UNIVERSITÀ DEGLI STUDI DI MILANO

Corso di Dottorato in Fisica, Astrofisica e Fisica Applicata

Dipartimento di Fisica

Ciclo XXXII

First indication of solar neutrinos from the CNO cycle reactions with the Borexino experiment

Supervisore: Dott. Barbara Caccianiga

Coordinatore: Professor Matteo Paris

Tesi di Dottorato di:
Davide Basilico

Anno Accademico 2018/19

External Referees:

Prof. Oliviero Cremonesi, Università degli Studi di Milano Bicocca

Dr. Aldo Serenelli, Institute of Space Sciences, Bellaterra, Barcelona

Commission of the final examination:

Prof. Ezio Previtali, Università degli Studi di Milano Bicocca

Prof. Francesco Villante, Università degli Studi dell'Aquila

Prof. Lino Miramonti, Università degli Studi di Milano

MIUR subjects (Settore Scientifico Disciplinare):

Settore FIS/01 - Fisica Sperimentale
Settore FIS/04 - Fisica Nucleare e Subnucleare

Final examination:

Date: 7 Feb 2020

Università degli Studi di Milano, Dipartimento di Fisica, Milano, Italy

Introduction

Motivation

The study of neutrinos is one of the most active sectors in modern astroparticle physics.

On the one hand, the discovery of neutrino oscillations, which was awarded the Nobel Prize in 2015, demostrated that neutrinos are unique tools to study particle physics. The long list of open questions regarding neutrinos (absolute masses, mass ordering, CP violation, neutrino Majorana or Dirac nature, existence of sterile neutrinos) is calling for answers which may have profound implications both on particle physics and cosmology.

At the same time, the neutrino elusiveness make them appealing from the astrophysical point of view. Thanks to their low cross-section, in fact, neutrinos can travel the cosmos almost undisturbed, carrying precious information on the astrophysical objects which have produced them. Ultra-high energy neutrino astronomy can provide insight in the long-standing problem of the origin and acceleration mechanism of high-energy cosmic rays, while supernova neutrinos can probe the theoretical predictions concerning the evolution and death of the stars.

Solar neutrinos play an irreplaceable role for the comprehension of the mechanisms powering our star. The study of solar neutrino fluxes has allowed to definitely prove that the Sun is powered by thermonuclear reactions occurring in its core, confirming the Standard Solar Model predictions.

Several questions concerning solar neutrinos are still unsolved. In particular, neutrinos emitted in the CNO cycle of reactions are the only undetected piece of the solar fusion puzzle. The CNO cycle is believed to be the main engine of very massive stars, while it is expected to contribute only $\sim 1\%$ to the solar luminosity. Observing neutrinos from the CNO cycle reactions would have therefore a striking importance in astrophysics, since it would provide the first experimental proof of the existence of this important source of energy in the core of the stars.

Moreover, since the CNO cycle is catalyzed by elements heavier than helium, the related neutrino flux is very sensitive to the metal abundance in the Sun. Recent determinations of the solar surface composition suggest that the solar metallicity might be lower than previously assumed. However, if lower abundances are given as input, the SSM shows a worse agreement with the helioseismic measurements. An accurate CNO neutrinos flux measurement would therefore be helpful in settling

this so-called metallicity problem.

This PhD thesis is mostly related to the determination of CNO solar neutrinos by means of the Borexino detector. Borexino is an ultrapure liquid scintillator-based detector located at the Laboratori Nazionali del Gran Sasso, Italy. Throughout a more than ten-years long data taking, it has achieved outstanding results about solar neutrino physics, measuring all the neutrino fluxes emitted from the pp chain nuclear reactions. The main difficulties of the CNO analysis analysis in Borexino are related to the very low signal/background ratio, and to the presence of a background whose energy spectral shape is highly correlated to the CNO neutrinos one. The main result, consisting in the the first direct indication of a signal from CNO neutrinos, has been achieved analyzing the latest three years of Borexino data taking.

Thesis overview

The present PhD thesis consists of eight Chapters. Chapter 1 and Chapter 2 are introductory to the neutrino physics and to the Borexino detector, respectively. Chapter 3 summarizes the strategy for the CNO neutrinos analysis followed throughout the thesis. Chapters 4-5-6-7 examine several topics of the CNO neutrinos analysis, spanning from the data selection sequence, to the multivariate fit techniques, to Monte Carlo simulations, to the ^{210}Bi and ^{210}Po backgrounds analysis.

Chapter 1: Neutrino physics. A basic introduction to the neutrino physics, emphasizing the processes involved in the generation of neutrinos from the Sun, the flavor oscillations, and the sterile neutrino hypotesis.

Chapter 2. The Borexino experiment: detector design and pp-chain results. A basic review of the Borexino detector: the detector design, the main background contaminants, the events reconstruction, the main physics achievements and next goals in the field of solar neutrino physics.

Chapter 3: Strategy for the search of CNO neutrinos with Borexino. The extensive sensitivity studies that will be reported in this Chapter indicate that Borexino has the chance to tackle CNO neutrinos provided that the interaction rate of ^{210}Bi isotope, the most annoying background, is determined independently by the multivariate fit used to disentangle signal and backgrounds. Later, the strategy adopted in this thesis to isolate a CNO neutrinos signal is outlined. The main techniques and ideas underlying the analysis are shortly summarized, while they will be separately and extensively described in the following chapters.

Chapter 4: Data selection and multivariate fit for the CNO neutrinos analysis. Basic data quality selection cuts, to disentangle as much as possible the neutrino signal from the radioactive background on an event-by-event basis, are discussed. Techniques to tag the ^{11}C cosmogenic background events, crucial in the context of CNO-ν determination, are reviewed. I have carefully analyzed the impact of these selection cuts on the new data, studying the selection stability as a function of time, and finding no evidences for anomalies or issues. The Phase-III-TMP dataset (July 2016 - June 2019), which will be analyzed in this thesis to search for the CNO-ν signal, will be defined, showing the related energy spectrum and stability plots. The multivariate fit of the energy spectra, radial distribution and e^+/e^- discriminator variable profile is outlined, along with the related methods and technicalities; the

two possible approaches to describe the detector energy response, on which the construction of the reference PDFs for the fit is based, are described in detail.

Chapter 5: The Monte Carlo simulation of Borexino detector for the CNO neutrinos analysis. The Monte Carlo simulation chain to generate physical events in Borexino is presented. The issue of the energy scale inconsistencies between Monte Carlo simulations and data, from 2016 on, is addressed. Together with the Borexino analysis team, I devoted a significant effort to solve this issue, which would have prevented us to perform the Monte Carlo fit on the Phase-III dataset, dedicated to CNO-ν analysis which started in July 2016. The problem was found to be related to the photomultiplier quantum efficiencies (QE) which are given as input to the simulations and are calculated from data on a weekly basis. After reviewing and improving the code which calculates the QE, I have performed a critical re-tuning of the Monte Carlo input parameters, based on a re-analysis on the calibration data; then, I have produced via simulations the reference energy shapes for signal and backgrounds to be given as inputs to the fit. This has solved the Monte Carlo problem and now it is possible to perform the Monte Carlo fit also on Phase-III data.

Chapter 6. Determination of the ^{210}Bi background rate for the CNO neutrinos analysis. A crucial ingredient for the CNO-ν determination consists in a accurate determination of ^{210}Bi background rate. The main strategy to determine the ^{210}Bi rate, based on the connection with its decay daughter ^{210}Po found in secular equilibrium, is reviewed. I will show the analysis of the ^{210}Po spatial and time distributions, as well as the related impact of temperature evolution and detector thermal stabilization. Then, the proposed ideas to determine the supported ^{210}Po rate (*radial analysis*, of which I have been the main contributor, and *clean bubble analysis*) are discussed in detail. Then, I show the analysis of the ^{210}Bi time stability and homogeneity in the scintillator, along with the study of the related systematic uncertainties to be associated to the ^{210}Bi independent constraint.

Chapter 7. First indication of a solar neutrinos signal from the CNO cycle reactions. The elements analyzed in the previous three Chapters (data selection and multivariate fit, ^{210}Bi -^{210}Po studies, Monte Carlo simulations) are combined together to achieve the first evidence of a CNO-ν signal. I present the multivariate analysis of the Borexino Phase-III-TMP dataset via the two analytical and Monte Carlo based approaches: the absence of CNO-ν signal is preliminarly excluded at a 3σ and 4σ confidence level respectively. Furthermore, I have exploited this measurement to infer additional information on solar physics, in terms of the solar metallicity discrimination, which has been slightly improved with respect to the previous Borexino results, and in terms of the first direct assessment of the C+N abundance in the solar core. Results need to be considered as preliminary, since Borexino is still taking data and the Phase-III dataset is not closed yet. The official release of the final results is expected to take place in a few months.

Separated from the rest of the thesis, Chapter 8 is fully dedicated to the SOX project, which has been a proposed continuation of Borexino experiment beyond the end of solar neutrino campaigns. The main physics goal of the project was the search for the light sterile neutrinos. Throughout my first year of PhD activity, I have been heavily involved in the sensitivity studies topic for the SOX project.

Chapter 8. SOX: Short distance neutrino Oscillations with BoreXino. SOX has

been one of the experiments planned to probe the sterile neutrino anomalies, and possibly test conclusively the sterile neutrino(s) hypothesis; the key idea relied on placing an intense artificial $\overline{\nu}_e$ source close to the Borexino detector. Unfortunately, the SOX project has been interrupted before its beginning, due to the technical issues to produce the $\overline{\nu}_e$ source with the required activity and purity levels. My work has been strongly focused on the investigation and impact of the main systematic factors affecting the SOX sensitivity to sterile neutrinos, both source-related and detector-related. In a nutshell, the studies shows how SOX could have been sensitive to the sterile hypotesis at the design expected levels, covering most of the allowed parameters region by the light sterile anomalies; the expertises and the analysis techniques which I adopted can be extended to a future SOX-like experiment, i.e. a large mass and liquid-based scintillator detector with low radioactive contamination aiming to investigate the light sterile neutrino hypothesis.

Contents

List of Figures

List of Tables

Chapter 1

Neutrino Physics

In this Chapter I will give a brief introduction to neutrino physics. Since my thesis is mostly devoted to the study of solar neutrinos with the Borexino experiment, I will focus in particular on the description of the Standard Solar Model and its predictions regarding the flux of neutrinos from the pp chain and the CNO cycle. After giving an overview of the current experimental status of solar neutrino studies, the open issues and experimental perspectives about solar neutrinos will be discussed, mainly connected to the detection of the still missing CNO neutrinos. I will also discuss the mechanism of neutrino oscillations, which have been discovered also thanks to solar neutrinos and can now be described in a rather complete theoretical framework. The well-estabilished three-flavors scenario is being challenged by persisting experimental anomalies which could be solved by the introduction of one or more families of sterile neutrinos. I will discuss the sterile neutrino issue and possible experimental ways to solve it; more on this subject will be elaborated in the last Chapter 8, where I will present the short-baseline SOX project and my studies on its sensitivity to sterile neutrinos.

This Chapter is organized as follows. After an historical introduction (Sec. 1.1), neutrinos are described in the frame of the Standard Model of particle physics (Sec. 1.2). Natural and artificial sources for neutrinos are listed and reviewed in Sec. 1.3. Focusing to the solar physics field, the Standard Solar Model and the associated predictions for solar neutrinos are introduced respectively in Sec. 1.4 and Sec. 1.5. The so-called *solar neutrino problem*, which is the inconsistency between measured and predicted solar neutrino fluxes, is presented in Sec. 1.6; the flavor oscillations phenomenon, confirmed as the solution of this problem, is then illustrated (Sec. 1.7). An introduction to the experimental way to constrain or determine the neutrino absolute mass scale will be outlined in Sec. 1.8. Moving to sterile neutrino physics (Sec. 1.9), the long standing experimental anomalies in the $\sim 1\,\mathrm{eV}$ mass region (Sec. 1.9.1) and the idea of a fourth flavor state as possible solution (Sec. 1.9.2) are addressed; future perspectives are outlined in Sec. 1.9.3.

1.1 Introduction

The existence of neutrinos has been proposed for the first time by Pauli in 1930 [1], trying to get an explanation for the continuous spectrum of electrons emitted in the β decay of several radioactive isotopes. Being the β decay supposed to be a two-body process, this observation was

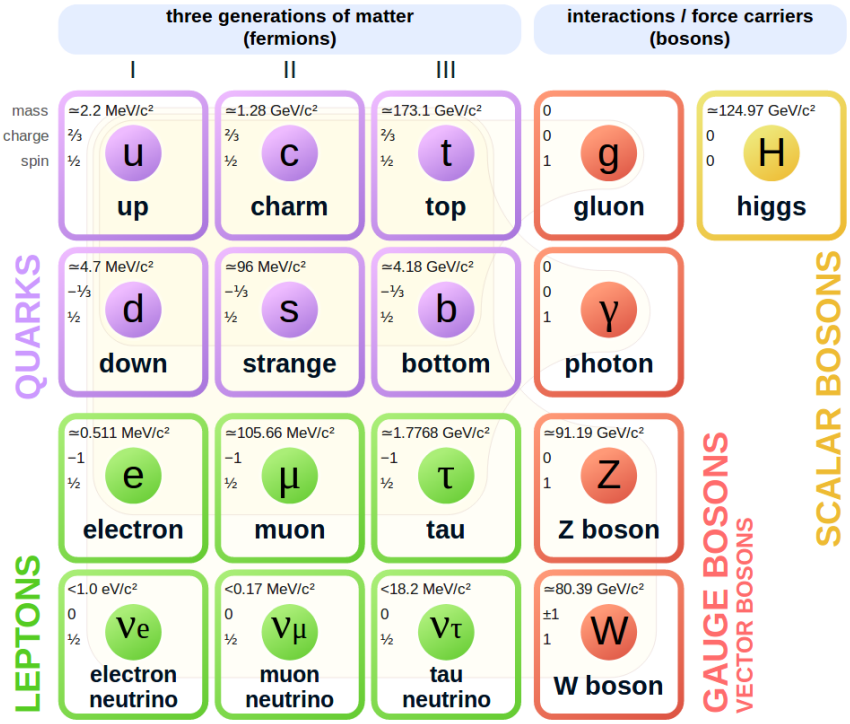


Figure 1.1: Schematic depiction of the **Standard Model** particles with related mass, electric charge and spin values.

not compatible with the energy conservation principle, which would have suggested instead a monoenergetic spectrum. In addition, assuming a two-body β decay, experimental results clashed with the the momentum and spin conservation laws. Pauli supposed the existence of an undetected neutral particle emitted in the β decay, along with the electron, allowing to recover the three conservation laws. In the following years, trying to describe the nuclear decay rates of β decays, Fermi proposed the first version of the weak interaction framework: it was based on a pointlike interaction of four fermions through two vector currents. Among the fermions involved, also the previously hypotized neutrino was present.

Nearly thirty years later, in 1956, Cowan and Reines observed the anti-neutrinos using a large water-based detector placed near a nuclear power plant [2]. They exploited a huge flux of electron antineutrinos emitted from the nuclear reactor, to induce the inverse β reaction $\overline{\nu}_e + p \rightarrow n + e^+$. In 1995, Reines was awarded for these results with the Nobel Prize *for pioneering experimental contributions to lepton physics* [3]. From the first detection, neutrinos have been observed from a wide range of natural and artificial sources: from the atmosphere, due to the cosmic rays interactions (atmospheric neutrinos), from the supernova SN-1987A (supernovae neutrinos), from the nuclear reactors, from particle accelerators, and above all from the nuclear reactions occuring in the Sun (solar neutrinos).

1.2 Neutrinos in the Standard Model of particle physics

The Standard Model of particle physics (SM) is a gauge theory which successfully describes in a common framework the strong, the weak and the electromagnetic fundamental interactions [4]. The theoretical predictions stemming from this model have been successfully confirmed by experiments throughout the years, making it one of the greatest achievements of the XX century physics. After the discovery of the Higgs boson at LHC, the Standard Model was established as a full, correct theory of fundamental physics in the electroweak energy scale. The Standard Model particles are firstly categorized according to their spin: fermions, spin–1/2 particles that are the fundamental blocks of matter, separated in three flavor families of leptons and quarks, and the bosons, spin–1 particles which are the mediators of the interaction fields. The six quarks (up, down, charm, strange, top, bottom), carrying a color charge, are able to interact also through the strong force. The dynamics of the six leptons instead is completely governed by the electroweak sector. Among the bosons, the strong interactions are mediated by eight massless gluons, while the three massive bosons $W^{\pm}$ and Z mediate the weak interaction. One massless photon is the electromagnetic force carrier. The Higgs boson plays a unique role in the Standard Model, by explaining why the other elementary particles, except the photon and gluon, are massive. Without the Higgs mechanism, all bosons would be massless. In the Standard Model fermions, such as leptons and quarks, acquire mass as a result of their interaction with the Higgs field.

Neutrinos are electrically neutral leptons with the lowest mass among the Standard Model's particles. Being leptons, they do not interact through the strong force. In the SM, neutrinos can be found in three orthogonal flavors, which are the eigenstates of the weak interaction, forming three doublets with the three charged leptons: the electron neutrino ν_e, the muon neutrino ν_μ, the tau neutrino ν_τ, along with their antiparticles (the three antineutrinos $\overline{\nu}_e$, $\overline{\nu}_\mu$, $\overline{\nu}_\tau$). Neutrinos can change flavor when they propagate, giving rise to the flavor oscillation phenomenon (Sec. 1.7). This represents the first evidence of physics beyond the Standard Model.

The concept of neutrino flavor is dynamical, defined through the reactions: $\nu_e/\nu_\mu/\nu_\tau$ is the neutrino produced along with $e^+/\mu^+/\tau^+$, or which produces an $e^-/\mu^-/\tau^-$ in a charged current weak interaction processes. A similar definition can be stated for the three anti-neutrinos $\overline{\nu}_e$, $\overline{\nu}_\mu$, $\overline{\nu}_\tau$. The experimental evidence studying neutrinos show that they are always produced in V-A weak interaction processes: this implies that they are found dominantly in a left-handed state.

According to the minimal formulation of the SM, the three neutrinos are considered as massless, and therefore as pure left-handed states [4]. Currently, there are no significant hints on the existence of right-handed neutrinos (sterile neutrinos), which would interact with the other SM particles only through gravitational force. Their possible interaction would be many orders of magnitude fainter than the ordinary left-handed neutrino weak interaction. Such particle cannot be included in the SM, because a right-handed component needs to be described as a singlet of the $SU(2)_L$ symmetry group, which cannot be coupled to the weak bosons. Right-handed neutrinos, if existing, could anyway assume a crucial role to understand many pending topics related to the physics beyond the SM, like the neutrino mass generation, the huge mass scale difference between the neutrino and the other SM particles, the matter-antimatter symmetry in the Universe. The hypothesis of the existence of light sterile neutrinos in the eV mass scale will be addressed in Sec. 1.9.

A summary of the Standard Model particles is depicted in Fig. 1.1 along with their mass, their electric charge and their spin values. In spite of the theoretical self-consistency and the experimental success, the SM leaves some phenomena unexplained, and it falls short of being a complete theory

of fundamental interactions [4]. For example, any attempt to describe the gravity in the Standard Model has revealed unsuccessful so far: the full theory of general relativity is not included. Dark matter candidates are not predicted. The fermion and boson masses, as well as the interaction couplings values, are free parameters of the model. For what concerns the neutrino physics field, some of the open questions are the neutrino mass ordering, which is still unknown, the Dirac or Majorana nature of neutrinos, their absolute mass scale determination.

1.2.1 Neutrino mass generation

Neutrinos are introduced in the Standard Model as truly massless fermions, due to their pure left-handed nature. In Sec. 1.7 we will describe how the existence of neutrino oscillations provides a clear evidence for the non-zero neutrino masses, giving the first sign of fundamental physics beyond the SM: thus, a mechanism for giving masses to neutrinos is needed.

The most general neutrino mass terms can be of two different types: Majorana or Dirac. At present, the question of the Dirac or Majorana nature for neutrino masses is still theoretically and experimentally open. In the case of Majorana neutrinos, the neutrino mass term is a linear combination of the product of left-handed and right-handed (charge-conjugate) components of neutrino fields:

$$\mathcal{L}_M = -\frac{1}{2} M_M \left(\overline{\nu}_L^c \nu_L^c + \text{h.c.}\right) \tag{1.1}$$

The Majorana mass term $\mathcal{L}_M$ mixes neutrinos and antineutrinos; it breaks the lepton number conservation by two units and Majorana neutrinos obey the so-called Majorana condition:

$$\nu_M = \nu_M^c \tag{1.2}$$

where c apex represent the transformed field under the charge inversion operator[1]. Besides a possible Majorana phase, this condition implies that one field describes both neutrino and anti-neutrino states, which are not distinguishable. Therefore, a Majorana neutrino is described as a system with two degrees of freedom, while Dirac particles are represented by four degrees of freedom. If neutrinos are indeed Majorana particles, neutrinoless double beta decay, as well as other lepton number violating phenomena, would be allowed.

In the case of Dirac neutrinos, the neutrino mass term is generated by the same standard Higgs mechanism that produces quark and charged lepton masses. One introduces the right handed fields, singlets under the SM gauge group: being not allowed to interact through weak interactions, a direct detection would be impossible. The Dirac mass term would be:

$$\mathcal{L}_D = -M_D \left(\overline{\nu}_R^c \nu_L^c + \text{h.c.}\right) \tag{1.3}$$

The total lepton number L is conserved, but the separate lepton flavor symmetries are broken. As a consequence, neutrino with definite mass are Dirac particles described by four-component spinors and Dirac neutrinos and anti-neutrinos have opposite L. It is important to note that the Dirac/Majorana distinction does not affect the flavor oscillation properties that will be described in Sec. 1.7, which are mostly independent from to the neutrinos mass generation mechanism.

[1] Quark and charged-lepton Majorana masses are forbidden by electric charge conservation.

1.3 Neutrino sources

Many different physical process in the universe, natural or anthropic, can yield neutrino production. In general, neutrinos can be classified according to their sources. The measured and expected (anti-)neutrino fluxes from currently known sources are shown in Fig. 1.2 as a function of their energy. For what concern their detection, the energy range from keV to several GeV is the domain of underground detectors, while the one from tens of GeV to about 100 PeV, with its much smaller fluxes, is addressed by Čerenkov light detectors underwater and in ice. The highest energies are only accessible with huge detector volumes, because of the very faint incoming flux.

Neutrinos and antineutrinos can be classified according to their production source as follows [5]:

- Cosmological neutrinos (relic neutrinos);
- Solar neutrinos;
- Reactor neutrinos;
- Radioactive source neutrinos;
- Accelerator neutrinos;
- Terrestrial neutrinos (geo-neutrinos);
- Atmospheric neutrinos;
- Supernovae neutrinos.
- Cosmogenic neutrinos (ultra-high energy neutrinos).

The lowest energy region (μeV and meV, see Fig. 1.2) is populated by the cosmological neutrinos, i.e. the 1.9 K neutrino counterpart (CνB) to the 2.7 K cosmic microwave background. This flux is expected to amount $\sim 10^{20}\,\mathrm{cm^{-2}\,s^{-1}}$, but with a cross section extremely low, $\sigma \sim (10^{-56} - 10^{-64})\,\mathrm{cm^{-2}}$ [6]. The direct detection of relic neutrinos is an extremely difficult task with present experimental techniques, and it was proposed in a paper by Steven Weinberg in 1962 [7]. At the current status, PTOLEMY [8] is the only project whose primary design goal is the CνB direct detection. The experimental technique is based on the process of neutrino capture on tritium $\nu_e + {}^3\mathrm{H} \rightarrow {}^3\mathrm{He} + e^-$: the smoking gun signature for CνB capture is the observation of a peak in the electron spectrum at $2m_\nu$ energy above the β-decay endpoint.

The keV-MeV range is populated by neutrinos from the Sun, from supernovae, from nuclear reactors, from artificial radioactive sources and from the interior of the Earth. Nuclear reactors are the major source of human-generated neutrinos: electron anti-neutrinos are emitted in the β^- decay of neutron-rich daughter fragments in the fission process from the ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu [9, 10]. On average, each nuclear fission releases about 200 MeV of energy, of which roughly 4.5% is radiated away as antineutrinos. The detectable antineutrinos from fission reactions show an energy peak between about 3.5 and 4 MeV, with a maximal energy of about 10 MeV. Nuclear power plants have been exploited by a several experiments aiming to study the neutrino oscillations in a wide source-to-detector distance range, or to investigate the existence of sterile neutrinos.

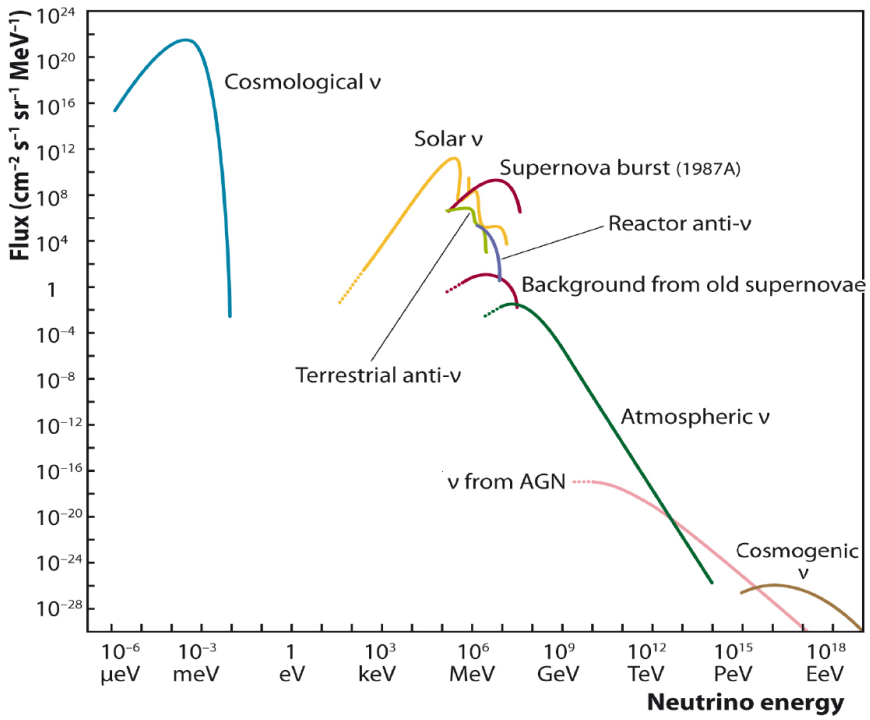


Figure 1.2: Measured or expected **(anti-)neutrino fluxes** from currently known sources as a function of their energy, in log-log scale [5].

The Sun is by far the most important neutrino source for current experiments, providing a very high neutrino flux at the Earth surface, estimated to be $\sim 6 \cdot 10^{10}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This topic will be addressed in detail in Sec. 1.5.

Intense radioactive sources can be used to study the neutrino oscillations. They can be produced from spent nuclear fuel in reactors, or via neutron irradiation. Due to the very small size, compared to the typical detector dimensions, they can be considered as pointlike sources.

Accelerators can be employed to make artificial high-statistics neutrino beams. The most common technique is to collide protons onto a fixed target, producing charged pions or kaons. These mesons are then magnetically focused into a tunnel where they decay in flight, emitting neutrinos or antineutrinos. Other decay products are then absorbed by a solid beam dump.

Geo-neutrinos [11] are $\overline{\nu}_e$ in the MeV energy range produced by β^- decays of long-lived isotopes, which are naturally present in the interior of the Earth, such as decays in the ^{238}U and ^{232}Th chains, and ^{40}K (which also produces neutrinos). Measuring the total geo-neutrino flux [12, 13, 14] through the inverse β reaction makes it possible to determine the contribution of radiogenic heat released in radioactive decays to the total terrestrial surface heat flux, currently estimated to be about 46 TW.

Atmospheric neutrinos are produced in hadronic showers resulting from collisions of cosmic rays with nuclei in the upper Earth atmosphere. These showers contain unstable charged mesons which decay before the contact with the earth surface, leading to the production of electronic or muonic

neutrinos and antineutrinos, according to the processes:

$$\pi^+ \to \mu^+ + \nu_\mu \qquad \mu^+ \to e^+ + \nu_e + \overline{\nu}_\mu \tag{1.4}$$

$$\pi^- \to \mu^- + \overline{\nu}_\mu \qquad \mu^- \to e^- + \overline{\nu}_e + \nu_\mu \tag{1.5}$$

resulting in an expected population ratio of $N(\nu_\mu + \overline{\nu}_\mu)/N(\nu_e + \overline{\nu}_e) \approx 2$. The energy range of atmospheric neutrinos lies between 1 GeV and 100 GeV.

