overview of the backgrounds for a neutrino search in Borexino is given (for a detailed discussion see Reference [141]). In particular, the backgrounds for the detection channel of elastic scattering is reviewed. The background expected for the IBD and thus the SOX analysis are discussed in Chapter 6. The background in Borexino can be classified into three categories: external background including contaminations on surfaces, internal background, and cosmic muons and muon-induced background.

External Background

All sources of radioactivity which are located outside of the scintillator are defined as external background. One example are radioactive contaminants in the glass of the PMTs (^{208}Tl,

^{214}Bi, and ^{40}K) and surface contaminants of the vessels (^{210}Bi). The buffer and the two vessels shield against external backgrounds and the external background is suppressed in the center of the detector. Hence, due to the ability to resolve the interaction position only data from the fiducial volume (FV) can be used for the analysis. The specific FV depends on the type of analysis and will be optimized for SOX in Chapter 7.

Internal Background

Despite the high radio-purity of the detector materials used for Borexino, the scintillator still contains residual internal contaminations. Among the components of the scintillator (mainly Hydrogen and Carbon), the beta-decaying ^{14}C is an irreducible background, as it is a long-lived isotope with a lifetime of 5732 y. ^{14}C is the main background component in Borexino and its Q-value of 156 keV hence defines the energy threshold for the neutrino detection.
After ^{14}C, ^{210}Bi and ^{210}Po are the most abundant background components in Borexino. ^{210}Bi decays via beta decay with a mean lifetime of 7.23 days and a Q-value of 1160 keV into ^{210}Po. ^{210}Po is a mono-energetic alpha emitter at 5410 keV with a mean lifetime of 200 days. Due to the strong quenching of alpha particles, the reconstructed energy is at ~ 400 keV. Both isotopes are daughters of the decay chain of the primordial isotope ^{238}U and could be technically determined by the fast ^{214}Bi-^{214}Po coincidence. However, both contaminations are higher and out of secular equilibrium with the rest of the decay chain. One reason is the contamination of the scintillator with the mother isotope ^{222}Rn during detector operations (for example during detector calibrations or scintillator purifications) that emanates from the materials that are in contact with the scintillator. Moreover, convective motions in the scintillator can bring the long-lived ^{210}Po from the inner vessel in the center of the detector. As ^{210}Bi has a similar spectral shape as the solar CNO neutrinos, its determination is crucial for the detection of the CNO neutrinos.
Another major background component is the beta-emitter ^{85}Kr. ^{85}Kr is contained in air and could have mixed with the scintillator during the detector filling. It decays with a mean lifetime of 15.4 years with an endpoint of 665 keV. The energy range is similar to the one of solar ^{7}Be neutrinos and is thus the main background for this analysis.

Cosmogenic Background

An additional background component are cosmogenic muons and muon-induced cosmogenic background. The muon flux on the sea level is $\sim 6.5 \cdot 10^5\, \mu/(\mathrm{m}^2 \cdot \mathrm{h})$ which is reduced by the rock overburden of 3800 m.w.e. by about six orders of magnitude. This corresponds to a detectable rate of ~ 4310 muons per day in Borexino [63]. The muons can through spallation processes on scintillator atoms produce neutrons, protons, and short-lived radioisotopes. The dominant background source is ^{11}C which is a β^+-emitter with a Q-value of 960 keV and a mean lifetime of 29.4 min. Such a large lifetime makes the tagging of this background component challenging. However, as ^{11}C is produced in coincidence with a cosmogenic neutron, a three-fold coincidence (TFC) of muon, neutron and ^{11}C, can be used to suppress the ^{11}C background. The rejection efficiency of the TFC is given by (92±4) % with a remaining exposure of (64.28±0.01) % [139].

2.1.5 Physics Results with Borexino

With a measurement time of more than 13 years, Borexino has produced a huge list of scientific results.[2] Previously to Borexino, solar neutrinos could be measured by the radio-

[2] A list of all publications can be found on the Borexino website [145].

chemical experiments (Homestake [24], SAGE [25], and GALLEX/GNO [26]) that have only measured an integral neutrino flux above a certain energy threshold ($\mathcal{O}(100\,\text{keV})$) or real-time measurements with water Cherenkov detectors (Kamiokande/Super-Kamiokande [27] and SNO [28]) that have an energy threshold of several MeV and are hence only accessible to the high-energy solar neutrinos (^{8}B and hep). Borexino is the first detector that can perform a real-time measurement with a low energy threshold ($\sim 150\,\text{keV}$). In the following, the latest results of solar neutrinos, including the first detection of the CNO neutrinos, and geo neutrinos are summarized.

Solar Neutrinos

A measurement of solar neutrinos does not only allow to probe the standard solar model (SSM) and the metallicity of the Sun [146], but also to test neutrino properties and physics beyond the standard model (SM).
In 2018, Borexino measured simultaneously all solar neutrinos which are produced in the pp-chain [32]. Previously, the individual neutrino types of the pp-chain have been determined individually which can be found in References [147–150]. The first direct evidence of CNO neutrinos has been released recently [33].

Methods In the simultaneous measurement of the pp-chain, the analyses are divided into three energy regions to treat the different backgrounds and systematics individually. First, the pp, ^{7}Be, and pep neutrinos are determined in the energy region of 0.19 MeV-2.93 MeV with a simultaneous fit in energy, radial distribution, and a pulse shape parameter. The latter is important to distinguish between β^- and β^+ particles and to handle the cosmogenic ^{11}C background. As described in the previous section, ^{11}C can be suppressed by the TFC-technique. The fit is performed on the TFC-subtracted as well as on the TFC-tagged energy spectrum. The TFC-subtracted energy spectrum and the fit of the individual neutrino and background components are shown in Fig 2.5a. The ^{8}B-rate is constrained to the result that is obtained in the high-energy region and the CNO-rate to the expectation of the SSM. The background rates are whenever possible measured independently and constrained in the fit. Two fit techniques based on MC simulations and an analytical response function are performed and produce consistent results. Second, ^{8}B neutrinos are determined in the energy region from 3.2 MeV-16 MeV with a fit of the radial distribution to separate the external background from the uniform distributed ^{8}B neutrinos[3]. Third, hep-neutrinos are analyzed in a counting analysis in the energy region 11 MeV-20 MeV. As the measurement event number is consistent with the expected background (mainly cosmogenic ^{11}Be and ^{8}B neutrinos), only an upper limit for the hep-neutrino rate could be computed. Hence, hep-neutrinos are the only solar neutrino species that are not yet observed.
The analysis performed for the CNO neutrinos [33, 151] is the same as the one described to extract the pp, ^{7}Be, and pep neutrinos. However, the similar energy spectrum of CNO, the background ^{210}Bi and the solar pep neutrinos lead to a correlation of these three components in the fit. For a measurement of CNO it is hence mandatory to constrain the pep and ^{210}Bi components. The pep component can be constrained within 1.4 % using the known solar luminosity and the ratio of the pp to pep neutrino rate. The main challenge is the constraint on ^{210}Bi, where the strategy is sketched in Figure 2.6. Thanks to the stabilization of the temperature of the detector after thermal insulation and the installation of an active temperature system, convective movements in the scintillator are reduced and an upper limit on the ^{210}Bi event number can be achieved. Due to the upper limit on ^{210}Bi, a lower limit for

[3] In addition, the analysis is performed in two sub-energy regions. For a detailed description see Ref [32].

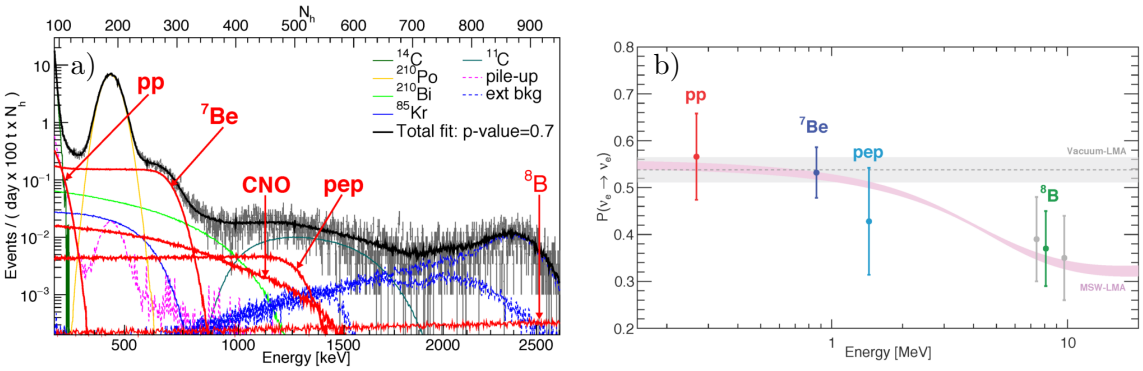
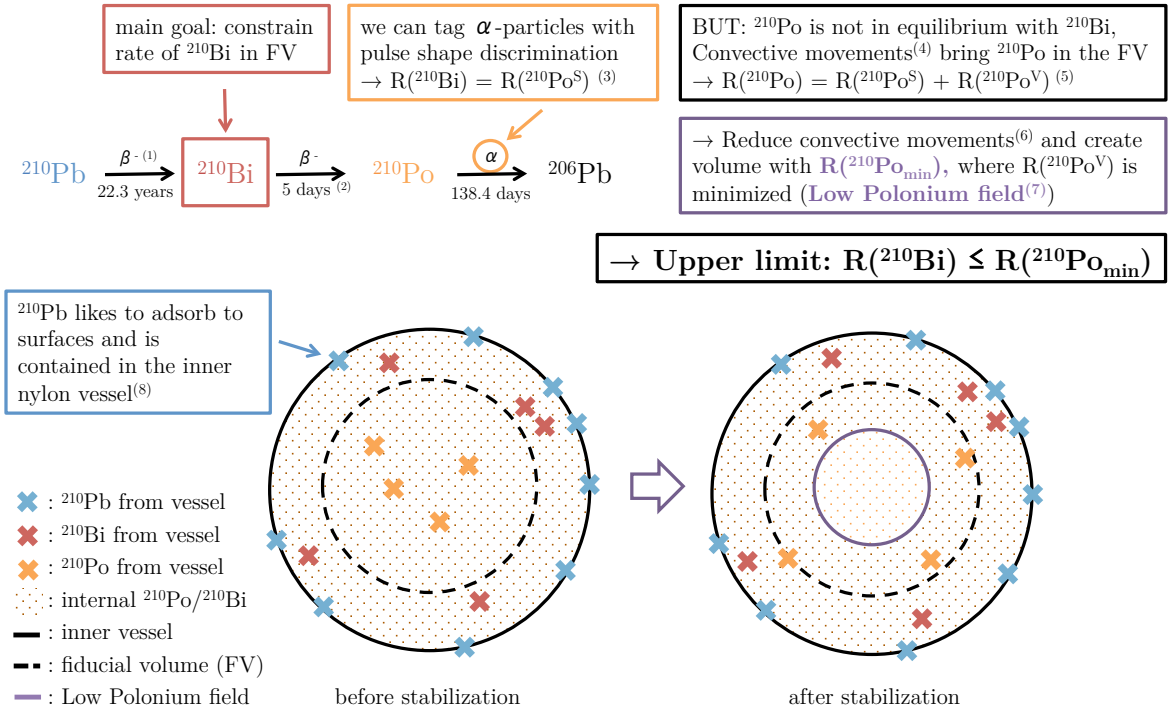


Figure 2.5: Solar Neutrino Measurement with Borexino a) Measured energy spectrum after data selection in Borexino and the individual neutrino and background components obtained from a multivariate fit. The solar neutrino species are colored in red, where the CNO and ^{8}B contributions are fixed in the fit. b) Survival probability computed from the solar neutrino fluxes measured with Borexino. The results agree well with the MSW-LMA prediction (light red). The energies correspond to the average neutrino energies of the specific neutrino type. Three values are obtained for the ^{8}B-neutrinos, as three analyses with different energy regions are performed. Both figures are taken from Reference [32].

CNO can be determined. The result is cross-checked with a simple rate analysis in the energy window in which the CNO rate is maximized compared to the background (780 keV-885 keV). A detailed discussion about these analyses, can be found in References [32, 33, 139, 151, 152].

Results The experimental results of the solar neutrino event numbers and the corresponding fluxes are listed in Table 2.1. The experimental uncertainty of the ^{7}Be flux is by a factor of two smaller than the theoretical uncertainties. The low energy threshold of the ^{8}B analysis is the lowest energy threshold used so far with respect to other experiments and the result agrees with the high-precision result from SuperKamiokande [153]. The CNO rate results to $7.2^{+3.0}_{-1.7}$ cpd/100t and is thus discovered for the first time at the 5.0σ level. In addition, the expected fluxes in the SSM for high (HZ) and low (LZ) solar metallicity are quoted in Table 2.1. The solar metallicity denotes the abundance of elements heavier than Helium and is a fundamental parameter for the determination of the physical properties of the Sun and is so far only poorly understood [146]. The solar metallicity affects the solar neutrino fluxes, mainly the CNO flux which differs about 40 % for the HZ and LZ-SSM. The results are compatible with the LZ and HZ metallically predictions from the SSM. However, the data seems to favor the HZ model. Given a better constraint on ^{210}Bi and a more precise measurement on CNO, the solar metallicity could possibly be determined in the future.
Given the measured event numbers and the expected fluxes (assuming high metallicity), the survival probability P_{ee} can be determined as a function of the neutrino energy, which is shown in Figure 2.5b. Borexino is the only experiment that can test the vacuum-dominated and matter-dominated regime at the same time. The results agree with the MSW-LMA solution and the vacuum-LMA solution is disfavored at a 98.2 %CL. With that measurement, the most precise measurement of the survival probability in the low energy region is achieved. By possibly increasing the accuracy of the pep measurement or lowering the energy threshold of the ^{8}B analysis in the future[4], the accuracy of the survival probability of the transition region from the vacuum dominated to matter-dominated regime could be enhanced. With

[4] The low energy threshold of the ^{8}B analysis is currently limited by the 2.6 MeV gamma from the ^{208}Tl background.



(1): 210Pb cannot be observed directly, as the Q-value of 63.5keV is below the detection threshold of Borexino. (2): The lifetime of 210Bi is too short too reach the FV.(3): R(210PoS) stands for the 210Po events that are supported from internal 210Bi contamination. (4): Convective movements are due to the temperature variations in hall C, as for example seasonal variations. (5): R(210PoV) stands for the 210Po events that are supported from the vessel contamination. (6): The temperature of the detector is stabilized through its insulation and an active temperature control system. (7): The LPoF is smaller than the FV. R(210Pomin) can be assumed to be uniformly distributed in the FV and can be extrapolated to the FV. To demonstrate this hypothesis, data of the energy region with the maximized 210Bi-rate has been analyzed: the 210Bi rate is found uniformly distributed in the FV and stable over time. (8): 210Pb is assumed to stick to the vessel, as 210Bi is uniformly distributed in the FV and stable over time.

Figure 2.6: Sketch of the Strategy to Constrain the ^{210}Bi-rate for the CNO Measurement in Borexino Due to the similar spectral shape to the CNO spectrum, the correlated ^{210}Bi-rate needs to be constrained for a measurement of CNO. The concentration of the alpha-decaying daughter isotope ^{210}Po can be determined via pulse shape discrimination techniques. However, additional ^{210}Po events are brought in from the vessel surface to the fiducial volume through convective movements of the scintillator and secular equilibrium is not given. Borexino has been stabilized through insulation and an active temperature control system. The reduced convective movements allowed to set an upper limit on the ^{210}Bi rate. For a detailed description see References [151] and [33].

that, physics beyond the SM could be tested, such as a sterile neutrino imprint at the meV scale [98].
From the Borexino measurement, the solar luminosity can be determined to $L = 3.89^{+0.35}_{-0.42} \cdot 10^{33}$ erg/s which agrees with the one obtained from photon measurements [154]. On the one hand, the agreement confirms the nuclear origin of the solar power and, on the other hand, it shows that the Sun has been in thermodynamic equilibrium over at least $\sim 10^5$ years. This can be concluded from the fact that the neutrino measurement mirrors the current luminosity. Instead, the photon measurement gives the luminosity of $\sim 10^5$ years ago which is the time needed by the photons (produced in the center) to reach the surface of the Sun. One can further extract the relative intensity of the ^{3}He-^{4}He and ^{3}He-^{3}He fusion rates using the measured fluxes of pp and ^{7}Be neutrinos to

$$R = \frac{2\Phi(^7Be)}{\Phi(pp) - \Phi(^7Be)} = 0.178^{+0.027}_{-0.023} \tag{2.8}$$

which agrees with the predicted values and confirms the SSM.

Table 2.1: Borexino Results on Solar Neutrinos The measured rate and the respective flux (using MSW-LMA predicted values) are listed in column two and three. The results are compared to the SSM-prediction for high (HZ) and low (LZ) solar metallicity in the fourth column. The results for pep neutrinos are obtained when CNO is constrained to HZ and LZ expectations. The results are taken from [32] and [33].

ν source	Rate [cpd/100 t]	Flux [$\mathrm{cm^{-2}s^{-1}}$]	Flux SSM-Prediction [$\mathrm{cm^{-2}s^{-1}}$]
pp	$134 \pm 10^{+6}_{-10}$	$6.1 \pm 0.5^{+0.3}_{-0.5} \cdot 10^{10}$	$5.98(1.0 \pm 0.006) \cdot 10^{10}$ (HZ)
			$6.03(1.0 \pm 0.005) \cdot 10^{10}$ (LZ)
^{7}Be	$48.3 \pm 1.1^{+0.4}_{-0.7}$	$4.99 \pm 0.11^{+0.06}_{-0.08} \cdot 10^{9}$	$4.93(1.0 \pm 0.06) \cdot 10^{9}$ (HZ)
			$4.50(1.0 \pm 0.06) \cdot 10^{9}$ (LZ)
pep (HZ)	$2.43 \pm 0.36^{+0.15}_{-0.22}$	$1.27 \pm 0.19^{+0.08}_{-0.12} \cdot 10^{8}$	$1.44(1.0 \pm 0.01) \cdot 10^{8}$ (HZ)
pep (LZ)	$2.65 \pm 0.36^{+0.15}_{-0.24}$	$1.39 \pm 0.19^{+0.08}_{-0.13} \cdot 10^{8}$	$1.46(1.0 \pm 0.009) \cdot 10^{8}$ (LZ)
^{8}B	$0.223^{+0.015+0.006}_{-0.016-0.006}$	$5.68^{+0.39+0.03}_{-0.41-0.03} \cdot 10^{6}$	$5.46(1.0 \pm 0.12) \cdot 10^{6}$ (HZ)
			$4.50(1.0 \pm 0.12) \cdot 10^{6}$ (LZ)
hep	< 0.002 (90% CL)	$< 2.2 \cdot 10^{5}$ (90% CL)	$7.98(1.0 \pm 0.30) \cdot 10^{3}$ (HZ)
			$8.25(1.0 \pm 0.12) \cdot 10^{3}$ (LZ)
CNO	$7.2^{+3.0}_{-1.7}$	$7.0^{+3.0}_{-2.0} \cdot 10^{8}$	$4.88(1.0 \pm 0.11) \cdot 10^{8}$ (HZ)
			$3.51(1.0 \pm 0.10) \cdot 10^{8}$ (LZ)

Geo Neutrinos

Geo neutrinos are electron (anti-)neutrinos that are produced in radioactive decays from the natural radioisotopes contained in the Earth: ^{238}U, ^{235}U, ^{232}Th, and ^{40}K (see Chapter 1.3) [35]. The energy produced in these decays is referred to as radiogenic heat and correlates to the number of produced geo neutrinos. Hence, with a measured geo neutrino signal, one can deduce the radiogenic heat of the Earth. Especially, the radiogenic heat production and the chemical composition in the different Earth layers can be studied with geo neutrinos. In a simplified scheme, the Earth can be separated into the core, the mantle, and the lithosphere. While the radiogenic heat in the lithosphere is relatively well known (and one does not expect any contribution from the core), the contribution from the mantle has large uncertainties and the geological models differ. The total terrestrial heat flux consists mainly of the radiogenic heat plus the left-over heat from Earth creation that is released during the cooling process of the Earth. Given a known radiogenic and total terrestrial heat allows thus to learn about the thermal and dynamical evolution of the Earth.

Geoneutrinos are detected via the IBD in Borexino. This detection channel is unaccessible to the ^{40}K contributions as their energy is below the IBD threshold of 1.8 MeV. The main backgrounds are reactor neutrinos, however, Borexino profits from the fact that there are no nuclear reactors in Italy. Geo neutrino measurements with Borexino are published in [35, 155–157], where the latest analysis in 2019 benefits from the preparation for the SOX analysis partially developed in this thesis. The exposure of $(1.29 \pm 0.05) \cdot 10^{32}$ protons $\times$ year could be increased by a factor of two in comparison to the previous result from 2015 with an increased precision from 24 % to 18 %. The geo neutrino signal is measured in a spectral fit to $52.6^{+9.4}_{-8.6}(\mathrm{stat})^{+2.7}_{-2.1}(\mathrm{syst})$ which corresponds to $47.0^{+8.4}_{-7.7}(\mathrm{stat})^{+2.4}_{-1.9}(\mathrm{syst})$ TNU (Terrestrial Neutrino Units). The obtained statistics of geo neutrinos are too small to be sensitive to the chemical composition, particularly to the Th/U mass ratio. However, the results obtained from an unconstrained and fixed Th/U mass ratio give compatible results, where the fixed mass ratio corresponds to the one measured from chondritic meteorites [158, 159]. As the signal from the lithosphere is known using geochemical and geophysical data [160], it can

be constrained in the fit and the unknown signal from the mantle can be determined. The mantle signal results to $21.2^{+9.5}_{-9.0}(\text{stat})^{+1.1}_{-0.9}(\text{syst})$ TNU which corresponds to a radiogenic heat of $24.6^{+11.1}_{-10.4}$ TW. With an assumption of the contribution of ^{40}K from the mantle and the total radiogenic heat of the lithosphere, a total radiogenic heat of the Earth of $38.2^{+13.6}_{-12.7}$ TW can be determined. This result corresponds to about 80 % of the total terrestrial heat flux and is in tension ($\sim 2.4\sigma$) with geological models that predict a low abundance of Thorium, Uranium and Kalium in the mantle. With the measured geoneutrino signal, one can test also the hypothesis of a georeactor at the Earth's center. One can exclude at the 95 % CL such a georeactor with a power larger than 2.4 TW. A similar analysis to the geo neutrino analysis is performed in Chapter 6 which discusses the background for the SOX experiment.

2.2 The SOX Project: Short Baseline Neutrino Oscillations with Borexino

The SOX project merges two experimental ideas described in Reference [132] and [133]. In a first phase, a ^{144}Ce-^{144}Pr source which is an $\bar{\nu}_e$-emitter is planned and is described in the next section. In case of a positive sterile neutrino signal and to validate the result, a second phase of the SOX project with a ^{51}Cr ν_e-source is scheduled. The properties of the ^{51}Cr source are shortly discussed in the following. For more details and its sensitivity to sterile neutrinos, see References [132, 161].
^{51}Cr decays via electron capture with a lifetime of 27.7 d and releases mono-energetic electron neutrinos with an energy of 752.7 keV and 432.6 keV [162]. As ν_e are detected via elastic scattering off electrons, the analysis with a ^{51}Cr source is complementary to the ^{144}Ce-^{144}Pr source. In comparison to an $\bar{\nu}_e$-source and the IBD detection channel, the activity of a ν_e-source needs to be larger as the cross section is smaller for the elastic neutrino electron scattering. In addition, the background is larger due to the single scattered electron signal. The main advantage for a ^{51}Cr source is that the source emits mono-energetic ν_e and the energy of the neutrinos is therefore well known. Moreover, the emitted energy per decay is also well characterized which is important for the determination of the activity (see Section 2.2.2). ^{51}Cr is produced by neutron irradiation inside a nuclear reactor via the process ^{50}Cr(n,γ)^{51}Cr. An advantage is that an enriched ^{50}Cr metal is available to the collaboration, however, the production process is challenging and the feasibility for the activity of ~ 500 PBq was under investigation [163]. Further challenging aspects are the short lifetime of the source which requires a quick transportation of the source and allows only for a short measurement time. In the following, the properties of the ^{144}Ce-^{144}Pr source and its characterization are discussed in Section 2.2.1 and 2.2.2. The chapter closes with the expected signature for the SOX experiment (Section 2.2.3).

2.2.1 The ^{144}Ce - ^{144}Pr Source

An intense search was performed to find the ideal isotope for a radioactive antineutrino source [133]. A long life time for the production, transportation and the measurement itself is needed, as well as a high Q-value to observe a high fraction above the energy threshold of the IBD of 1.8 MeV. ^{144}Ce fulfills all the requirements by beta-decaying with a long half-life of 250 d into ^{144}Pr, which decays shortly ($\tau_{\text{Pr}} = 17$ min) via another beta decay into the stable ^{144}Nd with a high $Q_{\text{Pr}} = 3$ MeV:

$$^{144}\text{Ce} \rightarrow {}^{144}\text{Pr} + e^- + \overline{\nu}_e \tag{2.9}$$

$$^{144}\mathrm{Pr} \rightarrow {}^{144}\mathrm{Nd} + e^- + \overline{\nu}_e \tag{2.10}$$

A simplified decay scheme can be seen in Figure 2.7. ^{144}Ce as well as ^{144}Pr have three main decay branches, where only the two main branches of ^{144}Pr have an energy endpoint above the IBD threshold and contribute hence to the SOX signal. With about 2 % ^{144}Pr decays into an excited state of ^{144}Nd, where high-energy gammas of 696 keV and 2186 keV are emitted while relaxation. These gammas need to be shielded, especially the 2186 keV gamma line that could, due to the similar energy, mimic the delayed signal of the IBD. The modeled neutrino spectrum of all decay branches is shown in Figure 2.8. One drawback of the ^{144}Ce-^{144}Pr source is that some of the beta transitions are classified as forbidden transitions that complicate the theoretical modeling (see for example [165]) and increase the uncertainties on the spectral shape.[5] Thus, several experimental setups have been developed to measure the spectral shape and are discussed in the next section.

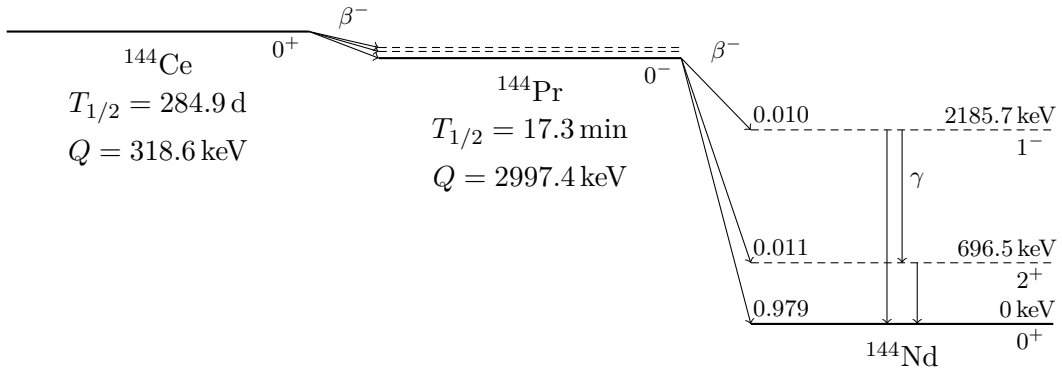


Figure 2.7: Simplified Scheme of ^{144}Ce Decay. The main decay branches along with their energies, branching ratios, and angular momenta and parity of the nuclear states are sketched. The values are taken from Reference [164].

Another criterion for the selection of the ideal source is that the production of the source has to be feasible and efficient. ^{144}Ce can be extracted from spent nuclear fuel with relatively-high fission yields of 5.5 % and 3.7 % from ^{235}U and ^{239}Pu per fission, respectively [133]. The Russian company "Federal State Unitary Enterprise Mayak Production Association" can extract ^{144}Ce over several complex separation steps from spent nuclear fuel using displacement complexing chromatography techniques (for the description of the individual processes see Reference [133, 167, 168] and references therein). In the end, 4 kg of CeO_2 are pressed with a density of 2.5 g/cm^3 and sealed into a stainless steel capsule of 150 mm height and diameter (see Figure 2.9a). The design of the capsule includes in addition copper disk radiators that are responsible for the heat transfer in the source. Moreover, a free volume of $\sim$ 25 % is taken into account to ensure that the pressure in the capsule does not exceed 6 bar.[6]

A 19 cm thick W-alloy shield (see Figure 2.9b) with the demand to attenuate the emitted gamma line of 2.185 MeV by a factor of at least $3 \cdot 10^7$ has been produced at Xiamen Tungsten Co. Ltd. in China. With a density of 18 g/cm^3, the shielding weights $\sim$ 2.5 t.

[5] The classification of a beta decay depends on the change of total angular momenta and parity of the mother and daughter nuclei (see for example Reference [166]). According to the selection rules, the three beta decays from ^{144}Ce and the main decay branch from ^{144}Pr are first forbidden non-unique transitions. ^{144}Pr decays in addition into two excited states of ^{144}Nd with an allowed and a first forbidden unique transition releasing the gamma lines with 696.5 keV and 2185.7 keV respectively.

[6] Four decays of ^{144}Ce produce one molecule of gaseous O_2 via the reaction: $4CeO_2 \rightarrow 2Pr_2O_3 + O_2$

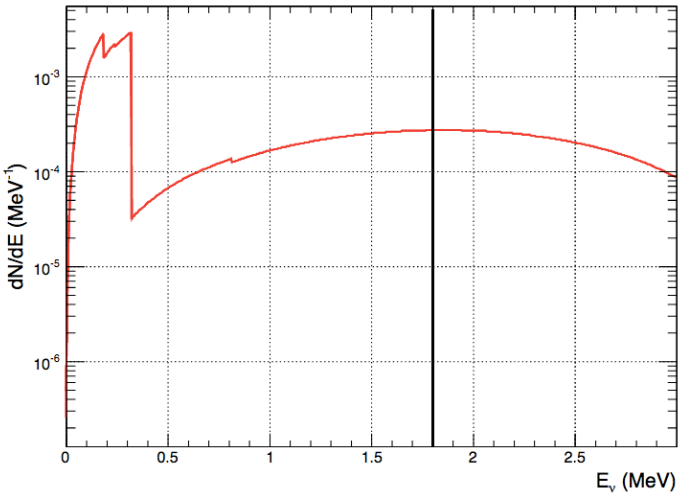


Figure 2.8: Modeled ^{144}Ce-^{144}Pr Neutrino Spectrum. The spectrum is modeled in Reference [165] and takes into account all beta branches from the ^{144}Ce and ^{144}Pr decay. The black line indicates the threshold of the IBD detection reaction.

The source has been ordered with the following specifications. First, the activity ranges between 100 and 150 kCi (3.7 and 5.5 PBq). Second, the level of radio impurities needs to be low to ensure, on the one hand, a safe source handling and to meet the radio protection requirements and, on the other hand, to minimize systematic effects on the analysis (for example to allow the determination of the activity). Hence, the ratio of the total thermal power released by controllable radionuclide impurities to the thermal power of ^{144}Ce needs to be smaller than $< 10^{-3}$ W/W. Moreover, the ratio of the activity of impurities that

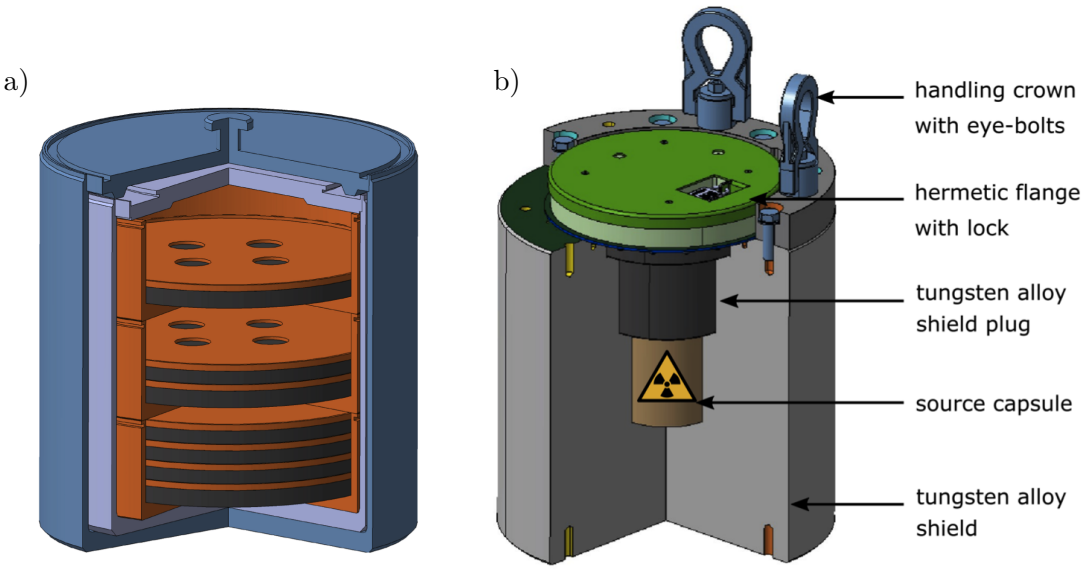


Figure 2.9: Source Capsule and Shielding. The source in form of CeO_2 powder is pressed into a sealed capsule (a) that consists of two stainless steel cylinders (diameter and height of 15 cm) around three individual copper cylinders with additional copper radiator disks. The tungsten-alloy shielding (b) ensures a thickness of at least 19 cm around the capsule. The schemes are taken from [167].

release gammas with an energy higher than 1 MeV with respect to the ^{144}Ce activity is $< 10^{-3}$ Bq/Bq. Due to LNGS restrictions, the maximum neutron flux has to be below 10^5 n/s, such that the activity of the main neutron emitter ^{244}Cm and ^{241}Am is $< 10^{-5}$ Bq/Bq and $< 5 \cdot 10^{-3}$ Bq/Bq with respect to the activity of ^{144}Ce. With the given requirements and an activity of 150 kCi, the gamma and neutron dose is simulated to <8 μSv/h and <5 nSv/h at a distance of 1 m [169]. The level of radio impurities of the ^{144}Ce source is measured by means of alpha, gamma, and mass spectroscopy at CEA Saclay using representative samples of the source provided by PA Mayak. A detailed description of the preparation for these measurements can be found in Reference [168]. Unfortunately, as these specifications could, due to technical problems in the source production, not be achieved, the SOX project was canceled.

2.2.2 Characterization of the Source

The characterization of the antineutrino source is a key ingredient of the experiment in order to predict the expected signature. As the activity directly correlates with the expected event number, it must be determined with high precision. The activity is determined via the power P_{source} emitted by the source divided by the mean energy $\langle E \rangle$ released per ^{144}Ce-^{144}Pr decay. Here, one assumes that all beta- and gamma-rays from the source are absorbed by the thick enough shielding and add to P_{source} which can be determined from calorimetric measurements. $\langle E \rangle$ is the mean energy of all beta- and gamma-rays per decay and depends on the energy spectra of all beta-decay branches of ^{144}Ce and ^{144}Pr which need to be determined precisely. Moreover, the spectral shapes of the two main ^{144}Pr-branches enter another two times in the analysis. The fraction of the neutrino spectrum above the IBD threshold affects, on the one hand, the expected event number, and, on the other hand, the expected energy spectrum in SOX. The experimental effort to measure P_{source} and the spectral shape of the source is summarized in the following.

Power Measurement

The power is measured before and after data taking in the cleanroom close to the source position (see Figure 2.2), where the initial power emitted from a 100-150 kCi source is about 1 kW ($\sim$ 215 W/PBq) [133]. To determine the power with high precision, two redundant thermal calorimeters have been developed. In both measurements, the power produced by the source is transferred to a circulating water system and is determined by

$$P_{\text{source}} = \dot{m} \cdot [h(p, T_{out}) - h(p, T_{in})] + P_{\text{loss}}. \tag{2.11}$$

Here $\dot{m}$ is the water mass flow and $h(p, T)$ denotes the enthalpy of the water as a function of the measured in- and outgoing values of the temperature and the pressure. The main difference between the two setups is the water circuit: in the one setup (CEA) the whole source and shielding are immersed in a water vessel, whereas in the second case (TUM-Genova) a water line circulates in a copper heat exchanger which encompasses the shielding. Heat losses (P_{loss}) through convection are minimized by operating the calorimeters in a vacuum tank. Moreover, thermal radiation and conduction are reduced using super insulation foils and a suspension structure with low thermal conductivity. A sketch of the TUM-Genova calorimeter is shown in Figure 2.10. With this setup an accuracy of 0.02 % on the power measurement could be achieved exceeding the design goal of 1 %. For a detailed description see References [167, 170].

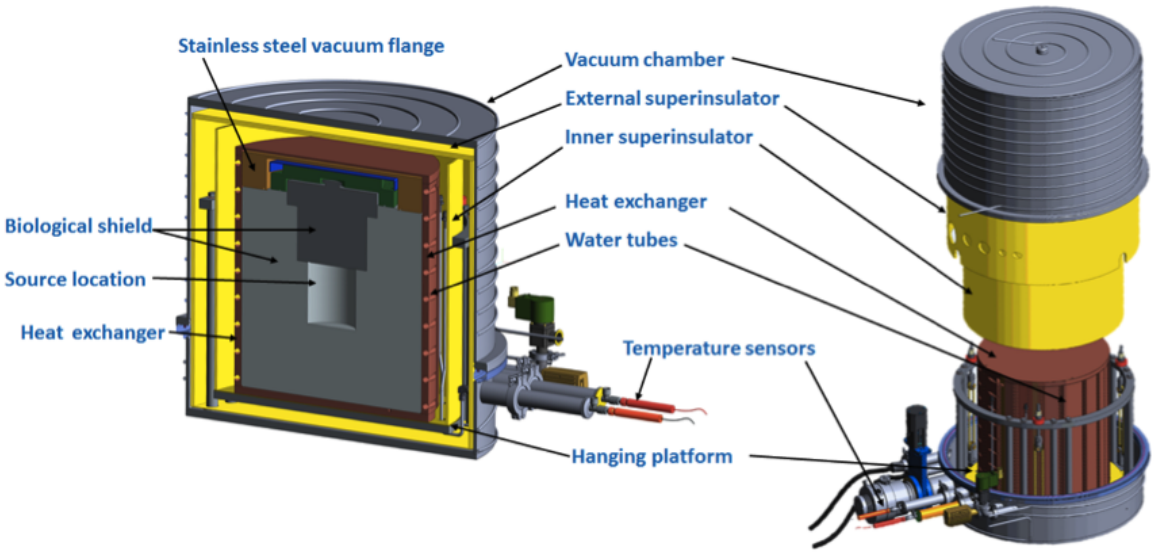


Figure 2.10: Crosssection of the TUM-Genova Calorimeter. The calorimeter measures the total power released by the source by monitoring the mass flow, temperature, and pressure of the circulating water in the copper heat exchanger. To minimize heat losses, the heat exchanger is suspended on a hanging platform via ropes with a low thermal conductivity. Further, the calorimeter contains two layers of super insulator foils and is operated in a vacuum tank. The figure is taken from [167].

Spectral Shape Measurement

To reach the expected sensitivity of the experiment, the spectral shapes of all neutrino branches need to be known. However, the most important decay branch is the main branch of ^{144}Pr which is a 0- to 0+ transition and hence classified as a unique first forbidden decay. Such a type of decay has large theoretical uncertainties and the last measurements which were performed decades ago disagree up to 10 % [167].
The goal is to determine the shape factor[7] with an absolute precision better than 0.03. Thus, several experimental setups have been developed in order to reach this accuracy. The concept of all measurements is similar: the energy of the electrons is measured and the neutrino energy spectrum is obtained as the complementary one. As one is especially interested in the part above the IBD threshold, the detectors need to be sensitive to small electron energies and have a small energy threshold. Thus, dead layers of the detector need to be minimized. To avoid backscattering and thus spectral distortions, several setups were designed with a 4Π-acceptance. Moreover, as the lifetime of ^{144}Pr is short and the separation of ^{144}Pr from ^{144}Ce is challenging, the combined spectrum of ^{144}Pr and ^{144}Ce has to be typically measured and the main branch of ^{144}Pr has to be extracted. In addition, the gamma lines that are produced with about $\sim 2\,\%$ in the ^{144}Pr decay can alter the spectrum. A second detector running in coincidence that is responsible for gamma tagging can be used to select and measure a specific decay branch into an excited state. On the other hand, a second detector which is running in coincidence can be used to suppress background and to veto the gamma rays produced from the source. The so-called TUM spectrometer uses such a technique which has been worked on in this thesis together with Simon Appel and will be described in more detail in the following. A complete list of the spectral measurements can be found in Reference [167].

[7] The energy spectrum can be expressed as a function of the shape factor b (see Chapter 7 and Appendix A.1).

Spectral Measurements with the TUM Spectrometer The TUM spectrometer consists of two detectors: a multi-wire chamber (MWC) and a plastic scintillator with a connected PMT. The density of the gas (CF_4, 1 bar) in the MWC is so low that gammas typically do not interact within the MWC volume and provide no signal. On the contrary, electrons do interact and, when operated in coincidence, gammas can be vetoed with an efficiency larger than 99.5 %. Therefore, the measured spectrum with the plastic scintillator plus the connected PMT can be considered as a pure electron spectrum.

The TUM spectrometer is sketched in Figure 2.11. The setup was originally developed to measure the spectrum of the fission products of ^{238}U [171, 172] and was designed for an energy region of 2 MeV-7.5 MeV. During the author's Master's thesis [161], the setup was optimized in order to become more sensitive to the low energy region. Therefore, the source position was moved directly above the MWC and the main path length of the electrons in the gas was reduced from ~ 10 cm to ~ 1 cm. A picture of the source position can be seen in Figure 2.12b. A detailed description of the detector setup and performance can be found in References [161, 171, 172].

Preliminary measurements were performed at CEA, Saclay, where calibration sources and a ^{144}Ce sample were produced by the Laboratoire National Henri Becquerel. A picture of the ^{144}Ce sample can be seen in Figure 2.12a. The radioactive source is a small drop of approximately 2 mm × 2 mm × 40 µm and is sealed in between two 18 µm thick mylar foils. One critical point of the samples is, that the volume of the radioactive drop can vary for each source. This can affect the backscattering of the electrons inside the source

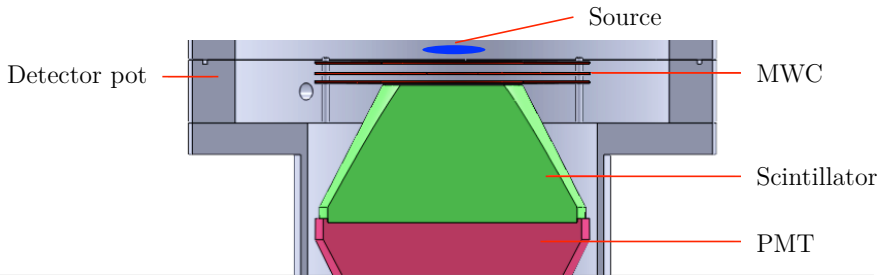


Figure 2.11: Sketch of the TUM Spectrometer. The spectrometer consists of two detectors: a plastic scintillator (BC404) attached to a PMT measures the energy spectrum of the source and a MWC vetoes gamma rays emitted by the source. The source is placed on top of the MWC with a distance of ~ 1 cm.

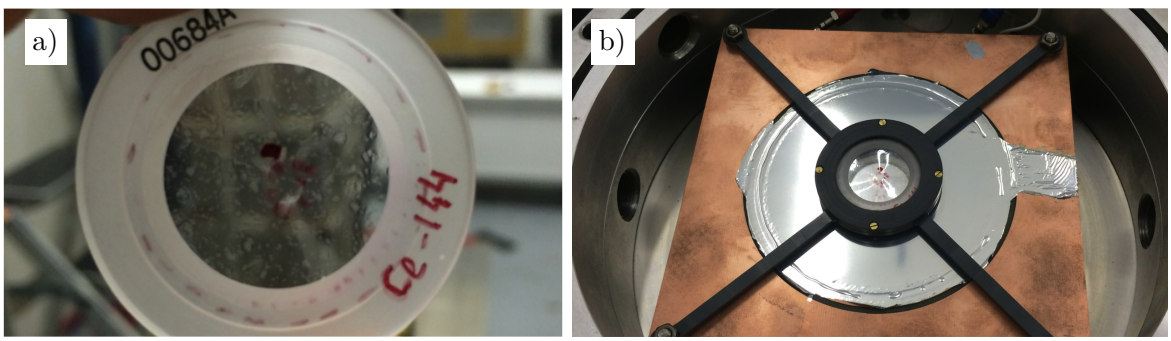


Figure 2.12: Picture of ^{144}Ce Sample and Sample Holder of the TUM Spectrometer. a) The ^{144}Ce source is a small radioactive active area of ~ 2 mm diameter and sealed between two mylar foils. b) The source is mounted in the source holder which is placed above the MWC.

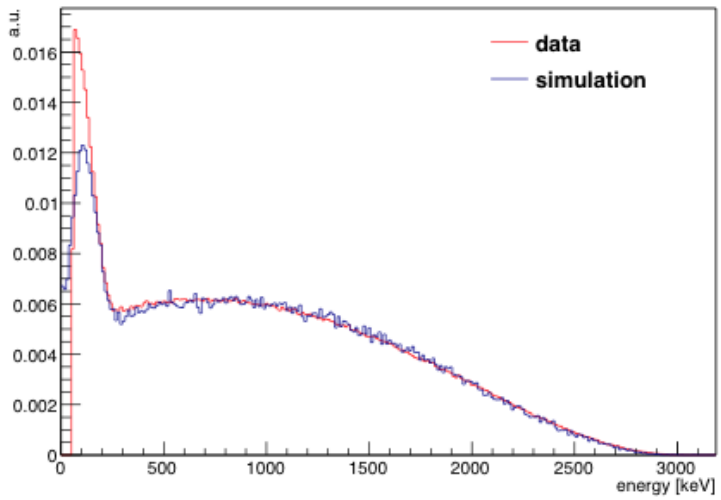


Figure 2.13: Preliminary Measurement of the ^{144}Ce Spectrum. The measurement obtained from the TUM spectrometer is shown in blue. The Geant4-based simulation is shown in red and takes into account all beta transitions. For the main ^{144}Pr branch the pure allowed beta spectrum is assumed. The discrepancy at small energies can be not explained.

and needs to be taken into account in the analysis. While the main results are described in the following, a complete description of the calibration measurements can be found in Appendix A.1. Using mono energetic conversion electrons at 624 keV from a ^{137}Cs source, an energy resolution of 6 % at this energy can be determined. Further, the energy threshold of the spectrometer was determined to about 150 keV. Thanks to the coincident measurement modus, the background for the measurement is small: the darknoise has a rate of 0.3 Hz which corresponds to 0.5 % of the activity of the ^{144}Ce sample ($\sim$ 60 Hz). The rate of accidental coincidences was determined to be $9.5 \cdot 10^{-3}$ times smaller than the rate of the ^{144}Ce sample. The measured energy spectrum of the ^{144}Ce-^{144}Pr sample can be seen in Figure 2.13 together with a MC simulation based on Geant4 [173]. While the simulation shows good agreement for high energies, the data observes more low energy events (<300 keV) than expected from the MC simulations. The discrepancy between data and simulation could not be resolved. The uncertainty of the source dimensions was studied by MC simulations and cannot cause such a discrepancy. Further, the same discrepancy is observed for all calibration sources (^{90}Sr, ^{36}Cl, ^{137}Cs) (see Appendix A.1) that do almost not emit gamma-lines. Hence, the additional events in the data can also not be explained from scattered gamma-rays. In contrast, the simulated energy deposition in the MWC agrees well with the measured spectrum (similar to a Landau distribution) and the mean deposited energy of $\sim$ 7 keV matches the expectation for minimizing ionizing particles [174] (see Appendix A.1).

It was also observed that the energy spectra are stable during the same data-taking run but are not reproducible, especially in the low energy region for two different runs, i.e. when the spectrometer is opened and closed between two runs. The instability can be seen in both detectors, however, it is not present in the electronics of the data acquisition chain that was tested independently with a pulser. The instability could not be explained but is most likely due to the relative placement of the source between the MWC and the plastic scintillator. Any deviation from a cylindrical setup in the source, the source holder, or the two detectors could affect the energy depositions in the detector at low energies. Due to the cancelation of the SOX project, the investigation of the systematic effects and the discrepancies between

simulation and data was not investigated further.

2.2.3 Signature of Sterile Neutrinos in SOX

The expected signature for the SOX experiment as a function of the ratio of the reconstructed distance and energy (L/E) can be seen in Figure 2.14. The spectral shape comes from the individual spectral shapes of the distance and energy spectrum. While the distance spectrum is given by the geometrical relation of the source and the detector, the energy spectrum is mainly given by the product of the neutrino spectrum and the cross section for the IBD (see Section 4). For a measurement time of 1.5 y, approximately 10^4 events are expected for SOX. In the case of a light sterile neutrino, the electron antineutrino oscillates into the sterile neutrino and the L/E spectrum is altered by the disappearance oscillation probability given in Equation 1.10. Due to the small size of the source and the ability to reconstruct the position and the energy, an oscillation pattern can be observed given a sterile neutrino mass of $\sim 1\,\text{eV}$. Such an oscillation signature, which is especially visible in the ratio of the oscillated (sterile neutrino) to un-oscillated (no sterile neutrino) L/E spectrum (see lower panel of Figure 2.14), would be a smoking gun signature for the existence of a sterile neutrino.

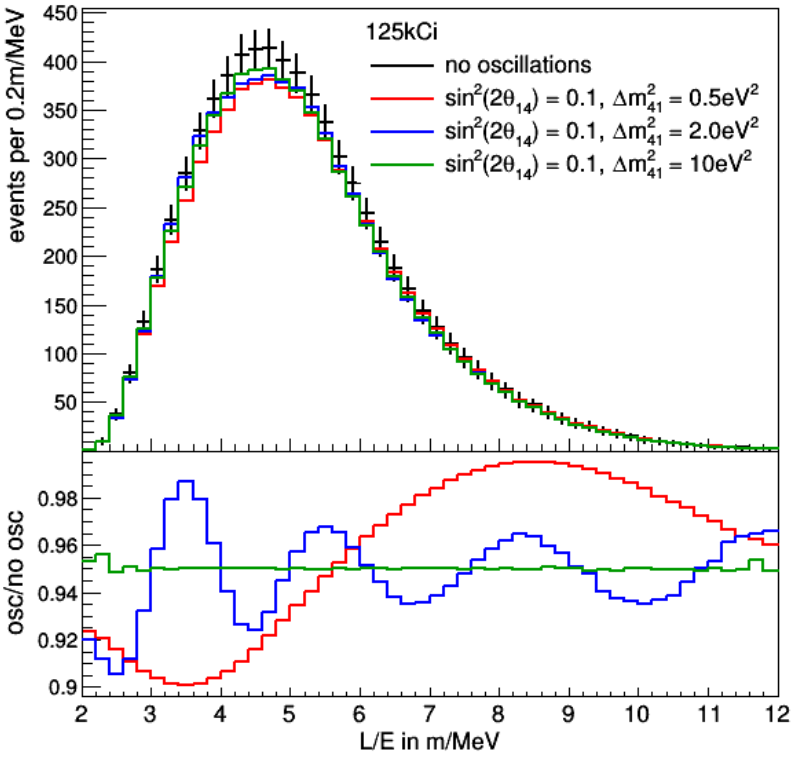


Figure 2.14: Expected Signature in SOX The signature is shown as a function of distance (L) over reconstructed energy (E) for a 125 kCi ^{144}Ce source and different sterile neutrino masses. On the top the total event number for a measurement time of 1.5 y is shown. On the bottom the ratio of the event numbers of oscillation to no-oscillation hypothesis is depicted. An oscillation signature is visible for $\Delta m^2_{41} \sim 1\,\text{eV}^2$. For larger masses ($\sim 10\,\text{eV}^2$) the detector resolution is worse than the oscillation lengths and produces an averaged deficit in the event number. The oscillations are damped for larger L/E values because of the worsening of the energy resolution for smaller energies. The observation of such oscillations would be a smoking gun signature for sterile neutrinos.

With a larger sterile neutrino mass ($\gtrsim$10 eV), the oscillation length becomes smaller than the detector resolution and an average rate deficit can be observed. In order to observe the averaged rate deficit, a precise knowledge of the activity as well as the detector volume is mandatory. In contrast, if the mass of the sterile neutrino becomes too small ($\lesssim$0.1 eV), the oscillation length becomes larger than the distance of the source to the detector and the disappearance signature vanishes. The background is not indicated in Figure 2.14, as it is too small to be visible. Indeed, due to the IBD reaction, the expected background is negligible in comparison to the signal event number (see Chapter 6). Thus, SOX would be a background-free experiment.

In summary, the SOX project could observe a sterile neutrino with a mass in the eV-scale and with the artificial ^{144}Ce-^{144}Pr source it is a complementary experiment in the sterile neutrino search program. Unfortunately, the SOX project was canceled during the work of this thesis. Nevertheless, the first part of this thesis covers the analysis preparation for the SOX experiment, as this work is still of interest for current and future short-baseline experiments. This work extends the previous existing feasibility studies and simplified analyses for SOX. First, Chapter 4 describes the developed software to generate the expected signal and perform a sterile neutrino analysis. Second, the expected signal with the detector response and detection efficiency is discussed in Chapter 5. The background for the SOX event sample is investigated using existing Borexino data in Chapter 6. Systematic effects and their impact on the sensitivity are studied in Chapter 7. The first part ends with the study of a hardware modification, namely the trigger system, in order to minimize the systematical error on the detection efficiency (Chapter 8).

3 Questions in Data Analysis

This chapter introduces common data analysis methods applied in the high-energy physics community, where the following questions are addressed: What are the most likely values for my parameters of interest? (Section 3.1); Which parameter values are compatible with my data? (Section 3.2); How significant is the signal? (Section 3.3); Which result can we expect for the experiment? (Section 3.4); Is the data compatible with the model? (Section 3.5). This chapter presents the basis for the following chapters and applies the notations of the methods directly to the field of short-baseline experiments. In Section 3.6 an overview of the different methods applied in current short-baseline experiments is given, which motivates the work for the second part of this thesis. As only Frequentist-based methods are currently used in short-baseline experiments, Bayesian methods are not discussed.

3.1 What are the Most Likely Values for the Parameters of Interest?

In experimental physics, one wants to determine the unknown parameters of the model out of the observed data. This can be done by comparing the observed with the expected data using a likelihood function. These ingredients are described shortly in the following.

3.1.1 Observed and Expected Data

Observed Data

Short-baseline experiments reconstruct the distance and energy of their events (see Chapter 1.5.5):

$$(\boldsymbol{L}_{\text{rec}}, \boldsymbol{E}_{\text{rec}}) = \{(L_{\text{rec}}, E_{\text{rec}})_1, (L_{\text{rec}}, E_{\text{rec}})_2, ..., (L_{\text{rec}}, E_{\text{rec}})_n\}, \tag{3.1}$$

where n is the total number of observed events. Since n is typically large, the data is grouped into distance and energy bins

$$\boldsymbol{N}^{\text{obs}} = \{N_{11}^{\text{obs}}, N_{12}^{\text{obs}}, N_{13}^{\text{obs}}, ..., N_{kl}^{\text{obs}}\}, \tag{3.2}$$

where $N_{ij}^{\text{obs}} = \sum_{m=1}^{n}\{(L_{\text{rec}}, E_{\text{rec}})_m : L_{\text{rec}} \in \text{bin}_i \cup E_{\text{rec}} \in \text{bin}_j\}$ is the observed number of events in the i-th distance and j-th energy bin. The number of bins depends on the range and the resolution of the experiments.

Expected Data and Model

The expected data is a function of the unknown parameters and, as a hypothesis is a statement about the population parameters [175], a function of the hypothesis. One distinguishes between the parameters of interest and nuisance parameters [176]. The parameters of interest

are the new parameters of the physical model. Such parameters of interest are in a sterile neutrino experiment assuming a (3+1) model the mixing angle $\sin^2(2\theta)$ and the mass splitting Δm^2 (see Chapter 1.5). Instead, nuisance parameters are unknown experimental parameters that affect the expected data but are not of primary interest, as for example the neutrino flux or the cross section. In the following, $\boldsymbol{\eta}$ denotes the vector of nuisance parameters. A hypothesis is hence a function of the oscillation parameters and the nuisance parameters:

$$\mathrm{H}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}). \tag{3.3}$$

The expected event number in each distance-energy bin ij is the sum of the expected neutrino events depending on $\sin^2(2\theta)$ and Δm^2 and the expected background events:

$$\begin{aligned} N_{ij}^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) = N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) \cdot pdf_{ij}^S(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) \\ + N_B^{\mathrm{exp}}(\boldsymbol{\eta}) \cdot pdf_{ij}^B(\boldsymbol{\eta}), \end{aligned} \tag{3.4}$$

where $N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})$ and $N_B^{\mathrm{exp}}(\boldsymbol{\eta})$ are the total expected event numbers of neutrino signal and background events. The probability distribution functions $pdf_{ij}^S(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})$ and $pdf_{ij}^B(\boldsymbol{\eta})$ are normalized and describe the spectral shapes of neutrino and background events. To this end, each ingredient can depend on nuisance parameters.
The expected number of neutrino events is the product of the integral oscillation probability $P_{\mathrm{int}}(\sin^2(2\theta), \Delta m^2)$ and the expected number of events without sterile neutrinos:

$$N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) = P_{\mathrm{int}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) \cdot N_S(\boldsymbol{\eta}). \tag{3.5}$$

Here, $P_{\mathrm{int}}(\sin^2(2\theta), \Delta m^2)$ is integrated over the L_{rec}- and E_{rec}- parameter range. $N_S(\boldsymbol{\eta})$ incorporates further the properties from the source and the detector, as for example the activity of the neutrino source, the cross section, or the detection efficiency.
The probability distribution functions and the expected event numbers have to be typically determined via Monte Carlo (MC) simulations that take care of the particle interaction, the propagation and the detection-related processes (see Chapter 4.1). A way to connect the observed to the expected data can be obtained using a likelihood function which is described in the following.

3.1.2 Likelihood Fit Function and Maximum Likelihood Estimator

The observed data is compared to the expected data using a likelihood function.[1] The probability to observe N^{obs} events for N^{exp} expected events is given by a Poisson distribution: $\mathcal{P}(N^{\mathrm{obs}}|N^{\mathrm{exp}})$. The likelihood function is defined by the product of the independent Poissonian probabilities for each bin [176]:

$$\mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) = \prod_{ij} \mathcal{P}(N_{ij}^{\mathrm{obs}}|N_{ij}^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})). \tag{3.6}$$

Additional auxiliary measurements ($\boldsymbol{\eta}^{\mathrm{obs}}$) can constrain the value of the nuisance parameters and therefore increase the sensitivity on the parameters of interest. For example, the activity of the neutrino source can be determined via a calibration measurement or the background rate can be determined via source-off measurements. The probability to observe η^{obs}, assuming the true parameter η and an uncertainty of σ_η in the auxiliary measurement, is given by a Normal distribution $\mathcal{N}(\eta^{\mathrm{obs}}|\eta, \sigma_\eta^2)$, when the value of the nuisance parameter itself is

[1] Given large event numbers per bin, a χ^2-function can be used (see Reference [176]).

large. Since the measurements are independent of each other, a joint likelihood function can be built [176]:

$$\mathcal{L}_{joint}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) = \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) \cdot \prod_{\eta_i} \mathcal{N}(\eta_i^{\text{obs}} | \eta_i, \sigma_{\eta_i}^2). \tag{3.7}$$

Given a specific data set, the most likely values of the free parameters are given by the maximum of the likelihood function

$$\mathcal{L}_{joint}(\widehat{\sin^2(2\theta)}, \widehat{\Delta m^2}, \widehat{\boldsymbol{\eta}}) = \max_{\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}} \mathcal{L}_{joint}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}) \tag{3.8}$$

and are called maximum likelihood estimators: $\widehat{\sin^2(2\theta)}$, $\widehat{\Delta m^2}$, and $\widehat{\boldsymbol{\eta}} = \{\widehat{\eta}_1, \widehat{\eta}_2, ...\}$ [176]. From the computational point of view, it is easier to minimize the negative log-likelihood function (NLL): $-2 \ln \mathcal{L}_{joint}(\widehat{\sin^2(2\theta)}, \widehat{\Delta m^2}, \widehat{\boldsymbol{\eta}})$. Maximum likelihood estimators have some desired properties, as they are for example asymptotically unbiased, meaning the expectation value of the maximum likelihood estimator corresponds to the true value [176]. The distribution of the maximum likelihood estimators in sterile neutrino experiments will be discussed in detail in Chapter 11.

3.2 Which Parameter Values are Compatible with the Data?

Additionally to the best-fit values, it is important to state the errors of the measured parameters. This is done by the construction of a confidence region. If the errors reach up or down to the border of the allowed parameter space, the confidence region naturally results in a limit. A confidence region contains by definition the true (but unknown) parameters with a certain probability (e.g. 95 %). Moreover, a confidence region depends on the specific data and will change for a different data sample. Nevertheless, if an experiment would be repeated many times, the true value would be contained by the confidence region with the desired probability.
This non-trivial property (also called coverage) can be obtained by the inversion of a set of hypothesis tests. To do so, the parameter of interest space is scanned with a fine raster. For each parameter-raster point a hypothesis test is performed and the tested hypothesis is accepted or rejected. The set of accepted hypotheses corresponds then to the confidence region which is illustrated in Figure 3.1 [175]. This procedure will be discussed in more detail in the following.

3.2.1 Hypothesis Test and Test Statistic

A hypothesis test defines the rule of accepting or rejecting a hypothesis in favor of an alternative hypothesis. In general, one differentiates between a simple hypothesis, where all parameters are fixed $\text{H}(\sin^2(2\theta) = x, \Delta m^2 = y)$ and a composite hypothesis, e.g. a hypothesis with free parameters $\text{H}(\sin^2(2\theta), \Delta m^2)$. The hypothesis test is illustrated in Figure 3.2a and consists of the following ingredients [175]:

- H_0: null hypothesis (tested hypothesis). H_0 is a simple hypothesis with the specific values $\sin^2(2\theta) = x$ and $\Delta m^2 = y$, where x and y are values from the defined parameter space $0 \leq x \leq 1$ and $0 \leq y$. In the presence of nuisance parameters, the nuisance parameters are not fixed and the hypothesis is not exactly simple: $\text{H}_0 : \{\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta} : \sin^2(2\theta) = x, \Delta m^2 = y, \boldsymbol{\eta} \in \boldsymbol{\eta}_{\text{range}}\}$, where $\boldsymbol{\eta}_{\text{range}}$ is vector of the defined parameter space for the nuisance parameters.

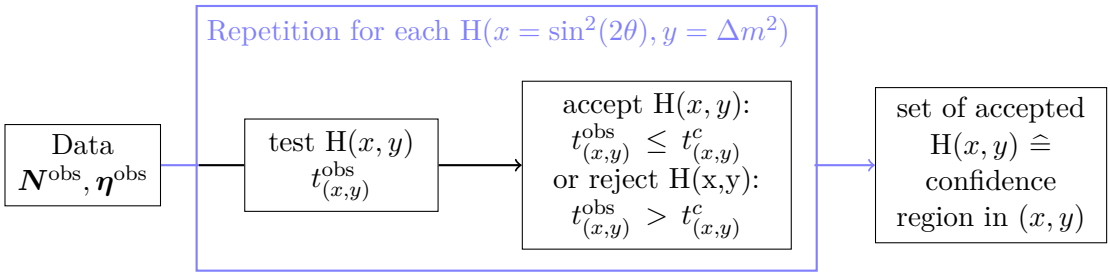


Figure 3.1: Construction of a Two-Dimensional Confidence Region The confidence region in (x, y) is the set of accepted hypotheses H(x, y) in a hypothesis test. To do so, a fine grid in the parameter space of (x, y) is defined and each hypothesis H(x, y) is tested using the test statistic $t_{(x,y)}$. The multidimensional observed data $\boldsymbol{N}^{\text{obs}}$ and possibly existing auxiliary measurements $\boldsymbol{\eta}^{\text{obs}}$ can in that way be transformed to a one-dimensional number $t^{\text{obs}}_{(x,y)}$. Depending on how the observed test statistic value compares to the critical value $t^{c}_{(x,y)}$, the tested hypothesis is accepted or rejected.

- H_1: alternative hypothesis. H_1 is typically the complementary hypothesis to H_0, namely H_1: $\{\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta} : 0 \leq \sin^2(2\theta) \leq 1, \Delta m^2 \geq 0, \sin^2(2\theta) \neq x, \Delta m^2 \neq y, \boldsymbol{\eta} \in \boldsymbol{\eta}_{\text{range}}\}$ and hence a composite hypothesis. The union of H_0 and H_1 is the full physically allowed parameter space.

- t: outcome of experiment. Typically a test statistic is defined that transforms the multidimensional data to a real-valued and one-dimensional number. A standard test statistic is the profile likelihood ratio that tests for a given data sample H_0 versus H_1:

$$t_{(x,y)} = -2 \ln \frac{\max\limits_{\boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta) = x, \Delta m^2 = y, \boldsymbol{\eta})}{\max\limits_{\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})}. \tag{3.9}$$

  The nominator is the maximum of the likelihood function (Equation 3.7) for the null hypothesis and the denominator the maximum of the likelihood in the full parameter space (union of H_0 and H_1). Small values of the profile likelihood ratio indicate good agreement between the data and H_0, large values instead disagreement. In the case, when H_0 and H_1 are two simple hypotheses, the Neyman-Pearson lemma states that a likelihood ratio test statistic maximizes the power of the test and is therefore the optimal test statistic [176]. Another advantage of the profile likelihood ratio is that it is approximately independent of the nuisance parameters [176].

- $f\left(t_{(x,y)} | (x_{\text{true}}, y_{\text{true}})\right)$: probability distribution function of $t_{(x,y)}$ when the hypothesis $\text{H}(\sin^2(2\theta) = x_{\text{true}}, \Delta m^2 = y_{\text{true}})$ is true. When $x = x_{\text{true}}$ and $y = y_{\text{true}}$, one calls the probability distribution the null distribution, otherwise alternative distribution.

- α: test size, probability to reject H_0, when it is actually true. This is also known as the type-I error. The test size is defined previously to the analysis and chosen to be small (e.g. 5 %). This is particularly important for testing the no-oscillation hypothesis, since rejecting it would result into a discovery (α is then reduced even further, for example $\alpha = 0.27\,\%$ for a 3σ level).

- $t^{\text{obs}}_{(x,y)}$: test statistic value for an observed data when the hypothesis H(x, y) is tested.

- $t^{c}_{(x,y)}$: critical value. The critical value is calculated from the null distribution and the test size: $\int_{t^{c}_{(x,y)}}^{\infty} f\left(t_{(x,y)} | (x, y)\right) = \alpha$.

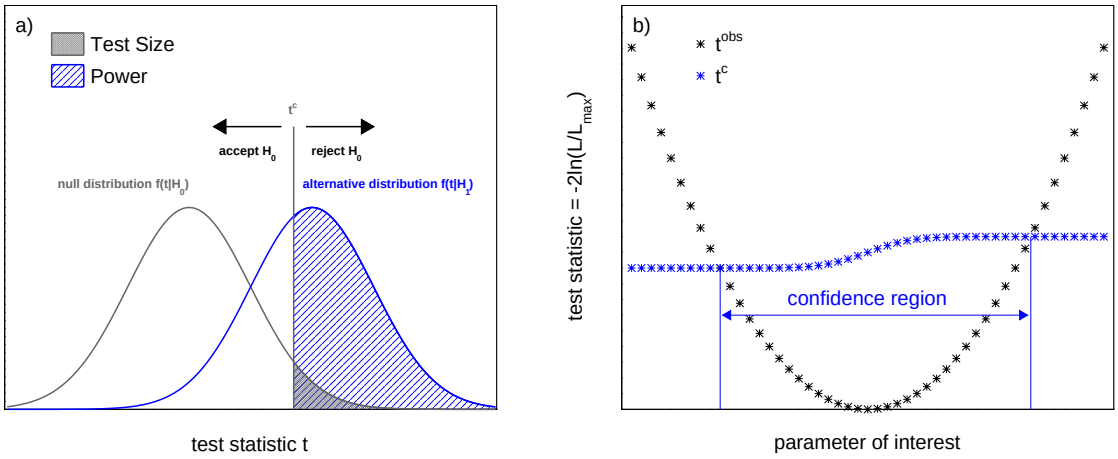


Figure 3.2: Connection of Hypothesis Test and Likelihood Function Illustration of a a) hypothesis test: In a hypothesis test, the null hypothesis H_0 is tested and either accepted or rejected and the alternative hypothesis H_1 is accepted. The null (alternative) distribution is the probability distribution of the outcome or the test statistic t when H_0 (H_1) is true. The critical value t^c defines the decision rule and is chosen so that the probability to reject the null hypothesis even though it is true is smaller than the test size. The power of a test is defined as the integral under the alternative distribution above t^c. b) test statistic as a function of the parameter of interest for a specific data sample. A fine grid in the parameter of interest space is defined and for each hypothesis a hypothesis test is performed (H_0 changes for each test) and a critical value exists (blue). As the test statistic is here defined via the negative log-likelihood function normalized to its absolute minimum (profile likelihood ratio), the confidence region corresponds to that parameter region with a profile likelihood value (black) smaller than the critical value.

- decision rule: the tested hypothesis $H(x, y)$ is rejected for $t^{obs}_{(x,y)} > t^c_{(x,y)}$, otherwise accepted. Similarly, one can compute the observed p-value: $p^{obs}_{(x,y)} = \int_{t^{obs}_{(x,y)}}^{\infty} f\left(t_{(x,y)}|(x, y)\right)$ and reject $H(x, y)$ for $p^{obs}_{(x,y)} < \alpha$.

- power: probability to accept the alternative hypothesis $H(x_{true}, y_{true})$, when it is actually true. The power is computed via the integral of the alternative distribution: $\int_{t^c_{(x,y)}}^{\infty} f\left(t_{(x,y)}|(x_{true}, y_{true})\right)$ and is the complementary probability of the type-II error (error of false acceptance). The higher the power of a test is, the higher the chance to observe new physics becomes.

3.2.2 Inversion of a Hypothesis Test

To obtain a confidence region in the parameter space, a hypothesis test at each $H(x, y)$ of the previously defined raster is performed. The set of accepted hypotheses can be interpreted as the confidence region. This is possible as the probability that the observed value falls into the acceptance region is 1-α. As this is given for each hypothesis, the true value is accepted with a probability of 1-α. Subsequently, the confidence region contains the true parameter with the desired probability (e.g. 95 % for $\alpha = 0.05$).
The probability, how often the confidence region contains the true value, is called coverage. The coverage must be fulfilled for each hypothesis. Given a coverage larger (smaller) than the nominal value, it is called "overcoverage" ("undercoverage").

With the inversion of the hypothesis test, the properties of the acceptance region are transferred to the confidence region. Further, the space of the alternative hypothesis defines the space of the confidence region, for example a two-dimensional alternative hypothesis will result into a two-dimensional confidence region [175].
Figure 3.2b shows the visualization of the construction of the confidence region, where the connection to the likelihood function is depicted. This approach is equivalent to a Neyman-Construction [176] with a likelihood-ratio test statistic that is also known as the Feldman-Cousins approach [177].

3.2.3 Probability Distributions of the Test Statistic and Wilks' Theorem

According to Wilks' theorem [178], the null distributions of a profile likelihood test statistic follow in the large sample limit asymptotically a chi-square distribution with the degrees of freedom corresponding to the difference of the number of free parameters in H_0 and H_1. To this end, the following regularity conditions have to be fulfilled [179]:

- H_0 is true;
- the hypotheses are nested: H_0 is contained in the hypotheses of the denominator in the likelihood ratio;
- the parameters of interest of H_0 are uniquely defined and are not at the borders of their allowed parameter space;
- there are enough data and asymptotic approximations are valid.

If Wilks' theorem can be applied, the critical values are the ones listed in Table 3.1. As a key step in Wilks' theorem, the maximum-likelihood estimators need to follow a normal distribution. Obviously, this cannot be achieved for the oscillation parameter $\sin^2(2\theta)$ at the parameter bound $\sin^2(2\theta) = 0$. The validity of Wilks' theorem in a sterile neutrino analysis is tested in Chapter 11.
When the regularity conditions are not met, a MC construction of the probability distribution has to be conducted. Figure 3.3 sketches the MC construction procedure. In order to construct the probability distribution $f\left(t_{(x,y)}|(x_{\text{true}}, y_{\text{true}})\right)$, pseudo-data samples from the assumed true hypothesis are generated. In this step, for each bin ij N_{ij}^{obs} are randomly drawn from a Poissonian distribution with the expectation value N_{ij}^{exp}. In addition, the auxiliary measurements η_i^{obs} are drawn from a Gaussian distribution with mean η_i^{exp} and width σ_{η_i}. For the pseudo data samples the respective value of the test statistic is determined. From the obtained distribution of the null hypothesis, the critical value can be calculated. As the critical value has to be determined for each tested hypothesis, the MC construction has to be performed for each hypothesis. Such a MC construction has by construction the proper coverage but can be time-consuming.

Table 3.1: Critical Values based on Wilks' Theorem The first column quotes the degrees of freedom of the chi-square distribution, whereas the second column gives the critical value for a 95 % CL. The values are taken from [1].

degree of freedom	critical value
1	3.84
2	5.99

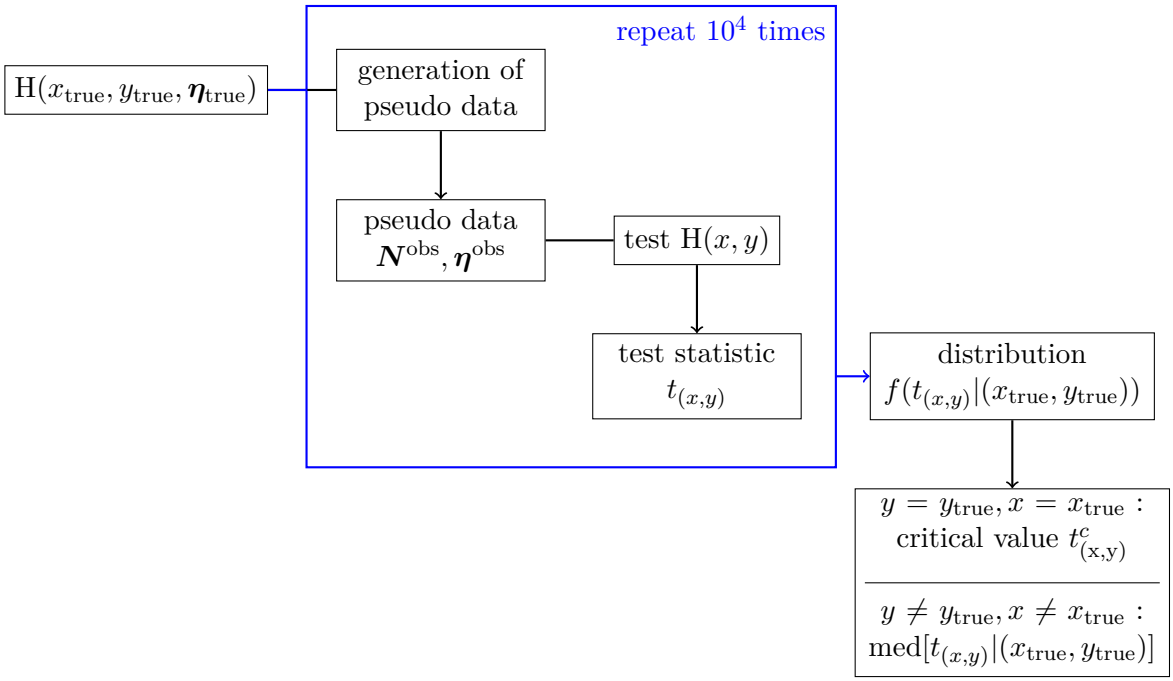


Figure 3.3: Generation of the Test Statistic Probability Distribution using MC Techniques The scheme visualizes how the null and the alternative distributions can be approximated assuming $\mathrm{H}(x, y)$ is tested and $\mathrm{H}(x_{\mathrm{true}}, y_{\mathrm{true}}, \boldsymbol{\eta}_{\mathrm{true}})$ is true. A large number of pseudo data $\boldsymbol{N}^{\mathrm{obs}}$ including pseudo data for the auxiliary measurements $\boldsymbol{\eta}^{\mathrm{obs}}$ is generated from $\mathrm{H}(x_{\mathrm{true}}, y_{\mathrm{true}}, \boldsymbol{\eta}_{\mathrm{true}})$ and the respective test statistic value for $\mathrm{H}(x, y)$ is computed. The critical value and/or the median value of the alternative distribution can then be determined from the distributions.

3.3 How Significant is a Signal?

The significance makes a statement about the compatibility of the data with the no-oscillation hypothesis and is defined as the p-value for the no-oscillation hypothesis:

$$p = \int_{t^{\mathrm{obs}}_{(0,0)}}^{\infty} f(t_{(0,0)}|(0,0))dt_{(0,0)}. \tag{3.10}$$

If the p-value is small, the probability that the no-oscillation hypothesis is true is small and it is likely that one has discovered new physics. To claim a discovery, the significance must be small, for example $p = 0.27\,\%$ for a 3σ level.
There are procedures which allow to approximately estimate the p-value without performing a full MC construction. These procedures are described and tested for the sterile neutrino search in Chapter 11.4.

3.4 Which Result can we Expect for an Experiment?

In addition to the observed confidence region, it is informative to compare it to the expected outcome of the experiment. Hence, statistical fluctuations of the result can be observed. Furthermore, before an experiment is built, it is important to study this expected reach. Moreover, individual experimental parameters can be optimized to achieve the maximal outcome of the experiment. In the following, the concept of the sensitivity and the Asimov data set are introduced.

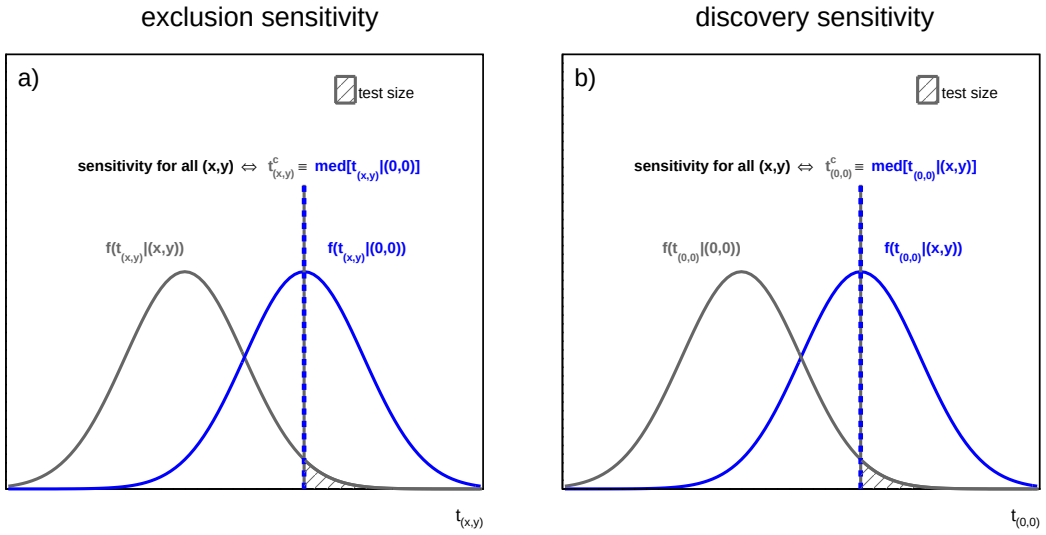


Figure 3.4: Illustration of Exclusion and the Discovery Sensitivity The exclusion sensitivity (a) delimits the parameter space that can be rejected (i.e. test H(x, y)) if the no-oscillation hypothesis is true. The discovery sensitivity (b) defines the parameter space that could be observed if they were true, i.e. the no-oscillation hypothesis is rejected (and tested). Both sensitivities are defined for that set of hypotheses with a power of 50 %.

3.4.1 Sensitivity

The expected outcome is given by the sensitivity. One can distinguish between two sensitivities:

- exclusion sensitivity: The exclusion sensitivity delimits the parameter space that can be excluded (rejected), when the no-oscillation hypothesis is true. The complementary parameter region can also be interpreted as the median confidence region, if the no-oscillation hypothesis is true. The exclusion sensitivity is obtained, when the median of the alternative distribution with $x_{\text{true}} = 0$ and $y_{\text{true}} = 0$ is equal to the critical value.
- discovery sensitivity: The discovery sensitivity indicates the hypotheses that can be discovered given they were true. In this construction, the tested hypothesis is always the no-oscillation hypothesis and the true hypothesis varies. Equivalently, the hypotheses that could be discovered have a median significance that is smaller than the test size.

Figure 3.4 illustrates the two types of sensitivities and their construction. In both cases, the power of the hypothesis tests is 50 % and one needs hence to determine the median of the alternative distributions. On the one hand, one can always perform a MC construction as previously described (see Figure 3.3). On the other hand, Wald [180] and Reference [181] state that the alternative distribution follows asymptotically a non-central chi-square distribution with k degrees of freedom and the non-centrality parameter λ. The degrees of freedom correspond - similar to Wilks' theorem - to the number of extra parameters in the alternative hypothesis in comparison to the null hypothesis. The non-centrality parameter can be estimated using the Asimov data set.

3.4.2 Asimov Data Set

The Asimov data set is the representative data set for a given hypothesis [181]. In other words, the Asimov data set is the data set without statistical fluctuations and corresponds to

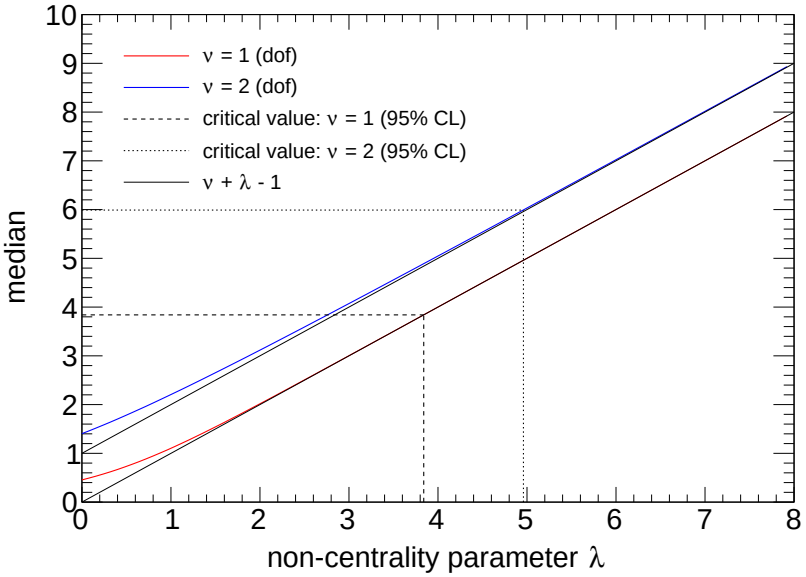


Figure 3.5: Median Value of Non-Central Chisquare Functions The median value is shown for one and two degrees of freedom (ν) as a function of the non-centrality parameter λ. The median values converge to $\lambda + \nu - 1$ (black lines). Thus, if the asymptotic probability distributions described by Wilks and Wald are valid, the sensitivity at the 95 % CL is reached for $\lambda = 3.84$ and $\lambda = 4.95$ for one and two degrees of freedom, respectively.

the expected event number: $\boldsymbol{N}_{\boldsymbol{A}}^{\text{obs}} = \boldsymbol{N}^{\text{exp}}(x_{\text{true}}, y_{\text{true}}, \boldsymbol{\eta}_{\text{true}})$. The non-centrality parameter λ is given by the value of the profile-likelihood ratio of the Asimov data set $A(x_{\text{true}}, y_{\text{true}})$ [181]:

$$t_{(x,y)}^{A(x_{\text{true}},y_{\text{true}})} = -2\ln \frac{\max\limits_{\boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta) = x, \Delta m^2 = y, \boldsymbol{\eta})}{\max\limits_{\sin^2(2\theta),\Delta m^2,\boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})} \tag{3.11}$$

$$= -2\ln \frac{\max\limits_{\boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta) = x, \Delta m^2 = y, \boldsymbol{\eta})}{\mathcal{L}(\sin^2(2\theta) = x_{\text{true}}, \Delta m^2 = y_{\text{true}}, \boldsymbol{\eta} = \boldsymbol{\eta}^{\text{obs}})}, \tag{3.12}$$

where $\boldsymbol{\eta}^{\text{obs}} = \boldsymbol{\eta}_{\text{true}}$. The median of the alternative distribution is of primary interest to predict the sensitivity of an experiment. Figure 3.5 depicts the median as a function of the non-centrality parameter λ for different degrees of freedom ν. The median converges for large non-centrality parameters (region of critical value) to $\lambda + \nu - 1$. Hence, the sensitivity for $\nu = 1$ is given, when the Asimov value of the test statistic corresponds to the critical value of one degree of freedom using Wilks' theorem (see Table 3.1). Further, for $\nu = 2$ the sensitivity is obtained for the hypotheses with $t_{(x,y)}^{A(x_{\text{true}},y_{\text{true}})} = 4.95$.
The maximum-likelihood estimators must converge in the large data sample to a normal distribution to apply this approximation. If the Asimov data set can be used in a sterile neutrino analysis is discussed in Chapter 11.