Neutrinos are also massively produced by Type Ib, Ic and Type II Supernovae explosions. During these events, the density at the star core increases to extreme values ($\sim 10^{17}\,\mathrm{kg/m^3}$), making it possible the electronic capture on protons, thus favouring the production of neutrons and electronic neutrinos. A prominent feature of the collapse is that $\approx 99\%$ of the gravitational binding energy of the resulting remnant is converted to a burst of neutrinos of all flavors, with energies of a few tens of MeV and over a timescale of a few tens of seconds. A second and more important neutrino source is the thermal energy ($10^{51-53}\,\mathrm{erg}$) of the newly formed neutron core, which is dissipated via the formation of neutrino-antineutrino pairs of all flavors. Supernova neutrinos from the supernova 1987A in the Large Magellanic Cloud have been recorded in 1987 by three detectors [15, 16, 17]: Kamiokande in Japan, IMB in the USA (both water Čerenkov detectors) and the Baksan scintillation detector in Russia.

The highest energies (starting from the TeV range) are domain of cosmogenic neutrinos, produced in or near extreme acceleration sites outside our galaxy. As for the atmospheric neutrinos, they are expected to be produced when charged particles interact with matter and light, resulting in charged mesons that decay into neutrinos and other particles. The possible sources candidates are the supernova remnants, the Gamma Ray Bursts or Active Galactic Nuclei [18]. The observations of high-energy neutrinos can provide insight into the long-standing problem of the origins and acceleration mechanisms of high-energy cosmic rays. High energy neutrinos can be detected by huge Čerenkov light detectors [19]. IceCube, a large volume detector located in Antarctica ice, detected for the first time neutrino events in the PeV energy region [20]. In 2017, IceCube detected a high-energy neutrino, IceCube-170922A, with an ~ 290 TeV energy, and with its arrival time and direction consistent with a γ-ray flare from the blazar TXS 0506+056[2] [21, 22], which was observed to be in a flaring state. The extensive photon multi-wavelength campaign which followed, ranging from radio frequencies to γ-rays, confirmed the spatial and temporal coincidence, suggesting that this blazars is a source of high-energy neutrinos at the 3.0σ significance level.

1.4 The Standard Solar Model

In the following I will shortly summarize the main concepts behind the formulation of the Standard Solar Model, necessary for a basic comprehension of the solar neutrinos production. For more details see Ref. [23, 24, 25, 26, 27, 28]. The model is based on a set of equations which describe the stellar evolution, which are shortly introduced hereafter:

- *Mass conservation:* in its expected main-sequence lifetime, expected to last $\sim 10^{10}$ y, the Sun will lose an irrelevant $\sim 0.01\%$ of its initial mass.

[2] Blazars are a class of active galactic nuclei with powerful relativistic jets pointed close to our line of sight.

- *Energy conservation:* energy is produced by nuclear fusion reactions, which take place in the core of the Sun. The related energy conservation accounts for both this nuclear-emitted energy, and for the energy contribution used for the star expansion or contraction.

- *Hydrostatic equilibrium:* a star burns in hydrostatic equilibrium, i.e., mantaining a local balance between the gravitational force and the pressure gradient resulting from the thermonuclear reactions.

- *Energy transport:* energy is transported by radiation in the innermost core of a star and by convection in the outernmost regions. In principle, it allows to determine the temperature at any point of the star.

An additional equation describes the time evolution of chemical composition: it depends on the nuclear reactions rates (which are an input of the model) and on the diffusion of helium and heavy elements. In addition, also an equation of state is needed, which specifies the relation between density, pressure, temperature and chemical composition. A proper equation of state must include the effects of ionization, radiation pressure, pressure ionization, degeneracy, and so on; it is usually provided by means of tabular form as functions of T , P (or ρ) and chemical composition [29]. Summarizing, the Standard Solar Model is obtained solving these equations in time imposing the boundary conditions measured experimentally for our Sun [25].

The metallicity of the Sun, i.e. the fraction of elements heavier than He, is one of the most crucial parameters given in input to the model; it is assumed to be the same as the current surface composition, which is measured through photosphere spectroscopy. Historically, the solar abundance values from Grevesse and Sauval (GS98, [30]) were used, adopting a ratio between heavy-element and hydrogen abundance $Z/X = 0.023$ (*high metallicity* or *HZ* scenario). Starting from 2001, several works (Ref. [31, 32, 33]) have reviewed the spectroscopic determinations of the solar photospheric composition, mainly thanks to the development of three dimensional hydrodynamic models of the solar atmosphere and to new analyses of solar emission spectral lines. The main result of this review is a significant lowering of the metal abundance (AGSS09, [31]) to $Z/X = 0.0165$, that is about 28% lower than the previous value of GS98 (*low metallicity* or *LZ* scenario).

Unfortunately, the SSM which makes use of the low metallicity abundances produces results which are in disagreement with helioseismology. Explanations have been put forward in terms of modifications of the physical processes involved in the energy transport (e.g. energy transport by diffusion [34, 28], solar accretion [26]) or of improvements related to the microphysics inputs of SSM (e.g. radiative opacities [25], nuclear reaction rates). Unfortunately, none of these possibilities leads to a complete solution of this so-called *solar abundance problem*.

A help to settle this tension could come from solar neutrinos: indeed, the predicted solar neutrino fluxes depend on the metallicity input, especially for neutrinos of the CNO cycle reactions (see Tab. 1.1 and Sec. 1.5.2). A direct and precise measurement of the CNO-ν flux could help to shed light on this issue; moreover, it would represent the first evidence of this energy production mechanism, which looks crucial for massive stars (see Sec. 1.5.2). The solar physics implications of a CNO-ν flux measurement will be discussed in Sec. 7.4.

1.5 Solar neutrinos

According to the SSM described in the previous Section, the Sun shines thanks to two main branches of thermonuclear reactions [35]. The proton-proton *pp chain* dominates the energy production in the Sun, contributing to $\approx 99\%$ of its luminosity. The still undetected *CNO cycle*, which employs Carbon, Nitrogen and Oxygen nuclei as catalysts, is supposed to be secondary for the Sun, providing $\approx 1\%$ of luminosity, while is expected to be dominant for massive stars $M \gtrsim 1.5 M_\odot$. The scheme of processes occurring in the *pp* chain and in the CNO cycle are reported respectively in the left and right panel of Fig. 1.3. The net process for both the two sequences is the conversion of four protons into a ^{4}He atom:

$$4p \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e + Q \tag{1.6}$$

emitting two electron neutrinos. An energy Q is released in this reaction[3]:

$$Q = 4m_p + 2m_e - m_{^4\mathrm{He}} = B(4,2) + 2m_e - 2(m_n - m_p) = 26.731\,\mathrm{MeV} \tag{1.7}$$

where $B(4,2)$ is the nuclear binding energy of ^{4}He. This energy is released as photon or neutrino kinetic energy. Instead, the e^+ emitted in the reaction annihilate with free electrons in the solar plasma, producing photons which travel and interact in the solar matter. Electron neutrinos are produced in eight of the reactions depicted in Fig. 1.3 (five for pp chain and three for CNO cycle), which will be described in more detail in Sec. 1.5.1 and Sec. 1.5.2. Considering both the pp chain and the CNO cycle processes, the overall ν_e energy spectrum lies in the $0\,\mathrm{MeV} < E_\nu < 18.8\,\mathrm{MeV}$ range, depending on the involved nuclear reactions. The resulting neutrino fluxes predicted by the B16 SSM[4], as a function of the ν_e energy, are depicted in Fig. 1.4.

The extremely low neutrino cross section makes solar neutrinos an ideal probe of the Sun interior [37]. Considering an average core solar density $\rho \approx (100 - 150)\,\mathrm{g/cm^3}$, which leads to an electron numerical density $n_e \sim 10^{26}\,\mathrm{cm^{-3}}$, and an average cross section of about $\langle\sigma\rangle \sim 10^{-43}\,\mathrm{cm^2}$ at 1 MeV energy, the ν_e free mean path λ_ν in the solar matter can be calculated as

$$\lambda_\nu = \frac{1}{n_e \langle\sigma\rangle} \sim \frac{1}{10^{26}\,\mathrm{cm^{-3}} \cdot 10^{-43}\,\mathrm{cm^2}} \sim 10^{17}\,\mathrm{cm} \tag{1.8}$$

which is more than six order of magnitude larger than $R_\odot \sim 6.9 \cdot 10^{10}\,\mathrm{cm}$. Thus, neutrinos are able to travel the solar matter almost undisturbed at almost the speed of light, reaching the Earth surface approximately in eight minutes. They act as perfect tracers of the Sun core, where they have been generated. Photons mean free path instead is much shorter, $\lambda_\gamma \sim 1\,\mathrm{cm}$, because of the higher cross section for electromagnetic processes $\sigma_\gamma \approx \sigma_{\mathrm{Thomson}} \sim 10^{-25}\,\mathrm{cm^2}$. Indeed, as hinted previously, they continuously interact through Thomson scattering (or inverse Bremsstrahlung), losing the initial energy and direction information. Because of these consecutive processes, photons wander in the Sun performing a random walk which reduces considerably their net speed: by means of brownian motion estimations, a photon emitted in the core of the Sun takes $\sim 10^5$ years to emerge from outside the photosphere. These considerations promote the neutrinos as the only real-time messengers from the Sun core.

[3] The Q-value $Q = M_i - M_f$ is defined as the difference between the initial state total mass M_i and final state total mass M_f, and it is equal, in natural units, to the total kinetic energy available to the final state particles.

[4] In the high metallicity scenario - see Sec. 1.5.2.

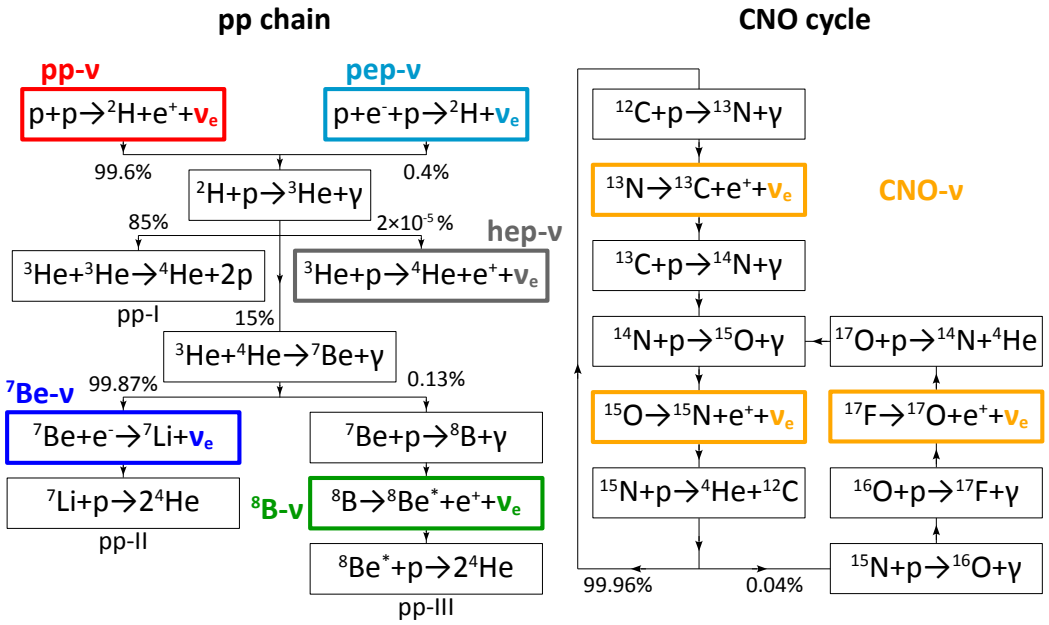


Figure 1.3: Scheme of the **sequence of processes** for the *pp* chain and CNO cycle according to the SSM. The traditional names of the neutrino-producing reactions and the corresponding neutrino names (pp-ν, pep-ν, hep-ν, ^{7}Be-ν, ^{8}B-ν, CNO-ν) are bolded and colored.

1.5.1 *pp* chain

Most stars, as the Sun itself, derive their luminosity from the conversion of hydrogen to helium. Approximately 99% of the Sun power is generated by means of the nuclear reactions present in the *pp* chain. Its first step consists in the fusion of two protons, producing a ^{2}He atom (see Fig. 1.3):

$$\textbf{pp)} \qquad p + p \rightarrow {}^2\mathrm{He} + e^+ + \nu_e \qquad Q = 0.42\,\mathrm{MeV} \qquad \text{B.R.} \approx 99.76\% \tag{1.9}$$

This pp reaction proceeds slowly, with a $\sim 10^{10}$ y timescale. Being the starting reaction of the chain, it keeps the Sun in the hydrogen burning stage of life for billion of years, allowing it to hold its dynamic equilibrium in the main sequence of the HR diagram.

Being the less model dependent among the solar reactions, the pp-ν determination is particularly relevant to test the hypothesis of hydrogen fusion as powering the Sun. This is also the most common reaction to give rise to the fusion processes, with the highest neutrino flux $\Phi(\text{pp-}\nu) \approx 6{\cdot}10^{10}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Neutrinos emitted through this reaction (pp-ν) show a continuous spectrum with very low Q-value $Q = 0.42\,\mathrm{MeV}$.

There is a secondary alternative way to start the *pp* chain, named the *pep* reaction

$$\textbf{pep)} \qquad p + e^- + p \rightarrow {}^2\mathrm{He} + \nu_e \qquad Q = 1.44\,\mathrm{MeV} \qquad \text{B.R.} \approx 0.24\% \tag{1.10}$$

The resulting neutrinos (pep-ν) are mono-energetic with 1.44 MeV energy. This reaction occurs with lower B.R. with respect to the *pp* one Eq. (1.9) [38], because of the reduced phase space

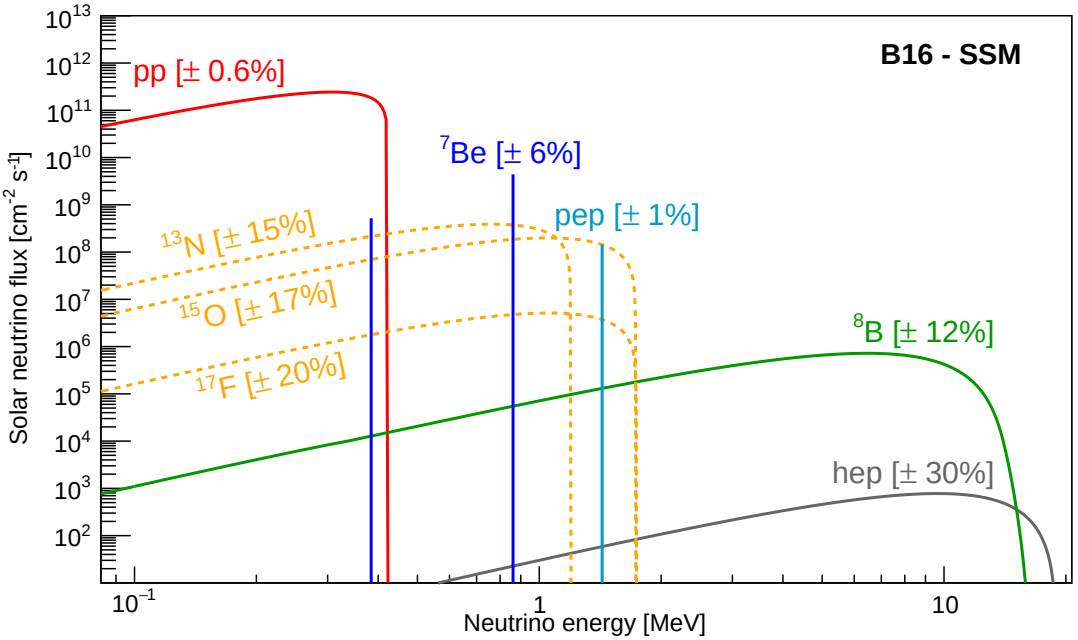


Figure 1.4: Fluxes of the solar neutrinos from the pp chain and the CNO cycle fusion sequences as a function of the neutrino energy. The solar neutrino energy spectrum is obtained from Ref. [36], using the updated fluxes taken from Ref. [25]. The flux (vertical scale) is given in $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{MeV}^{-1}$ units for continuum sources (pp-ν, ^{8}B-ν, CNO-ν, hep-ν) and in $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ units for monoenergetic sources (^{7}Be-ν, pep-ν).

available for the states in final configuration [38] [5]. The produced ^{2}He captures a proton and produces ^{3}He:

$$^{2}\mathrm{He} + p \rightarrow {}^{3}\mathrm{He} + \gamma \qquad Q = 5.49\,\mathrm{MeV} \tag{1.11}$$

The chain is now splitted into three terminations (pp-I, pp-II, pp-III) according to the reaction partner of the just produced ^{3}He. The pp-I reaction is the most frequent among the three alternatives, with B.R. $\approx 84.92\%$:

$$\textbf{pp-I)} \qquad {}^{3}\mathrm{He} + {}^{3}\mathrm{He} \rightarrow {}^{4}\mathrm{He} + 2p \qquad Q = 12.86\,\mathrm{MeV} \qquad \text{B.R.} \approx 84.92\% \tag{1.12}$$

The pp-I branch terminates without further neutrino emission.

With extremely low probability (B.R. $\sim 10^{-5}$), the ^{3}He undergoes a proton capture reaction, the *hep* neutrinos (hep-ν) are produced:

$$\textbf{hep)} \qquad {}^{3}\mathrm{He} + p \rightarrow {}^{4}\mathrm{He} + e^{+} + \nu_{e} \qquad Q < 18.8\,\mathrm{MeV} \qquad \text{B.R.} \approx 10^{-5} \tag{1.13}$$

[5] In particular, the relative probability of reactions Eq. (1.9) and Eq. (1.10) is approximately equal to the ratio of the final-state phase spaces.

Because of the very low production rate, the associated flux has not been detected yet, and only upper limits exist [39, 40].

The last way to continue the chain is given by the production of ^{7}Be through the strong interaction between a ^{3}He and an α particle:

$$^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma \qquad \text{B.R.} \approx 15.08\% \tag{1.14}$$

The ^{7}Be production, described by Eq. 1.14, leads to two possibilities: the pp-II and pp-III branches. The former, and also the most frequent one, produces monoenergetic neutrinos (^{7}Be-ν) through the electron capture on ^{7}Be:

$$\begin{aligned} &\textbf{pp-II)} \qquad {}^7\mathrm{Be} + e^- \rightarrow {}^7\mathrm{Li} + \nu_e \qquad Q = \begin{cases} 0.862\,\mathrm{MeV} & 90\% \\ 0.384\,\mathrm{MeV} & 10\% \end{cases} \\ &\qquad\qquad\quad {}^7\mathrm{Li} + p \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He} \end{aligned} \tag{1.15}$$

The ^{7}Li is produced in its ground state or in an excited state with 487 keV energy: in the former case, the related neutrino energy is 0.862 MeV, otherwise it is 0.384 MeV.

The *pp*-III branch begins with the proton capture on ^{7}Be, producing a neutrino continuous spectrum up to 16 MeV (^{8}B-ν) through the following ^{8}B decay:

$$\begin{aligned} &\textbf{pp-III)} \qquad {}^7\mathrm{Be} + p \rightarrow {}^8\mathrm{B} + \gamma \\ &\qquad\qquad\quad {}^8\mathrm{B} \rightarrow {}^8\mathrm{Be} + e^+ + \nu_e \qquad Q < 16\,\mathrm{MeV} \\ &\qquad\qquad\quad {}^8\mathrm{Be} \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He} \qquad \text{or} \qquad {}^8\mathrm{Be} + {}^4\mathrm{He} \rightarrow {}^{12}\mathrm{C} \end{aligned} \tag{1.16}$$

1.5.2 CNO cycle

Metals present in the Sun can act as catalysts for hydrogen burning, giving rise to sequences of nuclear fusions alternative to the pp chain ones [23]. The CNO cycle reactions make use of ^{12}C or ^{15}N nuclei as catalysts to convert four protons in ^{4}He: no net production of these elements takes place. This leads to the following net cycle balances in agreement with Eq. 1.6:

$$^{12}\mathrm{C} + 4p \rightarrow {}^{12}\mathrm{C} + {}^4\mathrm{He} + 2e^+ + 2\nu_e + Q \tag{1.17}$$

$$^{15}\mathrm{N} + 4p \rightarrow {}^{15}\mathrm{N} + {}^4\mathrm{He} + 2e^+ + 2\nu_e + Q \tag{1.18}$$

The CNO cycle is composed by two sub-cycles (see Fig. 1.3): the dominant *CNO-I* cycle and the secondary *CNO-II* cycle. The former, named also CN cycle, consists in a loop of reactions starting with the proton capture on ^{12}C nucleus:

$$^{12}\mathrm{C} + p \rightarrow {}^{15}\mathrm{N} + \gamma \qquad Q = 1.95\,\mathrm{MeV} \tag{1.19}$$
$$^{13}\mathrm{N} \rightarrow {}^{13}\mathrm{C} + e^{+} + \nu_e \qquad Q = 1.2\,\mathrm{MeV} \tag{1.20}$$
$$^{13}\mathrm{C} + p \rightarrow {}^{14}\mathrm{N} + \gamma \qquad Q = 7.54\,\mathrm{MeV} \tag{1.21}$$
$$^{14}\mathrm{N} + p \rightarrow {}^{15}\mathrm{O} + \gamma \qquad Q = 7.35\,\mathrm{MeV} \tag{1.22}$$
$$^{15}\mathrm{O} \rightarrow {}^{15}\mathrm{N} + e^{+} + \nu_e \qquad Q = 1.73\,\mathrm{MeV} \tag{1.23}$$
$$^{15}\mathrm{N} + p \rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He} \qquad Q = 4.96\,\mathrm{MeV} \tag{1.24}$$

It is worth to note that, summing over the Q values of the full CNO-I cycle (from Eq. 1.19 to Eq. 1.24), including the positron annihilation, the net released energy is 26.73 MeV, matching the pp-chain case. In case the CNO-I cycle last reaction, instead of returning to a ^{12}C nuclide as in Eq. 1.24, leads to ^{16}O, the cycle continues as CNO-II. Named also NO cycle, it provides only $\approx 0.04\%$ of the CNO-related solar luminosity:

$$^{15}\mathrm{N} + p \rightarrow {}^{16}\mathrm{O} + \gamma \qquad Q = 12.1\,\mathrm{MeV} \tag{1.25}$$
$$^{16}\mathrm{O} + p \rightarrow {}^{17}\mathrm{F} + \gamma \qquad Q = 0.6\,\mathrm{MeV} \tag{1.26}$$
$$^{17}\mathrm{F} \rightarrow {}^{17}\mathrm{O} + e^{+} + \nu_e \qquad Q = 1.74\,\mathrm{MeV} \tag{1.27}$$
$$^{17}\mathrm{O} + p \rightarrow {}^{14}\mathrm{N} + {}^{4}\mathrm{He} \qquad Q = 1.2\,\mathrm{MeV} \tag{1.28}$$
$$^{14}\mathrm{N} + p \rightarrow {}^{15}\mathrm{O} + \gamma \qquad Q = 7.35\,\mathrm{MeV} \tag{1.29}$$
$$^{15}\mathrm{O} \rightarrow {}^{15}\mathrm{N} + e^{+} + \nu_e \qquad Q = 1.73\,\mathrm{MeV} \tag{1.30}$$

The CNO cycle reactions rate strongly depends on the star metallicity and temperature. Metallicity indeed is strictly connected to the abundance of primordial metals in the stellar core. The core temperature plays a decisive role: in fact protons need an especially large thermal energy to tunnel the Coulombian barrier, which is approximately six times higher than the pp-reaction one (Eq. (1.9)), and therefore to activate the cycle (Eq. 1.17 or Eq. 1.18). The full cycle equilibrium can be achieved only in the most central regions, where $T > 1.33 \cdot 10^7$ K; for lower temperatures indeed, the $^{14}\mathrm{N} + p$ reaction of the cycle (see Eq. 1.22 and Eq. 1.29) acts as a inhibitor, draining the cycle [27].

Moving away from the Sun case, it is now clear that the CNO cycle importance in a star with respect to the pp chain is strictly related to the core temperature [41, 42]. The star luminosity in log scale as a function of matter temperature for the *pp* chain and CN cycle (respectively, solid and dashed lines) is displayed in Fig. 1.5. Results are normalized to the pp chain energy production in the Sun central core, assuming the solar metallicity and the core dynamical equilibrium. The critical temperature T_{cr} equalizing the pp chain and the CNO cycle contributions is $T_{cr} \approx 1.8 \cdot 10^7$ K. As previously mentioned, it can be checked that the CNO reactions play a secondary role for the Sun (black dot), contributing at the 1% level only to the energy production. Oppositely, CNO rate takes over for $T > T_{cr}$ stars which, if belonging to the main sequence, are also more massive than the Sun: according to the luminosity-mass relation, $L \sim M^{3.5}$.

As we discussed in Sec. 1.2, the SSM currently employs two possible sets of solar abundances, GS98 [30] and AGSS09met [31], (*high metallicity* and *low metallicity* hypotheses). Table 1.1 reports the neutrino fluxes from the reactions in the Sun [25], quoting the predictions from the B16-GS98 SSM

Solar ν	Flux B16-GS98 (HZ) [$cm^{-2}s^{-1}$]	Flux B16-AGSS09met (LZ) [$cm^{-2}s^{-1}$]	% diff.
pp	$5.98(1.0 \pm 0.006) \cdot 10^{10}$	$6.03(1.0 \pm 0.005) \cdot 10^{10}$	0.83
^{7}Be	$4.93(1.0 \pm 0.06) \cdot 10^{9}$	$4.50(1.0 \pm 0.06) \cdot 10^{10}$	8.7
pep	$1.44(1.0 \pm 0.01) \cdot 10^{8}$	$1.46(1.0 \pm 0.009) \cdot 10^{9}$	1.4
^{8}B	$5.45(1.0 \pm 0.12) \cdot 10^{6}$	$4.50(1.0 \pm 0.12) \cdot 10^{6}$	17.4
hep	$7.98(1.0 \pm 0.30) \cdot 10^{3}$	$8.25(1.0 \pm 0.12) \cdot 10^{3}$	3.4
^{13}N	$2.78(1.0 \pm 0.15) \cdot 10^{8}$	$2.04(1.0 \pm 0.14) \cdot 10^{8}$	26.6
^{15}O	$2.05(1.0 \pm 0.17) \cdot 10^{8}$	$1.44(1.0 \pm 0.16) \cdot 10^{8}$	29.8
^{17}F	$5.29(1.0 \pm 0.20) \cdot 10^{6}$	$3.26(1.0 \pm 0.18) \cdot 10^{6}$	38.6
All CNO	$4.88(1.0 \pm 0.16) \cdot 10^{8}$	$3.51(1.0 \pm 0.15) \cdot 10^{8}$	28.1

Table 1.1: Neutrino-producing reactions in the Sun [25]. The neutrino fluxes in $cm^{-2}s^{-1}$ units predicted by the B16-GS98 SSM and B16-AGSS09met SSM are listed respectively in the second and third columns, corresponding respectively to the HZ or LZ hypotheses [25], along with the related uncertainties. The percentage difference among the two predictions is reported in the fourth column.

ν fluxes correlation matrix - **B16-GS98 SSM (HZ)**

	pp	pep	hep	^{7}Be	^{8}B	^{13}N	^{15}O	^{17}F
pp	1	0.852	0.023	-0.649	-0.51	-0.34	-0.364	-0.278
pep	0.852	1	0.026	-0.595	-0.484	-0.349	-0.364	-0.291
hep	0.023	0.026	1	-0.039	-0.04	-0.016	-0.019	-0.034
^{7}Be	-0.649	-0.595	-0.039	1	0.847	0.321	0.388	0.428
^{8}B	-0.51	-0.484	-0.04	0.847	1	0.392	0.461	0.486
^{13}N	-0.34	-0.349	-0.016	0.321	0.392	1	0.985	0.282
^{15}O	-0.364	-0.364	-0.019	0.388	0.461	0.985	1	0.33
^{17}F	-0.278	-0.291	-0.034	0.428	0.486	0.282	0.33	1

ν fluxes correlation matrix - **B16-AGSS09met SSM (LZ)**

	pp	pep	hep	^{7}Be	^{8}B	^{13}N	^{15}O	^{17}F
pp	1	0.823	0.038	-0.617	-0.486	-0.294	-0.317	-0.256
pep	0.823	1	0.043	-0.551	-0.451	-0.28	-0.294	-0.23
hep	0.038	0.043	1	-0.04	-0.045	-0.039	-0.039	-0.037
^{7}Be	-0.617	-0.551	-0.04	1	0.85	0.354	0.42	0.447
^{8}B	-0.486	-0.451	-0.045	0.85	1	0.406	0.474	0.493
^{13}N	-0.294	-0.28	-0.039	0.354	0.406	1	0.981	0.307
^{15}O	-0.317	-0.294	-0.039	0.42	0.474	0.981	1	0.354
^{17}F	-0.256	-0.23	-0.037	0.447	0.493	0.307	0.354	1

Table 1.2: Correlation matrices for the solar neutrino fluxes as obtained from the B16-GS98 SSM (HZ hypothesis, top table) and B16-AGSS09met SSM (LZ hypothesis, bottom table) [36].