3.5 Is the Data Compatible with the Fit Model?

An important piece of information is if the assumed model can indeed describe the data. This can be done with a "goodness-of-fit"-test. In the current sterile neutrino searches, one assumes a (3+1) model. However, this model might be incomplete or even wrong from the theoretical and/or experimental point of view. To test the specific model, one

performs a hypothesis test, where the null hypothesis is the (3+1) model and the alternative hypothesis is any other model. To do so, one introduces a flexible alternative hypothesis, in which the event numbers are adjustable for each bin. The alternative hypothesis is given hence by $H_1 : \{\boldsymbol{N}^{\text{exp}}, \boldsymbol{\eta} : \boldsymbol{N}^{\text{exp}} \geq 0, \boldsymbol{\eta} \geq 0\}$ and the null hypothesis by $H_0 : \{\boldsymbol{N}^{\text{exp}} : \boldsymbol{N}^{\text{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})\}$, where $\boldsymbol{N}^{\text{exp}}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})$ is defined in Equation 3.4.
The corresponding profile-likelihood ratio test for a specific data sample with $\boldsymbol{N}^{\text{obs}}$ and $\boldsymbol{\eta}^{\text{obs}}$ is [176]:

$$t_{gof} = -2 \ln \frac{\max\limits_{\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})}{\max\limits_{\boldsymbol{N}^{\text{exp}}, \boldsymbol{\eta}} \mathcal{L}(\boldsymbol{N}^{\text{exp}}, \boldsymbol{\eta})} = -2 \ln \frac{\mathcal{L}(\widehat{\sin^2(2\theta)}, \widehat{\Delta m^2}, \widehat{\boldsymbol{\eta}})}{\mathcal{L}(\boldsymbol{N}^{\text{exp}} = \boldsymbol{N}^{\text{obs}}.\boldsymbol{\eta} = \boldsymbol{\eta}^{\text{obs}})}, \quad (3.13)$$

Here, the nominator is the maximum inside the null hypothesis. The denominator is the maximum inside the alternative hypothesis, which is given when the expected event numbers correspond exactly to the observed ones.
Similarly to a p-value, one can calculate how likely it is when the 3+1 model is true, to observe the calculated t_{gof} or even more extreme values:

$$p_{gof} = \int_{t_{gof}^{\text{obs}}}^{\infty} f(t_{(gof)}|(3+1))dt_{(gof)}. \quad (3.14)$$

According to Wilks' theorem, assuming the (3+1) hypothesis is true and the regularity conditions are fulfilled, the probability distribution $f(t_{(gof)}|(3+1))$ follows asymptotically a chi-square distribution with the degrees of freedom given by the difference of free parameters in H_1 and H_0. When the likelihood function contains auxiliary measurement terms of the nuisance parameters (see Equation 3.7), the degrees of freedom are given by the number of data bins minus the number of the parameters of interest. Instead, when the auxiliary measurement terms are absent in the likelihood function (see Equation 3.6), the alternative hypothesis is given by $H_1 : \{\boldsymbol{N}^{\text{exp}} : \boldsymbol{N}^{\text{exp}} \geq 0\}$. Subsequently, the degrees of freedom are further reduced by the number of nuisance parameters [182].
In comparison to a profile likelihood ratio test, a goodness-of-fit estimate can be already obtained in a chi-square test by the minimum of the chi-square function itself [176]. This is because, in the flexible alternative hypothesis model the minimum chi-square value of the alternative hypothesis will always be zero. The validity of Wilks' theorem in the goodness-of-fit test will be discussed in Chapter 11.5.

3.6 Methods and Results of Current Short-Baseline Experiments in 2018

Some of the current short-baseline experiments have already published first results. Table 3.2 lists the applied methods and results of these experiments. Note that only experiments where the oscillation probability is given by the two-flavor oscillation formula (see Equation 1.10) are selected. MiniBooNE and Neutrino-4 report a strong positive signal with a significance of 6×10^{-7} and 3.5σ, respectively. However, their best-fit values disagree with each other and the residual experiments cannot observe any signal and set hence exclusion limits on the oscillation parameter space. The current results and how they compare to each other are discussed in Chapter 13.
The table shows that the experiments use different statistical analysis approaches which prevent from the direct comparison of their results. On the one hand, the test statistic

Table 3.2: Results of Current Short-Baseline Experiments at the End of 2018. The names of the experiments are listed in the first column. The second column gives the name of the test statistic and the performed scan, where LR stands for likelihood ratio. The best-fit values and the respective goodness of fit (gof) are given in the third to fifth column. Note that the mixing angle of the appearance experiment MiniBooNE is not directly comparable to the disappearance experiments (see Equation 1.12). The significance is quoted in column six, where Neutrino-4 transfers the p-value to the corresponding σ-level. Whether the probability distributions are constructed with Monte Carlo techniques is indicated in column seven.

experiment	test statistic	best-fit values		gof	significance	MC
		$\sin^2(2\theta)$	$\Delta m^2[\text{eV}^2]$	$\Delta\chi^2$/dof		
MiniBooNE [120]	LR (global)	0.92	0.041	19.4/15.6	6×10^{-7}	yes
DANSS [126]	Gaussian CL_S	0.05	1.4	21.9/24	–	no
NEOS [71]	$\Delta\chi^2$ (raster)	0.05	1.73	57.5/59	0.22	yes
Neutrino-4 [127]	$\Delta\chi^2$ (global)	0.38	7.26	17/17	3.5σ	no
PROSPECT [128]	$\Delta\chi^2$ (global)	0.35	0.5	57.9/78	0.58	yes
STEREO [130]	$\Delta\chi^2$ (raster)	–	–	–	0.34	yes

differs and the way how the confidence regions are constructed, and on the other hand, the probability distributions functions of the test statistics are sometimes constructed with MC techniques and sometimes not. In addition, not all experiments report the same values and there is no consistent way in the presentation of the results.
It is worth mentioning that the reported best-fit values are all different from the no-oscillation hypothesis. As it will be shown in Chapter 9, this is an intrinsic feature of the short-baseline experiments and complicates the statistical analysis. This complication is one reason for the different choice of the statistical approaches of the experiments.

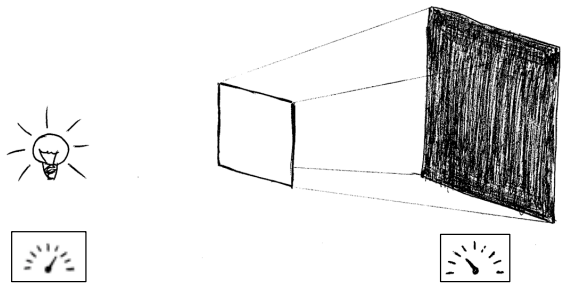
Motivated by this incoherent situation, the second part of this thesis discusses the statistical methods and issues in short-baseline experiments. The previous work from Reference [177] and [183] is extended for a large set of experimental parameters (i.e. disappearance and appearance experiments in a rate, shape and rate+shape analysis). Moreover, discovery and exclusion scenarios are studied. Chapter 9 starts with the statistical issues in short-baseline experiments including sensitivities, confidence regions, and maximum likelihood estimators. Chapter 10 compares the currently-applied test statistics and their interpretations. The validity of the asymptotic functions for the probability distributions of the test statistic is investigated in Chapter 11.

Part I

Data Analysis for the SOX Project

How to observe shadow?

Shadow can be measured **indirectly** by the lack of light



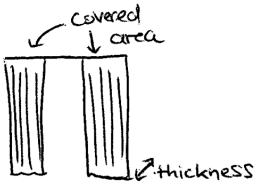
In a similar way sterile neutrinos can be observed:

 = **neutrino source**

 = **neutrino oscillations**

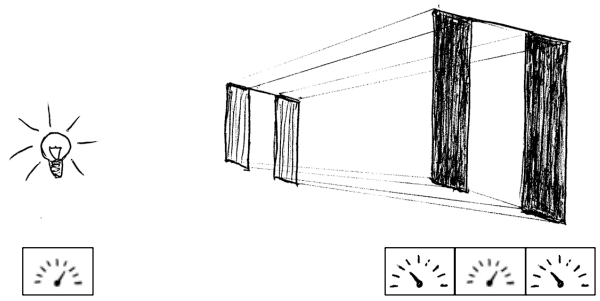
If less neutrinos than expected are measured, a sterile neutrino is found.

Actually the object looks more like a curtain and has two properties: **thickness** and **covered area**



The curtain can be for example:

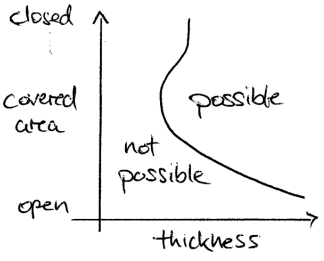
The lack of light depends then on the specific curtain



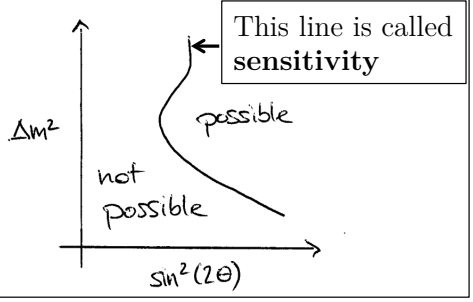
The thicker the curtain and the bigger the covered area, the larger the difference in the light measurements becomes.* The ability to observe shadow depends thus on the two properties of the curtain.
Similarly, the chance to observe a sterile neutrino depends also on two parameters.

* Too small differences cannot be resolved.

Can you observe shadow?



Can you observe a sterile neutrino?



4 Development of the Analysis Software

This chapter describes the analysis software of SOX, where a large fraction of it has been developed in the frame of this thesis. The main output of the analysis software is a confidence region in the oscillation parameter space for the SOX data sample. Moreover, using the analysis software, the expected sensitivity of the SOX experiment accounting for the experimental systematics can be determined. In contrast, before this thesis, the sensitivity of SOX was estimated based on toy-Monte Carlo (MC) simulations assuming ideal detector settings [132, 133, 161, 165].
The individual components of the analysis chain are schematically shown in Figure 4.1. The first step is to simulate the expected raw data which has the same format as the Borexino raw data (see Section 4.1). Therefore, all the subsequent data processing can be applied in the same way to MC and Borexino data. The inverse beta decay (IBD) events are selected from the raw data using the `bx-sox` filter (see Section 4.2). The result of SOX is displayed in a confidence region in the $\sin^2(2\theta)$-Δm^2-region that is obtained through the `sox-stats` analysis framework (see Section 4.3). The chapter closes with the comparison of the expected sensitivity obtained from the toy-MC simulations and the full analysis chain (see Section 4.4).

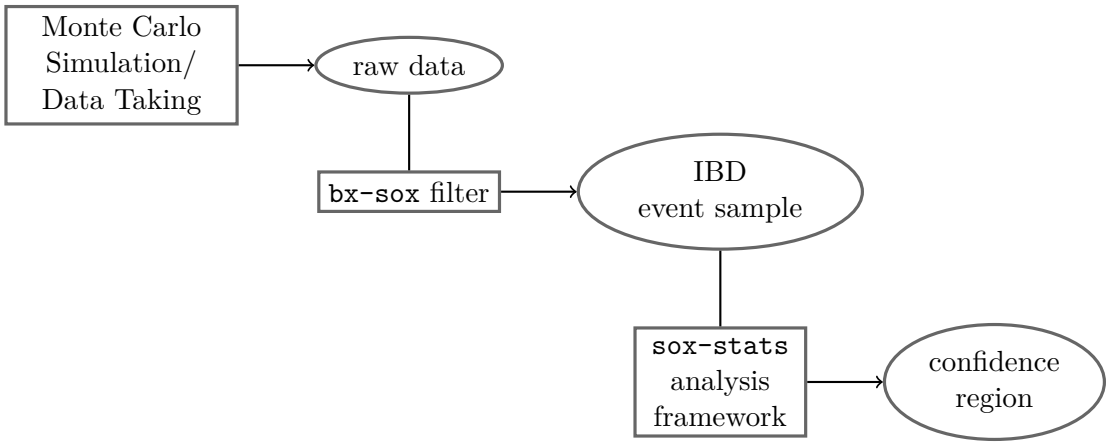


Figure 4.1: Scheme of the Analysis Structure of SOX The individual components are explained in the specific sections of this chapter.

4.1 Generation of the Expected Raw Data

In order to perform a sterile neutrino analysis, the expected signature as a function of the oscillation parameters $\sin^2(2\theta)$ and Δm^2 has to be precisely known. The specific detector

material and the status of the Borexino detector, such as the vessel shape and the number of working (photomultiplier tubes) PMTs, are crucial to predict the expected signature. The expected signature consists of the total event number and the probability distributions of the spectral distributions (PDFs) which are described in the following.

4.1.1 Event Number

The total event number N_{IBD} of IBD reactions within the inner vessel for a survival probability and detection efficiency of one can be determined analytically. Assuming a point-like source, N_{IBD} is given by

$$N_{\mathrm{IBD}} = \iiint \frac{A(t)}{4\pi L^2} \cdot \rho_p \cdot n_\nu(E) \cdot \sigma(E) \cdot F(L, R_{\mathrm{IV}}(L))\, dL\, dE\, dt, \tag{4.1}$$

where the integrals are computed over the full ranges of traveled distance (L) and energy (E) of the neutrinos, and the measurement time (t). Here, $A(t) = A_0 \exp(-\frac{\ln 2}{T_{1/2}} \cdot t)$ is the activity of the ^{144}Ce-source with its half-life $\mathrm{T}_{1/2}$, ρ_p the proton density in the liquid scintillator, $n_\nu(E)$ the neutrino energy spectrum of ^{144}Pr, and $\sigma(E)$ the cross section of the IBD reaction (see Equation 2.7). $F(L, R_{\mathrm{IV}}(L))$ is the surface area of the spherical cap with radius L centered at the source location that is contained within the inner vessel volume:

$$F(L, R_{\mathrm{IV}}(L)) = 2\pi L^2 \cdot (1 - \frac{d^2 + L^2 - R_{\mathrm{IV}}^2(L)}{2dL}). \tag{4.2}$$

Figure 4.2 sketches the geometrical relation of the inner vessel to the source location. Since the deformation of the inner vessel shape is cylindrical symmetric around the z-axis (see Chapter 2), the vessel radius R_{IV} can be expressed as a function of L. Given the small source dimensions of the source ($\sim 0.15\,\mathrm{m}$) in comparison to the detector dimensions, the assumption of a point-like source is appropriate.
The expected event number $N_{\mathrm{S}}^{\mathrm{exp}}$ of reconstructed IBD events depends on the detection efficiency ϵ for a given fiducial volume cut (FV) and the ratio τ of lifetime over measurement time of the detector:

$$N_S^{\mathrm{exp}} = N_{\mathrm{IBD}} \cdot \epsilon(FV) \cdot \tau. \tag{4.3}$$

The detection efficiency is determined from Borexino MC simulations and is defined here as the fraction of generated events in the whole inner vessel volume ($N_{\mathrm{true}}(IV)$) that are reconstructed in the fiducial volume ($N_{\mathrm{rec}}(FV)$)

$$\epsilon(FV) = \frac{N_{\mathrm{rec}}(FV)}{N_{\mathrm{true}}(IV)}. \tag{4.4}$$

The lifetime of the detector is reduced in comparison to the total measurement time, i.e. $\tau < 1$, due to for example calibration runs, maintenance operations, and detector instabilities. Moreover, software selection cuts (mainly the muon veto) further reduce the lifetime. The Borexino MC simulation used to determine the detection efficiency as well as the spectral distributions is described in the following.

4.1.2 Spectral Distribution

The expected spectrum from an antineutrino source is generated in two steps. First, the so-called antineutrino generator developed in Reference [165] produces the distributions of the true observables (e.g. E_{true} and L_{true}) of the antineutrinos in the Borexino detector. Second,

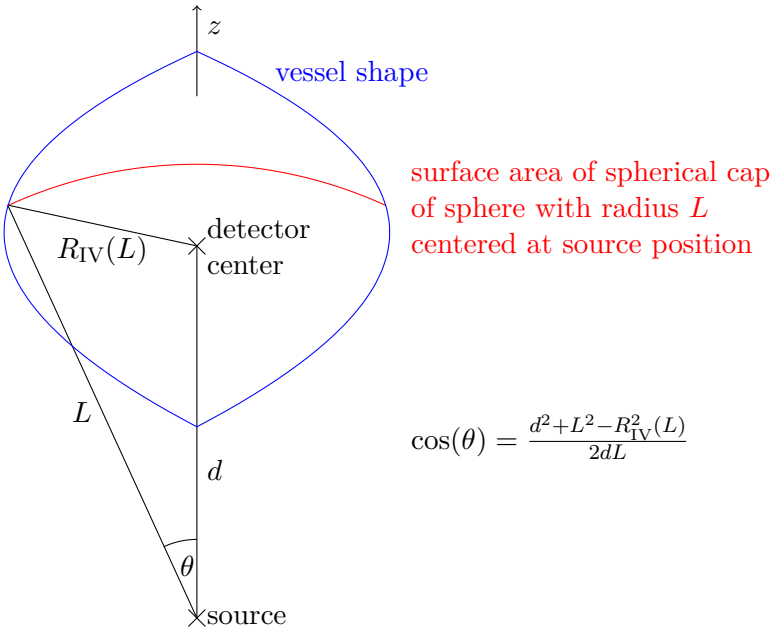


Figure 4.2: Scheme of the Computation of the Event Number The vessel shape which is cylindrical symmetric around the z-axis is indicated in blue. $R_{\mathrm{IV}}(L)$ denotes the radius of the inner vessel and can be expressed as a function of L which is the travelled distance of the neutrinos from the source. The distance between source and detector center is given by d. Using the law of cosines, the angle θ is a function of d, L, and $R_{\mathrm{IV}}(L)$ (not to scale).

the Borexino MC based on Geant4 [184] uses the output of the antineutrino generator and simulates the reconstructed observables (e.g. E_{rec} and L_{rec}).

The antineutrino generator takes into account the source position and geometry, and for each antineutrino event the true start position is uniformly distributed within the source. A possible interaction vertex is randomized within a sphere with the radius of the maximum baseline centered around the true start point. The possible interaction vertex is only saved when it appears inside the inner vessel of Borexino, where the specific vessel shape can be defined. The true baseline L_{true} is then given by the difference of start and interaction vertex. E_{true} is randomly drawn from the convolution of the antineutrino spectrum emitted by the ^{144}Ce source (see Figure 2.8) and the IBD cross section [143].

The Borexino MC is based on Geant4 [184] and starts at the true interaction vertex the two products of the IBD reaction: a positron and a neutron. The kinetic energies and the directions of the positron and the neutron follow from the kinematics of the IBD, where $E_{e^+} \approx E_\nu + m_p c^2 - m_n c^2$ and $E_n \lesssim 5.4\,\mathrm{keV}$ for $E_\nu = 3\,\mathrm{MeV}$ [185]. The MC simulation takes into account the full detector geometry and materials, and simulates the energy deposition of the particles and the resulting scintillation photons. Further, the propagation of these scintillation photons through the detector medium and the detection at the PMTs, including the electronic signal processing, is simulated. In particular, the vessel shape and the number of active PMTs, which are both unstable over the SOX data taking period, can be set to a specific configuration. Therefore, by running MC simulations with the time-dependent detector configurations, the time-dependent spectral distributions with the respective energy and position resolution can be determined. The output of the MC has the same format as the Borexino raw data. Hence, all further data processing and analysis can be performed in the same way to MC and Borexino data.

Figure 4.3 shows the reconstructed and true spectra of the positrons for the energy and the

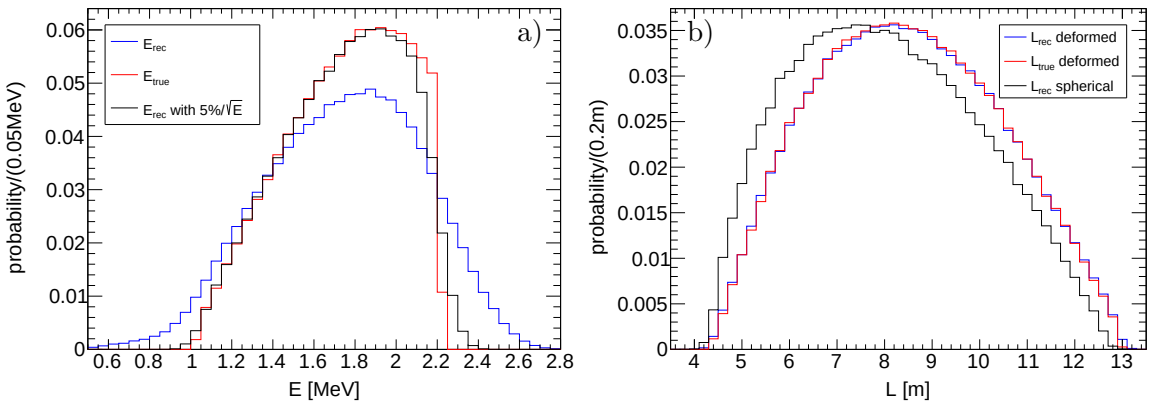


Figure 4.3: Simulated Energy and Distance Distributions of SOX with the Borexino MC The true (red) and reconstructed (blue) distributions are shown for (a) energy and (b) distance, respectively. The true (reconstructed) distributions are the output of the antineutrino generator (Borexino MC) using the detector configuration from October 2016 with the specific deformed vessel shape. In comparison, the distributions for the assumed ideal detector settings in the toy-MC sensitivity studies (spherical vessel, $\sigma_L = 10\,\text{cm}$, $\sigma_E = 5\,\%/\sqrt{E}$) are shown in black. Note that the true and the reconstructed observables almost overlap in panel b).

distance. The detector settings correspond to the ones of October 2016 (see Figure A.7 for the vessel shape). The energy spectrum is compared to the one used in the toy-MC sensitivity studies. Here, a spectrum with 5 % energy resolution at 1 MeV, which is the resolution in the innermost detector volume of Borexino, is assumed. As the resolution drops towards larger radii (see Figure 5.2), the average energy resolution in the full active volume with the Borexino MC is worse. In addition, the distance spectrum is compared to the one obtained with a spherical vessel which is used in the toy-MC sensitivity studies. The spectrum with the deformed vessel is shifted towards larger baselines. This is because of the balloon-like vessel shape with more active volume in the northern hemisphere of the detector (see Figure 2.3). Note that the deformation also affects the total expected event number, as the neutrino flux decreases with the baseline.

Within this thesis, the MC samples were studied in detail to investigate the detector response and the detection efficiency (see Chapter 5). Moreover, the proper functionality of the MC simulation was tested and possible problems were identified.

4.2 Event Selection with the `bx-sox` Filter

The `bx-sox` filter has been developed during this thesis and is based on the existing bx-antinu filter used in the previous geo neutrino analyses [155]. Two main improvements are implemented to increase the statistics for the SOX event sample. First, the algorithm is extended to search for double cluster events, i.e. events with a time difference $\Delta t < 16\,\mu\text{s}$, the gate length of one data acquisition (DAQ) window. Second, a new muon veto is implemented that differentiates between two different muon categories, increasing the life time by $\sim 10\,\%$. In turn, the `bx-sox` filter was the basis for the latest geo neutrino analysis [35]. In the following, the data selection cuts and the validation of the `bx-sox` filter are discussed.

Table 4.1: Selection Cuts in the bx-sox Filter. The first and second row lists the selection parameters and their definitions. The third row gives the specific cut, where the cut values are optimized in Chapter 5 and listed in Table 5.1.

parameter	definition	cut	comment
E_p	energy of prompt event	$E_p^{\min} < E_p < E_p^{\max}$	
E_d	energy of delayed event	$E_d^{\min,1} < E_d < E_d^{\max,1}$ or	n-capture on H
		$E_d^{\min,2} < E_d < E_d^{\max,2}$	n-capture on C
Δt	time difference between	$\Delta t^{\min,1} < \Delta t < \Delta t^{\max,1}$ or	2 clusters/gate
	prompt and delayed event	$\Delta t^{\min,2} < \Delta t < \Delta t^{\max,2}$	1 cluster/gate
ΔR	spatial distance between prompt and delayed event	$\Delta R < \Delta R_{\max}$	
mlp	pulse shape discrimination parameter of delayed event	$mlp > mlp_{\min}$	β-like events
vd	distance to vessel of prompt event	$vd > vd_{\min}$	
$\Delta t(\mu_{\text{int}}+n)$	vetoed time after $\mu_{\text{int}}+n$	$\Delta t(\mu_{\text{int}}+n) > 2\,\text{s}$	
$\Delta t(\mu_{\text{int}}-n)$	vetoed time after $\mu_{\text{int}}-n$	$\Delta t(\mu_{\text{int}}-n) > 2\,\text{ms}$	
$\Delta t(\mu_{\text{ext}})$	vetoed time after μ_{ext}	$\Delta t(\mu_{\text{ext}}) > 2\,\text{ms}$	
multiplicity cut	no event with $E > E_p^{\min}$ 2 ms before, after, or in between prompt and delayed event		

4.2.1 Selection Cuts

An IBD event needs to fulfill the following requirements which are also summarized in Table 4.1: The prompt and the delayed event have to trigger both the inner detector ($tt1$) and not the outer detector ($btb0$).[1] Thus, both events are not tagged as internal (μ_{int}) or external (μ_{ext}) muons[2]. Further, muon-induced background is removed by the following muon vetos: the prompt and the delayed event have to be at least 2 s after an internal muon, which is followed by neutrons ($\mu_{\text{int}}+n$), and 2 ms after an internal muon, which is not followed by neutrons ($\mu_{\text{int}}-n$) or μ_{ext}. The veto of 2 ms is sufficient to suppress cosmogenic neutron background (neutron capture time is $\sim 250\,\mu\text{s}$). The veto of 2 s suppresses most of the muon-induced hadronic background (see Chapter 6). To distinguish between ($\mu_{\text{int}}+n$) and ($\mu_{\text{int}}-n$), one searches for neutrons in the $tt128$-gate that is automatically issued after each internal muon with a gate length of 1.6 ms. If at least one neutron is found, the muon is tagged as ($\mu_{\text{int}}+n$). More details about the $tt128$ and the neutron identification can be found in Ref [140].
Further, the following IBD selection cuts are applied[3]: the energy of the prompt (E_p) and the delayed (E_d) event have to be reconstructed in the expected energy ranges. In addition, the prompt and the delayed events need to be correlated in time (Δt) and space (ΔR). A pulse shape discrimination cut on the delayed event is applied to suppress background induced by fast coincidences from the ^{238}U or ^{232}Th chain. Here, the so-called mlp parameter is used to discriminate neutrons against alpha particles [35]. A fiducial volume cut is applied to the reconstructed position of the prompt event. As the vessel shape is deformed, the fiducial volume cut is performed in terms of the distance to the inner vessel (vd).

[1] A list of the trigger types in Borexino and their definitions can be found in Appendix A.2.
[2] The exact definition for the internal muon is called "internal large" in Borexino and can be found in Reference [35].
[3] The values of the selection cuts are optimized in Chapter 5 and listed in Table 5.1.

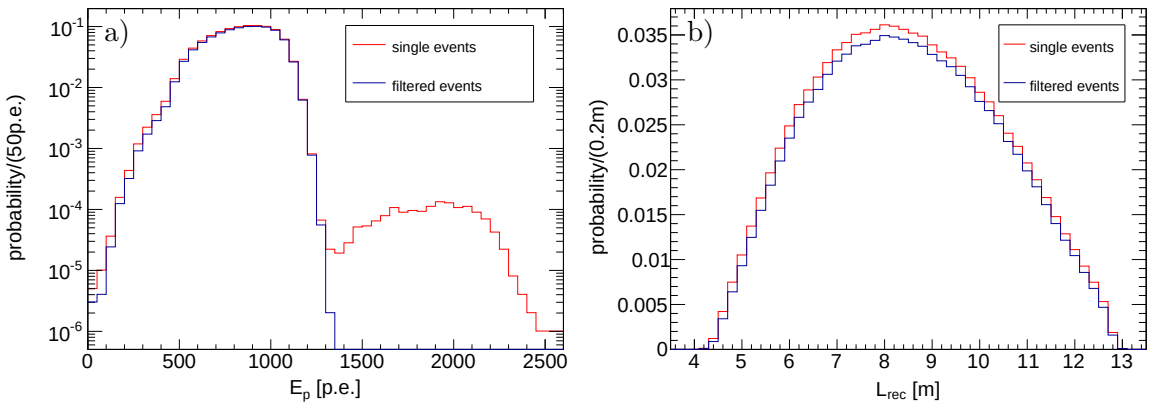


Figure 4.4: Simulated Energy and Distance Distribution before and after the IBD Event Selection with the `bx-sox` Filter. The distributions of the single prompt events are selected with the MC information on the primary particle and are normalized to one. The filtered events are selected with the `bx-sox` filter applying only a time cut (Δt<2000 µs) and normalized to the event number of the single prompt events. The reduced integral in the distributions of the filtered events comes mainly from the detection efficiency of the delayed event. The energy distributions (a) are shown in logarithmic scale to visualize the second component at high energies which can be traced back to the finite time resolution of the cluster algorithm (see text).

The selected IBD candidate must be a clear coincidence of two events. As the Borexino gate length is 16 µs long, several physical events (so-called clusters) can be reconstructed in the same gate. A cluster is defined as an accumulation of detected scintillation photons ($\gtrsim 30$) within a short time period ($\lesssim 1.5$ µs), and is reconstructed from a software algorithm [141]. It is required that the prompt and the delayed event are the only reconstructed clusters in the gate ($\#clusters = 1$), unless the prompt and the delayed event appear in the same gate ($\Delta t < 16$ µs and $\#clusters = 2$). Further, a so-called multiplicity cut is applied. Hence, if an event with an energy larger than the minimum energy of the prompt event appears 2 ms before the prompt event, 2 ms after the delayed event, or in between the prompt and the delayed event, the IBD candidate is rejected.

The energy and distance spectra obtained from the MC simulation after the event selection with the `bx-sox` filter is shown in Figure 4.4. In comparison, the spectra of the single prompt events are shown which are selected with the MC information on the primary particle. The total event number is reduced in comparison to the spectra of the single prompt events because of the detection efficiency of the IBD. Applying only an upper limit on the time difference ($\Delta t < 2000$ µs), the detection efficiency results to ~ 95 % that can be mainly explained due to the reduced detection efficiency of the neutrons close to the vessel border. The energy distribution of the single prompt events features a second distribution beyond the endpoint of the ^{144}Pr spectrum (~ 1300 p.e.) that does not appear in the `bx-sox` filtered spectrum. For these events (~ 0.2 %), the delayed event follows too shortly after the prompt event ($\Delta t \lesssim 300$ ns) and cannot be separated from the prompt event with the cluster algorithm. The reconstructed energy is hence given by the sum of the energy of the prompt and the delayed event.

The distributions of all selection parameters of the above filtered MC data sample is shown in Figure 4.5. In addition, the selection cuts of the previous geo neutrino analysis are indicated [155]. Chapter 5 discusses the improvement on the specific selection cut values in order to enhance the detection efficiency and hence the sensitivity in SOX.

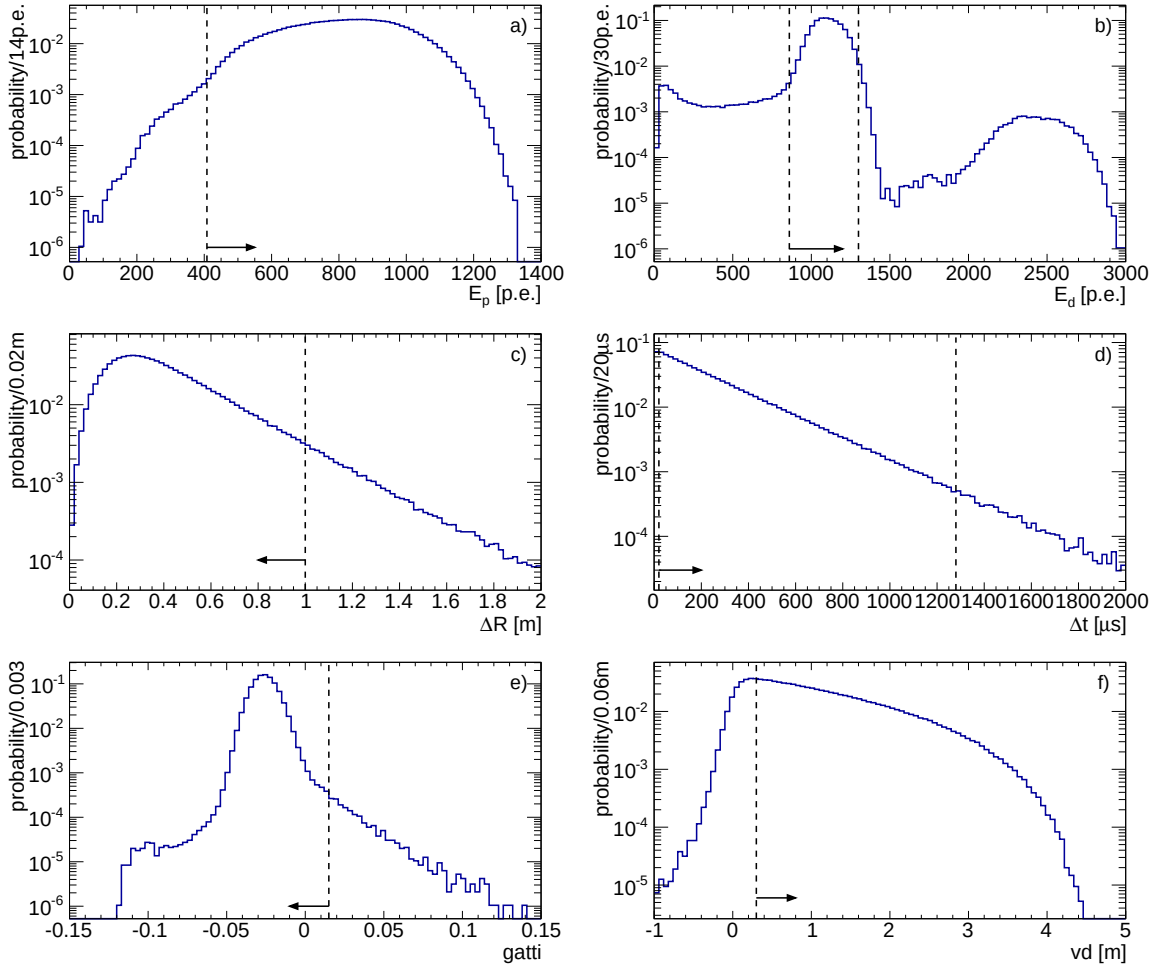


Figure 4.5: Distributions of the IBD Selection Parameters The distributions are simulated from 1×10^6 prompt and delayed events started in the whole inner vessel assuming the detector configuration from October 2016. The IBD events are selected with the `bx-sox` filter applying a minimal set of selection cuts (Δt<2000 µs). The dashed lines indicate the selection cuts used in the previous geo neutrino analysis [155]. The improvements on this selection cuts for the SOX analysis are discussed in Chapter 5.

4.2.2 Validation of the `bx-sox` Filter

The proper functionality of the `bx-sox` filter can be cross-checked using the saved information on the primary particles in the MC simulation. The Borexino MC saves the prompt event as an $\bar{\nu}_e$ ($pdg = -12$) and the delayed event as a neutron ($pdg = 2112$), where the particle data group (pdg) identifiers are used [176]. An additional filter algorithm (mctruth filter) is developed using the MC information to validate the `bx-sox` filter. To do so, IBD events are selected that fulfill the required particle types of prompt and delayed event. Further, both events produce enough scintillation light and are reconstructed as clusters. In both IBD selections (`bx-sox` and mctruth) only Δt<2000 µs is required and no further cuts on the energies, ΔR, and the pulse shape are applied.
With this minimal set of selection cuts, 953325 events (`bx-sox`) and 953150 events (mctruth) are found for the same MC data sample. The difference of 175 events can be explained by ^{14}N(n,p)^{14}C reactions that can appear on the nitrogen nuclei of the wavelength shifter PPO [186]. For such reactions, the MC information saves a proton ($pdg = 2212$) for the particle

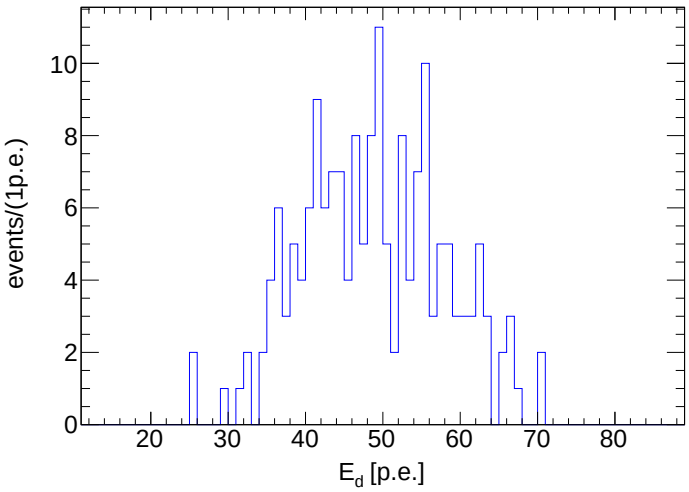


Figure 4.6: Validation of the `bx-sox` Filter The energy of the delayed events from MC data is shown for those IBD events which are selected with the `bx-sox`, but not with the mctruth filter. The mctruth filter selects IBD events according to the saved particle type in the MC. The shown events correspond to protons produced with a chance of $1.8 \times 10^{-2}\%$ in ^{14}N(n,p)^{14}C reactions with an energy of 626 keV, where the reconstructed energy is quenched by a factor of ~ 6.

type of the delayed event and are thus not selected with the mctruth filter. Figure 4.6 shows the energy of the delayed events for these ^{14}N(n,p)^{14}C reactions, which correspond to protons with an energy of 626 keV. As the scintillation light yield of protons is quenched, the energy is reconstructed at ~ 100 keV (50 p.e.). Therefore, as both filters give consistent results and the difference in the selected event numbers is understood, the proper functionality of the `bx-sox` filter could be demonstrated.

4.3 Construction of a Confidence Region with the `sox-stats` Framework

`sox-stats` is a collection of `C++` classes which provides interfaces and algorithms for the statistical analysis of the SOX data. The basic version has been implemented by Matteo Agostini. In the frame of this thesis, `sox-stats` has been extensively tested and modified to study the systematic uncertainties and the optimization of the analysis. In the following, a description of the basic version is given. The individual modifications are described in the respective sections of Chapter 7.

4.3.1 Description of `sox-stats`

The outcome of `sox-stats` is the minimum of the negative log-likelihood function ($-2 \ln \mathcal{L}$) in the $\sin 2(2\theta)$ - Δm^2 - plane from which the confidence region at a specific confidence level can be extracted. `sox-stats` is segmented in three different `C++` objects that take care of the generation of the expected data, the statistical model, and the scanning over the likelihood space. The minimization of the negative log-likelihood function is performed using the minimizer object from the `m-stats` package [187] which is based on Minuit [188]. A scheme of the structure of these three objects with the input and output parameters can be seen in Figure 4.7. These objects are described in more detail in the following.

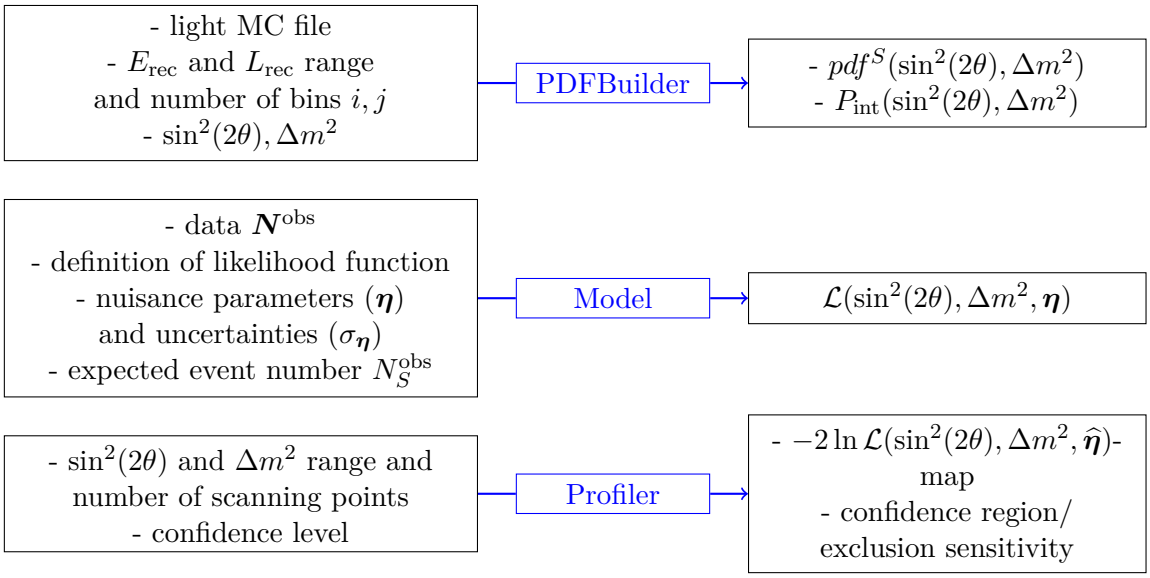


Figure 4.7: Scheme of the Statistical Analysis Framework for SOX (`sox-stats`) The statistical framework is divided into three objects: the PDFBuilder, the Model, and the Profiler. The specific input and output values are given on the left and right side, respectively. The minimization of the negative log-likelihood function is performed with the Minimizer based on Minuit from the external class package called `m-stats` [187].

1. **The PDFBuilder**

The output of the PDFBuilder are the normalized probability distribution functions ($pdf^S(\sin^2(2\theta), \Delta m^2)$) and the integral survival probability ($P_{\text{int}}(\sin^2(2\theta), \Delta m^2)$) for a certain oscillation hypothesis $\text{H}(\sin^2(2\theta), \Delta m^2)$. The data ($\boldsymbol{N}^{\text{obs}}$) and the probability distribution functions are represented by 2D histograms with bins (i, j) in reconstructed energy (E_{rec}) and distance (L_{rec}). For the generation of a specific probability distribution function $pdf^S(\sin^2(2\theta), \Delta m^2)$ of a given hypothesis, a high statistic (10^6) MC simulation sample of the no oscillation hypothesis is used (see Section 4.1). In order to speed up the analysis, a light MC file is used as input for the PDFBuilder which is extracted from the massive MC simulation using `bx-sox` and contains the necessary parameters for the analysis that are mainly:

(a) true energy E_{true}

(b) reconstructed energy E_{rec}, i.e. normalized charge

(c) true distance L_{true}

(d) reconstructed distance L_{rec}

$pdf^S(\sin^2(2\theta), \Delta m^2)$ is constructed by looping over the events in the light tree (n_{MC}) and filling the reconstructed parameters (L_{rec},E_{rec}) in a 2D histogram, where the fill weights correspond to the survival probability ($P_{ee}(L_{\text{true}}, E_{\text{true}}|\sin^2(2\theta), \Delta m^2)$) for the respective true parameters (L_{true},E_{true}) and oscillation parameters $(\sin^2(2\theta), \Delta m^2)$ according to Equation 1.10. $P_{\text{int}}(\sin^2(2\theta), \Delta m^2)$ is given by the ratio of the sum of the individual survival probabilities of all events in the light tree over n_{MC}:

$$P_{\text{int}}(\sin^2(2\theta), \Delta m^2) = \frac{1}{n_{\text{MC}}} \sum_{i=0}^{n_{\text{MC}}} P_{ee}(L_{\text{true}}, E_{\text{true}}|\sin^2(2\theta), \Delta m^2). \tag{4.5}$$

2. **The Model**

   In the Model the parameters of interest, the nuisance parameters, and the likelihood function are defined. The parameters of interest are $\sin^2(2\theta)$ and Δm^2. In the basic model one nuisance parameter is defined, namely the total event number for the no oscillation hypothesis N_S which is constrained by the number of expected events (N_S^{obs}) with the uncertainty σ_{N_S}. The likelihood function in a rate+shape analysis is given by (following the notation from Chapter 3):

$$\mathcal{L}(\sin^2(2\theta), \Delta m^2, N_S) = \prod_{ij} \mathcal{P}(N_{ij}^{\text{obs}} | N_{ij}^{\text{exp}}(\sin^2(2\theta), \Delta m^2, N_S)) \cdot \mathcal{N}(N_S^{\text{obs}} | N_S, \sigma_{N_S}^2) \quad (4.6)$$

   with

$$N_{ij}^{\text{exp}}(\sin^2(2\theta), \Delta m^2, N_S) = N_S^{\text{exp}}(\sin^2(2\theta), \Delta m^2, N_S) \cdot pdf_{ij}^S(\sin^2(2\theta), \Delta m^2) \quad (4.7)$$

   and

$$N_S^{\text{exp}}(\sin^2(2\theta), \Delta m^2, N_S) = P_{\text{int}}(\sin^2(2\theta), \Delta m^2) \cdot N_S. \quad (4.8)$$

   Instead, the likelihood function in a shape analysis is defined without the Gaussian constraint $\mathcal{N}(N_S^{\text{obs}} | N_S, \sigma_{N_S}^2)$ on the total event number. In a rate analysis, the number of bins in the data and the $pdf^S(\sin^2(2\theta), \Delta m^2)$ are set to one ($i = j = 1$).

3. **The Profiler**

   A scan over the whole $\sin 2(2\theta)$ - Δm^2 - plane (with uniform steps in the logarithmic space) is performed and at each point $-2 \ln \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})$ is minimized with respect to the nuisance parameters $\boldsymbol{\eta}$, i.e. N_S in the basic statistical model. The fit is executed by the minimization class of the `m-stats` package [187]. After the scan, the negative log-likelihood function is normalized for the absolute minimum to obtain the profile likelihood ratio as defined in Equation 3.9. Besides the negative log-likelihood function, the Profiler returns also the confidence region at a certain confidence level which is extracted with Wilks' theorem. If the data sample corresponds to the Asimov data set for the no-oscillation hypothesis, the exclusion sensitivity is obtained (see Chapter 3).

Note that the exclusion sensitivities obtained with `sox-stats` are using the Asimov data set and Wilks' theorem. In the second part of this thesis (see Chapter 11) the difference to a full MC construction is shown and discussed.

4.4 Sensitivity of the SOX Experiment

Figure 4.8 shows the sensitivity of SOX in a rate+shape analysis reconstructed from the full analysis chain, which is denoted as "real detector" in the following. In addition, the sensitivity from toy-MC studies using "ideal detector" settings is compared. The ideal detector assumes a spherical vessel with 4.25 m radius, an energy resolution of 5 % at 1 MeV, and a spatial resolution of 10 cm (see Figure 4.3). Further, a detection efficiency of 100 % is assumed. The sensitivity of the real detector uses the full analysis chain described in this chapter. Therefore, the expected data is generated with the Borexino MC assuming the vessel shape and the detector settings from October 2016. Further, IBD events are selected using the `bx-sox` filter with a minimal set of selection cuts (i.e. Δt<2000 µs, no cuts on energy and

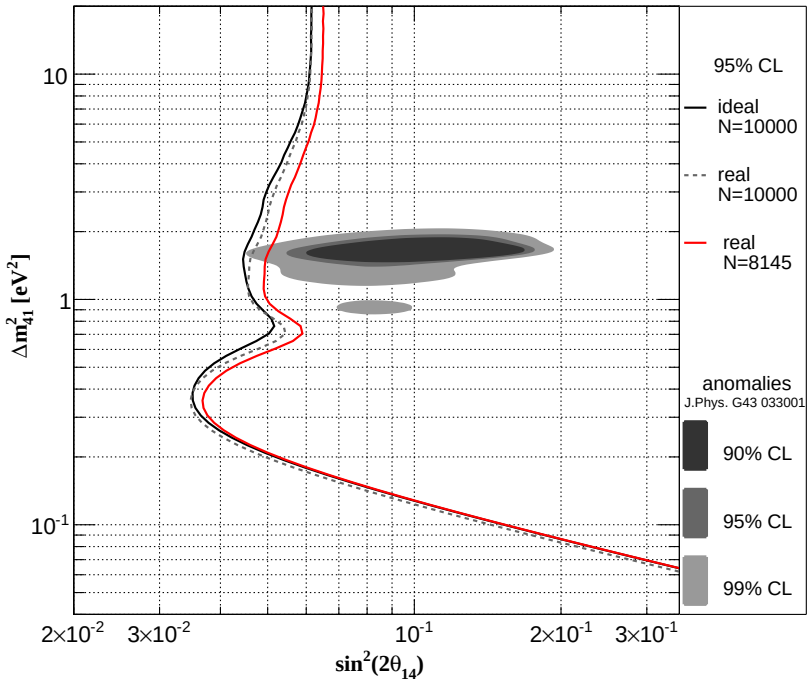


Figure 4.8: Comparison of the Sensitivity Reconstructed from an Ideal and Real Borexino Detector The ideal detector (black) assumes a spherical vessel shape and an energy resolution of 5 % at 1 MeV and a vertex resolution of 10 cm. The real sensitivity (red) is obtained with the whole analysis chain assuming a detector configuration from October 2016. The source activity corresponds to 115 kCi and the uncertainty on the total event number is 1 %. To visualize the differences in the shape of the sensitivity, the real sensitivity is in addition also computed for the same expected event number of the ideal detector (N=10000, black dashed line). The overall sensitivity is reduced for the real sensitivity, as the expected event number drops due to the detection efficiency ($\sim 95\,\%$) and the deformed vessel shape. The allowed regions for sterile neutrinos from a global analysis are added in gray. It has to be noted that the real sensitivity is still too optimistic, as not all systematic effects are taken into account.

ΔR). Both sensitivities are produced with the `sox-stats` framework applying no fiducial volume cut to the data.

The expected event number for the no-oscillation hypothesis for the ideal experiment is assumed to be 10000 events which corresponds to a source activity of ~ 115 kCi. The expected event number is reduced in the real setup by 18.5 % due to the deformed vessel shape and the detection efficiency. In both analyses, an uncertainty of 1 % on the total event number is taken into account.

The shape of the sensitivity is similar in both cases, where it is maximal for $\Delta m^2 \sim 3.5\,\mathrm{eV}^2$. At this Δm^2-value, the oscillation length is exactly twice the L_{rec}-value with the maximal expected event number L_{max} at ~ 8 m (see Figure 4.4). Hence, as the first oscillation minimum overlaps with the maximum in the expected event number, a large deficit in the expected event number is observable and the sensitivity is large. For smaller Δm^2-values, the oscillation length becomes larger than the L_{rec}-range and an oscillation signal and the sensitivity vanishes. In this mass regime ($\Delta m^2 \lesssim 2\,\mathrm{eV}^2$) the deficit in the event number and hence the sensitivity is approximately proportional to the product $\sin^2(2\theta) \times \Delta m^2$. On the contrary, for larger Δm^2-values ($1\,\mathrm{eV}^2 \lesssim \Delta m^2 \lesssim 10\,\mathrm{eV}^2$), the oscillation length becomes smaller than the detector dimensions and an oscillation pattern is observable. For even larger Δm^2-values ($\Delta m^2 \gtrsim 10\,\mathrm{eV}^2$), the oscillation length becomes smaller than the detector resolution and the

oscillation signature is washed out which leads to a constant rate deficit and hence constant sensitivity in terms of $\sin^2(2\theta)$ (see also the expected signature in Figure 2.14).
The overall sensitivity for the real detector is reduced in comparison to the ideal detector settings by $\sim 12\,\%$ at $\Delta m^2 = 2\,\mathrm{eV}^2$ which is mainly given due to the reduced expected event number. In addition, the shape of the sensitivity curve slightly changes with the real detector, which can be seen from the comparison of the sensitivities of the ideal and real detector settings with the same expected event number (black solid and dashed line). The sensitivity curve shifts towards smaller Δm^2-values for $\Delta m^2 \lesssim 1\,\mathrm{eV}^2$, as the deformed vessel shape shifts the L_{rec}-spectrum towards larger values for the real detector (see Figure 4.4). Moreover, the energy and spatial resolution is worse in the real Borexino detector than in the ideal assumptions. The ability to resolve small oscillation lengths is reduced in comparison to the ideal settings which decreases the sensitivity for $\Delta m^2 \gtrsim 1\,\mathrm{eV}^2$. It has to be noted that the shown sensitivity is too optimistic, as it does not take into account all systematic effects and the full detection efficiency (see Chapter 7).

This chapter described the developed analysis software for SOX which is used in the following chapters to characterize the Borexino detector and evaluate the systematic effects of SOX. The `bx-sox` filter is used, on the one hand, to improve the selection cuts for the IBD and determine its detection efficiency in Chapter 5. On the other hand, the expected background for SOX is studied with the `bx-sox` filter in Chapter 6. Further, the systematic effects of the SOX experiment and the final sensitivity projection are studied in Chapter 7 with the `sox-stats` framework.

5 Characterization of the Signal and the Borexino Detector

In order to predict the expected signal of SOX and run a sterile neutrino analysis, the Borexino detector needs to be well characterized. This chapter investigates the detector response of the Borexino detector to SOX antineutrino events in Section 5.1. Section 5.2 discusses the detection efficiency for SOX antineutrino events in the Borexino detector. Moreover, as the sensitivity of SOX increases with the expected event number, possibilities to improve the detection efficiency with respect to previous antineutrino analyses in Borexino are shown.

5.1 Detector Response

Borexino measures for each event the amount of scintillation light per photomultiplier tube (PMT) and the respective arrival times, which allows to reconstruct the energy and the position of the event. In the following, the expected energy and position reconstruction for the SOX event sample is studied. Moreover, their systematic uncertainties are discussed which are obtained from calibration measurements with radioactive sources [189].

5.1.1 Energy Reconstruction

In Borexino, three energy estimators N_p, N_h, and N_{pe} are defined, i.e. the number of PMTs which detected at least one photon, the number of reconstructed hits of all PMTs[1], and the total measured charge of the PMTs expressed in the number of photoelectrons (p.e.) [141]. As the number of working PMTs is decreasing with the lifetime of Borexino, each estimator is corrected for the time-dependent number of working PMTs and normalized to 2000 working PMTs, such that an energy deposition of 1 MeV constantly corresponds to ~ 500 p.e.. For large radii and/or high energies, i.e. in the case when the probability is high that a PMT can detect more than a single scintillation photon, N_{pe} gives better resolution than N_p or N_h. Thus, N_{pe} maximizes the sensitivity of SOX and will be therefore the choice for the SOX analysis.
The energy response is studied with the SOX Monte Carlo (MC) simulation[2] (see Chapter 4) assuming the expected energy and position distribution of SOX. From Figure 5.1 and Figure 5.2 it can be seen that the light yield and its resolution becomes worse for larger radii. While the energy resolution reaches 5 % for an energy of 2.2 MeV in the detector center, it decreases to 14 % in the outer radial shells. The main reason for the worsening for larger radii comes from the high fraction of inactive PMTs in the southern part of the detector, leading to a decreasing light yield for smaller z-values. As the SOX source is located below

[1] The Borexino electronics of one PMT can resolve two signals for time differences larger than 80 ns [140].
[2] Unless otherwise stated, the SOX MC simulation assumes the Borexino detector configuration of October 2016 and generates 10^6 positrons and neutrons each within the inner vessel.

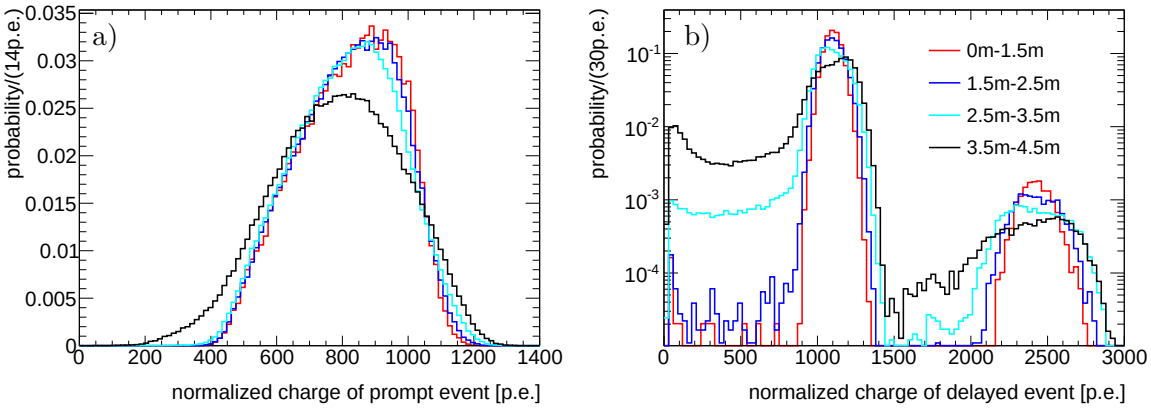


Figure 5.1: SOX Energy Spectra as a Function of the Detector Position The energy spectra for the prompt (a) and the delayed (b) event are obtained from SOX MC simulations and are grouped into radial shells of the true IBD interaction position. The integral of the distributions is normalized to one. The two peaks in the delayed spectra are due to the neutron captures on Hydrogen and on Carbon and the low energy tail comes from a partial energy deposition in the buffer.

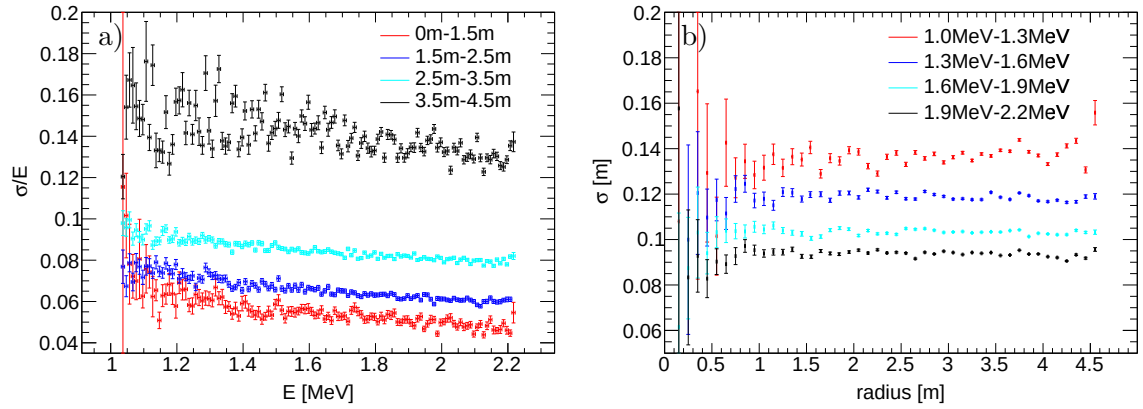


Figure 5.2: Energy and Position Resolution in the SOX Event Sample a) relative energy resolution as a function of the true energy for different radial shells. b) absolute position resolution as a function of the true radius for different energies. The results are obtained from SOX MC simulations. The error bars account for the statistical errors only. For the innermost detector volume and for the highest energies, the best resolutions are achieved.

the detector, the neutrino flux and the expected event number for SOX dominates in the southern part of the detector. Thus, the SOX event sample features mainly the detector response from this detector region. In addition, the energy scale decreases for large radii close to the inner vessel (IV) border, where a fraction of the energy can be deposited in the buffer and is quenched (see low energy tail in Figure 5.1).

The MC reproduces both the light yield and the resolution with a precision of $\leq 0.8\,\%$ in the detector center [184]. This is shown in Figure 5.3a which compares calibration data of radioactive gamma sources with MC simulations. In addition, the energy response as a function of the detector position has been studied with the 2.2 MeV gamma line from the ^{241}Am-^{9}Be source [184]. The relative agreement of the reconstructed energy is in the detector center $\lesssim 2\,\%$ and drops close to the IV border up to $4\,\%$, as depicted in Figure 5.3b. This discrepancy can be most likely explained due to the uncertainties in the vessel shape reconstruction.

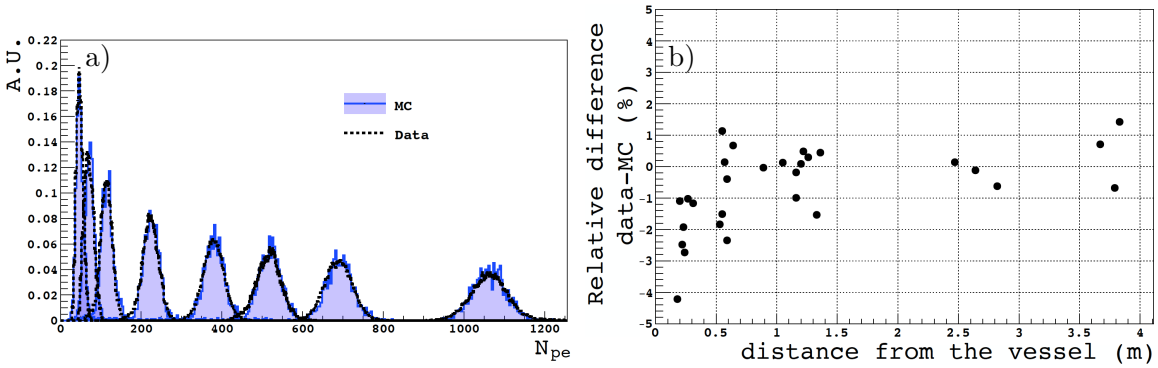


Figure 5.3: Accuracy of the Energy Response Calibration data is compared to MC simulation data for a) gamma sources in the detector center and b) the 2.2 MeV gamma line from an ^{241}Am-^{9}Be source as a function of the detector position. The agreement in the detector center is better than 0.8 % (panel a) and decreases close to the vessel border probably due to the uncertainty in the vessel shape reconstruction (panel b). Both plots are taken from [184].

5.1.2 Position Reconstruction

The position resolution of the reconstructed radius can be seen in Figure 5.2b for different true energies and radial positions. While the resolution is almost independent on the detector position, it worsens with a decreasing energy deposition. Hence, the position resolution is ~ 9 cm and ~ 14 cm for an energy of 2.2 MeV and 1 MeV, respectively. This behavior can be well reproduced with calibration data (see for example Reference [189]).

A comparison of the reconstructed and nominal positions of the ^{214}Po-α events from a ^{222}Rn calibration source shows an overall good agreement [189]. However, a bias in the z-coordinate is present which is depicted in Figure 5.4. The so-called "z-shift" is a function of the z-coordinate itself and reaches up to ~ 4 cm. The reason of the "z-shift" is unexplained and cannot be reproduced with MC simulations. The effect of this position shift on the sensitivity of SOX will be studied in Chapter 7.

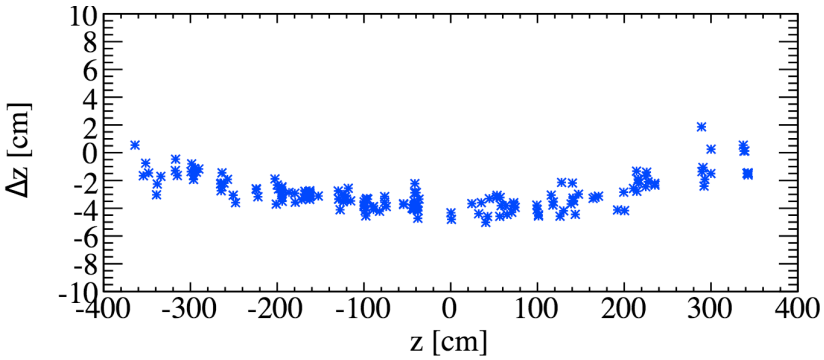


Figure 5.4: Position Reconstruction Shift Difference of the reconstructed and nominal z-coordinate determined from ^{214}Po-α events from a ^{222}Rn calibration source. This bias is only present in the z-coordinate and cannot be reproduced with MC simulations. The figure is taken from Reference [189].

5.2 Detection Efficiency

In the following, the detection efficiency for the SOX event sample is evaluated using MC simulations. The SOX source emits electron antineutrinos which are detected via the inverse beta decay (IBD) reaction in which positrons and neutrons are produced. Hence, first, the single detection efficiencies of positrons (Section 5.2.1) and neutrons (Section 5.2.2) are studied. In order to maximize the sensitivity of SOX, the selection cuts for IBD events are optimized with respect to the previous antineutrino analysis in Borexino (Section 5.2.3). In the last part, the resulting detection efficiency of the coincident IBD signal is discussed (Section 5.2.4).

5.2.1 Detection Efficiency of Positrons

From 10^6 generated positrons assuming the SOX energy spectrum and a uniform distribution within the IV, only eleven positrons are not detected with the MC simulation. These undetected positrons are located very close to the IV border so that both annihilation gammas can escape into the buffer medium. Given in addition a positron with a small kinetic energy, only a small amount of scintillation light is produced that is below the detection threshold of ∼30 hits. Figure 5.5a shows the differential detection efficiency for the positrons as a function of the true vessel distance for different cuts on the reconstructed energy and the fiducial volume (FV). A full detection efficiency is obtained in the detector center for positions with a vessel distance of at least 0.4 m. Close to the vessel border the annihilation gammas can spill out into the buffer medium and a fraction of the deposited energy is quenched. The closer the IV border, the higher is the chance for the gammas to spill out and the light yield is decreased (see also Figure 5.1). Hence, increasing the minimum energy cut to 400 p.e. decreases the detection efficiency to 86 % at the vessel border. Due to the position resolution, some positrons are reconstructed outside the IV. At the vessel border the detection efficiency is reduced by more than 50 % when a FV cut of 0 cm to the IV is required.
Figure 5.5b shows the detection efficiency integrated over the FV as a function of the FV cut for different minimum energy cuts. The integral detection efficiency is defined as the number of events detected for a FV cut over the number of events started in the same volume[3]. Note that the simulation generates only positrons within the IV. Hence the number of started events for negative FV cuts (outside the IV) stays constant. 99.3 % of the positrons started in the IV are detected within a distance of 0.4 m outside the IV for a minimum energy cut of 400 p.e.. Reducing the energy cut to 200 p.e. enlarges the detection efficiency to even more than 99.9 %. As the number of started events stays constant outside the IV, but the number of detected events are reduced for a certain FV cut, the detection efficiency decreases to 96.5 % at the vessel border for an energy cut of 400 p.e.. Thus, 2.8 % from the events above 400 p.e. are reconstructed outside the IV. The integral detection efficiency increases for larger FV cuts, as the differential detection efficiency increases as well towards the detector center (see Figure 5.5a). Moreover, the detection efficiency increases further beyond the volume with a constant differential detection efficiency (∼ 0.4 m) because of the spherical detector design and the position resolution which lead to a net inward shift.[4] The variation of the detection efficiency as a function of the minimum energy cut is largest (0.7 %) without a FV

[3] This definition of the detection efficiency is the one used in the previous geo neutrino analysis [155]. The detection efficiency expressed as the fraction of all started events which are detected for a specific FV cut is shown in Figure 5.14 for the detection efficiency of the IBD.

[4] Given a spherical uniform distribution, the event number N_{true} is proportional to r^2. N_{rec} denotes the distribution of N_{true} smeared with the position resolution Δr. The ratio $N_{\text{rec}}/N_{\text{true}}$ is largest for $r = 0$ and decreases for larger radii.

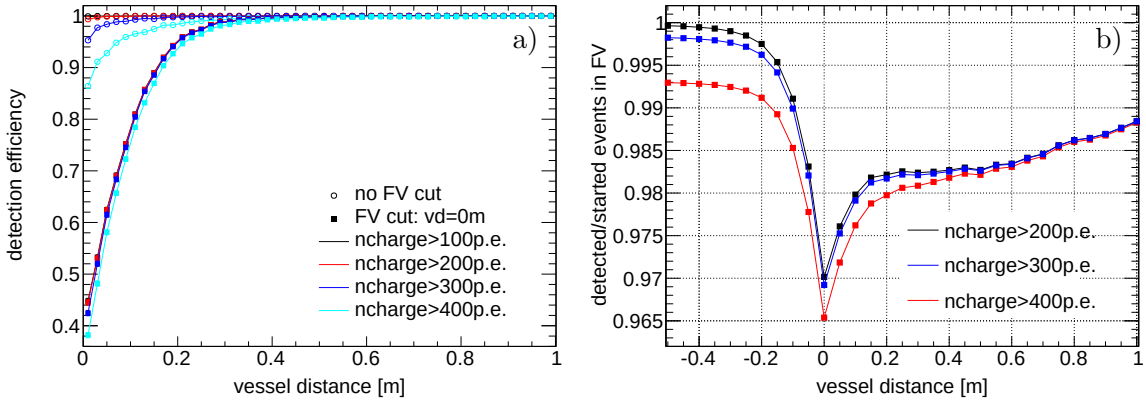


Figure 5.5: Detection Efficiency of Positrons a) The differential detection efficiency is shown as a function of the true vessel distance for different FV and minimum energy cuts. It is defined as the fraction of started events at the specific vessel distance that are detected for the given cuts. b) The integral detection efficiency is shown as a function of the FV cut for different minimum energy cuts. It is defined as the number of events detected for the specific FV cut over the number of events started in the same FV. The detection efficiencies are obtained from 10^6 positrons uniformly distributed in the IV with the energy distribution expected from SOX events.

cut and vanishes for larger FV cuts, as both annihilation gammas are completely detected within the active volume.

As the knowledge of the detection efficiency of positrons is mandatory for SOX and has not been calibrated yet in Borexino, a dedicated positron source was studied in Reference [190]. A ^{68}Ge-^{68}Ga source is an optimal candidate as the mother isotope ^{68}Ge has a long enough lifetime of 270.95 d for transportation issues and decays via electron capture [191]. ^{68}Ga produces with a probability of 88.9 % and a short half life of 67.8 min positrons with energies up to 1.9 MeV that cover the whole energy range of the prompt event [191] (the residual 11.1 % decay via electron capture). The ^{68}Ge-^{68}Ga source was planned to be deployed at several positions in the detector with a focus close to the IV border.

5.2.2 Detection Efficiency of Neutrons

The overall detection efficiency behavior is similar for neutrons as for positrons. However, in comparison to positrons, a full detection efficiency is reached for neutrons at larger vessel distances at about 1 m, which is visible in Figure 5.6a. This is due to the difference in the mean path length of the positron and the neutron (including the de-excitation gamma). The mean difference between the reconstructed position and the start position is about 2.4 times larger for neutrons (38 cm) than for positrons (see Figure 5.7a). While positrons annihilate close to the IBD reaction position and create two low energy gammas in the opposite direction, the mean difference is only 16 cm which is only slightly larger than the position resolution. Instead, neutrons need to thermalize before being captured and produce gammas with at least 2.2 MeV that have a path length of several tens of centimeters (the mean difference from reconstructed to neutron capture position is 35 cm).

The large mean free path length allows the gammas to spill out in the buffer medium and produce only a small amount of scintillation light that is below the detection threshold of ~ 30 hits. This leads to a differential detection efficiency of 70 % for neutrons started at the vessel border even though no energy cut is applied. The integral detection efficiency without an energy and a FV cut reaches 96.0 %. The detection efficiency is shown as a function of the

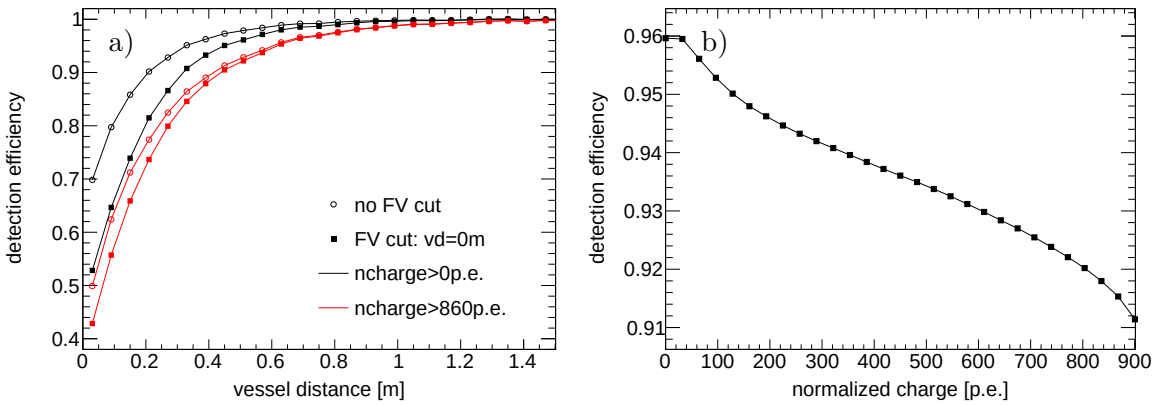


Figure 5.6: Detection Efficiency of Neutrons a) The differential detection efficiency is shown as a function of the true vessel distance for different FV and minimum energy cuts. b) The integral detection efficiency is given as a function of the minimum energy cut without applying a FV cut. The results are obtained from 10^6 started neutrons uniformly distributed in the IV. 4 % of the started neutrons remain undetected as they spill out in the buffer and produce a signal below the detection threshold.

minimum energy cut on the delayed event in Figure 5.6b. Applying a cut at 860 p.e., which is the one used for the latest geo neutrino analysis, results in 91.6 %. Hence, reducing the minimum energy cut would allow to increase the statistics by up to 4.4 % and would hence increase the sensitivity of the experiment. The optimization of the energy cut is discussed in the next section. Note that the detection efficiency of neutrons as well as the ones of positrons depends on the exact vessel shape which is discussed in more detail at the end of this chapter.

5.2.3 Optimization of the Selection Cuts for the Inverse Beta Decay

The latest geo neutrino analysis of Borexino (before the planned start of SOX) was performed in 2015 and achieved a detection efficiency of 84.2 % with a FV cut of 30 cm [155]. As the geo neutrino event rate is < 10 events per year, the selection cuts need to be strict in order to minimize the background for the analysis. However, the expected statistics of the SOX event sample is about 10^4 events and the selection cuts can hence be released in order to maximize the statistics and the sensitivity of the experiment. In the following, the optimization of the selection cuts is discussed. First, the energy cuts for positron and neutron events are studied. Second, the cut on the spatial and time difference is evaluated. Third, the pulse shape discrimination cut on the neutrons is optimized.

Energy of Prompt and Delayed Event

In the latest geo neutrino analysis, the selection cut on the energy of the prompt and the delayed event was set to E_p>408 p.e. and E_d=860 p.e.-1300 p.e., respectively.
The minimum energy cut for the energy of the prompt and the delayed event for the SOX analysis is re-defined because of the following considerations:

- Detection Efficiency: the lower the energy cut the higher the detection efficiency becomes (see Figure 5.5, 5.6, and 5.12 in Section 5.2.1, 5.2.2, and 5.2.4, respectively);
- Uncertainty on Detection Efficiency: the uncertainty reduces for smaller energy cuts (see Figure 5.12, 5.13 in Section 5.2.4, Figure 7.8 in Section 7.2.3, and Figure 8.10 in

Section 8.3.1);

- Spill In Events: a fraction of the IBD reactions in the buffer are reconstructed in the IV and contribute to the SOX data. Such spill in events produce typically a small amount of scintillation light and can be removed from the data by a high energy cut (see Figure 5.15 in Section 5.2.4);
- Background: the higher the energy cut the smaller the expected background events become (see Figure 6.1 in Section 6.1);
- Sensitivity: given the behavior of the detection efficiency and the background, the sensitivity increases for smaller energy cuts. This is mainly due to the larger expected event number for the SOX event sample (see Figure A.5 and A.6 in Appendix A.3);
- Algorithm Performance: the performance of the position reconstruction and pulse shape discrimination decreases for smaller energies (see Figure 8.2 in Section 8.1).

The cut value for the prompt event is set at 300 p.e., as the detection efficiency and sensitivity stays almost constant for smaller values. To achieve a sufficient algorithm performance, a conservative minimum energy cut of 200 p.e. is defined for the delayed event. The maximum energy cut of the delayed event is extended to 3000 p.e. in order to detect also the neutrons that get captured on Carbon with a probability of 1.1 % and produce a 4.95 MeV gamma [35]. A detailed discussion of the individual points is given in the referenced sections.

Spatial Distance

The distribution of ΔR, i.e. the reconstructed distance between the prompt and the delayed event, is shown in Figure 5.7a. The mean distance is 42.1 cm which is only slightly higher than the mean path length of the neutrons (see Section 5.2.2). The cut efficiency as a function of the ΔR_{max} cut parameter is depicted in Figure 5.7b which is defined as the fraction of the reconstructed IBD coincidences with $\Delta R < \Delta R_{\mathrm{max}}$. The cut efficiency can be increased from 96.2 % at 1 m (which is the cut parameter from the latest geo neutrino analysis) to 99.4 % at 1.5 m. As the number of accidental background events increases with ΔR_{max} (see next chapter), the sensitivity of SOX is studied as a function of ΔR_{max}. Figure 5.8a shows the expected signal and background event number as a function of ΔR_{max}. The background event number is obtained from Borexino data, which is discussed in the next chapter. Both event numbers increase with ΔR_{max}. However, as the signal to background rate is about 200, the sensitivity improves still for larger ΔR_{max} values. As the sensitivity improves only marginally beyond 1.5 m, the optimized ΔR_{max} cut-value is set to 1.5 m, where the sensitivity is improved by about 1 % compared to 1 m, which is depicted in Figure 5.8b.

Time Difference

The distribution of the time difference Δt between the prompt and the delayed event follows an exponential function with the neutron capture time of 254.5 ± 1.8 µs [35]. With the cuts of the previous geo neutrino analysis of Δt = 20 µs - 1280 µs, the Δt-cut efficiency is given by 91.8 %. Thus, using also the events with Δt < 20 µs, the detection efficiency can be enhanced. However, the minimum Δt cut depends on the specific Borexino data acquisition properties: the length of one data acquisition gate (called event) and the duration of one physical event (called cluster). The hit-time distribution of one event with two clusters is shown in Figure 5.9a. The gate length is set in Borexino to 16 µs, where the physical

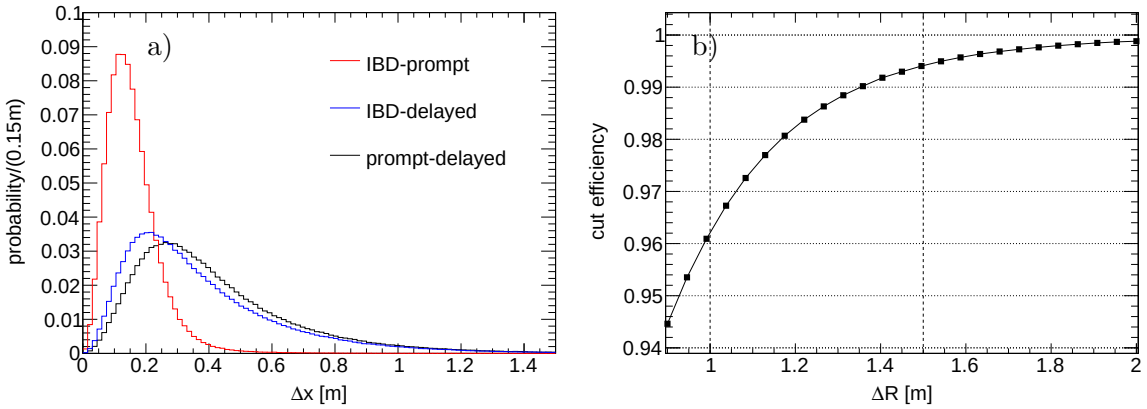


Figure 5.7: Spatial Distance Distributions and Cut Efficiency of ΔR a) The probability distributions of the difference of the reconstructed positions of the prompt and delayed event to the true IBD position and of ΔR is shown. b) The efficiency is shown as a function of the cut parameter ΔR. The dashed lines indicate the re-defined cut-value and the one used in the latest geo neutrino analysis.

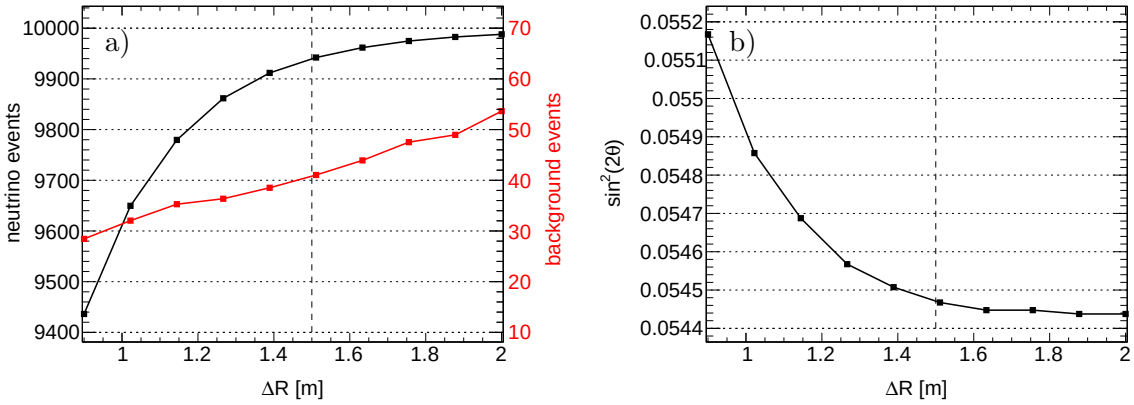


Figure 5.8: Optimization of the ΔR-cut a) The SOX and the background event numbers are given as a function of the ΔR-cut value. The SOX event number assumes a total event number of 10^4. The background event number is obtained from Borexino data selected for the optimized cut values without a FV cut (see next chapter). b) The sensitivity is shown in terms of the oscillation amplitude $\sin^2(2\theta)$ at $\Delta m^2 = 10\,\mathrm{eV}^2$. The sensitivity is computed as described in Appendix A.5 and assumes an uncertainty in the signal and background event number of 1 % and 10 %, respectively. The dashed line indicates the re-defined cut parameter.

event that issues the Borexino trigger is located 1 µs after the start of the gate. The cluster algorithm searches for an accumulation of hits in a short time period, where only the hits reconstructed in a cluster contribute to the energy estimators described at the beginning of this chapter. To achieve a high efficiency in the pulse shape discrimination and observe the full scintillation pulse, the length of the reconstructed cluster is pre-defined to 1.5 µs [141]. However, depending on the exact hit-time distribution, the duration can be smaller or larger than the nominal length. Figure 5.9b shows the distribution of the cluster duration for Borexino data in the energy range of SOX events (400 p.e. - 3000 p.e.), where the maximum value is 1.52 µs. Even though the cluster algorithm is able to resolve two clusters with a full efficiency already at a time difference of 0.55 µs (see Figure 5.10), $\Delta t^{\mathrm{min},1}$ is set to the maximum cluster duration of 1.52 µs to ensure an undistorted energy response.

Due to the data acquisition gate length of 16 µs and the following dead time of the data

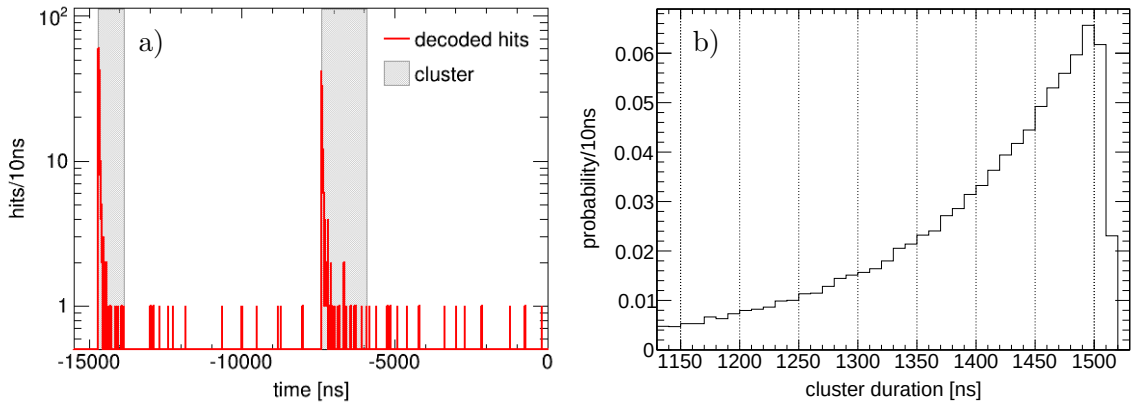


Figure 5.9: Event Structure and Cluster Duration a) Hit-time distribution of one Borexino event (gate length $\sim 16\,\mu s$) with two clusters, where the time is shown with respect to the trigger time. The energy and the duration of the first (second) cluster is reconstructed to 601 p.e.(436 p.e.) and 844 ns (1493 ns) respectively. b) The probability distribution of the cluster duration is shown for Borexino events with an energy in the prompt and delayed energy window (400p.e. - 3000p.e.). The duration reaches a maximum value of 1520 ns which defines thus the $\Delta t^{\mathrm{min},1}$-cut value.

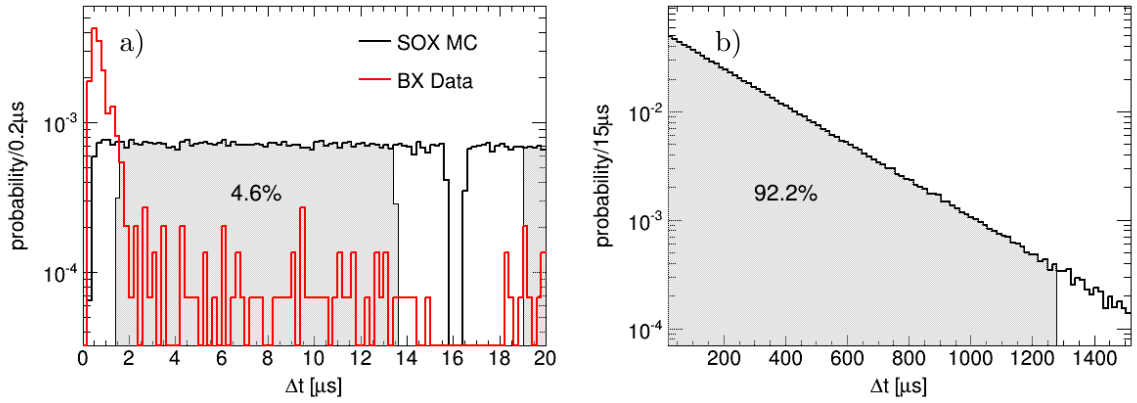


Figure 5.10: Cut Efficiency of Δt-cut The time difference between prompt and delayed event (Δt) is shown for double cluster events (Δt<15 µs in panel a)) and single cluster (Δt>18 µs in panel b)) events obtained from MC simulations. The distribution follows an exponential decay with the neutron capture time of 254.5 µs [35]. In addition, the time difference of correlated events (mainly ^{212}Bi-^{212}Po and ^{214}Bi-^{214}Po) recorded with Borexino is shown in red in arbitrary units. These events show the dead time after an gate from 15 µs - 18 µs, which is not fully implemented in the Borexino MC simulation. The dead time together with the cluster duration (see Figure 5.9) define the optimized cut parameters of Δt: 1.52 µs - 13.48 µs or 19 µs - 1280 µs which yield a total cut efficiency of 96.7 %.

acquisition system of $\sim 3\,\mu s$ (see Figure 5.10), coincidences with Δt: 13.48 µs - 19 µs are excluded from the analysis. The first value is chosen in order to observe the full cluster of the delayed event (i.e. 13.48 µs = 15 µs - 1.52 µs) and thus to ensure a full efficiency for the pulse shape discrimination of the delayed event. In comparison to the latest geo neutrino analysis, $\Delta t^{\mathrm{min},2}$ could be lowered from 20 µs to 19 µs as the exchange of the Borexino trigger board in May 2016 reduced the dead time after a data acquisition gate. The maximum time difference $\Delta t^{\mathrm{max},2}$ remains unchanged and is set to five times the neutron capture time, i.e. 1280 µs.

In summary, Δt is optimized to 1.52 µs - 13.48 µs or 19 µs - 1280 µs and the cut efficiency of Δt results into 96.7 %. This value is enhanced by 4.9 % in comparison to the latest geo neutrino analysis, which is mainly due to the inclusion of the double cluster events (i.e. $\Delta t <$ 15 µs).

Pulse Shape Discrimination

The main background for IBD-like events are fast coincidences from the ^{232}Th- and ^{238}U-chain (see Chapter 6). However, the delayed events in these fast coincidences are alpha particles and can be discriminated from neutrons by means of pulse shape discrimination techniques. The goal of the pulse shape discrimination is to achieve a high neutron acceptance while at the same time rejecting alphas with a large probability. In the latest geo neutrino analysis, the *gatti* parameter was used. The *gatti* parameter is defined as the weighted sum of the measured hit-time distribution, where the weights are defined as the difference of the probabilities that a photoelectron is detected at a certain time given an alpha or a beta particle over the sum of the two probabilities [141]. In the meantime, the new pulse shape discrimination parameter multi-layer perceptron (*mlp*) has been developed

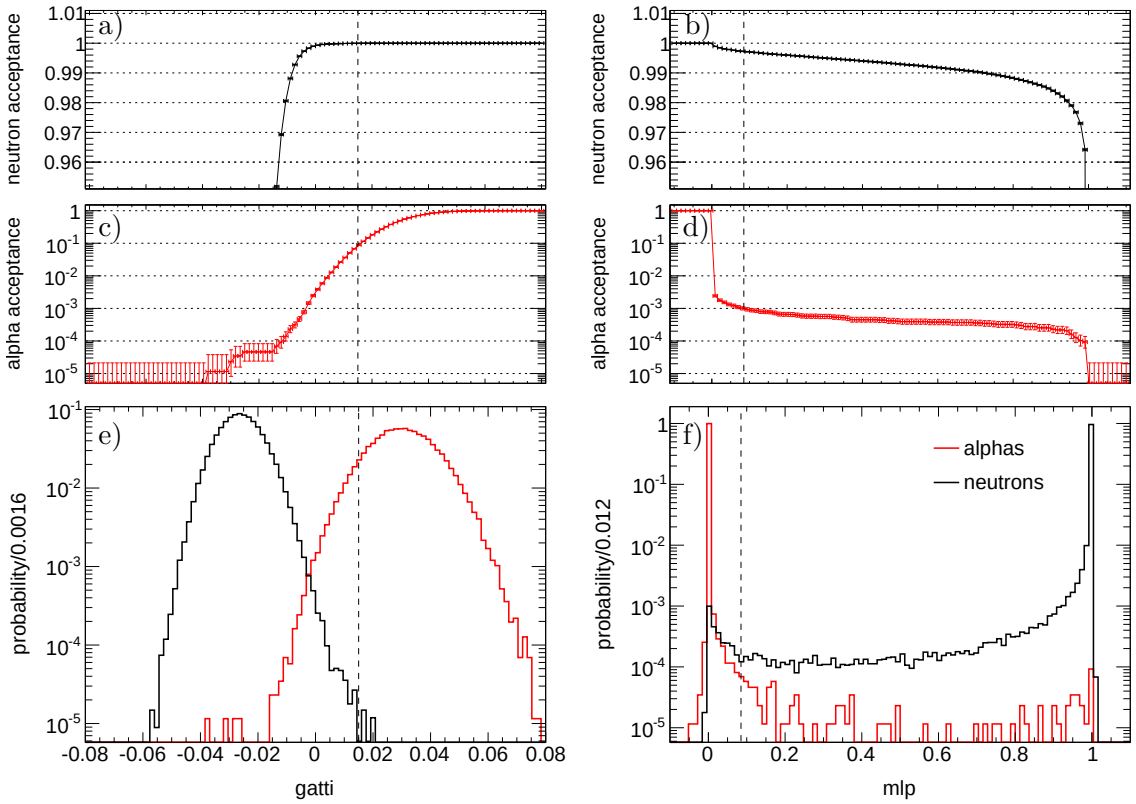


Figure 5.11: Performance of the *gatti* and *mlp* Parameter Panel e and f show the probability distributions of the pulse shape parameters for neutrons (MC simulation) and alpha particles (^{214}Po data during purification campaign [141]). The neutrons are selected with a minimum energy cut of 200 p.e. and the ^{214}Po events by the fast coincidence with ^{214}Bi (energy(^{214}Bi): 200 p.e.-1800 p.e., energy(^{214}Po): 200 p.e.-500 p.e., Δt: 20 µs-944 µs, $\Delta R <$ 1 m), where no FV cut is applied to both. Panel a and b (c and d) show the fraction of neutrons (alphas) accepted as a function of the pulse shape parameter cut ($gatti < gatti_{\mathrm{max}}$ and $mlp > mlp_{\mathrm{min}}$). The dashed lines indicate the cut values used in the latest geo-neutrino analysis (*gatti*) and the optimized value for the SOX analysis (*mlp*).

which is based on machine learning techniques and uses 13 characterizing input variables that are computed from the hit-time distribution of one event [192]. The performance of the pulse shape discrimination parameters is shown in Figure 5.11. The neutron acceptance is studied with simulated neutrons uniformly distributed in the whole IV. The rejection of alpha particles is studied with ^{214}Po data that has been selected by the fast coincidence with ^{214}Bi. The probability distributions of the *gatti* parameter follow almost Gaussian distributions, where for beta (alpha) particles the mean value is negative (positive). While the distributions of the *gatti* parameter partially overlap, the distributions of the *mlp* parameter are well separated (*mlp* is "0" for alpha particles and "1" for beta-like particles) which leads to an enhanced discrimination power.
The latest used *gatti* parameter cut of $gatti < 0.015$ yields a full neutron acceptance (99.996 %), but a rather high alpha acceptance of 8.3 %. The optimized *mlp* cut is defined as $mlp > 0.085$ so that an alpha acceptance of 0.1 % is achieved. However, the drawback is that the neutron acceptance is lowered to 99.72 %. It has to be noted that the neutron acceptance depends on the detector position and is higher in the detector center (99.92 % for $r < 3\,\mathrm{m}$ and 99.58 % for $r > 3\,\mathrm{m}$). As the current neutron acceptance of 99.72 % is obtained without applying a FV cut (and for neutrons above 200 p.e.), the performance of the *mlp* parameter would improve when a FV cut is required. Further, the *mlp* parameter could be improved by directly training the *mlp* parameter with neutrons instead of electrons, as for example from the ^{241}Am-^{9}Be calibration data.
Even though the obtained values from MC simulations agree with the ones from ^{241}Am-^{9}Be calibration data [35], the precise performance of the *mlp* parameter as a function of the detector position and energy has to be studied with new calibration data. While the background is not critical for the SOX analysis (see Chapter 7), a precise knowledge on the neutron acceptance is mandatory. Depending on the calibration data results and given the case that the calibration data agrees better with the *gatti* than with the *mlp* parameter, one could also reconsider a cut on the *gatti* parameter.

5.2.4 Detection Efficiency of the Inverse Beta Decay

The differential detection efficiency for the optimized selection cuts is shown in Figure 5.12 as a function of the true vessel distance. In the detector center a detection efficiency of 96.1 % is achieved which is given by the product of the cut efficiencies of ΔR (99.4 %) and Δt (96.7 %). The detection efficiency lowers towards the vessel border to $\sim 55\,\%$ ($\sim 33\,\%$ when a FV cut of 0 cm is applied) which is a combination of the loss of the detection efficiency of positrons and neutrons.
The detection efficiency is shown for four different detector regions divided by the polar angle of the detector. The different vessel shapes in the different regions lead to variations of the detection efficiency up to 6 % close to the vessel border. The variations are even larger when a higher energy cut of 860 p.e. on the delayed event (the value from the latest geo neutrino analysis) is applied. Thus, an unknown vessel shape results into an uncertainty of the detection efficiency close to the vessel border. The comparison of data and MC simulation using the ^{241}Am-^{9}Be calibration data shows that increasing discrepancy for smaller vessel distances and can be seen in Figure 5.13. The discrepancy reaches up to 10 % (4 %), when a delayed energy cut of 860 p.e. (200 p.e.) is applied. The uncertainty on the detection efficiency decreases with a lower energy cut, as the integral detection efficiency increases and is then effectively less affected by the exact vessel shape (see also Figure 7.8). In addition to an improved vessel shape reconstruction algorithm, it is mandatory to calibrate the detector and understand its behavior close to the vessel border. The idea of applying no minimum

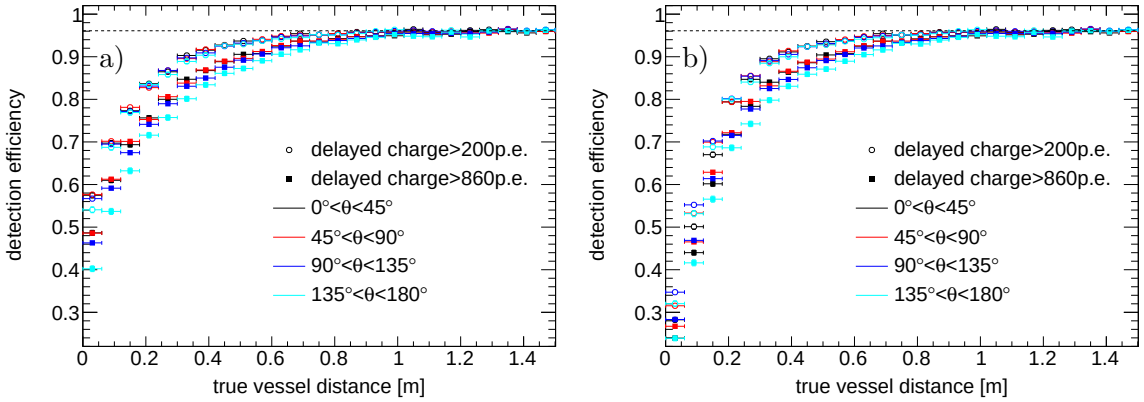


Figure 5.12: Differential Detection Efficiency The differential detection efficiencies of IBD events that survive the optimized selection cuts are shown as a function of the true IBD position. Panel a) shows the detection efficiency when no FV is applied, panel b) when the prompt event is reconstructed within the IV. The efficiencies are shown for a minimum energy cut of 200 p.e. and 860 p.e. and for four detector positions (separated by the polar angle θ). The results are obtained from MC simulations assuming a uniform distributed IBD distribution within the IV. Due to the different vessel shapes at the different detector positions, the detection efficiency differs close to the vessel border, where the difference is larger for the higher energy cut. The central value of 96.1 % can be explained by the ΔR and Δt cut efficiencies.

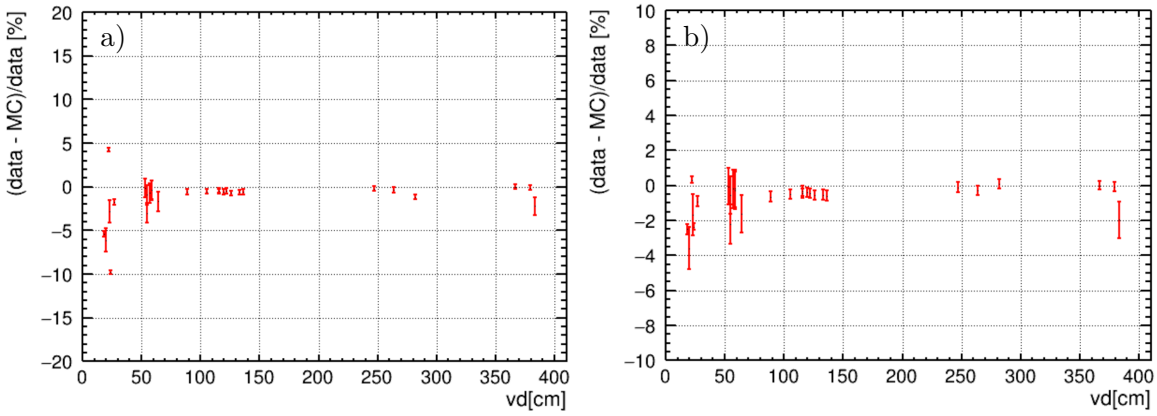


Figure 5.13: Accuracy of the Detection Efficiency The discrepancy of data and MC data from an ^{241}Am-^{9}Be source is shown for a minimum energy cut on the delayed event of 860 p.e. (a) and 200 p.e. (b) as a function of the vessel distance (vd). The discrepancy increases towards the vessel border, due to the uncertainty in the vessel shape reconstruction, and decreases for a lower energy cut. Note that the range of the y-axis is different in the two plots. The plots are taken from [193].

energy cut on the delayed event at all and hence reducing the uncertainty on the detection efficiency is extended in Chapter 8 by laying out a possible hardware modification of the Borexino trigger system.

The integral detection efficiency as a function of the FV cut is shown in Figure 5.14a in red, where the detection efficiency is computed as the ratio of the number of detected events over the number of started events for the same FV cut. Applying no FV cut, 90.8 % of all IBD reactions taking place in the IV can be detected. Due to spill out events, the detection efficiency reduces to 88.4 % when a FV cut of 0 cm is applied. The larger the FV cut, the higher the detection efficiency becomes. With the improved selection cuts, the detection

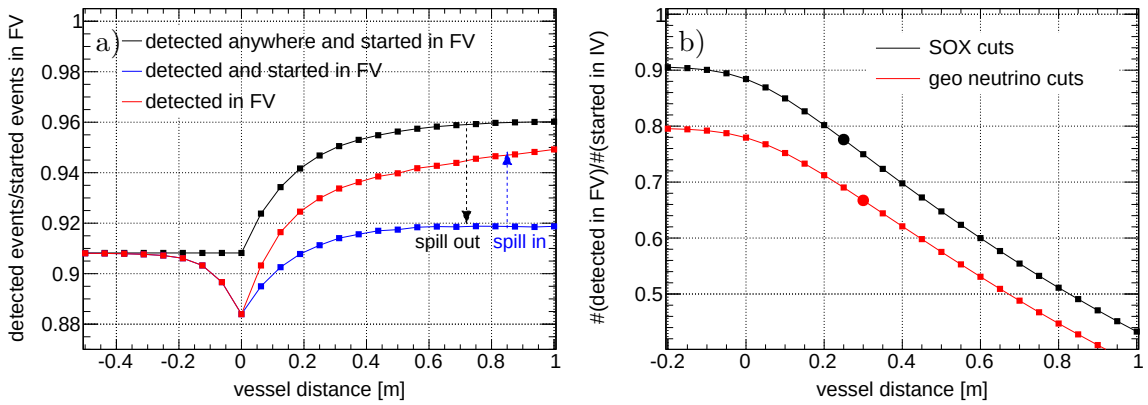


Figure 5.14: Integral Detection Efficiency The number of detected events is shown as a function of the FV cut with respect to the started event number in the a)FV and in the b) IV. In a) the number of detected events are given for the optimized selection cuts in red. The blue line requires in addition that the detected events are also started in the FV. The detected events for the black line can be detected anywhere, but are started in the FV. The detection efficiency in b) is computed for the optimized SOX cuts and the cuts from the latest geo neutrino analysis. The circles indicate the values for the respective FV cuts. The results are obtained from MC simulations assuming a a) uniform and b) SOX-like IBD distribution within the IV.

efficiency could be increased from 84.2 % to 93.3 % for the FV cut of the previous geo neutrino analysis of 30 cm. This corresponds to an improvement of 10.8 %.
The expected event number of the SOX event sample can be further increased by the optimization of the FV cut. One can define the integral detection efficiency then as the fraction of started events in the IV that are reconstructed in a certain FV which is shown in Figure 5.14b. The efficiency decreases for larger FV cuts and results into 77.6 % for the optimized FV cut of 25 cm. The optimization is discussed in Chapter 7 and takes into account the interplay between the increasing statistics and uncertainties for smaller FV cuts. On the one hand, one would like a small FV cut to increase as much as possible the statistics of the SOX event sample. On the other hand, the uncertainty of the detection efficiency increases close to the vessel border and decreases the sensitivity of the experiment. With the optimized selection cuts the statistics of the SOX event sample increases by 16.3 %.
For the optimized FV cut of 25 cm a detection efficiency of 93.0 % is achieved with respect to the same volume. Figure 5.14a shows in addition two more efficiencies: the fraction of events for a given FV cut that are detected anywhere and that are detected within the same FV. The differences in the efficiencies show the effect of spill in and spill out events, where the latter dominate. Moreover, it shows that spill in events need to be taken into account, and that in MC simulations the events need to be started in an extended volume by about 30 cm in comparison to the FV. Spill in events due to IBD interactions in the buffer are hence studied in the following.

Spill In Events from the Buffer

IBD reactions taking place in the buffer can be reconstructed within the FV and contribute to the observed data sample. IBD-like candidates are selected using the optimized selection cuts from MC simulations, where events are started in the volume up to the outer vessel. The increase in the event number due to the buffer events in comparison to the one expected from the IV is shown as a function of the FV cut for several minimum energy cuts on the prompt event in Figure 5.15. The higher the minimum energy cut, the smaller the amount

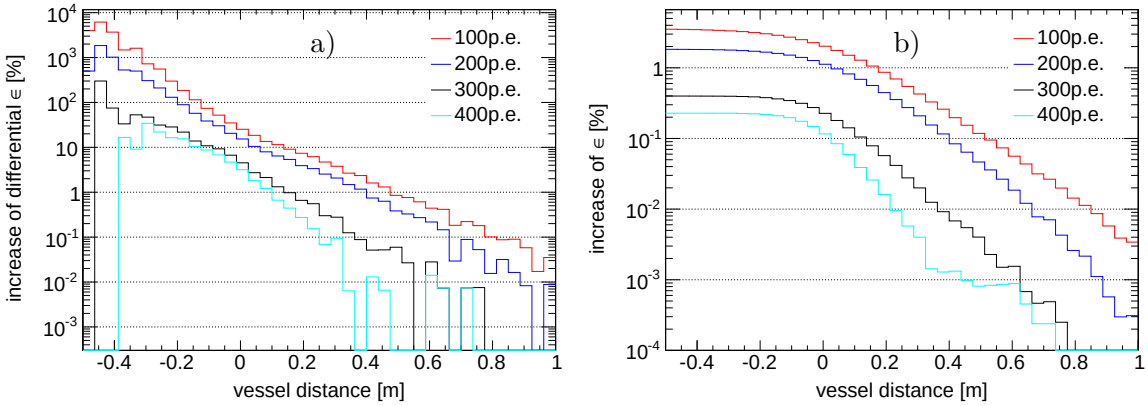


Figure 5.15: Increase of Event Number due to Spill in Events from IBD Reactions in the Buffer The increase of the a) differential and b) integral event number is shown as a function of the vessel distance and the minimum energy cut on the prompt event. The increase is computed with respect to the expected event number from IBD reactions in the IV from MC simulations. For the optimized cuts (300p.e. and 25 cm) an integral increase of 0.03 % is expected. The differential increase is rather high outside the IV, because of the low expected event number from the IBD reactions from the IV.

of spill in events becomes. While the increase of the integral event number is 0.03 % for the optimized energy and FV cut (300 p.e. and 25 cm), it is by a factor of ~ 10 larger for the same FV cut and an energy cut of 200 p.e. The lower the minimum energy cut, the more events are reconstructed within the FV even for large FV cuts, as the position reconstruction becomes worse for such small energies (see Figure 8.2).

Dead Time due to Borexino Events

It can be assumed that the detection efficiency is not affected by the dead time introduced by the Borexino standard events. The trigger rate in Borexino is ~ 19 Hz (October 2017). Hence, the rate of SOX events that accidentally fall into a standard Borexino event at the beginning of data taking can be computed to: $19\,\text{Hz} \times 36/\text{day} \times 19\,\mu\text{s} = 0.014$ events/day. Here, 36/day is the expected SOX rate at the beginning of data taking for a 125 kCi source and 19 µs is the data acquisition gate length including the following electronics dead time. For a total measurement time of 1.5 y, 4.1 events are expected which is negligible in comparison to $\sim 10^4$ expected SOX events.

5.3 Summary

The Borexino detector is well characterized and the systematics of the detector are small. However, the systematic uncertainties, which are obtained from the calibration campaign with radioactive sources in 2009 [189], increase towards the vessel border. Hence, a dedicated SOX calibration campaign was planned to especially characterize the part of the detector close to the vessel border and minimize the systematic uncertainties. The effect of the systematic uncertainties on the sensitivity of SOX is studied in Chapter 7.
Furthermore, the selection cuts of the IBD detection channel were optimized in this chapter and are summarized in Table 5.1. For the optimized FV cut of 25 cm, a detection efficiency of 93.0 % is achieved with respect to the expected event number of the same volume. This corresponds to an improvement of 10.8 % in comparison to the previous geo neutrino analysis,

Table 5.1: Optimized Selection Cuts for the SOX Analysis The selection parameters are listed in the first column, where their definitions can be found in Table 4.1. The second column quotes the selection cuts used in the previous geo neutrino analysis in Borexino [155]. The optimized selection cuts for the SOX analysis are given in column three. The detection efficiency in the last row is obtained from SOX MC simulations and is defined as the fraction of all generated events in the IV which survive the specified selection cuts.

parameter	geo neutrino analysis	SOX analysis
E_p	> 408 p.e.	300 p.e. - 1300 p.e.
E_d	860 p.e. - 1300 p.e.	200 p.e. - 3000 p.e.
Δt	20 µs - 1280 µs	1.52 µs - 13.48 µs or 19 µs - 1280 µs
ΔR	< 1.0 m	< 1.5 m
pulse shape	$gatti < 0.015$	$mlp > 0.085$
vd	> 30 cm	> 25 cm
efficiency	66.7 %	77.6 %

which comes mainly from the relaxed ΔR, Δt, and E_d cut. Due to the enlarged FV cut, the expected event number for the SOX event sample is increased in total by 16.3 %. The statistics of the SOX event sample can be even more increased by a new muon veto strategy which is discussed in the next chapter.

6 Investigation of Background Sources

This chapter evaluates the expected background for SOX which needs to be taken into account for the sterile neutrino analysis. As antineutrinos are detected through the inverse beta decay (IBD) and the corresponding coincidence signal, the background of single Borexino events is suppressed for an antineutrino search. However, a non-reducible background remains that is given by natural antineutrinos, namely geo and reactor neutrinos, and their backgrounds. An advantage of the SOX experiment is that the Borexino detector has been already taking data for more than ten years before the planned start of the SOX experiment and the backgrounds can be precisely characterized. As the selection cuts differ from the previous geo neutrino analyses, the detector-related background is studied for these selection cuts in Sections 6.1-6.6. First, the analyzed Borexino data and the fit method to determine the main background sources are introduced. Subsequently, the individual background sources of neutrino events, cosmogenic isotopes, accidental background, and fast coincidences from radioactive decay chains are evaluated. Moreover, additional detector-related background is summarized. Lastly, source-induced background is studied in Section 6.7.

6.1 Borexino Data Sample and Simultaneous Fit

In this analysis the data from December 15, 2011 until December 31, 2016 corresponding to a total measurement time of 1516.6 d (4.16 y) is investigated. This data is collected in the so-called Borexino Phase-II that denotes the period after an extensive purification campaign of the liquid scintillator in which the intrinsic background has been significantly reduced. The ^{238}U and ^{232}Th contaminations decreased to $< 9.4\times10^{-20}$ g/g (95 % CL) and $< 5.7\times10^{-19}$ g/g (95 % CL), respectively. Moreover, ^{85}Kr and ^{210}Bi concentrations were lowered by a factor of ~ 4.6 and ~ 2.3, respectively [139].

In comparison to a geo neutrino analysis with an expectation of $\mathcal{O}(10)$ events per year, the expected event number for SOX is $\mathcal{O}(10^4)$. Hence, from the background point of view it is affordable to soften the selection cuts in comparison to the geo neutrino analysis and improve at the same time the selection efficiency for the SOX event sample. In Chapter 5, the optimization of the selection cuts is discussed, which are summarized in Table 5.1.

Using the `bx-sox` filter (see Chapter 4) with the optimized cuts, 52 IBD events are found in the analyzed period. This corresponds to an expected background event number of 18.8 ± 2.6 for the SOX measurement time (1.5 y). Figure 6.1 shows the measured distributions of these events (black data points) for various parameters, namely the energy of the prompt (E_p) and delayed (E_d) event, the difference in space (ΔR) and time (Δt) in between prompt and delayed event, and the detection position of the prompt event given in distance (L_{rec}) to the source position (8.51 m below the detector center) and the vessel distance (vd).

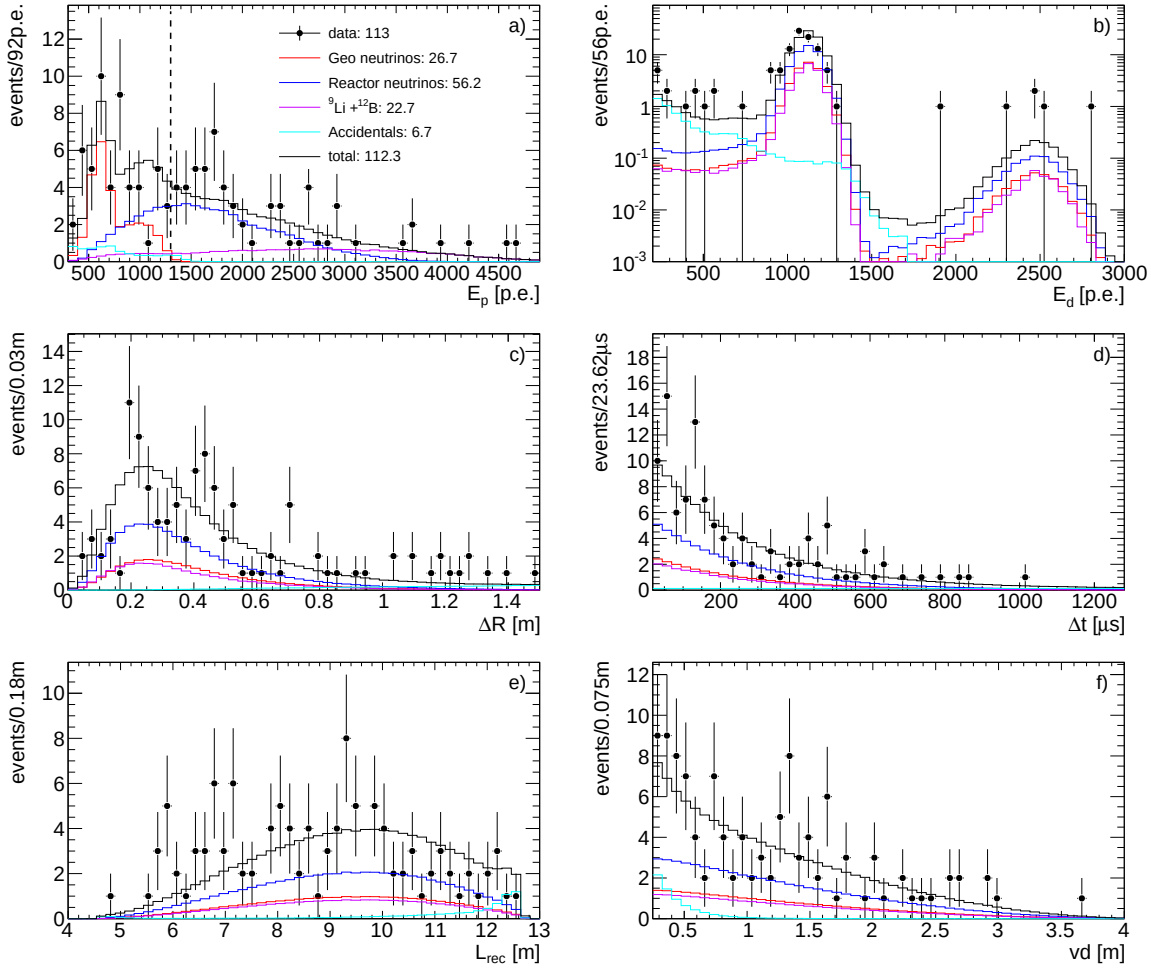


Figure 6.1: Simultaneous Fit of IBD Events in Borexino. The data (black data points) is selected from 4.16 y with the selection cuts from Table 5.1 and can be well described with a binned likelihood fit by the sum (black) of geo neutrinos (red), reactor neutrinos (blue), cosmogenic isotopes (violet), and accidental coincidences (light blue). The fit minimizes simultaneously the spectral distributions of the energy of the prompt (E_p) and delayed (E_d) event, the difference in space (ΔR) and time (Δt) in between prompt and delayed event, and the detection position of the prompt event given in distance (L_{rec}) to the source position and the vessel distance (vd). The event number of cosmogenic and accidental background is determined prior to the fit and fixed. The dashed line in panel a) indicates the maximum energy of the prompt event for the SOX experiment. The error bars give the statistical errors.

As the measurement time of the present data sample is about a factor of 2.8 larger than the SOX measurement time, the total expected background event number can be assumed accurate enough for a rate analysis in a sterile neutrino search. However, in a shape analysis, the expected spectral distributions (probability distributions), especially of the energy of the prompt event and the source distance, have to be precisely known. As the statistics of the selected data sample are too small that the distributions of the data could be directly used as the probability distributions for the SOX analysis, they are determined with a simultaneous fit. Several observables are fitted simultaneously: E_p, E_d, ΔR, Δt, L_{rec}, and vd. The fit includes the main background components, namely geo neutrinos, reactor neutrinos, cosmogenic background (mainly ^{9}Li), and accidental coincidences. A binned likelihood fit (see

Chapter 3 Equation 3.6) with the expected event number per bin i is performed

$$N_i^{\mathrm{exp}}(N_{\mathrm{geo}}, N_{\mathrm{rea}}, N_{\mathrm{cos}}, N_{\mathrm{acc}}) = \sum_{b\in\{\mathrm{geo,\ rea,\ cos,\ acc}\}} N_b \cdot pdf_i^b, \tag{6.1}$$

where i spans over all bins of all observables $j \in \{E_p, E_d, \Delta R, \Delta t, L_{\mathrm{rec}}, vd\}$: $i \in [1, \sum_j \mathrm{bins}_j]$ with bins_j the number of bins for the observable j. $N_{\mathrm{geo}}, N_{\mathrm{rea}}, N_{\mathrm{cos}}, N_{\mathrm{acc}}$ are the total event numbers of the geo neutrino, reactor neutrino, cosmogenic and accidental components. While the geo neutrino and reactor neutrino components are determined by the fit, the cosmogenic and accidental components are determined in an independent way and fixed in the fit (see Sections 6.3 and 6.4). The probability distributions pdf_i^{geo}, pdf_i^{rea}, pdf_i^{cos}, pdf_i^{acc} are normalized to one for each observable ($\sum_{i\in\mathrm{bins}_j} pdf_i = 1$). Note that the probability distribution of the cosmogenic background assumes only the spectrum of the dominating cosmogenic background ^{9}Li. However, N_{cos} is the sum of all cosmogenic contributions.
The probability distributions for geo neutrino, reactor neutrino, and cosmogenic background which were generated with the full Borexino Monte Carlo (MC) simulation (see Chapter 4) for the latest geo neutrino publication [35] are used. The simulation starts uniformly distributed events within the inner vessel, where the shape of the inner vessel and the detector configuration (e.g. number of active photomultiplier tubes (PMTs)) changes within the simulation to obtain the average expected detector response. The probability distribution of the accidental coincidences can be directly obtained from the data with an off-time window selection (Δt = 2-20 s).
To properly determine the individual components, the energy window of the prompt event is increased to the maximum energy of the reactor spectrum up to 4900 p.e. ($\sim$ 10 MeV). Afterwards, the fraction in the SOX energy range (below 1300 p.e.) is extracted.
Figure 6.1 shows the result of the simultaneous fit. The total event number of 112.3 ± 10.6 and the spectral distributions of the fit agree well with the data with a total event number of 113 ± 10.6. The fit result corresponds to an expected event number of 19.7 ± 2.7 for the SOX energy window and measurement time. The results of the individual components are listed at the end of this chapter in Table 6.2. Before, these individual background components are in detail discussed in the following sections.

6.2 Neutrino Events

Antineutrinos have been extensively studied with the Borexino detector [35,51,155–157,194]. These events are indistinguishable from SOX events and will contribute to the SOX data sample. The SOX analysis benefits from the experience of the previous antineutrino analyses. However, the relaxed selection cuts in SOX in comparison to the previous analyses make a reevaluation of these events necessary. After the discussion of geo and reactor antineutrinos, the expected background from atmospheric neutrinos, whose neutrinos can also produce an IBD-like signal, is studied.

6.2.1 Geo Neutrinos

Geo neutrinos are generated in radioactive decays from isotopes inside the Earth (see Section 1.3.1). Due to the IBD detection channel, only antineutrinos from the decay chains of ^{238}U and ^{232}Th with energies above 1.8 MeV can be detected. Their expected energy spectrum at the Borexino site was computed in [35] and extends up to $\sim$ 3.3 MeV which is slightly above the endpoint of ^{144}Pr of 3 MeV (see Figure 1.6). Hence, almost all (99.4 %) geo

neutrino events fall in the energy window of SOX and contribute hence to the background for SOX.
The expected geo neutrino event number is extracted from the simultaneous fit to 26.7 ± 6.5, where in the geo neutrino probability distribution the Th/U mass ratio is fixed (see Section 2.1.5). This event number corresponds to 9.6 ± 2.3 events for the SOX energy range and measurement time of 1.5 y. The determined value is in good agreement with previous measurements of $8.5^{+1.7}_{-1.3}$ (this value is rescaled using Equation 6.4 and the result from Reference [35].).

6.2.2 Reactor Neutrinos

Reactor antineutrinos are produced through beta decays from fission products in nuclear power plants (see Section 1.3.1). Even though there are no nuclear power reactors in Italy, Borexino observes the reactor neutrinos that are produced worldwide. The neutrino flux at the Borexino site was estimated in [35]. Here, the monthly energy production of each reactor site and the neutrino oscillation probability is taken into account. In this work, the energy spectrum of Mueller et al. [195] is used. The so-called '5 MeV bump' that was observed by several experiments [196–198] is not considered and can be assumed to be negligible as it is above the endpoint of the ^{144}Pr spectrum. Reactor neutrinos reach energies up to ~ 10 MeV and only 32.3 % fall into the SOX energy range.
The expected event number is extracted from the fit to 56.2 ± 7.0. In the SOX energy window and for a measurement time of 1.5 y, the expected event number corresponds to 6.6 ± 0.8. The result agrees with the rescaled value from previous measurements of $4.9^{+2.0}_{-0.5}$ [35]. Note that the reactor neutrino flux depends on the power of the nuclear power plants. Hence, the expected event number for reactor neutrinos could slightly differ in the SOX measurement period.

6.2.3 Atmospheric Neutrinos

Atmospheric neutrinos are secondary particles that are produced in interactions of primary cosmic rays with nuclei of the atmosphere (see Section 1.3.1). Atmospheric neutrinos can be neutrinos as well as antineutrinos with electron and muon flavor, respectively, and reach energies up to ∼TeV. In addition to the CC IBD channel itself, atmospheric neutrinos can interact via NC interactions mainly on Carbon atoms in which high-energy neutrons can be produced[1]. The IBD signal can be mimicked when for example the neutrons scatter off protons and produce a prompt-like event and the neutron itself is captured and provides the delayed signal. For a detailed description see for example Reference [35], in which the atmospheric neutrino events in Borexino are estimated. In this work, the previously obtained result is used and corrected for the different exposure and detection efficiency using Equation 6.4. The resulting value is 0.4 ± 0.2 which is negligible in comparison to the expected SOX event number.

6.3 Cosmogenic Background

Even though Borexino is located at a rock overburden of 3800 m.w.e. and the cosmic muon flux is reduced by about six orders of magnitude in comparison to the ground level, the residual muon flux measured in Borexino is $\sim 0.05\,\mathrm{s}^{-1}$. The inefficiency of the muon veto

[1] The measurement of the dominating NC interactions of atmospheric neutrinos in Borexino can be found in Ref [63].

was determined to 0.0013 ± 0.0005 and the background induced by untagged muons crossing the inner detector is negligible [35] (see Section 6.6). While the muons themselves can be well identified, they can produce fast neutrons and radioisotopes with lifetimes up to several minutes in nuclear spallation processes on Carbon nuclei of the liquid scintillator. The most dangerous backgrounds arise from the isotopes ^{9}Li and ^{8}He that decay via $\beta + n$ decays and can mimic the IBD signature. With the respective lifetimes of 257.2 ms and 171.7 ms, the previous geo neutrino analyses suppressed this background by applying a veto of two seconds after each internal muon (corresponding to more than seven times the lifetime of ^{9}Li). However, with a muon rate of $\sim 0.05\,\mathrm{s}^{-1}$, the dead time becomes $\sim 9.5\,\%$ of the total measurement time[2]. To increase the statistics for the SOX event sample by the reduction of the dead time, a new veto strategy was developed by Stefan Weinz [199] (see also Chapter 4). It is based on the fact that 82.2 % of all ^{9}Li/^{8}He events are produced in coincidence with at least one cosmogenic neutron. By vetoing only two seconds after muons followed by at least one neutron, the dead time can be reduced to 0.176 %, as these muons make up only 1.4 % of all muons. After all residual muons, internal muons without a neutron and external muons, the data is vetoed for 2 ms to get rid of the fast cosmogenic neutron background. With the new muon strategy, the lifetime and the statistics of the SOX signal increase by 11.4 %. Due to the relaxed muon veto, a few cosmogenic backgrounds, namely ^{9}Li, ^{8}He, and ^{12}B, will end up in the SOX data sample, which are evaluated in the following.

6.3.1 ^{9}Li, ^{8}He

^{9}Li and ^{8}He decay both via beta decay and can populate excited states of the daughter nuclides ^{9}Be and ^{8}Li. These excited states can emit a neutron while relaxation. The coincidence of the beta particle and the neutron is indistinguishable from an IBD signal. The Q-value of ^{9}Li and ^{8}He is 13.6 MeV and 10.7 MeV, respectively [140].
The ^{9}Li/^{8}He events obtained with the new veto strategy can be estimated by looking at the IBD-like events that appear after muons without a neutron (μ_{int-n}). In the energy region up to 4900 p.e. (~ 10 MeV) for the prompt event[3] 31 events are found and are shown in Figure 6.2a as a function of the time after the muon. The number of ^{9}Li/^{8}He events are determined using a binned likelihood fit and the following fit function:

$$f(t) = \Big(N_{\mathrm{Li/He}} \cdot \exp(-\tau_{\mathrm{Li}} \cdot t) + N_{\mathrm{B}} \cdot \exp(-\tau_{\mathrm{B}} \cdot t) + N_{\mathrm{acc}}\Big) \cdot \exp(-r_{\mu_{\mathrm{int}}-n} \cdot t). \tag{6.2}$$

This function consists of three components: $N_{\mathrm{Li/He}}$ is the amplitude of ^{9}Li/^{8}He events that are exponentially decaying after the muon and where the decay time is fixed to the lifetime of ^{9}Li. An extra component for ^{8}He can be neglected, as the ratio of IBD-like events from ^{9}Li:^{8}He was measured in Borexino to 6:1 [35]. N_{B} is the amplitude of accidental coincidences of two cosmogenic ^{12}B events that are produced by the same muon. The exponential decay time is given by the lifetime of ^{12}B of 29.1 ms. The ^{12}B events are described in the next section and constrained by the fit in Figure 6.3. The third component N_{acc} is the amplitude of all IBD-like events accidentally falling in the time window after a muon. These can be either real IBD events or any background source. All three components are multiplied with the term $\exp(-r_{\mu_{\mathrm{int}}-n} \cdot t)$ that gives the probability that no other muon falls in between the muon and the IBD-like event and is connected to the $(\mu_{\mathrm{int}} - n)$-rate $r_{\mu_{\mathrm{int}}-n} = 0.054\,\mathrm{s}^{-1}$.
N_{acc} can be constrained through the time distribution of the IBD-like events in the time window Δt(μ-IBD) = 2 s - 100 s (see Figure 6.2b). In this time window, the contributions of

[2]The probability to detect no muon for the time t is given by $\exp(-r_\mu \cdot t)$, where r_μ is the muon rate.
[3]In order to constrain the event number in the simultaneous fit, the same energy window as in the simultaneous fit is used.

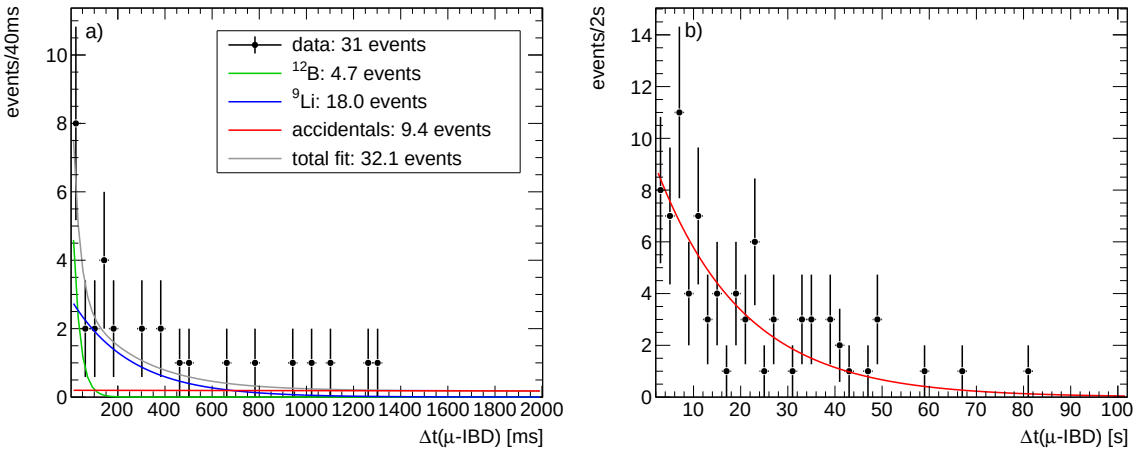


Figure 6.2: Determination of ^{9}Li/^{8}He Background IBD-like events from 4.16 y of Borexino data (black data points) as a function of the time between an internal muon without a cosmogenic neutron ($\mu_{\text{int}}-n$) and the prompt event. The IBD events are selected with the optimized selection cuts with an enlarged energy cut $E_p < 4900$ p.e.. The data can be fitted with the sum of ^{12}B (green), ^{9}Li/^{8}He (blue), and accidental (red) events. The latter are constrained for large time differences of $\Delta t(\mu\text{-IBD}) = 2\,\text{s}$ - 100 s in panel b. ^{12}B events are constrained in Fig 6.3.

^{9}Li,^{8}He, and ^{12}B are negligible, and the data can be solely described by $N_{\text{acc}} \cdot \exp(-r_{\mu_{\text{int}}-n} \cdot t)$. The fit returns a total expected event number for the accidental IBD-like events of 9.4 for the time interval 2 ms - 2 s. This value is consistent with the one expected due to the increase of the lifetime of 9.3 events.

Besides $N_{\text{Li/He}}$, all other parameters are fixed in the fit and the resulting ^{9}Li/^{8}He event number in the time interval 2 ms - 2 s is 18.0 ± 5.1. The total event numbers of data (31) and fit (32.1) agree well with each other.

The fraction of the ^{9}Li energy spectrum below 1300 p.e. is estimated to 15.9 % using the energy spectrum simulated with the Borexino MC. Therefore, the expected ^{9}Li/^{8}He background in the SOX energy window and measurement time corresponds to 1.0 ± 0.3 events. This result is compatible with a similar analysis from [199].

The contribution from ^{9}Li/^{8}He events that appear after the muon veto of two seconds is negligible and not considered.

6.3.2 ^{12}B

^{12}B is a cosmogenic isotope produced in the spallation processes induced by muons and decays via a beta decay with a lifetime of 29.1 ms and a Q-value of 13.4 MeV. As the production yield of ^{12}B is high in Borexino [140], it is possible that one muon creates several ^{12}B isotopes that generate a coincident signal. The coincidence of two ^{12}B events is the first time observed in Borexino in this work and is deduced from the fact that the exponential decay time after a muon and the prompt event as well as the time in between the prompt and delayed event correspond both to the lifetime of ^{12}B.

The background contribution of ^{12}B is determined similarly to the one of the ^{9}Li/^{8}He component. To increase the statistics of the ^{12}B events, IBD-like events with $\Delta t < 150$ ms (~ 5 times the lifetime of ^{12}B) are selected. Figure 6.3 shows the obtained data as a function of the time difference to the last preceding muon without a neutron. Equivalently to the determination of the ^{9}Li/^{8}He component, the data can be described with Equation 6.2. Again, N_{acc} is constrained through the data with $\Delta\text{t}(\mu\text{-IBD}) = 2\,\text{s}$ - 100 s to 11.6 events (see 6.3b).

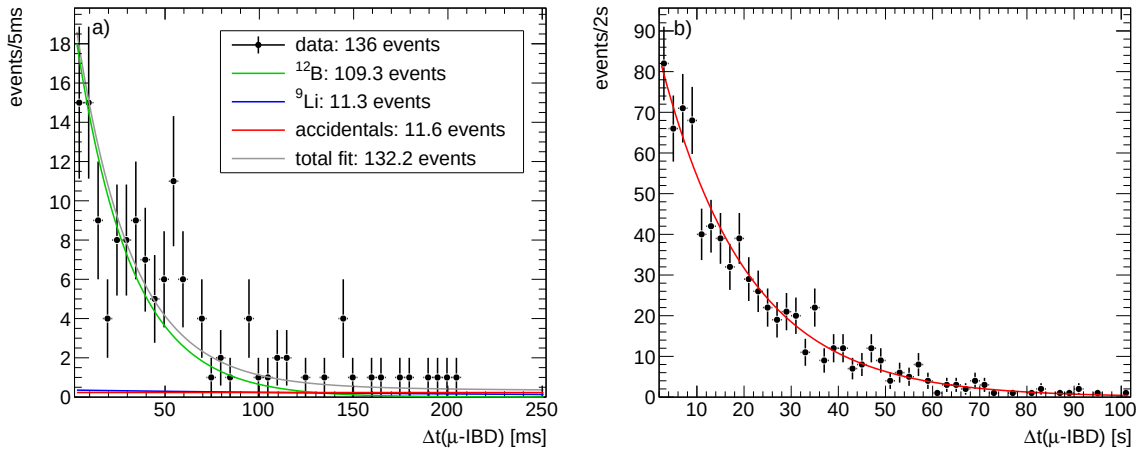


Figure 6.3: Determination of ^{12}B Background IBD-like events from 4.16 y of Borexino data (black data points) as a function of the time between an internal muon without a cosmogenic neutron ($\mu_{\text{int}} - n$) and the prompt event. Besides $E_p < 4900$ p.e. and Δt<150 ms, the IBD events are selected with the optimized selection cuts. The data can be fitted with the sum of ^{12}B (green), ^{9}Li/^{8}He (blue), and accidental (red) events. The latter are constrained for large time differences of $\Delta t(\mu\text{-IBD}) = 2$ s - 100 s in panel b. ^{9}Li/^{8}He events are constrained in Fig 6.2.

The ^{9}Li/^{8}He event number is constrained to the value obtained from the data with $\Delta t <$ 1280 µs to 11.3 events (see Figure 6.2). Note that the energy range is different in both fits and that the event number is corrected accordingly.

Leaving only N_{B} free in the fit, results into a total event number of 109.3 for ^{12}B. The total event number of data (136) and fit (132.2) are in agreement. The expected ^{12}B events in the SOX time window (Δt<1280 µs), reduces to only 4.3 %, i.e. 4.7 events. This event number is used in the fit to determine the ^{9}Li/^{8}He event number which was discussed in the previous section.

Out of these events only 7 % appear in the SOX energy window. This number is computed from the ^{12}B spectrum generated with the Borexino MC in [140]. Hence, for the SOX measurement time and energy window the ^{12}B background reduces to 0.12 events and is fully negligible.

In addition, one could fully get rid of this background by enlarging the dead time window from 2 ms to $\sim$ 150 ms after each muon without a neutron. However, such a veto would decrease the lifetime and the statistics of the SOX signal. Given the small expected event number of ^{12}B, such a modification is not necessary.

Note that the exact number of cosmogenic background events could slightly vary as they depend on the muon flux which modulates seasonally. Such modulations are mainly correlated to the temperature of the atmosphere and were measured in Borexino in References [63, 200].

6.4 Accidental Background

Uncorrelated single events can produce accidental coincidences and mimic an IBD signal. To validate the fit model for the accidental coincidences, data is selected for the muon veto used in the previous antineutrino analyses: i.e. two seconds are vetoed after each internal muon, independent on the fact whether there is a neutron. Further, to increase the statistics, a ΔR of 5 m is chosen. The rate of accidental coincidences as a function of Δt can be described by

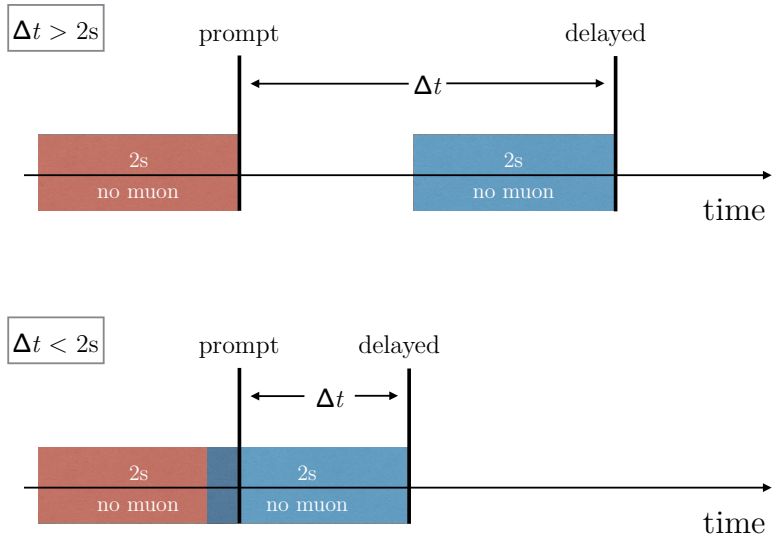


Figure 6.4: Scheme of the Time Dependency of the Accidental Coincidence Rate due to the Muon Veto An accidental coincidence is accepted, when no muon appears two seconds before the prompt and the delayed event (indicated in red and blue). As the muon veto overlaps for events with $\Delta t < 2\,\mathrm{s}$ (bottom), the time of the net muon veto of the delayed event is exactly Δt and hence smaller than 2 s. As a consequence, the accidental coincidence rate is enhanced for $\Delta t < 2\,\mathrm{s}$.

the following function:

$$r_{\mathrm{acc}}(\Delta t) = \begin{cases} r_{\mathrm{acc}}(\Delta t = 0\,\mathrm{s}) \cdot \exp(-r_\mu \cdot \Delta t) & \text{if } \Delta t \leq 2\,\mathrm{s}, \\ r_{\mathrm{acc}}(\Delta t = 0\,\mathrm{s}) \cdot \exp(-r_\mu \cdot 2\,\mathrm{s}) & \text{if } \Delta t > 2\,\mathrm{s} \end{cases} \tag{6.3}$$

where r_μ is the muon rate.
The accidental coincidence rate ($r_{\mathrm{acc}}(\Delta t = 0\,\mathrm{s})$) is altered by the probability that there is no muon before the delayed event which is a function of time. Figure 6.4 sketches this feature. A coincidence is only observed when there is no muon two seconds before the prompt and the delayed event. When the time difference between the prompt and the delayed event is smaller than two seconds, the muon vetos overlap and the net time without a muon is only Δt and hence smaller than two seconds. As the time is shorter than two seconds the probability is larger that no muon falls accidentally before the delayed event and corresponds to $\exp(-r_\mu \cdot \Delta t)$. After two seconds, the probability to observe no muon stays constant and is given by $\exp(-r_\mu \cdot 2\,\mathrm{s})$.
Figure 6.5a shows the fit of the collected data for IBD-like events up to Δt<4 s. The obtained muon rate of $0.061\,\mathrm{s}^{-1}$ in the fit agrees well with the actual internal muon rate of $0.055\,\mathrm{s}^{-1}$. As this model for the accidental coincidences has been developed during this work, the result was in addition cross-checked and validated with a toy MC.
For the optimized cosmogenic veto, only two seconds after each internal muon with at least one neutron are vetoed. Hence, r_μ reduces to $7.6 \cdot 10^4\mathrm{s}^{-1}$ and due to its small value it is fixed in the fit. Figure 6.5b shows the determination of the accidental coincidences for the selection cuts used in the simultaneous fit and for $\Delta t < 20\,\mathrm{s}$. The expected event number due

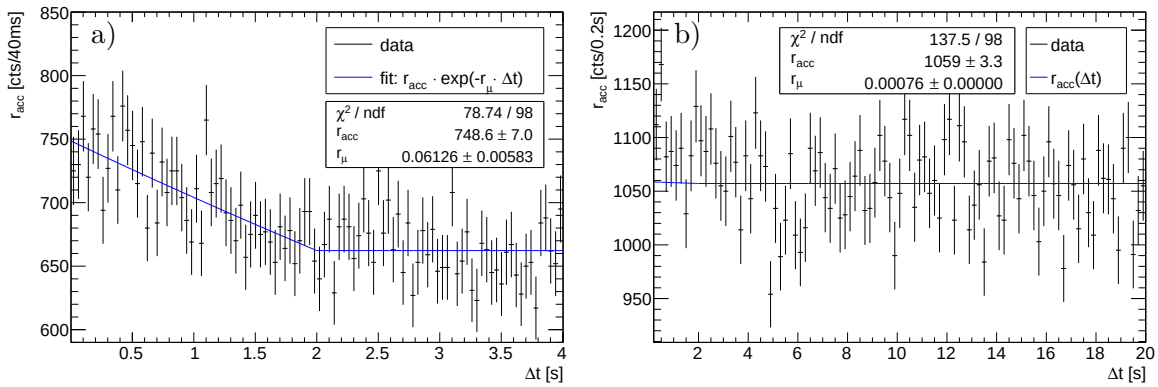


Figure 6.5: Determination of Accidental Coincidence Background a) Validation of the fit model for accidental coincidences for a muon veto of two seconds after each internal muon and ΔR=5m. The accidental coincidence rate changes in time due to the applied dead time after each muon. The probability of having no muon in between two events decreases in time with $\exp(-r_\mu \cdot \Delta t)$ and stays constant above 2 s. The fitted muon rate r_μ returns well the actually measured muon rate in Borexino. b) accidental coincidences selected for the optimized selection cuts with $E_p < 4900$ p.e.. Here, r_μ is fixed to the rate of muons which produce at least one neutron.

to accidental coincidences in the Δt window for SOX can then be determined to 6.73 ± 0.02 events. In addition, with the SOX energy window (91.5 %) and measurement time one expects 2.2 ± 0.01 events. The small error is given by the large event number in the off-time window. The fit in Figure 6.5b agrees well with the data, however, one could eventually observe a small excess in the data in comparison to the fit function for Δt<2 s (see Figure 6.5b). These events are most likely accidental coincidences of two uncorrelated cosmogenic events like ^{12}B or ^{8}B, as the excess vanishes when a muon veto of two seconds after each internal muon is applied. As the data after a muon within two seconds could be well described (see previous section), any extra contribution due to cosmogenic background is expected to be negligible and not discussed further.
Note that the spectral distributions of the accidental coincidences follow the spectral distributions of single Borexino events. This means that the energy spectrum peaks at small energies and that the accidental coincidences are mainly reconstructed close to the vessel border (especially at the poles). Further, the probability to observe two coincident events increases with the spatial distance ΔR (see Figure 6.1). Hence, more strict selection cuts on for example the fiducial volume, the minimum value of the energy, or ΔR could reduce this background component even further.

6.5 Fast Coincidences from the ^{238}U and ^{232}Th Decay Chains

The decay chains of ^{238}U and ^{232}Th have at least one radioisotope with a short lifetime that can produce an observable fast coincidence. These radioisotopes count to the intrinsic backgrounds of Borexino and are mainly located at the vessel surface and the detector structure material, e.g. the construction material at the poles.
The most critical backgrounds are ^{212}Bi-^{212}Po and ^{214}Bi-^{214}Po coincidences as the lifetimes of the Polonium isotopes are smaller or similar to the neutron capture time and appear hence in the SOX Δt-window. Table 6.1 lists the decay properties and the life times of the Polonium isotopes. In both cases the Bismuth isotope decays via a beta decay and the Polonium isotope via an alpha decay that allows to discriminate these events by means of pulse shape

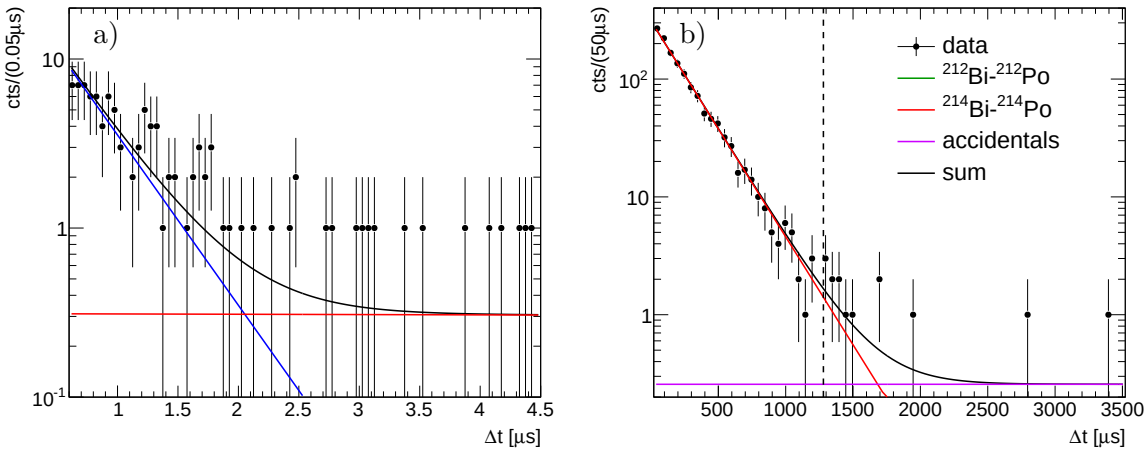


Figure 6.6: Determination of ^{212}Bi-^{212}Po and ^{214}Bi-^{214}Po Events IBD-like events (black data points) are selected for the optimized selection cuts, but without the *mlp*-cut on the delayed event. The two panels show different time scales for Δt. The time distributions can be well described by the sum (black) of the two Polonium lifetimes (blue and red) and accidental coincidences (violet). The dashed line indicates the maximum Δt cut used for the SOX analysis.

discrimination from the IBD signals. The alpha rejection efficiency was defined in Chapter 5 to $> 99.9\,\%$ for the optimized multi-layer perceptron (*mlp*) cut of $mlp > 0.085$.
Hence, an IBD signal can be mimicked by such fast coincidences if the alpha particle is not recognized given by the inefficiency of the pulse shape parameter. In addition, ^{214}Po decays with a probability of 1.04×10^{-4} into an excited state of ^{210}Pb that produces a gamma line of 799.7 MeV while relaxation. As the gamma and the alpha event happen almost at the same time, the pulse shapes overlap and the pulse shape discrimination loses power. The rejection efficiency for ^{214}Po is hence smaller and assumes that the pulse shape discrimination fails for that decay branch: $\epsilon(^{214}\text{Po}) = 99.89\,\%$.
These fast coincidences are searched for with the optimized selection cuts without applying a *mlp* cut on the delayed event. Figure 6.6 shows the Δt distribution of these events which follows the decay times of the two Polonium isotopes. Hence, their contributions can be determined from a fit assuming the two Polonium components plus accidental coincidences. As the fraction of real IBD events is small in the data sample and to obtain conservative results on the Polonium backgrounds, the IBD events are not included in the fit. Table 6.1

Table 6.1: Expected Event Numbers of IBD-like Signals due to Fast Coincidences with a Polonium Event The listed Polonium isotopes decay via an alpha decay within short lifetimes (second column) and can produce a fast coincidence signal as their energy (third column) is quenched and appear in the energy region of the delayed SOX event. IBD-like events are determined from a fit of the data of 4.16 y selected for the optimized selection cuts without the *mlp* cut on the delayed event (fourth column). The expected event number due to the inefficiency of the *mlp* cut for the analyzed period/ SOX measurement time are given in the last two columns.

isotope	lifetime	Q-value	events		
			4.16 y	4.16 y *mlp*-ϵ	1.5 y *mlp*-ϵ
^{212}Po [201]	425.1 ns	8.9 MeV	9.2	9.2×10^{-2}	3.3×10^{-3}
^{214}Po [202]	236.0 µs	7.8 MeV	1428.7	1.6	0.6

quotes the results for the SOX Δt-window for the two Polonium components. In order to obtain an upper limit on the Polonium component on the SOX data sample, the measured Polonium event numbers are multiplied with the inefficiency of the *mlp* cut and scaled to the SOX measurement time. Even though ^{214}Po dominates by two orders of magnitude in comparison to ^{212}Po, it contributes with only < 0.6 events to the SOX background and is hence negligible.
In Appendix A.3 Figure A.8, the same analysis is performed for the data sample when no fiducial volume cut is applied. Fast coincidences due to ^{215}Po- and ^{216}Po- decays need then to be taken into account as well.

6.6 Additional Detector-Related Background Sources

This section summarizes additional detector-related background sources that were studied in detail in the latest geo neutrino analysis [35]. In the following, these are described and their expected event numbers are estimated based on the previous results N_{geo}

$$N_{\text{SOX}} \approx \frac{\tau_{\text{SOX}}}{\tau_{\text{geo}}} \cdot \frac{\epsilon_{\text{SOX}}}{\epsilon_{\text{geo}}} \cdot N_{\text{geo}} \approx 0.164 \cdot N_{\text{geo}}, \tag{6.4}$$

where $\tau_{\text{SOX/geo}}$ and $\epsilon_{\text{SOX/geo}}$ is the exposure and detection efficiency used in the SOX/geo neutrino analysis, respectively. This estimation assumes that the backgrounds are uniformly distributed in the detector which is not true for external background components that increase towards the vessel border, as volumes at outer radii are less shielded. Note that also some of the selection cuts differ in the geo neutrino analysis, such as the larger energy window for the prompt event in the geo neutrino analysis. However, these small effects are negligible as the sum of these backgrounds is below one event for the total SOX measurement period.
In general, any process that produces neutrons can be a possible background candidate for an IBD signal. The neutrons will be thermalized and captured and produce exactly the same signal as the delayed IBD event. The prompt event depends on the process. For example, if fast neutrons are produced, they can scatter off protons and the deposited energy acts as the prompt signal.

Untagged Muons The inefficiency of the internal muon veto used for the SOX analysis is $(0.0013 \pm 0.0005)\,\%$ [35]. Hence, undetected muons can lead to non-vetoed cosmogenic background. On the one hand, the prompt event could be the muon itself. However, to observe a signal in the SOX energy window, the muon passed most likely only through the buffer medium. The delayed event could be one spallation neutron, where it has to be exactly one neutron as the multiplicity cut would veto this event otherwise. On the other hand, the IBD signal could be mimicked by coincidences of cosmogenic isotopes and/or neutrons.
The background due to untagged muons was estimated to (0.023 ± 0.007) events in [35]. Hence, the expected background for SOX corresponds to (0.004 ± 0.001) events and is negligible.

Fast Cosmogenic Neutrons The energy of cosmogenic neutrons produced through spallation processes of muons can reach several GeV. Such high energies allow the neutrons to penetrate large distances in the rock or the detector material. If a neutron reaches the inner detector volume, it can scatter off protons and produce a prompt signal. The neutron itself can give the delayed event. Untagged muons passing the water tank[4] or the rock can generate

[4]The background due to untagged muons crossing the inner detector is considered in the category of untagged muons.

such cosmogenic neutrons.
The rescaled upper limits at a 95 % CL from untagged muons in the water tank and in the rock result to 2.1×10^{-4} and 0.23 events, respectively.

(α,n) Events Alpha particles generated in the decay chains of ^{238}U, ^{235}U, ^{232}Th that contribute to the intrinsic background of Borexino can produce neutrons through (α,n) reactions. The main alpha source is ^{210}Po and the reaction occurs mostly on ^{13}C. The produced neutrons can achieve energies up to 7.3 MeV and scatter off protons and produce a prompt signal. After thermalization and capturing, the neutron is itself the delayed event. The prompt signal could also be mimicked by gamma events that can be emitted in the relaxation process of the excited states of the products of the reaction. However, the energy of the gamma particle is above the SOX energy range.
Detectable (α,n) reactions can arise from alpha decays that take place in the inner vessel or in the inner buffer. As the rate of alpha decays within the inner vessel can be accurately determined via pulse shape discrimination, the expected event number of (α,n) reactions in the inner vessel can be computed. The rescaled event number is 0.13±0.02. The determination of the ^{210}Po rate in the inner buffer is more uncertain and the upper limit of 0.43 events can be computed (95 % CL). Note that this limit is very conservative and that this background component is in addition expected to decrease with the ^{210}Po lifetime of ∼200 d.

(γ,n) Background Gamma particles with energies higher than 3 MeV can produce neutrons in (γ,n) reactions with the nuclei of the detector material. These high-energy gammas are in turn generated most likely in neutron capture reactions. The prompt signal can be mimicked by Compton scattering of the gamma and the delayed by the neutron itself.
The upper limit of this background can be estimated to 5.6×10^{-2} at the 95 % CL.

Spontaneous Fission Fast neutrons can be produced by spontaneous fission of the detector material and mimic an IBD signal. The ^{238}U contamination in the PMTs is the most important contribution to this background.
The upper limit of this background can be estimated to 9.3×10^{-3} at the 95 % CL.

In summary, the above-described background sources sum up to < 0.9 events and are therefore negligible for the SOX analysis.

6.7 Source-Induced Background

The antineutrino source for the SOX experiment has to be an extremely pure ^{144}Ce source. This is important, as a background component located at the same position as the source could generate a signal that is decreasing with the distance to the source. This could mimic a sterile neutrino signature especially with a long oscillation length, e.g. a small sterile neutrino mass below 1 eV^2. Moreover, impurities prevent from an accurate activity measurement and decrease the sensitivity of the experiment. Within the source specifications, there are two relevant types of background sources: gammas and neutrons (see Chapter 2).
The most serious gamma line is the 2.185 MeV line that is produced by the source itself with a probability of 0.7 %, when ^{144}Pr decays into an excited state of ^{144}Nd. Due to the very similar energy, such a gamma could in particular mimic a delayed event. In addition, the gamma can also produce a signal in the prompt energy window. However, these gammas can be easily shielded and actually define the thickness of the shielding: 19 cm of tungsten achieve

a total attenuation of 3.1×10^{-7} [133]. With the full Borexino MC based on Geant4 [184], a gamma flux per initial gamma at the source location with an energy deposit larger than 1 MeV in the inner vessel is determined to 1.4×10^{-18} [203]. Hence, at the beginning of data taking with a 100 kCi source activity, one expects 3.1 events/day within the inner vessel. The rate of an accidental coincidence of two such events in a 1280 µs time window is then 1.4×10^{-7}events/day which is a conservative upper limit, as no spatial correlation is required. For the total measurement time of 1.5 y one expects 7.8×10^{-5} events which is fully negligible.

In addition, a gamma-induced background signal can create an accidental coincidence with a single Borexino event. The Borexino rate above 1 MeV is $0.03\,\mathrm{s}^{-1}$ in the whole inner vessel which is more than ~ 800 times larger than the expected rate from gamma-induced background signals [35]. Assuming the same spatial and spectral distribution within the detector, the accidental coincidence rate would increase by 0.2 %. Using the value obtained in Section 6.4 of 2.2 events, one would expect 4.4×10^{-3} events for the full measurement time. As the intensity of the gamma line reduces with the source activity, the quoted event numbers are an upper limit.

Fast neutrons mainly produced by spontaneous fission of contaminants in the source are another background source. The most critical isotope is ^{244}Cm and the source is required to contain only 1×10^{-5} Bq of ^{244}Cm per one Bq of ^{144}Ce. The expected neutron flux of the source corresponds then to 1.4×10^{5} per second for a 100 kCi source [133]. In comparison to the gamma rays, about 40 % of the neutrons can escape the tungsten shielding. However, the neutrons cannot reach the inner vessel, but are captured before, e.g. in the steel plates or the water tank and can produce gammas up to 9 MeV. In turn, these gamma rays can reach the inner detector volume and contribute to the background. The probability that such a gamma deposits an energy in the prompt energy range is 2.4×10^{-7} per initial neutron. This number is determined from a TRIPOLI-4 simulation and a generic detector setup that is similar to the Borexino detector [133]. Preliminary results from the Borexino MC are in agreement to this result [203]. The rate of single events in the prompt energy window is then given by $\sim 0.03\,\mathrm{s}^{-1}$ which is similar to the single event rate of Borexino events above 1 MeV.

This event number is high enough that it contributes to the background for the SOX experiment. Again, the expected accidental coincidence rate is estimated by assuming the same spatial and spectral distribution within the Borexino detector. Given an increase of the single event rate by a factor of two, will lead to an increase of the accidental coincidence rate by a factor of four. Accidental coincidences due to source-induced neutron background (i.e. coincidence between two source-induced or one source-induced and one Borexino event) can be estimated to ~ 6.6 events for 1.5 years of data taking, assuming the result for the accidental coincidences from Section 6.4. As the half life of ^{244}Cm is 18.1 y, this background can be assumed to be constant over the measurement time of SOX. It has to be noted that this number gives an impression of the order of magnitude. However, the exact number depends on the selection cuts, the spatial and spectral distributions, and on the source activity.

Independent from the simulations and as source-induced background contributes only to the accidental coincidences, the increase of the accidental coincidence rate can be determined very precisely (< 0.03 events) directly from the data by looking at the off-time window for $\Delta t = 2\,\mathrm{s}$-$20\,\mathrm{s}$ (see Section 6.4). Vice versa, the source-induced increase of the accidental coincidence rate allows to draw conclusions on the contamination levels of the source and could also decrease the uncertainty on the activity measurement.

6.8 Summary of Background

The results for the expected background components using the optimized selection cuts are summarized in Table 6.2. The total expected background for the SOX analysis could be determined to $28.2^{+2.7}_{-4.2}$ and can be divided into detector- and source-related backgrounds. The detector-related background can be determined from existing Borexino data of 4.16 y to $21.6^{+2.7}_{-4.2}$ events. The uncertainty in the current analysis is only statistical and could be lowered by the analysis of a data sample with a longer measurement period. Source-related background is currently only estimated from MC simulations to 6.6 ± 0.03 events. However, source-related background is of accidental nature and can be precisely determined during data taking via an off-time window. In summary, the expected background event number is by a factor of ~ 200 smaller than the expected SOX signal event number for a source activity of 100 kCi and can be determined with a precision better than 10 %. Hence, the background is negligible and SOX can be called a background-free experiment.

Table 6.2: Expected Backgrounds for the SOX Experiment The detector-related background determined from 4.16 y of Borexino data is given in the first row. The data can be described by a simulations fit, where the individual components are grouped together and the sum of the fit is highlighted in gray. Additional negligible detector-related components and source-induced backgrounds are listed in the last rows. The second and third column give the measured events numbers in the energy range used for the simultaneous fit and for the SOX analysis, respectively. The last column quotes the expected background rescaled for the SOX measurement time. It has to be noted that the source-induced background is only an estimation, but can be precisely determined during data taking. The quoted numbers are valid for the optimized selection cuts from Table 5.1.

background	events		
	4.16 y (200-4900) p.e.	4.16 y (200-1300) p.e.	1.5 y (200-1300) p.e.
data	113 ± 10.6	52 ± 7.2	18.8 ± 2.6
background components included in simultaneous fit			
geo neutrinos	26.7 ± 6.5	26.6 ± 6.5	9.6 ± 2.3
reactor neutrinos	56.2 ± 7.0	18.6 ± 2.3	6.6 ± 0.8
^{9}Li/^{8}He	18.0 ± 5.1	2.9 ± 0.8	1.0 ± 0.3
^{12}B	4.7 ± 0.5	0.33 ± 0.04	0.12 ± 0.01
accidental coincidences	6.73 ± 0.02	6.16 ± 0.02	2.2 ± 0.01
total fit	112.3 ± 10.6	54.6 ± 6.9	19.7 ± 2.7
additional background components			
^{214}Bi-^{214}Po-coincidences			<0.6
atmospheric neutrinos			0.4 ± 0.2
residual background			<0.9
*source-induced background (*100 kCi*)*			
gammas			$\sim 4.4 \cdot 10^{-3}$
neutrons			$\sim 6.6 \pm 0.03$
total expectation			$\mathbf{28.2^{+2.7}_{-4.2}}$

7 Effect of the Systematic Uncertainties on the Sensitivity

In general, having any systematic uncertainty in an experiment not understood can lead to a misinterpreted signal. Hence, this chapter discusses the systematic uncertainties of the source (Section 7.1) and the detector (Section 7.2) and their effects on the sensitivity. The chapter ends with a summary including the sensitivity of the SOX experiment taking into account all systematic effects (Section 7.3). Note that this chapter relies on the asymptotic probability distributions of the test statistic, i.e. using Wilks' theorem and the Asimov data set (compare Chapter 3). However, it is assumed that the relative effects on the sensitivity are independent on the used method. The comparison of the absolute sensitivity can be found in Chapter 11.

7.1 Source-Induced Uncertainties

In order to draw conclusions on the sterile neutrino hypothesis, the properties of the source have to be precisely known. The two main sources of uncertainties of the neutrino source are the activity and the spectral shape of the neutrino spectrum (see Chapter 2 for the characterization of the source). An uncertainty in the activity transfers directly to an uncertainty in the expected event number that could mimic a sterile neutrino signature. The activity of the source at time t_0 is determined via

$$A(t_0) = \frac{h(t_0)}{\langle E \rangle}, \tag{7.1}$$

where $h(t_0)$ is the heat produced by the source at time t_0 and $\langle E \rangle$ is the mean deposited energy per ^{144}Ce-^{144}Pr decay in the tungstate shielding. The mean energy is given by the sum of the mean energies of the emitted electron and gamma lines per decay:

$$\langle E \rangle = \langle E_{\mathrm{Ce}}(e^-) \rangle + \langle E_{\mathrm{Ce}}(\gamma) \rangle + \langle E_{\mathrm{Pr}}(e^-) \rangle + \langle E_{\mathrm{Pr}}(\gamma) \rangle. \tag{7.2}$$

$\langle E_{\mathrm{Pr}}(e^-) \rangle$ and $\langle E_{\mathrm{Ce}}(e^-) \rangle$ depend on the spectral shape of the ^{144}Ce-^{144}Pr neutrino spectrum. The uncertainty in the spectral shape of the neutrino spectrum can hence, on the one hand, distort the neutrino spectrum and mimic a shape oscillation signal, and, on the other hand, affect the prediction of the expected event number and mimic a rate oscillation signal as well. In the following, the two sources of uncertainties, namely the uncertainty of the calorimetric measurement (Section 7.1.1) and the uncertainty on the spectral shape measurement (Section 7.1.2), are discussed independently. Moreover, effects of the size, the position, and the backgrounds of the source are investigated in Sections 7.1.3 and 7.1.4.

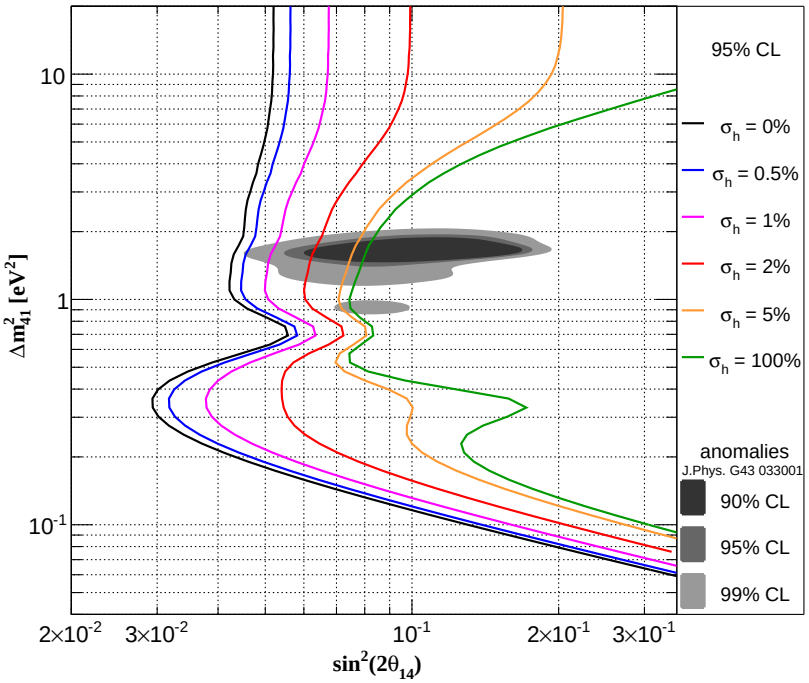


Figure 7.1: Impact of the Uncertainty of the Calorimetric Measurement on the Sensitivity The 95 % CL exclusion limits are shown for varying uncertainties σ_h, where an activity of 125 kCi and the optimized selection cuts (Table 5.1) are assumed. The achieved accuracy of $\sigma_h = 0.2\,\%$ allows to probe the full parameter space favored by the anomalous neutrino oscillation data (gray).

7.1.1 Calorimetric Power Measurement

The power (heat h) of the source is measured by two independent calorimetric setups (see Section 2.2.2) with an accuracy of 0.2 % [170].
The sensitivity as a function of the uncertainty σ_h is studied by treating the expected event number as a nuisance parameter and varying the width of the Gaussian pull term according to the specific uncertainty $\sigma_h \cdot N_S$. The likelihood function is given by Equation 4.6.
Figure 7.1 shows the 95 % CL contours for different uncertainty levels in a rate+shape analysis. With increasing uncertainty the sensitivity reduces in terms of $\sin^2(2\theta)$. Given an uncertainty of 100 %, the contour corresponds to the one of a shape only analysis, as defined in Chapter 4.3. The sensitivity of the shape only analysis is reduced to $\sin^2(2\theta)$=0.074 at Δm^2=1 eV^2 in comparison to the one with a perfectly-known heat ($\sin^2(2\theta)$=0.043). This reduction corresponds to a factor of ∼0.6. In contrast, at $\Delta m^2 \approx 0.35\,\text{eV}^2$ or $\Delta m^2 \gtrsim 10\,\text{eV}^2$, the sensitivity decreases by a factor of 0.06 or is even completely lost. At these Δm^2-values, the sterile neutrino signature is mostly given by a flat rate reduction (see Figure 2.14) that can be fully explained with a smaller but unknown activity.
The achieved accuracy of 0.2 % is better than the indicated 0.5 % in Figure 7.1 and allows therefore to probe the full parameter space preferred by the anomalous neutrino oscillation data.

7.1.2 Spectral Shape of the ^{144}Pr Spectrum

The main branch of the ^{144}Pr spectrum follows a non-unique first-forbidden decay that cannot be fully described by theory. In addition, the latest measurements were performed several

decades ago and disagree up to 10 % [167]. Therefore, new experiments have been planned within the collaboration (see Section 2.2.2) to determine the electron and thus the neutrino spectrum of ^{144}Pr.
The electron spectrum N_e, which is the complementary of the neutrino spectrum, can be described for a pure pseudoscalar transition[1] as a function of the shape factor b

$$N_e(W, b) = N_0(W) \cdot (1 + \frac{b \cdot m_e c^2}{W}). \tag{7.3}$$

Here, W denotes the total energy (i.e. the sum of the kinetic energy and the rest mass of the electron $m_e c^2$) and $N_0(W)$ represents the allowed spectrum that can be obtained from Fermi's theory (see for example Reference [165, 204]).
As already discussed, an uncertainty in the neutrino spectrum can affect the sensitivity in two ways. First, the neutrino spectrum itself deforms and can mimic an oscillation signature. Moreover, the fraction of the spectrum above the inverse beta decay (IBD) threshold can change and affect the estimation of the expected event number. Second, the mean emitted energy per decay changes which leads to a wrong activity estimation.
To study the effect of the uncertainty on the shape factor σ_b, b is treated as a nuisance parameter. The probability distributions and the expected event number depend then on b

$$pdf_{ij}^{\mathrm{S}}(\sin^2(2\theta), \Delta m^2) \rightarrow pdf_{ij}^{\mathrm{S}}(\sin^2(2\theta), \Delta m^2, b) \tag{7.4}$$

$$N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, N_{\mathrm{S}}) \rightarrow N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, N_{\mathrm{S}}, b) \tag{7.5}$$

and the likelihood function is extended with a Gaussian pull term for b with the measured value b^{obs}

$$\begin{aligned} &\mathcal{L}(\sin^2(2\theta), \Delta m^2, N_S) \rightarrow \mathcal{L}(\sin^2(2\theta), \Delta m^2, N_S, b) \\ &= \prod_{ij} \mathcal{P}(N_{ij}^{\mathrm{obs}} | N_{ij}^{\mathrm{exp}}(\sin^2(2\theta), \Delta m^2, N_S, b)) \cdot \mathcal{N}(N_S^{\mathrm{obs}} | N_S, \sigma_{N_S}^2) \cdot \mathcal{N}(b^{\mathrm{obs}} | b, \sigma_b^2). \end{aligned} \tag{7.6}$$

Technically, one can compute the expected probability distribution functions and event numbers with a single high statistics Monte Carlo (MC) data sample for one specific shape factor b^{obs}. The normalized neutrino spectrum $n_\nu(E, b)$ (and thus $pdf_{ij}^{\mathrm{S}}(\sin^2(2\theta), \Delta \mathrm{m}^2, b)$) with shape factor b can be obtained through the correction factor[2]

$$c(E, b) = 1 + \frac{b \cdot m_e c^2}{(Q - E) + m_e c^2}, \tag{7.7}$$

where Q is the endpoint of the ^{144}Pr spectrum. This results to

$$n_\nu(E, b) = n_\nu(E, b^{\mathrm{obs}}) \cdot \frac{c(E, b)}{c(E, b^{\mathrm{obs}})} \cdot \Big(\int_0^Q n_\nu(E, b^{\mathrm{obs}}) \cdot \frac{c(E, b)}{c(E, b^{\mathrm{obs}})} dE \Big)^{-1}. \tag{7.8}$$

The event number is given by

$$N_S^{\mathrm{exp}}(\sin^2(2\theta), \Delta \mathrm{m}^2, N_{\mathrm{S}}, b) = N_S(b^{\mathrm{obs}}) \cdot \frac{\int_{E_{thr}}^Q \int_L n_\nu(E, b) \sigma_{\mathrm{IBD}}(E) P_{ee}(E, L) dE dL}{\int_{E_{thr}}^Q \int_L n_\nu(E, b^{\mathrm{obs}}) \sigma_{\mathrm{IBD}}(E) dE dL} \cdot \frac{\langle E(b^{\mathrm{obs}}) \rangle}{\langle E(b) \rangle}, \tag{7.9}$$

where $N_S(b^{\mathrm{obs}})$ is the expected event number for the nominal shape factor b^{obs} and the no-oscillation hypothesis. The second term corrects the expected event number due to the

[1] The main decay branch of ^{144}Pr $\rightarrow$ ^{144}Nd is a $0^- \rightarrow 0^+$ transition.
[2] In the following $E = E_{\mathrm{true}}$ and $L = L_{\mathrm{true}}$.