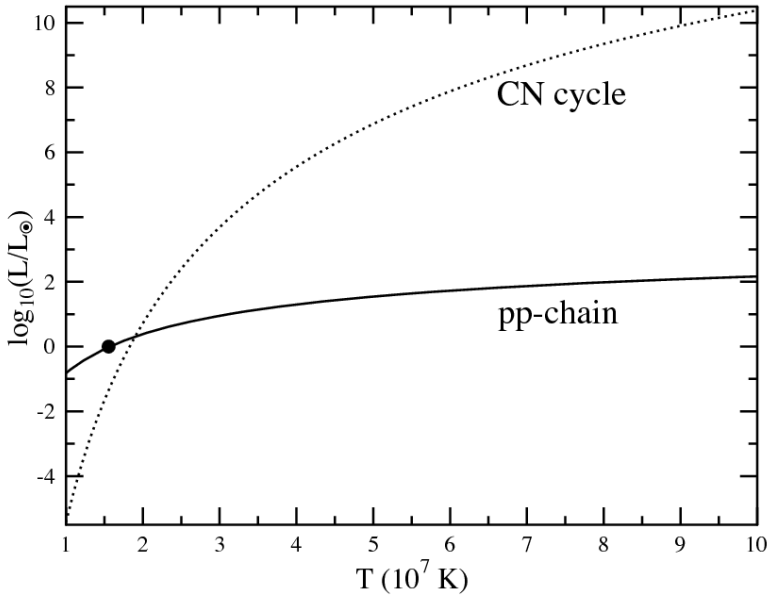


Figure 1.5: Stellar energy production: the star luminosity in log scale as a function of temperature for the *pp* chain and CN cycle (respectively, solid and dashed lines) [41, 42]. Luminosity is normalized to the pp chain energy production in the Sun central core $L_{\odot}$ and considering the solar metallicity, and also assuming the core dynamical equilibrium. The clear dominance of *pp* chain at solar temperatures range is shown: the black dot refers to the Sun core, which is powered dominantly by the *pp* chain reactions.

and B16-AGSS09met SSM respectively in the second and third columns, along with the related uncertainties; the percentage difference among the two predictions is reported in the fourth column. The correlation matrices among the several neutrino fluxes are reported in Tab. 1.2 [36].

It is crucial to underline that the two sets provide almost equivalent fluxes for the first stages of the *pp* chain (*pp*, *pep*, *hep*), while differences show up in the latest ones (^{7}Be-ν: $\approx 8\%$ difference, ^{8}B-ν: $\approx 17\%$ difference): these are the most affected reactions by the core temperature and by the metal composition in the solar matter. The CNO cycle bottleneck is the first reaction Eq. 1.19: the CNO-ν flux is strongly dependent both on the temperature and on the ^{12}C abundance, and more generally on the solar metallicity; the two flux predictions differ by about $\sim 28\%$. Thus neutrino fluxes, especially the CNO-related ones, can be used as a probe to test solar matter models. A precise measurement of the solar neutrino fluxes, including both the *pp* chain and the CNO cycle, would directly settle the controversy between high-metallicity and low-metallicity scenarios. In this thesis, we will show that Borexino is on its way to achieve the first direct indication of a CNO-ν signal: results are shown in Chapter 7. In particular, the implications of the CNO-ν measurements in terms on solar physics (metallicity discrimination and Carbon+Nitrogen abundances determination) are addressed in Sec. 7.4.

1.5.3 Helioseismology results and metallicity scenarios

The helioseismology is the observation and analysis of the acoustic waves of the solar surface in its normal modes [24, 25, 43, 44]. These oscillations provide a unique tool to investigate the interior properties of the Sun: for example, in the latest decades several predictions about the large-scale structure and the rotation of the solar interior have been inferred from helioseismology observations.

From the experimental side, the most accurate observation of solar oscillations is based on the Doppler shift measurements of emission lines in the solar spectrum. The typical velocities associated to the amplitudes of normal modes are about 10 cm/s or smaller; this implies that the dedicated experiments have to be extremely sensitive to frequency variations [43].

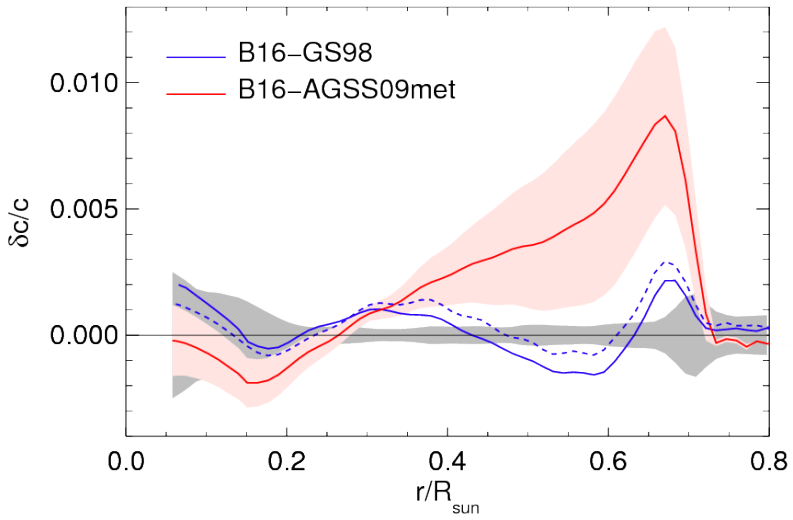


Figure 1.6: Fractional sound speed difference $\delta c = (c_{\rm data} c_{\rm mod})/c_{\rm mod}$ in the Sun [25]. The red line corresponds to B16-ASG09 SSM (low metallicity scenario) and the blue one to B16-GS98 SSMs (high metallicity scenario). The red shadow region correspond to 1σ theoretical uncertaities in SSM predictions, respectively. The black dotted curves are obtained by neglecting the opacity contribution to theoretical uncertainties.

The description of these movements on the solar surface is based on the spherical harmonics, since these functions are a natural description of the normal modes of a sphere [29]. In particular, thousand of pressure eigenmodes have been detected up to now, with multipole degrees from $l = 0$ up to $l \sim 10^3$, showing precisions at the 10^{-5} level. The identification of different oscillation modes is made possible because of the different multipole degree and frequencies associated to them; as a consequence, they probe different regions of the Sun. In particular, low degree modes ($l = 0, 1, 2, 3$) are associated to the oscillation of the innermost solar regions: they are used to probe the solar core, the region where solar neutrinos are produced.

As hinted in the previous Section, currently a tension between helioseismology and SSM predictions exists, when the photospheric low metallicity abundances are assumed. This conflict can be highlighted by analyzing the *sound speed profiles* extracted by helioseismic data or calculated by the SSM: indeed, the observed solar oscillations are in most cases of acoustic nature, consequently these frequencies are most sensitive to the sound speed [45]. Fig. 1.6 shows the fractional difference between the sound speed inferred from helioseismic frequencies and the one predicted by SSMs,

as a function of solar radius, for the two metallicity choices [25]. The B16-GS98 predictions agree better with the experimental results almost everywhere in the solar structure; on the contrary, the B16-AGSS09met model, which uses the most quoted metal abundance set, disagrees by $\sim 1\%$ with sound speed inferred from helioseismology at the bottom of the convective envelope ($r \sim 0.7 R_{\odot}$).

1.6 Solar neutrino experiments

Neutrinos interact only via weak forces[6], and therefore their typical cross sections are found in the order of $\sigma_\nu \sim 10^{-41} - 10^{-45}\,\mathrm{cm}^2$ (at $E \sim$ MeV) or lower. From the experimental point of view, these tiny cross sections put several constraints on the detector design, requiring high-mass targets or extremely intense neutrino fluxes. Neutrino experiments are often built underground, to isolate the detector from cosmic rays and other background radiation, which potentially mimic the signature of expected neutrino-induced signal. Neutrinos can not be detected directly: being electrically neutral, they are not able to ionize the materials they are crossing. We have to observe indirectly a neutrino from the secondary products emerging from its interaction with the target.

Two categories of detectors have been able to measure the solar neutrino flux: the radiochemical detectors and the real-time detectors.

- **Radiochemical detectors** pionereed the hunt for solar neutrinos. They consisted in large quantities of target liquid; neutrinos are captured on the target isotope X and produce a new isotope Y following the reaction $\overline{\nu}_e + {}^{A}_{Z}\mathrm{X} \rightarrow {}^{A}_{Z+1}\mathrm{Y} + e^-$. Chemical techniques are used to isolate and collect the atoms produced by the neutrino interactions on the target liquid. Counting techniques are used to get the number of produced atoms by exploiting the electron capture reaction on Y. This method allows to gather information only on the total rate of interactions occurring in a certain volume and beyond an energy threshold, defined by the reaction. One of the main drawbacks of this technique is that it is not possible to get information on the energy, the interaction position and on the arrival direction of the incoming neutrino.

- **Real-time detectors** are able to identify the neutrino interactions event by event, possibly measuring quantities than can be related to physical quantities of interest, like the neutrino energy or the interaction position. The employed target masses are usually water or scintillators, revealing neutrinos by means of elastic or inelastic scattering off electrons, inverse β reaction, or through processes creating Čerenkov light.

In the following, the experiments which have played a role in the field of solar neutrino physics are chronologically reviewed.

Homestake

The Homestake Solar Neutrino Observatory [2] was the first experiment to measure the solar neutrino fluxes. The experiment took place at the Homestake Gold Mine in South Dakota, USA,

[6] Being completely irrelevant, the gravitational force will be always neglected.

at 4200 mwe[7], taking data for 24 years, from 1970 to 1994. The detector was composed of a single horizontal steel tank containing 133 ton of ^{37}Cl in the form of 615 ton of tetrachloroethylene (C_2Cl_4). Homestake was a radiochemical experiment, detecting solar neutrinos by means of the inverse β reaction on ^{37}Cl atoms:

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^- \tag{1.31}$$

The related neutrino energy threshold $E_\nu^{\mathrm{thr}} = 0.814\,\mathrm{MeV}$ made the experiment sensitive only to the intermediate and high energy neutrinos, namely ^{7}Be-ν and ^{8}B-ν. The ^{37}Ar produced in the tank is extracted through chemical methods, and is counted through proportional counters, by detecting the Auger electron produced in the electron capture of the ^{37}Ar nuclei.

The collected data throughout the whole experiment lifetime led to an average neutrino rate equal to $R_{^{37}\mathrm{Cl}}^{\mathrm{Homestake}} = 2.56 \pm 0.16 \pm 0.16\,\mathrm{SNU} = 2.56 \pm 0.23\,\mathrm{SNU}$ [8], to be compared to the SSM expected rate $9.3 \pm 1.3\,\mathrm{SNU}$. In a nutshell, Homestake measured nearly one third of the solar neutrino rate predicted by the SSM, showing a more than 3σ discrepancy. This result is the first hint of the *solar neutrino problem*, consisting in the deficit of the neutrino interaction rate measured experimentally, with respect to the SSM theoretical predictions.

Gallium experiments

GALLEX [46], GNO [47, 48] and SAGE [49] have been three ^{71}Ga-based experiments which exploited the following inverse β reaction:

$$^{71}\mathrm{Ga} + \nu_e \rightarrow {}^{71}\mathrm{Ge} + e^- \tag{1.32}$$

Thanks to the reaction low energy threshold $E_\nu^{\mathrm{thr}} = 0.233\,\mathrm{MeV}$, they have been able to detect solar neutrinos from all the reactions occurring in the Sun's core: the gallium experiments are more sensitive to solar neutrinos with respect to Homestake, with the main contribution to the interaction rate given by pp-ν and ^{7}Be-ν. The ^{71}Ge atoms produced by solar neutrino interactions are extracted by means of chemical methods and are counted in proportional counters by observing their decay back to ^{71}Ga.

The GALLium EXperiment (GALLEX) was located in the Laboratori Nazionali del Gran Sasso (LNGS), in Italy, under a 3800 mwe mountain shielding. The detector consisted of a liquid gallium chloride ($GaCl_3 - HCl$) solution containing 30.3 ton of gallium. GALLEX has been active for almost six years, from 1991 to 1997; the Gallium Neutrino Observatory (GNO) is the GALLEX successor, and has been taking data from 1998 to 2003 employing the same detector setup with an improved extraction equipment. The Soviet-American Gallium Experiment (SAGE) is located in the Baksan Neutrino Observatory (BNO) in the northern Caucasus mountains, with an overhead shielding of about 4700 mwe. The averaged solar neutrino interaction rate measured by GALLEX, GNO and SAGE [48, 50, 51]) is $R_{\mathrm{Ga}} = 68.3 \pm 3.8\,\mathrm{SNU}$, which corresponds to nearly half of the SSM predicted rate. In the preliminary stage of the experiments, both the GALLEX and SAGE detector responses have been calibrated by means of powerful artificial ν_e sources.

[7] The meter water equivalent (mwe) is a standard measure of cosmic ray attenuation in underground laboratories. One mwe of a material is defined as the thickness of that material providing a shielding equivalent to one meter of water.

[8] A Solar Neutrino Unit (SNU) corresponds to a rate of 10^{-36} neutrino interactions per target atom per second.

Super-Kamiokande

Super-Kamiokande (SK) is a 50 kton mass water Čerenkov detector located in the Kamioka mine [40, 52]. From the historical side, the previous Kamiokande experiment introduced the following detection technique, which was later improved by SK due to the larger mass. SK is able to detect solar neutrinos in real time by means of the neutrino elastic scattering off water electrons:

$$\nu_x + e^- \to \nu_x + e^- \qquad x = e, \mu, \tau \tag{1.33}$$

The ν-e elastic scattering is sensitive to all active neutrinos, but the ν_μ-e^- and ν_τ-e^- total cross sections are nearly six times lower than the ν_e-e^- one ($\sigma^{\mathrm{ES}}_{\nu\text{-}\mu} = \sigma^{\mathrm{ES}}_{\nu\text{-}\tau} \approx 0.16\sigma^{\mathrm{ES}}_{\nu\text{-}e}$); this is due to ν_e additional scattering channels via W-boson exchange, not available for other flavors. The Čerenkov light emitted by the scattered electron, proportional to its path, is collected by an array of 11146 photomultipliers (PMTs) arranged on the tank surface. Interaction kinematics are reconstructed exploiting the time and charge of each PMT signal. One of the advantages of a water Čerenkov detector is given by the available information on the recoiled e^- direction, which is strictly correlated with the recoil ν direction. This feature is crucial in order to disentangle the solar neutrino signal from all the detector backgrounds.

The scattering reaction Eq. (1.33) has no intrinsic threshold; therefore the experiment threshold is dictated by the radioactive background, which dominates the low energy range. For what concerns SK, the threshold was set to about 5 MeV in the first stages of the experiment and later on, thanks to the improvement of the signal-background identification techniques, it was lowered to 3.5 MeV. Given this high threshold, SK is sensitive only to ^{8}B-ν[9]. The most recent ^{8}B-ν flux measured by SK is $\Phi^{\mathrm{SK}\text{-}^8\mathrm{B}} = (2.308 \pm 0.020 \pm 0.040) \cdot 10^6\,\mathrm{cm^2 s^{-1}}$ [52], which is nearly 0.4 times the SSM predicted flux. Beyond the solar physics, the major SK achievement is the first evidence for oscillation of atmospheric neutrinos [53].

SNO

The Sudbury Neutrino Observatory (SNO) [54, 55, 56, 57] was a real-time water Čerenkov experiment, located in the Creighton mine, near Sudbury (Canada) at a depth of 6010 mwe. Its detection mass was given by one kiloton of pure heavy water (D_2O), surrounded by an ultra-pure H_2O shield. The Čerenkov light was detected by a 9456 PMTs array, mounted on a spherical stainless steel structure. The SNO key advantage consists in the possibility to separate the neutral current contribution through the measurement of multiple channel components. Indeed, neutrinos can interact with D_2O target by means of three reactions:

$$\begin{cases} \mathrm{NC}: & \nu_x + {}^2\mathrm{H} \to n + p + \nu_x & E_\nu^{\mathrm{thr}} = 2.225\,\mathrm{MeV} & \\ \mathrm{CC}: & \nu_e + {}^2\mathrm{H} \to p + p + e^- & E_\nu^{\mathrm{thr}} = 1.4\,\mathrm{MeV} & x = e, \mu, \tau \\ \mathrm{ES}: & \nu_x + e^- \to \nu_x + e^- & & \end{cases} \tag{1.34}$$

- The charged current reaction (CC) is only allowed for ν_e; the emitted electrons can be detected collecting the produced Čerenkov light. The kinematics provide good mapping between the

[9] and also to the irrelevant flux of *hep*-ν.

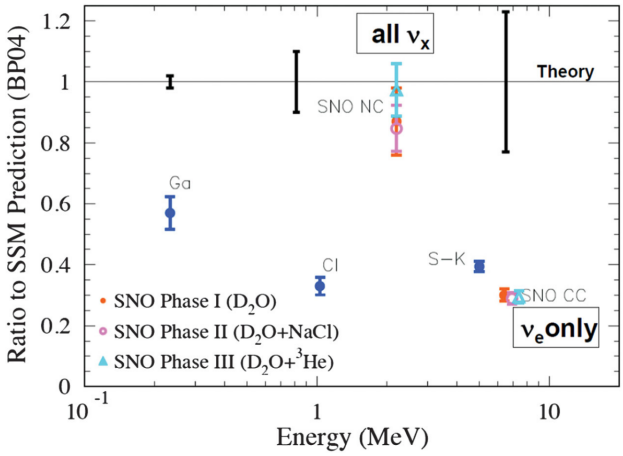


Figure 1.7: Ratios of measurements of solar neutrino fluxes (SNO, SK, Ga experiments) to the SSM predictions [58, 56].

electron energy and the incoming ν_e energy, making this channel useful to perform spectral measurements.

- The neutral current process (NC) generates a free neutron, which is typically captured on ^{2}H in a well-known mean time, resulting in the following emission of a monoenergetic de-excitation γ-ray. The outcoming signal is not sensitive to the neutrino energy or direction, but does not depend on the ν flavors: this reaction makes SNO sensitive to the full solar ν flux.
- The elastic scattering (ES) displays the lowest cross section among the three processes. As hinted previously, ν_e are more likely to interact via elastic scattering with respect to ν_μ and ν_τ. The scattered e^- has a direction strongly peaked in the forward direction oriented to the incident neutrino: thus it can be distinguished from the background thanks to its correlation with the direction of the Sun.

The fluxes determinations for the three above descripted channels are [57]:

$$\begin{cases} \Phi^{\mathrm{NC}}_{\mathrm{SNO}} = 5.54^{+0.33}_{-0.31}\,(\mathrm{stat.})^{+0.36}_{-0.34}\,(\mathrm{syst.}) \cdot 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} & \Phi^{\mathrm{NC}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} \simeq 1.00 \pm 0.08 \\ \Phi^{\mathrm{CC}}_{\mathrm{SNO}} = 1.67^{+0.05}_{-0.04}\,(\mathrm{stat.})^{+0.07}_{-0.08}\,(\mathrm{syst.}) \cdot 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} & \Phi^{\mathrm{CC}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} = 0.29 \pm 0.02 \\ \Phi^{\mathrm{ES}}_{\mathrm{SNO}} = 1.77^{+0.24}_{-0.21}\,(\mathrm{stat.})^{+0.09}_{-0.10}\,(\mathrm{syst.}) \cdot 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} & \Phi^{\mathrm{ES}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} = 0.31 \pm 0.05 \end{cases} \tag{1.35}$$

Two main conclusions can be inferred from these SNO measurements (see Fig. 1.7):

1. NC channel results match quite well the theoretical expectation, showing consistency with the SSMs neutrino fluxes prediction [59]: $\Phi^{\mathrm{NC}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} \simeq 1.00$;
2. CC and ES channels results are not compatible with the predicted fluxes from the SSM, showing significant rate deficits [60]: $\Phi^{\mathrm{CC}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} \approx 0.3$ and $\Phi^{\mathrm{ES}}_{\mathrm{SNO}}/\Phi_{\mathrm{SSM}} \approx 0.3$.

Since ν_μ and ν_τ cannot be produced in the Sun core, these results led to the confirmation of the flavor oscillation hypothesis as the solution of the solar neutrino problem, as it will be discussed in Sec. 1.7.

Borexino

Borexino [61] is a large volume liquid scintillator detector whose primary purpose is the real-time measurement of low energy solar neutrinos. It is located deep underground, in the Hall C of the Gran Sasso National Laboratory, in Italy, at approximately 3800 mwe. The Borexino design is driven by the principle of graded shielding: an inner scintillating ultra-pure core is located at the center of shielding concentric shells, with increasing radio-purity from outside to inside. The Gran Sasso mountain natural shielding, combined with the detector design, allows an extremely high muon flux suppression by a factor of $\sim 10^6$. Solar neutrinos of all flavors are detected by means of their elastic scattering off electrons of the scintillator (Eq. (2.2)).

Borexino has been taking data since 2007: designed to measure the ^{7}Be-ν flux [62], it achieved results well beyond the expectations thanks to its extreme radiopurity levels. Borexino achieved the first simultaneous measurement of all the *pp* chain neutrinos [39, 63], ruled out the day-night asymmetry of the 862 keV ^{7}Be-ν solar neutrino rate [64], and set the tightest upper limit so far on the flux of CNO solar neutrinos [65]. Borexino is the only experiment able to confirm the flavor oscillation framework for the whole range of solar neutrino energies, going also below the MeV region [39]. A more detailed description of the Borexino detector structure, scientific program, achieved results and future goals can be found in Chapter 2.

1.7 Flavor oscillations

The existence of the neutrino flavor oscillations is the solution to the solar neutrino problem. According to this hypothesis, the neutrinos can change their original flavor during their propagations. In the case of solar neutrinos, the emitted ν_e partly convert in ν_μ and ν_τ; if a detector is not sensitive to these two latter flavors, it necessarily will measure a deficitary solar neutrino flux. In the following, we review the main features of the flavor oscillation physics.

Neutrino oscillations [66, 67, 68] are a consequence of the flavor neutrino mixing, and of non-zero neutrino masses. In fact, regardless of their origin, neutrinos can only be produced and detected via weak interactions, namely, in a flavor eigenstate (ν_e, ν_μ, ν_τ). If neutrinos are massive, the states of definite flavor may not coincide with states of definite mass (ν_1, ν_2, ν_3): the newly produced neutrino is therefore a superposition of different mass eigenstates. Each of its components evolves in flight as a free quantum state separately and differently from each other. Therefore, the probability to detect different flavor eigenstates in the same neutrino beam varies with distance: this phenomenon is known as *neutrino oscillations*.

Solar neutrino experiments [2, 50, 52, 60, 62, 39, 63], atmospheric neutrino experiments [53, 69], reactor neutrino experiments [70, 71, 72, 73], long-baseline accelerator neutrino experiments [74, 75, 76, 77] and high-energy astrophysical neutrino experiments [78, 79] have provided compelling evidences for the existence of neutrino oscillations. Moreover, solar neutrino experiments point towards strong indications that the solar ν_e transitions are affected by the solar matter [80, 81] (see also Sec. 1.7.2). Flavor oscillations in atmospheric and solar neutrinos have been experimentally proved for the first time by Super-Kamiokande and SNO experiments, in 1998 and 2001 respectively; for this reason, the 2015 Nobel Prize for Physics[10] was awarded jointly to Takaaki Kajita and

[10] *The Nobel Prize in Physics 2015 was awarded jointly to Takaaki Kajita and Arthur B. McDonald for the discovery of neutrino oscillations, which shows that neutrinos have mass.*

Arthur B. McDonald, leaders of the two experiments [3].

In the following, the main phenomenology and experimental status of the neutrino oscillations will be summarized. Throughout the discussion, the plane-wave approximation for the neutrino eigenstates will be adopted, implying a well-defined and finite energy and momentum: this is often a good approximation, widely used in the practical calculations [82]. However, as neutrino production and detection are spatially localized, there must be finite intrinsic energy and momentum uncertainties, and a wave-packet description would be more general and more appropriate for a complete understanding of neutrino oscillations. The wave-packet decoherence and dispersion effects could introduce small corrections to the oscillation parameters. One can demonstrate that the plane wave approximation optimally holds if we consider stationary flux of neutrinos and/or experiments only looking at time-averaged observables [83, 84], which is our case of interest.

1.7.1 Vacuum oscillations

The existence of flavor oscillations [6] assumes that if a neutrino of a given flavor α ($\alpha = e, \mu, \tau$), with energy E, is produced in a weak interaction process, the probability to detect a neutrino of a different flavor β ($\beta = e, \mu, \tau$) at a distance L from the source is different from zero, defined as $P(\nu_\alpha \to \nu_\beta; E, L)$. This is called the $\nu_\alpha \to \nu_\beta$ oscillation or transition probability. Consequently, the probability that ν_α does not convert into a different flavor neutrino, i.e., the ν_α survival probability $P(\nu_\alpha \to \nu_\alpha; E, L)$, will be smaller than one. If only α flavor neutrinos ν_α are detected in a experiment and they take part in oscillations, a ν_α *disappearance* on the way from the source to the detector will be observed.

In the formalism of local quantum field theory, the flavor oscillations imply that the left-handed flavor eigenstates ν_α, which enter into the expression for the lepton current in the charged current weak interaction Lagrangian, are linear superpositions of the fields of three (or more) neutrinos ν_i, with definite masses $m_i \neq 0$:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^*_{\alpha i} \, |\nu_i\rangle \tag{1.36}$$

where ν_i is the left handed component of the neutrino field ν_i with mass m_i, while U is a 3×3 unitary matrix called neutrino mixing matrix, or *Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix*. $U_{\alpha i} = \langle \nu_\alpha | \nu_i \rangle$ represents the quantum amplitude for the ν_i to be detected as a ν_α. The sum is carried out on all three massive neutrino states; note that Eq. (1.36) implies that the family lepton numbers are not conserved.