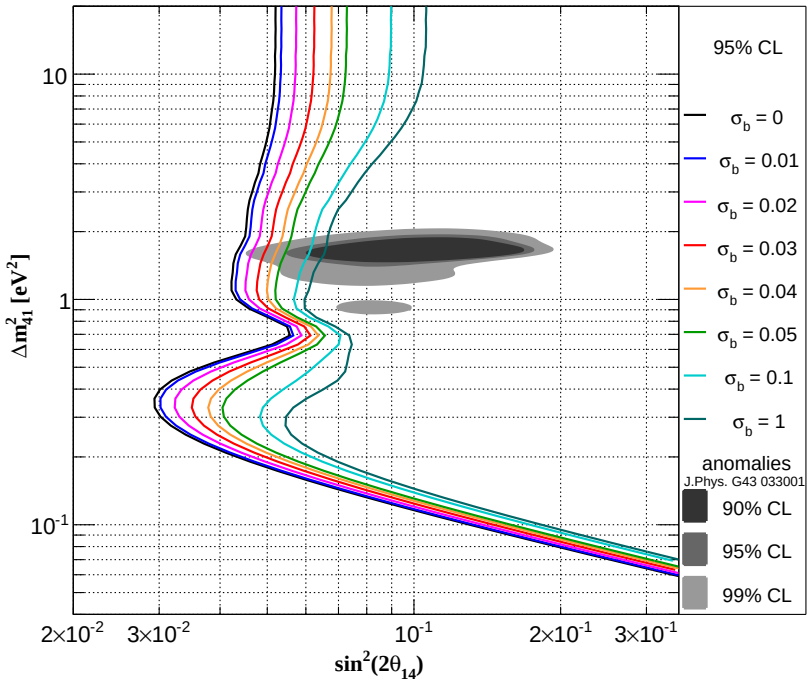


Figure 7.2: Impact of the Uncertainty of the Spectral Shape on the Sensitivity The 95 % CL levels are shown as a function of the uncertainty on the shape factor b. Note that the uncertainty is given in an absolute scale. The goal of the spectral shape experiments is to achieve an absolute error of 0.03 that could still cover the full preferred parameter space by global data (gray). The results are shown for the optimized selection cuts and an activity of 125 kCi.

spectral deformation, where $\sigma_{\text{IBD}}(E)$ is the cross section (Equation 2.7) and $P_{ee}(E, L)$ the survival probability (Equation 1.10). The third term $\frac{\langle E(b^{\text{obs}})\rangle}{\langle E(b)\rangle}$ accounts for the change in the mean emitted energy by the source, where $\langle E(b)\rangle$ is defined by Equation 7.2.[3]
The impact of the shape factor on the sensitivity can be seen in Figure 7.2, where N_S is fixed here in the likelihood function. The sensitivity obtained for an unconstrained shape factor reduces up to a factor of ~ 0.5 in comparison to the sensitivity obtained with a perfect knowledge on b. The main reason for the loss of the sensitivity is the changed expected event number that can mimic a flat disappearance signal. However, an oscillation signature can be slightly mimicked for Δm^2-values below 1 eV^2 which can be seen in Figure 7.3 that compares the sensitivities for a fixed and unconstrained shape factor in a shape only analysis. Hence, an uncertainty in the shape factor can only mimic large oscillation lengths in the order of the detector dimensions and the sensitivities are consistent above 1 eV^2. The desired smoking gun signature of the experiment can thus not be mimicked by a shape uncertainty.
The experimental setups have been designed to determine the spectral shape with an accuracy better than 0.03 (see Section 2.2.2). With that uncertainty, SOX can still probe almost the full 99 % CL parameter space preferred by the anomalous neutrino oscillation data.

7.1.3 Source Dimension and Source Position

Having a compact source is important to be able to resolve small oscillation lengths and be sensitive to sterile neutrino masses above 1 eV^2. In SOX the source consists of CeO_2

[3] In praxis, the correction factor enters into the weights of the PDFBuilder and the integral is computed via the weighted sum of the high statistics MC sample (see Section 4.3).

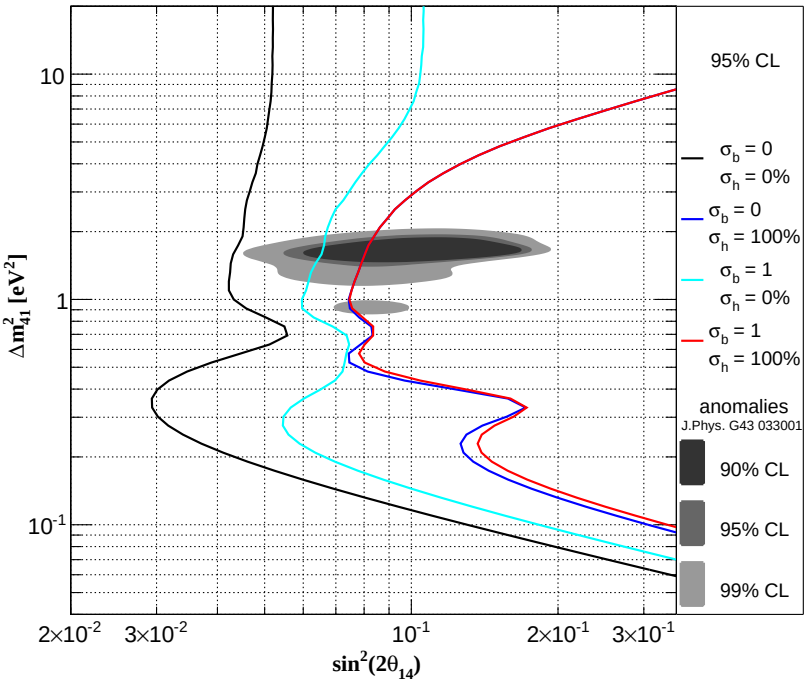


Figure 7.3: Interplay of the Uncertainties of the Heat and the Spectral Shape The sensitivities at the 95 % CL for fixed heat and shape (black), unconstrained shape and fixed heat (light blue), unconstrained heat and fixed shape (dark blue), and unconstrained heat and shape (red) uncertainties are shown. From the comparison of the curves one can observe that an uncertainty in the spectral shape can mimic shape oscillations only for Δm^2-values $< 1\,\mathrm{eV}^2$. The residual reduction of the sensitivity comes from rate oscillations. The results are shown for the optimized selection cuts and an activity of 125 kCi.

powder that is compressed into a cylindrical capsule with a diameter and height of 15 cm, respectively (see Section 2.2). The impact of the source dimensions on the sensitivity is studied by setting the source size to a sphere with a radius of 7 cm in the MC simulations. The sensitivity remains unchanged when it is compared to the one obtained from a point-like source. This is plausible as the source dimensions do not exceed the position resolution of the detector. These results are consistent with previous ones from Reference [133].
The source is located 8.51 m below the detector center with an uncertainty smaller than 1 cm [165]. Assuming a spherical vessel with 4.25 m radius and using Equation 4.2 leads to an uncertainty of 0.2 % in the event number. Note this number would be consistent with a position reconstruction shift of 1 cm in the z-coordinate (see Table 7.1).

7.1.4 Source-Induced Backgrounds

As discussed in Chapter 6, source-induced background is only from an accidental nature that can be accurately ($< 0.3\,\%$) measured during data taking. In addition, the total expected event number of source-induced background is less than 10 events in 1.5 y of data taking. As the signal event number is by a factor of more than 200 higher than the background, the sensitivity is almost not affected by source-correlated background, as will be discussed in Section 7.2.4 for the study of the detector-correlated background.
Note that this result is only valid given the source requirements (see Section 2.2). In contrast, a too large background of the source could, on the one hand, lead to a wrong activity measurement and, on the other hand, possibly mimic a sterile neutrino signature. The

too high contamination levels of the source produced end of 2016 were one reason for the cancelation of the SOX project.

7.2 Detector-Correlated Uncertainties

One advantage of the SOX experiment is that the Borexino detector has already been taking data for more than ten years and has been studied in detail. While the detector is well characterized in the fiducial volume (FV) used for the solar analysis (typically $r < 3\,\mathrm{m}$), the systematic uncertainties increase towards the vessel border. The systematic uncertainties arising from the comparison of calibration data of 2009 are summarized in Chapter 5. Given that the calibration campaign was performed several years ago and the main focus was the central detector, a new SOX-dedicated calibration campaign with positron and neutron sources (see Chapter 5) was planned and would have taken place before the start of SOX. In the following the effect of the systematic uncertainties on the sensitivity without the new calibration campaign are studied.

7.2.1 Energy Resolution

Chapter 5 discussed that the energy resolution degrades for larger radii (see Figure 5.1 and 5.2). Such a degradation smears the oscillation signature in the data for large radii, shown in Figure 7.4a, and reduces the sensitivity. In the default analysis, the data is binned in E and L which includes no information about the radius and thus averages the energy resolution. In order to increase the sensitivity, three-dimensional data (E, L, and R) and probability distribution functions are implemented

$$pdf^{\mathrm{S}}_{ij}(\sin^2(2\theta), \Delta m^2) \rightarrow pdf^{\mathrm{S}}_{ijk}(\sin^2(2\theta), \Delta m^2), \tag{7.10}$$

where k denotes the radial bin.
Figure 7.4b shows the comparison of the sensitivity for the two- and three-dimensional data. Only a minimal improvement in the Δm^2-region sensitive to small oscillation lengths ($\approx$ $1\,\mathrm{eV}^2$ - $5\,\mathrm{eV}^2$) is obtained for the three-dimensional data set. The small difference can be explained as, on the one hand, the data with a radius larger than 3 m account for already $\sim 60\,\%$ of the total data (see legend in Figure 7.4a) and therefore dominate statistically. On the other hand, the degradation of the energy resolution is mainly due to the relatively high number of broken photomultiplier tubes (PMTs) in the southern part of the detector that leads to a north-south difference in the energy response. However, the L-binning in the data implicitly separates the data with low and high z-coordinates so that a radial binning does not add any further information.
Due to the almost unchanged sensitivity, the SOX analysis stays with the two-dimensional data. However, a three-dimensional analysis could be reconsidered to implement radial dependent uncertainties.

7.2.2 Effect of Position Shifts

The z-shift

Calibration data indicates a bias in the z-coordinate reconstruction, the so-called "z-shift" (see Chapter 5), that cannot be reproduced with MC simulations. The z-shift can be parameterized as

$$\Delta z = a \cdot z^2 - c, \tag{7.11}$$

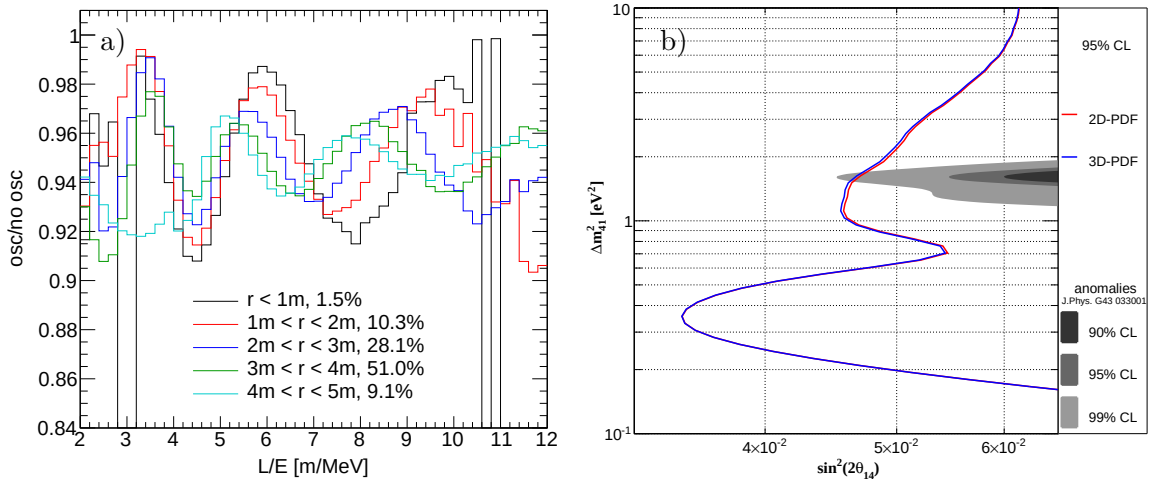


Figure 7.4: Effect of the Position-Dependent Energy Resolution on the Sensitivity a) Ratio of L/E-distributions for an oscillation ($\sin^2(2\theta)$= 0.1 and Δm^2= 2 eV^2) over the no oscillation hypothesis grouped for the radius in the detector. The energy resolution drops for larger radii and smears the oscillation signature. For each spherical shell the fraction contributing to the whole data sample is given in the legend. The small statistics for $r < 1\,m$ lead to an artifact in the signature. b) Comparison of the sensitivity for two- (E,L) and three- (E,L,R) dimensional probability distributions functions for the data with five radial bins. The sensitivity changes only slightly for $\Delta m^2 \approx 1\,\text{eV}^2$ - $5\,\text{eV}^2$.

where $a = 0.0025$, $c = 0.035$, and $[z]$ = m. The z-shift is studied in two ways (see also Reference [205]). First, it is assumed that the z-shift is present in the data, while the expected data used in the likelihood fit do not take the shift into account. The shifted data is generated by applying the z-shift, as defined in Equation 7.11, to the nominal MC simulations, which are used to build the expected data.

The position shift affects the rate and the shape analysis. With a FV cut of 25 cm to the inner vessel the event number in the data is reduced by 0.6 % compared to the expectation. Assuming an unchanged vessel shape, but applying a z-shift that reduces the reconstructed z-coordinate, events that were previously within the FV are now reconstructed outside the FV. The effect on the shape analysis is shown in Figure 7.5. The sensitivity including the z-shift fluctuates around the nominal sensitivity especially in the Δm^2-region below 2 eV^2. These fluctuations are in addition illustrated in Figure 7.6b which shows the ratio of the two sensitivity contours. They can be explained with the relative changes in the L and E spectra of the data and the expectation. The negative z-shift deforms the L spectrum to smaller values and at the same time the E spectrum to larger values. This is because the data from the north with the higher light yield is shifted towards the south. Both shifts result in an increase at small values of the L/E spectrum for the data compared to the expectation, which is shown in Figure 7.6a. Thus, the relative increase in the spectrum at ~ 3 m/MeV disagrees with a disappearance signature and the sensitivity becomes effectively improved for the respective oscillation length, i.e. ~ 5 m/MeV. Analogously, the relative reduction of the L/E spectrum above 5 m/MeV mimics exactly a sterile neutrino signature with a large oscillation length of ~ 10 m/MeV and reduces the sensitivity for the respective Δm^2-region at $\sim 2.5\,\text{eV}^2$.

In the second approach, the z-shift is implemented in the likelihood fit model, where a and c of Equation 7.11 become nuisance parameters. The expected event number and probability functions depend on a and c and the likelihood function minimizes in addition to the de-

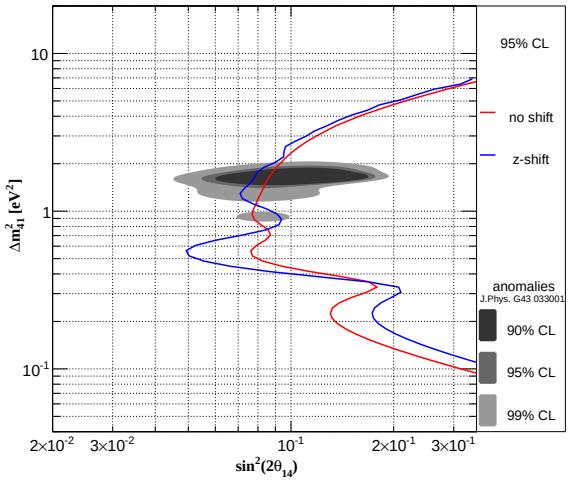


Figure 7.5: Effect of the z-shift on the Shape Analysis The sensitivity at the 95 % CL is shown when the data is modified with the z-shift (blue) and when the data is not modified (red). The expected data is in both cases not modified by the shift. The results are computed for a FV cut of 25 cm.

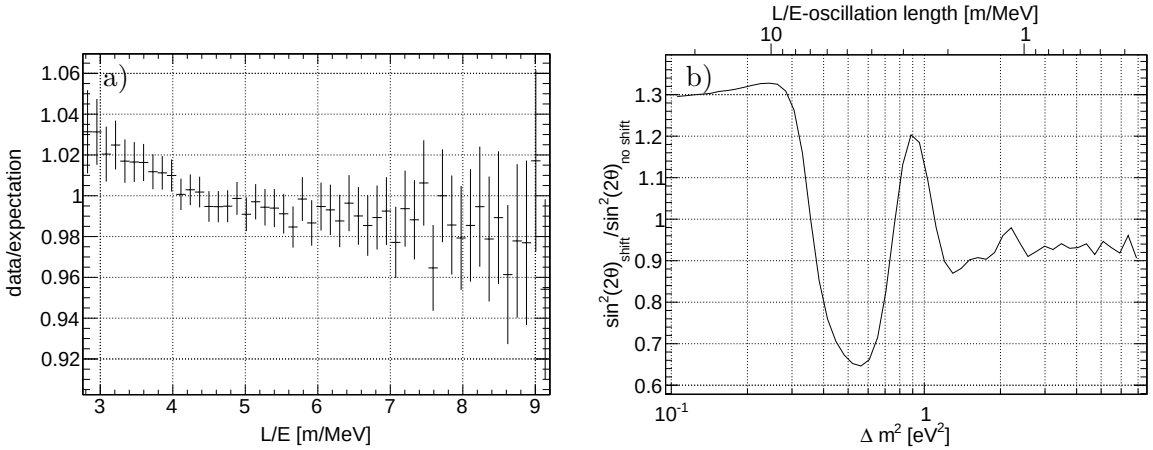


Figure 7.6: Effect of the z-shift on the Spectral Shape a) Relative difference in the L/E spectra for the data and the expected data. b) Relative difference in the sensitivity at the 95 % CL as a function of Δm^2. The absolute sensitivities are shown in Figure 7.5. The results are obtained, when the data is modified with the position shift and the expectation not. A FV cut of 25 cm is applied.

fault parameters over these two new parameters. Even though the shift parameters are not constrained in the fit, the fit always returns the true values for them. This implies that the z-shift constructed with the parametrization given in Equation 7.11 cannot mimic a sterile neutrino signature. Note that this study is performed with the Asimov data set. Further investigations with toy-MC data samples are needed in order to draw final conclusions. In addition, a more flexible model for the position shift could also affect the results.

Miscellaneous Position Shifts

Furthermore, the effect of various position shifts is discussed even though they are not a specific systematic uncertainty of Borexino. However, one would like to understand how much these shifts can affect the sensitivity. The same method as described for the z-shift is

Table 7.1: Uncertainty on the Event Number due to Systematic Position Shifts The results are obtained by modifying the MC simulation data with the respective position shifts and applying a FV cut of 25 cm. The results are supposed to give an estimate of the order of magnitude. Not all of the studied shifts are present in Borexino.

type of shift	parametrization [cm]	uncertainty [%]
"z-shift"	$0.25(z/\mathrm{m})^2$-3.5	0.6
absolute position shifts in x and y	5	<0.07
absolute position shifts in z	1	0.2
	2	0.5
	3	0.7
	4	0.9
	5	1.1
outward shift in x	2	0.7
	5	1.8
random shifts in x, y	2	<0.05

applied. The data is modified by a certain position shift, while the expected data remains unshifted. Table 7.1 summarizes the effect of these shifts on the uncertainty in the event number when a FV cut of 25 cm is applied. The following shifts are studied:

- absolute position shifts: Even though absolute position shifts up to 5 cm in x- and y-direction are negligible, a 5 cm shift along the z-direction can lead to an uncertainty in the event number up to 1.1 %. Figure 7.7a shows the impact on the shape only analysis for a negative 5 cm and 10 cm shift. Similarly to the previously discussed "z-shift", the relative deformation in the L/E spectra can be described by some oscillation pattern and leads to fluctuations around the nominal sensitivity;

- outward shifts: Outward shifts in the x- and y-coordinate can lead to uncertainties in the event number up to 1.8 %. However, fluctuations in the shape analysis appear mainly below 1.5 eV^2 (see Figure 7.7b);

- random position shifts: The calibration data shows an agreement with the nominal position within 2 cm [189]. Such random shifts are smaller than the size of the position resolution and have a negligible impact on the expected event number and the expected signature. Thus, the sensitivity is not affected.

To summarize, the uncertainty in the event number for position shifts up to 5 cm can lead to uncertainties up to 1.8 % in the event number. However, as only the extreme cases are beyond the 1 % level, the uncertainty due to a position shift can be assumed to be $< 1\,\%$. Position shifts mainly affect the low Δm^2-region below 1 eV^2 on the shape analysis. Thus, the studied position shifts cannot mimic a sterile neutrino signature above 1 eV^2 and hence not mimic the smoking gun signature of a sterile neutrino. In addition, the relative variations of the sensitivity due to a position shift are smaller or comparable to the expected statistical fluctuations of the sensitivity and are hence not critical (see Chapter 11). Note that the vessel shape is assumed to be known in this study. However, it is likely that a position shift also changes the estimated vessel shape, as the vessel shape reconstruction is applied to the reconstructed position, which would require further investigation.

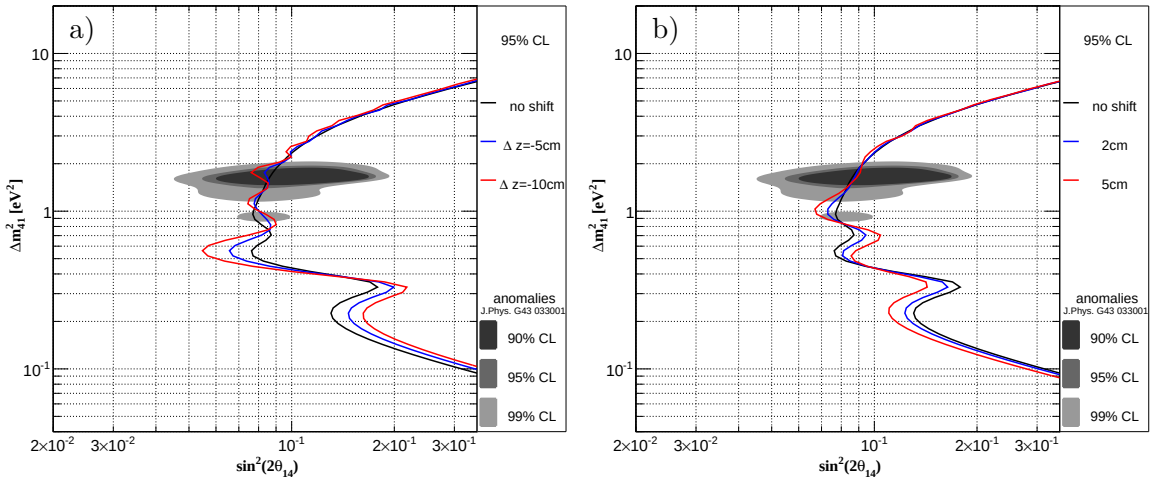


Figure 7.7: Effect of Systematic Position Shifts on the Shape Analysis The sensitivity at the 95 % CL is shown when the data is modified with a) absolute position shifts in the z-coordinate and an b) outward shift in the x-coordinate. The expected data is in both cases not shifted. A FV cut of 25 cm is applied.

7.2.3 Vessel Shape, Detection Efficiency, and the Fiducial Volume Cut

The largest uncertainty for the SOX experiment is the uncertainty in the vessel shape. As the vessel shape changes with time, it needs to be reconstructed on a weekly basis where the uncertainties reach up to 5 cm (see Section 2.1.2). To minimize the uncertainty, the algorithm for vessel shape reconstruction has been improved within the collaboration. An uncertain vessel shape leads not only to an uncertainty in the active volume and thus the expected event number. Moreover, the exact vessel shape affects the detector response and the detection efficiency close to the vessel border (see Chapter 5).

In the following, the effect of the vessel shape is studied and the FV is optimized with respect to the sensitivity taking into account the reducing uncertainties and statistics for larger FV cuts. To assess the effect of an uncertainty of 5 cm in the vessel shape, MC simulations are produced for three different vessel shapes [190]. Here, the vessel shape from October 2016 is shifted 5 cm inwards and outwards in comparison to the nominal one (see Figure A.7 in the Appendix). The uncertainty on the event number as a function of the FV cut is determined as the difference of the event number for the three data sets and the specific FV cut with respect to the nominal vessel shape (see also Reference [190] for a similar analysis). Thus, the uncertainty in the event number already includes the uncertainty in the active volume, the detection efficiency and the energy and position reconstruction. The resulting uncertainty is shown in Figure 7.8 for various energy cuts on the prompt and the delayed event. Applying no FV cut would result in an uncertainty of almost 4 % (see panel b), mainly due to the difference in the active volume. Such a large uncertainty is too large for the SOX analysis and would drastically reduce the sensitivity (see Figure 7.1). The uncertainty decreases with increasing FV cut and reaches a constant level of 0.2 % above a FV cut of 50 cm. Furthermore, the uncertainty reduces for smaller energy cuts on the prompt and the delayed event, respectively. This is because of the increase of the total detection efficiency and enters into the optimization of the selection cuts which was discussed in Chaper 5. In addition, the uncertainty on the event number computed only from the variation of the detection efficiency for a fixed vessel shape (see Figure 5.14) within 5 cm is compared and agrees well with the one obtained from the MC simulation (see the gray line in Figure 7.8b). However, this method

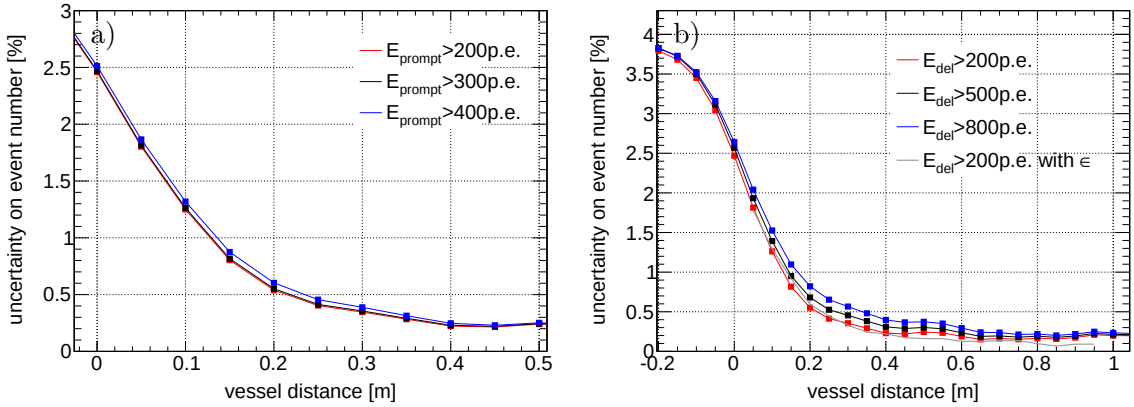


Figure 7.8: Uncertainty on the Event Number due to the Vessel Shape Uncertainty The uncertainty is shown as a function of the vessel distance for different cut values on the prompt (a) and delayed (b) energy. The uncertainty is obtained from MC simulations assuming three different vessel shapes in the range of the vessel uncertainty (see text). For comparison, the uncertainty of the detection efficiency computed only from the variation of the detection efficiency for a fixed vessel shape (see Figure 5.14) within 5 cm is added for the optimized delayed energy cut value of 200 p.e..

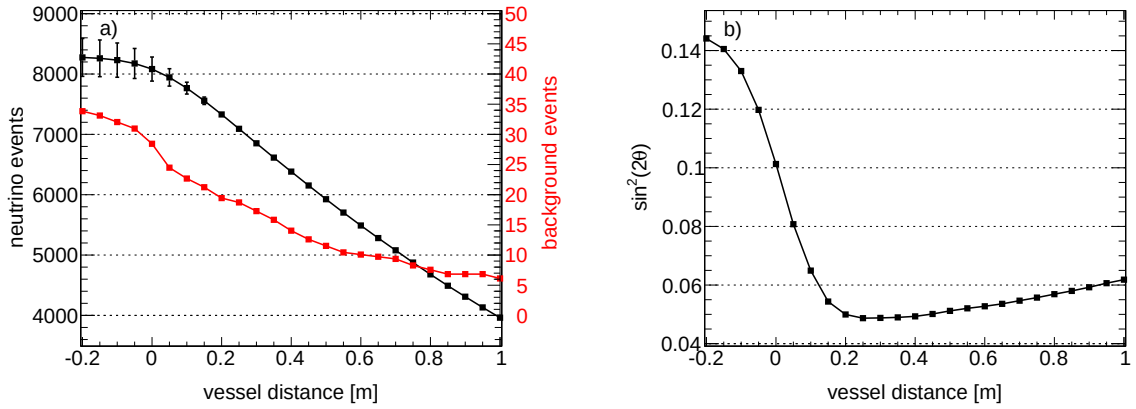


Figure 7.9: Optimization of the FV Cut a) The signal and background event numbers are shown as a function of the FV cut. The uncertainties on the event number are taken from Figure 7.8 and decrease for increasing FV cut. The background event numbers are determined from Borexino data (see Chapter 6). b) The rate sensitivity in terms of $\sin^2(2\theta)$ at $\Delta m^2 = 10\,\mathrm{eV}^2$ is shown as a function of the FV cut and is computed as described in Appendix A.5. The sensitivity is maximal for a FV cut of 25 cm. The results are computed for an activity of 125 kCi and the optimized selection cuts.

gives only meaningful results above a FV cut of 5 cm when the active volume is the same for all three vessel shapes. Here, the uncertainty comes then only from the detection efficiency and the energy and position response.

Using these uncertainties and the detector-related background event numbers studied in the previous chapter, the FV cut is optimized with respect to the maximum sensitivity. The rate only sensitivity is computed (see Appendix A.5) for the various FV cuts and depicted in Figure 7.9. As the uncertainty on the event number decreases, the sensitivity increases first for an increasing FV cut up to 25 cm. For further increasing FV cuts, the expected event number becomes too small and the sensitivity reduces. The FV cut with the maximum sensitivity at 25 cm is hence used as the optimized FV cut. The background does not affect

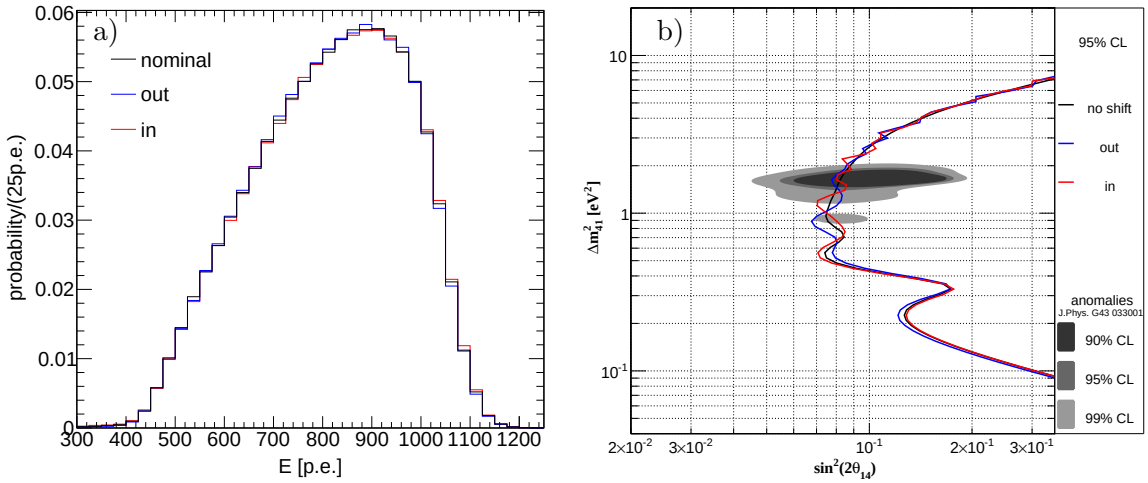


Figure 7.10: Effect of the Uncertainty in the Vessel Shape on the Spectral Shape a) The normalized energy spectra for the MC simulations with the three different vessel shapes (Figure A.7) are shown for a FV cut of 25 cm. b) The effect on the shape sensitivity at the 95 % CL due to the differences in the spectral shape is shown. While for the data the shifted vessel shape is used, the expected data assumes the nominal vessel shape. The results were produced in Reference [190].

the conclusions, but for completeness, it is taken into account.

The uncertainty in the vessel shape affects not only the expected event number, but also the energy and distance spectrum. Close to the vessel border the energy scale is affected (see Chapter 5), however, for a FV cut of 25 cm the difference is reduced. Moreover, the energy spectrum from the positrons produced in the center of the detector changes, as the light yield reduces slightly for an increasing vessel shape. This effect comes from the higher attenuation length of DMP in comparison to PPO in the wavelength region of the scintillation light [184]. The average distance traversed by the scintillation light through PPO increases for larger vessel shapes and the light yield is decreased. Figure 7.10a shows the normalized energy spectra for the three studied vessel shapes. The small differences in the energy distribution lead to small fluctuations around the nominal sensitivity, which is shown in Figure 7.10b. However, these fluctuations are small and well below the statistical fluctuations in the data set and the sensitivity.

In order to maximize the sensitivity of SOX, the collaboration studied to increase the FV even further and include at the same time nuisance parameters in the likelihood fit to account for the vessel shape uncertainty [190, 206]. On the one hand, a detection efficiency map as a function of the detector position and the radial information in the data, and on the other hand, the interpolation and extrapolation of the three vessel shape MC simulations are used. In both studies the likelihood fit returns the true vessel shape and a vessel shape uncertainty cannot be explained by any sterile neutrino signature. This is similar to what is obtained in the z-shift study (Section 7.2.2). Further studies with a more flexible model for the vessel shape in the likelihood function would have been required for the SOX analysis. The option to reduce the uncertainty on the detection efficiency and hence to allow for an increased FV by the hardware modification of the trigger system is discussed in Chapter 8.

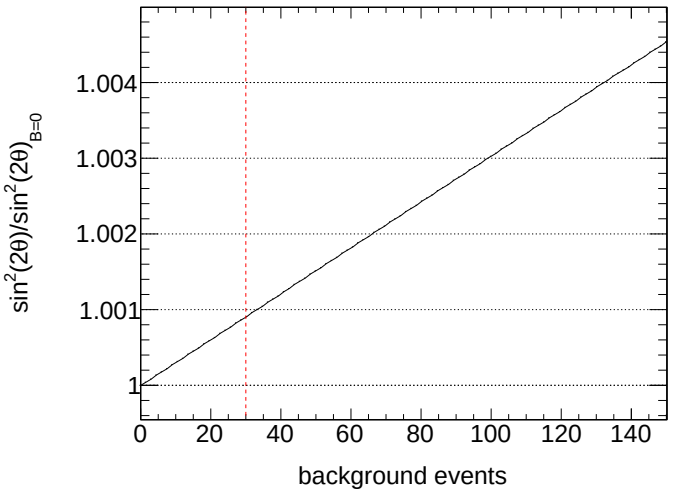


Figure 7.11: Effect of Background on Sensitivity The increase of the sensitive $\sin^2(2\theta)$-value at the 95 % CL at $\Delta m^2 = 10\,\mathrm{eV}^2$ in comparison to the sensitivity with zero background is shown as a function of the background event number. The sensitivity is computed in a rate analysis (see Appendix A.5) for an activity of 125 kCi with an uncertainty of 1.5 %. The background uncertainty is given by 10 %. The red line indicates the background level in SOX.

7.2.4 Detector-Related Backgrounds

The detector-related background can be determined from the running Borexino experiment with an accuracy better than 10 % (see Chapter 6). Therefore, the expected event number as well as the spectral shape are well known.
In order to study the background, the likelihood function incorporates the background component that is constrained at the 10 % level (see Equation 3.4). The inclusion of the background has no visible effect on the sensitivity of the SOX experiment, as the expected signal to background ratio is larger than 200. Figure 7.11 shows the sensitivity as a function of the background event number in a rate only analysis (using the formalism described in Appendix A.5). The sensitivity decreases on a minimal level (smaller than 0.1 %) for the expected background event number of 30 events which already includes the source-related background. Thus, the background in SOX is totally negligible and SOX can be called a "background-free" experiment.

7.3 Summary of Systematics and Sensitivity in SOX

Table 7.2 summarizes the systematics of the SOX experiment and their impact on the event number and the shape analysis, respectively. Notice, an uncertainty on the energy scale was not studied explicitly. However, the expected effect on the detection efficiency and therefore on the event number can be assumed to be small as the energy selection cuts are chosen in a conservative way. The effect on the shape analysis is indirectly studied by the position and vessel shape shifts that also alter the energy distribution. Small fluctuations around the nominal sensitivity arise mainly for Δm^2-values below $1\,\mathrm{eV}^2$ and the smoking gun signature for a sterile neutrino cannot be mimicked.
Figure 7.12 shows the sensitivities for the rate, shape, and rate+shape analysis for the optimized FV cut of 25 cm and the expected systematic uncertainties. The uncertainty on the event number is set to a conservative value of 1.5 % that includes the uncertainty on the power measurement, the detection efficiency, and position shifts. The sensitivities are shown

Table 7.2: Summary of Systematic Uncertainties in SOX The systematic parameter, its central value and the uncertainty are listed in the first three columns. The effect on the event number and on the shape analysis that are evaluated in this chapter are given in column three and four. The systematics are grouped in source and detector-correlated ones and are computed for a FV cut of 25 cm.

parameter	central value	uncertainty	uncertainty on event number	effect on shape analysis
Source-Related				
power	~ 1000 W	0.2 %	0.2 %	-
shape factor	0	0.03	< 1 %	< 10 % for $\Delta m^2 < 0.7\,\mathrm{eV}^2$
position	8.51 m	< 1 cm	< 0.2 %	< 2 % for $\Delta m^2 < 1\,\mathrm{eV}^2$
Detector-Related				
position shift		< 4.5 cm	~ 1 %	35 % for $\Delta m^2 < 1\,\mathrm{eV}^2$
vessel shape		5 cm	-	< 10 %
detection efficiency	93.0 %	0.5 %	0.5 %	-
background	~ 30	10 %	-	-

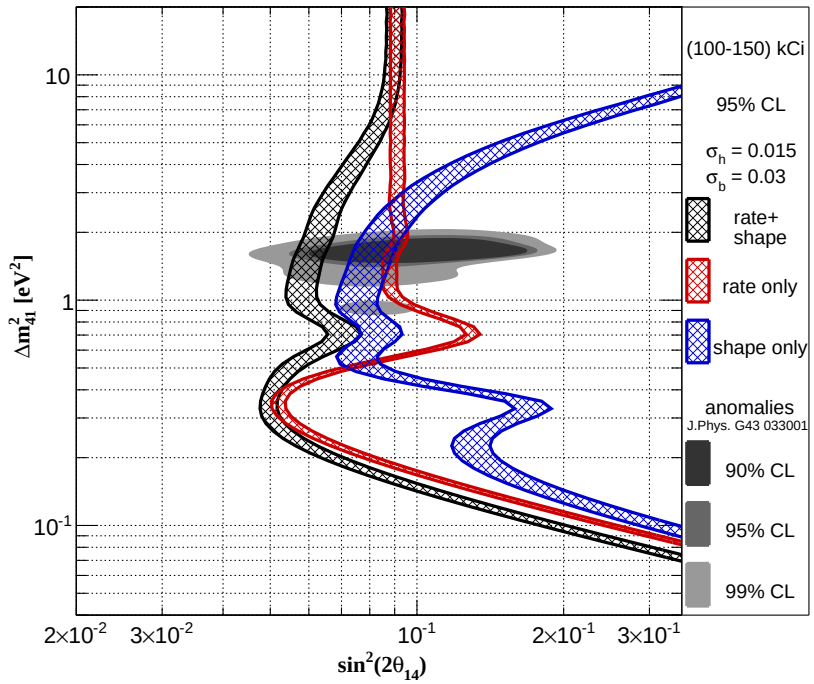


Figure 7.12: Sensitivity of the SOX experiment The sensitivities for the rate (red), shape (blue), and combined analysis (black) are shown for the optimized selection cuts (FV cut 25 cm) including the expected systematic uncertainties (b: spectral shape, h: event number). The sensitivity bands indicate the variation for the possible source activity range of 100-150 kCi. The parameter region preferred by the anomalous neutrino oscillation data is shown in gray.

in bands that span the possible activity range from 100-150 kCi corresponding to an event number between 5760 and 8640. The final rate+shape sensitivity can probe the full 90 % CL parameter region preferred by the anomalous neutrino oscillation data at the 95 % CL.

Note that the sensitivities are obtained assuming the asymptotic distributions of the profile likelihood ratio. In Chapter 11 the sensitivity of SOX is compared and discussed in detail to the ones obtained from toy-MC techniques.
Further, it has to be noted that a dedicated SOX-calibration campaign is mandatory to scan and characterize the detector especially close to the vessel border for positrons and neutrons. Moreover, the vessel shape reconstruction and position reconstruction close to the vessel border need to be improved. The resulting discrepancies have to be studied and taken into account in the final SOX analysis. In addition, the uncertainties on the detection efficiency could be lowered by a hardware modification of the trigger system which will be discussed in the next chapter.

8 Evaluation of an Optimized Trigger System

As discussed in the previous chapter, the shape of the inner vessel (IV) dominates the systematic uncertainties of SOX. As the uncertain vessel shape also increases the uncertainty of the detection efficiency close to the vessel border, a large fiducial volume (FV) cut needs to be applied. However, such a large FV cut decreases the statistics of the data sample and, hence, the sensitivity of the experiment. In this chapter, the modification of the Borexino trigger is studied to decrease the uncertainty on the detection efficiency induced by the vessel shape, and to improve the sensitivity of SOX in comparison to the one presented in the previous chapter. The modification is based on the fact that the light yield in the inner buffer (IB) is actually higher than designed due to a leakage of PPO from the IV to the IB. Provided a small detection threshold, this increased light yield enables the reconstruction of events interacting in the buffer medium, as discussed in Section 8.1. Two modifications of the hardware trigger are evaluated[1]: a reduced trigger threshold (Section 8.2) and an automatically issued neutron gate after a possible candidate for a prompt SOX event (Section 8.3).

8.1 Detector Response in the Buffer

While the active volume contained in the IV of the Borexino detector can be calibrated with radioactive sources, the buffer is out of reach of the calibration system [189]. However, cosmogenic neutrons can be used to determine the detector response in the buffer and to optimize Monte Carlo (MC) simulations, which is discussed in Section 8.1.1. Using the optimized MC simulation, the expected SOX signature due to events interacting in the buffer is studied in Section 8.1.2.

8.1.1 Light Yield of Cosmogenic Neutrons in the Buffer

Cosmogenic neutrons observed from 2008 to 2016 show a significantly higher light yield in the IB region than in the outer buffer (OB). This can be explained by the leak from the scintillator of the IV, especially the wavelength shifter PPO, to the IB [199]. The leak started in April 2008 with a rate of $1.33\,\mathrm{m}^3$/month. In order to reduce the density differences between IV and the buffer medium, the DMP concentration in the buffer was lowered from 5 g/l to 3 g/l and 2 g/l in April 2009 and January 2010, respectively. Thus, the leak rate was reduced to $0.56\,\mathrm{m}^3$/month and $1.5\,\mathrm{m}^3$/year, respectively [141]. Given the leak evolution, the PPO concentration in 2017 can be estimated to 0.12 g/l in the IB [207].
Figure 8.1 shows the energy response of the 2.2 MeV and 4.9 MeV gammas induced by cosmogenic neutron capture on Hydrogen and Carbon as a function of the reconstructed ra-

[1]The modifications were studied together with several people from the Borexino/SOX collaboration.

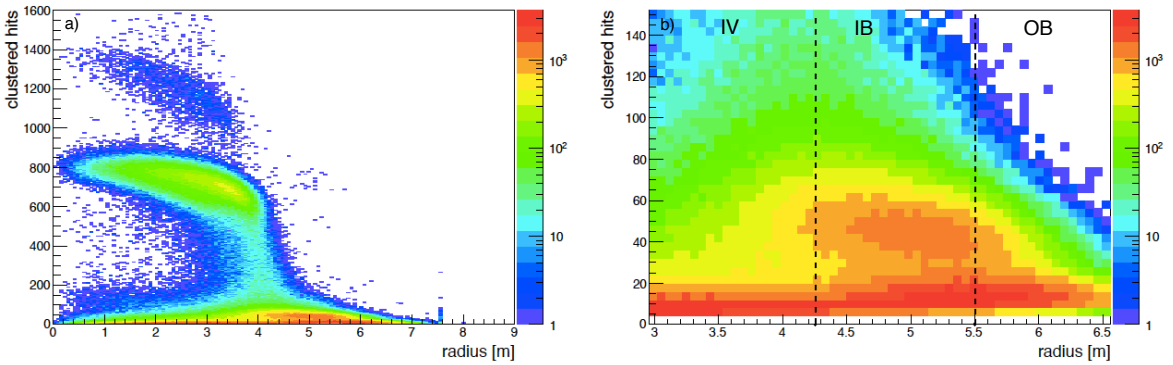


Figure 8.1: Energy Response of Cosmogenic Neutron Data The energy is shown in nhits (normalized to the number of working PMTs) versus the reconstructed radius in the a) full detector volume and b) inner buffer region. The events are selected in a time period up to 1.6 ms after a muon and with a distance of less than two meters to the reconstructed muon track. The dashed lines indicate the nominal radius of the nylon two vessels. Three regions can be identified that correspond to the neutrons captured in the IV ($r < 4.25\,\mathrm{m}$), IB ($4.25\,\mathrm{m} < r < 5.5\,\mathrm{m}$), and OB ($5.5\,\mathrm{m} < r < 6.3\,\mathrm{m}$), respectively. The light yield in the IB is increased in comparison to the nominal one in the OB, due to the IV leak. Moreover, substructures at low energies due to ^{14}C in the IV and mis-reconstructed neutrons are visible. Both plots are taken from [199].

dius. Three energy regions can be identified that correspond to neutrons captured in the IV (hits>500), IB (hits∼40), and OB (hits∼20). The latter light yield corresponds to the one expected without PPO in the buffer, i.e. the nominal value [199].
In an annual analysis of cosmogenic neutrons, it can be found that the light yield of the IB is similar to the one of the OB in 2008. From 2008 to 2010 the light yield increases and stays constant afterwards in the IB which is consistent with the leak evolution. Further, cosmogenic neutron data shows that the light yield is uniform in the whole IB, which leads to the conclusion that the scintillator and mainly the PPO mixed well within the IB [199].
As the light yield of the mixture of PC, DMP, and PPO of the IB has not yet been experimentally measured, the effective light yield of the IB is optimized in the MC simulation in order to be consistent with the cosmogenic neutron data. With this optimized configuration, the MC simulation can reproduce the response in the whole detector volume (IV, IB, and OB) as measured with the cosmogenic neutron data [207].
In summary, the high light yield in the buffer enables to detect neutrons also in the buffer medium. It allows therefore to exploit events interacting in the buffer for the SOX analysis.

8.1.2 SOX Events in the Buffer

With the optimized MC simulation that can reproduce the observed high light yield in the IB, 2×10^6 SOX events are simulated in the IV and in the IB (up to a radius of 5.5 m), where the detector status is set to October 2016. Figure 8.2 shows the energy response as a function of the reconstructed radius for the prompt and the delayed event[2]. The prompt (delayed) events with energies above 200 (300) hits deposit their full energy in the IV. The events reconstructed at $\sim 4\,\mathrm{m}$ with an energy in between 40-200 (100-300) hits deposit their energy partially in the IV and partially in the IB. Especially, one of the two 511 keV annihilation gammas of the prompt event can deposit its energy in the IV and the other one in the buffer. The substructure at 100 hits in panel a is mainly due to inverse beta decay (IBD) reactions

[2]The energy observable is in the following the actual observed value and not normalized to the number of 2000 photomultiplier tubes (PMTs), as described in Section 5.1.

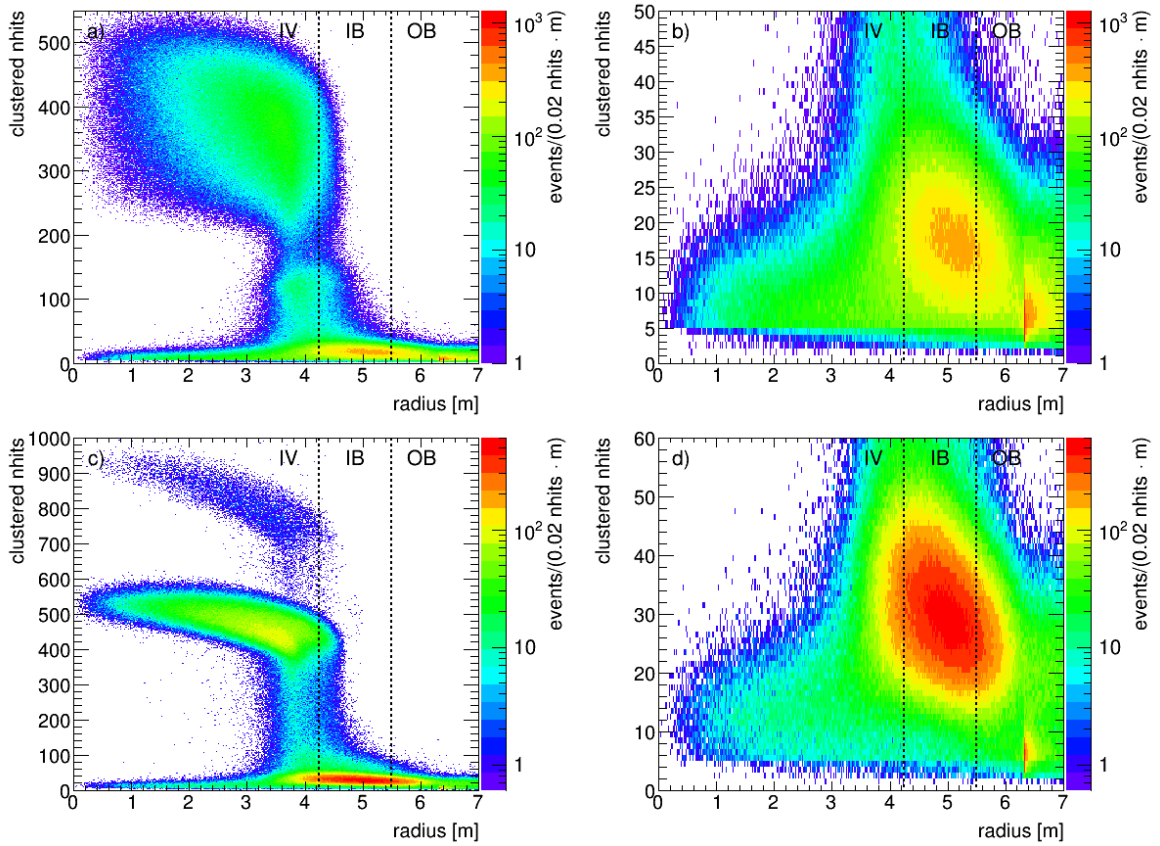


Figure 8.2: Energy Response of Simulated SOX Events in the IV and IB The energy is shown in (non-normalized) nhits versus the radius for 2×10^6 SOX events generated in the IV and IB. The prompt (panel a and b) and the delayed (panel c and d) events are selected with $\Delta t < 2000\,\mu s$. Panel b) and d) show a zoom in the buffer region, in which the energy spectrum of the prompt and the delayed event can be reconstructed. The dashed lines indicate the nominal radius of the inner and outer vessel. The cutoff at 5 hits is due to the efficiency threshold of the cluster algorithm. It can be seen that the position reconstruction becomes worse for low energies, where the structure at 6.3 m corresponds to an artifact in the reconstruction algorithm. Additional discussion about the individual substructures can be found in the text.

in the IB, where one gamma spills in the IV. The events below 50 (60) hits correspond to the ones interacting in the IB. The large deviation in the reconstructed radius from the true radial position for the IB-events shows the inefficiency of the position reconstruction algorithm at low energies. Another limitation of the position algorithm is the reconstructed structure at 6.3 m that corresponds to the position of the PMTs mounted on the stainless sphere. Moreover, the efficiency-threshold of the cluster algorithm can be seen at the kink at 5 hits in the prompt as well as in the delayed spectrum[3]. Within the collaboration the improvement on the position reconstruction and the cluster algorithm at such low energies was an ongoing effort.

As the Borexino trigger is issued when at least 20 hits are detected, a large fraction of the prompt and the delayed events in the buffer cannot be observed. The next two sections discuss hence two different options to modify the Borexino trigger system.

[3]The cluster algorithm looks for an accumulation of hits within a short time period. The reconstructed cluster is associated to a physical event.

8.2 Option 1: Reduction of the Trigger Threshold

An attractive and easy extension of the trigger system is to simply lower the trigger threshold so that the prompt and the delayed signal can be fully detected in the buffer. These additional events could be used for a sterile neutrino analysis and could improve the sensitivity of the experiment. In the following, the increase in the sensitivity and the feasibility from the data taking point of view are discussed.

8.2.1 Detection Efficiency, Signature, and Sensitivity

The detection efficiency in the buffer is reduced in comparison to the IV, as a lot of low energy events are mis-reconstructed. Figure 8.3 shows the comparison of the reconstructed and true radius for coincident SOX events selected with $\Delta R < 2\,\mathrm{m}$ and $\Delta t < 1280\,\mu\mathrm{s}$. While IBD reactions taking place in the IV can be well reconstructed, the reconstructed radius distribution from the events in the IB shows a large deviation from the true radial position (red and blue solid line). Moreover, the number of selected IBD events drops in the IB, as the requirement of spatial correlation is not fulfilled. With the selection cuts of $E_p = 10 - 40$ hits, $E_d = 10 - 60$ hits, $vd < 0.25\,\mathrm{m}$, $r < 5.75\,\mathrm{m}$ (radius of prompt event), $\Delta R < 1.75\,\mathrm{m}$, and $\Delta t = 1.52\,\mu\mathrm{s} - 13.48\,\mu\mathrm{s}$ or $19\,\mu\mathrm{s} - 1280\,\mu\mathrm{s}$, a detection efficiency of 25.6 % is reached in the buffer. The minimum energy cut on both events is given by the cluster inefficiency below 10 hits. Given this detection efficiency, a total event number of ~ 3600 events is expected for a measurement time of 1.5 years and an activity of 125 kCi.
The expected signature in the buffer as a function of energy and distance is shown in Figure 8.4. It can be seen, that a sterile neutrino with the oscillation parameters $\sin^2(2\theta)$=0.1 and Δm^2=0.5 eV^2 produces a detectable signal in the buffer. While the oscillatory pattern is mostly washed out by the poor energy resolution in the energy spectrum, it can be still observed in the distance spectrum. Moreover, the reconstructed distance range becomes larger

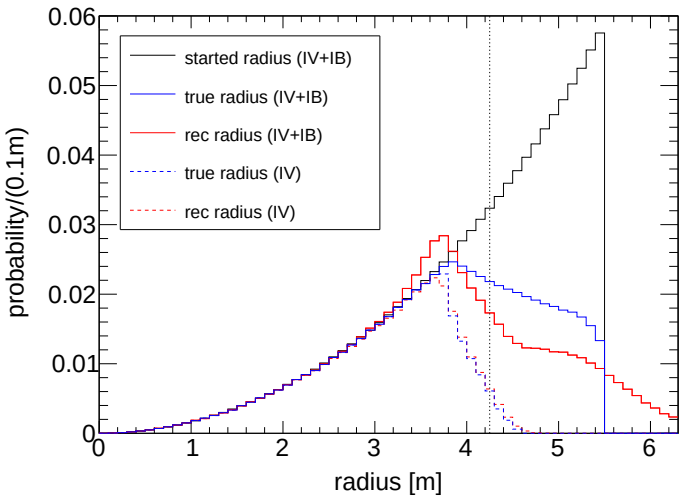


Figure 8.3: Position Reconstruction of SOX Events in the Buffer The reconstructed radius (red) is compared to the true radius (blue) of the prompt events generated in the IV (dashed) or IV+IB (solid). The distributions are obtained from MC simulations of 2×10^6 SOX events generated in the IV+IB (black) and selected with $\Delta R < 2\,\mathrm{m}$ and $\Delta t < 1280\,\mu\mathrm{s}$. The dashed line indicates the nominal radius of the IV. While the IBD reactions taking place in the IV are well reconstructed, a large fraction of the events in the buffer are mis-reconstructed and reduce the detection efficiency.

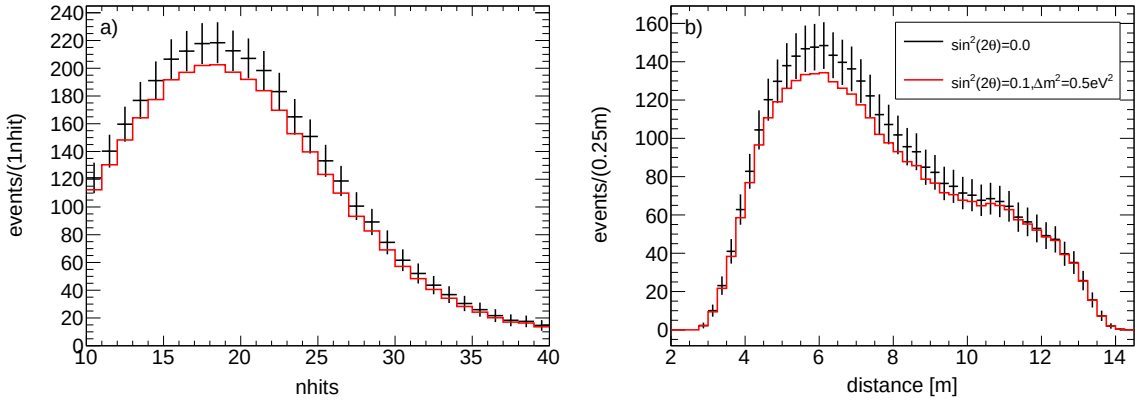


Figure 8.4: Sterile Neutrino Signature in the Buffer The energy and distance spectra are shown for the no-sterile and a sterile neutrino hypothesis. In total 3600 events are expected in the buffer for the no-sterile hypothesis, for which the statistical error bars are indicated. The distributions are obtained from MC simulations assuming the detector status of October 2016.

than for the nominal SOX experiment, which allows to probe a larger Δm^2-range.

Figure 8.5 shows the sensitivity with the additional buffer events in a rate+shape and shape analysis. Already the sensitivity of the neutrinos interacting within the IV improves in comparison to the sensitivity with the standard Borexino trigger. This is caused by the

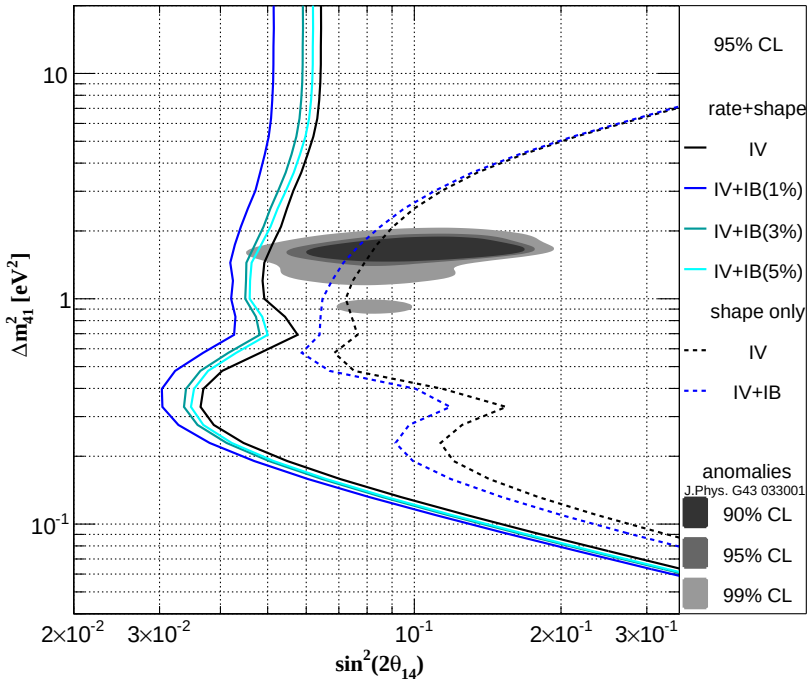


Figure 8.5: Sensitivity with a Reduced Trigger Threshold The sensitivity of the data selected in the IV (black) and in addition in the IB (blue) is shown in a rate+shape (solid) and a shape (dashed) analysis. The expected event number in the IV are 8500 events and in the IB 3600 events for a 125 kCi source. The sensitivity of the rate+shape analysis is shown for three different uncertainties on the expected event number reconstructed in the IB. The uncertainty on the event number in the IV is assumed to be 1 %. It has to be noted that no uncertainty on the spectral shape of the antineutrino spectrum and no background is taken into account.

increased detection efficiency and its accuracy close to the vessel border that allows to enlarge the FV (besides the optimized selection cuts of Table 5.1, a FV cut of 5 cm and a minimum energy cut on the delayed event of 10 hits is here assumed). Given in addition the events interacting in the buffer and an uncertainty of only 1 % on the event number, the full 99 % CL parameter space preferred by the anomalous neutrino oscillation data could be investigated. However, as radioactive sources cannot be inserted into the buffer and hence the buffer cannot be directly calibrated, the uncertainty on the detection efficiency will be large and the sensitivity is shown for increasing uncertainties on the event number. Even in the shape only analysis, the sensitivity improves with the buffer data. The improvement is mainly given in the Δm^2-region below $2\,\mathrm{eV}^2$ due to the increased distance range and the observable oscillation pattern in the distance spectrum.
It has to be noted that no background is taken into account which will be a crucial point and will be discussed in the next section. Moreover, closer distances to the source would also enhance the source-induced background components which would require further investigations.

8.2.2 Borexino Data Acquisition and Background

The nominal Borexino trigger is issued when at least 20 hits within 100 ns are detected [141]. In order to study the feasibility of the reduction of the trigger threshold, Borexino data is taken with a lowered threshold. Figure 8.6 shows the trigger rate versus the threshold which is exponentially increasing towards smaller thresholds. The trigger rate at a threshold of 6 hits is more than 10 times higher than the nominal trigger rate which leads to several problems in the data acquisition and prevents from a stable data taking. For example, due to the high trigger rate many Laben boards, where each board collects the data of eight PMTs [140], are automatically disabled and the run crashes after ~ 0.5 h.

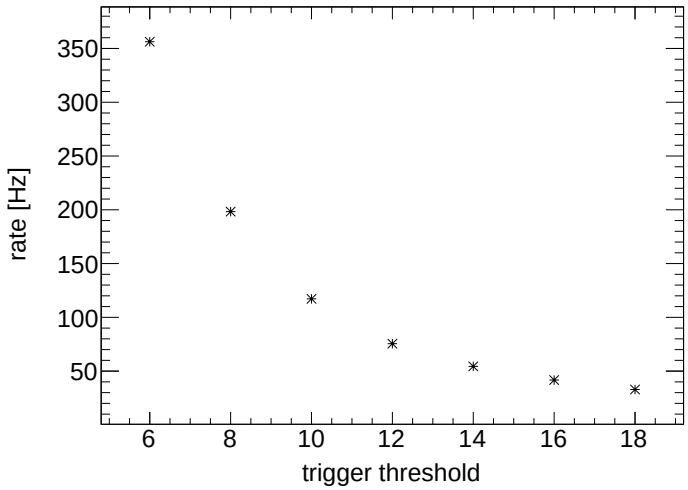


Figure 8.6: Borexino Trigger Rate as a Function of the Trigger Threshold The trigger threshold is defined as the minimum number of hits recorded within 100 ns, where the nominal Borexino trigger threshold is 20. The data was taken in June 2017.

However, the data from the run with the threshold of 6 is analyzed in order to study the feasibility from the background point of view.[4] Figure 8.7 shows the position of single events which are reconstructed within a radius of 6.3 m (i.e. position of PMTs). These events account

[4] A trigger threshold of 6 is required to achieve a full detection efficiency at 10 hits.

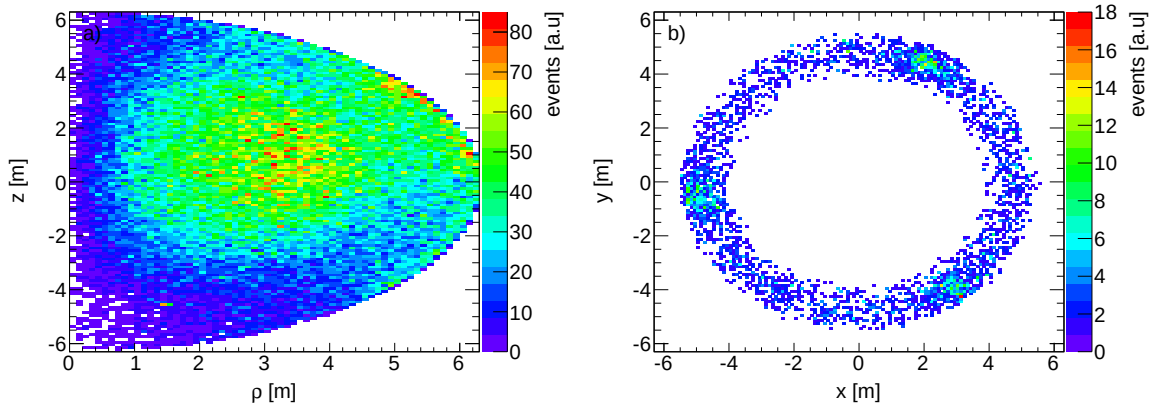


Figure 8.7: Position of Borexino Events with a Reduced Trigger Threshold The data was taken in June 2017 with a trigger threshold of 6. All events which have passed a muon cut and are reconstructed with r < 6.3 m are shown, where $\rho = \sqrt{x^2 + y^2}$. Panel b) shows the events with $z = 3\,\text{m} - 4\,\text{m}$ and $r = 5.5\,\text{m} - 6.3\,\text{m}$. The three substructures correspond to the positions of CCD cameras installed at the stainless steel sphere.

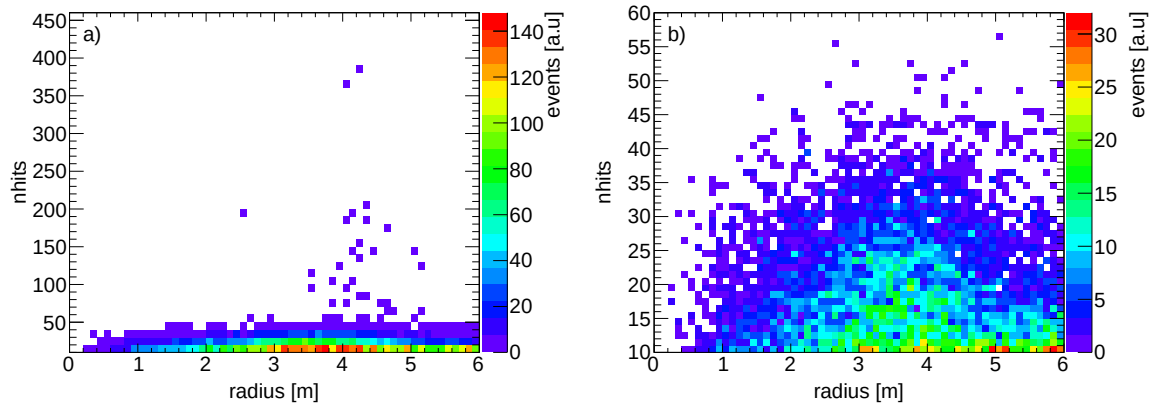


Figure 8.8: Energy versus Position of Borexino Events with a Reduced Trigger Threshold The data was taken in June 2017 with a trigger threshold of 6. All events which have passed a muon cut and are reconstructed with $r < 6.3\,\text{m}$ are shown. The energy range is shown above the cluster threshold of 10 hits. Panel b) shows a zoom in the energy region which is expected for SOX events interacting in the buffer.

only for $\sim 36\,\%$ of the trigger rate. Most of these events are in turn reconstructed within the IV and the event number in the buffer is therefore suppressed. Further, a north-south asymmetry in the event number is observed that can be explained by a larger number of working PMTs in the northern detector part. In addition, an increased event number can be observed at large radii in the north which can be traced back to the position of CCD cameras installed at the stainless steel sphere [141] (see Figure 8.7b). The energy versus the radius of the single events is shown in Figure 8.8. The energy distribution peaks at small energies due to the dominating ^{14}C background which is uniformly distributed within the detector. A few events which are mainly impurities from the IV are reconstructed above the endpoint of ^{14}C (~ 50 hits) close to the IV.

When the selection cuts from the previous section are applied, 33 events are found in the buffer for a measurement time of only 0.51 h. Obviously, the measurement time is too short to give an accurate background estimation, nevertheless, the expected high background rate

and the unstable data acquisition with the lowered trigger threshold make this modification of the trigger infeasible. Moreover, a calibration of the buffer medium with a precise knowledge on the detection efficiency and detector response would be challenging. As an alternative, the next section discusses the option to modify the trigger system so that the neutrons from the buffer can still be detected.

8.3 Option 2: Installation of a Neutron Trigger Gate

In Chapter 7, it was shown that the uncertainty of the detection efficiency increases towards larger radii, which is mainly caused by the uncertainty of the vessel shape. For all IBD reactions taking place close to the IV, a fraction of the neutrons, which depends on the specific vessel shape, are captured in the buffer and the detection efficiency is suppressed. As the increased light yield in the IB allows to detect also the neutrons from the buffer, all neutrons can be observed and the uncertainty on the detection efficiency can be decreased. However, as the energy of the neutrons captured in the buffer is partially smaller than the trigger threshold, the trigger system of Borexino has to be modified.
The idea is based on the successful working *tt*128-gate[5], which is automatically issued after muons to detect cosmogenic neutrons, and is sketched in Figure 8.9. Hence, after each high energy event that could be a possible prompt SOX event (nhits > sox-threshold), a 1.6 ms long gate is issued. To distinguish between the standard Borexino events (*tt*1), the high energy event is associated as a *tt*4-event and has the standard gate length of 1.6 µs. The automatically following neutron gate is associated as a *tt*16-event.

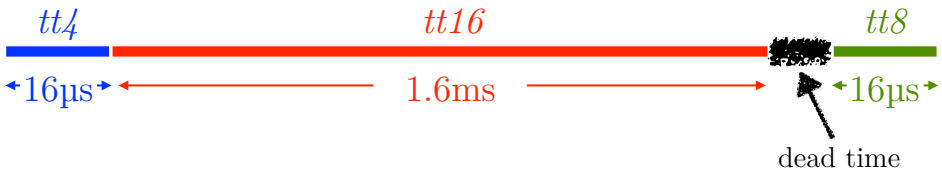


Figure 8.9: Sketch of the Gate Sequence in the SOX Trigger After each *tt*4-event a 1.6 ms long window (associated with *tt*16) is automatically issued to detect the delayed neutron signal. A *tt*4-event is assigned to events above a certain energy threshold. The event is a *tt*1-event, otherwise. The *tt*8-gate is introduced to reset an unexplained higher dark noise level after a *tt*16-event.

In the following, the improvement of the detection efficiency and the sensitivity using such a SOX trigger is compared to the standard Borexino trigger (Section 8.3.1). Moreover, the - in the Borexino hardware implemented - SOX trigger is characterized with respect to the detection and trigger efficiency, and the hit-time distributions in the *tt*4- and *tt*16-gates (Section 8.3.2-8.3.4). Subsequently, the *tt*4-threshold is optimized (Section 8.3.5) and Borexino data taken with the final configuration of the SOX trigger is analyzed in order to evaluate the expected background (Section 8.3.6).

8.3.1 Detection Efficiency and Sensitivity

With the SOX trigger and given that all spill out neutrons can be detected in the buffer, the uncertainty on the detection efficiency of the IBD depends mostly only on the detection efficiency of the prompt event. In turn, the uncertainty of the detection efficiency of the

[5] A summary of the Borexino trigger types can be found in Appendix A.2.

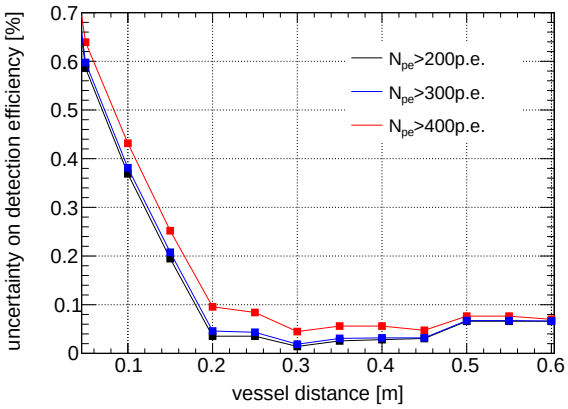


Figure 8.10: Uncertainty of the Detection Efficiency with the SOX Trigger The uncertainty is shown as a function of the FV cut and given for different energy cuts on the prompt event. The uncertainty is computed from the variation of the detection efficiency of the prompt event within ±5 cm as shown in Figure 5.5b.

prompt event depends on the uncertainty on the vessel shape that is 5 cm (see Chapter 5). However, the effect of the vessel shape on the detection efficiency is smaller for positrons than for neutrons and the overall uncertainty can be decreased. The uncertainty on the detection efficiency with the SOX trigger is shown in Figure 8.10 as a function of the FV cut. It is determined via the variations of the detection efficiency of the prompt event (see Figure 5.5) within ±5 cm for a given FV cut. This method has already been tested for the Borexino trigger and gives consistent results with the ones from full MC simulations that incorporate modified vessel shapes (see Figure 7.8). This method can only determine the uncertainties above a FV cut of 5 cm, as the active volume differs otherwise. The obtained uncertainty is largest for a FV cut of 5 cm with 0.6 %, which is about three times smaller than the uncertainty obtained with the standard Borexino trigger for the same FV cut, and drops quickly below 0.1 % at a FV cut of 20 cm.
The sensitivity with the SOX trigger is studied in Figure 8.11 with a rate only analysis considering the change in the total event numbers and their uncertainties. The sensitivity is maximal at a FV cut of 15 cm. The increasing uncertainty on the total event number and the decreasing total event number of the SOX event sample degrade the sensitivity for smaller and larger FV cuts, respectively. The sensitivity improves in comparison to the Borexino trigger, on the one hand, as the detection efficiency increases and hence the total expected event number increases. On the other hand, and more importantly, the reduction in the uncertainty of the detection efficiency improves the sensitivity and allows at the same time the enlargement of the FV. The FV cut with the SOX trigger results into 15 cm in comparison to the 25 cm of the Borexino trigger, where the sensitivity is increased by 9 %. It has to be noted that the computation does not include backgrounds. However, as studied in Chapter 7, the sensitivity is almost not affected by the background due to the large signal-to-background ratio. The backgrounds expected for the SOX trigger are discussed at the end of this chapter.

8.3.2 Trigger Efficiency and Decision Delay

In the following, the $tt4$-efficiency is studied which is defined as the fraction of $tt4$-events of all $tt1$-events and $tt4$-events as a function of the energy:

$$\epsilon(nhits) = \frac{\#tt4(nhits)}{\#tt1(nhits) + \#tt4(nhits)}. \tag{8.1}$$

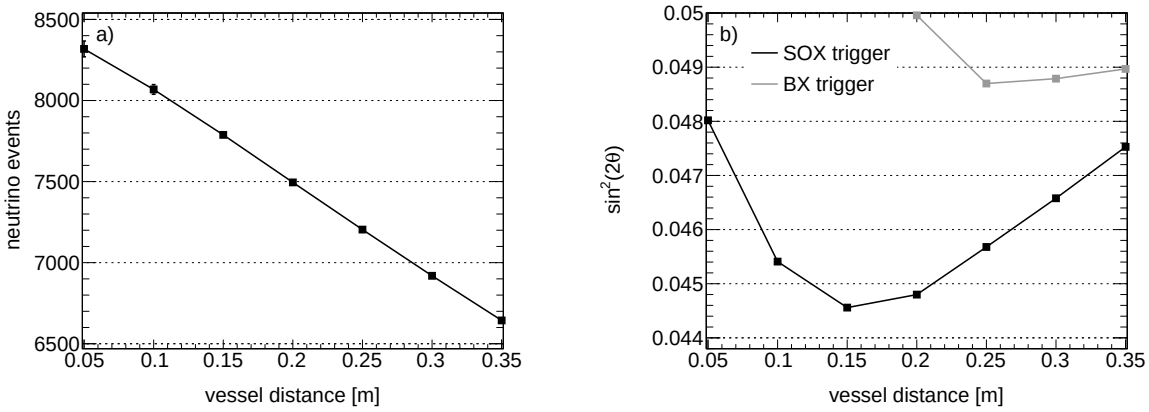


Figure 8.11: Sensitivity with the SOX Trigger The expected neutrino event number (a) and sensitivity (b) with the SOX trigger are shown as a function of the FV cut. The error bars shown for the event number take into account the uncertainty on the vessel shape and are smaller than the marker size for a FV larger than 0.15 m. The sensitivity with the SOX trigger increases in comparison to the one with the standard Borexino trigger (from Figure 7.9) and the FV cut can be enlarged to 15 cm. The sensitivity is computed for a rate analysis as described in Appendix A.5.

At first, a brief explanation of the Borexino trigger system is given, which is illustrated in Figure 8.12. The Borexino trigger is raised when a minimum number of hits ($tt1$-threshold) are detected within 100 ns[6]. The $tt1$-threshold is currently set to 20 hits. In praxis, each hit PMT creates a logical signal with a length of 100 ns. The sum of these logical signals is updated every 33 ns by the 30 mHz so-called sum clock. The trigger clock (60 MHz) evaluates every 16 ns the sum and issues a trigger if the sum is larger than the $tt1$-threshold. At this time the sum exceeds the $tt1$-threshold, but, due to the time spread of the arriving scintillation photons of several tens of nanoseconds, not necessarily the $tt4$-threshold (see black line in Figure 8.12). In order to achieve a full $tt4$-efficiency, the decision whether the event is associated with $tt1$ or $tt4$ and in addition to issue the following $tt16$-gate is delayed for several trigger clock cycles. A maximum delay time (in multiples of trigger clock cycles) is defined up to which at each clock cycle the sum is compared to the $tt4$-threshold. If the $tt4$-threshold is crossed, the event is assigned as $tt4$-event and the following $tt16$-gate is issued. If the sum remains below the $tt4$-threshold, the event is saved as $tt1$.

The optimum value of the maximum delay time is studied by varying its value and determining the $tt4$-efficiency using physical Borexino data and laser pulses. Here, the pulsed laser system is used that is installed in Borexino for the time calibration of the PMTs, where sub-ns 394 nm laser pulses are sent via optical fibers to each PMT [141] and produce a Gaussian energy spectrum with a mean of $\sim$ 70 hits (see Figure 8.14)[7]. Figure 8.13 shows the $tt4$-efficiency for a delay time between two and five clock cycles (33 ns - 83 ns) for several $tt4$-thresholds. The $tt4$-efficiency starts to rise at the set $tt4$-threshold. The larger the delay time the steeper the efficiency curve becomes and a full efficiency is reached for smaller energies. For example, in a physical run for a $tt4$-threshold of 50, a full efficiency is obtained at an energy of 70 hits for a delay time of 5 clock cycles in comparison to 80 hits for 3 clock cycles. Indeed, a delay of two clock cycles is too short to obtain full efficiencies. In comparison to the physical runs, the laser runs have a steeper efficiency transition curve due to the narrower

[6] In addition, the Borexino trigger is issued, when the threshold of the outer detector is crossed (see Appendix A.2).

[7] The following figures are shown at the end of this chapter.

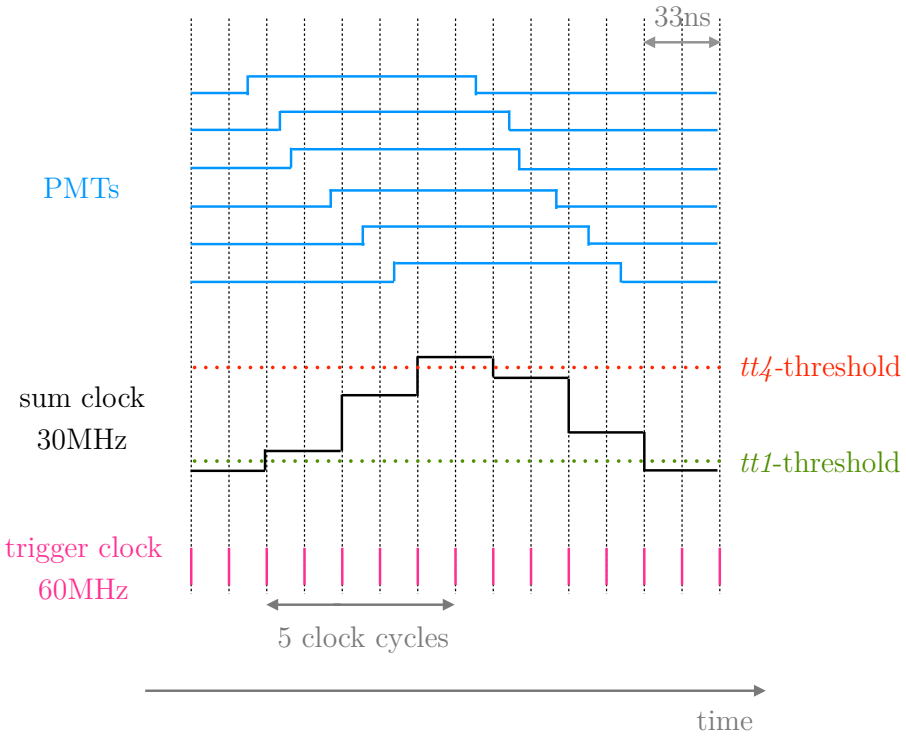


Figure 8.12: Sketch of the Borexino Trigger Logic Each hit PMT generates a 100 ns long logical signal. The sum of these channels is updated every 33 ns by the sum clock. The trigger clock evaluates the sum every 16 ns and issues a 16 µs gate when the sum exceeds the *tt*1-threshold. Whether the event is saved as *tt*1 or *tt*4 is decided with a delay of maximal five trigger clock cycles, i.e. 83 ns.

time spread of the signal. In conclusion, the delays of four and five clock cycles produce similar results and the optimized delay time is set to five clock cycles.

8.3.3 Detection Efficiency

Using the pulsed laser system, the detection efficiency of the laser pulses in the *tt*4- and *tt*16-gates is investigated. First, the detection efficiency in the *tt*4-gate is studied by the fraction of all sent laser pulses that are detected either in a *tt*1- or in a *tt*4-event. Given a high efficiency in observing the sum of *tt*1- and *tt*4-events, the efficiency of the *tt*4-events must be high as well. The number of sent pulses is determined from the known frequency of the pulser (100 Hz). The detection efficiency as a function of the *tt*4-threshold is listed in Table 8.1. For all thresholds, the efficiency reaches a sufficiently high level of at least 99.88 %. The small amount of unobserved pulses could be due to laser pulses that accidentally fall into other trigger gates. For example, the *tt*128-trigger or the service triggers (see Appendix A.2), which are not taken into account in the efficiency computation.
Second, the detection efficiency of events in the *tt*16-gate is determined with double-laser pulses, where the time difference between the two pulses is Δt. The efficiency is defined as the fraction of observed double-pulse events with a detected second pulse and is given in Table 8.1 for varying Δt. The efficiency is larger than 99.96 % for all time differences besides $\Delta t = 16\,\mu s$ (99.93 %) that lies in the transition region of the *tt*4- and the *tt*16-gate. The energy spectra from the laser pulses reconstructed in the *tt*4- and *tt*16-gate agree well with

Table 8.1: Detection Efficiency of Laser Pulses in the $tt4$- and $tt16$-Gate left: the $tt4$-efficiency as a function of the $tt4$-threshold. right: efficiency to detect the second laser pulse of double pulses as a function of the time difference between the two pulses (Δt). The second pulse appears in the $tt4$-gate for $\Delta t < 16\,\mu s$, and in the $tt16$-gate otherwise. In all tests, the delay time is set to five clock cycles and the $tt1$-threshold to 20.

$tt4$-threshold	efficiency [%]
41	99.91 ± 0.02
51	99.93 ± 0.02
61	99.96 ± 0.01
71	99.88 ± 0.01

Δt [µs]	efficiency [%]
13	$99.98^{+0.02}_{-0.05}$
14	$99.98^{+0.02}_{-0.05}$
15	$100.00^{+0.00}_{-0.04}$
16	$99.93^{+0.07}_{-0.06}$
17	$99.98^{+0.02}_{-0.05}$
18	$99.96^{+0.04}_{-0.06}$
19	$99.98^{+0.02}_{-0.05}$
20	$100.00^{+0.00}_{-0.05}$
50	$100.00^{+0.00}_{-0.04}$
100	$100.00^{+0.00}_{-0.04}$
200	$100.00^{+0.00}_{-0.05}$
400	$100.00^{+0.00}_{-0.04}$
800	$99.96^{+0.04}_{-0.05}$

each other and are shown in Figure 8.14.
In summary, the results from the laser tests indicate a proper functionality of the SOX trigger.

8.3.4 Hit-Time Distributions in the Trigger Gates

Figure 8.15a shows the cumulative hit-time distribution in the $tt4$- and $tt16$-gates from all events of a physical run. A significantly larger dark noise level is present in the $tt4$-gate than in the $tt16$-gate, which produces a step in the hit-time distribution at 16 µs. One can observe that (due to an unresolved problem) the dark noise level is increased in events that follow after a $tt16$- or $tt128$-event. Figure 8.16 shows the total number of all hits recorded in the gate for $tt1$- or $tt4$-events as a function of the time difference to the previous event. Two distinct structures can be observed: while one distribution shows a time-constant energy response, for the other one the number of reconstructed hits in one gate increases linearly with the time to the previous event. The latter distribution can be traced back to those events that follow after a $tt16$- or $tt128$-event. Excluding those events from the data, the higher dark noise level in the $tt4$-events vanishes and the transition in the hit-time distribution from the $tt4$- to $tt16$-gate is smooth (see 8.15b).
As the reason for the increased dark noise level after $tt16$/$tt128$-events could not be explained, a fake event with 16 µs length that automatically follows after each $tt16$/$tt128$-event is introduced. This fake event is not used for the data analysis, but resets the dark noise level and ensures that the following events are unaffected by a larger dark noise level. The associated trigger type is $tt8$. Figure 8.9 sketches the full sequence that is implemented in the SOX trigger.

The hit-time distribution recorded in the three gates (*tt*4, *tt*16, and *tt*8) can be seen in Figure 8.17. The dark noise level in *tt*4 and *tt*16 agree well with each other and are both slightly smaller than the one in *tt*8 (which is expected). The *tt*8-gate appears 50 µs after the end of the *tt*16-gate, even though the time difference is set to ten trigger clock cycles, i.e. ~ 160 ns. This indicates that the intrinsic dead time after a *tt*16/*tt*128-event is 50 µs. The dead time in between the *tt*8- and the next *tt*1- or *tt*4-event is negligibly small, which can be seen in Figure 8.18. Hence, after a total time of ~ 1.7 ms after a *tt*4-event the next *tt*1- or *tt*4-event can be recorded.
The dead time between the *tt*4- and the *tt*16-gate is shown in Figure 8.19 and is ~ 200 ns. In comparison, the dead time of the Borexino trigger after a *tt*1-event is ~ 3 µs (see Chapter 5). Having a minimized dead time increases the detection efficiency of the neutrons and increases hence the statistics of the SOX event sample and the sensitivity of the SOX experiment.

8.3.5 Optimization of the SOX-Trigger Threshold

The *tt*4-threshold is optimized in order to increase the signal-to-background ratio for the SOX event sample. Figure 8.20 shows the fraction of the total spectrum above the threshold of SOX and Borexino events, respectively. The spectrum of the SOX events is obtained from MC simulations, where the SOX events are generated within the IV. First events are visible at 100 hits. The Borexino spectrum is reconstructed from Borexino data in October 2017. The main contribution of the Borexino spectrum comes from the intrinsic ^{14}C background that reaches up to ~ 50 hits. Above 50 hits the fraction of Borexino events remains less than 1 % and decreases only slightly. Hence, the signal-to-background ratio stays almost constant above 50 hits. To remove, on the one hand, the full ^{14}C background and to avoid, on the other hand, the rejection of any signal events, the optimized *tt*4-threshold is set to 65.
With a Borexino rate of ~ 19 Hz (October 2017) and the *tt*4-threshold of 65, a *tt*4-rate of $7 \cdot 10^{-3} \cdot 19\,\text{Hz} = 0.13\,\text{Hz}$ is expected due to Borexino events. The rate of SOX events that accidentally fall into a *tt*4/*tt*16-event that is issued by a standard Borexino event is at the beginning of data taking: $0.13\,\text{Hz} \times 36/\text{day} \times 1.6\,\text{ms} = 7.5 \cdot 10^{-3}/\text{day}$. Here, 36/day is assumed to be the expected SOX rate at the beginning of data taking for a 125 kCi source. For the total measurement time of 1.5 years, 2.3 accidental *tt*4-events are expected. Similarly, the accidental coincidence rate of a SOX event in a *tt*1-gate can be computed: $18.87\,\text{Hz} \times 36/\text{day} \times 16\,\mu\text{s} = 0.01/\text{day}$ and 3.3 events in 1.5 years. Both rates are negligible in comparison to $\sim 10^4$ expected SOX events. In addition, there is a possibility to reconstruct these events, which depends on how the prompt and the delayed event fall relatively into these gates.

8.3.6 Final Configuration of the SOX Trigger and Background

With the final configuration of a *tt*4-threshold of 65, a delay time of five trigger clock cycles, and the fake *tt*8-event, physical Borexino data is taken. The resulting *tt*4-efficiency can be seen in Figure 8.21, where a full efficiency is obtained for 110 clustered hits.
The *tt*4-rate is measured to 0.11 Hz that agrees well with the prediction from the previous section. The total rate and energy spectrum of the runs with the final SOX trigger configuration are compared to the ones measured with the standard Borexino trigger without the SOX modification. An unresolved discrepancy in the rate and the spectrum at low energy events is observed. The average rate of *tt*1-events in a standard Borexino run (averaged over 5 runs in November 2017) is (18.85 ± 0.01) Hz. Instead, the rate of the sum of *tt*1- and *tt*4-events in the SOX trigger runs averages to only (17.81 ± 0.02) Hz, which is a deficit of 5.5 %. The observed difference cannot be explained by the fact that the effective measurement time in the SOX trigger runs is reduced by 0.018 % which is given by the dead time introduced by

the $tt16$-gate length times the $tt4$-rate. The comparison of the total energy spectra for $tt1$- and $tt4$-events is shown in Figure 8.22. The deviation is especially visible in the low energy region below 50 hits. The reason for the discrepancy has not been understood and due to the cancelation of the SOX project has not been further investigated.
The data taken with the final configuration and a total measurement time of 24.15 h is investigated for IBD-like events. This short measurement time can only provide preliminary results and a longer investigation would be necessary for the SOX analysis. For the optimized selection cuts (see Table 5.1), a FV cut of 15 cm, and without a minimum energy cut on the delayed event, seven events are found. The energy and the position of the prompt and the delayed events are shown in Figure 8.23. All events are reconstructed close to the vessel border and are likely either accidental coincidences or fast coincidences from the ^{238}U and ^{232}Th chain (see Chaper 6). It has to be noted that for all seven events, the pulse shape discrimination algorithm fails and the improvement of the pulse shape discrimination at such low energies would be mandatory to remove background events. Given more statistics, the selection cuts could be optimized especially for the region close to the vessel border to increase the signal-to-background ratio. Even though the background at the vessel border might be high, it can be precisely determined before data taking. In case of accidental coincidences, the background can also be constrained during data taking via an off-time window measurement (see Chapter 6).

In summary, in this chapter, the feasibility of using the events interacting in the buffer in order to increase the sensitivity of the SOX experiment was investigated. The most realistic option is the installation of a delayed neutron trigger gate to improve the detection efficiency of IBD reactions close to the vessel border. Such a modification of the Borexino trigger could increase the sensitivity by about 9 % in terms of $\sin^2(2\theta)$ and was extensively and successfully studied. However, an unresolved discrepancy in the event number and the energy spectrum in comparison to the nominal Borexino values would have required further investigations before implementation.

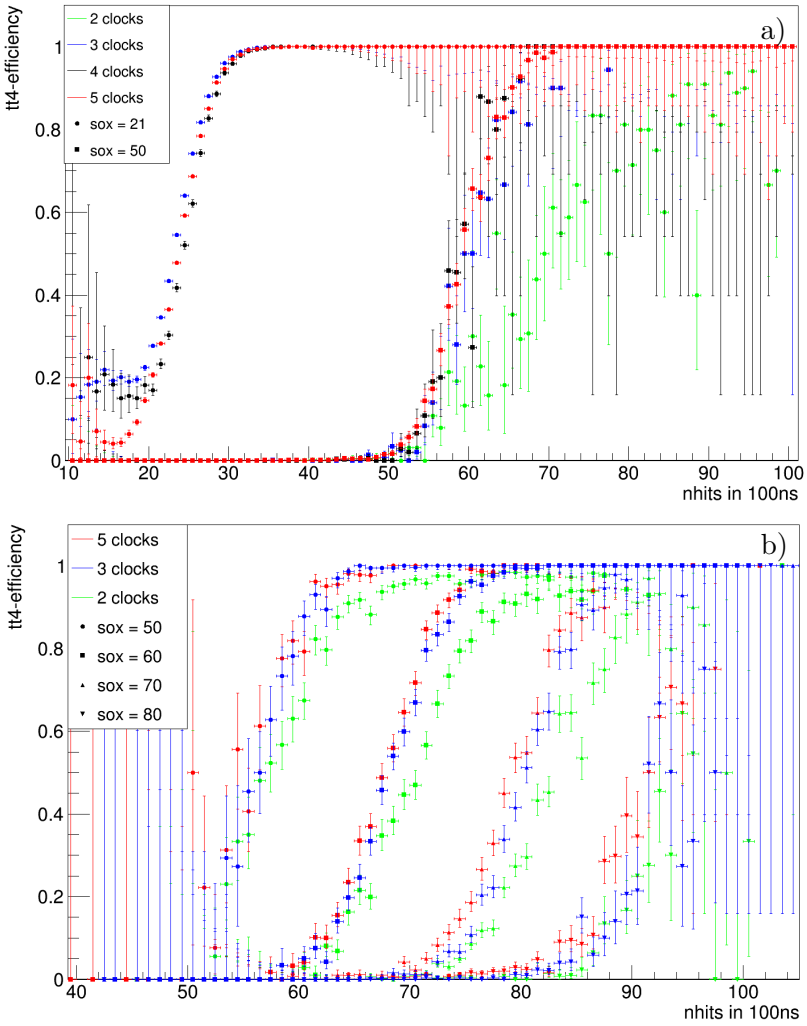


Figure 8.13: *tt*4-Efficiency as a Function of the Energy The efficiency is shown for physical events (a) and laser pulses (b) for the delay times of two to five clock cycles and several SOX (*tt*4) trigger thresholds. As the laser events are sharper in time in comparison to the physical events, the full efficiency is reached for lower energies. The *tt*1-threshold is for all settings the nominal one of 20. The shown errors are statistical errors.

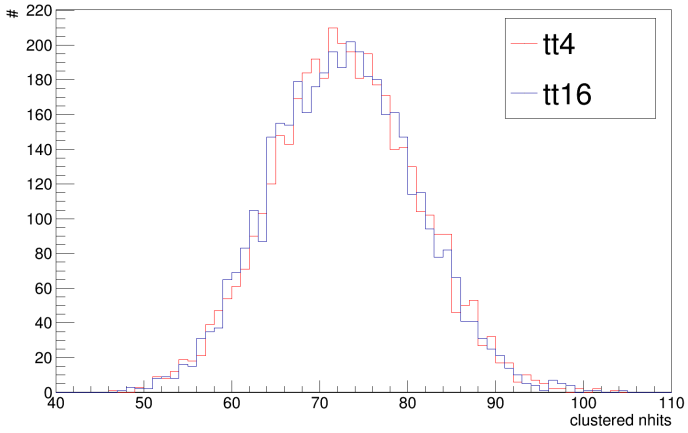


Figure 8.14: Energy Spectrum of Laser Pulses The energy spectrum of the laser pulses reconstructed in a *tt*4- and *tt*16-gate are shown in red and blue, respectively.

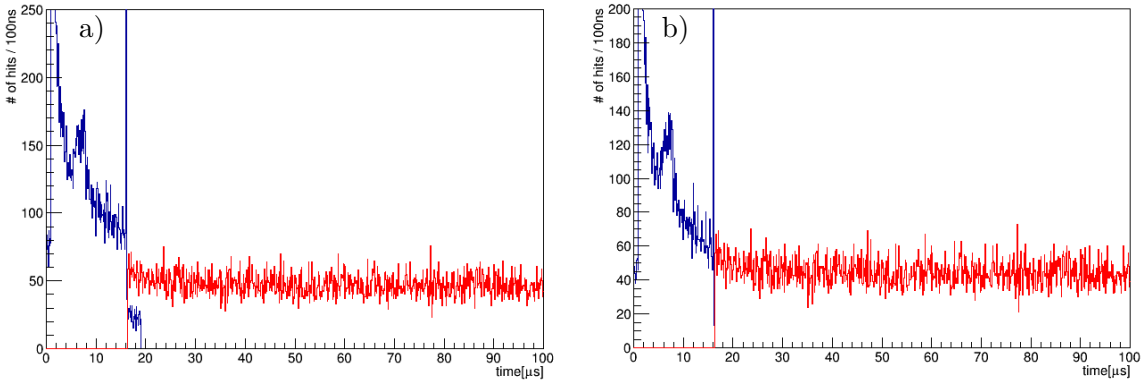


Figure 8.15: Hit-Time Distribution in the $tt4$- and $tt16$-Gate The cumulative hit time distribution of $tt4$- (blue) and the following $tt16$- (red) events of a physical run are shown for a) all events b) all events not preceded by a $tt16/tt128$-event. Due to an unsolved issue, $tt4$-events after a $tt16/tt128$-event have an increased dark noise level which is present up to $\sim 19.5\,\mu s$. The peak at $16\,\mu s$ corresponds to the trigger signal.

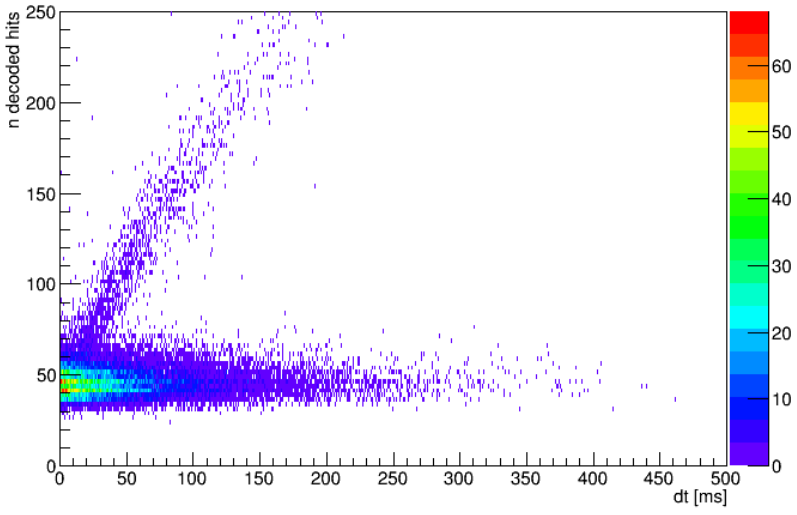


Figure 8.16: Dark Noise Level after $tt16/tt128$-Events The total number of hits in the $tt4$-gate is shown as a function of the time difference to the preceding event for a physical run. The linearly-increasing distribution can be traced back to those events that follow after a $tt16/tt128$-event.

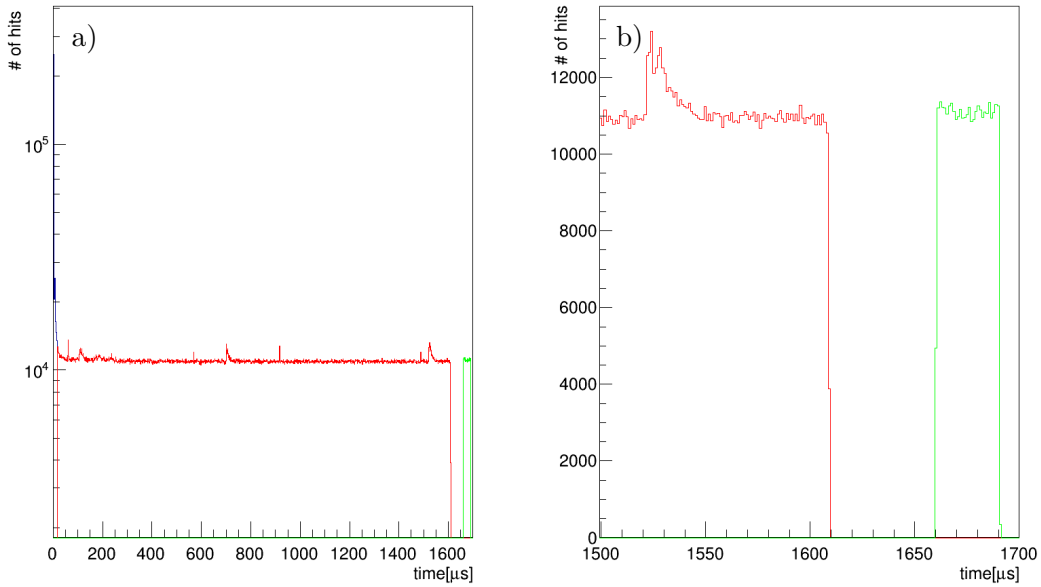


Figure 8.17: Hit-Time Distribution of the full SOX Trigger-Gate Sequence The cumulative hit time distribution is shown for $tt4$- (blue), $tt16$- (red), and $tt8$- (green) events of a physical run. The latter has been introduced to reset the dark noise level. The dead time after $tt16$-events is 50 µs long.

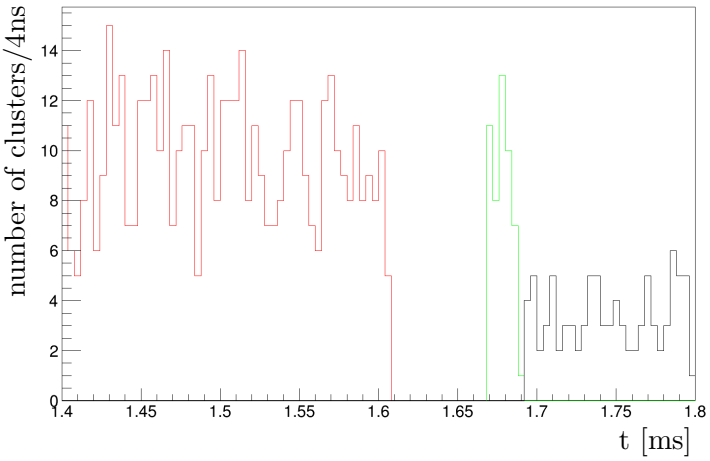


Figure 8.18: Dead Time after $tt16$-Events The time difference of the next physical event (cluster) to the previous $tt4$-event is shown. The events in the $tt16$, $tt8$, $tt1/tt4$-gate are indicated in red, green, and black, respectively. The total dead time after a $tt16$-gate results into ~ 0.1 ms. The event number is reduced in the black distribution, as events below the trigger threshold are not recorded.

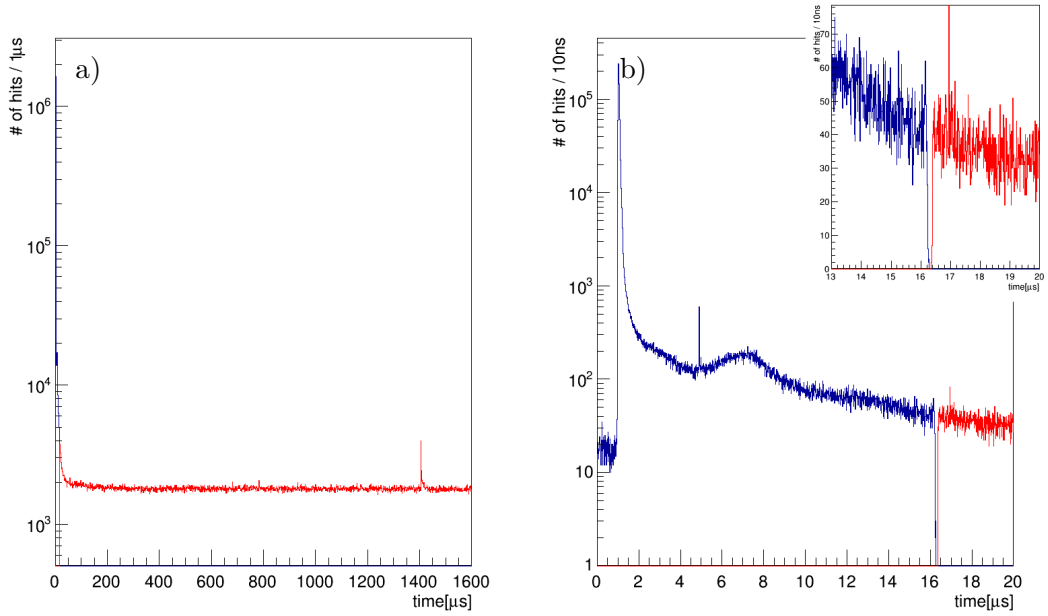


Figure 8.19: Hit-Time Distribution of $tt4$- and $tt16$-Events for the Final Configuration of the SOX Trigger The cumulate hit-times in the $tt4$-gate and the $tt16$-gate are shown in blue and red, respectively. Panel b) shows a zoom in the transition region of the $tt4$- and $tt16$-gates, where the dead time after the $tt4$-events of about 200 ns is visible.

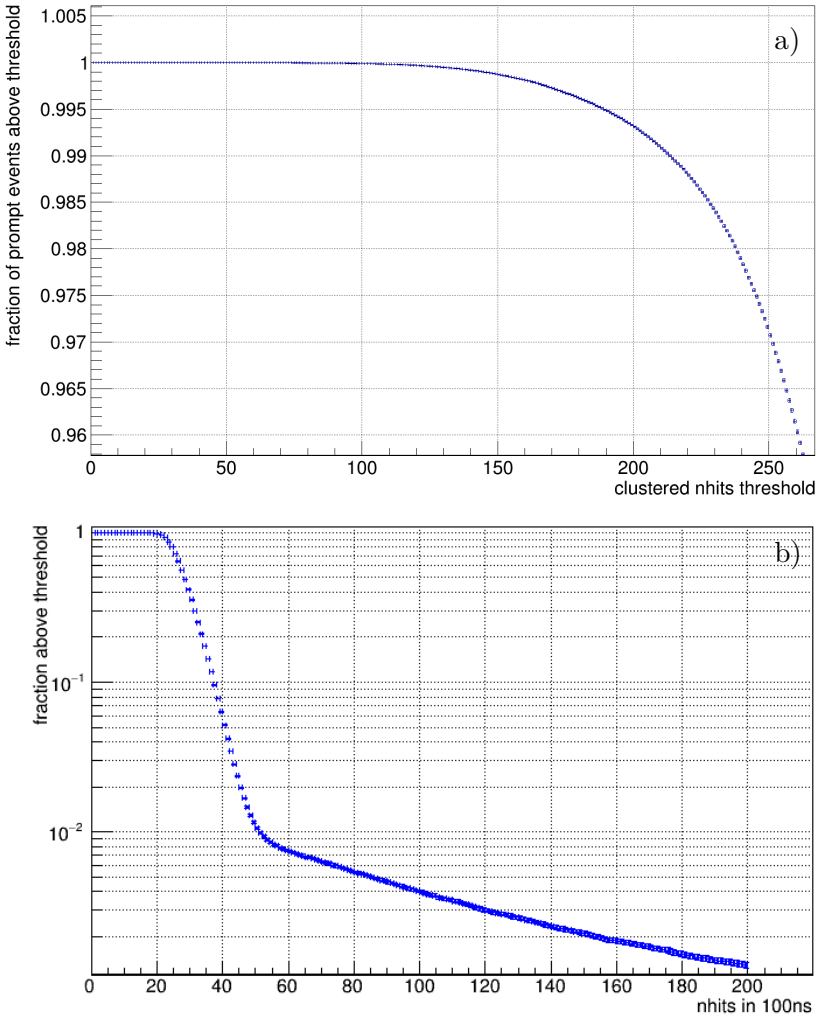


Figure 8.20: Fraction of Spectrum above the SOX Trigger Threshold The fraction is shown for SOX signal events (a) and background events (b). The function for the signal events is reconstructed from MC simulations with SOX events generated in the IV. The function for the background is obtained from Borexino data with a *tt*1-threshold of 20. The optimized *tt*4-threshold is set to 65, in order to be well above the ^{14}C background (below $\sim$ 50 hits) and be able to detect the low energy SOX events. Note that the energy estimator differs slightly. While in the top panel the energy is given by the hits reconstructed in the full cluster (1.5 µs), the energy in the bottom panel is given by the hits reconstructed within 100 ns.

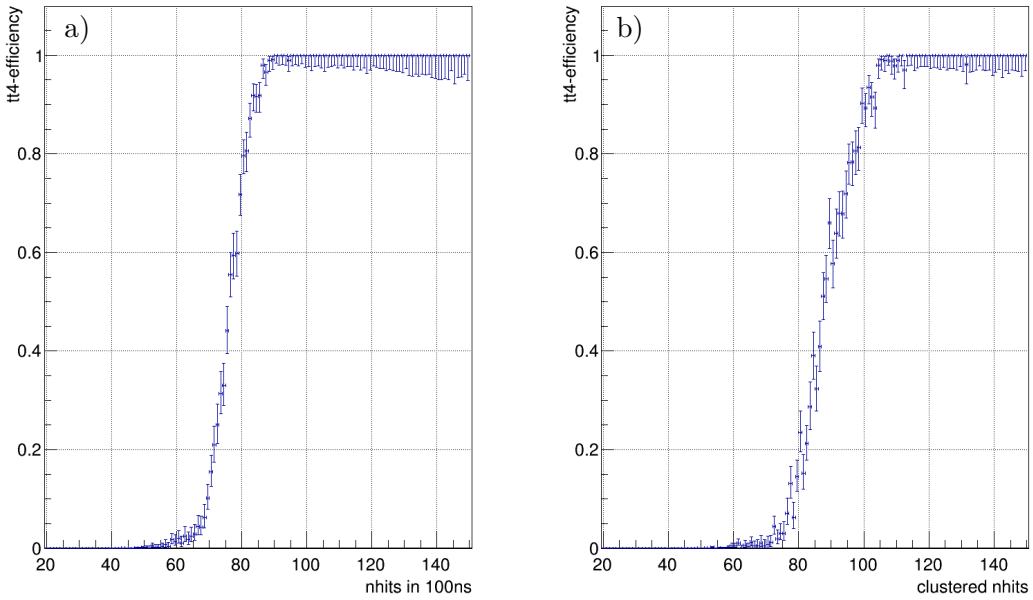


Figure 8.21: $tt4$-Efficiency for the Final Configuration of the SOX Trigger The efficiency is computed as a function of the number of hits reconstructed in a time window of 100 ns (a) and 1.5 µs (b), i.e. clustered nhits. The efficiency is obtained from one physical Borexino data run ($\sim$ 6 h). The error bars indicate the statistical error.

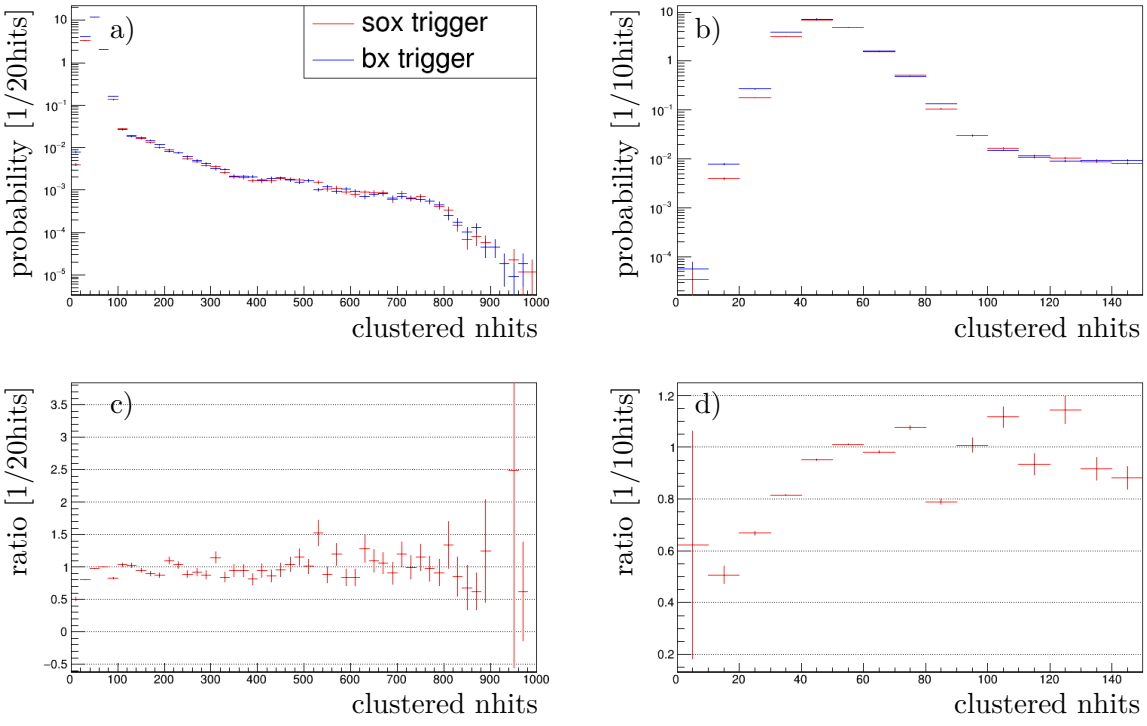


Figure 8.22: Comparison of the Energy Spectra Measured with the Borexino and the SOX Trigger The top panels show the energy spectra normalized to the measurement time for the full energy range (a) and the low energy region (b). The comparison is based on $tt1$ (Borexino trigger) and $tt1$+$tt4$ (SOX trigger) events with one cluster and no signal in the outer detector. Panel c) and d) show the ratio of the energy spectra of the SOX and the Borexino trigger. A significant deficit in the spectrum obtained with the SOX trigger is visible for energies smaller than 50 hits. The error bars account for the statistical uncertainties. Notice the different ranges of the y-axis in the bottom panels.

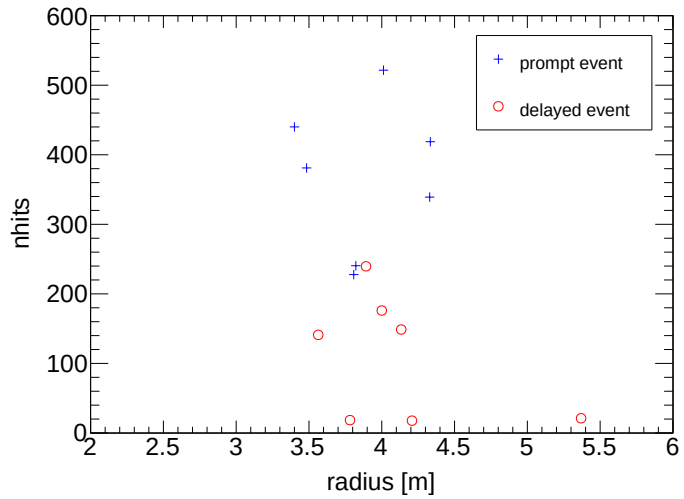


Figure 8.23: Energy and Position of Background Events with the SOX Trigger The energy (in clustered nhits) is shown as a function of the reconstructed radius for the prompt and the delayed event. The IBD-like events are selected with the optimized selection cuts, but without a minimum energy cut on the delayed event and with a FV cut of 15 cm. The data is taken with the final SOX trigger configuration with a measurement time of 24.15 h. Note that the pulse shape discrimination algorithm does currently fail for such low energies of the delayed event.

Part II

Statistical Methods and Issues in Sterile Neutrino Experiments

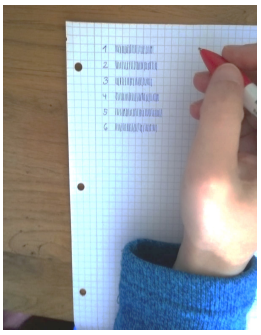
A large fraction of the following part was developed in collaboration with Matteo Agostini and was published in Reference [208].

Is my die fair?

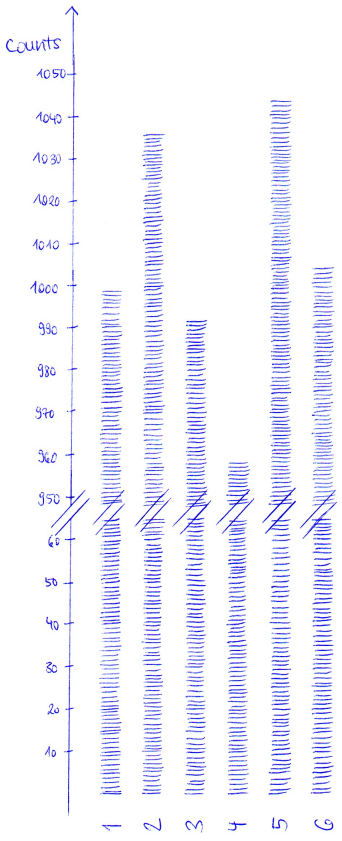
In order to check whether your die is fair you have to:

1. Roll the dice a lot of times while playing for example ''Mensch ärgere Dich nicht''.

2. Count the results.



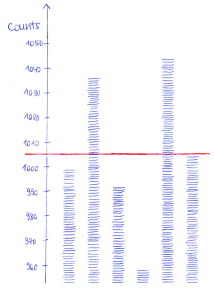
3. Have a look at your data.



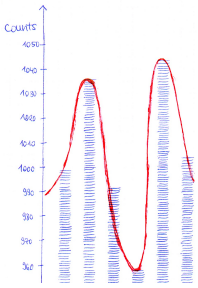
4. Choose a fit model and make a decision.

Depending on the fit model, the die is

fair



unfair



→ Given a flexible enough fit model, the fit will always return something different from a flat line.

As a consequence, you will almost never decide for a fair die, even though it actually is fair and you roll the dice really a lot of times.

In analogy, in the sterile neutrino search the signal of **no sterile neutrinos** corresponds to a **flat line**. Moreover, a **sterile neutrino signal** is represented by an **oscillation signature**, which is a **very flexible fit model**. Thus, the best-fit value in a sterile neutrino experiment is typically given by a sterile neutrino hypothesis, even though the no-sterile hypothesis might be true! Such best fit values are called **biased**.

9 Statistical Issues of Short-Baseline Experiments

As introduced in Chapter 1 and 3, a large number of experiments are currently looking for light sterile neutrinos. In the following chapter, the common statistical properties of these sterile neutrino experiments are studied. The focus is set on sterile neutrino experiments that assume a (3+1) model and search for an oscillation signature in the neutrino spectrum as a function of distance and energy. In particular, those experiments are considered for which the oscillation probability can be approximated with the two-flavor oscillation probability. The statistical issues are studied with two toy experiments and a toy model that are introduced in Section 9.1. Further, the sensitivities, the confidence regions, and the maximum likelihood estimators in a rate, shape, and rate+shape analysis are discussed in Section 9.2.

9.1 Toy Experiments and Toy Model

Short-baseline experiments look for a signature that is defined by the oscillation probability (see Section 1.5.5). In particular, the appearance probability that a muon neutrino converts to an electron neutrino is given by [116]:

$$P(\nu_\mu \to \nu_e) = \sin^2(2\theta_{\mu e}) \sin^2(k \cdot \Delta m^2 \cdot L/E), \tag{9.1}$$

where $k = 1.27\,\mathrm{MeV}/(\mathrm{eV}^2{\cdot}\mathrm{m})$. Analogously, the survival probability of an electron neutrino is given by [116]:

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta_{ee}) \sin^2(k \cdot \Delta m^2 \cdot L/E), \tag{9.2}$$

with the oscillation amplitudes $\sin^2 2\theta_{ee}$ and $\sin^2 2\theta_{\mu e}$ for the respective experiment. Both probabilities contain the same oscillation term: $\sin^2(k \cdot \Delta m^2 \cdot L/E)$ which produces an oscillation pattern in L/E and is shown in Figure 9.1a. As the Δm^2-value is correlated to the oscillation frequency, the oscillation length (period) becomes smaller for larger Δm^2-values. Due to this common term, the same statistical issues for both types of experiments are expected. In the following, a toy appearance and a toy disappearance experiment are defined with parameters representative of the currently running experiments (see Table 1.2). The parameters of the two toy experiments are summarized in Table 9.1.
The toy appearance experiment is designed to resemble the current MiniBooNE experiment [120]. Hence, as neutrino source a muon-neutrino beam is assumed that can lead to an excess of electron neutrino events through muon to electron neutrino oscillations, described by Equation 9.1. The reconstructed distance ranges from 500 m to 550 m, where the uncertainty is mainly given by the length of the decay region, in which the muon neutrinos are produced. For convenience, the energy spectrum is chosen to be flat from 200 MeV to 2000 MeV with a resolution of $\sigma_{E_{\mathrm{rec}}}$=10 MeV. The neutrino energy can be typically reconstructed using the kinematics of the scattering process and the energy and angle of the outgoing particle that

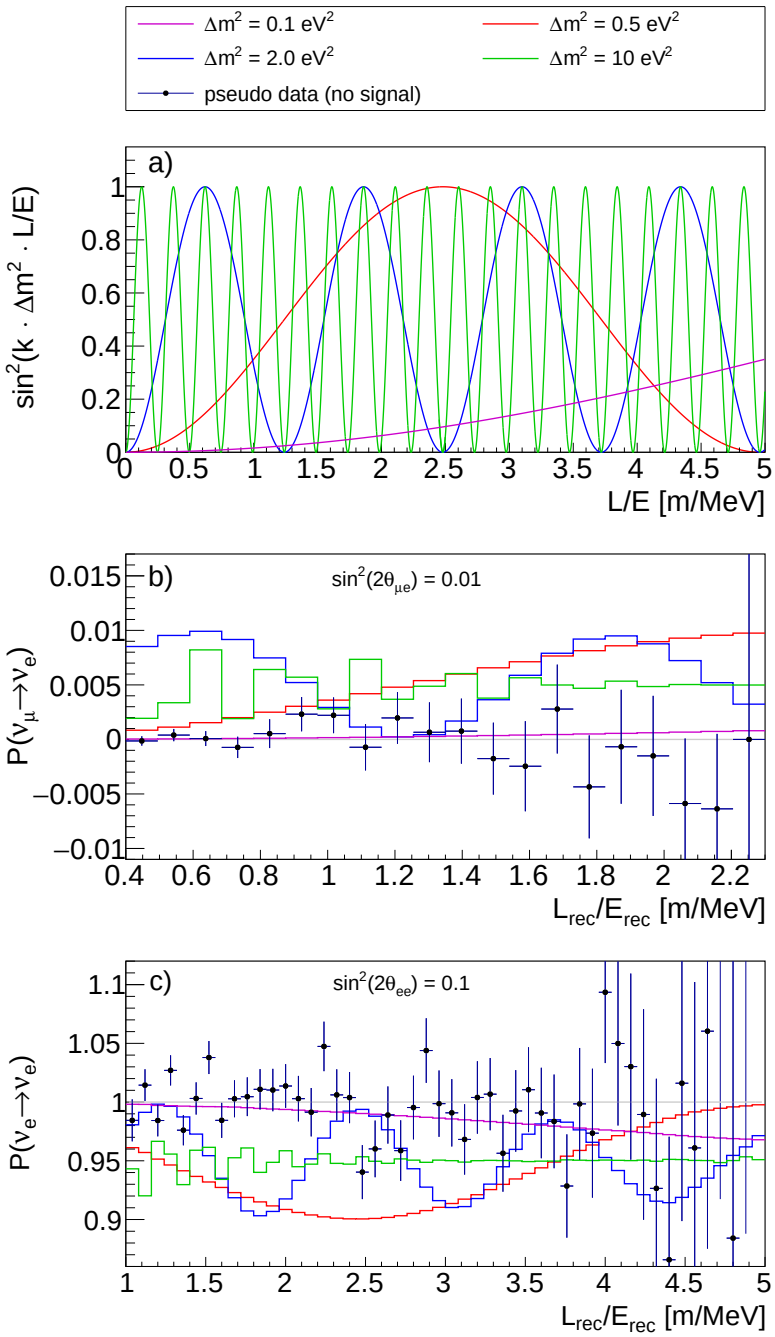


Figure 9.1: Signature of the Toy Experiments a) The oscillation term which is present in both oscillation probabilities is shown for different Δm^2-values. b) and c) The expected reconstructed oscillation probabilities for the toy appearance and toy disappearance experiment is shown for different Δm^2-values. The reconstructed oscillation probability is defined as the ratio of the difference of observed data and the expectation for the background over the expected data for the no-oscillation hypothesis. In addition, the reconstructed oscillation probability is shown for one pseudo-data sample from the no-oscillation hypothesis, where the error bars account for the statistical uncertainties before background subtraction.

is measured from Cherenkov and scintillation light. Thus, the $L_{\text{rec}}/E_{\text{rec}}$-range varies from 0.4-2.4 m/MeV with a resolution between 10 and 25 %. A total expected electron neutrino event number of 10^5 with an uncertainty of 10 % is expected. Here, the event number cor-

Table 9.1: Parameters of the Toy Appearance and Toy Disappearance Experiment The ranges of reconstructed distance (L_{rec}) and reconstructed energy (E_{rec}) are given along with their uncertainties ($\sigma_{L_{\text{rec}}}$, $\sigma_{E_{\text{rec}}}$). Further, the binning refers to the number of energy and distance bins in the data. The expected event numbers for the total neutrino event number (N_S^{exp}) and the background event number (N_B^{exp}) are quoted with the uncertainties (σ_{N_S}, σ_{N_B}), where N_S^{exp} corresponds to the event number for an oscillation probability of one.

parameter	unit	appearance	disappearance
ν - oscillation		$\nu_\mu \to \nu_e$	$\nu_e \to \nu_e$
L_{rec}	m	$500 - 550$	$7 - 10$
$\sigma_{L_{\text{rec}}}$	m	50	0.5
E_{rec}	MeV	$200 - 1200$	$2 - 7$
$\sigma_{E_{\text{rec}}}$	MeV	10	$0.05 \cdot \sqrt{E/(1\,\text{MeV})}$
$L_{\text{rec}}/E_{\text{rec}}$	m/MeV	$0.4 - 2.4$	$1 - 5$
$\sigma_{L_{\text{rec}}/E_{\text{rec}}}$	%	$10 - 25$	$5 - 10$
binning	$L_{\text{rec}} \times E_{\text{rec}}$	1×20	6×10
N_S^{exp}		10^5	10^5
σ_{N_S}	%	10	2
N_B^{exp}		10^3	10^4
σ_{N_B}	%	5	2

responds to the maximum possible event number for an appearance probability of one. The large uncertainty on the event number comes from several sources including the neutrino flux prediction, neutrino cross sections, uncertainties from nuclear effects, and the detector response. Further, the background is also assumed to be flat with an event number of 10^4 and an uncertainty of 5 %. As the traveled distance is larger than the detector size, the data is only binned in the energy observable.

The toy disappearance experiment is similar to a reactor- or source-based experiment that is looking for an electron neutrino disappearance oscillation (see Equation 9.2). The distance range is chosen to be 7-10 m with a resolution of 0.5 m, where the spectrum is proportional to $1/L^2$. The energy spectrum is flat from 2 MeV to 7 MeV with a resolution of 5 % at 1 MeV. Hence, the L/E-range spans from 1-5 m/MeV with a resolution between 5 to 10 %. As the detector is close to the source and the detector is capable to resolve the energy and the position (either by time-of-flight techniques or by the segmentation of the detector), the data is grouped into six L and ten E bins. The expected event number for a survival probability of one (and hence the no-sterile hypothesis) is 10^5 with an uncertainty of 2 %. The background event number is assumed to be 10^4 with 2 % uncertainty.

Figure 9.1b and c depict the expected reconstructed oscillation probability for the two toy experiments for several Δm^2-values. The signature clearly depends on the Δm^2-value:

(a) $\Delta m^2 < 0.1\,\text{eV}^2 \to (L/E)_{\text{osc}} > 10\,\text{m/MeV}$: The oscillation length is larger than the reconstructed L/E-range and can only partially be observed.

(b) $0.1\,\text{eV}^2 < \Delta m^2 < 10\,\text{eV}^2 \to (L/E)_{\text{osc}} \sim 1\,\text{m/MeV}$: The oscillation length is smaller than the reconstructed L/E-range and can be observed within the detector.

(c) $\Delta m^2 > 10\,\text{eV}^2 \to (L/E)_{\text{osc}} > 0.1\,\text{m/MeV}$: The oscillation length is smaller than the detector resolution. An overall increased (reduced) event number can be observed in an appearance (disappearance) experiment.

In addition, the reconstructed signature from a pseudo-data set from the no-sterile hypothesis is shown with the statistical error bars. From that signature, one can observe that statistical fluctuations in the data sample could possibly mimic a sterile neutrino signature. This will be discussed in more detail in this chapter and in Chapter 11.
In the following, three different kinds of analyses are performed: the rate, shape and rate+shape analysis. The likelihood function in a rate+shape analysis is defined by:

$$\mathcal{L}(\sin^2(2\theta), \Delta m^2, N_S, N_B) = \prod_{ij} \mathcal{P}(N_{ij}^{\text{obs}} | N_{ij}^{\text{exp}}(\sin^2(2\theta), \Delta m^2, N_S, N_B)) \cdot \mathcal{N}(N_S^{\text{obs}} | N_S, \sigma_S^2) \cdot \mathcal{N}(N_B^{\text{obs}} | N_B, \sigma_B^2), \tag{9.3}$$

where the notation follows the one introduced in Chapter 3. For a shape analysis the normal distribution $\mathcal{N}(N_S^{\text{obs}} | N_S, \sigma_S^2)$ (the constraint on the neutrino event number) is dropped in the likelihood function. Further, i and j are set to one in the likelihood function for a rate analysis and the data has hence only one bin in energy and distance.
For the test statistic, the profile likelihood ratio as defined in Equation 3.9 is used. The probability distributions of the profile likelihood ratio are constructed via the Monte Carlo (MC) construction introduced in Chapter 3.2.3. Details on the minimization algorithm of the negative log-likelihood function, can be found in Appendix A.4.
Note that the so-called "ratio - method" which is independent on the neutrino flux and therefore currently used by most of the reactor-based experiments is not considered in the following [128, 130]. However, the sensitivities of the ratio-method are expected to behave similar to the sensitivities in a shape analysis.

9.2 Sensitivities, Confidence Regions, and Maximum-Likelihood Estimators

As introduced in Chapter 3, one can distinguish between the exclusion and discovery sensitivity. The exclusion sensitivity delimits the parameter space that could be rejected if the no sterile hypothesis is true with a probability of at least 50 %. The parameter space on the right of the sensitivity is excluded and the parameter space on the left of the sensitivity corresponds to the median obtained confidence region when the no sterile hypothesis is true. Instead, the discovery sensitivity delimits the parameter space that could be discovered with 50 % probability if the sterile hypothesis is true. This means that for the hypotheses in the discovery region, if they are true, the no-signal hypothesis is rejected with a probability of at least 50 % (see Figure 3.4).
The sensitivities of the toy experiments are shown at the 95 % CL in Figure 9.2.[1] The appearance and disappearance experiment show similar behavior, which is summarized shortly in the following and discussed in detail in Section 9.2.1-9.2.3. As the signature depends on the Δm^2-value, the sensitivities can be grouped into the same three Δm^2-regions as defined previously. At $\Delta m^2 \sim 1\,\text{eV}^2$, the oscillation length can be resolved within the detector and the experiment is sensitive to the oscillation signature. This is the most sensitive Δm^2-region. For larger Δm^2-values the oscillation length becomes smaller than the detector resolution and only an overall change in the event number can be observed. At these Δm^2-values ($\Delta m^2 > 20\,\text{eV}^2$), $\sin^2(2\theta)$ and N_S become correlated and the sensitivity in the shape analysis is lost. Instead, the sensitivity in the rate analysis (i.e. when N_S is constrained) stays constant in terms of $\sin^2(2\theta)$ for large Δm^2-values. The sensitivity decreases for smaller Δm^2-values

[1] The following figures of this chapter can be found at the end of the chapter.

($\Delta m^2 \lesssim 0.1\,\mathrm{eV}^2$), as the oscillation length becomes larger than the reconstructed L/E range. In this Δm^2-region the change in the expected event number and hence the sensitivity is approximately proportional to the product $\sin^2(2\theta) \times \Delta m^2$.
The sensitivity in the appearance experiment is stronger in terms of $\sin^2(2\theta)$ than in the disappearance experiment. In the appearance experiment the no-sterile hypothesis is given by the background contribution only. Typically, the expected event number and the statistical uncertainties are hence smaller in an appearance experiment and the sensitivity is reached for smaller $\sin^2(2\theta)$-values. It should be emphasized that the mixing angles cannot be compared directly to each other (see Equation 1.12 for their connection).
In the following, the sensitivities obtained from the rate, the shape, and the combined rate + shape analysis are discussed in more detail. Additionally, the reconstructed confidence regions and maximum-likelihood estimator (MLE)-values from toy-data samples are investigated in the individual analyses.

9.2.1 Rate Analysis

A rate analysis is based on the integral event number independently from the spectral shape. Given an observed event number, any hypothesis with the same expected event number has the same probability to be true. Hence, the hypotheses at the sensitivity share all the same expected event number which deviates exactly by 2σ from the reference hypothesis. In turn, the likelihood function, which is shown in Figure 9.3a and b for pseudo-data samples of the toy appearance experiment[2], is constant for the set of hypotheses with the same expected event numbers. As a consequence, the reconstructed confidence regions are also restricted by these iso-event hypotheses. Such confidence regions are exemplarily depicted in black in Figure 9.4 for pseudo-data sets of the toy experiments under the no-signal and signal hypothesis, respectively. A signal in a rate analysis is represented by a confidence region that is a band in the $\sin^2(2\theta)$-Δm^2 parameter space (see black lines in Figure 9.4d). Since the data has only one bin, the oscillation parameters are correlated and only one parameter (the overall oscillation probability) can be deduced. Thus, the oscillation parameters $\sin^2(2\theta)$ and Δm^2 cannot be constrained at the same time and the analysis has only one degree of freedom.
In Figure 9.5a and b the MLE-values for pseudo data under the no-signal hypothesis are shown in black. Each point in the $\sin^2(2\theta)$-Δm^2 parameter plane stands for all hypotheses with the same oscillation probability. Hence for more visibility, the equivalent oscillation probability is shown in Figure 9.5c and d. On the one hand, the MLE-values are normally distributed around the true value in the physically allowed parameter space. The restriction of the parameter space, on the other hand, bounds the MLE-values to the value of the physical border. For instance, the no-signal hypothesis in a disappearance experiment has a survival probability of one ($P_{ee} = 1$). Given an observed event number larger than the expected one, the best-fit value results into the maximum possible oscillation probability. Therefore, due to statistical fluctuations, half of the pseudo-data samples have the no-signal hypothesis as the best-fit hypothesis. In Chapter 11 the effect of the distribution of the MLE-values on the distribution of the test statistic is discussed.
From the sensitivities in Figure 9.2 one observes, that there is a parameter space which can be on the one hand discovered but on the other hand not be excluded. This can be traced back to a boundary effect. In a signal discovery the tested hypothesis is the no-signal hypothesis, which lies directly on the edge of the allowed parameter space. The MLE-values are hence bounded as described above and the variance of the MLE is reduced on the boundary in

[2]The negative log-likelihood function of a disappearance experiment is shown in Figure 11.8.

comparison to hypotheses far from the boundary. This behavior is illustrated in Figure 9.6a, which shows the variance of the MLE in terms of the survival probability as a function of $\sin^2(2\theta)$ at Δm^2=1 eV^2. Testing the no-signal hypothesis with a smaller variance leads hence to better sensitivities. This is the reason why the discovery sensitivity is shifted to smaller $\sin^2(2\theta)$-values. Appendix A.5 shows how to compute the sensitivities for the rate analysis from the statistical and systematic uncertainties on the background and neutrino events only. The parameter bound also affects the power of the hypothesis tests in the rate analysis. The power is defined as the probability to reject the tested hypothesis when the alternative hypothesis is true (see Chapter 3). The exclusion power represents the probability to reject the signal hypothesis H($\sin^2(2\theta)$,Δm^2) when the no-signal hypothesis is true, while the discovery power gives the probability to reject the no-signal hypothesis when the signal hypothesis H($\sin^2(2\theta)$,Δm^2) is true. Figure 9.7a shows the exclusion and discovery power as a function of $\sin^2(2\theta)$ for $\Delta m^2 = 1\text{eV}^2$ for the toy disappearance experiment. In the discovery case, the tested hypothesis always stays the no-signal hypothesis and the power increases (as expected) for larger oscillation hypotheses. In the exclusion case instead, the power is minimal at $\sin^2(2\theta) \approx 0.15$ with a value smaller than the test size. Such a test, when the power becomes smaller than the test size, is called biased[3] [175]. Figure 9.7b visualizes the explanation of the reduction of the power using the distributions of the MLE-values of the oscillation probability for the no-oscillation hypothesis and the oscillation hypothesis close to the minimal power with $\sin^2(2\theta) = 0.015$ and $\Delta m^2 = 1\,\text{eV}^2$. Due to the parameter bound, the no-oscillation hypothesis and oscillation hypotheses close to the parameter bound can only be rejected in a disappearance experiment when the observed event number is smaller than the expected one. This behavior can be interpreted as a one-sided limit with one critical value. As the expected event number decreases for a signal hypothesis in a disappearance experiment, the critical value for the signal hypothesis decreases as well in comparison to the critical value of the no-signal hypothesis. Hence, the fraction of the expected event number under the no-signal hypothesis beyond the critical value of the signal hypothesis decreases in comparison to the fraction under the signal hypothesis, i.e. the test size. As this fraction corresponds to the power of the test, the power of the test of the signal hypothesis is reduced in comparison to the test size.[4] However, the power of the test of large signals increases, as these hypotheses are unaffected by the parameter bound. These hypotheses are that far from the parameter bound and can be rejected when the observed event number is larger or smaller than the expected one which corresponds then to a two-sided limit (with two critical values). Thus, the data sets with an observed event number larger than the expected one add also to the power of the test and the power increases.

9.2.2 Shape Analysis

In the definition of the shape analysis, the total neutrino event number is unconstrained. Such an analysis is for example performed when the absolute neutrino flux is unknown. The shape analysis is sensitive in the mass region from 0.1 eV^2 to 10 eV^2, where the oscillation length is around 1 m/MeV and can be directly reconstructed within the detector. Oscillation lengths smaller than the detector resolution change only the overall expected event number of the neutrino signal and an unknown neutrino flux could hence explain the change in the neutrino event number. Therefore, a shape analysis is not sensitive to large Δm^2-values.
For the disappearance experiment and given a true oscillation hypothesis, both oscillation

[3] A biased test implies that the confidence region will contain the tested hypothesis with a higher probability than designed by the test size, given the alternative hypothesis is true.

[4] The actual power is computed from the probability distributions of the test statistic.

parameters can be reconstructed (see red contour in Figure 9.4c). Hence, the degrees of freedom in a shape analysis are two. In contrast, it is worth mentioning that an appearance experiment has no exclusion sensitivity in a shape analysis. Remember in an appearance experiment the expected neutrino event number is given by:

$$N_S^{\text{exp}} = N_S \cdot P_{\text{int}}(\nu_\mu \to \nu_e) \propto N_S \cdot \sin^2(2\theta_{\mu e}). \tag{9.4}$$

Hence, the neutrino signal N_S and the oscillation amplitude $\sin^2(2\theta_{\mu e})$ are fully correlated and the degrees of freedom of the analysis is only one. Having no knowledge on N_S prevents thus from the reconstruction of $\sin^2(2\theta_{\mu e})$ and the likelihood function is independent on $\sin^2(2\theta_{\mu e})$ (see Figure 9.3c and d). Hence, given for example the no-oscillation hypothesis is true, any oscillation hypothesis could explain the data when N_S becomes arbitrarily small. This is also the reason why there is no exclusion sensitivity in the appearance experiment (see Figure 9.2b). Furthermore, no exclusion limit can be obtained from a data sample under the no-signal hypothesis (see Figure 9.4b).

In contrast, the discovery sensitivity in the appearance experiment is independent on the uncertainty of the total neutrino event number and overlaps with the rate + shape sensitivity in Figure 9.2b. Suppose the observed event number of a data set is larger than the expected one from the background, which can be constrained with some precision. In that case, the no-oscillation hypothesis is rejected independently on the uncertainty of the total neutrino event number. However, due to the correlation of N_S and $\sin^2(2\theta_{\mu e})$, the confidence region for a signal can only reconstruct the Δm^2-value (see red line in Figure 9.4d).

In comparison to the rate analysis, the ordering of the exclusion and discovery sensitivity is reversed in the shape analysis for the disappearance experiment (see Figure 9.2a). This behavior can be traced back to statistical fluctuations in the data that lead to best-fit hypotheses always different from the no-signal hypothesis, even though the true hypothesis is the no-signal hypothesis. This bias can be seen in the distribution of the MLE-values of $\sin^2(2\theta)$ and Δm^2 under the no-oscillation hypothesis for the shape analysis indicated in red in Figure 9.5a. These MLE-values are actually distributed close to the expected sensitivity of the shape analysis. In addition, one can observe that the MLE-values are completely unaffected by the boundary condition $\sin^2(2\theta) \geq 0$, which can be seen by the projection of the MLE-values on the respective survival probability shown in Figure 9.5c. The bias is that large that the MLE-values are shifted far away from the physical border at $P_{ee} = 1$. Further, the bias enlarges the total variance of the MLE which is shown in Figure 9.6b for the variance of the survival probability as a function of $\sin^2(2\theta)$ at $\Delta m^2 = 1\,\text{eV}^2$. One can clearly observe that the variance increases because of the bias by almost three orders of magnitude below $\sin^2(2\theta) < 0.05$, i.e. in the insensitive parameter region. Instead, when the oscillation signature is larger than the statistical fluctuations of the data set, the bias vanishes as the true hypothesis can be reconstructed. However, as the bias removes any dependency on the parameter bound for the variance (as discussed in the rate analysis, see Figure 9.6a) and the bias reduces with increasing $\sin^2(2\theta)$-value, the variance for the $\sin^2(2\theta)$-value close to the sensitivity is smaller than the variance of the no-oscillation hypothesis. Thus, the exclusion sensitivity is stronger in terms of $\sin^2(2\theta)$ in comparison to the discovery sensitivity. The observed distribution of the MLE-values is consistent with recent results from References [209–213]. Moreover, Reference [212] predicts the average value of the MLE of $\sin^2(2\theta)$ to $\langle \sin^2(2\theta) \rangle \approx 6.2/\sqrt{N_S^{\text{exp}}}$. Thus, in the case of the disappearance experiment a value of $\sin^2(2\theta) \approx 0.02$ is expected which agrees well with the observed values at $\sim 1\,\text{eV}^2$. Chapter 11 discusses how this bias affects the distributions of the test statistic.

Statistical fluctuations in the data and the sought-after oscillation signature lead to confidence regions which strongly fluctuate around the exclusion sensitivity in terms of $\sin^2(2\theta)$. This

behavior is especially visible for pseudo-data under the no-oscillation hypothesis from the toy disappearance experiment which is shown in red in Figure 9.4a. An oscillation length, which can describe the statistical fluctuations of a specific data set, results in a weaker exclusion limit at the respective Δm^2-value than the sensitivity. Moreover, the oscillation signature implies that typically also multiple values of the oscillation length can fit the statistical fluctuations of that data. Instead, the Δm^2-values with an oscillation length that are not compatible with the statistical fluctuations of the data, produce exclusion limits stronger than the sensitivity. Thus, the negative log-likelihood function features many local minima along the Δm^2-axis which can be seen for example in Figure 9.3 (and in Figure 11.8 of Chapter 11). These local minima complicate the minimization of the negative log-likelihood function and is one of the main challenges of a shape analysis.

9.2.3 Rate + Shape Analysis

In a rate + shape analysis, the total neutrino event number is known with the uncertainty σ_S and the analysis becomes hence sensitive to large Δm^2-values. The rate + shape analysis combines the individual sensitivities and confidence regions from the rate and the shape analyses, but is not a simple linear combination of them (see Figure 9.2 and 9.4). How the rate and the shape information exactly add depends on the parameters of the experiment. The bulk of the sensitivity steams from the rate analysis in the appearance experiment, while the shape part gives only little improvement. In the disappearance experiment, instead, the shape analysis dominates.
The rate + shape analysis has two degrees of freedom and can reconstruct both oscillation parameters under a signal hypothesis in the disappearance and appearance experiment (see blue lines in Figure 9.4c and d). The size of the confidence region becomes smaller in the rate + shape analysis in comparison to the shape analysis, as the constrained neutrino event number decreases the correlation to $\sin^2(2\theta)$.
For large Δm^2-values ($\Delta m^2 \gtrsim 20\,\mathrm{eV}^2$), the sensitivities and limits in the rate analysis are stronger than in a rate + shape analysis (see Figure 9.2 and 9.4a and b). This can be traced back to the degrees of freedom of the individual analyses. Since the data consists in a rate analysis of only one bin, $\sin^2(2\theta)$ and Δm^2 become totally correlated and the degree of freedom is one. While the degrees of freedom in a rate + shape analysis are two, and confidence regions increase with the number of degrees of freedom[5], the sensitivities do not overlap. In other words, in a rate analysis one loses all sensitivity to reconstruct Δm^2, but at the same time gains sensitivity in terms of $\sin^2(2\theta)$.
The MLE for the no-signal hypothesis is also biased. However, in comparison to the shape analysis, the parameter space of the MLE-values is narrowed and the bias is hence reduced (see Figure 9.5 in blue). Thus, the bias becomes smaller with decreasing uncertainty on the total neutrino event number. Moreover, the bias in the appearance experiment is found to be smaller than the one in the disappearance experiment.
The presence of the parameter bound and at the same time the bias of the MLE are responsible for the unintuitive effect that the exclusion sensitivity and limit become stronger with larger uncertainties on the event number for $\Delta m^2 \sim 1\,\mathrm{eV}^2$ (see Figure 9.2a and Figure 9.4a). To demonstrate this behavior in more detail, Figure 9.8 shows the exclusion and discovery power for several systematic uncertainties on the total neutrino event number as a function of $\sin^2(2\theta)$. The discovery power behaves as expected and is strongest for all $\sin^2(2\theta)$ - values for the smallest uncertainty. However, the ordering of the exclusion power reverses around the sensitive $\sin^2(2\theta)$-value: For $\sin^2(2\theta)$-values smaller than the sensitivity, the power is

[5] See Appendix A.8 for an illustration.

maximal for the largest uncertainty, while for $\sin^2(2\theta)$-values above the sensitivity the power is minimal for the smallest uncertainty (as expected). As discussed in the rate analysis, a parameter bound introduces a one-sided limit and reduces the exclusion power for the oscillation hypotheses close to the parameter bound. This is also the case for the rate + shape analysis. However, due to the bias, the MLE-values are shifted away from the parameter bound and are hence less affected by it which effectively produces a two-sided limit. Thus, as the exclusion power of two-sided limits increases with the oscillation hypothesis (also for the oscillation hypotheses close to the parameter bound), the bias counteracts the reduction of the power induced by the parameter bound. Given a large enough bias, the exclusion power effectively increases for the oscillation hypotheses close to the parameter bound and results into stronger sensitivities and exclusion limits for a shape analysis than for a rate+shape analysis. The above-described phenomenon is visually explained in Figure 9.9 and 9.10.

In summary, in this chapter the issues of the analysis of a short-baseline experiment were discussed. Some unintuitive features are found, which can be traced back to two reasons: the parameter bound on $\sin^2(2\theta)$ and the statistical fluctuations in the data sample that can mimic an oscillation signature and introduce a bias on the MLE. Moreover, the likelihood pattern in a shape and rate+shape analysis with its many local minima complicates the analyses and requires optimized minimization algorithms including a huge computational effort. Therefore, the current experiments try to simplify their analyses, but introduce at the same time results that cannot be compared directly to each other. In the next chapter, the methods that are currently used by the sterile neutrino search experiments will be discussed.

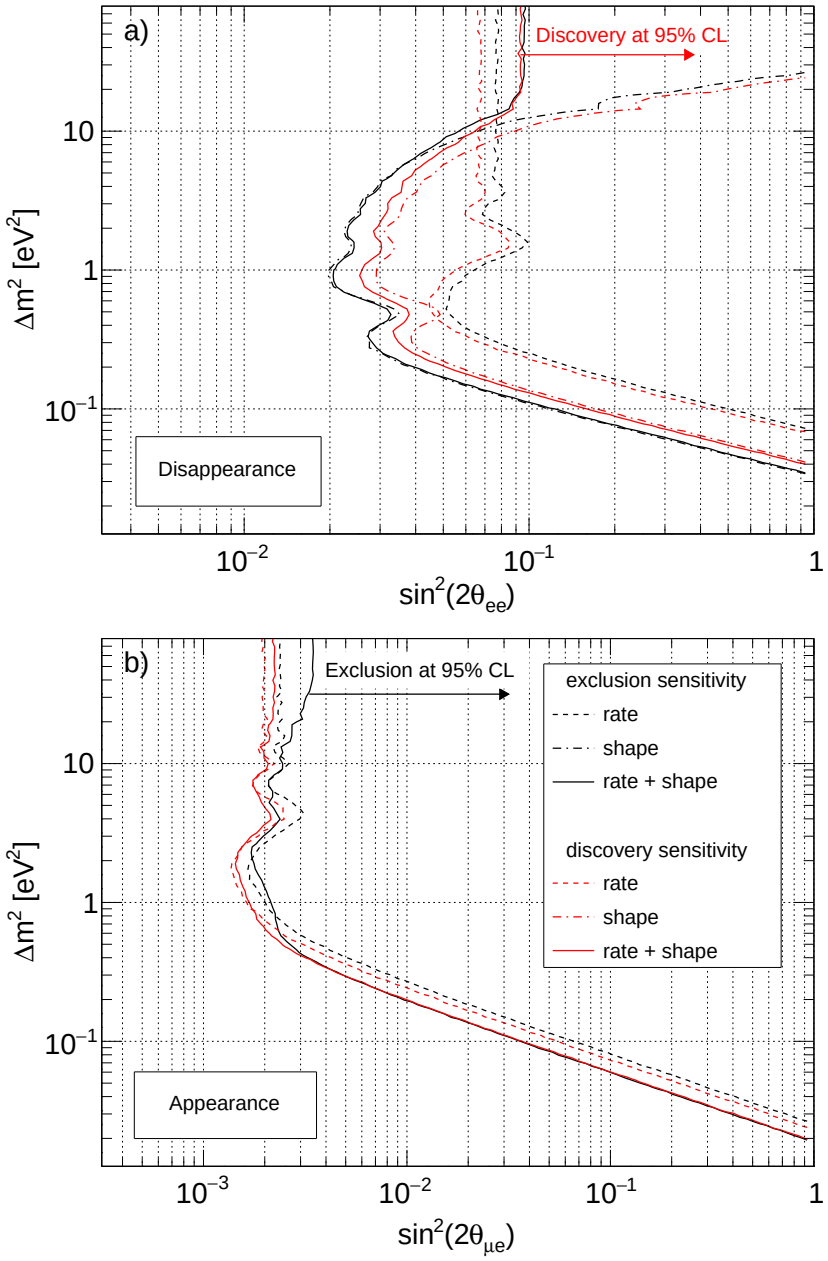


Figure 9.2: Sensitivities of Toy Experiments The sensitivities for the disappearance (a) and appearance (b) toy experiment are shown for exclusion (black) and discovery (red) in a rate, shape and rate + shape analysis at the 95 % CL. Note that the shape analysis in the appearance experiment is either consistent with the rate+shape analysis (discovery) or not sensitive at all (exclusion).

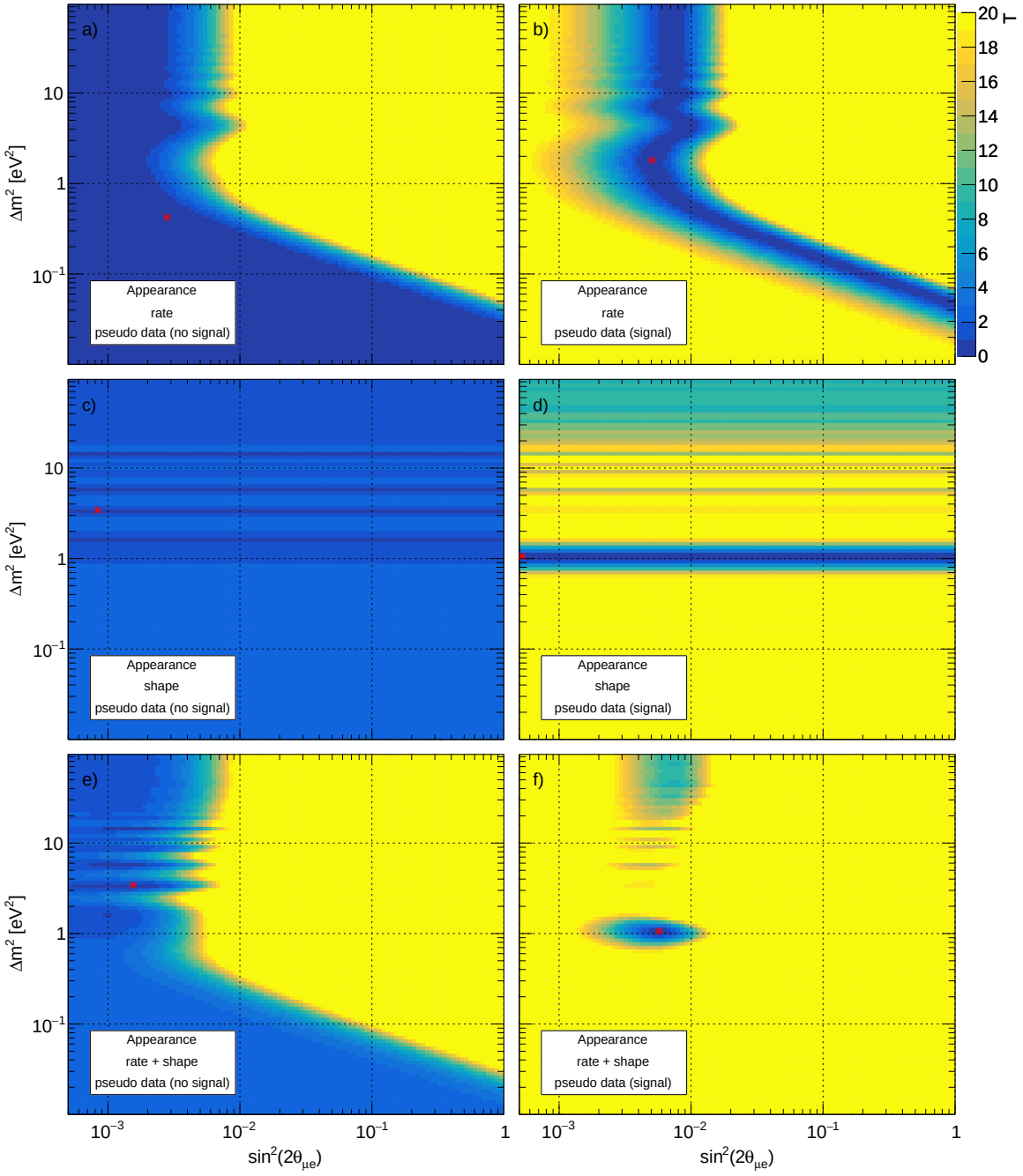


Figure 9.3: Negative Log-Likelihood Maps for the Toy Appearance Experiment The negative log-likelihood function is normalized to the absolute minimum and shown in a rate, shape and rate+shape analysis for pseudo data from the no-oscillation and from an oscillation hypothesis (signal $\sin^2(2\theta)_{\mu e} = 0.004$ $\Delta m^2 = 1\,\mathrm{eV}^2$), respectively. The negative log-likelihood function is clipped at a value of 20 and the the maximum likelihood estimator in the individual analysis is indicated in red.

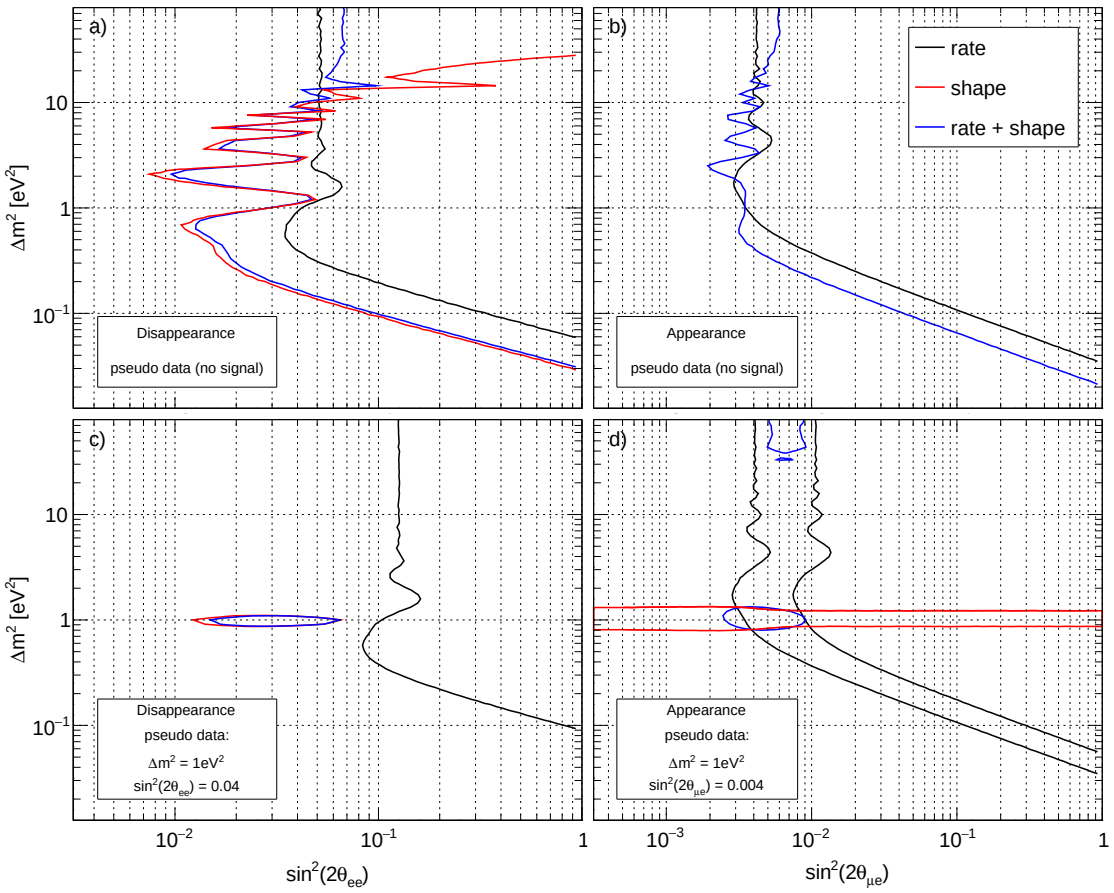


Figure 9.4: Reconstructed Confidence Regions from Pseudo-Data of the Toy Experiments The confidence regions at the 95 % CL obtained in a rate, shape and rate+ shape analysis are shown separately for the same pseudo-data either for the no-oscillation (a and b) or an oscillation hypothesis (c and d). Note that the rate analysis is not sensitive to the oscillation hypothesis of the disappearance experiment.

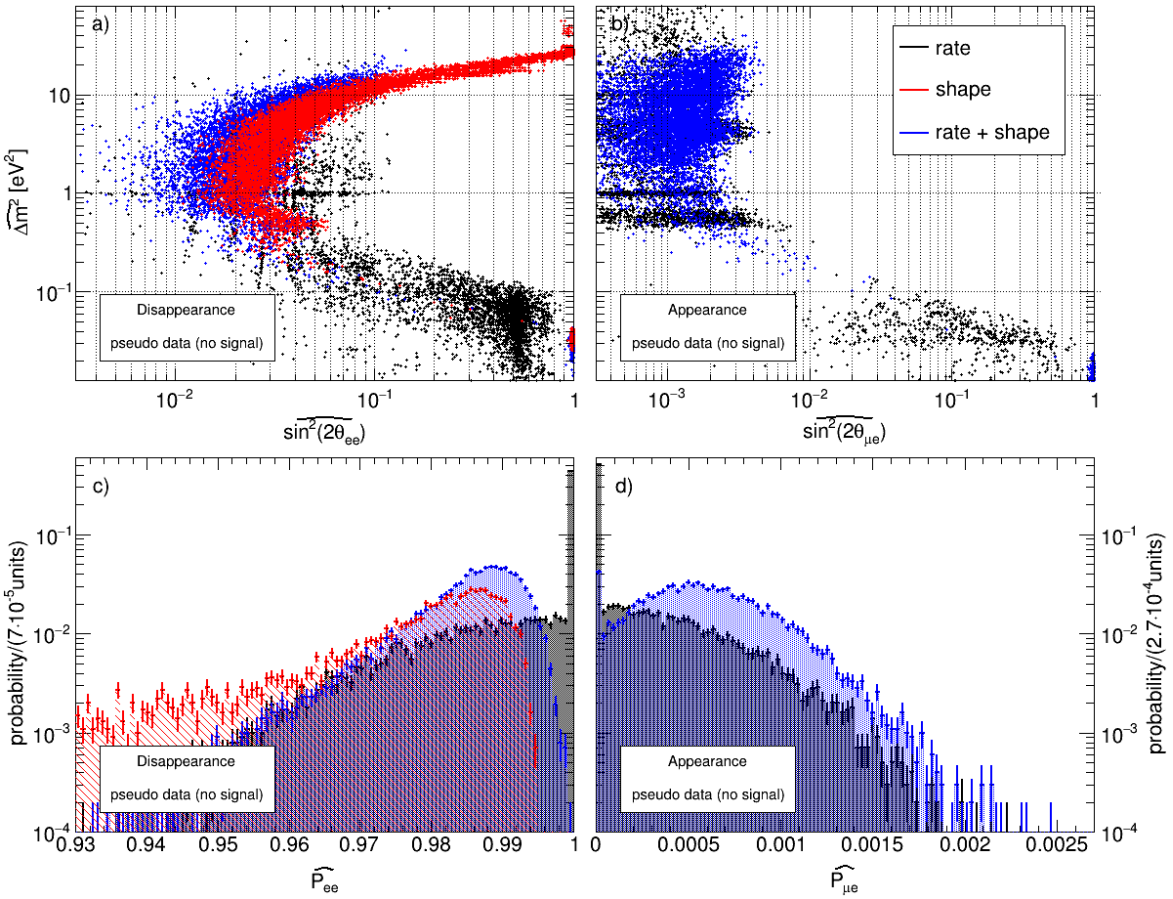


Figure 9.5: Maximum-Likelihood Estimators under the No-Oscillation Hypothesis for the Rate, Shape and Rate+Shape Analysis The MLE-values for $\sin^2(2\theta)$ and Δm^2 from 10^5 pseudo-data samples of the toy disappearance and appearance experiment are shown in a) and b). The distribution of the corresponding oscillation probabilities of a) and b) are shown c) and d), respectively. MLE-values which are reconstructed outside of the shown parameter range are not shown. The distribution of the MLE-values peak at the physical parameter bounds at one (disappearance) and zero (appearance). In the shape and rate+shape analysis, statistical fluctuations in the data sample mimic an oscillation hypothesis and introduce a bias in the analysis. The bias increases with increasing uncertainty on the total event number and moves the distribution away from the parameter bounds.

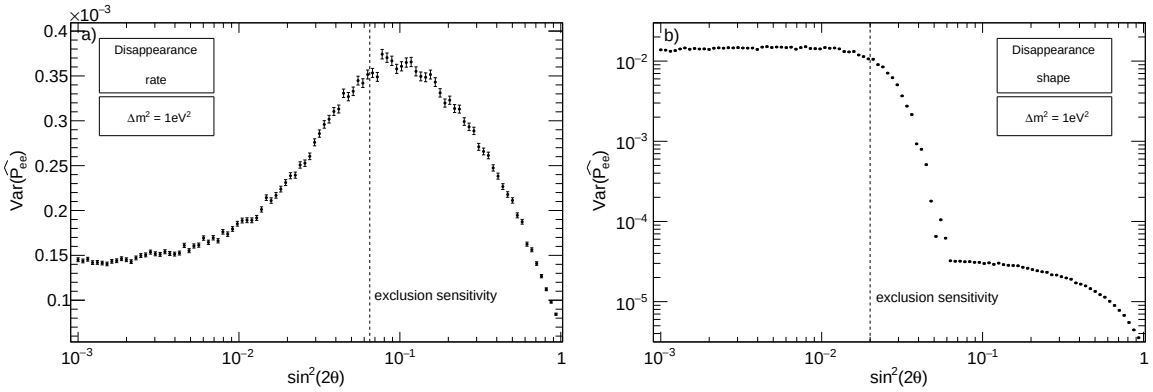


Figure 9.6: Variance of the Survival Probability. The variance as a function of $\sin^2(2\theta)$ for the toy disappearance experiment is shown in a rate (a) and a shape (b) analysis. The variance is determined from 10^5 toy experiments for each $\sin^2(2\theta)$-point. In general, the variance is decreasing with decreasing expected event number ($\sin^2(2\theta) > 0.1$). Due to the parameter bound at $\sin^2(2\theta) = 0$, the variance is suppressed for $\sin^2(2\theta)$-values smaller than the exclusion sensitivity in the rate analysis. In the shape analysis, the bias in the MLE enlarges the variance for small $\sin^2(2\theta)$-values and counteracts the suppression due to the parameter bound. Thus, in the rate analysis, the variance is larger at the $\sin^2(2\theta)$-value of the exclusion sensitivity than for the no-sterile hypothesis ($\sin^2(2\theta)$=0), in the shape analysis otherwise. This determines the ordering of the exclusion and discovery sensitivity (Figure 9.2).

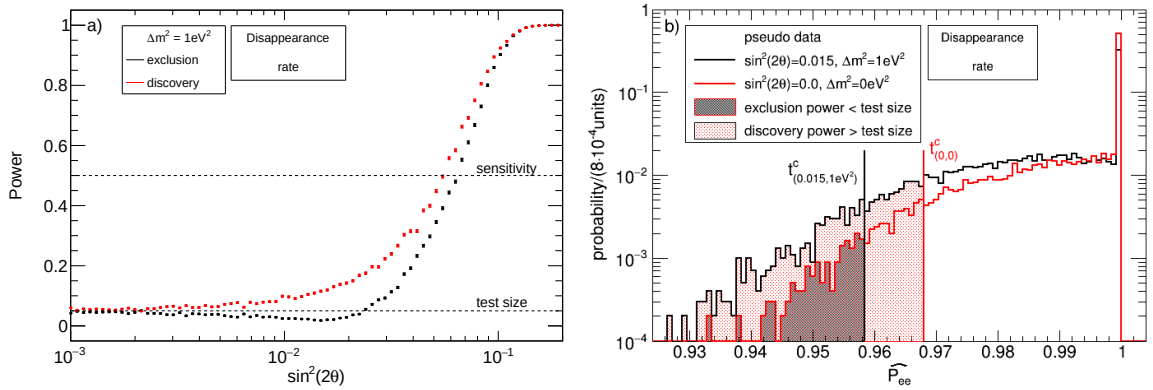


Figure 9.7: Power of a Test for Exclusion and Discovery for the Toy Disappearance Experiment in a Rate Analysis a) The power is shown as a function of $\sin^2(2\theta)$ for $\Delta m^2 = 1\,\mathrm{eV}^2$ and is computed from the probability distributions of the test statistic obtained from 10^4 pseudo-data samples each. The power increases with increasing $\sin^2(2\theta)$-value. However, due to the physical border, the power for exclusion drops below the test size of 0.05 for small $\sin^2(2\theta)$-values. b) The probability distribution of the MLE-values in terms of the survival probability is shown for the no-oscillation and the oscillation hypothesis with the minimal exclusion power of panel a. The shaded regions correspond approximately to the exclusion (black) and discovery (red) power which is defined by the fraction of the distributions beyond the critical values t^c of the alternative distributions.