This Eq. (1.36) superposition of neutrino mass eigenstates, produced in association with a given charged lepton of flavor α, is the state we refer to as the neutrino of flavor α. As hinted previously, the complete derivation of the most general neutrino oscillation probability would require a quantum wave packet description of the evolution of the massive neutrino quantum states [85]. Here we present a simplified version of this derivation corresponding to a plane-wave description, i.e.

$$|\nu_i(t)\rangle = \mathrm{e}^{-iE_i t} \, |\nu_i\rangle \tag{1.37}$$

where E_i is the ν_i energy, and τ_i is the proper time elapsed in the ν_i rest frame during its propagation. Suppose now that a flavor state $|\nu_\alpha\rangle$ is produced in a weak interaction process and after a lab-frame

elapsed time t it is observed by a neutrino detector. The latter is located at a distance L from the neutrino source and capable of detecting neutrinos of $\beta = e, \mu, \tau$ flavor. While the neutrino travels from the source to the detector, its superposition in terms of mass ν_i has changed:

$$|\nu_\alpha(t)\rangle = \sum_{i=1}^{3} U^*_{\alpha i}\, e^{-iE_i t}\, |\nu_i\rangle \tag{1.38}$$

Inverting Eq. (1.38) and combining it with Eq. (1.36), reminding also that U is an unitary matrix, one gets:

$$|\nu_\alpha(t)\rangle = \sum_{\beta=e,\mu,\tau} \left(\sum_{i=1}^{3} U^*_{\alpha i}\, e^{-iE_i} U_{\beta i} \right) |\nu_\beta\rangle \tag{1.39}$$

Eq. (1.39) implies that the superposition of massive neutrino states $|\nu_\alpha(t)\rangle$, which is originarily a pure flavor state, becomes a superposition of different flavor states at $t > 0$: this is the *neutrino flavor mixing*. The quantum amplitude for the oscillation $\nu_\alpha \to \nu_\beta$, named $\mathcal{A}(\nu_\alpha \to \nu_\beta)$, is a linear superposition over the contributions of all the mass eigenstates ν_i, given by

$$\mathcal{A}(\nu_\alpha \to \nu_\beta) \equiv \langle \nu_\beta | \nu_\alpha(t) \rangle = \sum_{i=1}^{3} U^*_{\alpha i}\, e^{-iE_i t}\, U_{\beta i} \tag{1.40}$$

The oscillation probability in vacuum $P(\nu_\alpha \to \nu_\beta) = |A(\nu_\alpha \to \nu_\beta)|^2$ can be written from Eq. (1.40) as

$$P(\nu_\alpha \to \nu_\beta) = \sum_i \sum_j U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\, e^{-i(E_i - E_j)t} \tag{1.41}$$

For the $\nu_\alpha \to \nu_\beta$ transition probability $P(\nu_\alpha \to \nu_\beta)$, only the relative phases $\delta\phi_{ij} = (E_i - E_j)t = \Delta E\, t$ of the propagation amplitudes for different mass eigenstates will induce physical consequences. According to the formulation above, and to the relativistic dispersion relation $E_i^2 = p_i^2 + m_i^2$, one can write:

$$E_i = \sqrt{p_i^2 + m_i^2} \implies E_i \simeq E + \frac{m_i^2}{2E} \implies \Delta E \simeq \frac{\Delta m_{ij}^2}{2E} \tag{1.42}$$

where $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$ is named *squared mass difference* or *squared mass splitting*. Experiments do not measure the transit time t, but one can show that to an excellent approximation, it can be taken to be $t = L/\overline{v}$, where

$$\overline{v} = \frac{p_i + p_j}{E_i + E_j} \tag{1.43}$$

is an approximation to the average of the velocities of the ν_i and ν_j components of the beam.

$$\delta\phi_{ij} \cong \frac{p_i^2 - p_j^2}{p_i + p_j} L - \frac{E_i^2 - E_j^2}{p_i + p_j} L \cong \left(m_i^2 - m_j^2\right) \frac{L}{2E} \tag{1.44}$$

where in the last step we have used the fact that for highly relativistic neutrinos $p_i \approx p_j \approx E$. Instead of using natural units, we can also express Δm^2_{ij} in eV^2, L in km, and E in GeV, making use of the following equivalence

$$\Delta m^2_{ij} \frac{L}{4E} \cong 1.27\, \Delta m^2_{ij}[\mathrm{eV}^2]\, \frac{L[\mathrm{km}]}{E[\mathrm{GeV}]} \tag{1.45}$$

An useful way to re-write the Eq. (1.41) is to separate the real and the imaginary parts of the product $U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}$. Following this idea, and making use of Eq. (1.44) also, one can show that the oscillation probability can be written as [6]:

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} + \\ - 4 \sum_{i>1} \mathcal{R}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin^2\left[1.27\, \Delta m^2_{ij} \frac{L}{E}\right] + \\ + 2 \sum_{i>j} \mathcal{I}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin^2\left[2.54\, \Delta m^2_{ij} \frac{L}{E}\right] \end{aligned} \tag{1.46}$$

The transition probability Eq. (1.46) shows an oscillatory trend with amplitude proportional to the elements in the mixing matrix, and with oscillation length L^{osc}_{ij} defined as

$$L^{\mathrm{osc}}_{ij} = \frac{4\pi E}{|\Delta m^2_{ij}|} = 2.48 \frac{E\ [\mathrm{MeV}]}{|\Delta m^2_{ij}\ [\mathrm{eV}^2]|} \tag{1.47}$$

L^{osc}_{ij} quantifies the distance scale over which the oscillation effects can be experimentally appreciable; in other words, the oscillation frequencies are proportional to the neutrino squared mass splittings. From Eq. (1.41), it is clear that a fundamental request in order to undergo flavor oscillations is that neutrinos must have different masses (i.e. $m_i \neq m_j$ and $\Delta m^2_{ij} \neq 0$ for a $\nu_i - \nu_j$ mass pair) and they must mix (i.e. $\sum_i U_{\alpha i} U_{\beta i} \neq 0$ for a given $\nu_\alpha \to \nu_\beta$ transition). The oscillation probability shown in Eq. (1.41) implies that neutrino oscillation experiments are only able to measure neutrino Δm^2_{ij} splittings, and not the absolute mass scale.

The convention widely used is to number the massive neutrinos in such a way that $m_1 < m_2$, so that Δm^2_{21} is positive. Consequently, according to Δm^2_{31} sign, the neutrino mass spectrum can be arranged in different schemes, named *orderings*:

- Normal Ordering (NO): $m_1 < m_2 < m_3$, $\Delta m^2_{31} > 0$, $\Delta m^2_{21} > 0$;
- Inverted Ordering (IO): $m_3 < m_1 < m_2$, $\Delta m^2_{32} < 0$, $\Delta m^2_{21} > 0$;

NO and IO schemes are depicted in Fig. 1.8. The ν_e, ν_μ, ν_τ flavor components for each mass eigenstate are reported respectively in red, blue and green blocks, as described by the U matrix. In the case of three neutrino mixing, only two independent neutrino mass splittings exist, since the three flavor sum rule $\Delta m^2_{21} + \Delta m^2_{31} + \Delta m^2_{23} = 0$ constraints one of the three mass splittings, removing one degree of freedom; these two free mass parameters are named Δm^2_{21} and Δm^2_{31}. The existing experimental data do not allow to claim the correct Δm^2_{31} sign, even if a combined analysis shows an overall preference for NO, quantified by the χ^2 difference $\Delta\chi^2_{\mathrm{IO-NO}} = 3.6$ [86].

From neutrino oscillation experiments, it follows that one of the two Δm^2 is two order of magnitude smaller than the other one. In particular, the neutrino mass splitting scale observed in the oscillation

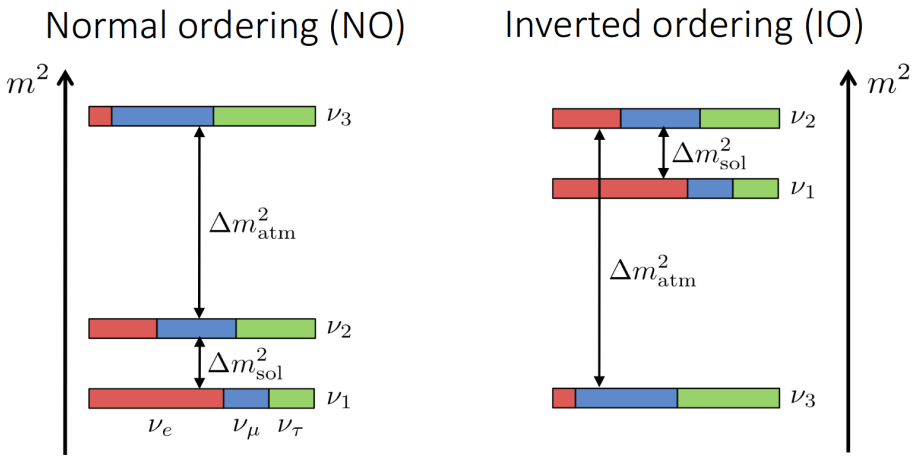


Figure 1.8: Schematic representation of normal and inverted **ordering of the neutrino mass spectrum**. In red, blue and green are reported respectively the ν_e, ν_μ, ν_τ flavor components for each mass eigenstate ν_1, ν_2, ν_3.

of atmospheric ν_μ and $\overline{\nu}_\mu$, often called $\Delta m^2_{\rm atm}$, is much larger than the neutrino mass squared difference responsible for the solar ν_e oscillations, often named $\Delta m^2_{\rm sun}$:

$$\Delta m^2_{\rm sun} = \Delta m^2_{21} \approx 7.37 \cdot 10^{-5}\,\text{eV}^2 \qquad |\Delta m^2_{\rm atm}| \approx 2.4 \cdot 10^{-3}\,\text{eV}^2 \tag{1.48}$$

It proves convenient, from the point of view of relating the mixing angles to observables, to identify $|\Delta m^2_{21}|$ with the smaller of the two neutrino masses squared splitting, $|\Delta m^2_{\rm sun}|$. Then the larger neutrino mass squared difference $|\Delta m^2_{31}|$ or $|\Delta m^2_{32}|$ (according to the ordering scheme) can be identified with $|\Delta m^2_{\rm atm}|$. The effects of Δm^2_{31} or Δm^2_{32} splittings in the oscillation of solar ν_e, and of Δm^2_{21} in the oscillation of atmospheric ν_μ and $\overline{\nu}_\mu$, and of accelerator ν_μ, are sub-dominant or marginal. In these cases the full three flavor mixing framework is not necessary and one can reduce to the simpler two flavor mixing formalism, which will be discussed in Sec. 1.7.1.

The PMNS mixing matrix can be conveniently parametrized[11] by three mixing angles θ_{12}, θ_{23}, θ_{13}, and, depending on the Dirac or Majorana nature of massive neutrinos, with a Dirac phase (δ), or with a Dirac phase and with two Majorana (δ, η_1, η_2) CP violation phases.

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \tag{1.49}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}\, e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}\, e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot U_M = \tag{1.50}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\, e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}\, e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\, e^{i\delta} & c_{23}c_{13} \end{pmatrix} \cdot U_M \tag{1.51}$$

[11] In a general case, a $n \times n$ unitary matrix can be parametrized by $n(n-1)/2$ angles and $n(n+1)/2$ phases.

Parameter	Mass ordering	Best fit value	$\pm 1\sigma$ range
$\sin^2\theta_{12} \,/\, 10^{-1}$	Normal/Inverted	2.97	$2.81 - 3.14$
$\sin^2\theta_{23} \,/\, 10^{-1}$	Normal	4.25	$4.10 - 4.46$
	Inverted	5.89	$5.67 - 6.05$
$\sin^2\theta_{13} \,/\, 10^{-2}$	Normal	2.15	$2.08 - 2.22$
	Inverted	2.16	$2.07 - 2.24$
$\Delta m^2_{21} \,/\, 10^{-5}\,\mathrm{eV}^2$	Normal/Inverted	7.37	$7.21 - 7.54$
$\lvert\Delta m^2_{31}\rvert \,/\, 10^{-3}\,\mathrm{eV}^2$	Normal	2.49	$2.46 - 2.53$
	Inverted	2.39	$2.36 - 2.43$
δ/π	Normal	1.38	$1.18 - 1.61$
	Inverted	1.31	$1.12 - 1.62$

Table 1.3: Best-fit values and 1σ allowed ranges of the 3ν oscillation parameters, derived from a global fit of the current neutrino oscillation data: mixing angles $\sin^2\theta_{12}$, $\sin^2\theta_{23}$, $\sin^2\theta_{13}$, squared mass splittings Δm^2_{21} and Δm^2_{31}, and Dirac phase δ data [86].

where:

$$U_M = \begin{pmatrix} e^{i\eta_1} & 0 & 0 \\ 0 & e^{i\eta_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.52}$$

and where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$. Consequently, in the Dirac neutrinos case, the parametrization of the PMNS matrix Eq. (1.51) is the usual for charged fermions (CKM quark mixing matrix [4]). In the case of three Dirac neutrinos, the Majorana phases are absorbed in the neutrino states, reducing to a single physical phase. In any case, the Majorana phases η_1 and η_2 never appear in oscillation phenomena since U_M is diagonal, i.e. their impact is cancelled out in transition/survival probability calculations through UU^* products.

Best-fit values and 1σ allowed ranges of the 3ν oscillation parameters, derived from a global fit of the current neutrino oscillation data combining several solar, atmospheric, accelerator experiments [86] are reported in Tab 1.3. While the four parameters Δm^2_{12}, $|\Delta m^2_{31}|$, $\sin^2\theta_{12}$ and $\sin^2\theta_{13}$ have been measured at the few % level, the remaining mixing angle $\sin^2\theta_{23}$ is less accurately known, at the $\sim 10\%$ precision level. Evaluating the PMNS matrix components with the reported angles and mass differences values, one obtains:

$$U = \begin{pmatrix} 0.82 \pm 0.01 & 0.54 \pm 0.02 & 0.15 \pm 0.03 \\ 0.35 \pm 0.06 & 0.70 \pm 0.06 & 0.62 \pm 0.06 \\ 0.44 \pm 0.06 & 0.45 \pm 0.06 & 0.77 \pm 0.06 \end{pmatrix} \cdot U_M \tag{1.53}$$

These values point to a completely different scenario with respect to the quark mixing sector: in the CKM matrix case, all the mixing angles are rather small, and the matrix is nearly diagonal [4]. In the neutrino scene instead, the flavor mixing is more enhanced. To get a more complete experimental overview, the squared-mass splittings Δm^2 and mixing angles $\tan^2\theta$ favored (solid regions) or excluded (open regions) by existing neutrino oscillation measurements [4] are shown in Fig. 1.9.

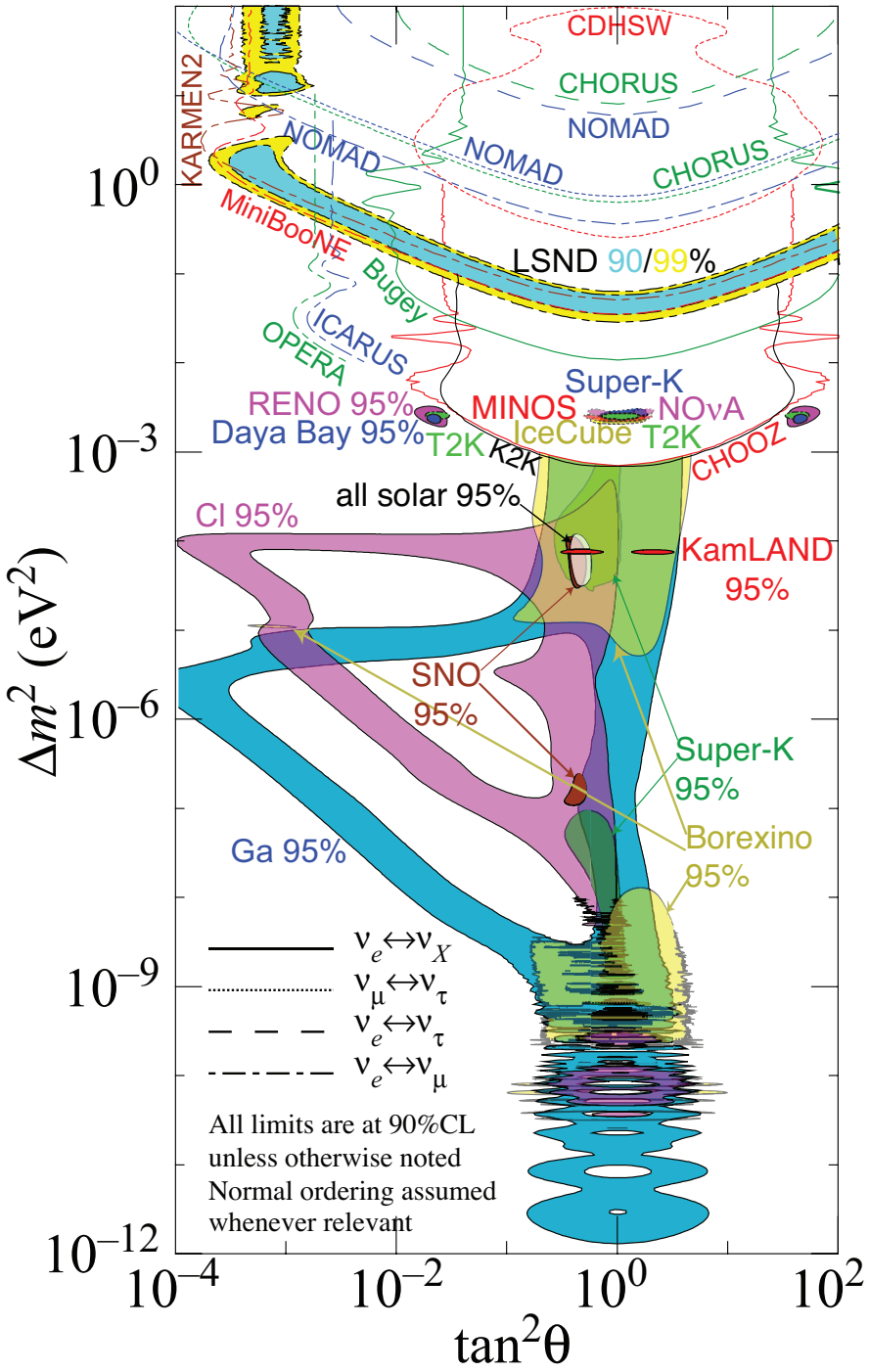


Figure 1.9: The squared-mass splittings Δm^2 [eV^2] and mixing angles favored $\tan^2\theta$ (solid regions) or excluded (open regions) by existing neutrino oscillation measurements [4]. Results are classified by channels: ν_e disappearance (solid lines), $\nu_\mu \leftrightarrow \nu_\tau$ (dotted lines), $\nu_e \leftrightarrow \nu_\tau$ (dashed lines), and $\nu_e \leftrightarrow \nu_\mu$ (dashed-dotted lines). The normal mass ordering is assumed where relevant.

Oscillation probabilities for neutrinos and antineutrinos can be related: assuming CPT invariance,

$$P\left(\nu_\alpha \to \nu_\beta\right) = P\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right) \tag{1.54}$$

From Eq. (1.41) we see that

$$P\left(\nu_\beta \to \nu_\alpha, U\right) = P\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta, U^*\right) \tag{1.55}$$

Thus, assuming that CP holds,

$$P\left(\nu_\alpha \to \nu_\beta\right) = P\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right) \tag{1.56}$$

consequently any probability for oscillation of an antineutrino is the same as that for a neutrino, except that the mixing matrix U has to be replaced by its complex conjugate. Thus, if U is not real[12], the neutrino and antineutrino oscillation probabilities can differ by having opposite values of the last term in Eq. (1.41). When CPT holds, any difference between these probabilities indicates a breaking of CP invariance by means of the non-null δ phase. The possibility of CP violation in the leptonic sector is still an open question.

Two flavor oscillation

The complete three-neutrino flavor mixing formalism is not necessary when L/E is such that an experiment observes only effects of the smaller or larger Δm^2 splitting. In this way, only two mass eigenstates are relevant, and the third one is mostly decoupled. Therefore, oscillation parameters information from experimental data can be often extracted following this simplification. The PMNS matrix U depends now on a single parameter θ, that represents a rotation angle, and reduces to a simple 2×2 form:

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{1.57}$$

where $0 \leq \theta \leq \pi/2$, and only an independent mass-squared difference $\Delta m^2 = m_2^2 - m_1^2$ exists. The oscillation probability $P\left(\nu_\alpha \to \nu_\beta\right)$ takes the well known form:

$$P\left(\nu_\alpha \to \nu_\beta\right) = \delta_{\alpha\beta} + \left(1 - 2\delta_{\alpha\beta}\right)\sin^2\left(2\theta\right)\sin^2\left(\Delta m^2 \frac{L}{4E}\right) \tag{1.58}$$

For a given neutrino flavor, Eq. (1.58) can be specified as a survival $P\left(\nu_\alpha \to \nu_\alpha\right)$ or conversion $P\left(\nu_\alpha \to \nu_\beta\right)$ probability:

$$P\left(\nu_\alpha \to \nu_\alpha\right) = 1 - \sin^2\left(2\theta\right)\sin^2\left(\Delta m^2 \frac{L}{4E}\right) \tag{1.59}$$

$$P\left(\nu_\alpha \to \nu_\beta\right)_{\alpha\neq\beta} = 1 - P\left(\nu_\alpha \to \nu_\alpha\right) = \sin^2\left(2\theta\right)\sin^2\left(\Delta m^2 \frac{L}{4E}\right) \tag{1.60}$$

[12] Clearly, all terms in U have to be different from zero.

There is a two-fold ambiguity in the interpretation of $P(\nu_\alpha \to \nu_\beta)$ in terms of neutrino mixing: the two sets of different physical parameters $\left(\theta, \Delta m^2\right)$ and $\left(\pi/2 - \theta, \Delta m^2\right)$ return the same transition probability in vacuum. One cannot tell from a measurement in vacuum whether the larger component of ν_α resides in the heavier or lighter neutrino mass eigenstate. This degeneracy is broken when there are more than two neutrinos mixed in the neutrino evolution and/or when neutrinos travel through regions of dense matter.

The probabilities Eq. (1.41) and Eq. (1.58) depend on two factors: on $\sin^2\left(\Delta m^2 \frac{L}{4E}\right)$, which exhibits oscillatory dependence on the distance L and on the neutrino energy E; and on $\sin^2(2\theta)$, which drives the absolute scale of the oscillations. To experimentally reveal the flavor oscillations, some conditions have to be fulfilled: the mixing angle θ should be different from zero, and at the same time should be large enough to allow the detection of the disappearance or appearance of neutrinos, and $E/L \geq \Delta m^2$. This last condition must be satisfied otherwise the oscillations do not have enough space to develop on the way to the neutrino detector. If the sizes of the neutrino source or of the detector are not negligible in comparison with the neutrino oscillation length, or the energy resolution of the detector is not high enough, the oscillating term will be averaged out, leading to a flat, energy independent survival probability:

$$P(\nu_\alpha \to \nu_\beta) = 1 - \frac{1}{2}\sin^2(2\theta) \tag{1.61}$$

1.7.2 Matter oscillations and MSW effect

In 1978, L. Wolfenstein [80] realized that neutrinos propagating in matter could be affected by the flavor mixing in a different way with respect to what occurs to neutrinos propagating in vacuum. The key idea relies on the different possibilities of ν_e to interact with the atoms constituting the matter with respect to ν_μ and ν_τ, since ordinary matter is composed by electrons, and not by muons or tauons. Indeed ν_e can interact both via charged current interactions (CC, virtual $W^\pm$ bosons exchange with e^-) and via neutral current interaction (NC, virtual Z boson exchange with e^-); ν_μ and ν_τ instead, only via NC processes. The flavor-dependent interaction potential generates a flavor-dependent velocity of neutrino flavor eigenstates, which can be interpreted as a phase difference of the neutrino wavefunction, leading in a not-trivial way to a modified version of flavor oscillation.

From the Feynman diagrams for the involved CC and NC processes it is possible to extract a flavor-dependent interaction potential V_α, with $\alpha = e, \mu, \tau$ [80, 81]:

$$V_\alpha = V_{\mathrm{CC}}\,\delta_{\alpha e} + V_{\mathrm{NC}} = \sqrt{2}\,G_F\left(N_e\delta_{\alpha e} - \frac{1}{2}N_n\right) \tag{1.62}$$

where $N_e = \rho\,\overline{Z/A}\,N_A$ is the electron number density of the matter, $\overline{Z/A}$ is the average ratio between matter electrons and nucleons. It is worth to underline that the NC process generates a common phase for all the flavors, and thus is not relevant for our purposes: consequently, only the CC component will be analyzed.

From now on, the case of solar matter will be considered, being the most relevant within this PhD thesis. In particular, we will consider only flavor oscillations in matter in homogeneous matter, i.e. with constant density: $N_e(t) = N_e$. This is clearly an oversimplification of the real density profile of the Sun, but allows anyway to get the more important features of the matter oscillations.

Numerically, in the Sun core $\rho \approx (100 - 150)\text{g/cm}^3$ and $\overline{Z/A} \approx 2/3$. Similarly to the vacuum case, the amplitude for the $\overline{\nu}_\alpha \to \overline{\nu}_\beta$ transition, and the probability that a α flavored neutrinos is detected as β flavored, at a time $t = x$ are

$$P(\nu_\alpha \to \nu_\beta)(t) = |\mathcal{A}(\nu_\alpha \to \nu_\beta)(t)|^2 = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 \tag{1.63}$$

Neglecting the flavor common phase shift due to the neutral currents, one gets the following Schrödinger equation for the amplitude $\mathcal{A}_{\alpha\beta}$ of Eq. 1.63:

$$i\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{A}_{\alpha\beta}(x) = \sum_i \left(\sum_j U_{\beta j} \frac{\Delta m^2_{i1}}{2E} U^*_{ij} + \delta_{\beta e}\delta_{ie} V_{\text{CC}} \right) \mathcal{A}_{\alpha i}(x) \tag{1.64}$$

where U is the flavor mixing matrix in vacuum, introduced in Sec. 1.7.1. We can reduce to the two flavor mixing case, since it is sufficient to highlight the main matter oscillations' characteristics. Eq. 1.64 can be re-written as:

$$i\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix} \mathcal{A}_{ee} \\ \mathcal{A}_{e\mu} \end{pmatrix} = \frac{1}{4E}\begin{pmatrix} -\Delta m^2 \cos 2\theta + A_{\text{CC}} & m^2 \sin 2\theta \\ m^2 \sin 2\theta & \Delta m^2 \cos 2\theta - A_{\text{CC}} \end{pmatrix}\begin{pmatrix} \mathcal{A}_{ee} \\ \mathcal{A}_{e\mu} \end{pmatrix} \tag{1.65}$$

with $A_{\text{CC}} \equiv 2EV_{\text{CC}} = 2\sqrt{2}\, E\, G_F\, N_e$. Note that it is analytically unfeasible to solve Eq. 1.65 in realistic conditions: numerical methods are required in case the potential has non-trivial dependences on position and/or energy. Let's consider briefly the simpler case of uniform solar matter, highlighting its main features. First of all, we can define an estimator ε^{Sun} quantifying the impact of the matter effect during the neutrino propagation in the Sun, as the ratio of the phase of the states propagating in matter with respect to the one for vacuum propagation:

$$\varepsilon^{\text{Sun}} \equiv \frac{\Phi_{\text{matter}}}{\Phi_{\text{vacuum}}} = \frac{V_{\text{CC}}}{\Delta m^2/2E} = \frac{\sqrt{2}G_F N_e}{\Delta m^2/2E} \approx \left(\frac{7.5 \cdot 10^{-5}\,\text{eV}^2}{\Delta m^2}\right)\left(\frac{E}{5\,\text{MeV}}\right)\left(\frac{\rho}{100\,\text{g/cm}^3}\right) \tag{1.66}$$

When $\varepsilon^{\text{Sun}} \sim 1$, the sizes of the two phases (vacuum and matter) are comparable. It is worth to point out that for $E \ll 1\,\text{MeV}$, e.g. for the pp-ν, the matter effect can be neglected since $\varepsilon^{\text{Sun}} \ll 1$; instead, for the more energetic component (i.e. ^{8}B-ν, whose spectrum extends beyond 5 MeV) it is relevant (see the neutrino energy spectrum in Fig. 1.4). From the experimental point of view, accurate measurements of solar neutrino fluxes can be exploited to probe the matter oscillation framework [63] (see Sec. 2.5.3).

Two new effective mixing parameters related to the matter flavor oscillations θ_M and Δm^2_M can be defined to diagonalize the Eq. 1.65 matrix:

$$i\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix} \mathcal{A}_{ee} \\ \mathcal{A}_{e\mu} \end{pmatrix} = \frac{1}{4E}\begin{pmatrix} -\cos 2\theta_M & \sin 2\theta_M \\ \sin 2\theta_M & \cos 2\theta_M \end{pmatrix}\begin{pmatrix} \mathcal{A}_{ee} \\ \mathcal{A}_{e\mu} \end{pmatrix} \tag{1.67}$$

Identifying $\sin 2\theta_M$ and $\cos 2\theta_M$ with the elements of Eq. 1.65 matrix, one gets a relation between θ_M and the other physical quantities involved:

$$\tan 2\theta_M = \frac{\sin 2\theta_M}{\cos 2\theta_M} = \frac{\Delta m^2 \sin 2\theta}{\Delta m^2 \cos 2\theta - A_{CC}} = \tan 2\theta \frac{1}{1 - \frac{A_{CC}}{\Delta m^2 \cos 2\theta}} \tag{1.68}$$

Similarly, for Δm^2_M:

$$\Delta m^2_M = \sqrt{(\Delta m^2 \cos 2\theta - A_{\rm CC})^2 + (\Delta m^2 \sin 2\theta)^2} \tag{1.69}$$

Once these definitions are set, we can describe the survival probability $P(\nu_e \to \nu_e)$ in a uniform density matter, which assumes a well-known structure:

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta_M) \sin^2\left(\Delta m^2_M \frac{L}{4E}\right) \tag{1.70}$$

which shows the same structure of the vacuum case survival probability Eq. (1.58), replacing $\theta \to \theta_M$ and $\Delta m^2 \to \Delta m^2_M$. Mikeyev and Smirnov pointed out that if $A_{\rm CC} = \Delta m^2 \cos 2\theta$ holds, a resonant behaviour in Eq. 1.68 is derived [81]. Consequently, the presence of matter can lead to a strong enhancement of the Eq. 1.70 oscillation probability, leading to a *maximal mixing* since $\sin^2(2\theta_M) \to 1$. This is called the Mikeyev-Smirnov-Wolfenstein (MSW) mechanism resonance. At fixed E, it takes place at the *resonance density* $N_e^{\rm res}$:

$$N_e^{\rm res} = \frac{\Delta m^2 \cos 2\theta}{2E\sqrt{2}G_F} \approx 6.65 \cdot 10^6 \frac{\Delta m^2[\mathrm{eV}^2]}{E[\mathrm{MeV}]} N_A \cos 2\theta \ \ [\mathrm{cm}^{-3}] \tag{1.71}$$

On the other side, for a fixed N_e and $\Delta m^2 \cos 2\theta$ couple, Eq. 1.71 holds at the *resonance energy* $E^{\rm res}$. In addition, since $A_{\rm CC} > 0$ holds, the resonance takes place only for $\theta < \pi/4$; the matter effects break the complementary angles invariance $\theta \to \pi/2 - \theta$ observed for the the vacuum oscillation case (see Sec. 1.7.1).

1.8 Neutrino mass scale determination

Since neutrino oscillations are not sensitive to the absolute scale of neutrino masses (m_i), dedicated experiments are needed to set relevant constraints. Hereafter, the constraints coming from double β decay experiments, from β spectrum measurements, and from cosmology are shortly reviewed.