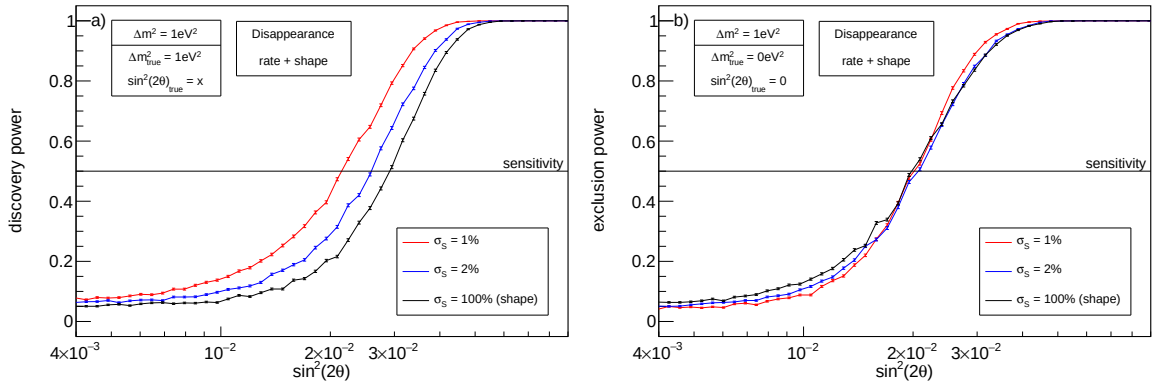


Figure 9.8: Exclusion and Discovery Power for Different Uncertainties on the Total Event Number in a Rate+Shape Analysis a) The discovery power is shown as a function of $\sin^2(2\theta)$ and decreases with increasing uncertainty. b) The exclusion power is shown as a function of $\sin^2(2\theta)$ and the ordering of the exclusion power is reversed for small $\sin^2(2\theta)$-values and leads to stronger limits in a shape analysis than in a rate+shape analysis (see Figure 9.2a and Figure 9.4a). This effect is induced by the presence of the parameter bound and the bias of the MLE (see Figure 9.9 and 9.10).

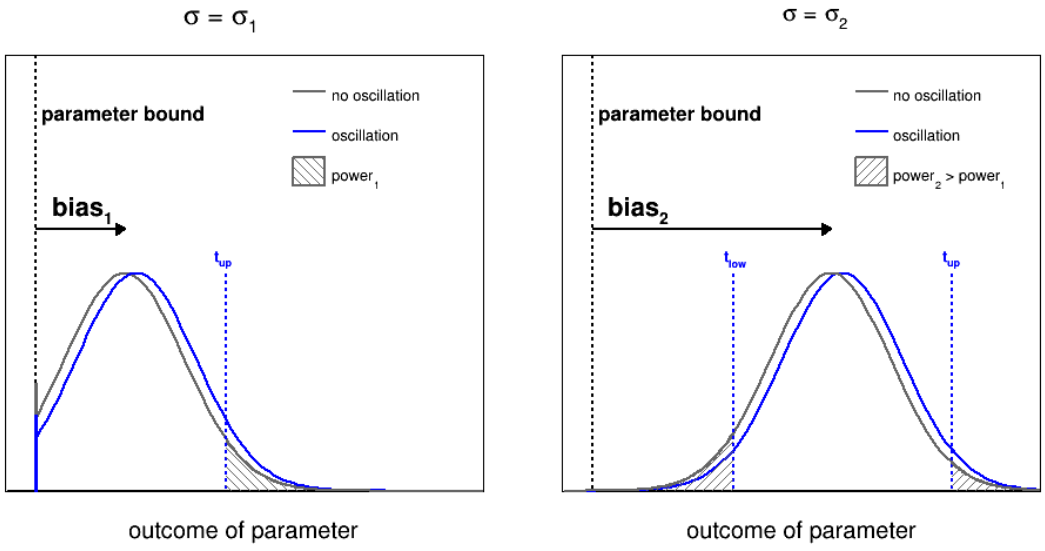


Figure 9.9: Sketch of the Impact of the Bias on the Power of a Test Two pseudo-probability distributions for the no-oscillation and a small oscillation hypothesis for two different uncertainties on the total event number σ_1 (left) and σ_2 (right), where $\sigma_1 < \sigma_2$, are shown. The presence of a parameter bound leads only to a one-sided limit (upper limit t_{up}) that induces a reduced power for small oscillation hypotheses (see also Figure 9.7). Given a bias that is large enough to shift the probability distribution away from the parameter bound, a two-sided limit arises (upper and lower limit t_{up}/t_{low}). The fraction of the probability distribution of the no-oscillation hypothesis below the lower limit contributes then also to the power and the power increases.

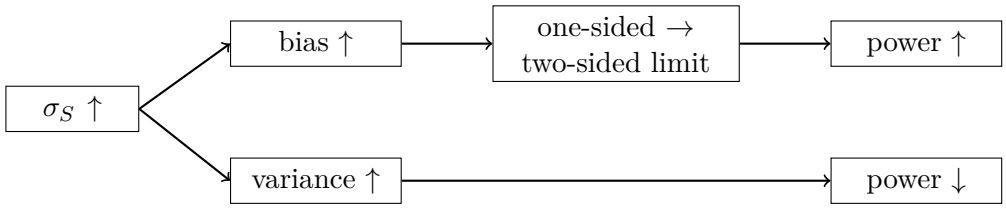


Figure 9.10: Interplay Between Bias and Variance of the Maximum-Likelihood Estimators The sketch explains the flip flop of the exclusion power in Figure 9.8 and is valid for small (insensitive) oscillation hypotheses. On the one hand, as the systematic uncertainty on the expected neutrino event number increases, the bias on the MLE becomes larger. Due to the bias, the presence of the parameter bound effectively vanishes, and the limit becomes two-sided (see Figure 9.9). As the power of a two-sided limit is larger than for a one-sided limit for the test of a small oscillation hypothesis, the power increases. On the other hand, the systematic uncertainty increases the statistical uncertainty (variance) and reduces the power of a test. For hypotheses close to the parameter bound, the bias effect can dominate and reverse the ordering of the exclusion power.

10 Comparison of Statistical Methods

As discussed in Chapter 3, current short-baseline experiments apply different analysis methods, which prevents from a direct comparison of the results. The goal of the following chapter is to compare the different methods and discuss their properties and interpretations. Here, the work of Reference [177] and [183] is extended by studying a large set of experimental parameters, as the rate, shape and rate+shape analysis in appearance and disappearance experiments. Further, exclusion and discovery scenarios are discussed.
Section 10.1 introduces the statistical methods that are afterwards compared in Section 10.4. Beforehand, the individual components of a method, namely the test statistic and the so-called $\mathrm{CL_S}$ method, are discussed separately in Section 10.2 and 10.3.

10.1 Statistical Methods

In the following, a method is decomposed into two ingredients. The first one is the choice of a test statistic that is mandatory. The second one is the optional combination with the $\mathrm{CL_S}$ method.

10.1.1 Choice of the Test Statistic and the Alternative Hypothesis

To test the hypothesis $\mathrm{H}(\sin^2(2\theta) = x, \Delta m^2 = y)$, all experiments use a likelihood ratio[1] [176]

$$\mathrm{T} = -2\ln \frac{\max\limits_{\boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta) = x, \Delta m^2 = y, \boldsymbol{\eta})}{\max\limits_{(\sin^2(2\theta),\Delta m^2)\in \mathrm{H}_1, \boldsymbol{\eta}} \mathcal{L}(\sin^2(2\theta), \Delta m^2, \boldsymbol{\eta})}, \tag{10.1}$$

where $\boldsymbol{\eta}$ represents the vector of nuisance parameters (see Chapter 3). The nominator is the maximum value of the likelihood function (see Equation 3.7) in the allowed parameter space of the nuisance parameters. The denominator maximizes the likelihood function additionally over the oscillation parameters in the parameter space defined by the alternative hypothesis H_1.
The difference in the test statistics can be traced back to the definition of the alternative hypothesis. The definition of the alternative hypothesis changes the denominator in Equation 10.1 and thus the reference hypothesis and the reference value of the test statistic. Table 10.1 lists the different test statistics and their alternative hypotheses.
When the alternative hypothesis corresponds to the physically-allowed parameter space, the test statistic is called $\mathrm{T_2}$. As the alternative hypothesis is two-dimensional, the inversion

[1] Depending on the event number, the ratio of a χ^2-teststatistic is used instead. However, the following discussion applies to both cases.

Table 10.1: Test Statistics Used in Sterile Neutrino Searches The first column specifies the nomenclature used throughout this thesis. The test statistics differ by the definition of the alternative hypothesis which is specified in the second and third column. The name of associated techniques is given in the fourth column. Experiments using different approaches are summarized in column five. Note that the table represents the status in 2018. Some of these experiments adapted their methods in the meantime (see Chapter 12).

name	H_1:{$\sin^2(2\theta)$, Δm^2} $\sin^2(2\theta)$	Δm^2	techniques associated	experiments
T_2	$0 \leq \sin^2(2\theta) \leq 1$	$\Delta m^2 \geq 0\,\mathrm{eV}^2$	2D/ global scan or global p-value	LSND [117] MiniBooNE [120] PROSPECT [128]
T_1	$0 \leq \sin^2(2\theta) \leq 1$	$\Delta m^2 = y$	raster scan or local p-value	NEOS [71] STEREO [130]
T_0	$\sin^2(2\theta) = 0$	$\Delta m^2 = 0\,\mathrm{eV}^2$	Gaussian CL_S	DANSS [126]

of the set of hypothesis tests with T_2 results into a two-dimensional confidence region (see Figure 10.1) that contains the $\sin^2(2\theta)$-Δm^2-values that are most compatible with the data within the 3+1 model. This test statistic corresponds to the discussed test statistic in the previous chapter and is agnostic towards the two oscillation parameters which is the current situation in the sterile neutrino experiments.

If the Δm^2-value could be constrained by measurements or a theoretical model in the future, the test statistic T_1 can be used. Given a known $\Delta m^2 = y$, $\sin^2(2\theta)$ remains the only free oscillation parameter. The alternative hypothesis is therefore H_1:{$\sin^2(2\theta)$, Δm^2:$0 \leq \sin^2(2\theta) \leq 1, \Delta m^2 = y$}. This test produces a one-dimensional confidence region for $\Delta m^2 = y$. Even though Δm^2 is currently not known, T_1 is applied by NEOS and STEREO[2]. As the negative log-likelihood function has several local minima along the Δm^2-axis (see for example Figure 9.3 or Figure 11.8), the minimization process for T_2 requires a large computational effort. Fixing the Δm^2-value simplifies the minimization process and reduces therefore the computation time. To obtain a two-dimensional region, the construction of a one-dimensional confidence region is repeated for all possible Δm^2-values. Hence, the alternative hypothesis changes for each Δm^2-value and always corresponds to the one-dimensional $\sin^2(2\theta)$-space at the Δm^2-value of the null hypothesis. The confidence region[3] obtained with T_1 corresponds to the union of one-dimensional confidence regions, where each one-dimensional confidence region contains the $\sin^2(2\theta)$-values most compatible with the data for the respective Δm^2-value.

The alternative hypothesis for T_0 is the no-oscillation hypothesis. This test statistic could be used if both oscillation parameters are motivated by a theory or a measurement. Besides the nuisance parameters, all parameters are fixed and the test compares two (almost) simple hypotheses. Hence, the computational effort is low. The obtained confidence regions are only point-like. This means for each hypothesis in the $\sin^2(2\theta)$-Δm^2-plane, the construction of the confidence region needs to be repeated to obtain a two-dimensional region. When testing

[2]STEREO applied T_1 in their first analysis [130]. In a recent update, the results are in addition also shown for T_2 [211].

[3]Even though the two-dimensional region does not correspond to a confidence region by definition, the regions reconstructed by any method will be referred to as a confidence region in the following.

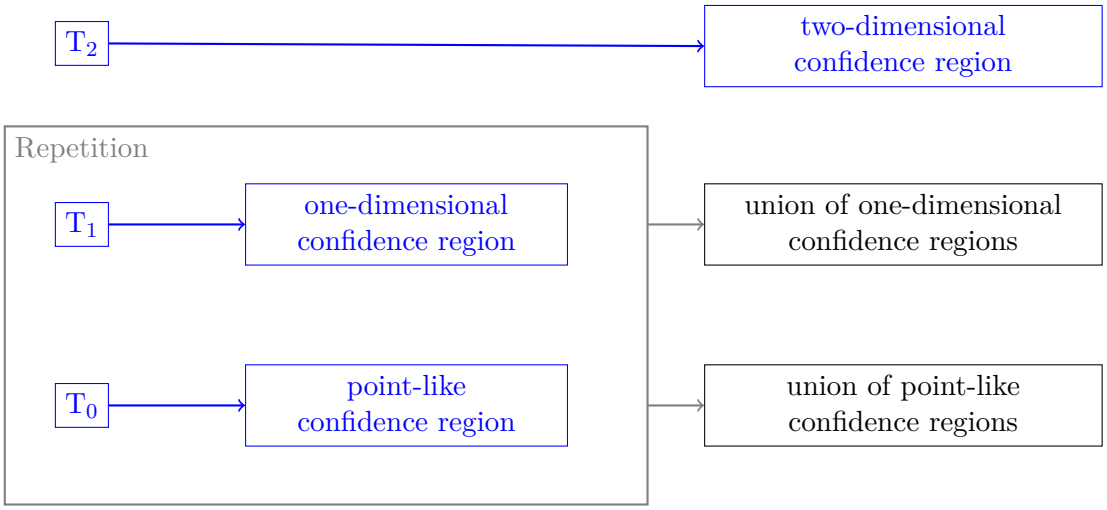


Figure 10.1: Construction of Two-Dimensional Confidence Regions The construction and interpretation of the confidence region depends on the test statistic. The construction using the text-book approach (see Chapter 3.2) is indicated with blue. Through the repetition of the text-book approach, two-dimensional regions are obtained for T_1 and T_0.

the no-oscillation hypothesis, the null and the alternative hypothesis are the same. Hence, the no-oscillation hypothesis is always accepted and no discovery can be made. Further, the confidence region will contain all $\sin^2(2\theta)$-Δm^2-values that are more likely than the no-oscillation hypothesis.
It has to be noted that all three test statistics are conceptually correct for the test of a single hypothesis. However, the interpretation differs and some issues, which will be discussed in this chapter, arise when the test statistics T_0 and T_1 are used to generate two-dimensional confidence regions.

10.1.2 Combination with the CL_S Method

In order to avoid to reject hypotheses beyond the sensitivity, the CL_S method can be applied. This method results into more conservative sensitivities and limits and is especially used in the searches for small signals [176].
Figure 10.2 sketches the CL_S method: when one is testing a hypothesis with a small oscillation signal $H(x, y)$, the distributions of the test statistic under the oscillation hypothesis $f(t|(x, y))$ and the no-oscillation hypothesis $f(t|(0, 0))$ typically overlap. Given an observed test statistic value t^{obs}, one can compute the p-value under the oscillation hypothesis and under the no-oscillation hypothesis via

$$p(x, y) = \int_{t^{obs}}^{\infty} f(t|(x, y)) \quad \text{and} \tag{10.2}$$

$$p(0, 0) = \int_{-\infty}^{t^{obs}} f(t|(0, 0)). \tag{10.3}$$

Instead of the standard rejection criteria $p(x, y) < \alpha$, the hypothesis $H(x, y)$ is rejected, when

$$CL_S = \frac{p(x, y)}{1 - p(0, 0)} < \alpha \tag{10.4}$$

with the test size α.

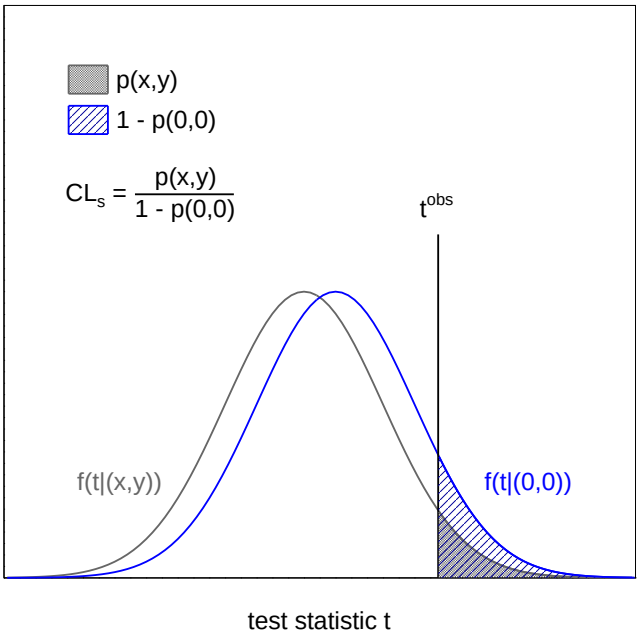


Figure 10.2: Sketch of the CL_S Method The hypothesis $H(x, y)$ is rejected when the CL_S-value is smaller than the test size α.

For small oscillation signals, $p(0, 0)$ is smaller than one and leads to a more conservative limit. When the oscillation signal is large, $f(t|(x, y))$ and $f(t|(0, 0))$ are well separated and $p(0, 0)$ becomes negligible. The CL_S limit converges thus to the standard one.
In short-baseline experiments the CL_S method is combined with the test statistic T_0. This method is known as the "Gaussian CL_S method" [183].

10.2 Comparison of the Applied Test Statistics

The test statistics are compared using the toy disappearance experiment (see Chapter 9.1) and the toy-Monte Carlo (MC) approach to generate the probability distributions of the test statistics (see Chapter 3.2.3). How these results compare to the ones under the assumption of the asymptotic probability distributions of the test statistic, is shown in Chapter 11. The test statistics are discussed in exclusion and discovery scenarios exemplarily for the rate+shape analysis. However, similar results are obtained in a rate and shape analysis.

10.2.1 Setting a Limit

Figure 10.3a shows the exclusion sensitivities for the different test statistics. The sensitivity of T_0 is stronger than T_1, and the sensitivity of T_1 is in turn stronger than T_2. This ordering comes from the number of free oscillation parameters in the alternative hypothesis that define the degrees of freedom in the analysis. Fixing a parameter before the analysis, increases the ability to distinguish between the hypotheses of the residual parameters. Hence, the power (the probability to reject the tested hypothesis when the alternative is true) follows the same ordering[4] and is largest for T_0. This power is shown in Figure 10.5a assuming that the no-

[4]Below $\sin^2(2\theta)$=0.01, the power of T_1 is suppressed because of the parameter bound at $\sin^2(2\theta)$=0. Due to the bias in T_2, the suppression of the power is reduced in comparison to T_1 (see Chapter 9.2.1).

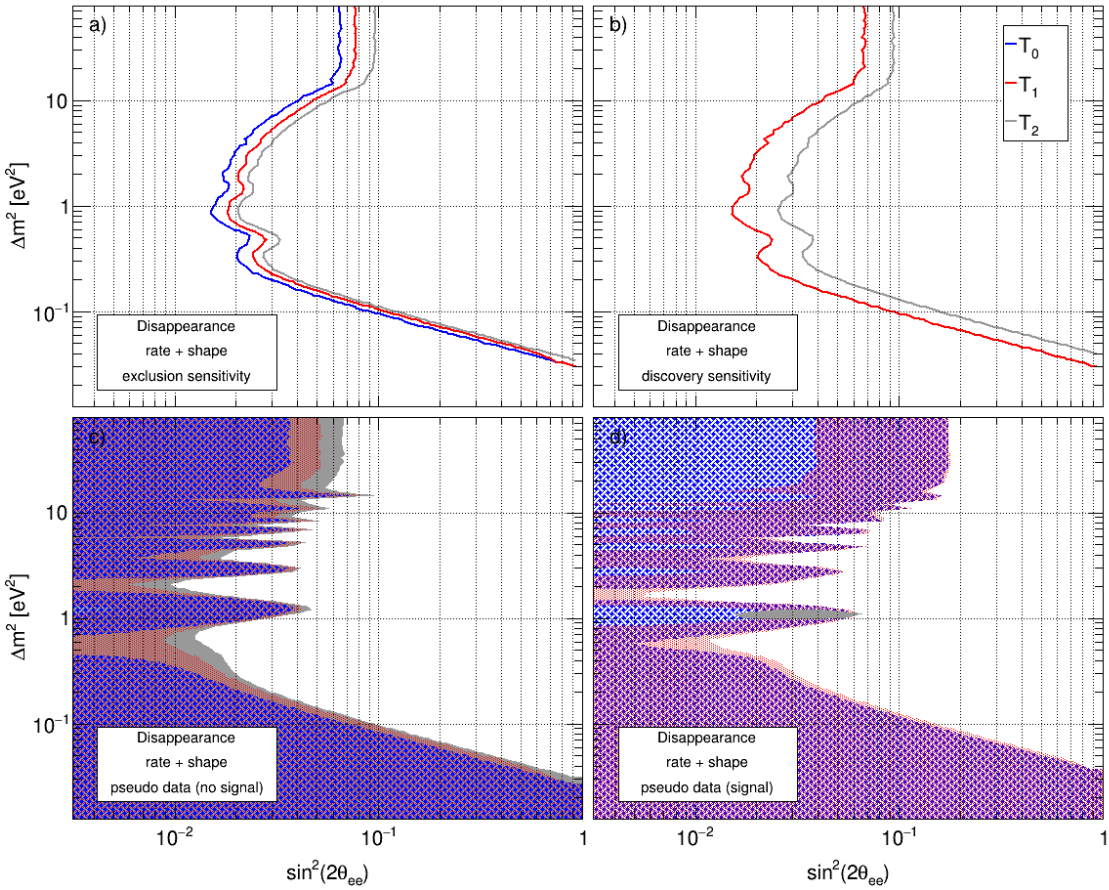


Figure 10.3: Comparison of the Test Statistics a) exclusion and b) discovery sensitivities are shown at the 95 % CL, where the parameter space right of the sensitivities can be excluded/discovered. Note that there is no discovery sensitivity for T_0 in panel b. c) and d) confidence regions at the 95 % CL are reconstructed from pseudo data from the no-oscillation and an oscillation hypothesis ($\sin^2(2\theta)$=0.04 and Δm^2=1 eV^2).

signal hypothesis is true. The ordering of the power is also the reason why the confidence regions typically get larger with the degrees of freedom in the analysis, which can be seen in Figure 10.3c for pseudo data from the no-oscillation hypothesis. As a limit corresponds to a confidence region with the lower bound at zero and the upper bound at the limit, a smaller confidence region corresponds to a stronger limit. Even though the limits and sensitivities reconstructed with T_0 and T_1 are typically stronger, one has to consider that the oscillation parameters are currently unknown. Therefore, results obtained with T_0 and T_1 can be overestimated up to a factor of two. Note that this factor depends on the experimental parameters and the type of analysis, which will be shown at the end of the chapter.
This overestimation can be very large in the case of T_0 for single Δm^2-values (e.g. 0.5 eV^2 and 2 eV^2 in Figure 10.3c). Such a strong overestimation happens when the no-oscillation hypothesis is the most likely hypothesis for that specific Δm^2-value. Since T_0 has almost no degree of freedom[5], it already rejects all hypotheses that are slightly more unlikely than the no-oscillation hypothesis. To weaken such strong limits, T_0 is typically used together with the CL_S method.
It has to be added that the ordering of the limits of T_1 and T_2 can be reversed in the case

[5]The presence of nuisance parameters introduce a tiny degree of freedom.

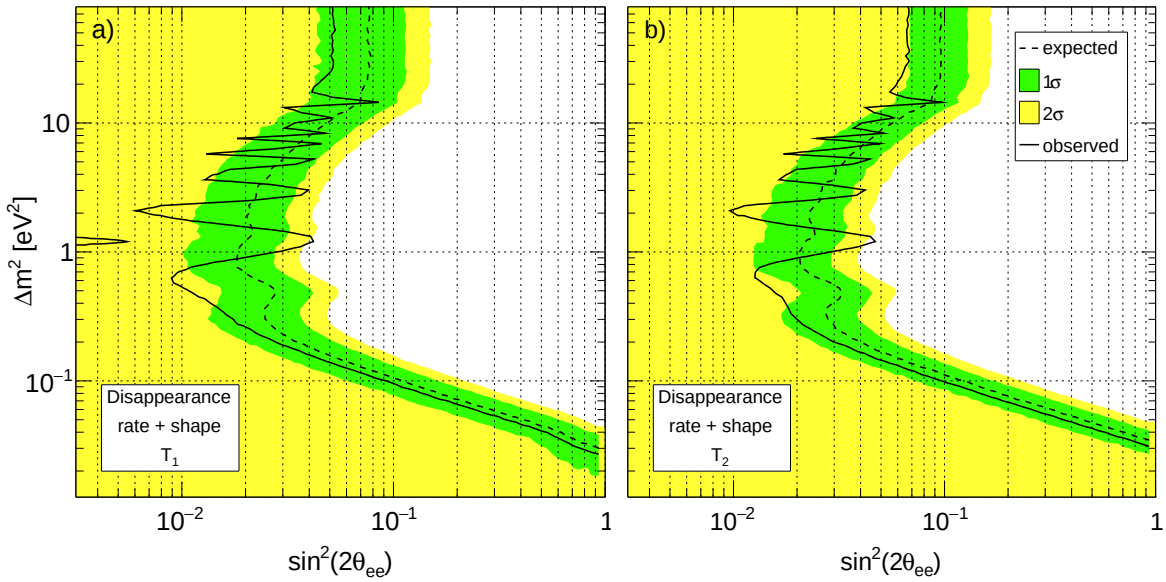


Figure 10.4: Statistical Variance of the Exclusion Sensitivity The 1σ and 2σ variance is shown for the 95 % CL level of the exclusion sensitivity for a) T_1 and b) T_2. In addition, the limits reconstructed from the same pseudo-data sample from the no-oscillation hypothesis are shown. Note, that T_1 can reject the no-oscillation hypothesis for single Δm^2-values (here at $\Delta m^2 \sim 1.1\,\mathrm{eV}^2$).

when a signal hypothesis is strongly favored over the no-oscillation hypothesis, however, the no-oscillation hypothesis cannot be excluded yet. Thus, such a situation can be connected to the p-value of the no-oscillation hypothesis. For example, this ordering is observed for p-values smaller than ~ 0.25 and ~ 0.15 in the shape and rate+shape analysis of the toy disappearance experiment, respectively.[6] However, as the signal is strongly preferred over the no-oscillation hypothesis, the no-oscillation hypothesis at the specific Δm^2-value of the best-fit hypothesis is typically rejected with T_1. Such a result is reported by the latest STEREO publication, where the p-value of the no-oscillation hypothesis corresponds to 0.09 [211].

Due to the above-mentioned overestimation, it is useful to compare a limit to the expected sensitivity and its variance. In Figure 10.4 the 1σ and 2σ variation of the exclusion sensitivities for T_1 and T_2 are shown. The variation of T_2 is slightly larger than for T_1, due to the higher number of degrees of freedom in the analysis. In addition, a pseudo-data sample from the no-oscillation hypothesis is shown. The observed event number of this data sample is smaller than the expected one and causes overall stronger limits than expected. Moreover, the limits fluctuate around the sensitivity especially for $0.1\,\mathrm{eV}^2 \lesssim \Delta m^2 \lesssim 20\,\mathrm{eV}^2$. The size of the fluctuations are connected to the 1σ and 2σ variation of the exclusion sensitivities, as they contain the limit with a probability of 68 % and 95 %, respectively. These fluctuations of the limit are due to the statistical fluctuations in the data sample which can be described by an oscillation signature for some specific Δm^2-values (see Chapter 9.2.2). In comparison to T_2 and as the Δm^2-values are tested individually in T_1, the no-oscillation hypothesis can be rejected for single Δm^2-values (for example at $\Delta m^2 \sim 1.1\,\mathrm{eV}^2$). The results of this section are compatible with References [177, 183, 214].

[6] Of course the p-value needs to be larger than the test size of 5 %, as the confidence region is no limit otherwise.

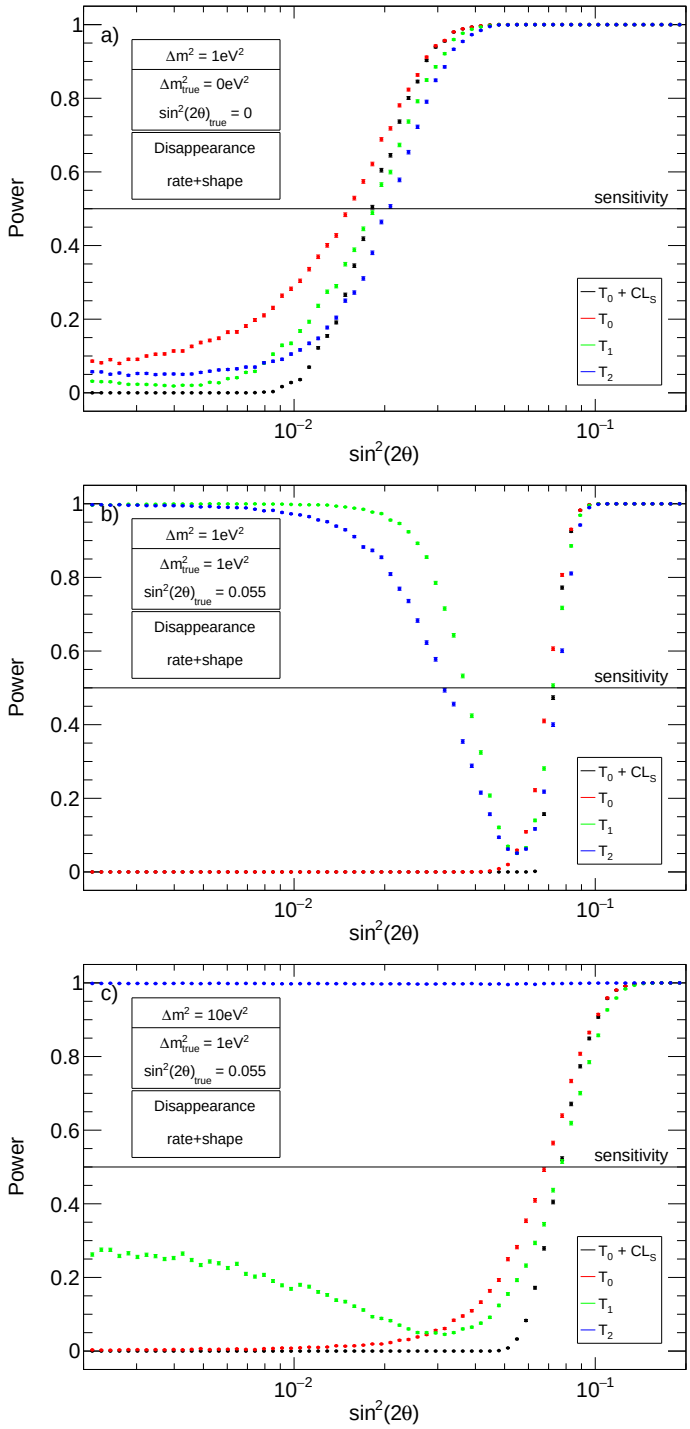


Figure 10.5: Power for the Different Statistical Methods The power is computed for the 95 % CL, under the assumption that a) the no-oscillation hypothesis is true b) an oscillation hypothesis is true c) an oscillation hypothesis is true, but a different Δm^2-value is tested ($\Delta m^2 \neq \Delta m^2{}_{\text{true}}$). Notice that the difference between the true and the tested hypothesis is so large in panel c that the power of T_2 is always one.

10.2.2 Looking for a Signal

Figure 10.3b compares the discovery sensitivities of the test statistics. T_0 has no sensitivity to discover a sterile neutrino, as T_0 cannot reject the no-oscillation hypothesis. Hence, the power for the no-oscillation hypothesis is zero and the test is biased. T_1 can discover a larger oscillation parameter space than T_2. This follows again from the different number of free parameters in the alternative hypothesis. However, due to the fixed Δm^2-value in the test statistic, the discovery sensitivity of T_1 is about a factor of 1.5 overestimated.

The difference becomes more visible in the confidence regions reconstructed from pseudo data from a signal hypothesis, which is shown in Figure 10.6 for $\sin^2(2\theta)= 0.04$ and Δm^2=1 eV^2. T_2 creates a narrow confidence region around the true oscillation parameters. The significance of such a signal corresponds to $\sim 3.5\sigma$ (red line in the right panel). Instead, the significance of T_1 becomes a function of the Δm^2-value, as each Δm^2-value is tested independently from each other. Therefore, for the present pseudo data the no-sterile hypothesis can be rejected several times at the 3σ-level (e.g. 1 eV^2, 2.5 eV^2, 5 eV^2, 9 eV^2). As previously discussed, constraining the Δm^2-value before testing the no-sterile hypothesis, leads to a larger discrimination power between the hypotheses and hence to a larger, but overestimated significance at the true Δm^2-value of $\sim 4.5\sigma$. In high energy physics, a popular method exists to correct the significance obtained from T_1 and account for the so-called look-elsewhere effect. This method is applied to the sterile-neutrino search and discussed in Chapter 11.4.1. In addition to the overestimation of the significance, the confidence region of T_1 cannot reconstruct the true Δm^2-value, as it is the union of one-dimensional $\sin^2(2\theta)$-confidence regions for each Δm^2-value. These results are in agreement with References [177, 214].

Even though T_0 cannot be used to claim a discovery, a confidence region can be reconstructed in any case. Fig 10.3d shows the confidence region of T_0 for the same pseudo data in comparison to T_1 and T_2. The confidence region corresponds only to a limit, as it contains the no-oscillation hypothesis and all hypotheses that are more likely than it. Hence, neither $\sin^2(2\theta)$ nor Δm^2 can be reconstructed with T_0.

As just mentioned, T_0 and T_1 cannot pinpoint an oscillation signal and the confidence regions contain a lot of $\sin^2(2\theta)$-Δm^2-values that are indeed not true. The smaller the so-called probability of false acceptance is, the more desirable, informative, and accurate the tests and the obtained confidence regions become [215]. As the probability of false acceptance of a test is the complementary probability of the power, a large power is desirable. The power for the different test statistics are compared in Figure 10.5. As already discussed, the power is largest for T_0 when the true hypothesis is the no-oscillation hypothesis (panel a). In this case, the power grows monotonically with the mixing angle for all test statistics. Given a true sterile-neutrino hypothesis with $\Delta m^2{}_{\text{true}}$, the power is largest for T_1 when the tested Δm^2-value corresponds to the true one (panel b). Instead, T_0 loses power because the true hypothesis is defined neither in the null nor in the alternative hypothesis. While the power has a minimum for T_1 and T_2 at the true value of the mixing angle and reaches a maximal value of one for the no-oscillation hypothesis, the power of T_0 is minimal for the no-oscillation hypothesis. Further, testing Δm^2-values different from the true one, T_2 becomes the most powerful test (panel c), as T_1 in turn does not contain the true hypothesis within the definition of the null or the alternative hypothesis. Here, the power of T_0 and T_1 is below 0.5 (i.e. the sensitivity) for many $\sin^2(2\theta)$-values. The power of T_2 is on the contrary independent on the mixing angle and is always one due to the large difference between the tested and the true hypothesis.

In case of a true signal (see panel b and c), the power of T_0 for small $\sin^2(2\theta)$-values is zero and hence smaller than the test size of 0.05. T_0 is therefore a biased test that also leads

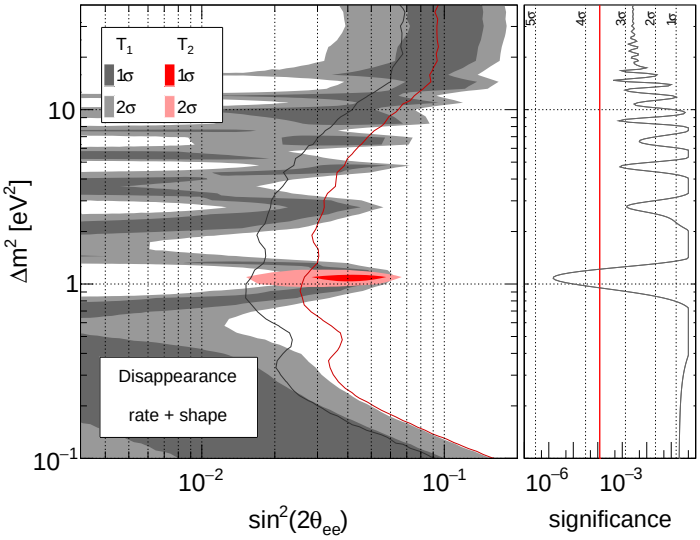


Figure 10.6: Reconstructed Confidence Regions for a Pseudo-Data Sample from an Oscillation Hypothesis The 1σ and 2σ confidence regions of T_1 and T_2 are compared. In addition, the discovery sensitivities at the 95 % CL are shown. On the right panel, the significance (p-value for the no-oscillation hypothesis) of T_1 and T_2 are compared. As T_1 tests each Δm^2-value individually, the confidence region of T_1 contains for each Δm^2-value a $\sin^2(2\theta)$-interval and the significance is a function of the Δm^2-value.

to biased confidence regions, i.e. the confidence region always contains these small $\sin^2(2\theta)$-values. This bias confirms that T_0 can only produce limits.

10.3 Impact of the CL_S Method on the Test Statistic

The CL_S method introduces an artificial overcoverage to penalize exclusion limits which are stronger than the sensitivity. This behavior is shown in Figure 10.7a that compares the exclusion limits obtained from a pseudo-data sample from the no-oscillation hypothesis for T_0 and T_0+CL_S. The overall limit is more conservative, especially in the parameter region, in which T_0 is not sensitive. In the sensitive parameter region, the limit of T_0 converges to the one of T_0+CL_S. To achieve this weakening effect, the power is zero for small $\sin^2(2\theta)$-values (see Figure 10.5). Thus, these small $\sin^2(2\theta)$-values will always be within the confidence region. This is equivalent to a coverage of 100 % for these hypotheses, which is shown in Figure 10.7b. Thus, the confidence regions obtained with T_0+CL_S are biased and their interpretations are not trivial. For large $\sin^2(2\theta)$-values the power and the coverage of T_0+CL_S converges with the ones from T_0.

For the same reason, also the exclusion sensitivity of T_0+CL_S is more conservative as T_0 (see Figure 10.7). Actually, the sensitivity with a 95 % CL of T_0+CL_S corresponds to the 97.5 % CL of T_0 which is proven in Appendix A.7.

The combination of T_0+CL_S adapts the interpretation of T_0. Hence, only limits can be reconstructed and the confidence region contains all hypotheses that are more likely than the no-oscillation hypothesis. These results are consistent with Reference [183].

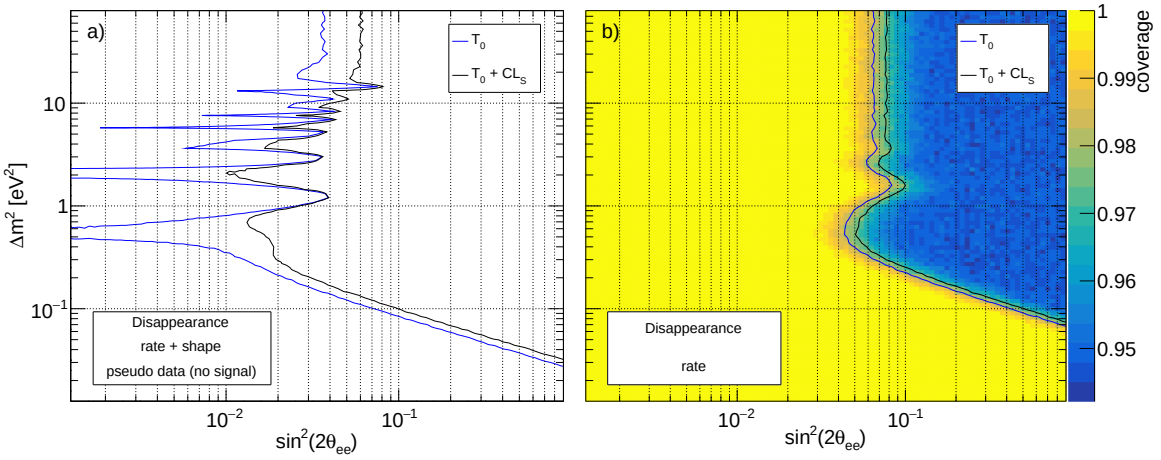


Figure 10.7: Impact of the CL$_S$ Method a) The limits of T_0 and T_0+CL$_S$ are compared using pseudo data from the no-oscillation hypothesis at the 95 % CL. b) The coverage of the T_0+CL$_S$ method is computed for the nominal level of 0.95. In addition, the exclusion sensitivities of T_0 and T_0+CL$_S$ are compared at the 95 % CL.

10.4 Comparison of the Applied Methods

In addition to the previously discussed properties, a full comparison of the currently used methods, namely T_0+CL$_S$, T_1, and T_2 for a rate, shape, and rate+shape analysis, is shown in the following. Figure 10.8 shows a comprehensive overview of the differences by comparing the sensitivities and reconstructed confidence regions from pseudo data. A detailed discussion on the ingredients of the methods can be found in the previous sections. In the following, additional information will be given.

The exclusion sensitivity of T_0+CL$_S$ overlaps with the sensitivity of T_1 which can be seen in panel a-b and is consistent with Reference [183]. The fact that the sensitivities of the two methods are the same is related to the fulfilled asymptotic behavior of the test statistic distributions and is discussed in detail in Appendix A.7. However, the limits obtained with T_0+CL$_S$ can be weaker or stronger than T_1, depending on how the obtained result compares to the sensitivity. Stronger limits than expected for a given Δm^2-value also cause stronger limits for T_1 than for T_0+CL$_S$, as the CL$_S$ method weakens the limit for T_0+CL$_S$ (see for example Δm^2=50 eV^2 in panel f). In case of a weaker limit than expected, the situation is reversed, as the CL$_S$ method has no effect on T_0 and the difference in the degrees of freedom in the test remains (see for example Δm^2=1.5 eV^2, 3 eV^2 in panel f).

Furthermore, the sensitivities and confidence regions are the same for T_2 and T_1 in a rate analysis, as both analyses have effectively only one degree of freedom (see Chapter 9.2.1).

An interesting observation is that the differences among the exclusion sensitivities become almost negligible in a shape analysis. The reason for this behavior is the bias of T_2 for the no-oscillation hypothesis (see Figure 9.2 and Figure 9.5) that produces on average stronger limits for T_2 than expected for a two-dimensional analysis. These stronger limits are hence comparable to the overestimated ones from T_1 and T_0+CL$_S$. The bias is reduced for a rate+shape analysis, which leads, on the other hand, to larger and expected differences between the methods.

The same bias that is reducing the differences between the exclusion sensitivities of T_1 and T_2, is responsible for an increased difference in the discovery sensitivities of a shape and rate+shape analysis of T_1 and T_2 (see panel h and i). The oscillation signal has to overcome the (due to the bias) enlarged variance and the discovery sensitivity is shifted to larger

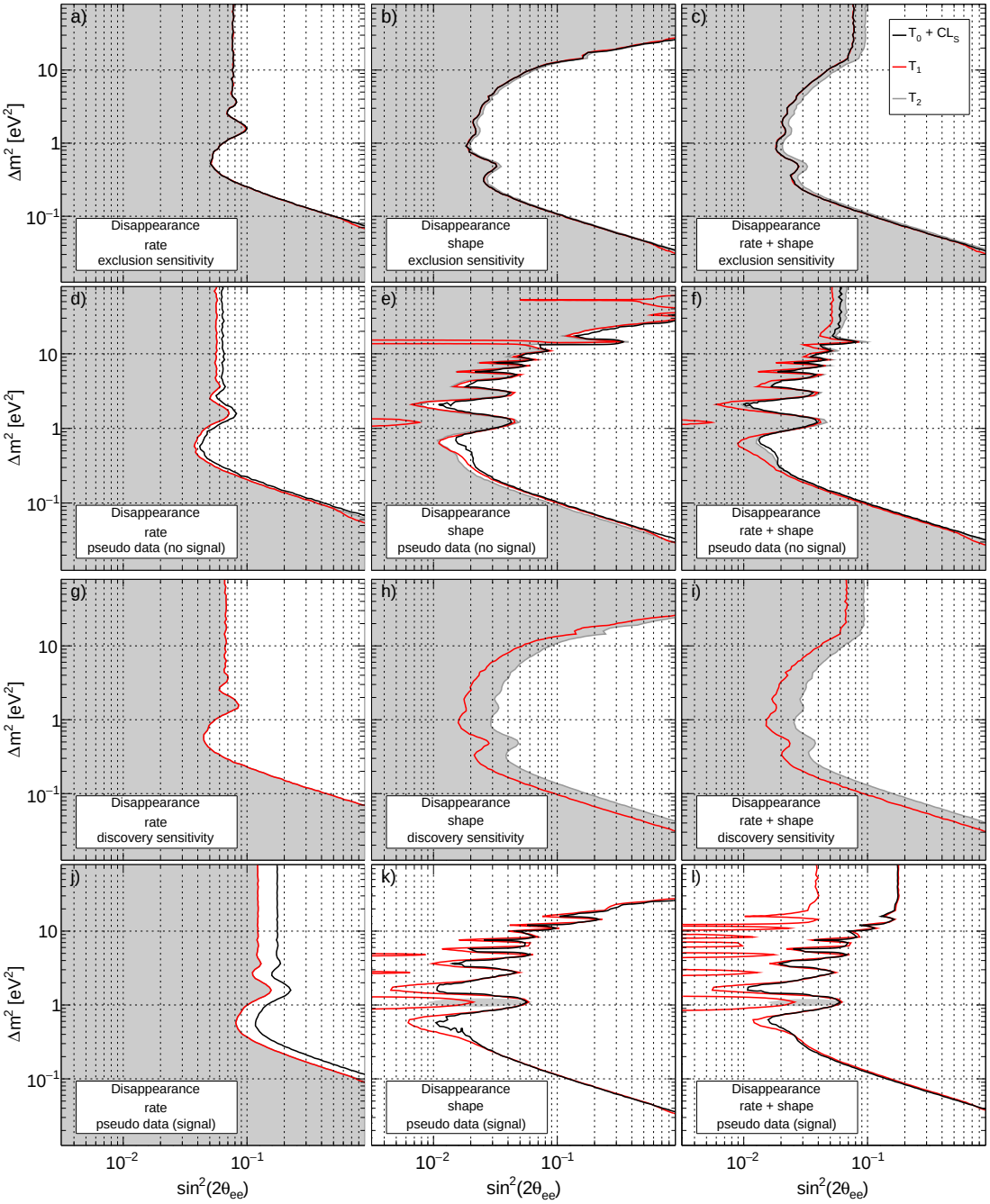


Figure 10.8: Comparison of the Applied Methods The methods are compared in a rate (left), shape (middle) and rate+shape analysis (right) using exclusion (a-c) and discovery (g-i) sensitivities. Further, confidence regions reconstructed from pseudo data from the no-oscillation hypothesis (d-i) and an oscillation hypothesis with $\sin^2(2\theta)$= 0.04 and Δm^2=1 eV^2 (j-l) are compared. All results are computed for a 95 % CL. The gray-shaded areas indicate the parameter space within the confidence region or beyond the sensitivity of T_2.

$\sin^2(2\theta)$-values for T_2.

In summary, in this chapter the currently applied methods in short-baseline experiments were compared. The differences can be mainly traced back to the definition of the alternative hypothesis in the hypothesis test that changes the interpretation of the confidence regions. The largest deviations arise, when a sterile neutrino hypothesis is true, and T_0 and T_1 become not suitable to reconstruct confidence regions. However, the differences are reduced, when the no-oscillation hypothesis is true. This may be a reason for the wide acceptance of these various methods. In addition, the computational effort is reduced for T_0 and T_1 and the asymptotic probability distributions of the test statistic are valid, which will be shown in the next chapter.

11 Validity of the Asymptotic Formulae for the Probability Distributions

As described in Chapter 9, the maximum-likelihood estimator (MLE) for the no-oscillation hypothesis is biased in a sterile neutrino search experiment. This bias in addition with the physical parameter bound on $\sin^2(2\theta)$ leads to non-intuitive issues in the analysis. Thus, one expects the probability distributions of the test statistic to deviate from their asymptotic formulae.
Therefore, this chapter discusses the probability distributions of the test statistics T_0, T_1, and T_2 and the validity of the asymptotic formulae in more detail (Section 11.1-11.3). In particular, the reconstructed results based on the assumptions of the asymptotic formulae are compared to the ones from the previous chapter. The sensitivity of SOX using a Monte Carlo (MC) construction is discussed and compared to the one based on the asymptotic formulae. Approximations for the p-value and the goodness-of-fit test are also discussed in the sterile neutrino framework in Section 11.4 and 11.5.

11.1 Probability Distributions of T_0

The probability distributions of a likelihood ratio of two simple hypotheses asymptotically tend to Gaussian functions in the large sample limit[1] [179]. Even in the presence of nuisance parameters[2], the distributions can be approximated with [183]

$$f\left(t_{(x,y)}|(x,y)\right) \sim \mathcal{N}\left(t_{(x,y)}^{A(x,y)}, 4t_{(x,y)}^{A(x,y)}\right) \quad \text{and} \tag{11.1}$$

$$f\left(t_{(x,y)}|(0,0)\right) \sim \mathcal{N}\left(t_{(x,y)}^{A(0,0)}, 4t_{(x,y)}^{A(0,0)}\right), \tag{11.2}$$

where $\mathcal{N}$ is the normal distribution and $t_{(x,y)}^{A(x,y)}$ ($t_{(x,y)}^{A(0,0)}$) is the value of the test statistic for the Asimov-data set of $\mathrm{H}(x,y)$ (H(0,0)) as defined in Formula 3.12. The asymptotic distributions hold under rather mild conditions, namely, the data size of each bin is large, the parameter range of the nuisance parameters is not constrained, and the difference between the two tested hypotheses H(0,0) and $\mathrm{H}(x,y)$ is small. Further, for small differences between the two hypotheses, $t_{(x,y)}^{A(x,y)} \approx -t_{(x,y)}^{A(0,0)}$ holds, and the probability distributions are symmetric [179].
A good agreement between the asymptotic expectations and the toy-MC distributions for the toy disappearance and toy appearance experiment is observed. Figure 11.1a compares

[1] As the null and the alternative hypothesis are not nested, Wilks' theorem does not apply here.

[2] As nuisance parameters are free parameters, the hypotheses in the hypothesis test are not exactly "simple" anymore.

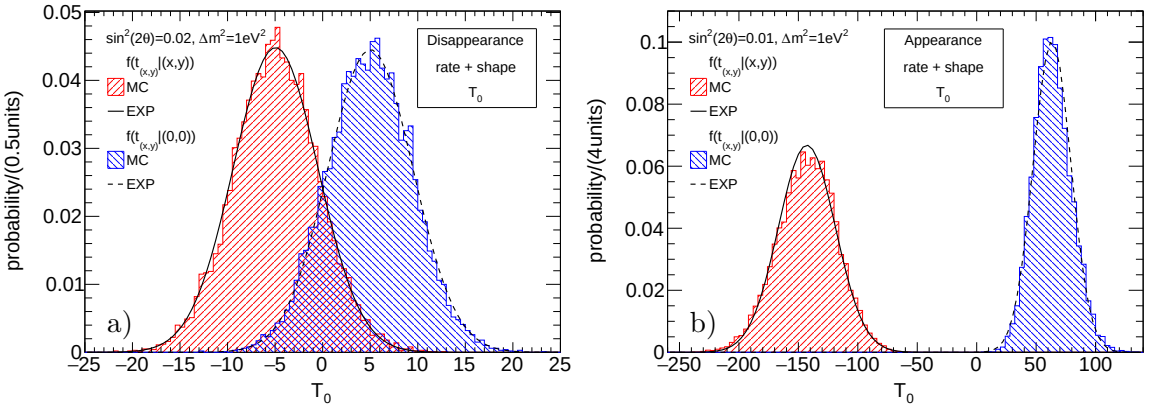


Figure 11.1: Probability Distributions of T_0 The probability distributions under the null (red) and the alternative hypothesis (i.e. no-oscillation hypothesis, blue) are produced from 10^4 toy-MC data samples each. Panel a shows the distributions for the test of an oscillation hypothesis close to the sensitivity ($\sin^2(2\theta)$=0.02 and Δm^2=1 eV^2) in a rate+shape analysis of the toy disappearance experiment, where the distributions are symmetric. Instead, in panel b, which shows the test of a well-observable oscillation hypothesis ($\sin^2(2\theta)$=0.01 and Δm^2=1 eV^2) in a rate+shape analysis of the toy appearance experiment, the symmetry is broken. In both cases, the probability distributions agree well with the expected Gaussian distributions (black lines).

exemplarily the distributions for $\sin^2(2\theta)$=0.02 and Δm^2=1 eV^2 for the toy disappearance experiment in a rate+shape analysis. At this hypothesis, which is close to the 95 % CL-sensitivity, the probability distributions are symmetric. In Figure 11.1b, the difference between the tested ($\sin^2(2\theta)$=0.01 and Δm^2=1 eV^2 for the toy appearance experiment in a rate+shape analysis) and the no-sterile hypothesis is so large that the symmetry is broken. However, the MC-approach and the expectation agree still reasonably well. The results agree with Reference [183], to which the reader is referred for more details.

11.2 Probability Distributions of T_1

In comparison to T_0, the alternative hypothesis of T_1 is a composite hypothesis, i.e. a hypothesis with free parameters. The null hypothesis is nested in the alternative hypothesis and T_1 is a profile likelihood ratio with the free oscillation parameter $\sin^2(2\theta)$. If some regularity conditions are met (see Chapter 3), the probability distribution under the null hypothesis $\left(f\left(t_{(x,y)}|(x,y)\right)\right)$ follows a χ^2 distribution with one degree of freedom (Wilks' theorem [178]). According to Wald [180], the probability distribution under the alternative hypothesis $\left(f\left(t_{(x,y)}|(0,0)\right)\right)$ follows a non-central χ^2 distribution with one degree of freedom. Here, the non-centrality parameter can be estimated with the test-statistic value using the Asimov data set (see Formula 3.12). Both approximations typically come along with normally distributed MLE-values.

Figure 11.2 shows the MLE-values for $\sin^2(2\theta)$ for the no-sterile hypothesis (blue) and for a sterile neutrino hypothesis (red) with $\sin^2(2\theta)$=0.045 and Δm^2=1 eV^2. In the physically-allowed parameter space, the MLE-values are normally distributed around the true value. In contrast, for the no-sterile hypothesis, statistical fluctuations in the data-sample that would prefer negative $\widehat{\sin^2(2\theta)}$-values are projected inside the allowed parameter space, namely to $\sin^2(2\theta)$=0, the no-sterile neutrino hypothesis. Similarly, the border at $\sin^2(2\theta)$=1 affects the distributions of the MLE-values. Hence, the physical border prevents the MLE-values

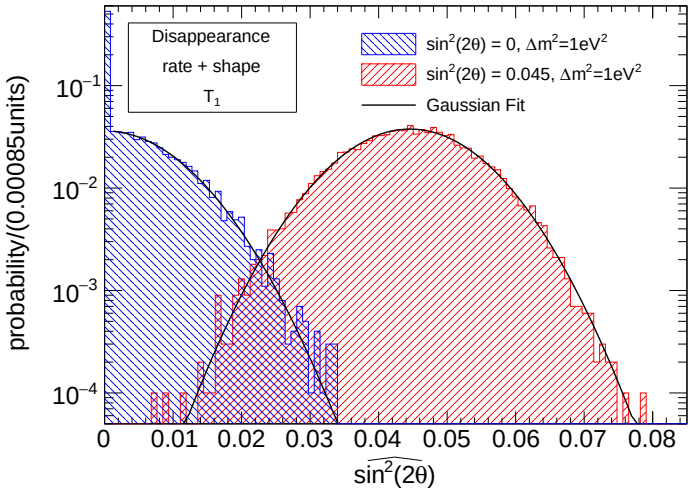


Figure 11.2: Maximum-Likelihood Estimator Values of T_1 The MLE-values of $\sin^2(2\theta)$ are obtained from 10^4 toy-MC data samples from the no-signal hypothesis (blue) and a signal hypothesis (red) for the toy disappearance experiment in a rate+shape analysis. The MLE-values are normally distributed around the true $\sin^2(2\theta)$-value. Due to the physical parameter border, negative values are projected to zero.

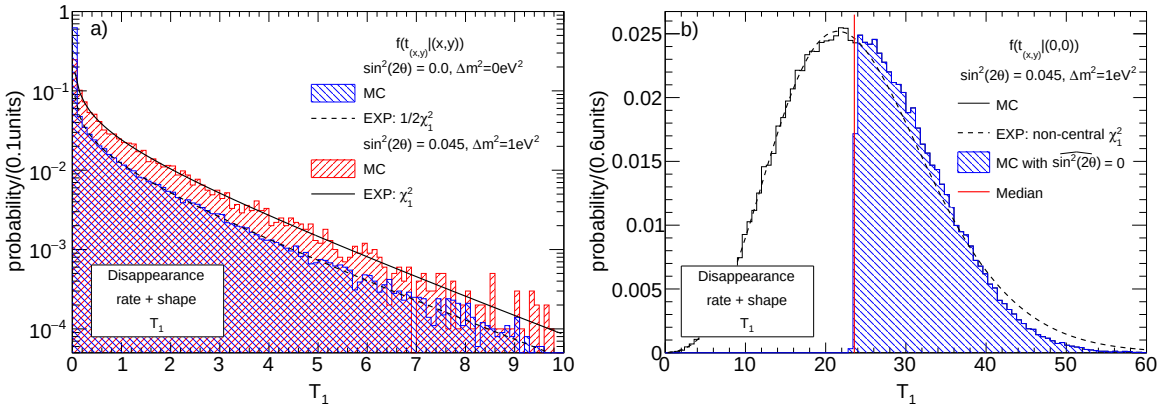


Figure 11.3: Probability Distributions of T_1 a) The distributions under the null hypothesis for the no-signal (blue) and a signal (red) hypothesis can be described by the expected distributions (black lines). b) The distribution under the alternative hypothesis (i.e. no-signal hypothesis and $\sin^2(2\theta)_{\text{test}}$=0.045, $\Delta m^2{}_{\text{test}}$=1 eV^2) is shown in solid black for the MC construction and in dashed black for the expectation. The median of both distribution overlaps and is indicated in red. The fraction of the MC distribution, where the no-signal hypothesis is the best-fit hypothesis for the respective toy-data samples, is shaded in blue. Both distributions are constructed from 10^4 toy-MC data samples each for the toy disappearance experiment in a rate+shape analysis.

from being normally distributed, when the true value is close to it.

In the case, when the MLE-values are normally distributed, the probability distributions under the null hypothesis follow a one-dimensional χ^2 distribution. Figure 11.3a shows the agreement of the asymptotic and toy-MC distribution for an exemplary sterile neutrino hypothesis with $\sin^2(2\theta)$=0.045 and 1 eV^2. For all other hypotheses with normally distributed MLE-values, the same result is obtained. This is visualized using the coverage (the probability that the true value is inside the confidence region) which would be obtained when Wilks' theorem is applied. When Wilks' theorem is valid, the coverage corresponds to the size of

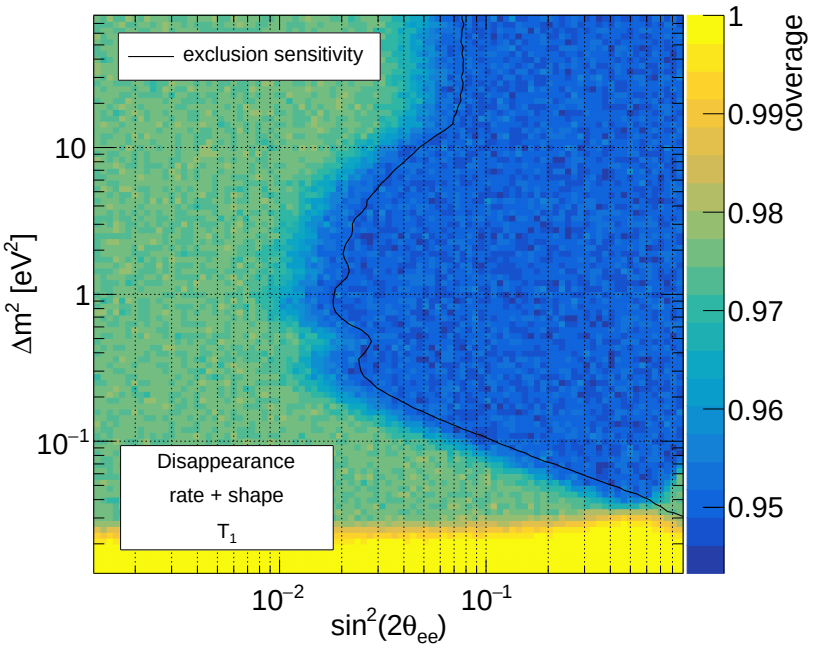


Figure 11.4: Coverage of T_1 The coverage is computed from toy-MC probability distributions (10^4 toy-MC data samples each) under the assumption of Wilks' theorem, where the nominal value is 0.95. In addition, the exclusion sensitivity at the 95 % CL is shown.

the nominal confidence level. In Figure 11.4 the coverage for a 95 % CL for the rate+shape analysis of the toy disappearance experiment is shown. It can be seen that the coverage is correct in the sensitive parameter region, and Wilks' theorem can be hence applied for T_1. However, the coverage is larger than the nominal value (overcoverage) in the parameter regions close to the physical border, as the border decreases the effective degrees of freedom and leads to non-normally distributed MLE-values. For the no-sterile hypothesis the distribution under the null hypothesis follows a $1/2\chi^2$ distribution [181] (see Figure 11.3a). This is a superposition of a χ^2 distribution and a Dirac delta function at zero, where the latter comes from the toy-data samples with the no-oscillation hypothesis as the best-fit hypothesis ($\widehat{\sin^2(2\theta)} = 0$). Such a distribution has a smaller critical value (2.7 instead of 3.84 for a 95 % CL) and applying Wilks' theorem rejects only 2.5 % of the probability distribution (i.e. the test size is 2.5 %). Thus, a confidence region based on Wilks' theorem contains the no-sterile hypothesis with a probability of 97.5 %, i.e. the coverage is 97.5 %. The probability distributions for hypotheses close to the border are a mixture of the $1/2\chi^2$ and the χ^2 distribution, and lead hence also to overcoverage. The overcoverage due to a physical border is present for both borders of $\sin^2(2\theta)$ at zero and at one (e.g. for $\Delta m^2 \sim 0.05\,\mathrm{eV}^2$). As there is no sensitivity for $\Delta m^2 < 0.03\,\mathrm{eV}^2$, the respective hypotheses will always be accepted and the coverage is 100 %.

The construction of the exclusion sensitivities relies on the probability distributions for the alternative hypothesis when the true hypothesis is the no-sterile hypothesis. Since the MLE-values of the no-sterile hypothesis are not normally distributed, the toy-MC probability distribution deviates from the asymptotic distribution, which is shown in Figure 11.3b. However, in the parameter region close to the sensitivity, differences between the two distributions can be observed above their median values which almost perfectly overlap. The values above the median correspond to the events with $\widehat{\sin^2(2\theta)} = 0$. Hence, the median of the distribution and thus the median exclusion sensitivity can be obtained using the Asimov-data set.

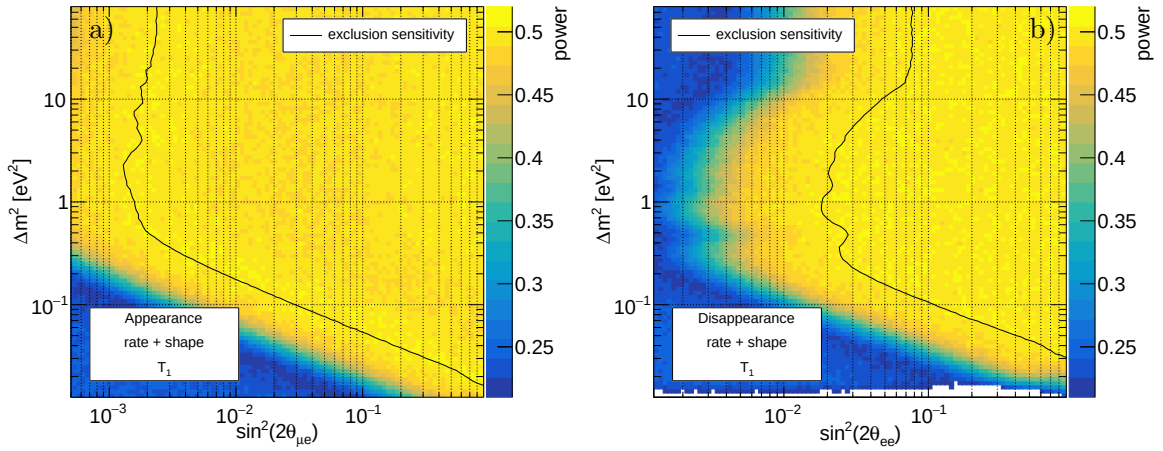


Figure 11.5: Power of T_1 The power is defined here as the fraction above the expected median value of the probability distribution when the no-signal hypothesis is true. The power is shown for the rate+shape analysis for the appearance (a) and disappearance (b) experiment and is computed from 10^4 toy-MC data samples each. The nominal value of the power corresponds to 0.5. In addition, the exclusion sensitivities at the 95 % CL are shown.

Note that the median of the asymptotic and toy-MC distribution does not overlap for all hypotheses. The distributions are compared using the power which would be obtained when the expected median value is used, i.e. the fraction of the toy-MC distribution above the asymptotic median. A power of 50 % indicates that the median of the toy-MC distribution agrees with the one from the asymptotic distribution. Figure 11.5 shows the power for the toy appearance and toy disappearance experiment in a rate+shape analysis. The median values match for the hypotheses close to the sensitivity, otherwise they can differ. As the distribution for the no-sterile hypothesis is given by a $1/2\chi^2$ distribution, the power is only half of the nominal value for the no-sterile hypothesis, namely 25 %. Similarly, the power is reduced for hypotheses close to the no-sterile hypothesis.
Similar results are obtained for the rate and shape analysis for both toy experiments. Even though the MLE-values are not normally distributed for all hypotheses, the median exclusion sensitivity can be obtained with the asymptotic distributions. This result agrees with previous studies from Reference [177] and was also recently observed in Double Chooz [209].

11.3 Probability Distributions of T_2

In the following, the asymptotic properties of the test statistic T_2 are discussed in detail (Section 11.3.2 and 11.3.3). If Wilks' theorem is valid, the probability distribution under the null hypothesis ($f\left(t_{(x,y)}|(x,y)\right)$) follows a χ^2 distribution with two degrees of freedom. However, Wilks' theorem requires some regularity conditions (see Chapter 3). These regularity conditions, as sketched in Figure 11.6, ensure normally distributed MLE-values and, in turn, the validity of Wilks' theorem. Similarly, the asymptotic probability distribution under the alternative hypothesis (e.g. $f\left(t_{(x,y)}|(0,0)\right)$) follows a non-central χ^2 distribution with two degrees of freedom with the non-centrality parameter estimated with the Asimov data set (see Formula 3.12) [180, 181].
As the sensitivities shown in the first part of this thesis rely on the asymptotic distributions, the following discussion focuses on the SOX experiment and the difference of the respective sensitivity is shown in Section 11.3.4. The chapter closes with the comparison to the obtained

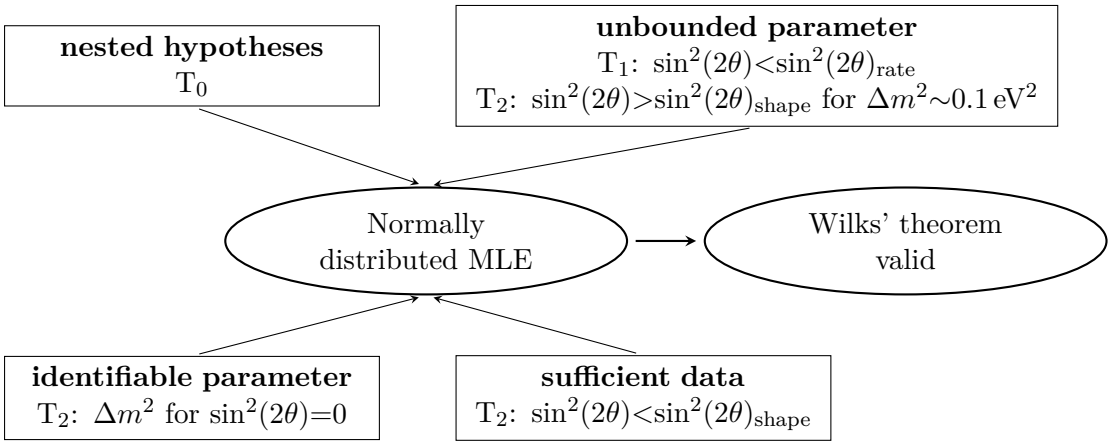


Figure 11.6: Requirements for the Validity of the Asymptotic Probability Distributions The requirements [216] (bold) and the specific case (test statistic and parameter region) when the requirement is not fulfilled are given in the boxes, where $\sin^2(2\theta)_{\text{rate/shape}}$ defines the sensitive $\sin^2(2\theta)$ value in a rate/shape analysis. A non-fulfilled requirement prevents from normally distributed MLE-values and thus the validity of Wilks' theorem.

results from Feldman and Cousins (Section 11.3.5). As the validity of the asymptotic formulae come typically along with normally distributed MLE-values, the distributions of the MLE-values as a function of the tested hypothesis are studied first.

11.3.1 Normality of Maximum-Likelihood Estimator Values

The proof of Wilks' and Wald's asymptotic approximations relies in a key-middle step on the MLE property of being normally distributed. As already discussed in Chapter 9, the distribution of the MLE-values under the no-signal hypothesis is affected by the presence of the physical border at $\sin^2(2\theta)$=0 and by the oscillatory nature of the signal leading to biased MLE, i.e. the best-fit hypothesis is likely not to be the true one.
To illustrate this bias, Figure 11.7 shows a pseudo-data sample from the no-sterile hypothesis of the SOX experiment as a function of the distance and the energy. In addition, the best-fit hypotheses for this pseudo-data sample are shown in a rate, shape and rate+shape analysis. While the best-fit hypothesis for the rate analysis is the true no-sterile hypothesis, the best-fit hypothesis deviates significantly from the no-sterile hypothesis in the shape and rate+shape analysis. Statistical fluctuations in the data can mimic an oscillation signature and therefore a sterile neutrino signature. This effect can be in particular observed in panel e) and f) of Figure 11.7. Note the constraint on the total event number in a rate+shape analysis typically reduces the size of the best-fit value of the oscillation amplitude $\sin^2(2\theta)$. Thus, in the present example, the MLE-value of $\sin^2(2\theta)$ is 0.03 in a rate+shape analysis and 0.07 in a shape analysis, respectively.
Normally distributed MLE-values often come along with a likelihood function that is itself asymptotically a Gaussian function [176]. Figure 11.8 shows the negative log-likelihood function for pseudo data from the no-sterile and a sterile neutrino hypothesis for the rate, shape, and rate+shape analysis, where the data sample from the no-sterile hypothesis corresponds to the one from Figure 11.7. In the rate analysis, the likelihood function is asymptotically Gaussian around the MLE-value, where all hypotheses with the same survival probability are MLE-values. As $\sin^2(2\theta)$ and Δm^2 are degenerate in a rate analysis, the likelihood function is only a one-dimensional Gaussian function of the survival probability (see Chapter 9).

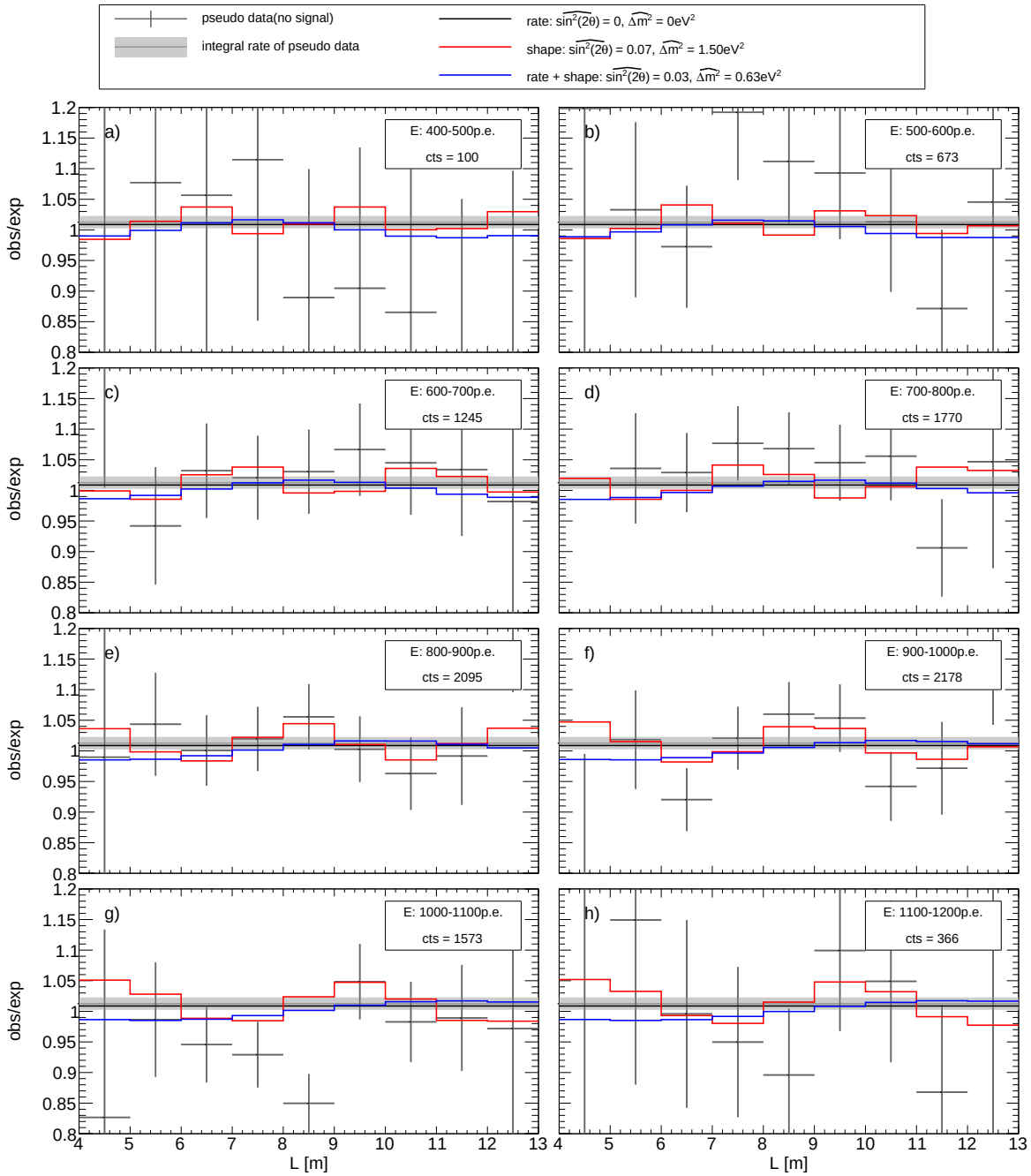


Figure 11.7: Pseudo Data for the SOX Experiment The ratio of the pseudo-data, which is generated from the no-signal hypothesis, to the expected data for the no-signal hypothesis is shown as a function of the distance for eight energy ranges, where the energy is reconstructed in the number of photoelectrons (p.e.). The ratio of the total event event number of the pseudo data to the total event number of the expected data is depicted in gray with the statistical 1σ error. The total event number per energy range is given in the respective panels. In addition, the best-fit hypothesis in a rate, shape, and rate+shape analysis are shown.

The likelihood function in the shape and rate+shape analysis is for all Δm^2-values a monotonic function in $\sin^2(2\theta)$ with a well-defined maximum. On the contrary, the likelihood function as a function of Δm^2 for a fixed $\sin^2(2\theta)$-value is a non-trivial function with several

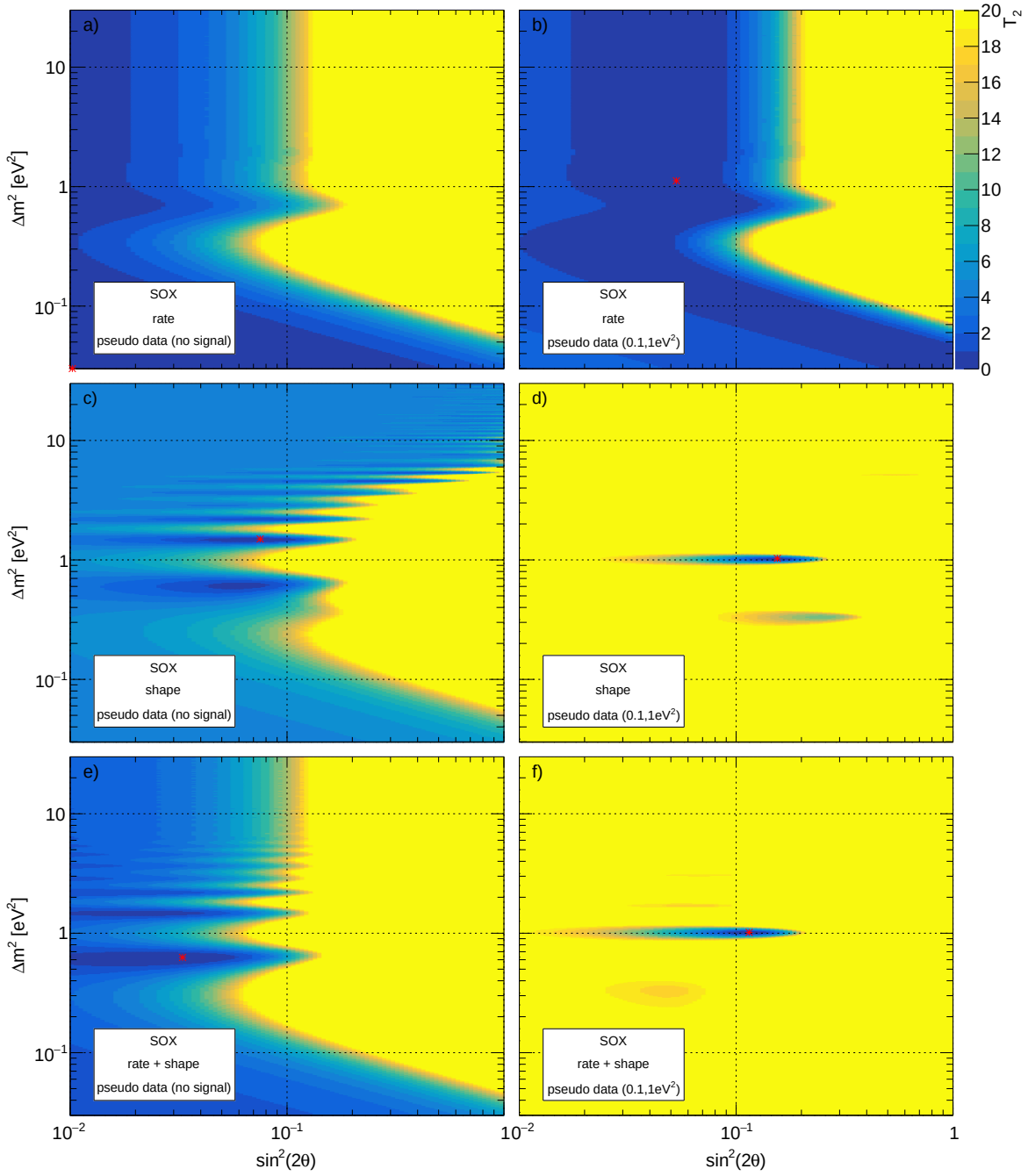


Figure 11.8: Negative Log-Likelihood Function for Pseudo Data in the SOX Experiment The negative log-likelihood function is shown in a rate (top), shape (middle) and rate+shape (bottom) analysis for the same pseudo data from the no-signal (left) and signal (right) hypothesis. The negative log-likelihood function is normalized to the absolute minimum and is cut at 20. The best-fit hypotheses are indicated in red. Note that in a shape and rate+shape analysis the likelihood function has several local minima in Δm^2.

local maxima. This behavior clearly deviates from a Gaussian function and hints at the non-validity of the asymptotic distributions. In addition, it complicates the minimization of the negative log-likelihood function that becomes challenging and time-consuming (see Appendix A.4). However, an almost Gaussian likelihood function is obtained for the data

sample of the signal hypothesis. The size of this signal is already larger than the statistical fluctuations in the data samples and the true hypothesis can be reconstructed. This implies that the bias and hence the normality of the MLE become a function of the oscillation hypothesis.
Figure 11.9 shows the MLE-values for different hypotheses in the rate, shape and rate+shape analyses. In a rate analysis with its effectively only one degree of freedom, the MLE-values are normally distributed around the $\sin^2(2\theta)$-Δm^2-values with the true survival probability (see top panel of Figure 11.9a). However, for a hypothesis close to the physical parameter border, theMLE-values are bounded and the MLE-values peak at the value of the parameter border (see the discussion of T_1 in Section 11.2). In a shape analysis, one can distinguish between four different parameter regions:

(A) parameter region below sensitivity ($\sin^2(2\theta)$<$\sin^2(2\theta)_{\text{shape}}$): the $\sin^2(2\theta)$-value is smaller than the size of the statistical fluctuations of the data samples. Hence, these hypotheses can typically not be reconstructed and the best-fit hypothesis is different from the true one which leads to biased MLE. Actually, the MLE-values are distributed close to the sensitivity of the shape analysis which is consistent with recent results from References [209–213]. Moreover, Reference [212] predicts the average value of the MLE of $\sin^2(2\theta)$ to $\langle\sin^2(2\theta)\rangle \approx 6.2/\sqrt{N_S^{\text{exp}}}$. Thus, a value of $\sin^2(2\theta) \approx 0.06$ is expected which agrees approximately with the observed values at $\sim 1\,\text{eV}^2$.

(B) parameter region above sensitivity ($\sin^2(2\theta)$>$\sin^2(2\theta)_{\text{shape}}$): the oscillation signature is larger than the statistical fluctuations and can be reconstructed. The MLE-values follow hence a two-dimensional normal distribution in $\sin^2(2\theta)$ and Δm^2.

(C) parameter region at sensitivity ($\sin^2(2\theta)$∼$\sin^2(2\theta)_{\text{shape}}$): the size of the statistical fluctuations are similar to the sterile neutrino signature. When the signature dominates over the statistical fluctuations the true hypothesis can be reconstructed, otherwise not. Therefore, the MLE-values are a combination between normally distributed values around the true hypothesis and randomly-distributed $\sin^2(2\theta)$-Δm^2-values in the sensitivity region (see parameter region (A)).

(D) parameter region at Δm^2∼$0.1\,\text{eV}^2$ and $\sin^2(2\theta)$> 0.5: for these hypotheses, the oscillation length becomes larger than the distance between the neutrino source and the detector and the integral survival probability and hence the sensitivity is approximately proportional to the product $\sin^2(2\theta)\times\Delta m^2$. Therefore, the variance of the MLE is large and the distribution of the MLE-values would exceed the physical border at $\sin^2(2\theta)$=1. Due to the physical border, the MLE-values are bounded and peak at $\sin^2(2\theta)$=1.

The rate+shape analysis adopts the features of the shape analysis. However, as the total event number is constrained, the bias and variance of the MLE are reduced. Moreover, the no-sterile hypothesis can become reconstructed as the best-fit hypothesis.

11.3.2 Null Distribution and Validity of Wilks' theorem

Figure 11.9 shows the probability distributions of T_2 obtained from MC constructions for different hypotheses in a rate, shape and rate+shape analysis. The rate analysis with its effectively one degree of freedom behaves similar to the test statistic T_1 which was discussed in Section 11.2. Hence, the probability distributions of the hypotheses close to the sensitivity can be described by one-dimensional χ^2 functions and Wilks' theorem can be used for a rate analysis. However, the physical border at $\sin^2(2\theta)$=0 reduces the effective degrees of

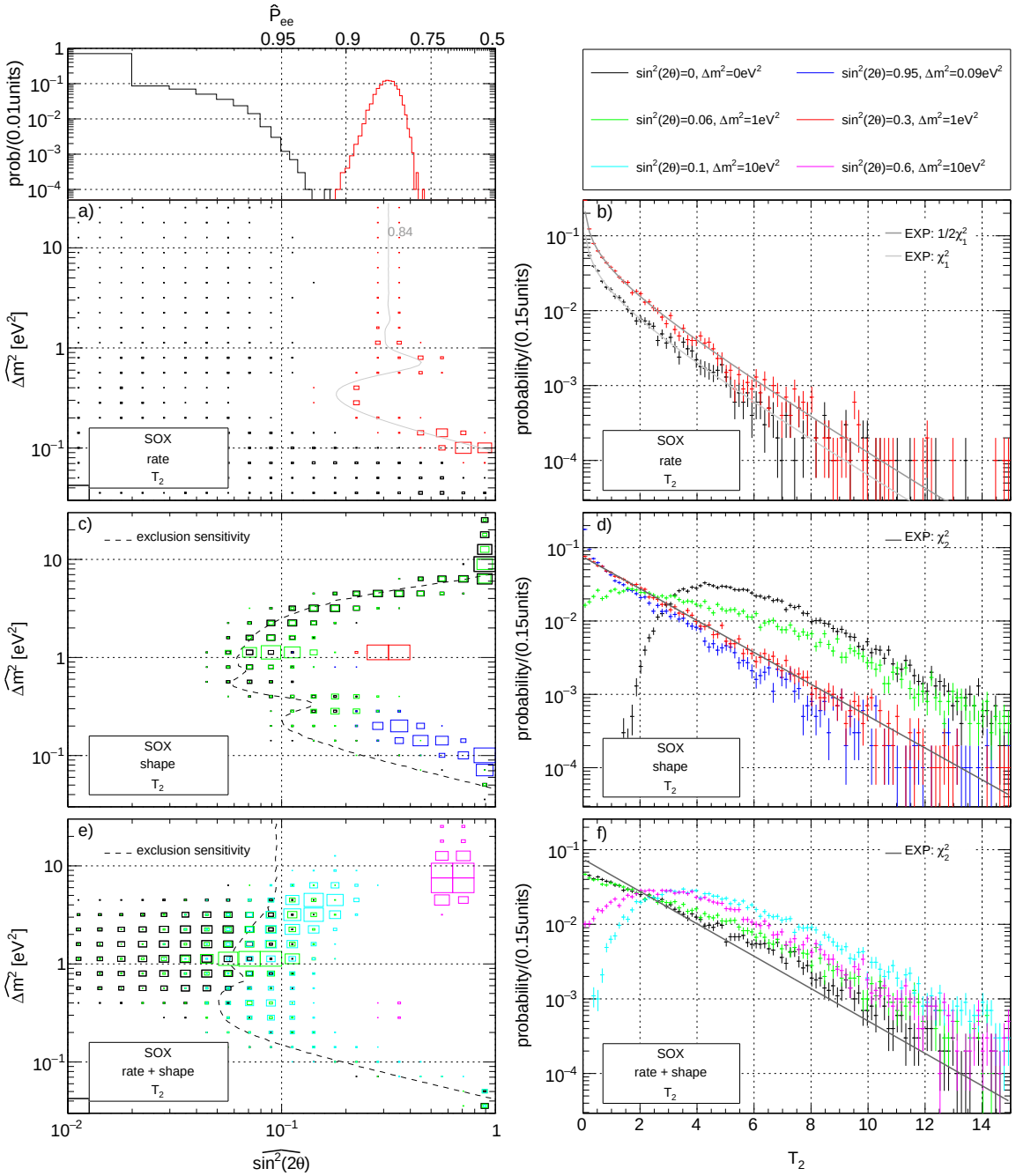


Figure 11.9: Maximum-Likelihood Estimator Values and Probability Distributions of T_2 for the SOX Experiment left panels: The MLE-values for the oscillation parameters are shown for different oscillation hypotheses in a rate, shape and rate+shape analysis and grouped into bins, where the size of the box correlates to the number of MLE-values in the specific bin. The no-signal hypothesis is filled into the smallest $\sin^2(2\theta)$-Δm^2 bin. The top left panel shows the projection of the MLE-values on the survival probability in the rate analysis. The $\sin^2(2\theta)$-Δm^2-points with the same survival probability of 0.84 is indicated in gray in panel a. In addition, the exclusion sensitivity at the 95 % CL is shown in panel c and e. right panels: The probability distributions under the null hypothesis in the rate, shape and rate+shape analysis are shown for the same oscillation hypotheses. The expected probability distributions according to Wilks' theorem are indicated in gray, where χ^2_n stands for a chi-square distribution with n degrees of freedom. The probability distributions can be described by the expected distributions, when the MLE-values are normally distributed. Note that redundant distributions are not shown: cyan and magenta are similar to black in panel c and d, blue and red of panel c and d are almost the same for panel e and f.

freedom of the analysis for hypotheses close to the border and the probability distribution of in particular the no-signal hypothesis becomes a $1/2\chi^2_1$ distribution.

In the shape analysis, the distributions depend on the true hypothesis and on the distributions of the MLE-values. Therefore, the same parameter regions as previously described can be identified. Given normally distributed MLE-values (parameter region (B)), the distributions of T_2 follow a χ^2 distribution with two degrees of freedom and Wilks' theorem is valid. The bias which is present for the no-signal hypothesis and the hypotheses in parameter region (A) and (C) increases the effective degrees of freedom of the analysis and shifts the distribution to larger test statistic values, where a larger bias results into a larger shift of the probability distribution. Hence, the critical value increases compared to Wilks' theorem and applying Wilks' theorem would result into an undercoverage of these hypotheses in the confidence region. The non-validity of Wilks' theorem for this parameter region can be traced back to the non-fulfilled requirement of sufficient data. Even though the event number per data bin is sufficient, the difference of the event number between the oscillation and no-oscillation hypothesis is too small. The shift of the probability distributions in particular for the no-oscillation hypothesis was also observed by several groups [71, 177, 183, 209–212, 217]. The hypotheses of the parameter region (D) are affected by the physical border which decreases the effective degrees of freedom. The probability distributions of T_2 are thus shifted to smaller test statistic values and the critical value becomes smaller in comparison to Wilks' theorem. Moreover, using Wilks' theorem would lead to an overcoverage of these hypotheses.

The rate+shape analysis is a combination of both features from the rate and the shape analysis. In particular, the parameter bound at $\sin^2(2\theta)$=0 counteracts the bias and reduces the shift of the test statistic distribution towards larger test statistic values for hypotheses in parameter region (A). However, the probability distributions of the rate+shape analysis cannot be described by a simple combination of the two single analyses.

In order to show a comprehensive comparison to Wilks' theorem for all hypotheses, the coverage and the critical values for a 95 % CL is depicted in Figure 11.10 for the SOX experiment in a rate, shape, and rate+shape analysis (see Figure A.10 for the toy experiments). In comparison to a critical value of 5.99, the critical values vary from ~ 3 up to ~ 10 and the coverage from ~ 60 % up to ~ 97.5 %. Again, the previously defined parameter regions can be identified. The hypotheses at the sensitivities in a shape and rate+shape analysis are in parameter region (C) and still affected by the biased MLE. Therefore, the critical value is larger and the coverage is smaller than expected from Wilks' theorem. Wilks' theorem can hence not be applied in these analyses. It is worth to mention that the critical values at the sensitivity for a rate+shape analysis are not constant which affords special attention when confidence regions are reconstructed.

Figure 11.11a compares the exclusion sensitivity obtained from a MC construction to the one assuming Wilks' theorem in the rate+shape analysis of the toy disappearance experiment. The sensitivity with Wilks' theorem overestimates the actual sensitivity by almost a factor of two, due to the previously discussed undercoverage. In the shape analysis the sensitivity is even more overestimated (not shown here). Therefore, reconstructed confidence regions using Wilks' theorem overestimate the exclusion limits, or can even lead to the rejection of the no-signal hypothesis even though it is true.

The probability of rejecting the no-signal hypothesis (i.e. p-value) and thus claiming a signal is connected to the coverage of the no-signal hypothesis, which is listed in Table 11.1 for all studied experiments at the 95 % CL. Hence, instead of 5 %, this probability is 23 % in the rate+shape analysis and 35 % in the shape analysis for the toy disappearance experiment. These probabilities and therefore the overestimation of the significance applying Wilks' theorem are illustrated as a function of the significance under Wilks in Figure 11.11b for the SOX

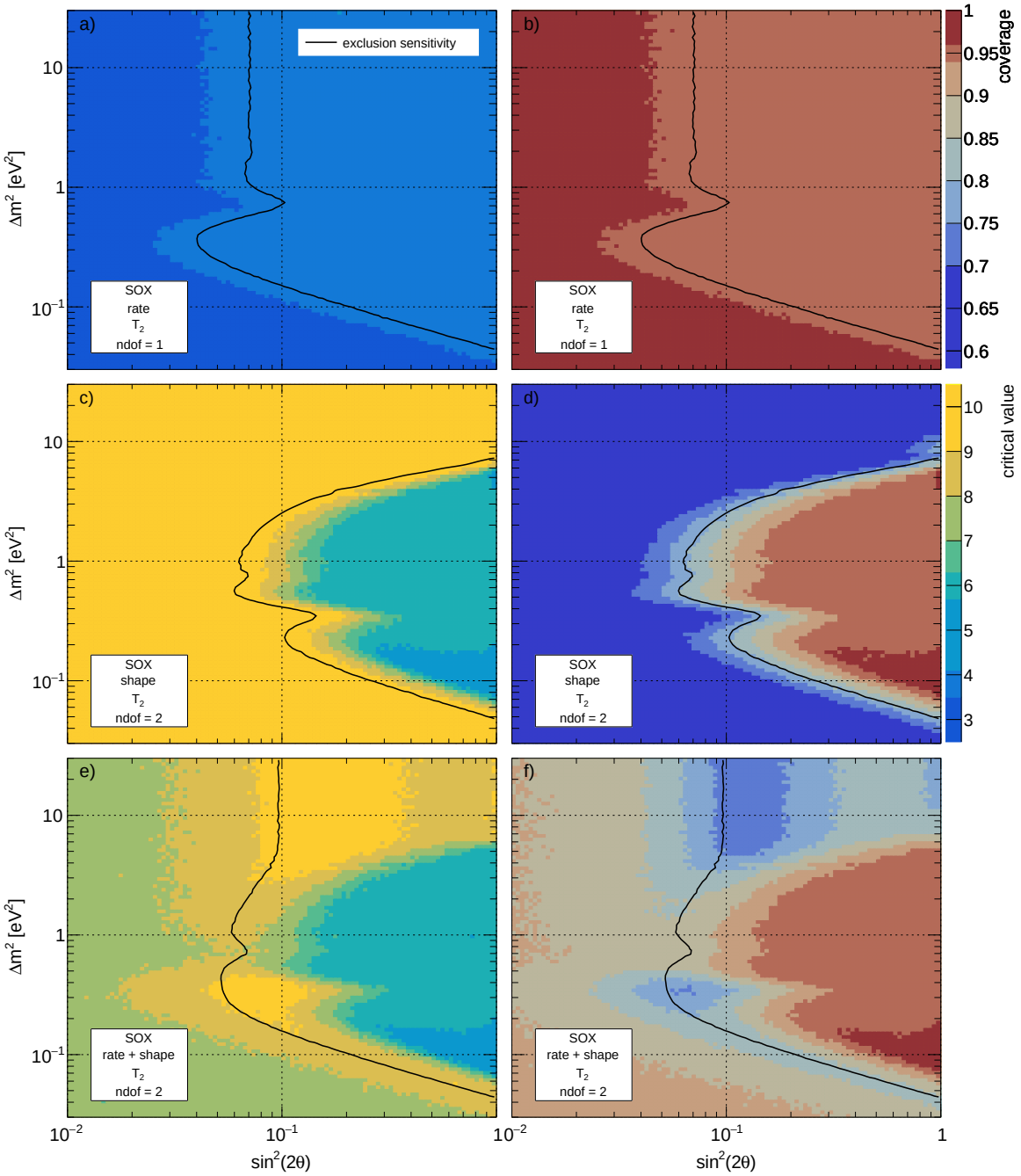


Figure 11.10: Critical Values and Coverage of T_2 for the SOX Experiment The critical values (left) and the coverage using Wilks' theorem (right) are shown at the 95 % CL in a rate, shape, and rate+shape analysis, where the top (bottom) color scale belongs to the coverage (critical value). The nominal value for the critical value is 3.84 and 5.99 for one and two degrees of freedom (ndof), respectively. The exclusion sensitivities at the 95 % CL are shown in black. The assumed degrees of freedom of the χ^2 functions are defined as the number of uncorrelated parameters in the respective analysis.

and the toy experiments. The significance is overestimated for all studied experiments up to the $\sim 1\sigma$ level. For instance, a 3σ significance assuming Wilks corresponds actually only to a 2.3σ-level for the shape analysis in the toy disappearance experiment. Equivalently, the

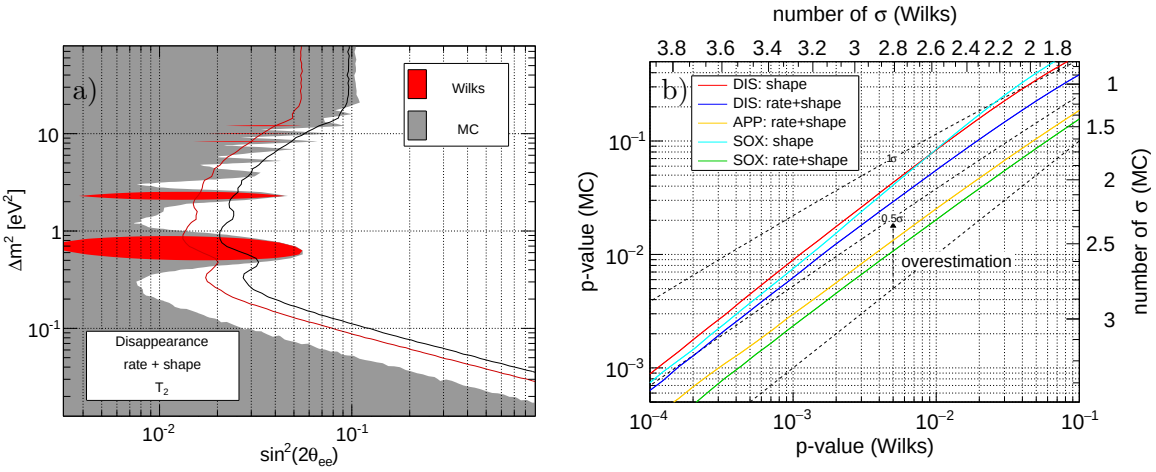


Figure 11.11: Validity of Wilks' Theorem a) The exclusion sensitivities at the 95 % CL based on Wilks theorem and the full MC construction are shown for the toy disappearance experiment in a rate+shape analysis. In addition, the reconstructed confidence regions at the 95 % CL from a pseudo-data sample from the no-sterile hypothesis is shown for Wilks' theorem and the full MC construction. b) The p-value reconstructed from 10^6 pseudo-data samples is shown as a function of the significance when Wilks' theorem is assumed. The significance is shown for the rate+shape analysis of the toy appearance experiment, and the shape and rate+shape analysis of the toy disappearance and SOX experiment, respectively. The right and top axis transfer the p-value to the corresponding σ-level. The dashed lines indicate whether the p-value using Wilks' theorem is consistent (bottom) or overestimated by 0.5σ (middle) and 1σ (top), respectively.

probability to observe a signal with a significance of 3σ or larger with Wilks, is 2.8 % instead of 0.27 %. The overestimation depends on the experiment and might be correlated to the size of the statistical and systematic uncertainties on the total event number (see Table 11.1). Such a correlation would require further investigation. Indeed, a similar hint for a correlation was recently published by Reference [212].

11.3.3 Alternative Distribution and the Asimov Data Set

According to Wald [180] and Reference [181], the alternative distribution follows asymptotically a non-central χ^2 distribution with the degrees of freedom of the analysis. The non-centrality parameter can be determined from the test statistic value of the Asimov data set, i.e. the data set without statistical fluctuations (see Chapter 3.4). In the following, the no-sterile hypothesis is assumed as the true hypothesis, as these distributions (mainly the median) are needed for the determination of the exclusion sensitivities.

Similar to the validity of Wilks' theorem, this approximation holds when the MLE-values are normally distributed. However, as already discussed, this is not the case for the no-sterile hypothesis. The alternative distributions under the no-signal hypothesis of T_2 are shown in Figure 11.12 for the SOX experiment for a rate, shape, and rate+shape analysis when the hypothesis with $\sin^2(2\theta)$=0.1 and Δm^2=1 eV^2 is tested. All three distributions differ from the asymptotic distributions. However, in the rate analysis, the distribution is only altered above the median value and the median values of both distributions are consistent. This is the same behavior that was observed for the distributions of T_1 (see Section 11.2). The distributions of the shape and rate+shape analysis differ significantly due to the biased MLE and are shifted towards larger test statistic values in comparison to the expected distributions. As the bias

Table 11.1: Difference between Asymptotic and Toy-MC Probability Distributions The differences in the specific analyses for the SOX, toy appearance, and toy disappearance experiment are listed in the individual columns. The coverage of the null distributions for the no-oscillation hypothesis are listed in the first row assuming a nominal value of 95 %. The comparison for the alternative distribution is given by the difference in the median values at the sensitivity in the second row. In addition, the ratio of the statistical to the systematic uncertainty on the total event number for the no-oscillation hypothesis is quoted in the third row.

	SOX		toy appearance	toy disappearance	
	shape	rate+shape	rate+shape	shape	rate+shape
Coverage of H(0,0)	0.60	0.92	0.89	0.65	0.77
ΔMedian at sensitivity	4.5	1	1.4	4.1	3
$\sigma_{\text{stat}}/\sigma_{\text{syst}}$ of H(0,0)	0	0.67	0.64	0	0.16

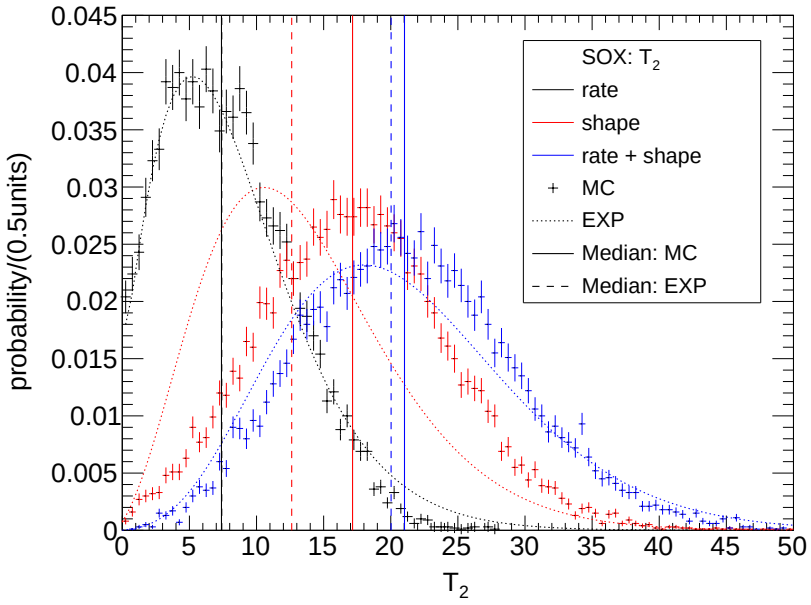


Figure 11.12: Alternative Distribution of T_2 for the SOX Experiment The distributions and the respective median values are shown in a rate, shape and rate+shape analysis for the true no-sterile hypothesis and tested hypothesis with $\sin^2(2\theta)$=0.1 and Δm^2=1 eV^2. The expected distributions and their median values are indicated with the dashed lines. The median values overlap in the rate analysis.

is larger for the shape analysis, the difference between the median values is also larger in the shape analysis than in the rate+shape analysis.

It can be observed that the difference between the median obtained from a MC construction and the expectation changes as a function of the tested hypothesis, because of the physical border at $\sin^2(2\theta)$=0. The difference for the tested hypotheses close to the the physical border is reduced, as it is shown in Figure 11.13 for the toy experiments in a rate+shape analysis. The difference of the medians at the sensitivities are listed for the studied experiments in Table 11.1. Due to the bias, all values are positive, i.e. the values obtained through the MC constructions are larger than the ones from the expectations, where the size of the difference correlates to the bias of the MLE of the no-signal hypothesis.

Instead of the (almost) constant shift, one can compare the distributions in terms of the power, i.e. the fraction of the alternative distribution above the expected median. Fig-

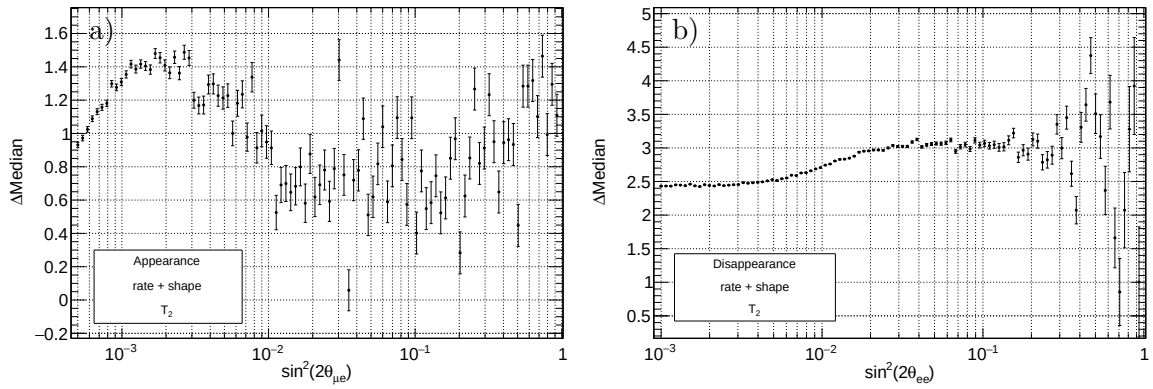


Figure 11.13: Difference of the Median Values of the Asymptotic and MC Alternative Distributions of T_2 The difference is shown as a function of $\sin^2(2\theta)$ for Δm^2=1 eV^2 for the toy appearance (a) and toy disappearance (b) experiment in a rate+shape analysis. The statistical error increases with increasing $\sin^2(2\theta)$-value, as the absolute test statistic values become larger.

ure 11.14 shows the power in a rate, shape, and rate+shape analysis of the SOX experiment for the full oscillation parameter space (the power for the toy experiments can be found in Appendix A.6 Figure A.12). Due to the positive shift in the median, the power is larger than the nominal value of 0.5 for the hypotheses close to the sensitivity. Thus, the sensitivity would be underestimated when the Asimov data set is used. Despite the constant shift in the medians, the power reaches for large $\sin^2(2\theta)$-values the desired power of 0.5. This feature is simply given because of the increasing values of the test statistic, where the shift of the median becomes comparably small.

11.3.4 Comparison of SOX Sensitivities

In the first part of this thesis, the exclusion sensitivities for the SOX experiment were computed based on the asymptotic distributions. Here, these sensitivities are compared to those obtained from a full MC construction, which is shown in Figure 11.15. The comparison assumes 10^4 expected events for the no-oscillation hypothesis with an uncertainty of 1.5 %. Further, the optimized selection cuts are applied (see Chapter 5) and no background is taken into account. The data is grouped in total in 72 bins in distance (9,4 m,13 m) and energy (8,400 p.e.,1200 p.e.). In the following, the three different kinds of analyses are discussed separately.

As discussed in the previous sections, the critical and median values of a rate analysis agree in the parameter region of the sensitivity with the expected values from the asymptotic distributions. Hence, the approximations can be used in a rate analysis. However, in the SOX sensitivities displayed in Figure 7.12, the degeneracy of $\sin^2(2\theta)$ and Δm^2 is not taken into account and two degrees of freedom are assumed in the analysis. Hence, the rate sensitivity of is underestimated by about 20 %.

The shape sensitivity is underestimated with the asymptotic distributions by about 8 %. However, the difference is comparably small, as the probability distributions under the null and the alternative hypothesis are shifted in a similar way towards larger test statistic values. The underestimation can be explained with the bias that is larger for the no-sterile hypothesis than for a hypothesis close to the sensitivity, which results into a relatively larger shift for the median values than for the critical values. In other words, given the no-signal hypothesis is true, the best-fit hypothesis is a random oscillation hypothesis. However, the best-fit

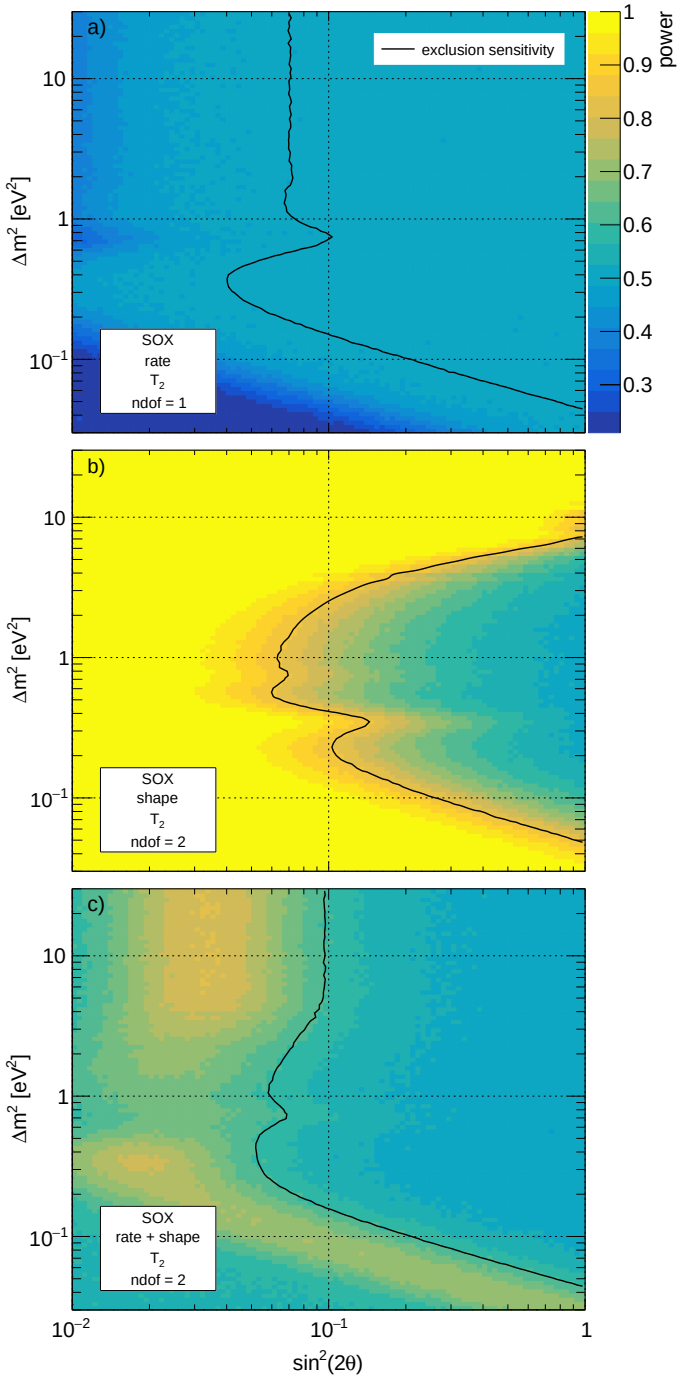


Figure 11.14: Power of T_2 for the SOX Experiment The power is computed here as the fraction of the alternative distribution above the expected median value for the rate, shape, and rate+shape analysis. The nominal value corresponds to 0.5. In addition, the exclusion sensitivities at the 95 % CL are shown.

hypothesis is unlikely the specific tested hypothesis. Thus, as the local maxima in the likelihood function have typically large values, the chance to reject the tested hypothesis is higher than expected and the sensitivity effectively increases.

Although the sensitivities of the rate and the shape analyses are underestimated, the sensitivity for the combined rate+shape analysis is actually overestimated by 15 %. However, the

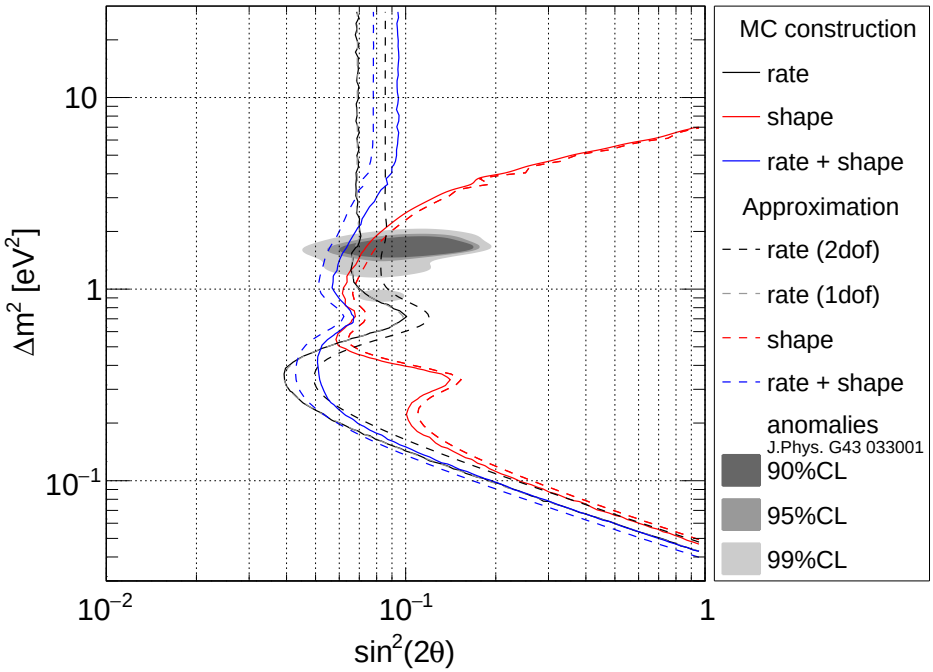


Figure 11.15: Comparison of the Sensitivities for the SOX Experiment The sensitivities are shown for the 95 % CL in the rate, shape and rate + shape analysis based on a full MC construction (solid) and the asymptotic probably distributions (dashed). The sensitivities assume an expected event number of 10^4 events with an uncertainty of 1.5 %. In addition, the oscillation parameter space allowed by the anomalous neutrino oscillation data is indicated in gray. Note that the gray dashed line overlaps with the black solid line.

precise value of the overestimation depends on the Δm^2-value. The overestimation of the sensitivity comes from the fact that the bias of the no-sterile hypothesis is reduced in comparison to a hypothesis close to the sensitivity, as the border at $\sin^2(2\theta)$=0 reduces the effective degrees of freedom for the no-sterile hypothesis in a rate+shape analysis. This results into a relatively larger shift for the critical values than for the median values. In comparison to the shape analysis, the best-fit hypothesis is still often a random oscillation hypothesis, if the no-signal hypothesis is true. However, the constraint on the total event number constrains the parameter space of the MLE-values as well. Therefore, the difference of the best-fit hypothesis and the specific tested hypothesis (close to the sensitivity) is small. Thus, the chance to reject the tested hypothesis reduces and becomes even smaller than expected from the asymptotic distributions. This results into an effective decrease of the sensitivity.

Note that the differences in the sensitivities depend on the specific experimental parameters. The comparison of the sensitivities of the toy experiments can be found in Appendix A.6. Even though the sensitivities deviate, the assumption of the asymptotic distributions is a suitable and convenient tool to study systematic uncertainties of an experiment and the respective relative changes of the sensitivities. However, a full MC construction has to be performed for a final analysis.

11.3.5 Comparison to the Results from Feldman and Cousins

The analysis with T_2 using a toy-MC construction is based on the method of Feldman and Cousins [177]. In order to validate and compare the presented results, the toy experiment

from Reference [177] is studied, namely an appearance experiment with $L = 600 - 1000\,\mathrm{m}$ and $E = 10 - 60\,\mathrm{GeV}$. The data is grouped into five energy bins, where the expected background event number per bin is 100 ($N_B^{\mathrm{exp}} = 500$). The expected neutrino event number is $N_S^{\mathrm{exp}} = 50000$ for an appearance probability of one. Further, N_S^{exp} and N_B^{exp} are assumed to be known and fixed in the likelihood function.

The coverage at the 90 % CL is compared to the results from Reference [177] in Figure 11.16. The overall agreement is very good and the same parameter regions of under- and overcoverage can be identified. Moreover, the absolute values of the coverage vary from 75 % to 96 % which is consistent with the range of 76 % to 94 % from Reference [177]. However, the parameter region with an overcoverage at $\Delta m^2 \sim 100\,\mathrm{eV}^2$ and $\sin^2(2\theta) > 0.1$ reported by Feldman and Cousins is not observed with the present analysis. The overcoverage could be explained by an artifact in the analysis introduced by the minimization algorithm of the negative log-likelihood function. In that parameter region, the oscillation signal can be well reconstructed and the MLE-values are normally distributed with a small variance around the true values. As the minimization of the negative log-likelihood function is not trivial, the minimization algorithm typically scans over a set of Δm^2-values at which the negative log-likelihood function is minimized over $\sin^2(2\theta)$. Thus, the absolute minimum can be missed if the Δm^2-steps are not fine enough. Therefore, effectively only one parameter ($\sin^2(2\theta)$) is minimized and the probability distributions have only one degree of freedom which leads to an artificial overcoverage. To properly reconstruct the absolute minimum of Δm^2 in that parameter region, additional Δm^2-raster points around the specific tested hypothesis are defined, which is described in Appendix A.4.

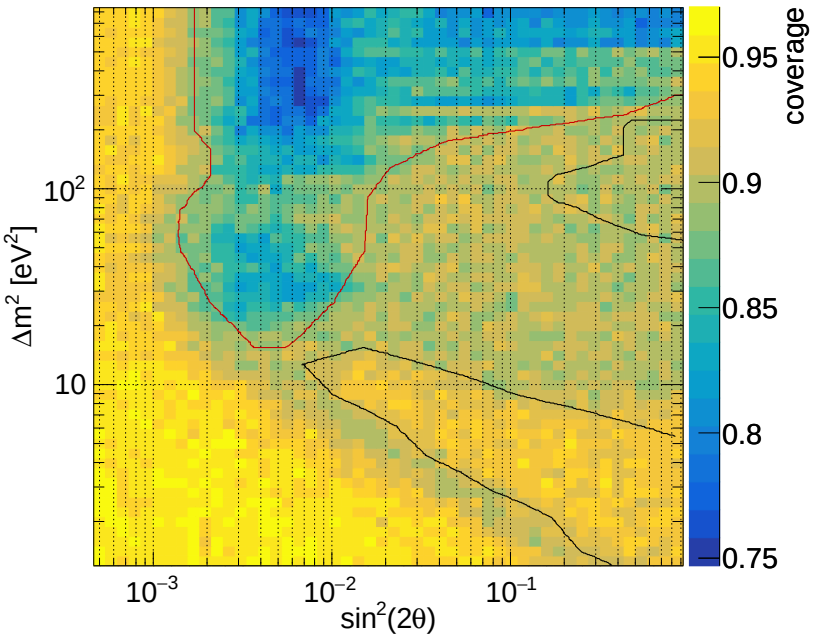


Figure 11.16: Comparison of the Coverage to Feldman and Cousins The coverage is computed at the 90 % CL for the toy experiment defined in Reference [177] using 10^3 toy-MC samples at each raster point. The lines indicate the results of Reference [177], where regions of overcoverage are drawn in black and regions of undercoverage in red.

11.4 Estimation of the P-Value

In addition to the previously discussed bias, another requirement for Wilks' theorem is not fulfilled for T_2 under the no-sterile hypothesis. Namely, the non-identifiability problem [218] exists when a (nuisance) parameter is present only under the alternative hypothesis. In sterile neutrino searches, Δm^2 is not defined for the no-signal hypothesis, as it can have any value when $\sin^2(2\theta)$=0. As the no-sterile hypothesis is of special interest for claiming a discovery, the distribution has to be known with high precision. For example, claiming a discovery at 5σ would require a toy-MC with the statistics of the order of 10^7 [176], where the p-value of the no-signal hypothesis and thus the significance can be computed from the probability distribution:

$$p(T_2) = \int_{T_2^{\text{obs}}}^{\infty} f(T_2|(0,0)). \tag{11.3}$$

As the computation of $f(T_2|(0,0))$ is computationally intensive, alternative procedures have been developed to evaluate the significance of a discovery. In the following, the estimation with the Gross-Vittels correction and a recently published method based on fitting Gaussian white noise are applied to the toy disappearance and the SOX experiment.

11.4.1 Gross-Vittels Correction

The Gross-Vittels correction [219] is a common procedure performed in high energy physics, which can give an upper bound on the tail probability of the test statistic distribution, i.e. the p-value. This procedure corrects the underestimated p-value obtained with T_1 (local p-value) for the fact that a sterile neutrino could have any Δm^2-value and is called the look-elsewhere effect.
Thus, the upper bound for the global p-value can be approximated with

$$p(T_2) \approx \min_{\Delta m^2} p(T_1) + \langle N_c \rangle, \tag{11.4}$$

where $\langle N_c \rangle$ is the mean number of the so-called upcrossings above the level c. Each upcrossing corresponds to a Δm^2-value for which the respective signal hypothesis is preferred over the no-signal hypothesis at a certain level c. The number of upcrossings $\langle N_c \rangle$ can be estimated from a small number of MC simulations ($\mathcal{O}(100)$) for a low level c_0 with $c_0 \ll c$ to

$$\langle N_c \rangle = \langle N_{c_0} \rangle e^{-(c-c_0)/2}. \tag{11.5}$$

Figure 11.17a visualizes the determination of the number of upcrossings above the level c_0 for one data sample from the no-signal hypothesis.
As the no-signal hypothesis is on the boundary of the parameter space, the probability distribution under the no-signal hypothesis of T_1 follows a $1/2\chi^2_1$-distribution. The local p-value is hence given by (see also Section 11.2) [216]

$$p(T_1) = 0.5 \cdot P(\chi^2_1 > c). \tag{11.6}$$

This procedure has been tested for the sterile neutrino search with the toy disappearance experiment. The p-value as a function of the test-statistic value is shown in Figure 11.17b for the MC construction and the Gross-Vitells correction. The estimation overestimates the actual p-value with a factor of ~ 1.5 (~ 1.6 for the shape analysis). This can be traced back to the in turn overestimated number of upcrossings. The individual upcrossings are correlated to each other because of the oscillation signature that produces a harmonic series in the likelihood function of Δm^2 (see for example Figure 11.17a). If it would be possible

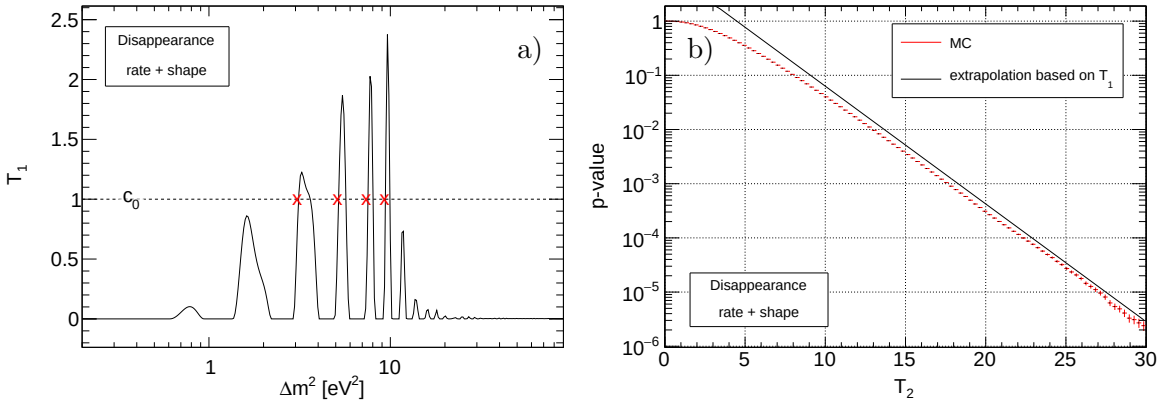


Figure 11.17: Gross-Vitells Correction a) Example of the number of upcrossings (red crosses) from a pseudo-data sample from the no-signal hypothesis. The y-axis gives the value of the test statistic T_1, when the no-sterile hypothesis is tested. b) The p-value obtained from a MC construction of T_2 (red, 10^8 toy-MC samples) is compared to the estimation of the global p-value using the Gross-Vitells correction (black) as a function of the observed test statistic value of T_2. The mean number of upcrossings is determined from 1000 MC samples and the level $c_0 = 1$ to $\langle N_u \rangle = 5.6$. Due to the correlation of the individual upcrossings the estimation of the p-value is too high.

to determine the uncorrelated number of upcrossings, this procedure might be an attractive tool. However, first attempts to determine the fundamental frequencies in Δm^2 did not produce the desired results, and further studies are required.

11.4.2 Fitting Gaussian White Noise

Reference [212] reinterprets the minimization of short-baseline experiments, in particular disappearance experiments, with the problem of fitting Gaussian white noise. The authors show that the distribution of the square-root of the test statistic T_2 under the no-oscillation hypothesis can be described (under some assumptions) as the one of the maximum of N standard normal random variables, where N denotes the effective number of bins. Such a distribution can be easily obtained from numerical simulations and is defined as "max. Gauss distribution" in Reference [212]. Moreover, for $N \to \infty$ an analytical expression for the cumulative probability distribution of $\sqrt{T}$ is given:

$$F(x) = \exp\left\{-\exp[-A_N(x - B_N)]\right\}, \tag{11.7}$$

where

$$A_N = \sqrt{2\log N} \quad \text{and} \quad B_N = A_N - \frac{\log\log N + \log 4\pi}{2A_N}. \tag{11.8}$$

Figure 11.18 compares the p-value as function of $\sqrt{T_2}$ obtained from the MC construction to the ones using Equation 11.7 for the toy disappearance and the SOX experiment in a shape and rate+shape analysis. The distributions in the shape analysis can be well described by the estimations up to large p-values of ~ 0.02 and ~ 0.07 for the toy disappearance and SOX experiment, respectively. The enhanced agreement in the toy disappearance experiment could be traced back to a better fulfillment of the assumptions for the estimation. Namely, the estimation requires the same value for the statistical error of each bin, which is indeed given by the flat energy spectrum of the disappearance-experiment. From Figure 1 of Reference [212],

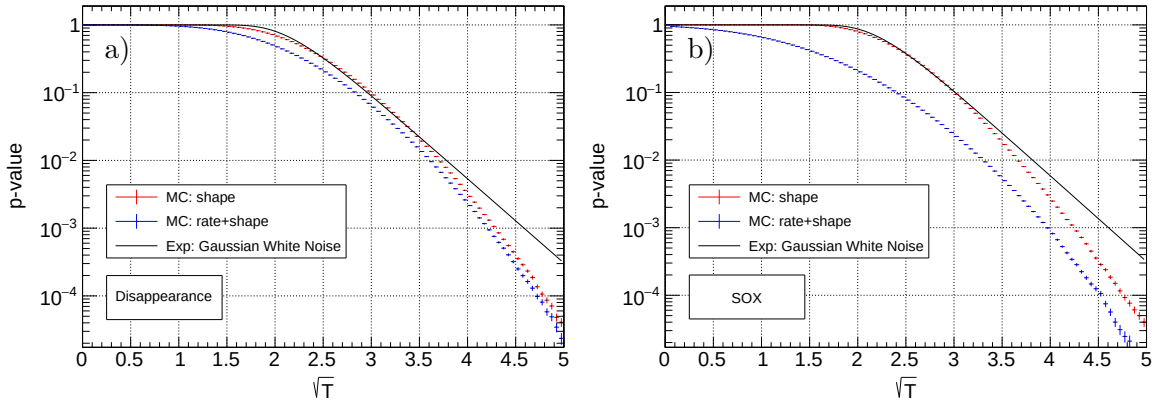


Figure 11.18: Estimation of P-Value P-value as a function of the square-root of the test statistic value T_2 for the a) disappearance and b) SOX experiment. The p-value is computed from 10^6 toy-MC samples for the shape (red) and rate+shape (blue) analysis each. The estimation is based on Equation 11.7 using $N = 60$ and $N = 72$ for the disappearance and SOX experiment, respectively.

it looks promising that the agreement for large p-values is even more improved using the "max. Gauss distribution".
While the approximation holds especially in the shape analysis, the effective number of bins is reduced in a rate+shape analysis. This implies that the p-value is overestimated with the estimation in a rate+shape analysis. The overestimation depends on details of the experimental parameters, such as the ratio of the statistical to the systematic uncertainty on the total event number (see also Table 11.1). The obtained results agree and confirm the ones of Reference [212]. Thus, this estimation is a promising method to determine the p-value in a shape analysis which can be especially useful for the current reactor-based experiments.

11.5 Goodness-of-Fit Test

A goodness-of-fit test returns a measure of the discrepancy of the data and the model [176], where the assumed model (here the (3+1) model) is tested against any other model. As introduced in Chapter 3.5, the probability distributions of the goodness-of-fit test statistic (see Formula 3.13) follows asymptotically a χ^2 distribution, where the degrees of freedom depend on the number of bins, free parameters, and the presence of auxiliary-measurement terms in the likelihood function [182]. For example, for the toy experiment 58 and 57 degrees of freedom are expected in a rate+shape and shape analysis, respectively (see Chapter 3.5). Note, as the asymptotic distributions follow from Wilks' theorem, normally distributed MLE-values are required that the approximation holds.
In the following, the properties of the goodness-of-fit test for the toy disappearance experiment is studied. Figure 11.19 shows the probability distribution of the goodness-of-fit test statistic when the no-signal hypothesis and a sterile neutrino hypothesis with $\sin^2(2\theta)$=0.1 and Δm^2=1 eV^2 is true. When the latter is true, the asymptotic distributions are valid, as the MLE-values are normally distributed. Moreover, the difference of one degree of freedom is visible for the shape and rate+shape analysis.
Instead, when the no-signal hypothesis is true, the MLE-values are not normally distributed and the bias of the MLE shifts the probability distribution to smaller test statistic values.[3]

[3]The absolute values of the global minima of the negative log-likelihood function are smaller than expected for normally distributed MLE-values.

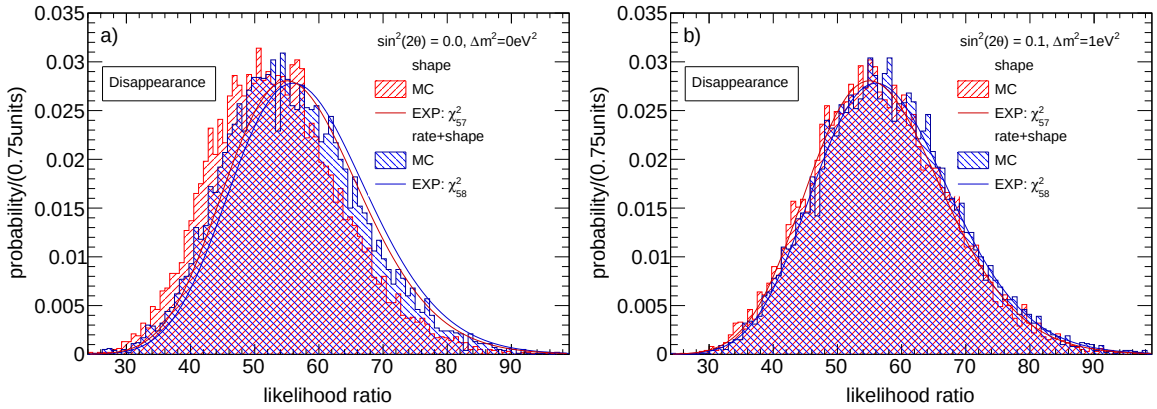


Figure 11.19: Goddness-of-Fit Distribution of the goodness-of-fit test statistic, when the true hypothesis is the no-sterile hypothesis (a) and a sterile neutrino hypothesis (b). The lines indicate the expected distributions assuming Wilks' theorem. The degrees of freedom of the χ^2 distribution is reduced by one in the shape analysis, as the likelihood function does not contain the Gaussian constraint term on the total neutrino event number.

Table 11.2: Comparison of Goodness-of-Fit Parameters The distributions of the goodness-of-fit test are compared for the toy-MC construction and the asymptotic distribution for the toy disappearance experiment assuming the no-signal hypothesis is true. The first row lists the mean value of the respective distributions. The second row gives the p-value one would quote when the mean value of the MC distribution is observed.

	shape		rate+shape	
	MC	EXP	MC	EXP
mean	53.5	57	55.8	58
p-value at mean(MC) [%]	47.4	60.7	47.2	55.8

Due to the larger bias the shift becomes larger in the shape analysis. The goodness-of-fit values based on the asymptotic distributions are hence overestimated, as listed in Table 11.2. Even though the asymptotic distributions are not valid for the no-sterile hypothesis, the difference of the goodness-of-fit values is comparably small due to the large total number of degrees of freedom. Thus, quoting a goodness-of-fit value based on the asymptotic distributions might be applicable, as these results are in addition more conservative.

In this chapter, the asymptotic distributions of the test statistics T_0, T_1, T_2, and the goodness-of-fit test were extensively discussed. While the asymptotic distributions are valid for T_0 and T_1, they do not apply for T_2. Thus, the usage of Wilks' theorem with T_2 causes too strong limits and overestimates the significance which can even lead to a wrong claim of a signal. In addition, the sensitivities reconstructed from the Asimov data set in combination with Wilks' theorem differ from the true sensitivities. Two estimations for the p-value of the no-oscillation hypothesis were discussed. First, the Gross-Vitells correction was tested for sterile neutrino experiments. It was found that the correction does not apply and that the p-value is overestimated, i.e. the significance is underestimated. Second, the recently published method on fitting Gaussian white noise gives promising results for a shape analysis. However, the p-value in a rate+shape analysis is as well overestimated. Finally, the degrees of freedom in a goodness-of-fit test are likely to be underestimated which results, however, only into an overestimated and more conservative goodness-of-fit p-value.

12 Summary

A vast experimental program is currently ongoing to test the existence of a sterile neutrino at the eV-scale, which is motivated by anomalous results from neutrino oscillation experiments. This hypothetical, beyond the standard model, particle cannot interact weakly. However, due to its mass, it mixes with the active neutrinos and can thus be observed indirectly. This can be realized by a disappearance or an appearance experiment, where the sterile neutrino mixing leads either to a reduction of the initial neutrino flux or the appearance of a different neutrino flavor. In both cases, an oscillation signature could be observed, which cannot be explained by the standard 3-neutrino model. This thesis covers two aspects of the sterile neutrino search. The first part discusses the data analysis of the SOX project, which is one specific experiment with the goal of testing the sterile neutrino hypothesis. The second part evaluates the non-trivial statistical features of the sterile neutrino search and compares the analysis methods applied in the current short-baseline (SBL) experiments.

12.1 Data Analysis for the SOX Project

The source-based SOX project is a disappearance experiment and exploits the solar neutrino detector Borexino at the Laboratory Nazionali del Gran Sasso (LNGS) in Italy. Borexino is a 270 t-liquid scintillator detector with an unprecedented low background level. The ^{144}Ce-^{144}Pr-source, with an activity in between 100 kCi and 150 kCi, produces electron antineutrinos ($\bar{\nu}_e$) up to 3 MeV via β-decays and is located 8.5 m below the center of the spherical Borexino detector. As the $\bar{\nu}_e$ are detected via the inverse beta decay (IBD) and the Borexino detector can reconstruct the deposited energy and interaction position of the neutrinos, an oscillation pattern induced by active to sterile neutrino oscillations can be observed. For instance, a sterile neutrino with a mass of ~ 1 eV results into an oscillation length of ~ 1 m which can be directly resolved in the detector and would be a smoking gun signature. Unfortunately, the SOX project was canceled in December 2017 due to problems in the source production. However, the work performed in this thesis had, on the one hand, a direct impact on the understanding of the Borexino detector itself and enhanced in particular the $\bar{\nu}_e$-analyses in Borexino [35, 51]. On the other hand, these results will be relevant for current or future experiments, especially the proposed ^{144}Ce source-based experiment JUNO [42].
This thesis focuses on the data analysis of SOX and a large fraction of the analysis software was developed or improved in order to predict the expected signature, select the IBD events, and produce a confidence region in the parameter space of the oscillation parameters $\sin^2(2\theta)$ and Δm^2.
The expected signal, including the detector response and detection efficiency, was discussed. Moreover, the detection efficiency of the IBD was optimized with respect to the previously-performed Borexino analysis of geo neutrinos, which are also detected via the IBD [155]. Here, the individual selection cuts were optimized concerning the expected signal and background event numbers, and the systematic uncertainties in order to maximize the overall sensitivity.

The improved detection efficiency results in $(93.0 \pm 0.5)\%$ for a fiducial volume (FV) cut at 25 cm distance from the inner vessel. This corresponds to an enhanced detection efficiency of 10.8 % using the same FV cut of the previous geo neutrino analysis. Moreover, the statistics of the SOX event sample could be increased with an enlarged FV by a total of 16.3 %. This still relatively large FV cut of 25 cm is mandatory to minimize the uncertainty on the detection efficiency due to the uncertain inner vessel shape. A large fraction of this optimization strategy entered in the latest geo neutrino analysis of Borexino [35].

Furthermore, the expected detector- and source-related backgrounds for SOX were investigated. The detector-related background was studied using 4.16 y of Borexino data. The expected event number for a measurement time of 1.5 y of SOX was evaluated to $21.6^{+2.7}_{-4.2}$, where the main contribution comes from the irreducible geo and reactor neutrinos. During this work, a new model for the determination of accidental coincidences was developed. Moreover, the coincidence of two μ-induced ^{12}B-decays can mimic an IBD signal in Borexino, which was observed for the first time in the frame of this thesis. These results have been published as part of the latest geo neutrino analysis [35] and the search for low-energy neutrinos from astrophysical sources [51]. The source-related background was estimated from Monte Carlo (MC) simulations to $\sim 6.6 \pm 0.03$ events. These events are mainly induced from neutrons produced by spontaneous fission of ^{244}Cm, a contamination of the ^{144}Ce-source which is extracted from spent nuclear fuel. While the neutrons are captured on their way to the detector, the resulting relaxation gamma reaches the FV and can accidentally coincide with another single detector- or source-related event. Thus, as the source-related background is only of accidental nature, it can be well characterized during the data taking of SOX. In total, $28.2^{+2.7}_{-4.2}$ background events are expected for a source activity of 100 kCi, which corresponds to a signal to background ratio of ∼200. In conclusion, the expected event number as well as the spectral shape of the background for SOX can be well-constrained with a precision of more than 10 %.

The sensitivity of SOX and the systematic effects on the sensitivity were also studied. SOX is sensitive to oscillation amplitudes $\sin^2(2\theta) \gtrsim 0.06$ at $\Delta m^2 = 1\,\mathrm{eV}^2$ at the 95 % CL and covers most of the parameter space which is preferred from the anomalies. Due to the well-characterized source and detector, the expected event number is known and a sensitivity for $\sin^2(2\theta) \gtrsim 0.09$ at large mass splittings ($\Delta m^2 \gtrsim 10\,\mathrm{eV}^2$) is achieved (95 % CL). Here, an uncertainty on the total event number of 1.5 % and an absolute uncertainty of 0.03 on the spectral shape of the neutrino spectrum is assumed. The uncertainty on the event number mainly consists of the measurement of the power of the source (0.2 %), uncertainties on the position reconstruction (∼ 1 %), and the uncertainty on the detection efficiency (0.5 %). An uncertainty on the event number decreases the sensitivity for $\Delta m^2 \gtrsim 10\,\mathrm{eV}^2$ and $\Delta m^2 \sim 0.3\,\mathrm{eV}^2$, where the sterile neutrino signature is mostly given by a flat reduction of the event number. The uncertainty on the spectral shape of the neutrino spectrum affects not only the spectral shape of the neutrino signal itself, but also the total expected event number. This comes from the power-to-activity conversion using the mean emitted energy per decay which depends on the spectral shape. However, within the required accuracy of the spectral shape, the effect on the total expected event number is below 1 %. Besides this additional uncertainty on the event number, an uncertainty in the spectral shape can only mimic an oscillation signature with $\Delta m^2 < 1\,\mathrm{eV}^2$, which corresponds to an oscillation length larger than ∼ 5 m. Similar results are obtained for the uncertainties in the position reconstruction and the source position. Thus, the desired smoking gun signature of an oscillation length of ∼ 1 m cannot be mimicked by any detector- or source-related systematic in SOX.

Moreover, the sensitivity and the impact of the systematic uncertainties in SOX were studied using the asymptotic distributions of the test statistic, described by Wilks' theorem and the

Asimov data set. As shown in the second part of this thesis, these methods approximate within 15 % the results of a full analysis based on MC techniques.
In order to reach the expected sensitivity, a well-characterized source and detector are required. Thus, in the frame of this thesis, preliminary measurements on the spectral shape of the ^{144}Ce-^{144}Pr spectrum were performed. Moreover, a dedicated SOX calibration campaign with radioactive sources at various positions in the Borexino detector was prepared by the collaboration.
In order to further increase the sensitivity of the SOX experiment, two hardware modifications of the Borexino trigger system were investigated. The modifications are based on the fact that the light yield of the inner buffer in Borexino is higher than actually designed due to a leak in the inner vessel. This enhanced light yield allows detecting SOX events also in the buffer. As the amount of scintillation light is close to the trigger threshold of Borexino, a modification of the nominal Borexino trigger system is required. First, the reduction of the trigger threshold was studied. However, the resulting high trigger rate prevented from a stable data acquisition system, and this modification was hence discarded. Second, the installation of an automatically issued neutron capture gate after each potential prompt SOX candidate was investigated. The idea is to detect all the neutrons from the IBD reactions close to the inner vessel border and decrease thus the uncertainty on the detection efficiency induced by the uncertain vessel shape. The proper behavior of this trigger modification in terms of detection efficiencies and hit-time distributions in the data acquisition gates were successfully tested and validated. However, the trigger rate with the modified trigger was reduced by 5.5 % in comparison to the nominal Borexino trigger rate. This problem could not be resolved and was not further investigated due to the cancelation of the SOX project. During this study, an artifact of the Borexino data acquisition system was identified for the first time. This artifact is now taken into account in the data analysis of Borexino.

12.2 Statistical Methods and Issues in Sterile Neutrino Experiments

The statistical issues in a (3+1) model for SBL experiments were studied for a large set of experimental parameters using a toy appearance and a toy disappearance experiment. Moreover, three different statistical analyses were discussed. The analysis based on the integral event number only (rate analysis) and the analysis based on the relative differences in the spectral shape as a function of the distance L and the energy E (shape analysis) are studied. In addition, the analysis given by the shape analysis with a constraint on the total event number (rate+shape analysis) is investigated. Further, exclusion and discovery scenarios were evaluated and compared.
Due to the same sought-after signature in appearance and disappearance experiments, the same behavior of the sensitivities and the statistical issues are observed. The rate analysis has effectively one degree of freedom and can only reconstruct the integral oscillation probability, as the two oscillation parameters $\sin^2(2\theta)$ and Δm^2 become correlated. The confidence region and the expected sensitivity can be obtained from Wilks' theorem and the Asimov data set assuming one degree of freedom for the asymptotic test statistic distributions. In contrast, the shape and rate+shape analysis can observe an oscillation signature in L and E and can hence reconstruct - in case of a signal - both oscillation parameters. However, with the effective two degrees of freedom in the analysis, weaker limits in terms of $\sin^2(2\theta)$ are obtained than in the one-dimensional rate analysis for $\Delta m^2 \gtrsim 10\,\mathrm{eV}^2$. Statistical fluctuations in the spectral data typically lead to best-fit hypotheses different from the no-oscillation hypothesis, even

when it is true. This can be explained by the flexibility of the fit model which can always describe the statistical fluctuations in the data sample with a certain oscillation signature. Thus, the maximum-likelihood estimators are biased in a shape and rate+shape analysis. This bias results into several unexpected effects. For instance, the bias reverses the ordering of the discovery and exclusion sensitivity for the same confidence level in comparison to the rate analysis. Moreover, the bias in combination with the presence of the parameter bound for the mixing angle at $\sin^2(2\theta) = 0$ changes the ordering of the exclusion power for different constraints on the total event number. From that follows that the limit and sensitivity obtained from a rate+shape analysis are weaker than the ones from a shape analysis at $\Delta m^2 \sim 1\,\mathrm{eV}^2$. The above-described features imply that the rate+shape analysis is not a simple linear combination of the individual rate and shape analyses in a SBL experiment.

The bias is also responsible for the non-validity of the asymptotic probability distributions in a shape and rate+shape analysis. In particular, Wilks' theorem does not hold for the no-oscillation hypothesis and leads to an undercoverage. The undercoverage, which depends on the experimental parameters, is for the toy disappearance experiment 65 % and 77 % for a nominal value of 95 % in a shape and rate+shape analysis, respectively. Thus, the limit and the significance reconstructed from Wilks' theorem are overestimated by up to 1σ and could also lead to the claim of a signal even though the no-oscillation hypothesis is true. The tool from Reference [212] for the estimation of the probability distribution of the no-oscillation hypothesis and hence the significance was tested and gave promising results for the shape analysis. However, the significance is underestimated in a rate+shape analysis which is similar to what is obtained for the tested look-elsewhere effect correction in both the shape and the rate+shape analysis. Further, the asymptotic distribution for the goodness-of-fit test statistic cannot be described by Wilks' theorem if the no-oscillation hypothesis is true. The above-described bias decreases the effective degrees of freedom of the test, and using Wilks' theorem results into overestimated goodness-of-fit values.

The methods applied to the current SBL experiments were compared and discussed. Their differences can be traced back to the definition of the test statistic and the alternative hypothesis used in the hypothesis test. The natural test statistic T_2 is based on the most general alternative hypothesis, i.e. the physically allowed parameter space, and carries the above-described properties. The reconstructed confidence region with T_2 contains those hypotheses which describe the data best in the (3+1) model. This is the natural interpretation of a confidence region, which is obtained via the text-book approach. The test statistic T_2 can reconstruct a limit as well as a signal. However, the discussed bias and the non-trivial likelihood function require a large computational effort. Moreover, as the asymptotic formulae for the probability distributions of the test statistic are not valid, an extensive MC construction is required.

Two alternative test statistics T_0 or T_1 are applied in current SBL experiments. These test statistics restrict the alternative hypothesis and simplify the minimization of the negative log-likelihood function. The asymptotic formulae of the probability distributions of the test statistics are valid and the computational effort is reduced in comparison to T_2. Even though both oscillation parameters are currently unknown and a restriction of them in the alternative hypothesis is not natural, two-dimensional regions can be obtained via the repetition of the text-book approach of the construction of a confidence region. T_1 fixes the mass splitting Δm^2 to the value of the tested hypothesis in the alternative hypothesis and constructs therefore a one-dimensional confidence region via the text-book approach. Instead, T_0 defines the alternative hypothesis as the no-oscillation hypothesis and fixes hence both oscillation parameters. The reconstructed confidence region with T_0 is therefore point-like. The resulting two-dimensional regions are hence the union of one-dimensional or point-like confidence

Table 12.1: Comparison of the Confidence Regions based on the Different Test Statistics applied to Current Short-Baseline Experiments The first row describes the interpretation of the hypotheses contained in the specific confidence region. The second row indicates whether the two-dimensional confidence region is constructed via the inversion of a hypothesis test. The third and the fourth row show whether a signal or a limit can be reconstructed. The fifth column lists the validity of the asymptotic formulae for the probability distributions of the test statistic. Whether the amount of computation time is relatively low is indicated in row six.

	T_0	T_1	T_2
statement of confidence region	hypotheses more likely than the no-signal hypothesis	hypotheses describing data best for each Δm^2-value	hypotheses describing data best for 3+1 model
text-book approach	–	–	+
signal	–	–	+
limit	+	+	+
asymptotic formulae	+	+	–
computation time	+	+	–

regions, respectively. Even though the test of a single hypothesis with T_0 and T_1 is correct in terms of a hypothesis test and coverage, the reconstructed two-dimensional regions are not obtained via the text-book approach and carry hence a different interpretation. In particular, the two-dimensional region reconstructed with T_1 contains for each Δm^2-value the $\sin^2(2\theta)$-region which describes the data best for that specific Δm^2-value. Hence, all Δm^2-values are contained in the two-dimensional region and Δm^2 cannot be reconstructed with T_1. The two-dimensional region reconstructed with T_0 always contains the no-oscillation hypothesis and only those hypotheses which are more likely than the no-oscillation hypothesis. Hence, T_0 cannot reject the no-oscillation hypothesis and therefore claim a discovery. Moreover, neither $\sin^2(2\theta)$ nor Δm^2 can be reconstructed with T_0. Thus, both test statistics cannot be used to reconstruct a signal.

In contrast, all three test statistics can set a limit in the $\sin^2(2\theta)$ and Δm^2 parameter region, where the differences to each other are reduced. Nevertheless, the different alternative hypothesis prevents from a direct comparison and defines the ordering of the limits. Thus, as the restriction of the alternative hypothesis increases the power of a test, the limit with T_0 is stronger than T_1, which is in turn typically stronger than T_2. As T_0 can produce very strong limits for some Δm^2-values, it is used in combination with the CL_S-method which weakens the overall limit. However, the interpretation of the two-dimensional region remains unchanged. The three test statistics and their properties are summarized in Table 12.1.

In order to ease the comparison of the results, a standardized approach in the presentation of the results would be beneficial for the field. In Reference [208], in which a large part of this thesis was published, an analysis based on T_2 using a MC construction was proposed. In particular, if the no-oscillation hypothesis is accepted, the reconstructed confidence region will extend down to vanishing $\sin^2(2\theta)$ values and its upper limit can be plotted along with its median value (i.e. the exclusion sensitivity) and 68/95 % central intervals expected under the no-oscillation hypothesis. If the no-oscillation hypothesis is instead rejected, the confidence region can be plotted for different confidence levels along with the discovery sensitivity. Moreover, a p-value for the significance and a goodness-of-fit test are supposed to be reported. Furthermore, the non-validity of Wilks' theorem and the non-trivial interplay between the rate and the shape analysis implies that a MC construction is also required for a global sterile

neutrino analysis. Therefore, we encouraged the individual experiments to provide as much as possible information needed for a global analysis.
Fortunately, this proposed standardized approach has been adopted by several groups in their recent analyses. In particular, STEREO [211] updated their analysis using not only T_1, but also T_2 with a MC construction. Moreover, STEREO as well as PROSPECT [210] extended their results by extra information, as for example the map of their critical values. In addition, stimulated by the results of this thesis, groups performing global fits are aware of the overestimation of the significance when Wilks' theorem is applied. Therefore, they already perform a MC construction [217] or plan to do so in the future [64].

12.3 Outlook

The work of this thesis on the statistical issues in SBL experiments could be extended by the study of a wide range of the experimental parameters. In particular, it would be interesting to investigate the distribution of the test statistic under the no-oscillation hypothesis as a function of the systematic and statistical uncertainties of the experiment. Moreover, given a possible correlation, as already hinted at in Chapter 11, it might be possible to approximate the significance of a signal also in a rate+shape analysis.[1] In addition, as Reference [212] observes a difference between the probability distributions for a Poisson- or Gaussian-distributed test statistic, the investigation of such a dependency is worthwhile as well.
From the published critical values of STEREO [211] and PROSPECT [210] it is visible that the applied ratio-method does not fully agree with a shape analysis. Hence, a detailed study of the ratio-method and its comparison to the shape analysis is desirable.
Moreover, similar statistical issues were found in experiments beyond the SBL program, as for example in the sterile neutrino search in the long-baseline experiment Double Chooz [209] and the test of new physics in the coherent elastic neutrino-nucleus scattering experiment COHERENT [220]. Therefore, this work could be extended to study the statistical issues for experiments beyond the SBL program.

[1] The distribution in a shape analysis can already be well described by the approximation in Reference [212].

13 Current Status of Sterile Neutrino Search

The current situation of light sterile neutrinos is unclear. Several experiments have already released first results, where some of them could exclude a large fraction of the preferred oscillation parameter region of the anomalies. In contrast, a few experiments hint at a positive signal which, however, cannot be confirmed by the other experiments. The results are summarized in more detail in the following.

An interesting coincidence of the positive $\sim 2\sigma$ confidence regions of the reactor experiments NEOS [71] and DANSS [126] is observed. The best-fit value of a combined analysis is given by $\sin^2(2\theta) = 0.049$ and $\Delta m^2 = 1.29\,\mathrm{eV}^2$ at a significance, which is dominated by the DANSS data set, of 3.7σ [221]. The observed mixing is smaller than the expected one from the reactor antineutrino anomaly (RAA) and Gallium anomaly and is in tension with them at the $\sim 2\sigma$ level [116]. As this result is based on the ratio of the measured energy spectra at different baselines, it is independent on the reactor neutrino flux, and hence more robust than the RAA [7]. Moreover, this observation is also called the "spectral anomaly" of the RAA [222]. However, DANSS presented preliminary results in 2019 and the combined significance with NEOS [71], Bugey-3 [223], and PROSPECT [128] reduces to, assuming Wilks' theorem, 2.6σ with a best-fit value of $\sin^2(2\theta) = 0.026$ and $\Delta m^2 = 1.29\,\mathrm{eV}^2$ [224]. Based on our work in Reference [208], this significance is reevaluated to 1.8σ using a toy-Monte Carlo (MC) construction [217]. The corresponding confidence regions are depicted in Figure 13.1a. In addition to the ones from the toy-MC construction, the overestimated confidence regions reconstructed from Wilks' theorem are shown. DANSS preliminarily presented another updated result at the Neutrino conference 2020 which shows no indication for a sterile neutrino signal and states a corresponding significance of 1.5σ [225]. Thus, one can expect that the significance of the combined analysis will reduce further.

Another positive signal at the 3.0σ level is reported by the Neutrino-4 collaboration, where their best-fit value is $\sin^2(2\theta) \sim 0.26$ and $\Delta m^2 \sim 7.25\,\mathrm{eV}^2$ [228]. However, this result is heavily debated (see for example Reference [229]) and is in tension with the residual experiments, which can be seen in Figure 13.1b. In particular, the preferred oscillation parameters are inconsistent with the RAA and the above-discussed results from NEOS and DANSS. Further, the limits of PROSPECT [210] and solar neutrino data [123] reject the best-fit value of Neutrino-4. In addition, STEREO [211] and the tritium-endpoint measurement KATRIN [138] are about to probe this parameter region in the near future. Moreover, Neutrino-4 uses Wilks' theorem and the significance of the signal is therefore overestimated, as discussed in Chapter 11. Based on the work of this thesis, PROSPECT and STEREO emphasize for another time the non-validity of Wilks' theorem and request the Neutrino-4 collaboration to publish their results using a toy-MC construction [229]. The significance of the Neutrino-4 signal is reestimated in Reference [212] to 2.6σ which, however, treats only the statistical error. Hence, it is likely that the significance decreases even further when the

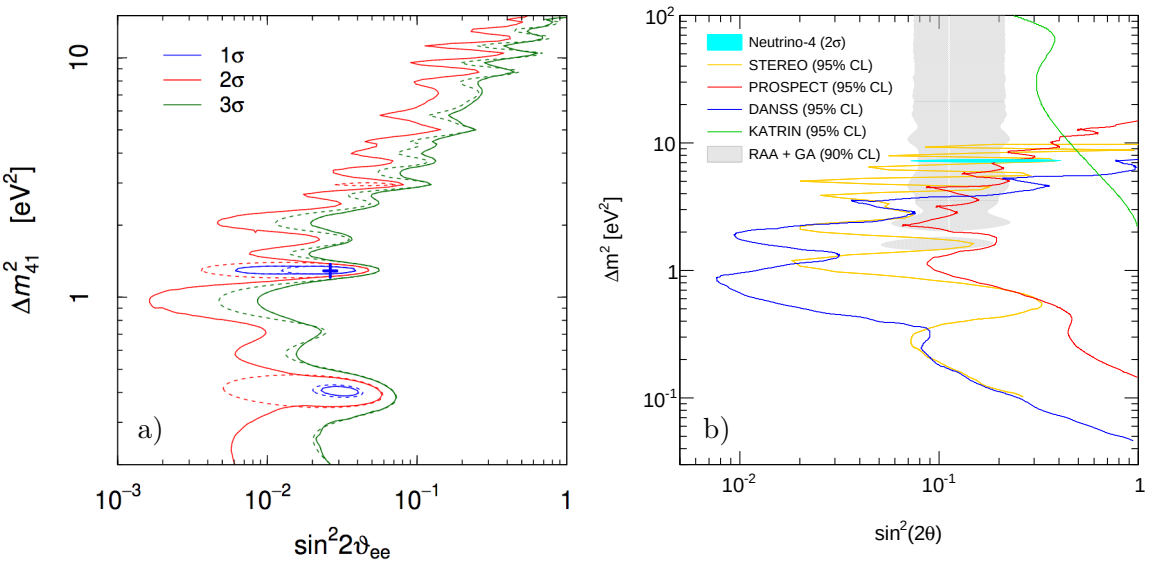


Figure 13.1: Current Status of Sterile Neutrinos a) The preferred parameter space in $\sin^2(2\theta)$-Δm^2 for a combined analysis of NEOS [71], DANSS [226], Bugey-3 [223], and PROSPECT [128] data are shown for the 1σ (blue), 2σ (red), and 3σ (green) level. The confidence regions reconstructed from a toy-MC construction (Wilks' theorem) are shown in solid (dashed). The cross indicates the best-fit value. The figure is taken from Reference [217]. b) Results on the oscillation parameters $\sin^2(2\theta)$-Δm^2 from the Neutrino Conference 2020. The limits from STEREO [227], PROSPECT [210], DANSS [225], and KATRIN [138] are shown at the 95 % CL. The signal observed from Neutrino-4 [228] is shown at the 2σ level. The gray area highlights the 90 % CL region allowed by the reactor and gallium anomaly [67]. Remind that the experiments use different methods which prevents from the direct comparison of the results.

systematic effects are taken into account.

In a recent update, MiniBooNE reports an increased excess of 4.8σ in a combined neutrino and antineutrino measurement with a best-fit value at $\sin^2(2\theta) = 0.807$ and $\Delta m^2 = 0.043\,\text{eV}^2$ [230].

The preferred parameter region in a global analysis in 2019 becomes small and is shifted towards slightly smaller mixing angles than initially hinted at by the anomalies [231]. The best-fit value in a (3+1) model is $\sin^2(2\theta) = 0.053$ and $\Delta m^2 = 1.32\,\text{eV}^2$ and is favored over the no-oscillation hypothesis with more than 5σ. However, the tension between appearance and disappearance experiments remains and corresponds to a 4.5σ level. The strong tension comes actually from the fact that a signal is observed in $\nu_e/\bar{\nu}_e$-disappearance and in $\nu_\mu \to \nu_e/\bar{\nu}_\mu \to \bar{\nu}_e$-appearance experiments, but not in the $\nu_\mu/\bar{\nu}_\mu$-disappearance channel [7]. The strongest bounds come here from Minos&Minos+ [232] and IceCube [233]. In order to make the appearance and disappearance results compatible with each other, alternative models are studied. However, a (3+2) model cannot reduce the tension. In contrast, the (3+1+decay) model, where the fourth neutrino mass state can decay, lowers the tension to a 3.2σ level [231]. Moreover, such a model could also decrease the tension with cosmological data which strongly disfavor a sterile neutrino at the eV-scale. The number of relativistic neutrinos N_{eff} is consistent with three neutrinos and the sum of the respective neutrino masses is $m_\nu \lesssim 0.1\,\text{eV}$ [17]. However, this result is based on the cosmological model which assumes for example that the sterile neutrino is in thermal equilibrium with the active neutrinos at the time of neutrino decoupling. Therefore, a model that avoids the thermal equilibrium can hence reduce this tension (see e.g. References [116, 231]).

In order to explain the anomalous data, models of new physics, also beyond the sterile neutrino hypothesis, are proposed [234–240]. Moreover, the anomalies themselves are reevaluated and crosschecked for systematic effects. For instance, the Gallium anomaly was recently reestimated based on new cross-section calculations to only 2.3σ [241]. Doubts on the RAA arise from two observations that show a deviation from the nominal Huber-Mueller flux prediction [65, 66]: First, the so-called "5 MeV bump" in the spectral shape which is a discrepancy of $\sim 10\,\%$ was observed by several experiments [196–198]. Second, Daya Bay [242, 243] and RENO [244] directly observed that the IBD event rate and energy spectrum evolve and depend on the reactor fuel composition during burnup. According to their results, a misprediction of the reactor neutrino flux, mainly the ^{235}U contribution, may be the main source for the RAA. Therefore, the RAA is reevaluated using not only the Huber-Mueller flux, but also two updated flux predictions [72, 204]. The anomaly depends then on the flux prediction and results into a significance varying in between 0.95σ and 2.8σ [222], where the smallest significance is obtained for the neutrino flux based on ab initio calculations (see Chapter 1). The MiniBooNE anomaly is also disputed, as the excess is only present in the lowest energy bins which is not fully compatible with the (3+1) model [7, 120]. MiniBooNE reinvestigates hence carefully their individual backgrounds. In particular, single photons from misidentified π^0-decays could mimic a neutrino signal. However, π^0-decays can be well constraint with in-situ measurements and are disfavored as an explanation for the excess [230]. Moreover, the future short-baseline program at Fermilab uses liquid Argon time-projection chambers and can discriminate between photons and electrons and can hence independently test this possible background source [135].

14 Conclusions

Light sterile neutrinos are one of the current hot topics in neutrino physics, and a large number of experiments are looking for them. However, the present status is confusing, and more experimental data is needed. Here, the complementary SOX project with its well-characterized source and detector would have provided a helpful measurement to disentangle this unclear situation. Even though SOX was canceled, the results obtained within this dissertation project are a valuable input for the design and analysis of current and future experiments. Moreover, the optimized event selection and background modeling were used to improve other antineutrino analyses in Borexino. In order to describe the present global picture of light sterile neutrino searches, either new physics or so far unknown systematic effects need to be considered. Furthermore, due to the different applied analysis methods, special attention is required for the comparison of the results and the interpretation of the significance. The standardized analysis approach proposed in the frame of this work helps to clarify the situation from a statistical point of view and has been adopted by several experiments. Within the next years, the currently running experiments will expand their reach and test the full parameter space opened by anomalous neutrino oscillation data. However, recent results may hint at an allowed parameter region towards smaller mixing angles and next-generation experiments with an intense and well-characterized neutrino source (e.g. IsoDAR [245]) might be necessary to give conclusive results on the existence of light sterile neutrinos. A discovery of sterile neutrinos would have deep repercussions on both particle physics and cosmology.

Appendix

A.1 Calibration of the TUM Spectrometer

The energy spectrum of the produced electron in a beta decay can be described with:

$$N(W) \propto W \cdot p \cdot (W - W_0)^2 \cdot F(Z, W) \cdot C(W), \tag{A.1}$$

where W is the total energy and p the momentum of the electron [246]. W_0 is the endpoint of the beta spectrum. $F(W, Z)$ is the Fermi function and a numerical solution can be found for example in [247]. The shape factor $C(W)$ modifies the allowed spectrum and depends on the classification of the transition.[1] The calibration sources and their transitions are listed in Table A.1.

Table A.1: Calibration Sources for the TUM Spectrometer The decays of the calibration sources are listed along with their probabilities and energies in the first four columns. Given a beta decay, the energy corresponds to the Q-value. The classification and the shape factor used for the simulation of the beta transition is given in the last two columns. Here, q denotes the momentum of the neutrino and (1) stands for $3(W^2-1)^2+3(W_0-W)^4+10(W^2-1)(W_0-W)^2$.

Source	Decay	Probability [%]	Energy [keV]	Type of Forbidden Transition	Shape Factor
^{90}Sr [248]	β^-	100	545.9	unique 1st	p^2+q^2 [246]
^{90}Y [248]	β^-	99.983	2278.7	unique 1st	p^2+q^2 [246]
	β^-	0.017	518.0	unique 1st	p^2+q^2 [246]
^{36}Cl [191]	β^-	98.1	709.5	non-unique 2nd	(1) [249]
^{137}Cs [248]	β^-	94.4	514.0	unique 1st	p^2+q^2 [246]
	β^-	5.6	1175.6	non-unique 2nd	$0.03p^2+q^2$ [250]
	γ	85.0	661.7		
	conversion e^-	7.6	624.2		
	conversion e^-	1.4	656.0		
	conversion e^-	0.3	660.5		

The comparison of the Geant4-based simulated and measured spectra with the TUM spectrometer are shown in Figures A.1, A.2, and A.3. Figure A.4 compares the measured spectrum of the same ^{137}Cs source for two different measurement runs.

[1] One differentiates forbidden decays between unique and non-unique transitions, where non-unique transitions are typically more difficult to describe as they consist of more than one ("unique") nuclear matrix element.

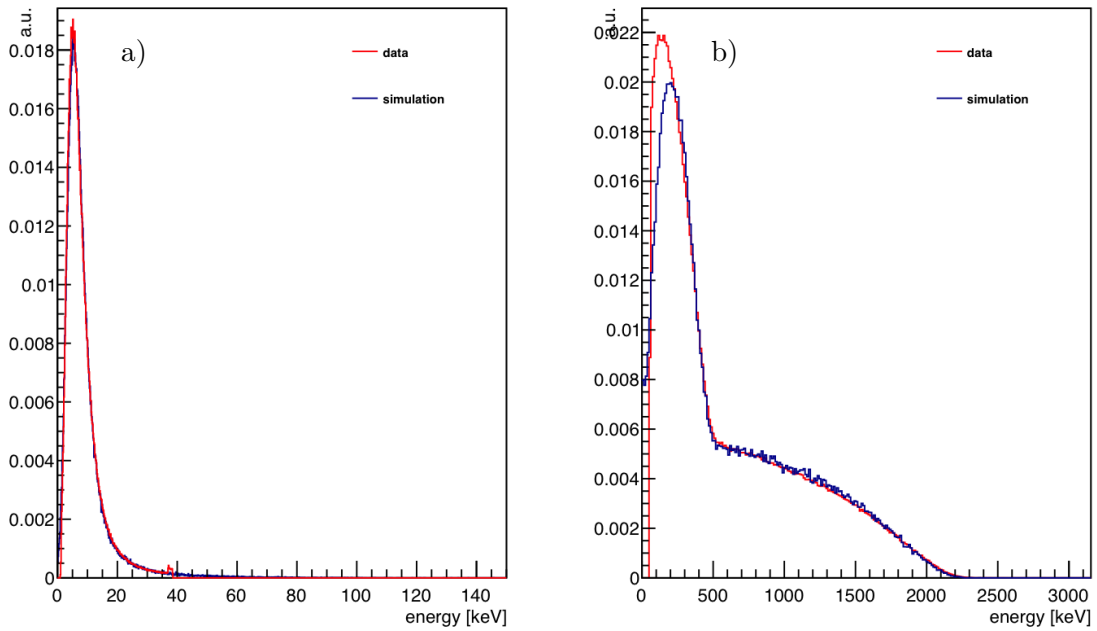


Figure A.1: Calibration of the TUM Spectrometer with a ^{90}Sr-^{90}Y Source a) spectrum measured with the multi-wire chamber b) spectrum measured with the photo multiplier tube.

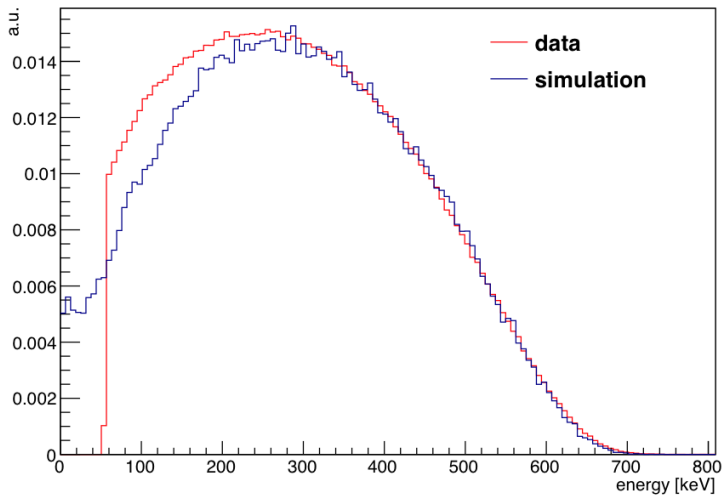


Figure A.2: Calibration of the TUM Spectrometer with a ^{36}Cl Source Spectrum measured with the photo multiplier tube.

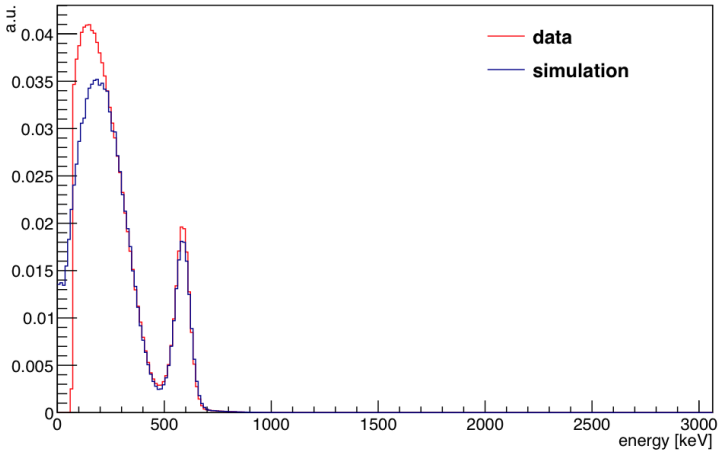


Figure A.3: Calibration of the TUM Spectrometer with a ^{137}Cs Source Spectrum measured with the photo multiplier tube.