- Double β^- decays are isobaric transitions from a $(A, Z+2)$ nucleus to a (A, Z) daughter nucleus mediated by weak interaction: $(A, Z+2) \to (A, Z) + 2e^- + 2\overline{\nu}_e$. The available energy $Q_{\beta\beta}$ is shared between the decay products. If neutrinos are Majorana fermions, processes such as neutrinoless double-beta decay ($0\nu\beta\beta$) would be allowed to take place, violating the lepton number by two units. In fact, in a $0\nu\beta\beta$ decay, $(A, Z) \to (A, Z+2) + 2e^- + Q_{\beta\beta}$ the two emitted ν_e would annihilate, proving their Majorana nature. For a given double β decaying isotope, the $0\nu\beta\beta$-decay half-life $T^{0\nu}_{1/2}$ can be estimated as:

$$T^{0\nu}_{1/2} = \left(G|\mathcal{M}|^2 \langle m_{\beta\beta}\rangle^2\right)^{-1} \simeq 10^{27-28} \left(\frac{0.01\,\mathrm{eV}}{\langle m_{\beta\beta}\rangle}\right)^2 \ \ [\mathrm{y}] \tag{1.72}$$

  where G is the the phase-space factor, $\mathcal{M}$ is the nuclear matrix element, and the effective mass $\langle m_{\beta\beta}\rangle = \sum_i U^2_{ei} m_i$ is strictly related to the neutrino absolute mass scale by means of PMNS element combination. From Eq. (1.72) it follows that ton-scale experiments are needed to probe the physically interesting region $T^{0\nu}_{1/2} \sim 10^{28}$ y and $\langle m_{\beta\beta}\rangle \sim 0.01$ eV. Several

experiments, addressing complex technological challenges and employing different double-beta decaying isotopes, have strongly contributed in recent years to lower the $\langle m_{\beta\beta} \rangle$ upper limit [87, 88, 89, 90, 91]. The best sensitivity so far has been achieved by GERDA detector [92], located at LNGS and searching for ^{76}Ge $0\nu\beta\beta$ decay: no signal has been observed, and the derived half-life limit $T_{1/2} > 0.9 \cdot 10^{26}$ y (90% C.L.) corresponds to a sensitivity to the absolute mass scale in β decay of $0.15 - 0.44$ eV.

- Direct and model-independent information about $m_{\overline{\nu}_e}$ can be obtained, in laboratory, by means of kinematic studies of the β^- decay spectrum of tritium $^3\mathrm{H} \rightarrow {}^3\mathrm{He}^+ + e^- + \overline{\nu}_e$ [93, 94, 95]. The smoking gun for the mass measurement is given by the small, characteristic β spectrum shape distortion close to its endpoint, due to $m_{\overline{\nu}_e} \neq 0$. The most stringent m_ν upper limit has been derived by the KATRIN experiment [96], $m_{\overline{\nu}_e} < 1.1$ eV at 90% C.L. KATRIN design goal is to improve its sensitivity of one order of magnitude with respect the latest release, down to 0.2 eV (90% C.L.) after 5 years of data taking.

- Cosmology experiments return the strongest mass limits in the sum of neutrino masses $\sum_i m_i$, which represent the flavor-blind total *gravitational charge*. As a drawback, the resulting limits are strongly model dependent, according on the model complexity and number of free parameters, on the input dataset, and are especially sensitive to systematic effects. Assuming the existence of three light massive neutrinos and the validity of the ΛCDM (Lambda-Cold Dark Matter) model, the best limit comes from the Planck CMB survey [97], $\sum_i m_i < 0.26$ eV at 95% CL. Combining this result with the Baryon Acoustic Oscillations data and the CMB lensing, the previous limit is as low as

$$\sum_i m_i < 0.12\,\mathrm{eV} \qquad 95\%\,\mathrm{CL} \tag{1.73}$$

1.9 Sterile Neutrinos

Sterile neutrinos are hypotetical light fermions that do not take part to Standard Model interactions with the exception of the gravitational one [98]. Currently, compelling constraints on the number of active[13] neutrinos exist. First of all, the Standard Model predicts the decay width of the Z^0 boson into neutrino couples, which is expected to be proportional to the number of light[14], active, left-handed neutrinos. At present the measurement of the invisible Z^0 width carried out at LEP yields $N_\nu = 2.984 \pm 0.008$ [4], hence it follows that there are only three light active neutrinos. Moreover, the standard three-flavor neutrino oscillation paradigm has been established by several solar atmospheric [53], reactor [70, 71, 72] and accelerator experiments [74, 75, 77].

However, long standing anomalies coming from several experiments led to the tantalizing hint of the existence of at least one additional sterile neutrino state, which guides the oscillations at short distances. The *sterile* term usually refers to neutrinos with right-handed chirality, which may be added in a minimal extension of the Standard Model. The existence of right-handed neutrinos is theoretically well-motivated, as all other known fermions have been observed with both chiralities; furthermore, they could explain the mass generation mechanism for active states in a natural way. The idea of the fourth neutrino state [99] has been proposed by Pontecorvo[15], and current experimental results do not allow to definitely prove it or reject it. If the existence of sterile neutrinos will be confirmed, a major revolution in the landscape of particle physics will take place.

1.9.1 Experimental anomalies

In the following, the main experimental results indicating the possible existence of sterile neutrinos will be shortly described.

Gallium anomaly

GALLEX and SAGE have been introduced in Sec. 1.6 as two radiochemical neutrino detection experiments designed to measure the solar neutrino flux. To approve the experiment concept and directly determine the ν_e detection efficiency, the experiments employed intense artificial ^{51}Cr and ^{37}Ar ν_e radioactive sources placed below the detectors [100, 101, 49, 51]. The radioactive nuclei ^{51}Cr and ^{37}Ar decay through electron capture emitting mono-energetic ν_e with hundreds of keV energy:

$$e^- + {}^{51}\mathrm{Cr} \rightarrow {}^{51}\mathrm{V} + \nu_e \qquad E_\nu = \begin{cases} 427\,\mathrm{keV} & 9.0\% \\ 432\,\mathrm{keV} & 0.9\% \\ 747\,\mathrm{keV} & 81.6\% \\ 752\,\mathrm{keV} & 8.5\% \end{cases} \tag{1.74}$$

[13] An *active neutrino* is the left-handed neutrino described in the Standard Model.
[14] Neutrinos are defined as *light* if $m_\nu < m_{Z^0}/2 \approx 45.6\,\mathrm{GeV}$.
[15] From Ref. [99]: *neutrino oscillations can convert potentially active particles into particles that are, from the point of view of ordinary weak interactions, sterile, i.e. practically unobservable, since they have the incorrect helicity.*

$$e^- + {}^{37}\mathrm{Ar} \rightarrow {}^{37}\mathrm{Cl} + \nu_e \qquad E_\nu = \begin{cases} 811\,\mathrm{keV} & 90.2\% \\ 813\,\mathrm{keV} & 9.8\% \end{cases} \tag{1.75}$$

The emitted ν_e were then detected by means of the same inverse β reaction Eq. 1.32 exploited for the solar neutrinos detection $\nu_e + {}^{71}\mathrm{Ga} \rightarrow {}^{71}\mathrm{Ge} + e^-$. The ratio of measured versus predicted interaction ^{71}Ge rate in GALLEX and SAGE is

$$\overline{R} = 0.86 \pm 0.05 \tag{1.76}$$

where $\overline{R}$ is the averaged result of the two experiments. Consequently, the number of measured events is about 2.8σ far from than the expectations. This events deficit is named *gallium anomaly*, and could be a manifestation of short-baseline neutrino oscillations of $\overline{\nu}_e$ in sterile state. The non-oscillation hypotesis is disfavoured at the 99.23% C.L. Being the source-detector distance $L < 2\,\mathrm{m}$, this hypotetical oscillation would require a high mass splitting to be appreciable, roughly located between $1\,\mathrm{eV}^2$ and $10\,\mathrm{eV}^2$. Combining GALLEX and SAGE results, and naming as θ_{new} and Δm_{new} the mixing parameters related to the new hypotetical oscillation, one can extract their best fit values:

$$\sin^2(2\theta_{\mathrm{new}}) = 0.5 \qquad \Delta m^2_{\mathrm{new}} = 2.2\,\mathrm{eV}^2$$

while the 99% C.L. lower limits with respect to the null signal hypothesis are $\sin^2(2\theta_{\mathrm{new}}) > 0.07$ and $\Delta m^2_{\mathrm{new}} > 0.35\,\mathrm{eV}^2$.

LSND and MiniBooNE anomalies

LSND [102] (Liquid Scintillator Neutrino Detector) has been an experiment designed to search $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$ oscillation from $\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu$ decay at rest with high sensitivity.

The source of neutrinos for the measurement is the interaction of the intense 798 MeV proton beam at the Los Alamos Neutron Science Center (LANSCE) with a fixed target, which produces a large number of pions, mostly π^+ . The π^- are mainly absorbed: only a small fraction decay to μ^-, which in turn are largely captured. Thus, the resulting neutrino source is dominantly due to $\pi^+ \rightarrow \mu^+ + \overline{\nu}_\mu$ and $\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu$ decays, most of which decay at rest. Such a source has a small contaminants of $\overline{\nu}_e$, and is therefore ideal to search for $\overline{\nu}_e$ appearing from oscillations $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$. Clearly, the measurement of the ratio of expected events and observed ones requires an accurate knowledge of the initial antineutrino flux. The LSND detector consisted of a $8.3\,\mathrm{m} \times 5.7\,\mathrm{m}$ cylindrical tank with center located 29.8 m from the neutrino source. 1220 photomultiplier tubes covered $\approx 25\%$ of the tank surface area, which was filled with 167 tons of liquid scintillator consisting of mineral oil. The $\overline{\nu}_e$ component found in the beam, either due to signal or background, could be detected via inverse β reaction $\overline{\nu}_e + p \rightarrow n + e^+$ in the mineral oil target of the LSND detector.

Left panel of Fig. 1.10 compares the number of expected and predicted events as a function of the positron energy. A total excess of $87.9 \pm 22.4 \pm 6.0$ inverse β reaction events with e^+ energy between 20 MeV and 60 MeV is observed above expected neutrino-induced backgrounds. This excess is hardly justifiable in the well estabilished three flavor framework, but would be compatible with an oscillation to a fourth state with mass squared difference in the range $(0.2 - 10)\,\mathrm{eV}^2$. Right panel of Fig. 1.10 shows favored regions for $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$ oscillations in the mixing parameters

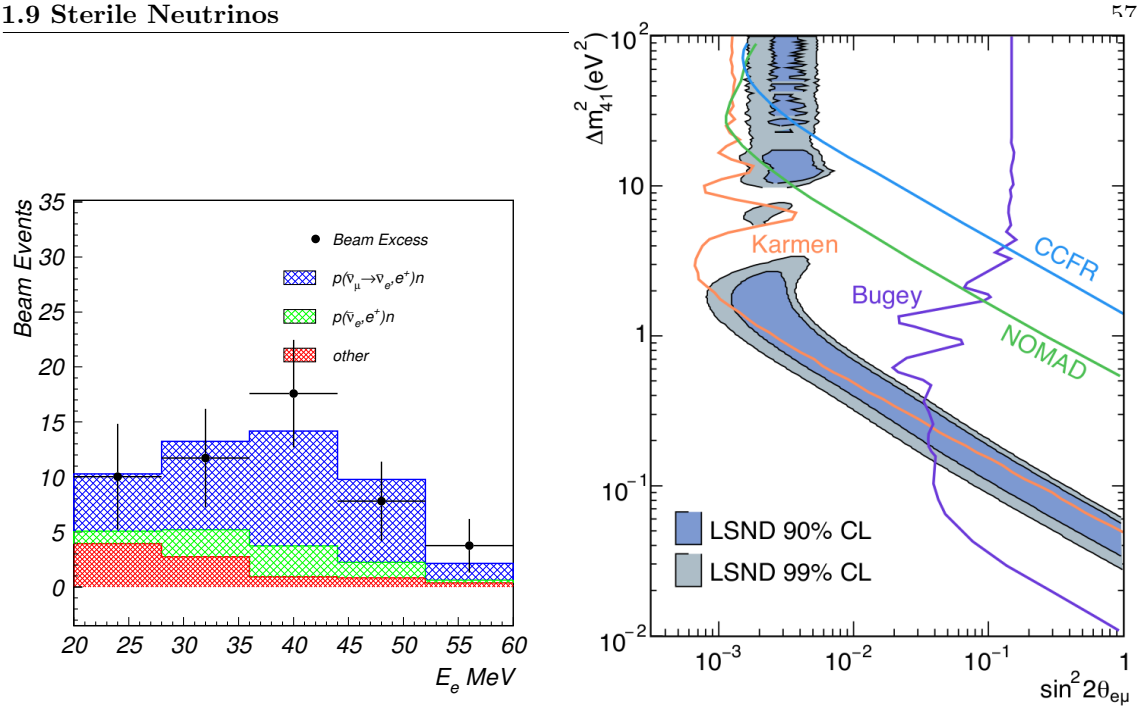


Figure 1.10: LSND final results. Left panel: the $\overline{\nu}_e$ events energy distribution collected between 1993 and 1998 at LSND experiment as a function of the positron energy. Right panel: favored regions for $\overline{\nu}_\mu \to \overline{\nu}_e$ oscillations in the mixing parameters $\left(\sin^2(2\theta_{\text{new}}), \Delta m^2_{\text{new}}\right)$ space at 90% C.L. or 99% C.L. from LSND neutrino analysis (blue and grey areas respectively) [102] compared with the 95% C.L. exclusion limits obtained by other short-baseline experiments (Karmen, NOMAD, Bugey, CCFR).

$\left(\sin^2(2\theta_{\text{new}}), \Delta m^2_{\text{new}}\right)$ space at 95% CL from LSND neutrino analysis (solid line) and antineutrino analysis (dashed line).

The need to unambiguously verify or refute the LSND controversial results led to the proposal of MiniBooNE experiment [103] at Fermilab. The searched oscillations were $\nu_\mu \to \nu_e$ and $\overline{\nu}_\mu \to \overline{\nu}_e$. The detector location was chosen to satisfy $L[\text{m}] \,/\, E[\text{MeV}] \approx 1$, similarly to the LSND case, maximizing the sensitivity to oscillations at $\Delta m^2_{\text{new}} \approx 1\,\text{eV}^2$. The detector consisted of a 12.2 m diameter sphere filled with pure mineral oil. Neutrino interactions in the detector produce final-state electrons or muons, inducing scintillation and Čerenkov light detected by the 1520 photomultiplier tubes covering the detector interior.

MiniBooNE results until 2010 are displayed in Fig. 1.11. In the antineutrino channel, no $\overline{\nu}_e$ excess are found in the LSND energy range, i.e. 475 MeV and 1250 MeV, but in the energy region below 475 MeV. This excess is hardly justifiable within the three neutrino flavor model, but is consistent with antineutrino oscillations in the $0.2\,\text{eV}^2 \leq \Delta m^2_{\text{new}} \leq 1\,\text{eV}^2$ range and have some overlap with the evidence for antineutrino oscillations from LSND. The neutrino mode running also shows an excess at low energy of 162.0 ± 47.8 events, pointing to a deviation from the no oscillation hypotesis of 3.4σ, but the energy distribution of the excess is marginally compatible with a simple two neutrino oscillation formalism. While this incompatibility might be explained by unexpected systematic uncertainties and backgrounds, expanded oscillation models with several sterile neutrinos can reduce the discrepancy by allowing CP violating effects (see Sec. 1.9.2). In May 2018, the MiniBooNE experiment recently reported an anomalous 4.5σ excess of electron-like events, consistent with ν_e

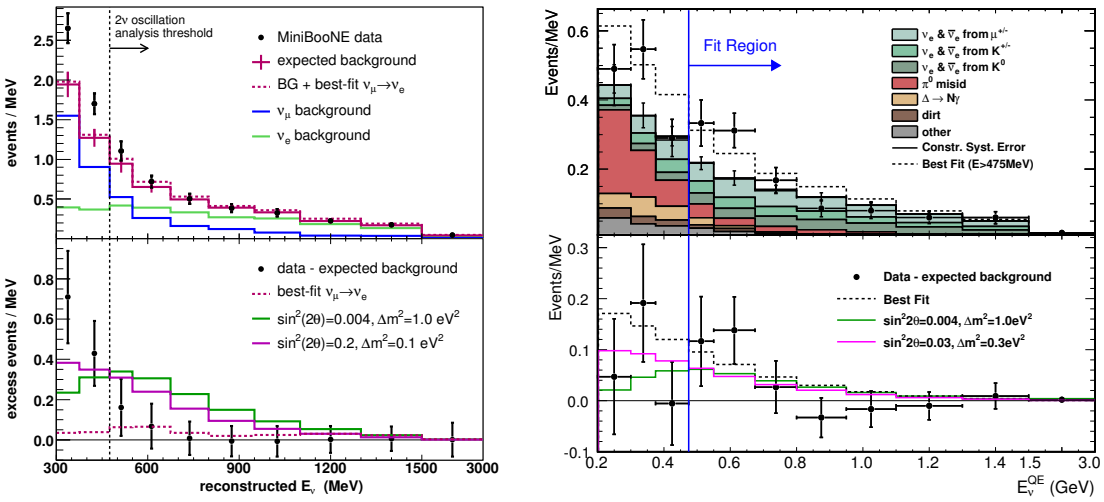


Figure 1.11: MiniBooNE main results until 2010. On the top: ν_e and $\overline{\nu}_e$ reconstructed energy distributions, respectively on left panel and right panel. On the bottom: event excess as a function of reconstructed energy (black), and the two expected distributions according to LSND new oscillation hypotesis at high and low Δm^2_{new} (green and purple histograms).

appearance from a ν_μ beam at short-baseline [104]. The acquired data are consistent in energy and magnitude with the excess of events reported by LSND.

The MiniBooNE successor is the MicroBooNE experiment [105], which currently operates a large 170 ton Liquid Argon Time Projection Chamber (LAr TPC) that is located on the Booster neutrino beam line at Fermilab. The experiment started collecting the first neutrino data in October 2015. MicroBooNE investigates the low energy excess events observed by the MiniBooNE experiment and will measure low energy neutrino cross sections on argon in the 1 GeV energy range.

On the other hand, there are a number of neutrino oscillation experiments that do not support evidences for sterile hypotesis. The KARMEN experiment [106], displaying a very similar design to LSND one, observed no such indications; however, a joint analysis of the two experiments shows that their data sets are compatible with oscillations occuring either in a mass band from $0.2\,\text{eV}^2$ to $1\,\text{eV}^2$ or in a region around $7\,\text{eV}^2$. Also, a number of ν_μ disappearance searches in the relevant mass range, as MINOS experiment, returned negative results [107].

Reactor anomalies

Further hints for the existence of sterile neutrinos come from measurements of the neutrino flux from short baseline reactor (SBL) experiments. In preparation for the Double Chooz reactor experiment [73], the Saclay reactor neutrino group re-evaluated the specific reactor antineutrino flux for ^{235}U, ^{239}Pu, ^{241}Pu, and ^{238}U, using the new existing nuclear data. In 2011, they reported results [108], corresponding to a flux that is a few percent higher than the previous prediction. The net result is a +3% shift in the predicted emitted spectra. The origin of these biases was found in the approximate treatment of nuclear data and corrections to the Fermi theory.

This also required a re-analysis of the ratio of observed event rate to predicted rate for 19 published

neutrino oscillation experiments at reactor-detector distances below 100 m. After including the re-analysis of these experiment, the ratio for all these reactor experiments is $R = 0.927 \pm 0.023$, with a significance above 3σ from $R = 1$ hypothesis. This evidence has been called *reactor anomaly*: a fraction of antineutrinos could oscillate in the sterile state, becoming undetectable. The ratios of observed and expected antineutrinos fluxes are displayed in Fig. 1.12 as a function of the reactor-detector distance for each reactor experiment; the hypotesis of absence and existence of sterile state are compared in solid lines. The allowed regions in the mixing parameters $(\sin^2 2\theta_{\rm new}, \Delta m^2_{\rm new})$ space obtained from the fit of the reactor neutrino data are reported in Fig. 1.13.

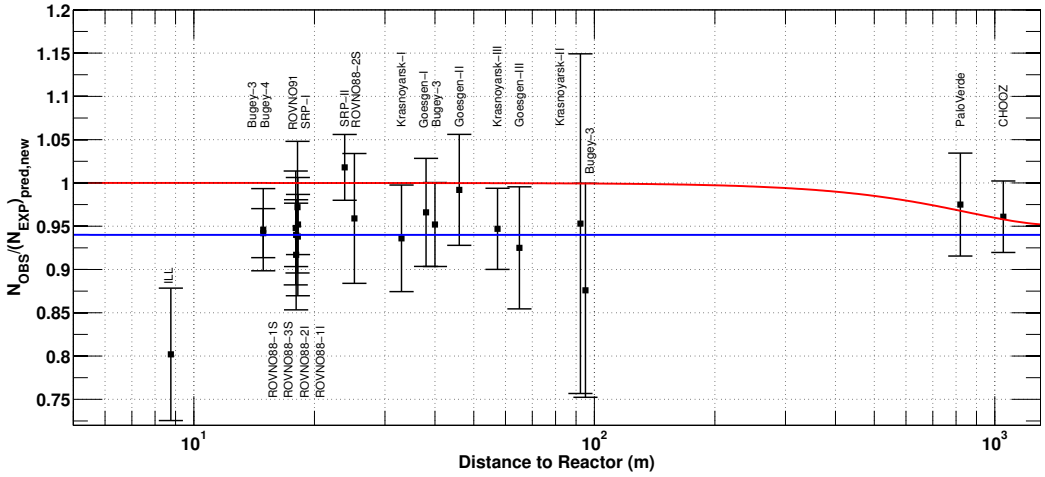


Figure 1.12: Significance of reactor experiments anomaly: in red solid line the three active neutrinos mixing model prediction is reported, while in solid blue line the $3+1$ one (assuming a mixing angle $\sin^2(2\theta_{\rm new}) = 0.12$ and $\Delta m^2_{\rm new} > 1\,{\rm eV}^2$).

1.9.2 Interpretation of the anomalies

The experimental anomalies described in the previous section could be interpreted by assuming the existence of one or more sterile neutrinos in the $\Delta m^2_{\rm new} \approx {\rm eV}^2$ range, which mixes with active ones. The missing $\overline{\nu}_e$ (ν_e) flux would be due to the impossibility to detect $\overline{\nu}_s$ (ν_s). We will interpret LSND, MiniBooNE, reactor experiments and other experimental results within sterile neutrino oscillation models, analyzing the two most plausible ones: 3+1 and 3+2[16] schemes[17].

3+1 model

The simplest extension of three-neutrino flavor mixing introduces one massive sterile neutrino which mixes with the three active flavors. In this four-neutrino mixing framework, the flavor neutrino basis is given by the three active neutrinos ν_e, ν_μ, ν_τ and by a sterile neutrino ν_s. The three active

[16] It is worth to anticipate that the existence of more than one sterile neutrino is strongly disfavoured by cosmological $\sum m_i$ constraints [97].

[17] Three sterile neutrino states, one for each active flavor generation, seem most natural to exist; however, a 3+3 model is not found to be better than the 3+2 fits, and the implications would be very similar. Therefore, a discussion of 3+3 model is not included here (see Ref. [109]).

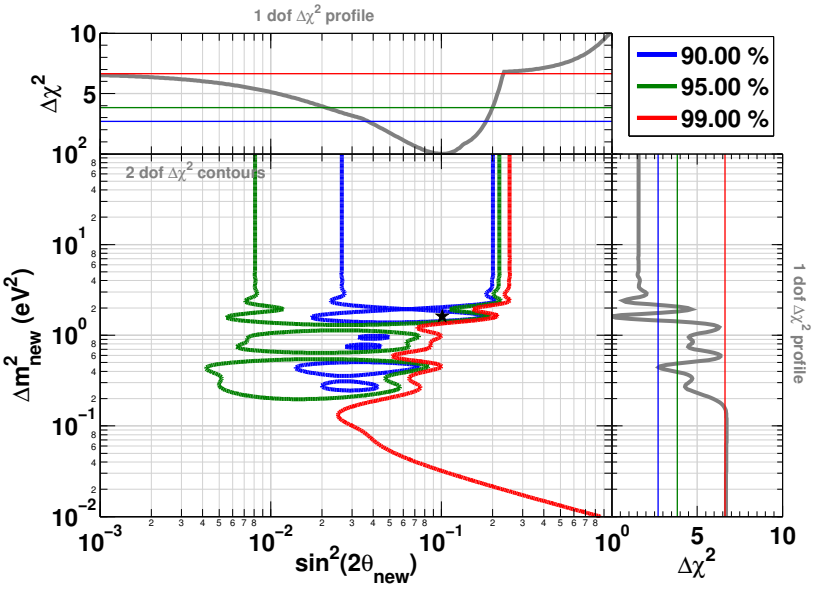


Figure 1.13: Reactor anomaly global fit. Allowed regions in the mixing parameters $(\sin^2(2\theta_{\text{new}}), \Delta m^2_{\text{new}})$ space are obtained to the fit of the reactor neutrino data, without any energy spectra information, in $3+1$ neutrino hypothesis, with $\sin^2(\theta_{13}) = 0$.

ones are mainly composed of three massive neutrinos ν_1, ν_2, ν_3 with light masses m_1, m_2, m_3, while the latter is mainly composed of a heavy neutrino ν_4 with mass m_4 such that $\Delta m^2_{\text{new}} = \Delta m^2_{41}$ and $m_1 \approx m_2 \approx m_3 \ll m_4 \implies m_4 \approx \sqrt{|\Delta m^2_{41}|}$. This leads to an extended U 4×4 PMNS mixing matrix that connects the 3+1 flavors to the four mass states[18]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} & U_{\mu 4} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} & U_{\tau 4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{pmatrix} \tag{1.77}$$

The sterile state mixing with the active neutrinos has to be sufficiently small in order not to spoil the successful three flavor mixing explanation of solar, atmospheric and long-baseline neutrino oscillation measurements. In other words, the sterile neutrino mass eigenstate m_4 must be mostly sterile:

$$|U_{\alpha 4}| \ll 1 \qquad \alpha = e, \mu, \tau \tag{1.78}$$

while $|U_{s4}| \approx 1$. Under these hypotheses, the effects of active-sterile neutrino mixing in solar [110] and atmospheric [111] neutrino experiment results are tiny and out of current detector precisions, but should be revealed sooner or later. The two possible mass schemes, normal and inverted, are depicted in Fig. 1.14. Since the inverted scheme assumes three massive neutrinos at the eV^2 scale, it is highly disfavored by cosmological constraints [97] over the normal one.

[18] So-called $2+2$ four-neutrino mixing schemes are strongly disfavored by the absence of any macroscopic signal of sterile neutrino effects in solar and atmospheric neutrino data.

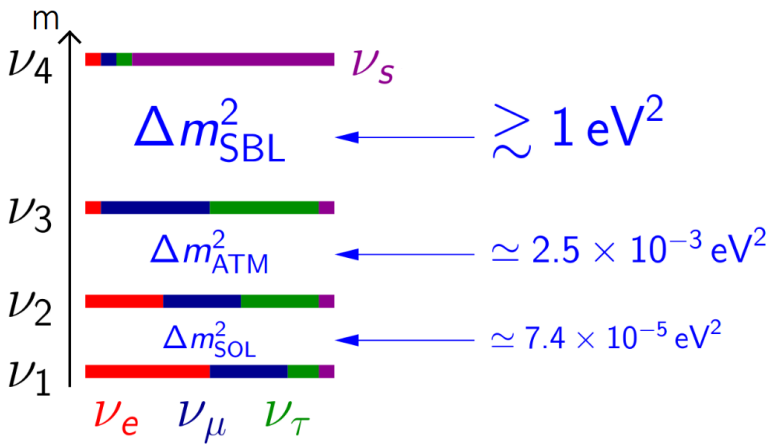


Figure 1.14: Schematic illustration of the 3+1 mass ordering, taking into account the $\Delta m^2_{21} \approx \Delta m^2_{32} \approx \Delta m^2_{31} \equiv 0$ approximation.