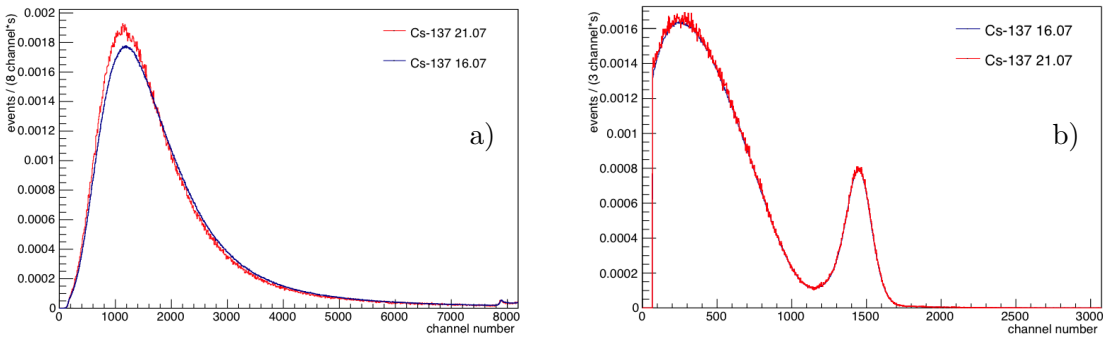


Figure A.4: Stability of the TUM Spectrometer The spectrum of a ^{137}Cs source measured in a) the multi-wire chamber b) the photo multiplier tube is shown for two different measurements. In between the two measurements, the spectrometer was opened and closed, which could affect the relative source position to the two detectors and thus explain the difference in the spectra.

A.2 Borexino Trigger Types

Table A.2: Borexino Trigger Types The Borexino trigger types (*tt*) and the corresponding Borexino trigger board inputs (*btb*) with their conditions and gate lengths are listed. The combination of *tt* and *btb* is important to distinguish between events only triggering the Inner Detector (ID), the Outer Detector (OD), or both. The service triggers (calibration and random) are automatically issued every two seconds.

Trigger Name	*tt*	*btb*	Condition	Gate Length
Neutrino	1	0	≥20 ID PMT hits within 100 ns	16 µs
Muon	2	4	≥6 OD PMT hits within 150 ns	16 µs
Neutron	128	8	after *tt*1 & *btb*4	1.6 ms
SOX Prompt	4	16	≥65 ID PMT hits within 100 ns	16 µs
SOX Delayed	16	32	after *tt*4	1.6 ms
Fake Signal	8	64	after *tt*16/*tt*128	1.6 µs
Calibration	32	64	issued every 2 s	1.6 µs
Random	64	64	issued every 2 s	1.6 µs

A.3 Optimization of Selection Cuts

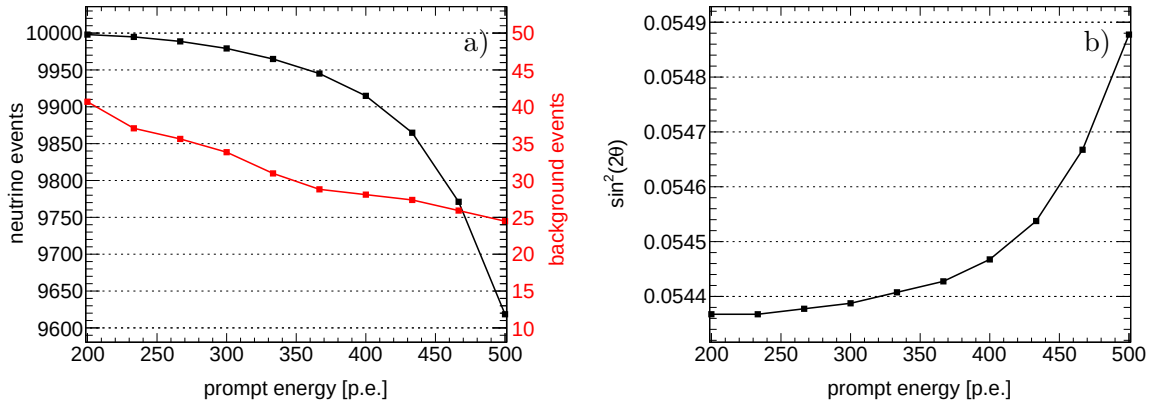


Figure A.5: Optimization of the E_p-Cut a) the SOX and the background event numbers are given as a function of the E_p-cut value. The SOX event number assumes a total event number of 10^4. The background event number is obtained from Borexino data selected for the optimized cut values without a FV cut. b) the sensitivity is shown in terms of $\sin^2(2\theta)$ at $\Delta m^2 = 10\,\mathrm{eV}^2$. The sensitivity is computed as described in Appendix A.5 and assumes an uncertainty on the signal and background event number of 1 %, respectively.

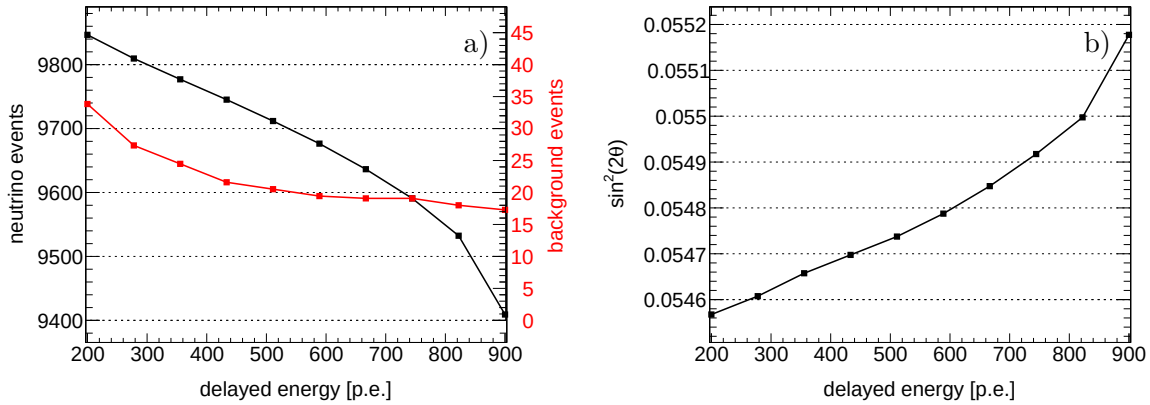


Figure A.6: Optimization of the E_d-Cut a) the SOX and the background event numbers are given as a function of the E_d-cut value. The SOX event number assumes a total event number of 10^4. The background event number is obtained from Borexino data selected for the optimized cut values without a FV cut. b) the sensitivity is shown in terms of $\sin^2(2\theta)$ at $\Delta m^2 = 10\,\mathrm{eV}^2$. The sensitivity is computed as described in Appendix A.5 and assumes an uncertainty on the signal and background event number of 1 %, respectively.

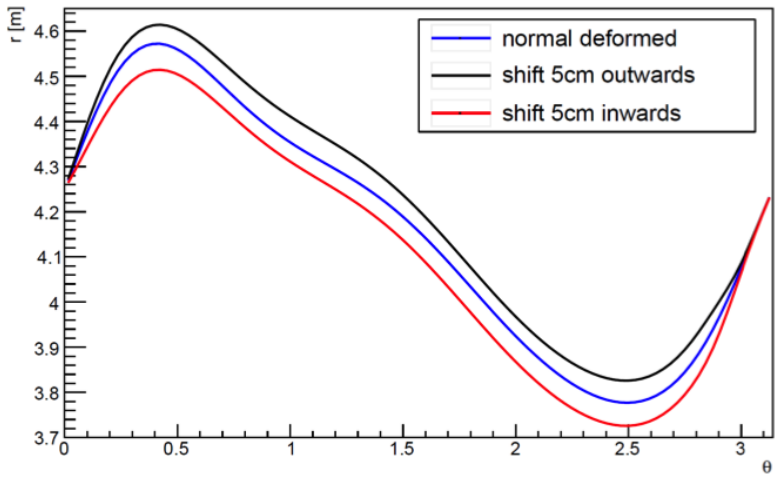


Figure A.7: Reconstructed Inner Vessel Shape The blue line shows the reconstructed radius of the vessel shape from October 2016 as a function of the polar angle θ. One assumes a cylindrical symmetry of the vessel shape. The red and black vessel shapes are generated by applying a 5 cm shift which is the size of the uncertainty of the vessel shape. The plot is taken from Reference [190].

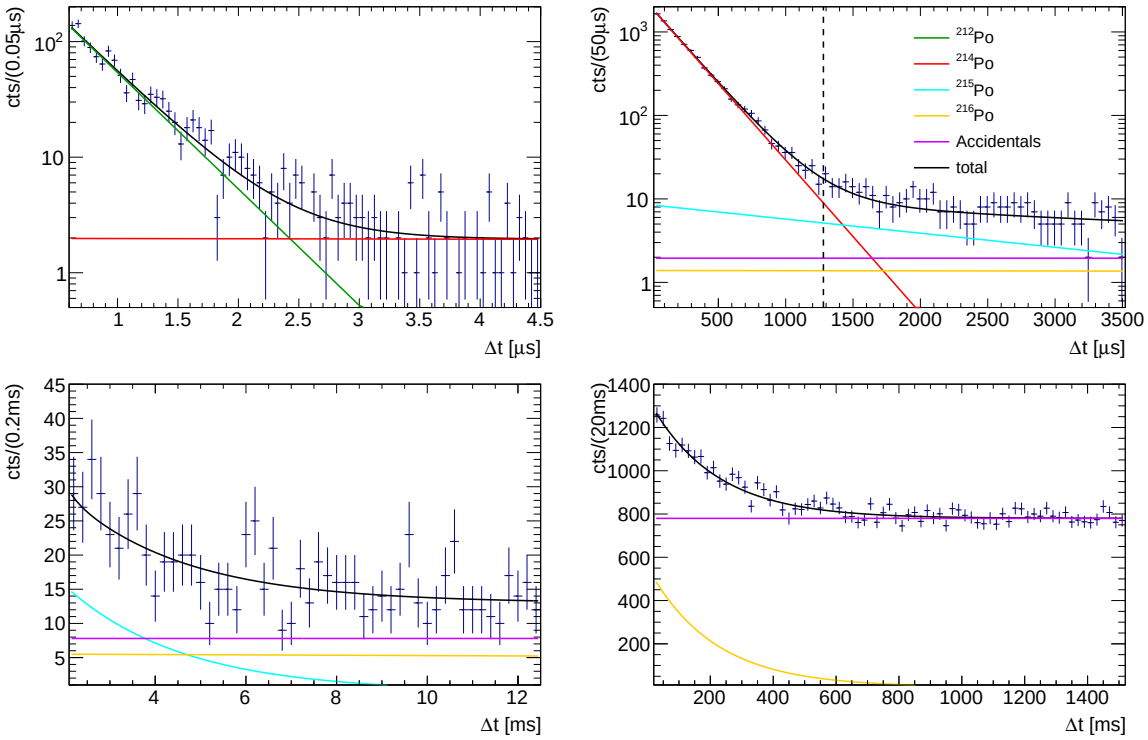


Figure A.8: Fast Coincidences with Polonium Isotopes The IBD-like events are selected for the optimized selection cuts, but without the *mlp*-cut on the delayed event and for $E_p = 200$ p.e.. The four panels show different time scales for Δt. The measured time distribution can be well described by the sum of four exponential decay functions with the respective Polonium lifetimes ($\tau(^{212}\text{Po}) = 425.1$ ns, $\tau(^{214}\text{Po}) = 236.0$ µs, $\tau(^{215}\text{Po}) = 2.6$ ms, $\tau(^{216}\text{Po}) = 209$ ms) and a flat contribution due to accidental coincidences. The dashed line indicates the maximum Δt-cut value for SOX.

A.4 Minimization Algorithm of the Negative Log-Likelihood Function

To minimize the non-trivial negative log-likelihood function (see for example Figure 9.3 and 11.8) and determine the maximum-likelihood estimator (MLE) for a specific data sample, the Δm^2-range is divided into a fine raster of 500 Δm^2-values. These 500 raster points are equally-spaced in a logarithmic scale. Further, 100 individual raster points are added for the assumed true Δm^2-value. These additional values cover a Δm^2-region up to a deviation of $\pm 5\%$ for the specific Δm^2-value in linear spacing. As the likelihood function is a monotonic function in $\sin^2(2\theta)$ for a fixed Δm^2-value, the minimization algorithm MINUIT [188] converges properly for fixed Δm^2-values. After the individual minimization at these 600 raster points, the smallest local minimum out of these defines then the global minimum.
To speed up the minimization process, one can use the advantage that only one probability distribution $pdf^S(\sin^2(2\theta) = x_1, \Delta m^2)$ and integrated oscillation probability $P_{\text{int}}(\sin^2(2\theta) = x_1, \Delta m^2)$ has to be produced per Δm^2-value. All other probability distributions of the signal hypotheses with $\sin^2(2\theta) = x_2$ for the same Δm^2-value can be obtained via rescaling and the probability distribution of the no-signal hypothesis $pdf^S\,(0,0)$. In case of the appearance experiment, the normalized probability distribution is unchanged ($pdf^S(\sin^2(2\theta) = x_2, \Delta m^2) = pdf^S(\sin^2(2\theta) = x_1, \Delta m^2)$) and the integral survival probability becomes

$$P_{\text{int}}(\sin^2(2\theta) = x_2, \Delta m^2) = \frac{x_2}{x_1} \cdot P_{\text{int}}(\sin^2(2\theta) = x_1, \Delta m^2). \tag{A.2}$$

For the disappearance experiment, the product of the probability distribution and the integral survival probability can be rescaled using:

$$\begin{aligned} &P_{\text{int}}\left(\sin^2(2\theta) = x_2, \Delta m^2\right) \cdot pdf^S\left(\sin^2(2\theta) = x_2, \Delta m^2\right) = \\ &pdf^S\,(0,0) - \frac{x_2}{x_1} \cdot \Delta pdf^S\left(\sin^2(2\theta) = x_1, \Delta m^2\right), \end{aligned} \tag{A.3}$$

where

$$\begin{aligned} &\Delta pdf^S\left(\sin^2(2\theta) = x_1, \Delta m^2\right) = \\ &pdf^S\,(0,0) - P_{\text{int}}\left(\sin^2(2\theta) = x_1, \Delta m^2\right) \cdot pdf^S\left(\sin^2(2\theta) = x_1, \Delta m^2\right). \end{aligned} \tag{A.4}$$

Notice that such a construction works only, when there are no bin-correlated uncertainties.

A.5 Calculation of Sensitivities in a Rate Analysis

In the following, the formalism for an easy computation of the sensitivities in a rate analysis is given. To reject a hypothesis with the expected event number N_{exp} at the $n\sigma$ level, the observed event number N_{obs} needs to be apart exactly $n\sigma$:

$$\Delta N = |N_{\text{exp}} - N_{\text{obs}}| = n \cdot \sigma, \tag{A.5}$$

where σ is the total uncertainty of statistical and systematic ones of the signal (s) and the background (b), respectively:

$$\sigma^2 = \sigma^2_{\text{stat}}(N_{\text{exp}}) + \sigma^2_{\text{syst}}(N_{\text{exp}}) = N^s_{\text{exp}} + N^b_{\text{exp}} + (\sigma^s_{\text{r}} \cdot N^s_{\text{exp}})^2 + (\sigma^b_{\text{r}} \cdot N^b_{\text{exp}})^2. \tag{A.6}$$

Notice that these equations are valid, when the expected event numbers of the signal (N^s_{exp}) and the background (N^b_{exp}) are large, where $N_{\text{exp}} = N^s_{\text{exp}} + N^b_{\text{exp}}$. Moreover, σ^s_{r} and σ^b_{r} denote the relative systematic uncertainties on the signal and background event number, respectively. One has to distinguish between one-sided and two-sided limits, as sketched in Figure A.9. If a hypothesis can be rejected for both a positive and a negative difference in the event numbers (i.e. $N_{\text{obs}} > N_{\text{exp}}$ or $N_{\text{obs}} < N_{\text{exp}}$), n corresponds to ≈ 2 for a 95 % CL. This is the case for the exclusion sensitivity (two-sided limit). When a hypothesis can only be rejected for an excess or deficit in the event number as for the no-oscillation hypothesis in the appearance and disappearance experiments, the 95 % CL corresponds to $n = 1.645$. This is the case for the discovery sensitivity (one-sided limit).
For large Δm^2-values, when the integral oscillation probability averages out, the difference in the event numbers needed for the sensitivity can be connected to the sensitive value of the mixing amplitude:

$$\Delta N = 0.5 \cdot \sin^2(2\theta) \cdot N_\nu, \tag{A.7}$$

where N_ν denotes the neutrino event number in case of an oscillation probability of one.
Table A.3 lists the computed values for the sensitivities with $N_\nu = 10^5$ for the toy appearance and the toy disappearance experiment. They are consistent with the values which are obtained through a full MC approach (see Figure 9.1). Given a known background event number, the 1σ signal discovery limit for the appearance experiment reduces to $\Delta N/\sqrt{N^b_{\text{exp}}}$which is equivalent to $s/\sqrt{b}$ from Reference [181].

Table A.3: Assumed and Computed Values in a Rate Analysis for the Toy Experiments The assumed values are the ones from Table 9.1. The values indicated in gray are computed using Equations A.5-A.7. The last column shows the computed sensitive $\sin^2(2\theta)$ value in a rate analysis for large Δm^2-values at the 95 % CL.

	n	N^s_{exp}	N^b_{exp}	σ^s_{r}	σ^b_{r}	σ	N_{obs}	ΔN	$\sin^2(2\theta)$
toy appearance experiment									
exclusion	1.96	120	10^3	0.1	0.05	61.4	1000	120	$2.40 \cdot 10^{-3}$
discovery	1.645	0	10^3	0.1	0.05	59.2	1097.3	97.3	$1.95 \cdot 10^{-3}$
toy disappearance experiment									
exclusion	1.96	96157	10^4	0.02	0.02	1960.8	110000	3843	0.077
discovery	1.645	100000	10^4	0.02	0.02	2037.2	106649	3351	0.067

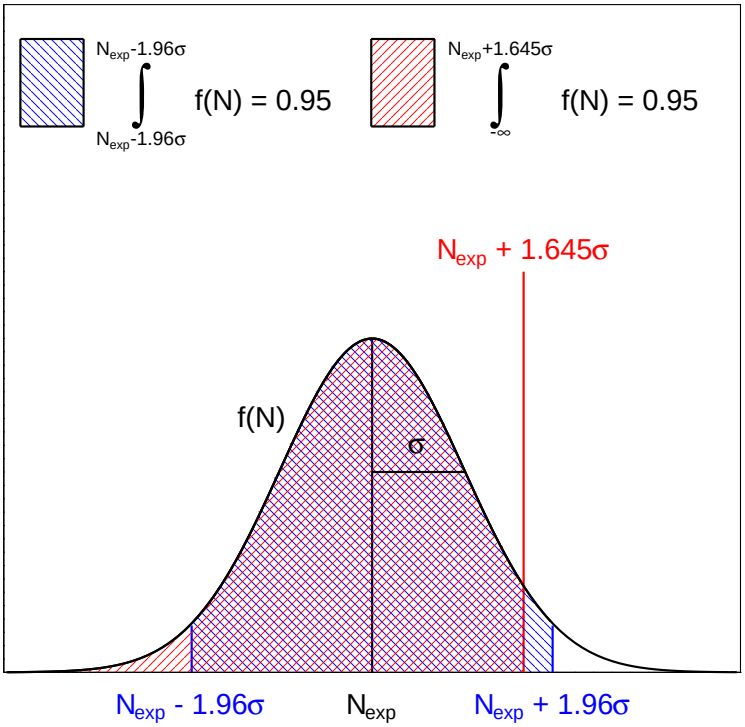


Figure A.9: Visualization of One-Sided and Two-Sided Limits. $f(N)$ denotes the Gaussian distribution with the mean N_{exp} and the standard deviation σ. The critical values at the 95 % CL for a one-sided and two-sided limit are indicated in red and blue, respectively.

A.6 Asymptotic Properties of T_2

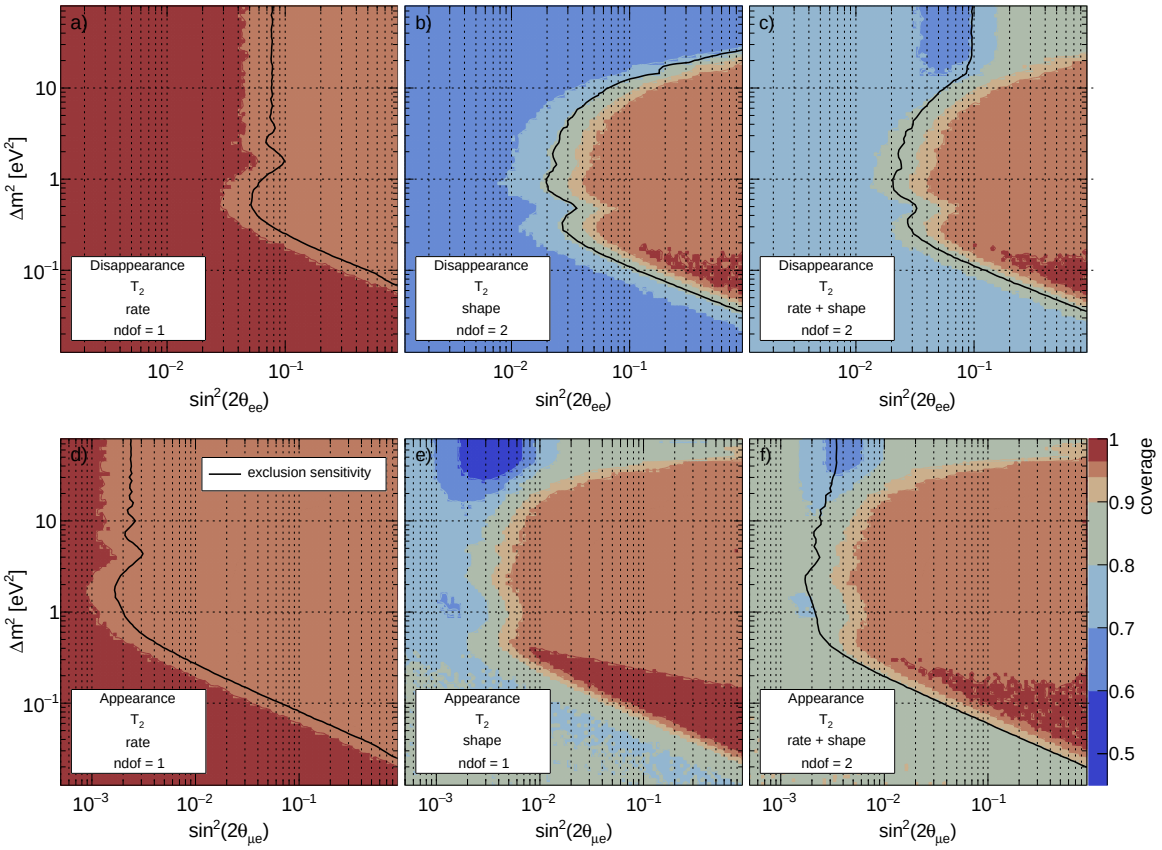


Figure A.10: Coverage of T_2 for the Toy Eperiments The coverage assuming Wilks' theorem are shown at the 95 % CL in a rate, shape, and rate+shape analysis for the toy disappearance and the toy appearance experiment. In addition, the exclusion sensitivities at the 95 % CL are shown in black. The assumed degrees of freedom (ndof) of the chi-square functions are defined as the number of uncorrelated parameters in the respective analysis.

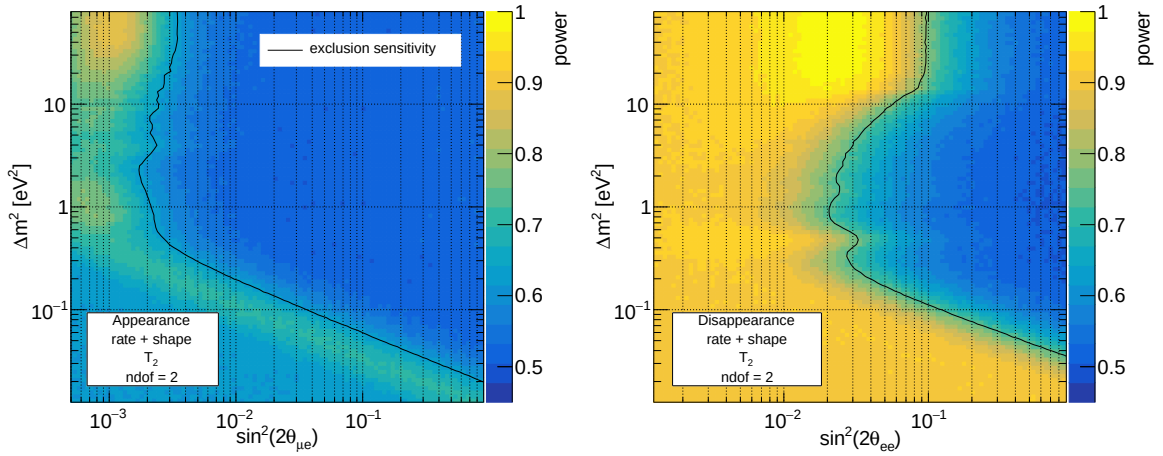


Figure A.11: Power of T_2 for the Toy Experiments The power is computed as the fraction of the alternative distribution above the expected median value and is shown for the rate+shape analysis for the toy appearance and toy disappearance experiment. The nominal value corresponds to 0.5. In addition, the exclusion sensitivities at the 95 % CL are shown.

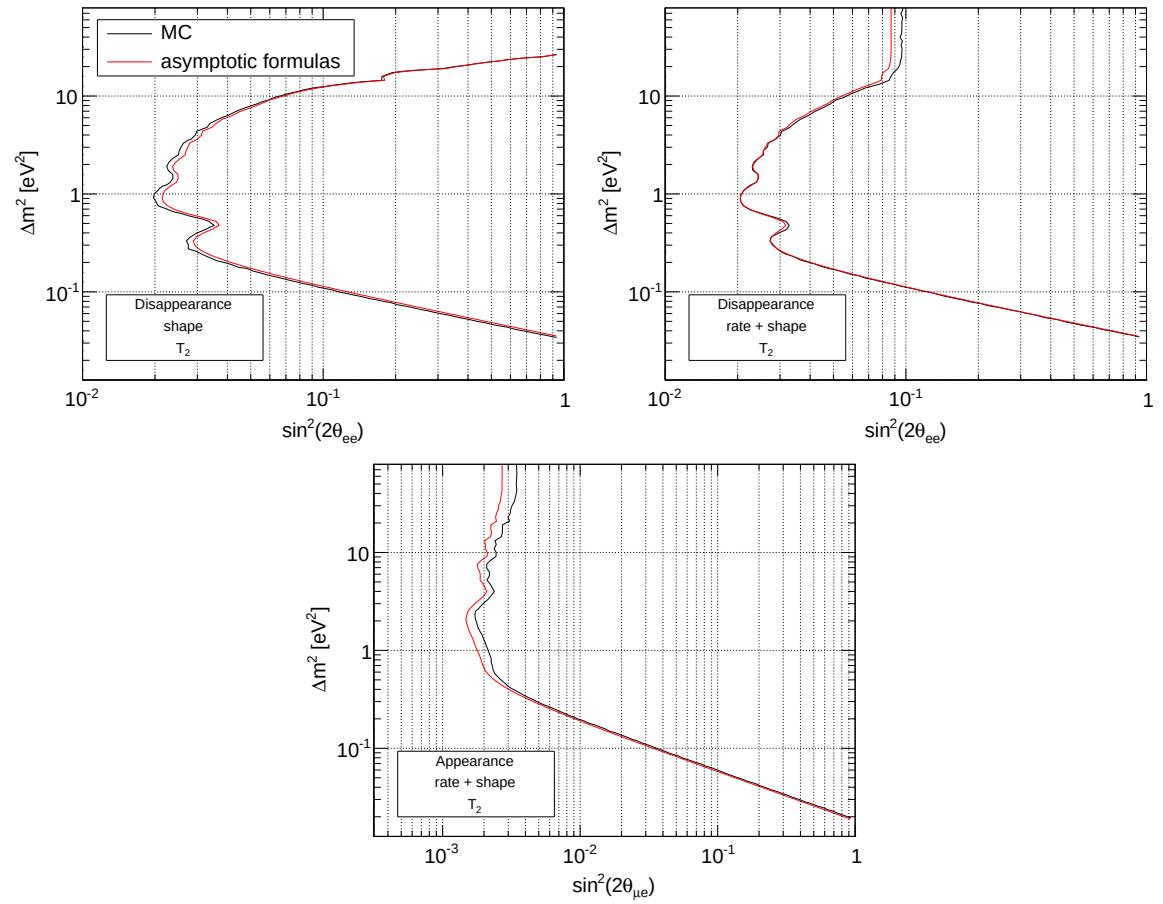


Figure A.12: Comparison of Sensitivities of Toy Experiments The exclusion sensitivities of T_2 at the 95 % CL are shown using a MC construction and the asymptotic probability distributions of the test statistic for a shape and rate+shape analysis of the toy disappearance experiment and the rate+shape analysis of the toy appearance experiment.

A.7 Connection of the Sensitivities of T_0 and T_1

In Figure 10.8 it was shown that the exclusion sensitivities for T_0+CL_S and T_1 overlap, which will be proven in the following. The proof is based on the asymptotic formulae for the probability distributions of T_0 and T_1 which are valid for both test statistics in the oscillation parameter space close to the sensitivity (see Chapter 11). Thus, the null and the alternative distributions of T_0 follow approximately Normal distributions

$$f\left(t_{(x,y)}|(x,y)\right) \sim \mathcal{N}\left(t_{(x,y)}^{A(x,y)}, 4t_{(x,y)}^{A(x,y)}\right) \tag{A.8}$$

$$f\left(t_{(x,y)}|(0,0)\right) \sim \mathcal{N}\left(t_{(x,y)}^{A(0,0)}, 4t_{(x,y)}^{A(0,0)}\right), \tag{A.9}$$

where $t_{(x,y)}^{A(x,y)}$ ($t_{(x,y)}^{A(0,0)}$) is the value of the test statistic for the Asimov-data set of $\mathrm{H}(x,y)$ ($\mathrm{H}(0,0)$) as defined in Formula 3.12 and $t_{(x,y)}^{A(x,y)} \approx t_{(x,y)}^{A(0,0)}$. Moreover, the null and the alternative distribution of T_1 can be described by a central and non-central chi-square distribution with one degree of freedom χ_1^2, respectively. Here, the non-centrality parameter of the alternative distribution can be approximated with $t_{(x,y)}^{A(0,0)}$. Therefore, the median of the alternative distributions for the test of the same hypothesis is the same for T_0 and T_1 and is given by $t_{(x,y)}^{A(0,0)}$, which will be abbreviated with t in the following.
The exclusion sensitivity is defined for these hypotheses, where the integral of the null distribution above the median value of the alternative distribution is given by the test size α:

$$\int_{\mathrm{med}}^{\infty} f\left(t_{(x,y)}|(x,y)\right) = \alpha. \tag{A.10}$$

In the case of T_0 one obtains:

$$\alpha(\mathrm{T}_0) = \int_{\mathrm{med}}^{\infty} f\left(t_{(x,y)}|(x,y)\right) = \int_t^{\infty} \mathcal{N}(-t, 4t) \tag{A.11}$$

$$= \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{t-(-t)}{2\sqrt{2t}}\right)\right) = \frac{1}{2}\left(1 - \mathrm{erf}(\sqrt{\frac{t}{2}})\right) \tag{A.12}$$

$$= \frac{1}{2}\left(1 - \frac{1}{\sqrt{\pi}}\gamma(\frac{1}{2}, \frac{t}{2})\right), \tag{A.13}$$

with $\int_x^{\infty} \mathcal{N}(\mu, \sigma^2) = \frac{1}{2}\left(1 - \mathrm{erf}(\frac{x-\mu}{\sqrt{2}\sigma})\right)$ and $\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}}\gamma(\frac{1}{2}, x^2)$, where $\mathrm{erf}(x)$ and $\gamma(a,x)$ denote the error and the lower incomplete gamma function, respectively.
In the case of T_1 one finds:

$$\alpha(\mathrm{T}_1) = \int_{\mathrm{med}}^{\infty} f\left(t_{(x,y)}|(x,y)\right) = \int_t^{\infty} \chi_1^2 = 1 - \frac{\gamma(\frac{1}{2}, \frac{t}{2})}{\Gamma(\frac{1}{2})} = 1 - \frac{\gamma(\frac{1}{2}, \frac{t}{2})}{\sqrt{\pi}}, \tag{A.14}$$

with the gamma function $\Gamma(x)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
Thus,

$$\alpha(\mathrm{T}_0) = \frac{1}{2}\alpha(\mathrm{T}_1). \tag{A.15}$$

This means that the confidence levels of the sensitivities of T_0 and T_1 are connected: e.g. the sensitivity of T_1 with a 95 % confidence level corresponds to the 97.5 % confidence level for T_0.

Moreover, the test size of the T_0+CL_S method is connected to the test size of T_0 and hence to the one of T_1:

$$CL_S = \frac{\int_{med}^{\infty} f\left(t_{(x,y)}|(x,y)\right)}{\int_{med}^{\infty} f\left(t_{(x,y)}|(0,0)\right)} = \frac{\alpha(T_0)}{0.5} = \alpha(T_1). \tag{A.16}$$

Here, it is assumed that the median of the distribution of T_0+CL_S corresponds to the ratio of the integrals of the null and alternative distribution above the median value of the alternative distribution. This connection proves that the exclusion sensitivities are the same for T_0+CL_S and T_1 independent on the specific confidence level. Note that this connection holds only for the sensitivities but is not given for the reconstructed limits and confidence regions from a single data set.

A.8 Degrees of Freedom and Size of Confidence Region

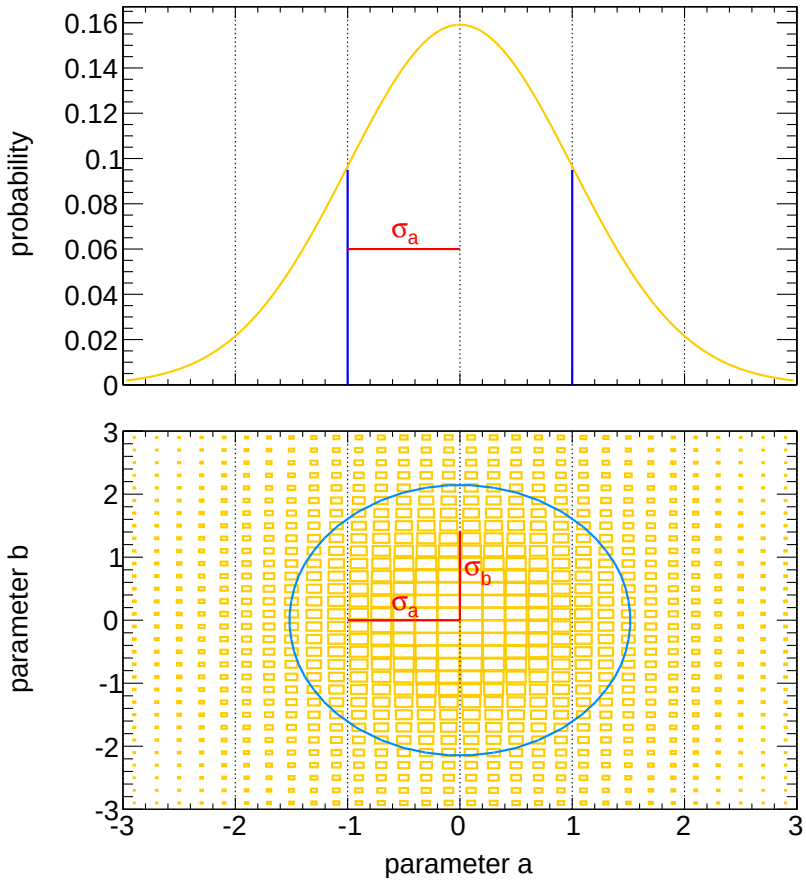


Figure A.13: Degrees of Freedom and Size of Confidence Region The Normal probability distributions are shown in yellow for one (top) and two (bottom) degrees of freedom, respectively. The standard deviation of the parameters a and b are indicted in red. The blue lines restrict the parameters which contain 68% of the probability distributions around the mean value and can be interpreted as the confidence regions. The confidence region of parameter a for $b = 0$ is larger than its standard deviation σ_a for two degrees of freedom. Hence, the confidence region of parameter a increases for two degrees of freedom in comparison to the confidence region with one degree of freedom. This illustration is supposed to visualize that the size of a confidence region increases with the number of free parameters and therefore with the degrees of freedom of a test.

List of Figures

List of Tables

Bibliography

[1] Particle Data Group collaboration, P. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[2] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [arXiv:1207.7214].

[3] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [arXiv:1207.7235].

[4] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group collaboration, S. Schael et al., *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257 [hep-ex/0509008].

[5] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes and R.P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, *Phys. Rev.* **105** (1957) 1413.

[6] M. Goldhaber, L. Grodzins and A.W. Sunyar, *Helicity of Neutrinos*, *Phys. Rev.* **109** (1958) 1015.

[7] C. Giunti and T. Lasserre, *eV-scale Sterile Neutrinos*, *Ann. Rev. Nucl. Part. Sci.* **69** (2019) 163 [arXiv:1901.08330].

[8] M. Drewes, *Auf dem Weg zu neuen Ufern: Sterile Neutrinos könnten Fragen aus Teilchenphysik und Kosmologie klären*, *Physic Journal 18* **2** (2019) 28.

[9] S. Bilenky, *Neutrino in Standard Model and beyond*, *Phys. Part. Nucl.* **46** (2015) 475 [arXiv:1501.00232].

[10] L. Wolfenstein, *Neutrino oscillations in matter*, *Phys. Rev. D* **17** (1978) 2369.

[11] S. Mikheyev and A. Smirnov, *Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos*, *Sov. J. Nucl. Phys.* **42** (1985) 913.

[12] A. Smirnov, *The MSW effect and matter effects in neutrino oscillations*, *Phys. Scripta T* **121** (2005) 57 [hep-ph/0412391].

[13] E. Vitagliano, I. Tamborra and G. Raffelt, *Grand Unified Neutrino Spectrum at Earth: Sources and Spectral Components*, *Rev. Mod. Phys.* **92** (2020) 045006 [arXiv:1910.11878].

[14] P. Padovani, *Gamma-ray emitting AGN and unified schemes*, in *32nd Rencontres de Moriond: High-Energy Phenomena in Astrophysics*, p. 7, 1997, astro-ph/9704002.

[15] IceCube, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift NuSTAR, VERITAS, VLA/17B-403 collaboration, M. Aartsen et al., *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, *Science* **361** (2018) 1378 [arXiv:1807.08816].

[16] IceCube collaboration, M. Aartsen et al., *Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert*, *Science* **361** (2018) 147 [arXiv:1807.08794].

[17] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [arXiv:1807.06209].

[18] IceCube collaboration, H. Niederhausen, *Recent IceCube Measurements Using High Energy Neutrinos*, in *18th Conference on Elastic and Diffractive Scattering*, 2019, arXiv:1909.12182.

[19] E. Adelberger et al., *Solar fusion cross sections II: the pp chain and CNO cycles*, *Rev. Mod. Phys.* **83** (2011) 195 [arXiv:1004.2318].

[20] C. Pena-Garay and A. Serenelli, *Solar neutrinos and the solar composition problem*, arXiv:0811.2424.

[21] Particle Data Group collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys. C* **40** (2016) 100001.

[22] J.N. Bahcall, A.M. Serenelli and S. Basu, *New solar opacities, abundances, helioseismology, and neutrino fluxes*, *Astrophys. J. Lett.* **621** (2005) L85 [astro-ph/0412440].

[23] J.N. Bahcall, W.F. Huebner, S.H. Lubow, P.D. Parker and R.K. Ulrich, *Standard solar models and the uncertainties in predicted capture rates of solar neutrinos*, *Rev. Mod. Phys.* **54** (1982) 767.

[24] R. Davis, *Nobel Lecture: A half-century with solar neutrinos*, *Rev. Mod. Phys.* **75** (2003) 985.

[25] J. Abdurashitov, E. Faizov, V. Gavrin, A. Gusev, A. Kalikhov, T. Knodel et al., *Results from SAGE (The Russian-American gallium solar neutrino experiment)*, *Physics Letters B* **328** (1994) 234 .

[26] P. Anselmann, W. Hampel, G. Heusser, J. Kiko, T. Kirsten, E. Pernicka et al., *Solar neutrinos observed by GALLEX at Gran Sasso*, *Physics Letters B* **285** (1992) 376 .

[27] K.S. Hirata, T. Kajita, T. Kifune, K. Kihara, M. Nakahata, K. Nakamura et al., *Observation of* ^{8}B *solar neutrinos in the Kamiokande-II detector*, *Phys. Rev. Lett.* **63** (1989) 16.

[28] SNO collaboration, Q.R. Ahmad et al., *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301.

[29] KamLAND collaboration, A. Gando et al., *^{7}Be Solar Neutrino Measurement with KamLAND*, *Phys. Rev. C* **92** (2015) 055808 [arXiv:1405.6190].

[30] Borexino collaboration, C. Arpesella et al., *First real time detection of ^{7}Be solar neutrinos by Borexino*, *Phys. Lett. B* **658** (2008) 101 [arXiv:0708.2251].

[31] KamLAND collaboration, S. Abe et al., *Measurement of the ^{8}B Solar Neutrino Flux with the KamLAND Liquid Scintillator Detector*, *Phys. Rev. C* **84** (2011) 035804 [arXiv:1106.0861].

[32] BOREXINO collaboration, M. Agostini et al., *Comprehensive measurement of pp-chain solar neutrinos*, *Nature* **562** (2018) 505.

[33] BOREXINO collaboration, M. Agostini et al., *Experimental evidence of neutrinos produced in the CNO fusion cycle in the Sun*, *Nature* **587** (2020) 577 [arXiv:2006.15115].

[34] *John Bahcall's Solar Neutrinos Web Page*, Accessed: 2020-11-27, http://www.sns.ias.edu/~jnb/.

[35] Borexino collaboration, M. Agostini et al., *Comprehensive geoneutrino analysis with Borexino*, *Phys. Rev. D* **101** (2020) 012009 [arXiv:1909.02257].

[36] KamLAND collaboration, T. Araki et al., *Experimental investigation of geologically produced antineutrinos with KamLAND*, *Nature* **436** (2005) 499.

[37] H.T. Janka, *Neutrino Emission from Supernovae*, arXiv:1702.08713.

[38] K. Hirata, T. Kajita, M. Koshiba, M. Nakahata, Y. Oyama, N. Sato et al., *Observation of a neutrino burst from the supernova SN1987A*, *Phys. Rev. Lett.* **58** (1987) 1490.

[39] R.M. Bionta, G. Blewitt, C.B. Bratton, D. Casper, A. Ciocio, R. Claus et al., *Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud*, *Phys. Rev. Lett.* **58** (1987) 1494.

[40] E. Alexeyev, L. Alexeyeva, I. Krivosheina and V. Volchenko, *Detection of the neutrino signal from SN 1987A in the LMC using the INR Baksan underground scintillation telescope*, *Physics Letters B* **205** (1988) 209 .

[41] J. Sawatzki, M. Wurm and D. Kresse, *Detecting the Diffuse Supernova Neutrino Background in the future Water-based Liquid Scintillator Detector Theia*, *Phys. Rev. D* **103** (2021) 023021 [arXiv:2007.14705].

[42] JUNO collaboration, F. An et al., *Neutrino Physics with JUNO*, *J. Phys. G* **43** (2016) 030401 [arXiv:1507.05613].

[43] J. Sawatzki, *Towards the Detection of the Diffuse Supernova Neutrino Background in the Large Volume Scintillator Experiment JUNO*, Ph.D. thesis, Technische Universität München, 2020.

[44] Hyper-Kamiokande collaboration, K. Abe et al., *Hyper-Kamiokande Design Report*, arXiv:1805.04163.

[45] K. Abe et al., *Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —*, arXiv:1109.3262.

[46] DUNE collaboration, B. Abi et al., *The DUNE Far Detector Interim Design Report Volume 1: Physics, Technology and Strategies*, arXiv:1807.10334.

[47] A. Cocco, A. Ereditato, G. Fiorillo, G. Mangano and V. Pettorino, *Supernova relic neutrinos in liquid argon detectors*, *JCAP* **12** (2004) 002 [hep-ph/0408031].

[48] K. Møller, A.M. Suliga, I. Tamborra and P.B. Denton, *Measuring the supernova unknowns at the next-generation neutrino telescopes through the diffuse neutrino background*, *JCAP* **05** (2018) 066 [arXiv:1804.03157].

[49] A. Mirizzi, I. Tamborra, H.T. Janka, N. Saviano, K. Scholberg, R. Bollig et al., *Supernova Neutrinos: Production, Oscillations and Detection*, *Riv. Nuovo Cim.* **39** (2016) 1 [arXiv:1508.00785].

[50] P. Antonioli, R.T. Fienberg, F. Fleurot, Y. Fukuda, W. Fulgione, A. Habig et al., *SNEWS: the SuperNova Early Warning System*, *New Journal of Physics* **6** (2004) 114.

[51] Borexino collaboration, M. Agostini et al., *Search for low-energy neutrinos from astrophysical sources with Borexino*, *Astropart. Phys.* **125** (2021) 102509 [arXiv:1909.02422].

[52] C.V. Achar et al., *Detection of muons produced by cosmic ray neutrinos deep underground*, *Phys. Lett.* **18** (1965) 196.

[53] F. Reines, M.F. Crouch, T.L. Jenkins, W.R. Kropp, H.S. Gurr, G.R. Smith et al., *Evidence for High-Energy Cosmic-Ray Neutrino Interactions*, *Phys. Rev. Lett.* **15** (1965) 429.

[54] Kamiokande-II collaboration, K. Hirata et al., *Observation of a small atmospheric muon-neutrino / electron-neutrino ratio in Kamiokande*, *Phys. Lett. B* **280** (1992) 146.

[55] D. Casper, R. Becker-Szendy, C.B. Bratton, D.R. Cady, R. Claus, S.T. Dye et al., *Measurement of atmospheric neutrino composition with the IMB-3 detector*, *Phys. Rev. Lett.* **66** (1991) 2561.

[56] Frejus collaboration, K. Daum et al., *Determination of the atmospheric neutrino spectra with the Frejus detector*, *Z. Phys. C* **66** (1995) 417.

[57] NUSEX collaboration, M. Aglietta et al., *Experimental study of atmospheric neutrino flux in the NUSEX experiment*, *Europhys. Lett.* **8** (1989) 611.

[58] Super-Kamiokande collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562 [hep-ex/9807003].

[59] MACRO collaboration, M. Ambrosio et al., *Matter effects in upward going muons and sterile neutrino oscillations*, *Phys. Lett. B* **517** (2001) 59 [hep-ex/0106049].

[60] Soudan 2 collaboration, M.C. Sanchez et al., *Measurement of the L/E distributions of atmospheric neutrinos in Soudan 2 and their interpretation as neutrino oscillations*, *Phys. Rev. D* **68** (2003) 113004 [hep-ex/0307069].

[61] IceCube collaboration, M. Aartsen et al., *Measurement of Atmospheric Tau Neutrino Appearance with IceCube DeepCore*, *Phys. Rev. D* **99** (2019) 032007 [arXiv:1901.05366].

[62] ANTARES collaboration, A. Albert et al., *Measuring the atmospheric neutrino oscillation parameters and constraining the 3+1 neutrino model with ten years of ANTARES data*, *JHEP* **06** (2019) 113 [arXiv:1812.08650].

[63] D. Jeschke, *Modulations of the Cosmic Muon Flux & Identification of Atmospheric Neutrino Interactions in Borexino*, Ph.D. thesis, Technische Universität München, 2018.

[64] J.M. Berryman and P. Huber, *Sterile Neutrinos and the Global Reactor Antineutrino Dataset*, arXiv:2005.01756.

[65] T.A. Mueller, D. Lhuillier, M. Fallot, A. Letourneau, S. Cormon, M. Fechner et al., *Improved predictions of reactor antineutrino spectra*, *Phys. Rev. C* **83** (2011) 054615.

[66] P. Huber, *Determination of antineutrino spectra from nuclear reactors*, *Phys. Rev. C* **84** (2011) 024617.

[67] G. Mention, M. Fechner, T. Lasserre, T. Mueller, D. Lhuillier, M. Cribier et al., *The Reactor Antineutrino Anomaly*, *Phys. Rev. D* **83** (2011) 073006 [arXiv:1101.2755].

[68] Daya Bay collaboration, D. Adey et al., *Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay*, *Phys. Rev. Lett.* **121** (2018) 241805 [arXiv:1809.02261].

[69] RENO collaboration, G. Bak et al., *Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO*, *Phys. Rev. Lett.* **121** (2018) 201801 [arXiv:1806.00248].

[70] Double Chooz collaboration, H. de Kerret et al., *Double Chooz θ_{13} measurement via total neutron capture detection*, *Nature Phys.* **16** (2020) 558 [arXiv:1901.09445].

[71] NEOS collaboration, Y. Ko et al., *Sterile Neutrino Search at the NEOS Experiment*, *Phys. Rev. Lett.* **118** (2017) 121802 [arXiv:1610.05134].

[72] M. Estienne et al., *Updated Summation Model: An Improved Agreement with the Daya Bay Antineutrino Fluxes*, *Phys. Rev. Lett.* **123** (2019) 022502 [arXiv:1904.09358].

[73] C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse and A.D. McGuire, *Detection of the Free Neutrino: a Confirmation*, *Science* **124** (1956) 103.

[74] KamLAND collaboration, T. Araki et al., *Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion*, *Phys. Rev. Lett.* **94** (2005) 081801 [hep-ex/0406035].

[75] Double Chooz collaboration, Y. Abe et al., *Indication of Reactor $\bar{\nu}_e$ Disappearance in the Double Chooz Experiment*, *Phys. Rev. Lett.* **108** (2012) 131801 [arXiv:1112.6353].

[76] Daya Bay collaboration, F. An et al., *Observation of electron-antineutrino disappearance at Daya Bay*, *Phys. Rev. Lett.* **108** (2012) 171803 [arXiv:1203.1669].

[77] RENO collaboration, J. Ahn et al., *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, *Phys. Rev. Lett.* **108** (2012) 191802 [arXiv:1204.0626].

[78] K2K collaboration, M. Ahn et al., *Measurement of Neutrino Oscillation by the K2K Experiment*, *Phys. Rev. D* **74** (2006) 072003 [hep-ex/0606032].

[79] MINOS collaboration, P. Adamson et al., *Combined analysis of ν_μ disappearance and $\nu_\mu \to \nu_e$ appearance in MINOS using accelerator and atmospheric neutrinos*, *Phys. Rev. Lett.* **112** (2014) 191801 [arXiv:1403.0867].

[80] OPERA collaboration, N. Agafonova et al., *Final Results of the OPERA Experiment on ν_τ Appearance in the CNGS Neutrino Beam*, *Phys. Rev. Lett.* **120** (2018) 211801 [arXiv:1804.04912].

[81] T2K collaboration, K. Abe et al., *Observation of Electron Neutrino Appearance in a Muon Neutrino Beam*, *Phys. Rev. Lett.* **112** (2014) 061802 [arXiv:1311.4750].

[82] NOνA collaboration, M. Acero et al., *First Measurement of Neutrino Oscillation Parameters using Neutrinos and Antineutrinos by NOνA*, *Phys. Rev. Lett.* **123** (2019) 151803 [arXiv:1906.04907].

[83] Borexino collaboration, P. Alvarez Sanchez et al., *Measurement of CNGS muon neutrino speed with Borexino*, *Phys. Lett. B* **716** (2012) 401 [arXiv:1207.6860].

[84] Y. Nakajima, *Recent results and future prospects from Super-Kamiokande*, Talk at Neutrino Conference, https://doi.org/10.5281/zenodo.3959640, 2020.

[85] I. Esteban, M. Gonzalez-Garcia, M. Maltoni, T. Schwetz and A. Zhou, *NuFIT 5.0* , http://www.nu-fit.org, 2020.

[86] IceCube collaboration, M. Aartsen et al., *Development of an analysis to probe the neutrino mass ordering with atmospheric neutrinos using three years of IceCube DeepCore data*, *Eur. Phys. J. C* **80** (2020) 9 [arXiv:1902.07771].

[87] F. Capozzi, E. Lisi and A. Marrone, *Probing the neutrino mass ordering with KM3NeT-ORCA: Analysis and perspectives*, *J. Phys. G* **45** (2018) 024003 [arXiv:1708.03022].

[88] T2K collaboration, K. Abe et al., *Search for CP Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with 2.2×10^{21} Protons on Target*, *Phys. Rev. Lett.* **121** (2018) 171802 [arXiv:1807.07891].

[89] I. Esteban, M. Gonzalez-Garcia, M. Maltoni, T. Schwetz and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, *JHEP* **09** (2020) 178 [arXiv:2007.14792].

[90] M.J. Dolinski, A.W. Poon and W. Rodejohann, *Neutrinoless Double-Beta Decay: Status and Prospects*, *Ann. Rev. Nucl. Part. Sci.* **69** (2019) 219 [arXiv:1902.04097].

[91] GERDA collaboration, M. Agostini et al., *Final Results of GERDA on the Search for Neutrinoless Double-β Decay*, *Phys. Rev. Lett.* **125** (2020) 252502 [arXiv:2009.06079].

[92] KamLAND-Zen collaboration, A. Gando et al., *Search for Majorana Neutrinos near the Inverted Mass Hierarchy Region with KamLAND-Zen*, *Phys. Rev. Lett.* **117** (2016) 082503 [arXiv:1605.02889].

[93] EXO-200 collaboration, G. Anton et al., *Search for Neutrinoless Double-β Decay with the Complete EXO-200 Dataset*, *Phys. Rev. Lett.* **123** (2019) 161802 [arXiv:1906.02723].

[94] M. Agostini, G. Benato and J. Detwiler, *Discovery probability of next-generation neutrinoless double- β decay experiments*, *Phys. Rev. D* **96** (2017) 053001 [arXiv:1705.02996].

[95] KATRIN collaboration, M. Aker et al., *Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN*, *Phys. Rev. Lett.* **123** (2019) 221802 [arXiv:1909.06048].

[96] L. Gastaldo et al., *The electron capture in^{163}Ho experiment –ECHo*, *Eur. Phys. J. ST* **226** (2017) 1623.

[97] S. Gariazzo, C. Giunti, M. Laveder, Y. Li and E. Zavanin, *Light sterile neutrinos*, *J. Phys. G* **43** (2016) 033001 [arXiv:1507.08204].

[98] P. de Holanda and A. Smirnov, *Solar neutrino spectrum, sterile neutrinos and additional radiation in the Universe*, *Phys. Rev. D* **83** (2011) 113011 [arXiv:1012.5627].

[99] P. Bakhti and Y. Farzan, *Constraining Super-light Sterile Neutrino Scenario by JUNO and RENO-50*, *JHEP* **10** (2013) 200 [arXiv:1308.2823].

[100] KATRIN collaboration, S. Mertens et al., *A novel detector system for KATRIN to search for keV-scale sterile neutrinos*, *J. Phys. G* **46** (2019) 065203 [arXiv:1810.06711].

[101] NA62 collaboration, E. Cortina Gil et al., *Search for heavy neutral lepton production in K^+ decays to positrons*, *Phys. Lett. B* **807** (2020) 135599 [arXiv:2005.09575].

[102] ATLAS collaboration, G. Aad et al., *Search for heavy neutral leptons in decays of W bosons produced in 13 TeV pp collisions using prompt and displaced signatures with the ATLAS detector*, *JHEP* **10** (2019) 265 [arXiv:1905.09787].

[103] CMS collaboration, A.M. Sirunyan et al., *Search for heavy neutral leptons in events with three charged leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **120** (2018) 221801 [arXiv:1802.02965].

[104] P. Minkowski, *$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?*, *Phys. Lett. B* **67** (1977) 421.

[105] M. Gell-Mann, P. Ramond and R. Slansky, *Complex Spinors and Unified Theories*, *Conf. Proc. C* **790927** (1979) 315 [arXiv:1306.4669].

[106] R.N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Nonconservation*, *Phys. Rev. Lett.* **44** (1980) 912.

[107] T. Yanagida, *Horizontal Symmetry and Masses of Neutrinos*, *Prog. Theor. Phys.* **64** (1980) 1103.

[108] J. Schechter and J. Valle, *Neutrino Masses in SU(2) x U(1) Theories*, *Phys. Rev. D* **22** (1980) 2227.

[109] R. Slansky, S. Raby, J. Goldman and G. Garvey, *The oscillating neutrino: An introduction to neutrino masses and mixings*, *Los Alamos Sci.* **25** (1997) 28.

[110] T. Asaka, S. Blanchet and M. Shaposhnikov, *The νMSM, dark matter and neutrino masses*, *Physics Letters B* **631** (2005) 151 .

[111] T. Asaka and M. Shaposhnikov, *The νMSM, dark matter and baryon asymmetry of the universe*, *Physics Letters B* **620** (2005) 17 .

[112] L. Canetti, M. Drewes and M. Shaposhnikov, *Matter and Antimatter in the Universe*, *New J. Phys.* **14** (2012) 095012 [arXiv:1204.4186].

[113] M. Drewes et al., *A White Paper on keV Sterile Neutrino Dark Matter*, *JCAP* **01** (2017) 025 [arXiv:1602.04816].

[114] K. Abazajian et al., *Light Sterile Neutrinos: A White Paper*, arXiv:1204.5379.

[115] L. Funcke, G. Raffelt and E. Vitagliano, *Distinguishing Dirac and Majorana neutrinos by their decays via Nambu-Goldstone bosons in the gravitational-anomaly model of neutrino masses*, *Phys. Rev. D* **101** (2020) 015025 [arXiv:1905.01264].

[116] S. Böser, C. Buck, C. Giunti, J. Lesgourgues, L. Ludhova, S. Mertens et al., *Status of Light Sterile Neutrino Searches*, *Prog. Part. Nucl. Phys.* **111** (2020) 103736 [arXiv:1906.01739].

[117] LSND collaboration, A. Aguilar-Arevalo et al., *Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam*, *Phys. Rev.* **D64** (2001) 112007.

[118] KARMEN collaboration, B. Armbruster et al., *Upper limits for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ from muon decay at rest*, *Phys. Rev. D* **65** (2002) 112001 [hep-ex/0203021].

[119] MiniBooNE collaboration, A. Aguilar-Arevalo et al., *A Search for Electron Neutrino Appearance at the $\Delta m^2 \sim 1eV^2$ Scale*, *Phys. Rev. Lett.* **98** (2007) 231801 [arXiv:0704.1500].

[120] MiniBooNE collaboration, A.A. Aguilar-Arevalo et al., *Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment*, *Phys. Rev. Lett.* **121** (2018) 221801.

[121] D. Frekers et al., *The Ga-71(He-3, t) reaction and the low-energy neutrino response*, *Phys. Lett. B* **706** (2011) 134.

[122] C. Giunti, M. Laveder, Y. Li and H. Long, *Pragmatic View of Short-Baseline Neutrino Oscillations*, *Phys. Rev. D* **88** (2013) 073008 [arXiv:1308.5288].

[123] M. Dentler, A. Hernández-Cabezudo, J. Kopp, P.A. Machado, M. Maltoni, I. Martinez-Soler et al., *Updated Global Analysis of Neutrino Oscillations in the Presence of eV-Scale Sterile Neutrinos*, *JHEP* **08** (2018) 010 [arXiv:1803.10661].

[124] J. Conrad, C. Ignarra, G. Karagiorgi, M. Shaevitz and J. Spitz, *Sterile Neutrino Fits to Short Baseline Neutrino Oscillation Measurements*, *Adv. High Energy Phys.* **2013** (2013) 163897 [arXiv:1207.4765].

[125] M. Archidiacono, N. Fornengo, C. Giunti, S. Hannestad and A. Melchiorri, *Sterile neutrinos: Cosmology versus short-baseline experiments*, *Phys. Rev. D* **87** (2013) 125034 [arXiv:1302.6720].

[126] DANSS collaboration, I. Alekseev et al., *Search for sterile neutrinos at the DANSS experiment*, *Phys. Lett.* **B787** (2018) 56.

[127] NEUTRINO-4 collaboration, A.P. Serebrov et al., *First Observation of the Oscillation Effect in the Neutrino-4 Experiment on the Search for the Sterile Neutrino*, *JETP Letters* **109** (2019) 213.

[128] PROSPECT collaboration, J. Ashenfelter et al., *First search for short-baseline neutrino oscillations at HFIR with PROSPECT*, *Phys. Rev. Lett.* **121** (2018) 251802.

[129] SoLid collaboration, Y. Abreu et al., *SoLid: A short baseline reactor neutrino experiment*, arXiv:2002.05914.

[130] STEREO collaboration, H. Almazán et al., *Sterile Neutrino Constraints from the STEREO Experiment with 66 Days of Reactor-On Data*, *Phys. Rev. Lett.* **121** (2018) 161801.

[131] V. Barinov, V. Gavrin, D. Gorbunov and T. Ibragimova, *BEST sensitivity to O(1) eV sterile neutrino*, *Phys. Rev. D* **93** (2016) 073002 [arXiv:1602.03826].

[132] Borexino collaboration, G. Bellini et al., *SOX: Short distance neutrino Oscillations with BoreXino*, *JHEP* **08** (2013) 038.

[133] J. Gaffiot et al., *Experimental Parameters for a Cerium 144 Based Intense Electron Antineutrino Generator Experiment at Very Short Baselines*, *Phys. Rev. D* **91** (2015) 072005 [arXiv:1411.6694].

[134] S. Ajimura et al., *Technical Design Report (TDR): Searching for a Sterile Neutrino at J-PARC MLF (E56, JSNS2)*, arXiv:1705.08629.

[135] MicroBooNE, LAr1-ND, ICARUS-WA104 collaboration, M. Antonello et al., *A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam*, arXiv:1503.01520.

[136] IceCube collaboration, M.G. Aartsen et al., *Search for sterile neutrino mixing using three years of IceCube DeepCore data*, *Phys. Rev.* **D95** (2017) 112002.

[137] X. Qian, C. Zhang, M. Diwan and P. Vogel, *Unitarity Tests of the Neutrino Mixing Matrix*, arXiv:1308.5700.

[138] KATRIN collaboration, M. Aker et al., *Bound on 3+1 active-sterile neutrino mixing from the first four-week science run of KATRIN*, arXiv:2011.05087.

[139] Borexino collaboration, M. Agostini et al., *First Simultaneous Precision Spectroscopy of pp, ^{7}Be, and pep Solar Neutrinos with Borexino Phase-II*, *Phys. Rev. D* **100** (2019) 082004 [arXiv:1707.09279].

[140] J.Q. Meindl, *Reconstruction and Measurement of Cosmogenic Signals in the Neutrino Experiment Borexino*, Ph.D. thesis, Technische Universität München, 2013.

[141] Borexino collaboration, G. Bellini et al., *Final results of Borexino Phase-I on low energy solar neutrino spectroscopy*, *Phys. Rev. D* **89** (2014) 112007 [arXiv:1308.0443].

[142] A. de Gouvea and J. Jenkins, *What can we learn from neutrino electron scattering?*, *Phys. Rev. D* **74** (2006) 033004 [hep-ph/0603036].

[143] A. Strumia and F. Vissani, *Precise quasielastic neutrino/nucleon cross-section*, *Phys. Lett. B* **564** (2003) 42 [astro-ph/0302055].

[144] J. Benziger, M. Johnson, F. Calaprice, M. Chen, N. Darnton, R. Loeser et al., *A scintillator purification system for a large scale solar neutrino experiment*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **417** (1998) 278 .

[145] Borexino collaboration, *Borexino Experiment: Official Web Site*, Accessed: 2020-11-27, http://borex.lngs.infn.it.

[146] N. Vinyoles, A.M. Serenelli, F.L. Villante, S. Basu, J. Bergström, M. Gonzalez-Garcia et al., *A new Generation of Standard Solar Models*, *Astrophys. J.* **835** (2017) 202 [arXiv:1611.09867].

[147] Borexino collaboration, G. Bellini et al., *Precision measurement of the ^7Be solar neutrino interaction rate in Borexino*, *Phys. Rev. Lett.* **107** (2011) 141302 [arXiv:1104.1816].

[148] Borexino collaboration, G. Bellini et al., *First evidence of pep solar neutrinos by direct detection in Borexino*, *Phys. Rev. Lett.* **108** (2012) 051302 [arXiv:1110.3230].

[149] Borexino collaboration, G. Bellini 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* **82** (2010) 033006 [arXiv:0808.2868].

[150] BOREXINO collaboration, G. Bellini et al., *Neutrinos from the primary proton–proton fusion process in the Sun*, *Nature* **512** (2014) 383.

[151] BOREXINO collaboration, M. Agostini et al., *Sensitivity to neutrinos from the solar CNO cycle in Borexino*, *Eur. Phys. J. C* **80** (2020) 1091 [arXiv:2005.12829].

[152] Borexino collaboration, M. Agostini et al., *Improved measurement of ^8B solar neutrinos with $1.5 kt \cdot y$ of Borexino exposure*, *Phys. Rev. D* **101** (2020) 062001 [arXiv:1709.00756].

[153] Super-Kamiokande Collaboration collaboration, K. Abe, Y. Haga, Y. Hayato, M. Ikeda, K. Iyogi, J. Kameda et al., *Solar neutrino measurements in Super-Kamiokande-IV*, *Phys. Rev. D* **94** (2016) 052010.

[154] C. Fröhlich and J. Lean, *The Sun's total irradiance: Cycles, trends and related climate change uncertainties since 1976*, *Geophysical Research Letters* **25** (1998) 4377.

[155] Borexino collaboration, M. Agostini et al., *Spectroscopy of geoneutrinos from 2056 days of Borexino data*, *Phys. Rev. D* **92** (2015) 031101 [arXiv:1506.04610].

[156] Borexino collaboration, G. Bellini et al., *Measurement of geo-neutrinos from 1353 days of Borexino*, *Phys. Lett. B* **722** (2013) 295 [arXiv:1303.2571].

[157] Borexino collaboration, G. Bellini et al., *Observation of Geo-Neutrinos*, *Phys. Lett. B* **687** (2010) 299 [arXiv:1003.0284].

[158] A. Rocholl and K. Jochum, *Th, U and other trace elements in carbonaceous chondrites: Implications for the terrestrial and solar-systemTh/U ratios*, *Earth and Planetary Science Letters* **117** (1993) 265 .

[159] S.A. Wipperfurth, M. Guo, O. Šrámek and W.F. McDonough, *Earth's chondritic Th/U: Negligible fractionation during accretion, core formation, and crust-mantle differentiation*, *Earth and Planetary Science Letters* **498** (2018) 196.

[160] M. Coltorti, R. Boraso, F. Mantovani, M. Morsilli, G. Fiorentini, A. Riva et al., *U and Th content in the Central Apennines continental crust: A contribution to the determination of the geo-neutrinos flux at LNGS*, *Geochimica et Cosmochimica Acta* **75** (2011) 2271 .

[161] B. Neumair, *Sensitivitätsstudien und Messung des ^{144}Ce -^{144}Pr Antineutrinospektrums für das SOX Experiment*, Master's thesis, Technische Universität München, 2014.

[162] M.M. Bé, V. Chisté, C. Dulieu, E. Browne, V. Chechev, N. Kuzmenko et al., *Table of Radionuclides*, vol. 1 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2004), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol1.pdf.

[163] D. Bravo-Berguño, *Precision Background Stability and Response Calibration in Borexino: Prospects for Wideband, Precision Solar Neutrino Spectroscopy and BSM Neutrino Oscillometry Through a Deeper Detector Understanding*, Ph.D. thesis, Virginia Tech., Blacksburg, 2016.

[164] M.M. Bé, V. Chisté, C. Dulieu, M. Kellett, X. Mougeot, A. Arinc et al., *Table of Radionuclides*, vol. 8 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2016), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol8.pdf.

[165] M. Meyer, *SOX - Towards the Detection of Sterile Neutrinos in Borexino : Beta Spectrum Modeling, Monte Carlo Development and Sensitivity Studies for the Sterile Neutrino Search in Borexino*, Ph.D. thesis, Universität Hamburg, 2016.

[166] K.S. Krane, *Introductory nuclear physics*, John Wiley & Sons, New York, NY (1988).

[167] K.M. Altenmüller, *Search for sterile neutrinos in beta decays*, Ph.D. thesis, Technische Universität München, 2019.

[168] T. Houdy, *Testing non-standard neutrino interactions from ^{8}B solar neutrino rate measurement with Borexino to characterization of the ^{144}Ce source of the SOX experiment.*, Ph.D. thesis, Université Sorbonne, Paris, 2017.

[169] T. Lasserre, *CeANG Status*, Talk at SOX Collaboration Meeting, 2017.

[170] K. Altenmüller et al., *A calorimeter for the precise determination of the activity of the ^{144}Ce-^{144}Pr anti-neutrino source in the SOX experiment*, *JINST* **13** (2018) P09008.

[171] N. Haag, A. Gütlein, M. Hofmann, L. Oberauer, W. Potzel, K. Schreckenbach et al., *Experimental Determination of the Antineutrino Spectrum of the Fission Products of ^{238}U*, *Phys. Rev. Lett.* **112** (2014) 122501 [arXiv:1312.5601].

[172] N.H. Haag, *Experimental Determination of the Antineutrino Spectrum of the Fission Products of U-238*, Ph.D. thesis, Technische Universität München, 2013.

[173] GEANT4 collaboration, S. Agostinelli et al., *GEANT4–a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[174] A. Sharma, *Properties of some gas mixtures used in tracking detectors*, preprint SLAC-JOURNAL-ICFA-16-3, 1998.

[175] G. Casella and R.L. Berger, *Statistical inference*, Thomson Learning, Australia Pacific Grove, CA (2002).

[176] Particle Data Group collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev.* **D98** (2018) 030001.

[177] G.J. Feldman and R.D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys. Rev.* **D57** (1998) 3873.

[178] S.S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (1938) 60.

[179] L. Lyons, *A Paradox about Likelihood Ratios?*, arXiv:1711.00775.

[180] A. Wald, *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, *Trans. Am. Math. Soc.* **54** (1943) 426.

[181] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554.

[182] H. Niederhausen, *Measurement of the High Energy Astrophysical Neutrino Flux and Tau Neutrinos Observed in Four Years of IceCube Data*, Ph.D. thesis, State University of New York at Stony Brook, 2018.

[183] X. Qian et al., *The Gaussian CL_s method for searches of new physics*, *Nucl. Instrum. Meth.* **A827** (2016) 63.

[184] Borexino collaboration, M. Agostini et al., *The Monte Carlo simulation of the Borexino detector*, *Astropart. Phys.* **97** (2018) 136 [arXiv:1704.02291].

[185] P. Vogel and J.F. Beacom, *Angular distribution of neutron inverse beta decay, $\bar{\nu}_e + p \to e^+ + n$*, *Phys. Rev. D* **60** (1999) 053003 [hep-ph/9903554].

[186] J. Coon and R. Nobles, *Disintegration of He^3 and N^{14} by Thermal Neutrons*, *Physical Review - PHYS REV X* **75** (1949) 1358.

[187] M. Agostini, *M-STATS: framework for frequentist statistical analysis*, https://github.com/mmatteo/m-stats.

[188] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343.

[189] Borexino collaboration, H. Back et al., *Borexino calibrations: Hardware, Methods, and Results*, *JINST* **7** (2012) P10018 [arXiv:1207.4816].

[190] M. Nieslony, *Optimization of the antineutrino detection efficiency in the frame of the sterile neutrino search with SOX*, Master's thesis, Karlsruhe Institute of Technology, 2017.

[191] M.M. Bé, V. Chisté, C. Dulieu, X. Mougeot, V. Chechev, F. Kondev et al., *Table of Radionuclides*, vol. 7 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2013), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol7.pdf.

[192] BOREXINO collaboration, M. Agostini et al., *Seasonal Modulation of the ^{7}Be Solar Neutrino Rate in Borexino*, *Astropart. Phys.* **92** (2017) 21 [arXiv:1701.07970].

[193] K. Choi et al., *Detector Response*, Internal SOX Report, 2017.

[194] Borexino collaboration, G. Bellini et al., *Study of solar and other unknown anti-neutrino fluxes with Borexino at LNGS*, *Phys. Lett. B* **696** (2011) 191 [arXiv:1010.0029].

[195] T. Mueller et al., *Improved Predictions of Reactor Antineutrino Spectra*, *Phys. Rev. C* **83** (2011) 054615 [arXiv:1101.2663].

[196] RENO collaboration, J. Choi et al., *Observation of Energy and Baseline Dependent Reactor Antineutrino Disappearance in the RENO Experiment*, *Phys. Rev. Lett.* **116** (2016) 211801 [arXiv:1511.05849].

[197] Daya Bay collaboration, F.P. An et al., *Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay*, *Phys. Rev. Lett.* **116** (2016) 061801 [arXiv:1508.04233].

[198] Double Chooz collaboration, Y. Abe et al., *Improved measurements of the neutrino mixing angle θ_{13} with the Double Chooz detector*, *JHEP* **10** (2014) 086 [arXiv:1406.7763].

[199] S. Förster, *Improved Veto for Cosmogenic Background and Calibration Studies for the Borexino/SOX Experiment*, Ph.D. thesis, Johannes Gutenberg-Universität Mainz, 2018.

[200] Borexino collaboration, G. Bellini et al., *Cosmic-muon flux and annual modulation in Borexino at 3800 m water-equivalent depth*, *JCAP* **05** (2012) 015 [arXiv:1202.6403].

[201] M.M. Bé, V. Chisté, C. Dulieu, E. Browne, V. Chechev, N. Kuzmenko et al., *Table of Radionuclides*, vol. 2 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2004), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol2.pdf.

[202] M.M. Bé, V. Chisté, C. Dulieu, E. Browne, V. Chechev, N. Kuzmenko et al., *Table of Radionuclides*, vol. 4 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2008), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol4.pdf.

[203] A. Caminata, *Source-Induced Background*, Talk at SOX Collaboration Meeting, 2017.

[204] L. Hayen, J. Kostensalo, N. Severijns and J. Suhonen, *First-forbidden transitions in the reactor anomaly*, *Phys. Rev. C* **100** (2019) 054323 [arXiv:1908.08302].

[205] D. Basilico, B. Neumair and J. Martyn, *Search for sterile neutrinos with SOX: Monte Carlo studies of the experiment sensitivity and systematic effects*, *Journal of Physics: Conference Series* **1342** (2020) 012107.

[206] J. Martyn, *SOX Sensitivity Study on Detection Efficiency at the Vessel Edge and Improvement of Position Reconstruction*, Master's thesis, Johannes Gutenberg-Universität Mainz, 2016.

[207] M. Nieslony, *Monte Carlo Simulation in the Inner Buffer*, Talk at SOX Collaboration Meeting, 2017.

[208] M. Agostini and B. Neumair, *Statistical Methods Applied to the Search of Sterile Neutrinos*, *Eur. Phys. J. C* **80** (2020) 750 [arXiv:1906.11854].

[209] Double Chooz collaboration, T. Abrahão et al., *Search for Signatures of Sterile Neutrinos with Double Chooz*, arXiv:2009.05515.

[210] PROSPECT collaboration, M. Andriamirado et al., *Improved Short-Baseline Neutrino Oscillation Search and Energy Spectrum Measurement with the PROSPECT Experiment at HFIR*, arXiv:2006.11210.

[211] STEREO collaboration, H. Almazán et al., *Improved sterile neutrino constraints from the STEREO experiment with 179 days of reactor-on data*, *Phys. Rev. D* **102** (2020) 052002 [arXiv:1912.06582].

[212] P. Coloma, P. Huber and T. Schwetz, *Statistical interpretation of sterile neutrino oscillation searches at reactors*, *Eur. Phys. J. C* **81** (2021) 2 [arXiv:2008.06083].

[213] S. Silaeva and V. Sinev, *Simulation of an experiment on looking for sterile neutrinos at nuclear reactor*, arXiv:2001.10752.

[214] L. Lyons, *Raster scan or 2-D approach?*, arXiv:1404.7395.

[215] E.L. Lehmann and J.P. Romano, *Testing statistical hypotheses*, Springer Texts in Statistics, third ed., Springer, New York (2005).

[216] S. Algeri, D. van Dyk, J. Conrad and B. Anderson, *On methods for correcting for the look-elsewhere effect in searches for new physics*, *Journal of Instrumentation* **11** (2016) P12010.

[217] C. Giunti, *Statistical Significance of Reactor Antineutrino Active-Sterile Oscillations*, *Phys. Rev. D* **101** (2020) 095025 [arXiv:2004.07577].

[218] A.C. Davison, *Statistical Models*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press (2003), 10.1017/CBO9780511815850.

[219] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *Eur. Phys. J.* **C70** (2010) 525.

[220] P.B. Denton and J. Gehrlein, *A Statistical Analysis of the COHERENT Data and Applications to New Physics*, arXiv:2008.06062.

[221] S. Gariazzo, C. Giunti, M. Laveder and Y. Li, *Model-independent $\bar{\nu}_e$ short-baseline oscillations from reactor spectral ratios*, *Phys. Lett. B* **782** (2018) 13 [arXiv:1801.06467].

[222] J.M. Berryman and P. Huber, *Reevaluating Reactor Antineutrino Anomalies with Updated Flux Predictions*, *Phys. Rev. D* **101** (2020) 015008 [arXiv:1909.09267].

[223] B. Achkar, R. Aleksan, M. Avenier, G. Bagieu, J. Bouchez, R. Brissot et al., *Search for neutrino oscillations at 15, 40 and 95 meters from a nuclear power reactor at Bugey*, *Nuclear Physics B* **434** (1995) 503 , http://www.sciencedirect.com/science/article/pii/055032139400513E.

[224] C. Giunti, Y. Li and Y. Zhang, *KATRIN bound on 3+1 active-sterile neutrino mixing and the reactor antineutrino anomaly*, *JHEP* **05** (2020) 061 [arXiv:1912.12956].

[225] Y. Shitov, *Recent results from the DANSS experiment*, Talk at Neutrino Conference, https://doi.org/10.5281/zenodo.4124171, 2020.

[226] DANSS collaboration, M. Danilov, *Recent results of the DANSS experiment*, *PoS* **EPS-HEP2019** (2020) 401 [arXiv:1911.10140].

[227] S. Schoppmann, *Accurate Measurement of Electron Antineutrinos of U-235 Fissions from the STEREO experiment*, Talk at Neutrino Conference, https://doi.org/10.5281/zenodo.3959676, 2020.

[228] A. Serebrov et al., *Preparation of the Neutrino-4 experiment on search for sterile neutrino and the obtained results of measurements*, arXiv:2005.05301.

[229] PROSPECT, STEREO collaboration, M. Andriamirado et al., *Note on arXiv:2005.05301, 'Preparation of the Neutrino-4 experiment on search for sterile neutrino and the obtained results of measurements'*, arXiv:2006.13147.

[230] MiniBooNE collaboration, A. Aguilar-Arevalo et al., *Updated MiniBooNE Neutrino Oscillation Results with Increased Data and New Background Studies*, arXiv:2006.16883.

[231] A. Diaz, C. Argüelles, G. Collin, J. Conrad and M. Shaevitz, *Where Are We With Light Sterile Neutrinos?*, *Phys. Rept.* **884** (2020) 1 [arXiv:1906.00045].

[232] MINOS+ collaboration, P. Adamson et al., *Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit*, *Phys. Rev. Lett.* **122** (2019) 091803 [arXiv:1710.06488].

[233] IceCube collaboration, M. Aartsen et al., *eV-Scale Sterile Neutrino Search Using Eight Years of Atmospheric Muon Neutrino Data from the IceCube Neutrino Observatory*, *Phys. Rev. Lett.* **125** (2020) 141801 [arXiv:2005.12942].

[234] A. Merle, S. Morisi and W. Winter, *Common origin of reactor and sterile neutrino mixing*, *JHEP* **07** (2014) 039 [arXiv:1402.6332].

[235] K. Babu, D.W. McKay, I. Mocioiu and S. Pakvasa, *Light sterile neutrinos, lepton number violating interactions, and the LSND neutrino anomaly*, *Phys. Rev. D* **93** (2016) 113019 [arXiv:1605.03625].

[236] M. Carena, Y.Y. Li, C.S. Machado, P.A. Machado and C.E. Wagner, *Neutrinos in Large Extra Dimensions and Short-Baseline ν_e Appearance*, *Phys. Rev. D* **96** (2017) 095014 [arXiv:1708.09548].

[237] G. Magill, R. Plestid, M. Pospelov and Y.D. Tsai, *Dipole Portal to Heavy Neutral Leptons*, *Phys. Rev. D* **98** (2018) 115015 [arXiv:1803.03262].

[238] D. Döring, H. Päs, P. Sicking and T.J. Weiler, *Sterile neutrinos with altered dispersion relations as an explanation for neutrino anomalies*, *Eur. Phys. J. C* **80** (2020) 1202 [arXiv:1808.07460].

[239] J. Liao, D. Marfatia and K. Whisnant, *MiniBooNE, MINOS+ and IceCube data imply a baroque neutrino sector*, *Phys. Rev. D* **99** (2019) 015016 [arXiv:1810.01000].

[240] P.B. Denton, Y. Farzan and I.M. Shoemaker, *Activating the fourth neutrino of the 3+1 scheme*, *Phys. Rev. D* **99** (2019) 035003 [arXiv:1811.01310].

[241] J. Kostensalo, J. Suhonen, C. Giunti and P.C. Srivastava, *The gallium anomaly revisited*, *Phys. Lett. B* **795** (2019) 542 [arXiv:1906.10980].

[242] Daya Bay collaboration, F. An et al., *Evolution of the Reactor Antineutrino Flux and Spectrum at Daya Bay*, *Phys. Rev. Lett.* **118** (2017) 251801 [arXiv:1704.01082].

[243] Daya Bay collaboration, D. Adey et al., *Improved Measurement of the Reactor Antineutrino Flux at Daya Bay*, *Phys. Rev. D* **100** (2019) 052004 [arXiv:1808.10836].

[244] RENO collaboration, G. Bak et al., *Fuel-composition dependent reactor antineutrino yield at RENO*, *Phys. Rev. Lett.* **122** (2019) 232501 [arXiv:1806.00574].

[245] J. Conrad and M. Shaevitz, *Electron Antineutrino Disappearance at KamLAND and JUNO as Decisive Tests of the Short Baseline $\overline{\nu}_\mu \to \overline{\nu}_e$ Appearance Anomaly*, *Phys. Rev. D* **89** (2014) 057301 [arXiv:1310.3857].

[246] H. DANIEL, *Shapes of Beta-Ray Spectra*, *Rev. Mod. Phys.* **40** (1968) 659.

[247] H. Behrens and J. Jänecke, *Numerical Tables for Beta-Decay and Electron Capture*, Springer-Verlag Berlin Heidelberg (1969).

[248] M.M. Bé, V. Chisté, C. Dulieu, E. Browne, C. Baglin, V. Chechev et al., *Table of Radionuclides*, vol. 3 of *Monographie BIPM-5*, Bureau International des Poids et Mesures (2006), http://www.bipm.org/utils/common/pdf/monographieRI/Monographie_BIPM-5_Tables_Vol3.pdf.

[249] C.S. Wu and L. Feldman, *Beta-Spectrum of* Cl^{36}, *Phys. Rev.* **82** (1951) 457.

[250] L.M. Langer and R.J.D. Moffat, *The Twice-Forbidden Transition of* Cs^{137} *and the Law of Beta-Decay*, *Phys. Rev.* **82** (1951) 635.

Danksagung

So eine Doktorarbeit schreibt sich natürlich nicht von alleine und wäre ohne die Unterstützung vieler Personen nicht möglich gewesen. Daher möchte ich mich insbesondere bedanken bei

- Stefan Schönert und meinem Doktorvater Lothar Oberauer für die Aufnahme am Lehrstuhl und das Angebot zur Promotion. Neben der arbeitsbezogenen Unterstützung und den unterhaltsamen Geschichten aus dem Nähkästchen, habe ich sehr den herzlichen Umgang am Lehrstuhl genossen. Außerdem, habe ich sehr die Möglichkeit geschätzt auf viele Dienstreisen fahren zu können und dabei wertvolle Erfahrungen zu sammeln.
- meinem Mentor Matteo Agostini für die immer wieder nützlichen Hilfestellungen und Gedanken. Die enge Zusammenarbeit hat deutlich zum zweiten Teil dieser Arbeit beigetragen.
- unseren Sekretärinnen Paola und Sabine(s) für die Erleichterung aller bürokratischen Hürden und die große Auswahl an schönen Büroutensilien.
- den einzelnen Mitgliedern der Borexino/SOX Kollaboration für die familiäre Arbeitsatmosphäre und die stetige Unterstützung inklusive teilweiser auch notwendiger Motivation.
- allen aktuellen sowie auch ausgeschiedenen Arbeitskollegen des Lehrstuhls für die abwechslungsreichen, lustigen und guttuenden Mittags- und Kaffeepausen.
- den Korrektoren Julia, Simon, Konstantin, Viktoria, Thibaut, Christoph, Matteo und Lothar für die zahlreichen hilfreichen Anmerkungen und Verbesserungsvorschläge zu meiner Arbeit.
- meiner besten Freundin Barbara für das Zeichnen des liebevollen und witzigen Familienoszillations-Comics.
- meinen Eltern für die stetige freie Studien- und Berufswahl, meinem Freund fürs um mich Kümmern, Bekochen und das Aushalten meiner Launen (vor allem in den letzten Monaten) und meiner ganzen Familie fürs da sein.