Neutrino oscillations in the sterile state are generated by a squared-mass difference $\Delta m^2_{\mathrm{new}} \gtrsim 0.1\ \mathrm{eV}^2$, which is, at least, three order of magnitude higher than the two measured solar and atmospheric splittings Δm^2_{21} and Δm^2_{31}. For this reason, in experiments sensitive to this squared mass difference one can approximate $\Delta m^2_{21} \approx \Delta m^2_{32} \approx \Delta m^2_{31} \equiv 0$, so that we can reduce to a two flavor oscillation problem:

$$\sin^2\left(\frac{\Delta m^2_{21}L}{4E}\right) \approx \sin^2\left(\frac{\Delta m^2_{31}L}{4E}\right) \ll \sin^2\left(\frac{\Delta m^2_{41}L}{4E}\right)$$

For a given flavor, the conversion and survival probabilities in 3+1 schemes in these experiments are [112]:

$$P_{\nu_\alpha\to\nu_\alpha}(L,E) = 1-4\left(1-|U_{\alpha4}|^2\right)|U_{\alpha4}|^2\sin^2\left(\frac{\Delta m^2_{41}L}{4E}\right) \equiv 1-\sin^2(2\theta_{\alpha\alpha})\sin^2\left(\frac{\Delta m^2_{41}L}{4E}\right) \quad (1.79)$$

$$P_{\nu_\alpha\to\nu_\beta}(L,E) = 4|U_{\alpha4}|^2|U_{\beta4}|^2\sin^2\left(\frac{\Delta m^2_{41}L}{4E}\right) \equiv \sin^2(2\theta_{\alpha\beta})\sin^2\left(\frac{\Delta m^2_{41}L}{4E}\right) \quad (1.80)$$

for $\alpha,\beta = e,\mu,\tau,s$. Therefore all oscillation probabilities in sterile state depend, with excellent approximation, only on the absolute value of the largest squared mass difference, that is Δm^2_{41}. The oscillation amplitudes depend only on the absolute values of the elements in the fourth column of the mixing matrix $|U_{\alpha4}|$. From Eq. (1.79) and Eq. (1.80) one can see that the electron and muon neutrino and antineutrino appearance and disappearance in short-baseline experiments depend on $|U_{e4}|^2$ and $|U_{\mu4}|^2$. Constraints on their product $\sin^2(2\theta_{e\mu}) \approx \sin^2(2\theta_{ee})\cdot\sin^2(2\theta_{\mu\mu}) = 4|U_{e4}|^2|U_{\mu4}|^2$ can be extracted from the experiments searching for $\nu_e \to \nu_\mu$, as LSND, MiniBooNE and in future MicroBooNE. The disappearance experiments determine separately $|U_{e4}|^2$ (reactor or Ga-based) and $|U_{\mu4}|^2$ (MINOS) through the measurements of $\sin^2(2\theta_{ee})$ and $\sin^2(2\theta_{\mu\mu})$ (see Eq. (1.79)).

The allowed regions in the $(\sin^2(2\theta_{\mu e}), \Delta m^2_{41})$, $(\sin^2(2\theta_{ee}), \Delta m^2_{41})$ and $(\sin^2(2\theta_{\mu\mu}), \Delta m^2_{41})$ planes obtained in the 3+1 global fit [112] are shown in Fig. 1.15. The included experiments are:

- $\nu_\mu \to \nu_e$ and $\overline{\nu}_\mu \to \overline{\nu}_e$ appearance data: LSND, MiniBooNE until 2010[19], BNL-E776, KARMEN, NOMAD, ICARUS, OPERA;
- $\nu_e \to \nu_e$ and $\overline{\nu}_e \to \overline{\nu}_e$ disappearance data: experiments at nuclear reactors, GALLEX, SAGE;
- $\nu_\mu \to \nu_\mu$ and $\overline{\nu}_\mu \to \overline{\nu}_\mu$ disappearance data: CDHSW, atmospheric neutrino experiments, MINOS, MiniBooNE.

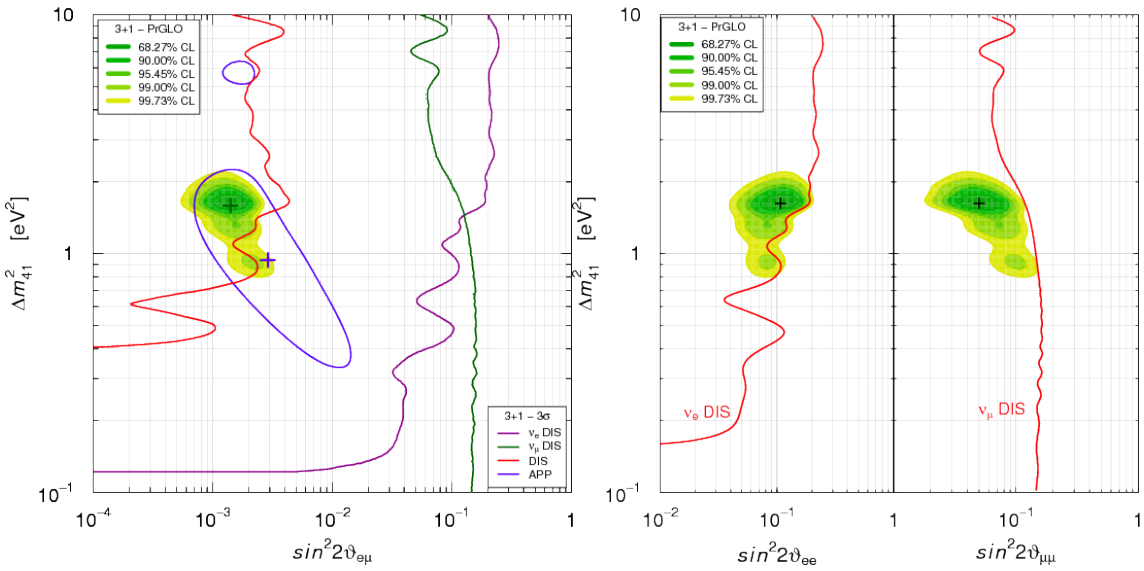


Figure 1.15: Sterile anomalies global fit. Allowed regions in $(\sin^2(2\theta_{\mu e}), \Delta m^2_{41})$, $(\sin^2(2\theta_{ee}), \Delta m^2_{41})$ and $(\sin^2(2\theta_{\mu\mu}), \Delta m^2_{41})$ planes obtained in a 3+1 global fit of neutrino oscillation data sensitive to $\Delta m^2_{41} \approx \text{eV}^2$. These are compared with the 3σ allowed regions obtained from $\overline{\nu}_\mu/\nu_\mu$ short-baseline appearance data (APP) and the 3σ constraints obtained from $\overline{\nu}_e/\nu_e$ short-baseline disappearance data (ν_e DIS), $\overline{\nu}_\mu/\nu_\mu$ short-baseline disappearance data (ν_μ DIS) and the combined short-baseline disappearance data (DIS). The best-fit points of the global and APP fits are indicated by crosses.

The MiniBooNE low-energy anomaly is incompatible with neutrino oscillations, because it would require a tiny value of Δm^2_{41} and a large value of $\sin^2(2\theta_{e\mu})$, which are excluded by other experiments results. It is likely that the explanation for this anomaly lies independently on neutrino oscillations: therefore it has not been included in the global fit. The oscillation parameters best-fit extracted values are:

$$\Delta m^2_{41} = 1.6\,\text{eV}^2 \qquad |U_{e4}| = 0.167 \qquad |U_{\mu 4}| = 0.114 \tag{1.81}$$

implying $\sin^2(2\theta_{e\mu}) = 0.0014$, $\sin^2(2\theta_{ee}) = 0.11$ and $\sin^2(2\theta_{e\mu}) = 0.050$. The combined disappearance constraint in the $(\sin^2(2\theta_{\mu e}), \Delta m^2_{41})$ plane excludes a large part of the region allowed by $\nu_\mu \to \nu_e$ ($\overline{\nu}_\mu \to \overline{\nu}_e$) appearance data, leading to the well-known appearance-disappearance tension.

Let's consider now the MiniBooNE anomalous ν_e excess from the 2018 data release, one of the most recent hints to sterile neutrinos. If interpreted as evidence of a sterile neutrino in a 3+1 mixing

[19] MiniBooNE data from the 2018 release are not included here.

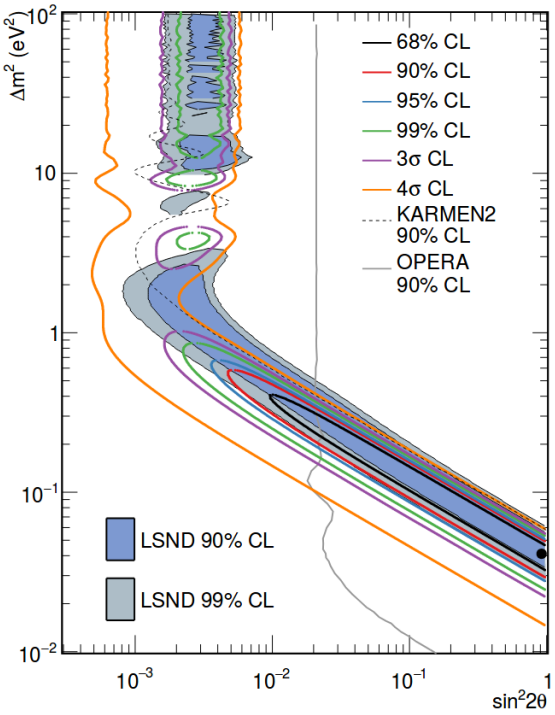


Figure 1.16: MiniBooNE data until 2018. The allowed regions for a combined neutrino mode and antineutrino mode 200 MeV $< E_\nu <$ 3000 MeV within a two-neutrino oscillation model [104] are displayed. The best neutrino oscillation fit in neutrino mode occurs at $\sin^2(2\theta_{41}) = 0.84$, $\Delta m^2_{41} = 0.039\,\text{eV}^2$. The shaded areas show the 90% and 99% C.L. LSND $\overline{\nu}_\mu \to \overline{\nu}_e$ allowed regions. The black point shows the MiniBooNE best fit point. Also the 90% C.L. limits from the KARMEN [106] and OPERA [76] experiments are shown.

scenario, the significance of the combined LSND and MiniBooNE excesses yields a 6.0σ discrepancy with respect to the three flavor SM prediction. Fig. 1.16 shows the MiniBooNE allowed regions for a combined neutrino mode and antineutrino mode 200 MeV $< E_\nu <$ 3000 MeV within a two-neutrino oscillation model. With this technique, the best neutrino oscillation fit in neutrino mode occurs at $\sin^2(2\theta_{41}) = 0.84$, $\Delta m^2_{41} = 0.039\,\text{eV}^2$. This event excess will be extensively studied by the Fermilab short-baseline neutrino (SBN) program [113].

In a nutshell, there is a puzzling tension among data released by several experiments. Given the lack of corresponding ν_μ disappearance observations, required in the case of oscillations involving a sterile flavor, there is strong motivation for alternative explanations of this sterile anomaly [114, 115, 116]. One of these alternatives assumes the existence of a new dark sector containing a dark neutrino and a dark gauge boson, both with masses between a few tens and a few hundreds of MeV [117]. Other possibilities could be quite exotic, as the existence of sterile resonant neutrino oscillations [118], sterile neutrino non-standard interactions [119], extra dimensions for neutrinos [120] and so on.

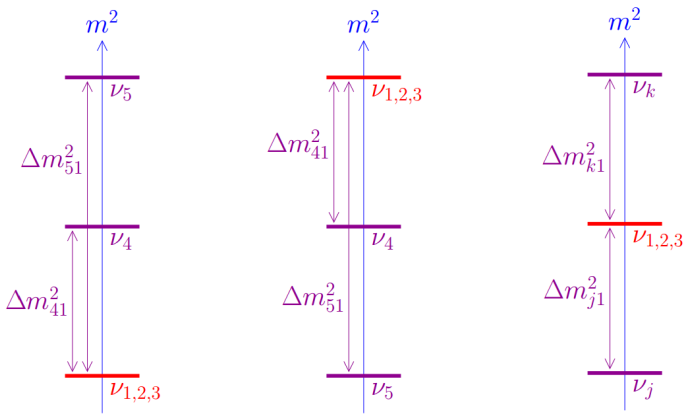


Figure 1.17: Sterile neutrinos: schematic illustration of the three possible 3+2 mass orderings, taking into account that $|\Delta m^2_{21}| \ll |\Delta m^2_{31}| \ll |\Delta m^2_{41}| < |\Delta m^2_{51}|$. In the third scheme the identification of the labels k and j is chosen in order to satisfy the inequality $|\Delta m^2_{41}| < |\Delta m^2_{51}|$.

3+2 model

The 3+2 mixing scheme was considered to be interesting in 2010, when the MiniBooNE neutrino and antineutrino data showed a CP-violating tension [112]. The model introduces a second sterile neutrino in the eV^2 mass region, thus experiments sensitive at the eV^2 scale depend on seven parameters. In addition to the two mass-squared splittings (Δm^2_{41} and Δm^2_{51}) and the magnitude of the mixing matrix elements ($|U_{e4}|$, $|U_{\mu 4}|$, $|U_{e5}|$, $|U_{\mu 5}|$), there is now also a physical complex phase δ CP-violating. The three possible mass hierarchy schemes are depicted in Fig. 1.17. The 2+3 and 1+3+1 scheme, since they involve three or four massive neutrinos at the eV scale, are highly disfavored over the normal scheme by cosmological constraints on the sum of neutrino masses [97].

Considering the normal scheme, the best-fit values for these parameters from global fit to all data sets are [121]:

$$\Delta m^2_{41} = 0.47\,\text{eV}^2 \quad \Delta m^2_{51} = 0.87\,\text{eV}^2 \quad \delta/\pi = 1.64$$

$$|U_{e4}| = 0.128 \quad |U_{\mu 4}| = 0.165 \quad |U_{e5}| = 0.138 \quad |U_{\mu 5}| = 0.148 \tag{1.82}$$

The CL contours in marginalized $(\Delta m^2_{41}, \Delta m^2_{51})$ space can be seen in Fig. 1.18. It can be shown that appearance-disappearance tension in the 3+2 global fit, including low energy MiniBooNE data until 2010, is even worse than that of the 3+1 fit. One can see also that there is little improvement of the goodness of global fit for the 3+2 model with respect to the 3+1 case, in spite of the four additional parameters and of the additional possibility of CP violation [112, 122]. This fact does not justify or support the larger complexity of the 3+2 scheme over the 3+1 one. Furthermore, the 3+2 model is even more disfavoured than the 3+1 one by cosmological constraints [97].

1.9.3 Future perspectives

From the experimental point of view, the hunt for the sterile neutrinos is a very active field of research: more than ten experiments are currently dedicated to this hard task [123]. The

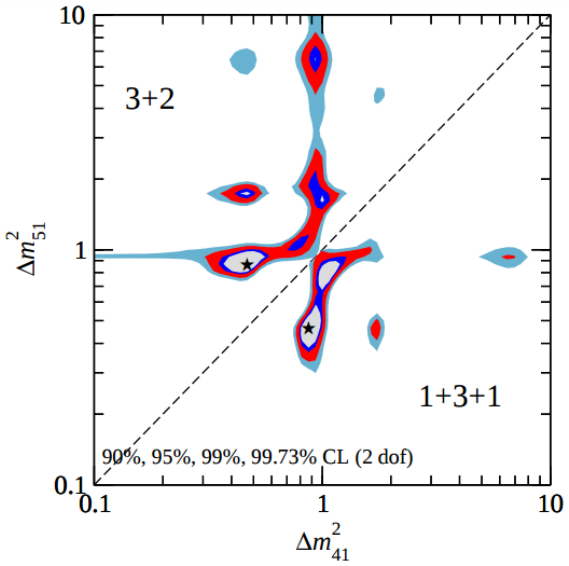


Figure 1.18: Sterile neutrinos: global fit for 3+2 and 1+3+1 scenarios (upper left and lower right respectively) in the the neutrino mass squared differences Δm^2_{41} and Δm^2_{51} plane. The two black indicate the best-fit values.

experimental program has the amibitious goal to solve conclusively this open problem in a few years thanks to a broad range of dedicated experiments employing different techniques and looking in different flavor oscillation channels.

Six very short baseline reactor experiments are taking data just now. They employ a reactor as $\overline{\nu}_e$ source with a detector placed at a distance of the order of 10 m. The DANSS [124, 125] and NEOS [126] experiments, based on neutrino flux coming from industrial reactors, are the most promising ones, benefitting from a high counting rate up to $5000\,\overline{\nu}_e$/day. They show the highest sensitivity at $\Delta m^2_{41} \sim (1-2)\,\mathrm{eV}^2$, reaching regions with $\sin^2(2\theta_{41}) < 0.01$. DANSS is located at the Kalinin nuclear power plant in Russia, and its key advantage is the use of a movable solid scintillator detector, which allows to change the baseline L. The NEOS experiment makes use of a Gd loaded liquid scintillator detector located at a distance 24 m from the core of a reactor in South Korea. Other dedicated reactor experiments which are currently taking data are NEUTRINO-4 [127], PROSPECT [128], SoLid [129] and STEREO [130].

Also the sector of sterile neutrino physics via accelerators is very encouraging. MINOS and MINOS+ [74, 131] look for ν_μ disappearance in two detectors at baselines of 1.04 km and 735 km. No evidence of mixing between active and sterile neutrinos has been observed until now, and a stringent limit on θ_{24} is set for all values of $\Delta m^2_{24} \gtrsim 10^2\,\mathrm{eV}^2$ [131]. The Short Baseline Neutrino Program at Fermilab [113] will consist of three detectors each employing LAr TPC technology, exposed to a booster neutrino beam. Among the three, Icarus is the largest, and sensitivity studies foresee that it will guarantee a 5σ sensitivity to the LSND excess, if it really exists. Icarus has been recently delivered to Fermilab from CERN, and its installation is being completed. The detector closest to the beam source, SBND (Short Baseline Near Detector), is currently under construction. Finally, MicroBooNE is currently operating and taking data using the neutrino beam. The possibility of probing sterile neutrino oscillations using the DUNE experiment Near Detector has already been

promisingly considered [132]. There are also proposals to use a future β-beam or a low-energy neutrino factory to search for short-baseline disappearance [133]. To close the topic, competitive improvements on sterile neutrino mass may come also from the next future generation of the β-spectrum experiments (KATRIN, [134], Troitsk [135]).

Chapter 2

The Borexino experiment: detector design and pp-chain results

Borexino is a large volume liquid detector, whose primary purpose is the real-time measurement of low energy solar neutrino fluxes [39, 61, 65]. It is located deep underground (3800 m mwe) in the Hall C of the Laboratori Nazionali del Gran Sasso (Italy), where the atmospheric muon flux is suppressed by a factor of 10^{-6} with respect to the Earth surface. The Borexino detector started its data taking on 15 May 2007, which it is currently still ongoing.

The original goal of the experiment, i.e. the measurement of solar neutrinos emitted from ^{7}Be reaction, has been completed successfully [62]. The excellent Borexino radiopurity levels made possible to broaden the scientific scope of the experiment: Borexino measured the fluxes of all the pp chain neutrino components [39, 65, 136, 137, 138]. Currently, the collaboration is strongly focusing on the possible detection of CNO neutrinos, the only missing piece of the solar fusion puzzle.

This Chapter is organized as follows. The detector experimental design is outlined in Sec. 2.1. The reactions by which the solar neutrinos and the anti-neutrinos interact in the detector, along with the expected solar neutrino rates according to Standard Solar Model, are shown in Sec. 2.2. The background issues and the low contamination levels achieved in Borexino are critically reviewed in Sec. 2.3. The data acquisition steps, from the electronic chain system to the event offline reconstruction, are summarized in Sec. 2.4; particular emphasis is dedicated to the event position and energy reconstructions algorithms and characterizations, which will be extensively used in the main analysis of this thesis. The Borexino solar neutrino achievements up to 2016 are described in Sec. 2.5. In particular, the comprehensive measurement of solar neutrino rates from the pp chain, published in Ref. [39], will be extensively described in Sec. 2.5.1 and Sec. 2.5.2, together with its implications on solar and particle physics (Sec. 2.5.3). The SOX project for the search of sterile neutrinos, which was planned to use the Borexino detector, is shortly introduced in Sec. 2.6. Conclusions are outlined in Sec. 2.7.

2.1 The detector design

Borexino has been designed to provide a large fiducial volume of ultra-clean scintillator [61]. In the current configuration, it contains $\approx 278\,\mathrm{t}$ of liquid scintillator that convert the energy deposited by neutrino interactions into scintillation photons. The detector is instrumented with 2112 photomultiplier tubes (PMTs) that measure the intensity and the arrival time of this light, allowing the reconstruction of the energy and position of the events.

The Borexino detector is schematically depicted in Fig. 2.1. Its design is driven by the principle of graded shielding: the inner scintillating core is located at the center of a set of concentric spherical layers of increasing radiopurity from outside to inside. This configuration has been established in order to suppress as much as possible the external background due to the surrounding rocks and to the remaining cosmic muons, which represent a serious obstacle to solar neutrino fluxes measurements. In the following, the detector main features and component are listed, from inside to outside.

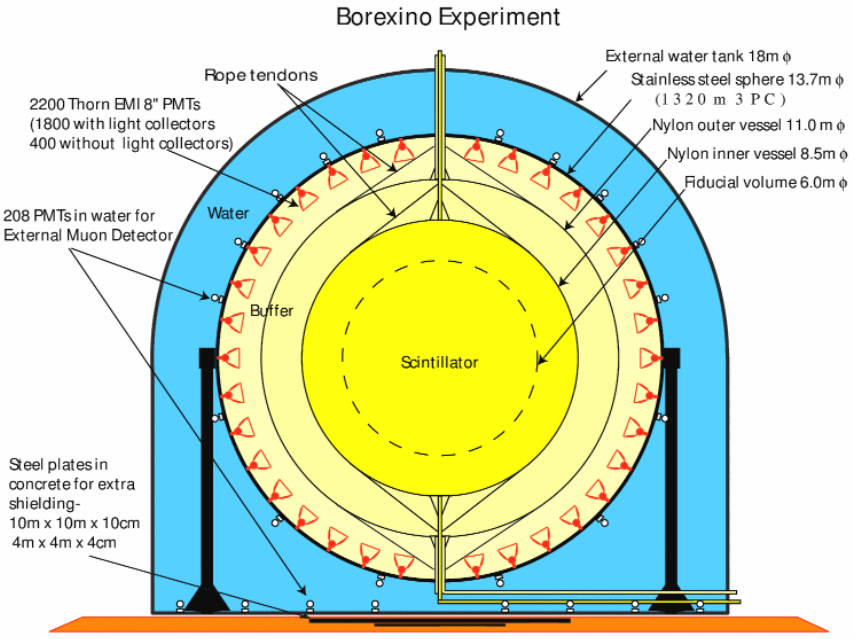


Figure 2.1: Schematic view of the **Borexino detector**.

After careful studies of different organic scintillators, both in small-scale lab tests and in larger scale in the CTF detector, the liquid scintillator was chosen as a mixture of pseudocumene (PC, 1,2,4-trimethylbenzene, $C_6H_3(CH_3)_3$) as solvent and PPO (2,5-diphenyloxazole, $C_{15}H_{11}NO$) as fluor, with a $1.5\,\mathrm{g/l}$ concentration. This scintillator has an electron density of $(3.307 \pm 0.003) \cdot 10^{29}\mathrm{t}^{-1}$ and a mass density of about $0.879\,\mathrm{g/cm^3}$. The light yield is about $\approx 1.2 \cdot 10^4$ photons/MeV [139], while its emission spectrum peaks at 360 nm (see Fig. 2.2). This fact, together with the low absorption of the scintillator and buffer fluid, and to the photoelectric effect probability of photons on PMTs, leads to an overall light collection of ≈ 500 photo-electrons per MeV of deposited energy, corresponding to a 5% energy resolution at 1 MeV. The very high light yield value allows the detection energy

threshold to be set down to a level of a few tens of keV, even if the unavoidable contamination of ^{14}C present in any organic liquid practically limits the *neutrino window* above about 200 keV.

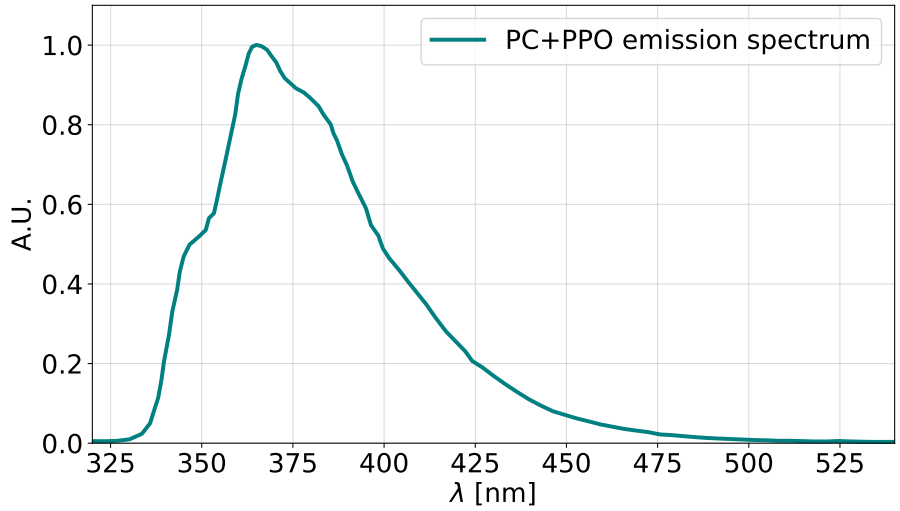


Figure 2.2: Wavelength emission spectrum of the PC + PPO mixture used as liquid scintillator in Borexino.

The scintillator is contained in a transparent spherical nylon membrane of 8.5 m diameter and 125 μm thickness. This envelope is named *inner vessel* (IV) and must be tight as it separates fluids which should never mix, i.e. the scintillator and the buffer liquid (see later on). Nylon has approximately the same refraction index (1.53) of pseudocumene, and thin nylon sheets can be made to be very transparent: in this way, optical processes of scintillation photons crossing this surface are avoided (reflections or refractions).

The inner vessel is surrounded by 1040 t shielding liquid, named *buffer liquid*. Its density is similar to the scintillator one, to balance the hydrostatic pressure and to avoid severe modifications of the vessel spherical shape which would lead, in the worst case, to IV mechanical ruptures. Since the PC density is about 0.88 g/cm^3, water could not be used as liquid buffer: it was decided to make use of pure pseudocumene plus a solute named DMP (dimethylphtalate), an organic light quencher, in a 5 g/l concentration. The main role of the quencher is to suppress the PC scintillation in the buffer, which mainly comes from background events due to PMTs contaminations.

A leak of scintillator from the IV to the buffer region within the OV started approximately on April 9th 2008, for reasons which are still unclear. The leak rate was estimated to be about 1.33 m^3/month, thus the need to halt the fluid rate was urgent. A partial removal of DMP was carried out to reduce the density difference between the scintillator and the buffer fluid, and thus the pressure difference across the leak. Further reduction of the DMP concentration until the current value of 1.5 g/l helped to minimize the leak rate, which then stabilized at $\sim$ 1.5 m^3/year. The lost scintillator volume in the IV was compensated by several refilling operations with the pseudocumene. Since the IV is made of a thin layer on nylon, and therefore is not rigid, the buoyancy forces between the fluids have been slowly developing a deformation of the IV shape, which is now not spherically symmetric nor stable in time. The presence of the leak enhances the deformation evolution in time. Fig. 2.3 shows example of a recent IV shape and the initial

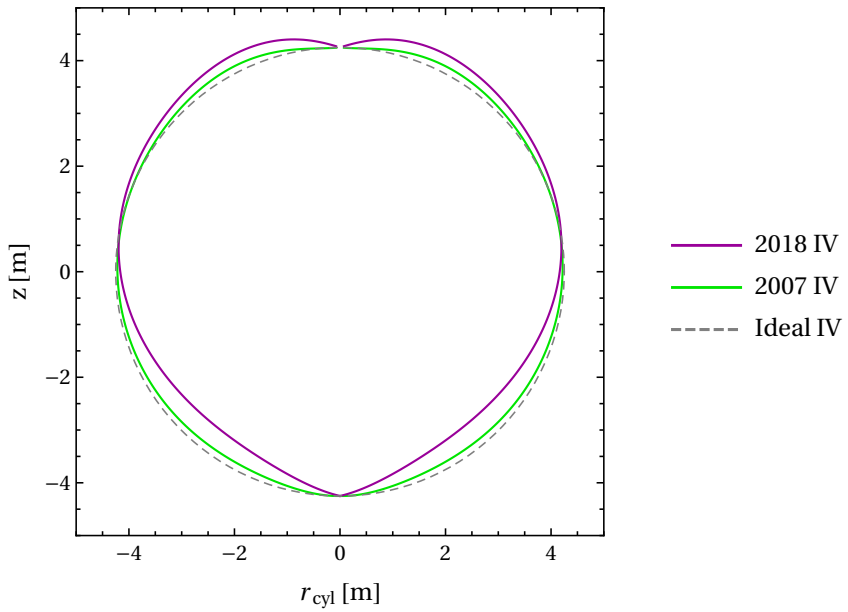


Figure 2.3: Inner Vessel shapes determined by an algorithm developed by the collaboration based on the background contamination present on the Inner Vessel, in 2007 (green) and in 2018 (purple), compared to the ideal spherical one (grey, $r = 4.25\,\text{m}$). The x-axis and the y-axis refer respectively the $r_{\text{cyl}} = \sqrt{x^2 + y^2}$ and the z-coordinates.

one, determined through an algorithm developed by the collaboration based on the background contamination present on the Inner Vessel.

A second nylon membrane, named *outer vessel* (OV), was added in the buffer liquid region to stop the radon atoms diffusing from outer parts of the detector toward the IV. Its diameter is about 12.6 m and its thickness is the same as the one of the IV. Left panel of Fig. 2.4 is a photography of the two nylon vessels installed and inflated with nitrogen.

All the components described so far are included in a stainless steel sphere (SSS) $8 - 10$ mm thick, with a 13.7 m diameter. This sphere contains the buffer liquid and is the support structure for the PMTs. It is held by 20 legs welded to the external water tank floor. A view of the internal surface of the SSS with PMTs mounted inside can be seen in right panel of Fig. 2.4.

Borexino is provided with both an inner and an outer muon veto. The Inner Muon Veto consists of about 400 PMTs without light concentrators (see next subsection). These PMTs are more efficient with respect to those with concentrators in collecting light from the buffer liquid; therefore, they are particularly sensitive to light emitted by muons crossing the buffer. The Outer Muon Veto is composed of 208 additional PMTs installed in the water tank and pointed towards the external, in order to detect the Čerenkov light produced by muons in the shielding water. Water tank inner walls, as well as the outer SSS surface, are covered with some Tyvek sheets, a high reflectivity material that enhances the muon detection probability by maximizing light collection.

The full experimental setup is housed in a dome shaped steel tank, with a 18 m diameter cylindrical base and a hemispherical top of 16.9 m maximum height. It is filled with 2400 t of ultra pure water

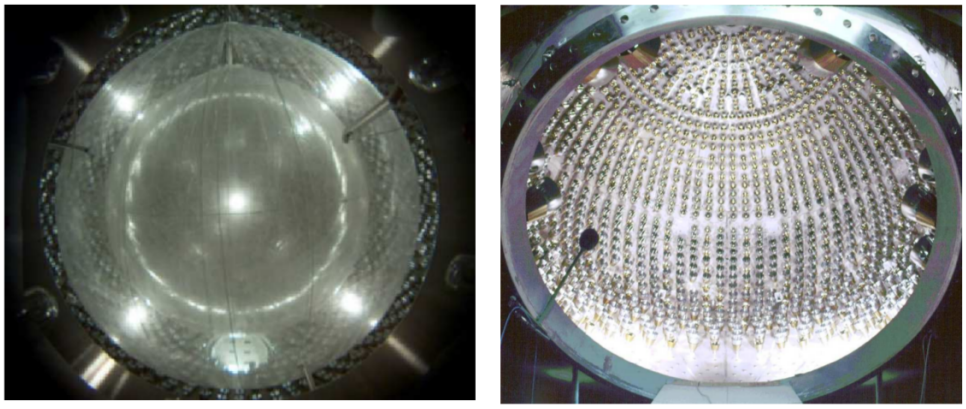


Figure 2.4: Borexino vessels. On the left, the Inner and Outer nylon Vessels (IV and OV) installed and inflated with nitrogen in the stainless steel sphere. On the right, the inner surface of the Stainless Steel Sphere (SSS). The picture is taken from the main SSS door, and shows the internal surface of the sphere with PMTs evenly mounted inside.

as the outermost shielding of the detector from external γ and neutrons.

Photomultiplier tubes

Borexino is instrumented with 2212 8" ETL-9351 PMTs [61]. Employing a special low radioactivity glass, with a complex sealing system, they have been designed to resist the pseudocumene corrosive action on the front and water action on the back. The necessity to measure the energy of each scintillation event and to reconstruct its position by means of a time of flight technique puts severe constraints on the selection of the most suitable photomultiplier model for Borexino. In order to have a large number of collected scintillation photons, high quantum efficiency (QE) values are needed; to achieve a precise time definition, and thus a precise position resolution, a low transit time jitter is required. The selected PMTs are sensitive to light within a wavelength range between 350 nm and 500 nm, and their quantum efficiency peaks at 420 nm to $\approx$ 26%. 1838 PMTs are equipped with aluminum light concentrators (named *cones*) in order to increase light collection capability. The overall geometrical coverage is around 30%. The photomultipliers are installed on the internal surface of the SSS (right panel of Fig. 2.4) in dedicated feed-throughs. PMTs have to satisfy low radioactivity requirements: the photocatode glass is the main radioactivity source, with concentrations equals to $3 \cdot 10^{-8}$g/g for ^{238}U, $1 \cdot 10^{-8}$g/g for ^{232}Th, $2 \cdot 10^{-5}$g/g for ^{40}K.

Monte Carlo simulations showed that the mean number of photoelectrons (*p.e.*) detected by one PMT is found in the $0.02 - 2.0$ range for events with energy between 250 keV and 800 keV. It is therefore important that PMTs provide a good single electron performance both for the signal amplitude and for the timing response. Furthermore, to minimize the probability of random hits during the data acquisition, the chosen PMTs feature a low dark rate. More detailed manufacturer specifications are reported in Ref. [61].

Not all the 2212 PMTs are currently active and operative. During the filling of the SSS with ultra-pure water, about 80 PMTs suddenly stopped working, most likely because of defects in the sealing. The total number of inactive PMTs at the beginning of data taking was 175. The current failure rate, now quite stable in time, is found at about 70 lost PMTs per year. The number of

alive PMTs, at October 2019, is 1127 out of 2212.

2.2 Neutrinos and anti-neutrinos detection in Borexino

Neutrino detection

Solar neutrinos of all flavors are detected by means of their elastic scattering off scintillator electrons (Eq. (1.33)):

$$\nu_x + e^- \to \nu_x + e^- \qquad x = e, \mu, \tau \tag{2.1}$$

Being all scattering angles allowed, the kinetic energy transferred to the electron varies from zero to a large fraction of the neutrino energy. From the relativistic energy and momentum conservation laws, one can relate the transferred energy and the scattering angle θ for any given interaction:

$$E'_\nu = \frac{E_\nu}{1 + \frac{E_\nu}{m_e c^2}(1 - \cos\theta)} \qquad T_e = E_\nu - E'_\nu \implies T_e^{\max} = E_\nu - E_\nu^{'\min} = \frac{E_\nu}{1 + \frac{m_e c^2}{2E_\nu}} \tag{2.2}$$

where $m_e c^2$ is the electron rest energy, E_ν is incoming neutrino energy, and E'_ν is the neutrino energy after the interaction. The kinetic energy of the recoiled electron is $T_e = E_\nu - E'_\nu$, and therefore shows a continuous spectrum even in the case of monoenergetic neutrinos, extending up to a $T_e^{\max}$ maximum value (endpoint). The differential cross section for the $\nu_x - e^-$ scattering, according to the Standard Model prediction at the lowest order in perturbation theory[1], is [140]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}T_e} = \frac{2G_F^2 m_e}{\pi}\left[g_L^2 + g_R^2\left(1 - \frac{T_e}{E_\nu}\right)^2 - g_L g_R \frac{m_e T_e}{E_\nu^2}\right] \qquad \sigma_0 \equiv \frac{2G_F^2 m_e}{\pi} \tag{2.3}$$

$$g_L = \begin{cases} +\frac{1}{2} + \sin^2\theta_W & \text{for } \nu_e \\ -\frac{1}{2} + \sin^2\theta_W & \text{for } \nu_\mu, \nu_\tau \end{cases} \qquad g_R = \sin^2\theta_W \qquad \sin^2\theta_W \approx 0.231 \tag{2.4}$$

where $\theta_W = \arccos(M_W/M_Z)$ is the Weinberg angle, g_L and g_R are respectively the left and right chirality electroweak couplings. The leading constant σ_0, which gauges the cross section magnitude, is $\sigma_0 \approx 8.80 \cdot 10^{-46}\,\mathrm{cm}^2$. The total cross section can be obtained integrating Eq. (2.3) from $T_e = 0$ to $T_e = T_e^{\max}$.

Borexino can detect neutrinos of all flavors, but it is worth to underline that for the elastic scattering $\sigma_{\nu-\mu} = \sigma_{\nu-\tau} \approx 0.16\sigma_{\nu-e}$ (see Sec. 1.6). Electron recoils induced by different neutrino flavors cannot be distinguished on an event-by-event basis [2]. Furthermore, the neutrino induced events are intrinsically indistinguishable from β and γ backgrounds: the several signal and background contributions can be disentangled only by means of a spectral fit (see Sec. 2.5.1 or the more detailed Sec. 4.2).

[1] Radiative corrections are neglected in Eq. (2.3) for the sake of simplicity, but they are accounted for in the Borexino full calculations of ν expected interaction rates. They impact at the $\approx 1\%$ level on the total cross section budget, which is a relevant precision level for the Borexino physics goals.

[2] In principle, the slight differences of recoil energy spectra between ν_e and ν_μ, ν_τ might allow a statistical separation, but this is practically not feasible with a realistic amount of data [65]

Solar ν		E_ν [keV]	$T_e^{\max}$ [keV]	R_ν B16 (GS98) HZ [cpd/100t]	R_ν B16(AGS09) LZ [cpd/100t]	Main bkg
pp chain	pp	≤ 420	261	131.1 ± 1.4	132.2 ± 1.4	^{14}C
	^{7}Be	384 862	231 662	47.9 ± 2.8	43.7 ± 2.5	^{85}Kr, ^{210}Bi
	pep	1440	1220	2.74 ± 0.04	2.78 ± 0.04	^{11}C, ^{210}Bi
	^{8}B	$\leq 1.50 \cdot 10^4$	14500	0.211 ± 0.003	0.174 ± 0.002	^{208}Tl, ext. γ
CNO cycle	^{13}N	≤ 1199	988			
	^{15}O	≤ 1732	1509	4.92 ± 0.78	3.52 ± 0.52	^{11}C,^{210}Bi
	^{17}F	≤ 1740	1517			

Table 2.1: Expected neutrino interaction rates in Borexino according to the B16 Standard Solar Model predictions, in the high or low metallicity scenarios (HZ and LZ) and assuming MSW+LMA flavor transformation [25]. E_ν is the neutrino energy (end-point for continous energy spectra) and $T_e^{\max}$ is the maximal energy of the scattered e^- according to Eq. (2.2). The expected interaction rates in Borexino, reported respectively in the third and fourth columns, are calculated for the high-metallicity scenario (B16 GS98) and for the low-metallicity scenario (B16 AGSS09). Note that the CNO-ν rate is quoted as a single component (given by the sum of ^{13}N-ν, ^{15}O-ν and ^{17}F-ν) since the current experimental sensitivity does not allow to disentangle the three components. The reported R_ν errors include also the uncertainty on the flavor oscillation probability. The fifth column indicates the main backgrounds present in Borexino in the same energy region of interest of a given neutrino species.

Accounting for the effects of the neutrino oscillations, which partly convert the ν_e emitted by the Sun into ν_μ and ν_τ, the expected solar neutrino interaction rate in Borexino R_ν is given by:

$$R_\nu = N_e \int \mathrm{d}E_\nu \frac{\mathrm{d}\Phi}{\mathrm{d}E_\nu} \int \mathrm{d}T \left[\frac{\mathrm{d}\sigma_e}{\mathrm{d}T} P_{ee}(E_\nu) + \frac{\mathrm{d}\sigma_\mu}{\mathrm{d}T} (1 - P_{ee}(E_\nu)) \right] \tag{2.5}$$

where N_e is the number of target electrons, $\mathrm{d}\Phi_\nu/\mathrm{d}E_\nu$ is the SSM differential energy spectrum for solar neutrinos (as discussed in Sec. 1.5) and $P_{\mathrm{ee}}(E_\nu)$ is the solar ν_e survival probability. The expected interaction rates of solar neutrinos in Borexino according to the high-metallicity and the low-metallicity scenarios, using the oscillation parameters from Ref. [86], along with the related energies of the reactions, are reported in Tab. 2.1, separately for the pp chain and for the CNO cycle.

The solar neutrino count rate in the Borexino range is found between a few and a few tens of counts per day for one hundred tons of target material (*cpd/100t* from now on), and defines the required background rates and the needed radio-purity levels of the detector. These low rates, corresponding to an equivalent activity of 10^{-9} Bq/kg, drive the radiopurity requirement for Borexino: since the typical natural radioactivity is in the order of $10^1 - 10^2$ Bq/kg, the Borexino scintillator needs to be 9-10 order of magnitude less radioactive than anything on Earth.

Anti-neutrino detection

Borexino detects $\overline{\nu}_e$ via the inverse β reaction (IBD) [12],

$$\overline{\nu}_e + p \rightarrow e^+ + n \tag{2.6}$$

which features a $\overline{\nu}_e$ energy threshold of $E_{\rm thr} = m_e + m_n - m_p \approx 1.806\,\text{MeV}$. The generated e^+ promptly comes to rest in the liquid scintillator, annihilates with a scintillator electron, inducing two 511 keV γ-rays; these yield a prompt event with a visible energy of $E_{\rm vis} = T_{e^+} + 2m_e = E_{\overline{\nu}_e} + m_p - m_n + m_e = E_{\overline{\nu}_e} - 0.782\,\text{MeV}$. Instead, the free neutron emitted, after a free mean path of the order of cm or tens of cm, thermalizes and is typically captured by PC protons[3], in a mean time of 256 μs. This capture is followed by the emission of a 2.22 MeV γ-ray, providing a coincident delayed event[4]. The characteristic time, energy and spatial coincidence of prompt and delayed events offers a clean and unmistakable signature of $\overline{\nu}_e$ detection. The IBD total cross section ($\sigma_{\rm IBD}$) can be inferred from the Standard Model predictions:

$$\sigma_{\rm IBD}(E) = \frac{\sigma_0}{m_e^2}(E - m_n + m_p)\sqrt{(E - m_n + m_p)^2 - m_e^2} \qquad \sigma_0 = 9.455 \cdot 10^{-48}\,\text{m}^2 \tag{2.7}$$

2.3 Backgrounds

The achievement of extremely low background levels in Borexino played a crucial role for the experiment success. According to the Standard Solar Model, the order of magnitude of sub-MeV solar neutrino interactions in Borexino is a few tens cpd/100t (see Tab. 2.1), corresponding to an equivalent activity of a few 10^{-9} Bq/kg from natural radioactivity of materials. Moreover, the neutrino induced events are intrinsically indistinguishable from β and γ radioactivity: this requires very restrictive requirements in terms of radiopurity of the scintillator and of the detector construction materials.

The Borexino collaboration has been addressed the radiopurity issue in several ways: developing suitable purification techniques for scintillator, water, and nitrogen, performing careful material selections, developing innovative cleaning techniques for metal surfaces. In particular CTF, the Borexino protoype, has played a crucial role in the long R&D phase: it was built with the task of demonstrating the feasibility of large scale purification of organic liquid scintillators to the levels required for the ^{7}Be-ν detection. Following this experience, the basic strategy employed for Borexino has been to purify the scintillator with a combination of distillation, water extraction, and nitrogen gas stripping procedures. All these techniques are employed in common industrial processes, but they have been adapted and optimized for Borexino, bringing to radiopurity levels well beyond the original design goals.

In spite of the purification campaign success, still relevant background contaminations are present in the Borexino detector, and represent the main obstacle to the solar neutrino flux determinations. Events generated by the intrinsic contaminants are referred to as internal background, while those reaching the scintillator from outside are referred to as external background.

[3] The scintillation light related to the proton recoil is highly quenched and negligible.
[4] In $\approx 1\%$ of cases, the capture takes place on a PC carbon atom, resulting in a 4.9 MeV de-excitation γ.

2.3.1 Internal backgrounds

In the following, we review the internal backgrounds present in Borexino. The contamination level will be indicated in g/g units, i.e. parts of contaminant isotope per unit.

^{14}C

The ^{14}C isotope (β^- decay, $\tau = 8270\,\mathrm{y}$, $Q = 156\,\mathrm{keV}$) is by far the most abundant background in Borexino, and dominates the very low energy range. Its extremely high average rate is about $3.5 \cdot 10^6$ cpd/100t, corresponding to $\approx 40\,\mathrm{Bq/100t}$.

At the Earth surface level, it is cosmogenically produced via the reaction $^{14}\mathrm{N} + n \rightarrow p + {}^{14}\mathrm{C}$, leading to a relative abundance of $1.2 \cdot 10^{-12}$ g/g in living organic matter. ^{14}C is chemically identical to ^{12}C: once it is produced, it cannot be separated by any purification process. Therefore, it unavoidably accompanies the ^{12}C isotope, with concentrations depending on the site of production and storage.

Fortunately, since organic liquids are derived from oil, which remained underground for millions of years, the concentration of ^{14}C in the liquid scintillators from old oil reservoirs is found at the 10^{-18} g/g level. For the solar neutrino low energy window, the acceptable ^{14}C relative abundance is $3 \cdot 10^{-18}$ g/g, since a larger concentration would spoil the neutrino signal through a high rate of pile-up signal. The current Borexino ^{14}C contamination, measured analyzing the lowest energy part of the events spectrum, is $2 \cdot 10^{-18}$ g/g.

^{238}U and ^{232}Th chains

A large contribution of the internal background is given by radio-isotopes belonging to the natural ^{238}U (see scheme in Fig. 2.5) and ^{232}Th chains. The current contamination level, of about 10^{-17} g/g for ^{238}U and 10^{-18} g/g for ^{232}Th, has been achieved through the distillation and is about two order of magnitude better than original design specifications.

Unfortunately, some of the isotopes of the radioactive ^{238}U and ^{232}Th chains are found out of the equilibrium. The first element is the ^{210}Pb, which can be absorbed by metal and plastic surfaces, and gives rise to the following sub-chain:

$$^{210}\mathrm{Pb} \xrightarrow[Q_\beta = 63.5\,\mathrm{keV}]{\tau \approx 32\,\mathrm{y}} {}^{210}\mathrm{Bi} \xrightarrow[Q_\beta = 1160\,\mathrm{keV}]{\tau \approx 7.23\,\mathrm{d}} {}^{210}\mathrm{Po} \xrightarrow[Q_\alpha = 5.3\,\mathrm{MeV}]{\tau \approx 199.1\,\mathrm{d}} {}^{206}\mathrm{Pb} \tag{2.8}$$

^{210}Pb cannot be detected in Borexino, because of its very low end-point ($Q = 63.5\,\mathrm{keV}$). Its daughters instead, ^{210}Bi and ^{210}Po, are a major source of background in Borexino. ^{210}Bi spectrum spans the energy range of interest for ^{7}Be-ν, pep-ν, CNO-ν, showing a smooth featureless profile.

^{210}Po is a 5.3 MeV α emitter. Because of the high quenching effects (see Sec. 2.4.1), its decay reconstructed energy is found around 500 keV, and therefore its spectrum extends in the typical ^{7}Be-ν energy region. Clearly visible also by eye in the energy Borexino spectrum, because of the evident gaussian spectrum shape, its rate can be easily isolated thanks to α-β discrimination techniques. Beside the contribution coming from the ^{210}Pb chain, the ^{210}Po itself is found outside of secular equilibrium at the beginning of the experiment and after the purification stages. The reason is probably due to a ^{210}Po contamination of the purification pipes, which have been partly washed out by the liquid flow movement thus entering in the scintillator. The link between ^{210}Bi

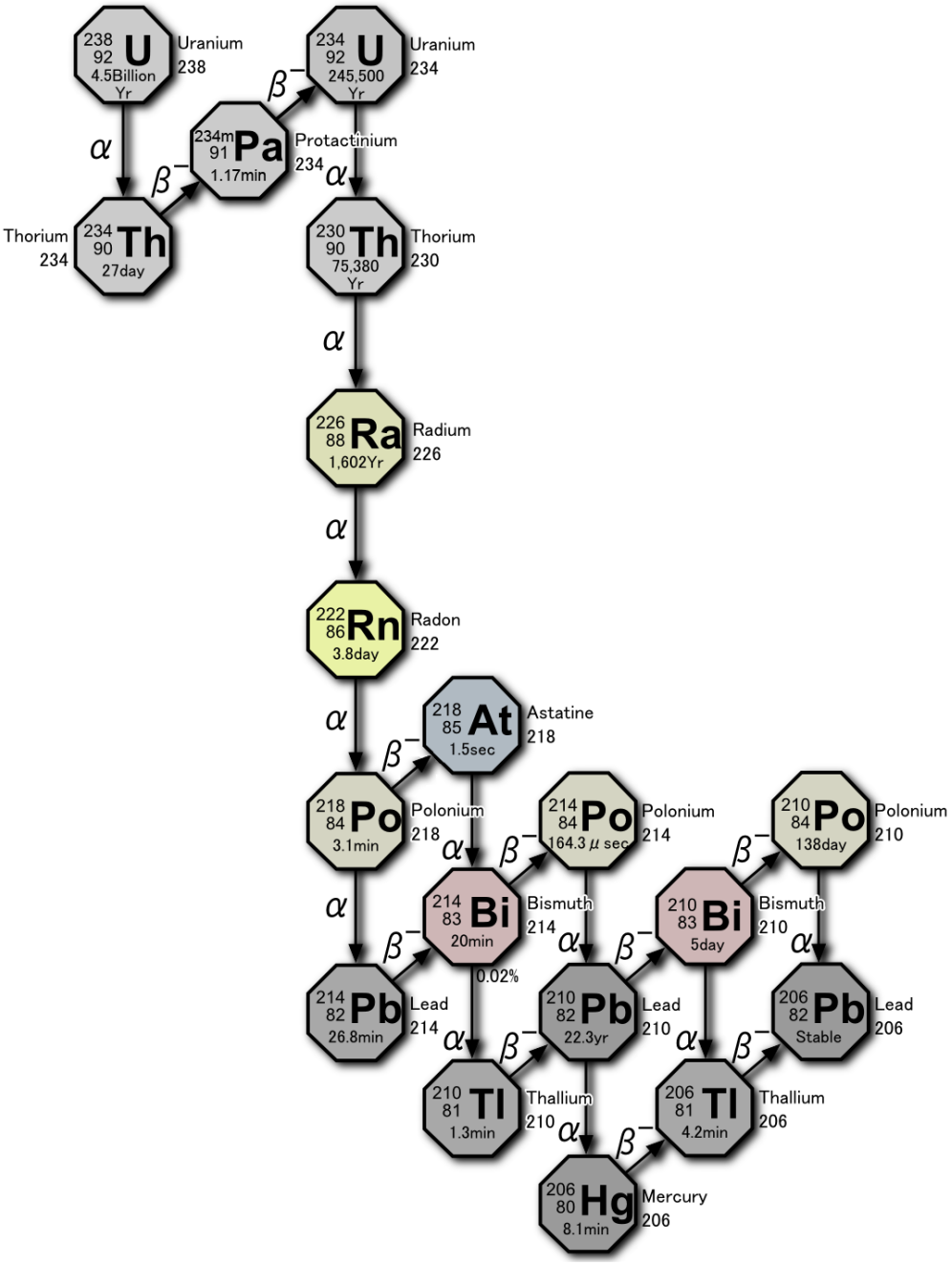


Figure 2.5: Scheme of the natural decay chain of ^{238}U.

and ^{210}Po, their spatial distribution and their evolution in time, play a crucial role in the CNO-ν detection analysis, and will be described extensively in Chapter 6.

^{222}Rn

The ^{222}Rn isotope, member of the ^{238}U chain, is a chemically inert gas that diffuses easily in air and in materials. It is formed within the ^{238}U radioactive chain, and decays α with a half life of $\tau_{\mathrm{Rn}} = 3.8\,\mathrm{d}$ converting into ^{218}Pb. The ^{222}Rn air activity in Gran Sasso Hall C is kept acceptable thanks to a continuous venting of the halls with fresh outside air piped in through the highway tunnel, reaching an average concentration $\sim 50\,\mathrm{Bq/m^3}$ [141]. The air circulation is fundamental to keep safe working conditions. ^{222}Rn contamination is one of the reasons for which the detector must remain leak tight to air; it is especially dangerous for Borexino because it is more soluble in pseudocumene scintillator than in air. During the initial filling of the detector, a small leak in a valve has caused a ^{222}Rn spike, which decayed away after few months since the beginning of data-taking.

^{40}K

The ^{40}K isotope (β^- decay or EC, $\tau = 1.85 \cdot 10^9\,\mathrm{y}$) is a primordial nuclide naturally present at the 0.012% level with respect to natural potassium, decaying via two different channels:

$$^{40}\mathrm{K} \rightarrow {}^{40}\mathrm{Ca} \qquad \beta^- \quad Q = 1.32\,\mathrm{MeV}, \mathrm{B.R.} = 89.3\% \tag{2.9}$$

$$^{40}\mathrm{K} \rightarrow {}^{40}\mathrm{Ar} + \gamma \qquad \mathrm{EC} \quad E_\gamma = 1.46\,\mathrm{MeV}, \mathrm{B.R.} = 10.7\% \tag{2.10}$$

It is present into the scintillator either through micron dust particulates or as a PPO contamination, at the level of parts per million. Currently, ^{40}K concentration in the detector is below the maximum concentration allowed ($10^{-14}\,\mathrm{g/g}$).

^{85}Kr

The isotope ^{85}Kr (mainly β^- decay, $Q = 687\,\mathrm{keV}$, $\tau = 15.4\,\mathrm{y}$) is present in the air mostly because of human-caused nuclear explosions in atmosphere, at the average concentration of about $1\,\mathrm{Bq/m^3}$. Its energy spectrum lies in the same region of ^{7}Be-ν, showing a smooth featureless profile. Since even extremely small air exposures during the detector-filling operations would yield a significant contamination, avoiding air exposure of fluids is extremely important in order to limit the ^{85}Kr background level.

With very low probability (B.R. = 0.43%) ^{85}Kr can also decay into the metastable $^{85\mathrm{m}}$Rb. This state generates a coincidence of the prompt electron ($Q = 173\,\mathrm{keV}$) and the following de-excitation to its fundamental state through a 514 keV γ, with a mean time of $\tau \approx 2\,\mu\mathrm{s}$. This chain provides a method to get an estimation of the ^{85}Kr abundance present in the scintillator, which during the Phase-I (2007-2010) was found to be $30.4 \pm 5.3\,(\mathrm{stat.}) \pm 1.5\,(\mathrm{syst.})\,\mathrm{cpd/100ton}$. After the purifications, i.e. from the Phase-II data taking stage, there is no evidence of ^{85}Kr contamination by using the coincidence method; an upper limit for ^{85}Kr interaction rate 4.9 cpd/100t at 95% C.L. is set. Analyzing the Phase-III-TMP dataset (July 2016 - June 2019), the analysis returns instead an higher upper limit for ^{85}Kr rate of 10.9 cpd/100t at 95% C.L.

^{39}Ar

The isotope ^{39}Ar (β^- decay, $Q = 565\,\text{keV}$, $\tau = 269\,\text{y}$) is a cosmogenic pure β^- emitter and its energy endpoint is close to the ^{7}Be-ν shoulder. Therefore, great care was taken in ensuring that argon levels in the specially prepared nitrogen used for the stripping of the scintillator was as low as possible. These operations translate in an expected ^{39}Ar upper limit for the interaction rate of $0.02\,\text{cpd}/100\text{t}$ at 95% C.L., which is therefore negligible and not considered in all the solar neutrino analysis.

2.3.2 External background

The main source of external background is the radioactivity of the materials that surround the scintillator: the PMTs and the light cones cause the major contaminations, due to the ^{208}Tl, ^{214}Bi and ^{40}K isotopes, which are $\beta\gamma$ decaying. Only the emitted γ have a sufficiently high mean free path to reach the inner part of the scintillator; the typical energy region is found in the $1 - 3\,\text{MeV}$ range. The determination of the external background rate is crucial for the solar neutrino analysis, and its reduction is the main motivation for the definition of a Fiducial Volume (see Sec. 4.1.2).

2.3.3 Cosmogenic background

Muon and muon-induced background

Even if the underground location of the Gran Sasso laboratory strongly suppresses the muon flux thanks to the rock shielding, not all the high energy muons are stopped. The overall muon flux is reduced, with respect of the sea level, by a factor of 10^6. About 4500 μ/day (equivalently, $1.2\,\mu\,\text{m}^{-2}\text{h}^{-1}$) enter into the the inner part of the detector, and have to be efficiently identified along with their induced background. Along their trajectory through the detector, indeed, crossing muons can generate unstable nuclei by direct interactions (e.g. spallation processes) or through the generation of hadronic showers. The most important cosmogenic isotopes affecting the Borexino analysis are listed in Tab. 2.2, together with their main characteristics.

Besides the external muon veto, several software criteria are applied to tag and remove muons, most of which are based on pulse shape techniques. The overall muon veto identification efficiency is found to be 99.992%. Details about the muon tagging can be found at Ref. [142]. The data taking for the whole detector is vetoed for 300 ms after a muon is tagged. This allows to suppress drastically the interaction rates of the many short-lived cosmogenic isotopes, i.e. the ones with $\tau < 300\,\text{ms}$, as shown in the last column of Fig. 2.2. The only three not negligible isotopes for what concerns the low energy solar neutrino analysis are ^{11}C, ^{10}C and ^{6}He.

Muon events are classified into three different scenarios, according to the detector regions they cross:

- **Muons crossing the IV** (IVμ) deposit hundreds of MeV of ionization energy into the scintillator and therefore release a huge amount of scintillation light.
- **Muons crossing the buffer liquid without entering the IV** (Bμ) represent the most concerning case for the solar neutrino analysis. The generated light, partly due to Čerenkov process and partly to buffer scintillation, is low enough to mimic a solar neutrino signal.

Isotope	Mean life	Energy [MeV]	Decay mode	Residual rate after 300 ms μ veto [cpd/100t]
n	255 μs	2.23	capture γ on ^{1}H	< 0.005
^{12}N	15.9 ms	17.30	β^+	$< 5 \cdot 10^{-5}$
^{13}B	25.0 ms	13.40	$\beta^-\gamma$	$< 5 \cdot 10^{-5}$
^{12}B	29.1 ms	13.40	β^-	$(7.1 \pm 0.2) \cdot 10^{-5}$
^{8}He	171.7 ms	10.70	$\beta^-\gamma$ n	0.004 ± 0.002
^{9}C	182.5 ms	16.50	β^+	0.020 ± 0.006
^{9}Li	257.2 ms	13.60	$\beta^-\gamma$ n	0.022 ± 0.002
^{8}B	1.11 s	18.00	$\beta^+\alpha$	0.21 ± 0.05
^{6}He	1.16 s	3.51	β^-	0.31 ± 0.04
^{8}Li	1.21 s	16.00	$\beta^-\alpha$	0.31 ± 0.05
^{11}Be	19.9 s	11.50	β^-	0.034 ± 0.006
^{10}C	27.8 s	36.50	$\beta^+\gamma$	0.54 ± 0.04
^{11}C	29.4 min	0.96	β^+	~ 28
^{7}Be	76.9 d	0.48	EC γ	0.36 ± 0.05

Table 2.2: Cosmogenic isotopes in Borexino: for each isotope (first column) we report its mean life τ (second column), its energy (third column) and its decay mode (fourth column). In case of β decays, the Q value is reported in the third column. The last column shows the expected residual rate after the 300 ms time veto applied after a muon passing through ID is identified [65]; the isotopes returning a non negligible contribution to the Borexino low energy spectrum are highlighted (^{11}C, ^{10}C, ^{6}He.

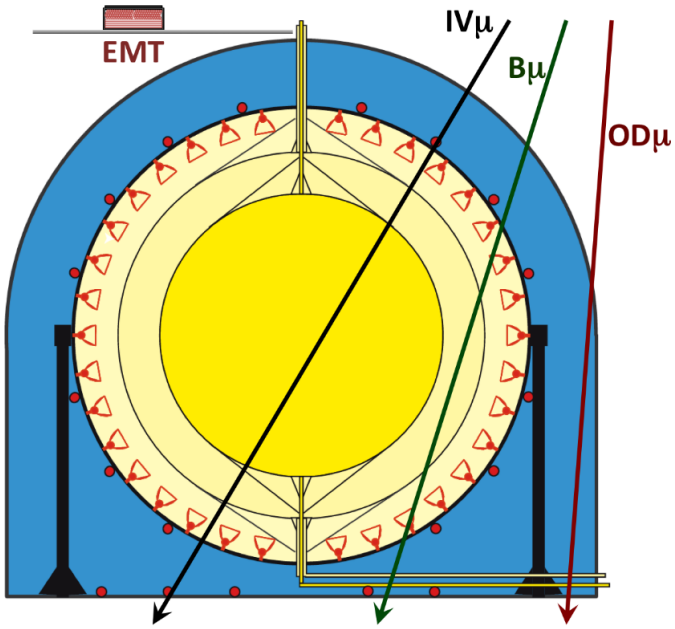


Figure 2.6: Scheme of the three **muon crossing scenarios in the Borexino detector**: crossing the IV (IVμ), crossing the buffer liquid without entering the IV (Bμ), crossing the OD only (ODμ's).

- **Muons crossing the OD only** (ODμ) are tagged by means of a special trigger identification flag, indicating that no light is present in the ID.

In the neutrino analysis, the most concerning muon-induced background is by far the ^{11}C cosmogenic isotope, whose long lifetime is $\tau_{^{11}\mathrm{C}} = 29.4\,\mathrm{min}$. It is generated by the interaction of a cosmic muon with a ^{12}C atom:

$$\mu + {}^{12}\mathrm{C} \rightarrow {}^{11}\mathrm{C} + n \tag{2.11}$$

$$^{11}\mathrm{C} \rightarrow {}^{11}\mathrm{B} + e^{+} + \nu_e \tag{2.12}$$

The total energy released in the detector is found between 1020 keV and 1980 keV, made by the sum of the β^+ decay ($0 \leq E \leq 960\,\mathrm{keV} = Q$) and of the 1020 keV from γs coming from the e^+e^- annihilation. The associated energy spectrum lies in the energy region of interest for the detection of electron recoils from pep and CNO neutrinos. Its long mean decay time makes the short 300 ms veto useless for its reduction.

As for the other cosmogenic isotopes, ^{11}C is continuously produced by the cosmic muons, therefore purifications cannot help to reduce it in the scintillator. Only a dedicated ^{11}C tagging method (TFC), based on its spatial and time coincidence with muons and captured neutrons, allows to limit this background significantly. The TFC algorithm has been developed by the Borexino collaboration in order to identify the ^{11}C background, and will be reviewed in Sec. 4.1.3.

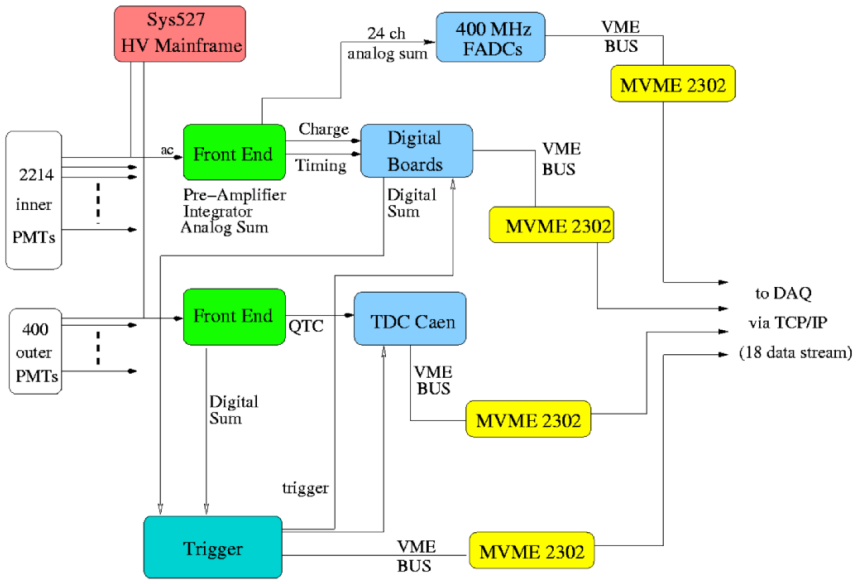


Figure 2.7: Block diagram of the **Borexino electronics layout**.

Neutron background

High-energy neutrons generated by muons in the interactions with the surrounding rock or with the detector material are detected by a combined identification of the outer and inner detector. The 300 ms veto from the muon passage allows to strongly suppress the neutron contribution at the $< 5 \cdot 10^{-3}$ cpd/100t level (see Tab. 2.2).

The rock walls at LNGS generate a neutron flux of about $(1.08 \pm 0.02) \cdot 10^{-6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for the thermal component and $(1.98 \pm 0.05) \cdot 10^{-6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for the epithermal one[5]. The thermal neutrons are mainly produced from fission and (α, n) reactions and the typical energy is about 0.025 eV. A strong reduction of flux inside the detector, of a factor $\sim 10^{-8}$, is provided by the water shielding; in case of fast neutrons, they are firstly moderated and then, when the associated cross sections increases, captured.

2.4 Data acquisition

The Inner Detector (ID), that is the Borexino scintillator surrounded by the 2212 PMTs [61] and the Outer Detector (OD), that is the water Čerenkov muon veto [142], have two independent data acquisition systems. As discussed in Sec. 2.1, the OD plays an important role for the muon and external background rejections: it fires when at least six OD PMTs detect light within a time window of 150 ns.

The ID is composed of the 2212 PMT installed on the Stainless Steel Sphere, pointed inward: it collects information from the events taking place within the IV volume. In particular, it detects the interaction time and the charge of the events, which are fundamental for the events position and energy reconstructions respectively. In this section we focus on the ID only, which plays the major

[5] Neutrons with energy in the $E_n = (0.5 - 1) \cdot 10^3$ eV range.

role for the neutrino detection. The collected signal is processed by the electronics chains of the Borexino data acquisition systems. The Borexino electronic signal processing scheme, designed to achieve the required timing properties, is shown in Fig. 2.7. The ID electronics provides amplitude and time-to-digital conversion for each front-end channel, and is organized in 14 identical racks, each of them handling 160 PMTs. The racks provide the high voltage, process the signal through the front-end boards and record the related information with backplane boards acting as an interface to the trigger system (Laben boards).

The pulse signal due to a single photoelectron at the front-end input shows an amplitude of about $12 - 15\,\mathrm{mV}$ and a total width of 15 ns. The number of photoelectrons collected by the Borexino detector is approximately 500 per each MeV of released energy. PMTs operate mainly in a single photoelectron regime; the multiple hit probability for a 1 MeV event in the detector center is $\approx 10\%$. This probability significantly increases getting closer to the IV and for higher energy events, due to geometric acceptance effects.

The signal from PMTs is AC coupled to the front-end board that performs noise filtering, pre-amplification, shaping and integration of the input signal: it guarantees both a linear response used for time measurement and a voltage signal proportional to the total charge. Each front-end board performs also an analog sum of twelve linear output signals; moreover, a flash ADC system extends the dynamic range of the system to 3 MeV. When data reach the front-end board, they are temporarily stored in an internal memory. If a trigger signal arrives before $6\,\mu$s, data are kept and copied to a specifically designed VME readable card, which provides the analog to digital conversion of the charge signal. On the contrary, they are lost when the internal memory is overwritten.

The main requirement for the Borexino triggering system is to be able to identify, quickly and efficiently, the scintillation events over the spurious dark noise [61]. To activate the ID trigger, the emitted scintillation photons have to fire at least 25 PMTs hits within a 99 ns time window (*trigger threshold* in the Trigger Time Window). Indeed, the TTW must be set to be larger than the maximum arrival time spread of photons at the PMTs, which is at most about 50 ns [6] [61]. The hits threshold set corresponds to approximately 50 keV of released energy. The trigger rate, of about 30 Hz is mainly driven by the ^{14}C contamination, which gives by far the largest contribution. The time and charge information collected by the main acquisition system for each fired PMT defines the *hits*, which are acquired within a $16.5\,\mu$s gate.

The data taking is organized in runs of about six hours each. Runs with more than twenty thousands events must be validated by the shifter, who is the responsible supervisor of the data taking for the specific run. The validation procedure allows to guarantee the good quality of the just acquired data: the shifter has to check information about the number of events detected, the stability plots, the ID/OD behaviour and the laser and pulser triggers. Once the run is validated, a raw data file is produced and uploaded on the Borexino INFN-CNAF storage area.

2.4.1 Data offline reconstruction

The offline data processing is carried out by Echidna, a ROOT based C++ software code developed by the Borexino collaboration. Based on the charge and the arrival times of the detected PMT hits, Echidna reconstructs the basic physics observables of an event, such as the energy and the position of the interaction vertex. The Echidna output is saved in a ROOT file, which includes

[6] From the hardware side, the trigger board system allows the TTW to be set from a minimum of 48 ns up to a maximum value of 99 ns

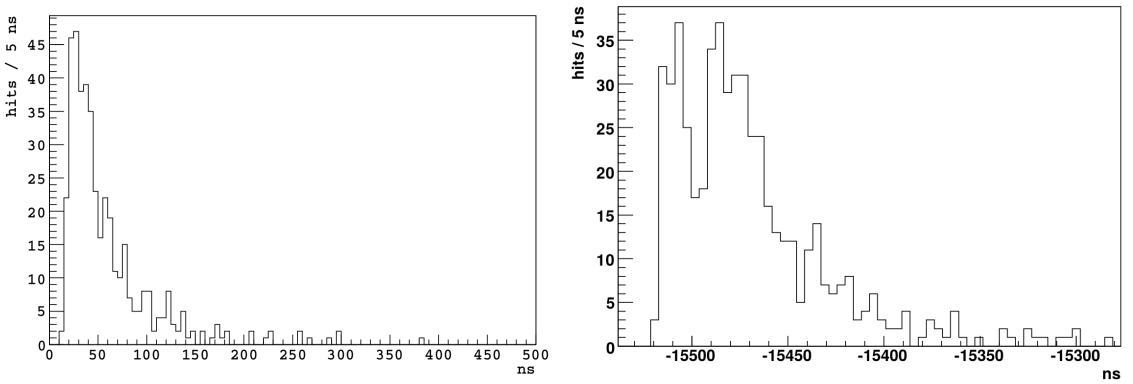


Figure 2.8: Borexino clustering. The time distribution of the PMT hits in a single cluster event is shown in left panel. The time distribution of the PMT hits for a two-cluster event is shown in right panel: the clusters are partially overlapped, but are disentangled by the clustering algorithm. In this case, their time distance is about 25 ns. Clusters as close as 15 ns can be efficiently identified [61].

the raw and reconstructed information in a customized ROOT Tree (named BxTree). The data analyses presented in this thesis are completely based on the Echidna framework. All the runs of each week are also organized in pre-filtered structures (DSTs), where a basic selection is performed and service triggers are removed. In the following, the main features and performance of the data offline reconstruction will be described.

Clustering

The *clustering* is the offline software procedure applied to identify the cluster, i.e. a group of photons belonging to the same scintillator event [61]. The clustering conditions are shortly summarized. In order to identify a cluster, once the decoded hit time distribution has been binned in a 16 bins distribution, two conditions need to be satisfied. At least one time bin has to overcome the average dark noise at a 3σ level, and at least 20 decoded hits (~ 50 keV) have to be located in the first 48 ns. Clusters usually last between 100 ns and 1 μs.

The largest fraction of events in Borexino show *one cluster*. A 0-*cluster event* occurs if the trigger fires but no clusters are identified: this means that no physical events have taken place. In case more than one cluster is recognized in the trigger time window, we define a *multi-cluster event*. *Pile-up events* are the result of two uncorrelated events that fall in the same trigger acquisition window and are too close in time (< 100 ns) that cannot be disentangled, and are therefore detected as a single event. The associated energy is approximately the sum of the energies of the two overlapping events. This category of events is generated mainly by two ^{14}C events overlapping, by a combination of ^{14}C and ^{210}Po events or by a combination of ^{14}C and external background events. More details about the pile-up events are illustrated in Sec. 4.2.1 and Sec. 4.2.2.

Time and charge equalization

All the PMTs response must be aligned in time and charge on a regular time basis [61]. The Borexino equalization system is based on an external 394 nm laser source, whose emitted light is

carried simultaneously to each PMT by a dedicated system of multiplexed optical fibers.

The PMT time alignment is performed via software at least once a week. The achieved accuracy of the time equalization is 0.5 ns, below the PMTs time jitter (≈ 1 ns).The computed *time calibration constants* are written in the Borexino database and are used during the data processing to correct the response of each phototube on a channel-by-channel basis. [61].

The charge calibration is performed through the illumination of all the channels at the single photoelectron level by the 384 nm laser; the PMTs parameter for the single electron response are then measured and aligned. The gain has been demonstrated to be uniform within a gaussian uncertainty lower than 1 ADC count. Moreover the PMT illumination level is adjustable, to check the linearity of the charge response.

Energy reconstruction

The energy scale of the scintillation events must be known with high precision: indeed, the neutrino and background rates can be disentangled only by means of a spectral fit to the recoiled e^- energy. Also, the large dimensions of the detector require a careful mapping of its energy response in different positions within the scintillating volume.

Several variables are defined to estimate the energy of the events. Two basic information can be used to construct the energy estimators: the number of hits for each PMT, and the related charge collected by the PMTs [61]. Within a trigger window, the following estimators are constructed. N_p is the number of PMTs that have detected at least one hit (multiple hits are not included), while N_h is the total number of detected hits. Another possible way of estimating the energy of an event is by summing up the charge information of all PMT signals. This third energy variable, N_{pe}, is named *charge*. The main advantage of *charge* is its linearity at higher energy, being less affected by saturation effects with respect to *npmts*, in spite of a larger dark noise as a drawback.

As discussed previously, the number of active PMTs decreases in time. One way to account for this variation is to normalize the energy estimators on an event-by-event basis, to a fixed number of 2000 PMTs working channels:

$$N^n = f_n N \qquad f_n = \frac{2000}{N_{\mathrm{livePMTs}}} \tag{2.13}$$

where N is a generic energy estimator, and $N_{\mathrm{livePMTs}}(t)$ is the number of working channels for each detected event. Thus, three *normalized energy variables* are defined: N_p^{n}, N_h^{n}, N_{pe}^{n}. To get a feeling of the energy conversion scale, an energy deposit of 1 MeV in the detector center corresponds to about $N_{pe}^{\mathrm{n}} \approx 500\,\mathrm{p\,e}$ and $N_h^{\mathrm{n}} \approx N_p^{\mathrm{n}} \approx 350$. These variables allow to compare and align energy spectra from different years.

On the other side, in Borexino non-homogeneities in the light collection are present, because of the non-isotropic coverage of working PMTs on the Stainless Steel Sphere. This effect can alter the detected energy scale in the regions far from the detector center, but can be attenuated defining a set of *geometrically normalized estimators*. These variables take into account the geometrical weight of each PMTs, quantified by the relative position between the PMT and the event:

$$N^{\mathrm{gn}} = f_{gn}(t) N \qquad f_{gn}(t) = \frac{2000}{\sum_i \sin^2\theta_i} \qquad \theta_i = \frac{\pi R_{\mathrm{p,i}}^2 \cdot \vec{x}_{\mathrm{p,i}} \cdot (\vec{x}_{\mathrm{p,i}} - \vec{x})}{d^2 |\vec{x}_{\mathrm{p,i}}|\,|\vec{x}_{\mathrm{p,i}} - \vec{x}|} \tag{2.14}$$

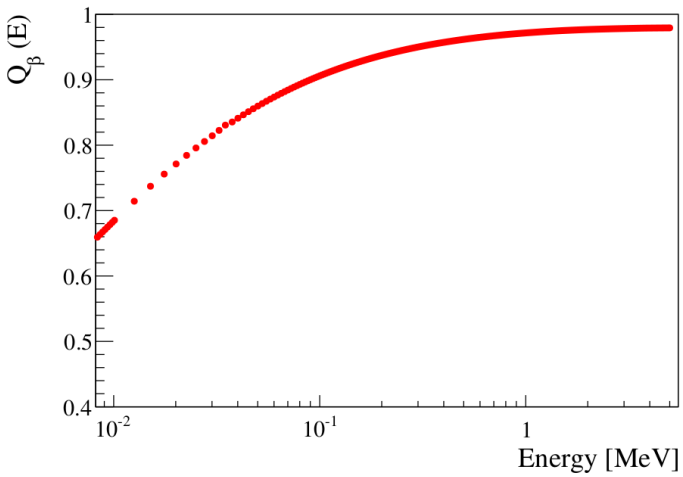


Figure 2.9: The **quenching factor** $Q_\beta(E)$ calculated from Birks formula Eq. (2.16) with the Borexino quenching parameter for β events ($k_\beta = 0.011\,\mathrm{cm/MeV}$).

where $\vec{x}_{\mathrm{p,i}}$ and $\vec{x}$ refer respectively to the PMT position and the event position, $R_{\mathrm{p,i}}$ is the distance of PMTs from the center, $d = \max_i R_{\mathrm{p,i}}$, and the sum runs over the active PMTs.

A parallel class of energy variables is given by the *fixed duration estimators*. They are built from the three previously described ones, but including all the available information up to dt1 = 230 ns or dt2 = 400 ns time intervals, respectively, after the beginning of each cluster (e.g. N_p^{dt1}, N_p^{dt2}, $N_p^{\mathrm{dt1-n}}$, $N_p^{\mathrm{dt2-n}}$ and so on). A fixed cluster estimator allows to study more accurately the pile-up event and to lower the dark noise rate, in spite of a loss of linearity for higher energies. In particular, $N_p^{\mathrm{dt1-n}}$ will be extensively used as reconstructed energy variable for the analytical multivariate fit, as illustrated in Chapter 4 and Chapter 7.

In principle, the energy deposited by a particle interacting in the Borexino scintillator is proportional to the number of photons collected by the PMTs, but the ionization quenching [143] introduces an intrinsic non-linear correction between the energy deposited in the scintillator E and the emitted light L_p:

$$L_p = Y_0 \cdot Q_p(E) \cdot E \tag{2.15}$$

where $Q_p(E) \leq 1$ is called quenching factor. The suffix p recalls that $Q_p(E)$ and L_p depend, for a fixed energy, on the particle type (α, β, or γ). Y_0, which is about $\sim 10^4$ photons/MeV, is the scintillation light yield in absence of quenching (i.e. $Q_p(E) = 1$). The knowledge of the quenching factor allows to convert the energy expressed in photons emitted to the value in MeV. The $Q_p(E)$ trend can be parametrized by the Birks model [143], but the involved parameters have been determined experimentally. Consequently, the detector calibrations played a fundamental role, allowing to get the energy conversion. The quenching factor $Q_\beta(E)$ for $\beta^\pm$ decay can be parametrized according to the Birks formula as:

$$Q_\beta(E) = \frac{1}{E}\int_0^E \frac{\mathrm{d}E}{1 + k_\beta \cdot \frac{\mathrm{d}E}{\mathrm{d}x}} \tag{2.16}$$

where $\frac{dE}{dx}$ is the specific energy loss and k_β is the Birks quenching parameter[7]. The energy dependence of $Q_\beta(E)$ calculated with $k_\beta = 0.011\,\text{cm/MeV}$ is reported in Fig. 2.9. The non-linear effect is more relevant as long as the energy deposit is below a few hundreds keV.

Even though α decays produce particles with energy above 4 MeV, which is well above the energy range of interest for the determination of solar neutrino interaction rates, the high quenching effects drastically lower the reconstructed energy. Indeed, the high density of ionization produced by these particles leads to a number of scintillation photons reduced by a factor of the order of ten with respect to an e^- with the same energy. The major α background contaminant in Borexino is the ^{210}Po isotope: its mono-energetic 5.3 MeV decay corresponds to a $\approx 500\,\text{keV}$ e^- recoil energy, populating the ^{7}Be-ν spectrum region.

Position reconstruction

The reconstruction of the events position in Borexino is based on the time distribution of the collected photons. It is performed by a time-of-flight algorithm [65] present in Echidna, the offline reconstruction code. An accurate position reconstruction performance plays a crucial role in the possibility to perform a fiducial volume (FV) software cut, that is a subregion of the scintillator where external background is minimized.

The position reconstruction algorithm is illustrated hereafter. We define as t_i the photon hit time on the i-th PMT fired, located at $\vec{r}_i = (x_i, y_i, z_i)$, and we assume that the scintillation occurred at an unknown position $\vec{r}_0 = (x_0, y_0, z_0)$ and time t_i' in the active volume. The time of emission of each photon, the event time, is named t_0 and has to be determined. T_i^{flight} is defined as the photon time of flight from the origin $\vec{r}_0$ of the photon to the position $\vec{r}_i$ of the i-th PMT. The goal of the position reconstruction is the determination of the $(\vec{r}_0, t_0)$ couple. For the generic i-th photon, the time of flight T_i^{flight} can be computed as:

$$T_i^{\text{flight}}(\vec{r}_0, \vec{r}_i) = \frac{|\vec{r}_i - \vec{r}_0|}{v_g} \qquad v_g = \frac{c}{n_{\text{eff}}} \tag{2.17}$$

where v_g is the group velocity of the photon wave packet emitted in the scintillation event. n_{eff} is the *effective refraction index*, defined as the mean of the group refraction index $n_{\text{group}}(\lambda)$ weighted on the energy distribution of photons produced by scintillation[8]:

$$n_{\text{group}}(\lambda) = n(\lambda) - \lambda \cdot \frac{dn(\lambda)}{d\lambda} \tag{2.18}$$

$$n_{\text{eff}} = \frac{\int_{-\infty}^{+\infty} E_{\text{PPO}}(\lambda) \cdot n_{\text{group}}(\lambda)\, d\lambda}{\int_{-\infty}^{+\infty} E_{\text{PPO}}(\lambda)\, d\lambda} \tag{2.19}$$

where $E_{\text{PPO}}(\lambda)$ is the emitted photon distribution for a scintillation event (see Fig. 2.2). In the wavelength range of photons in Borexino, between approximately 350 nm and 600 nm, $dn/d\lambda \approx \Delta n/\Delta\lambda \approx 3\%$: nevertheless the small value, the $\Delta n/\Delta\lambda$ term relevantly impacts on the effective velocity of the propagating scintillation wave packet. This has to be taken into account in the

[7] The k_β parameter depends on the considered scintillator.

[8] It is used to define an effective velocity of the photons, [61] taking into account that they travel as a wave packet and the scattering and reflections along their path.

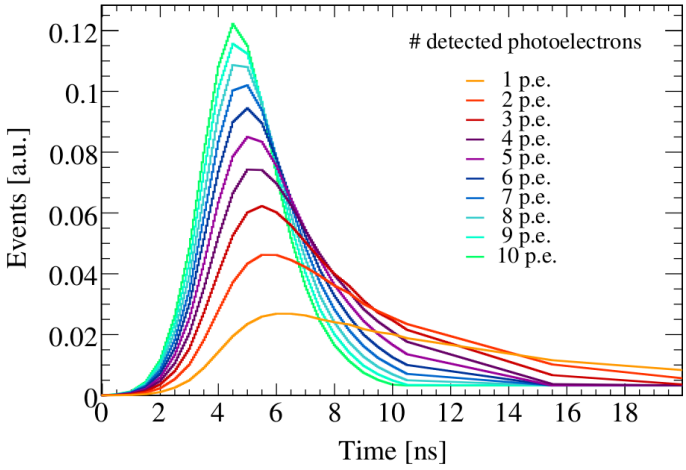


Figure 2.10: Position reconstruction algorithm: time distributions at which the first event fires a Borexino PMT, after time-of-flight subtraction. By comparing the events probability density functions, where either a single p_e or multiple p_e are detected, the time-of-arrival of the first p_e is pushed to earlier time in the latter cases.

reconstruction algorithm: the value currently used for the data reconstruction, i.e. $n_{\text{eff}}^{\text{data}} = 1.68$, is significantly larger than the pseudocumene refraction index measured one at 600 nm ($n_{\text{PC}} = 1.50$). Calibration data with radioactive sources have been crucial in determining the exact value of n_{eff} needed for a correct reconstruction of the event positions.

The reconstruction algorithm considers, for each photon, the emitted photon time distribution and it compares it with the reference probability density function of the profile of scintillation light, named $p(t)$. It maximizes the likelihood $\mathcal{L}(\vec{r_0}, t_0 \,|\, \vec{r_i}, t_i)$ with respect to t_0 and r_0, given the measured hit space-time set $(\vec{r_i}, t_i)$:

$$\mathcal{L}_{\text{pos}}(\vec{r}, \vec{t}) = \mathcal{L}\left(\vec{r}_0, t_0 | \vec{r}_i, t_i\right) = \prod_{i=1}^{N_{\text{livePMTs}}} p(t_i') = \prod_{i=1}^{N_{\text{livePMTs}}} p\left(t_i - \frac{|\vec{r_i} - \vec{r_0}|}{v_g} - t_0\right) \tag{2.20}$$

where n_{PMT} is the number of PMTs fired in the event. In case multiple photons hit a PMT, only the first one p_e produced is recorded. This introduces an energy bias in the algorithm, as a channel's hit time probability distribution from a specific interaction position depends on the output number of p_e. The larger the number of p_e, the earlier, on average, the time of the hit: if not corrected for, this effect would cause an outward reconstruction of events far from the center of the detector. To account for this effect, the reconstruction algorithm has been corrected introducing different scintillator response functions $p(t)$ according to the number of hitting photoelectrons: Fig. 2.10 shows the time distributions adopted in the computation of the likelihood.

α/β and e^+/e^- discriminations

The testing and tuning of α/β or e^+/e^- discrimination techniques are fundamental for the background rejection, and are based on the position reconstruction algorithm output.

The discrimination between α and β particle relies on the intrinsic time response of the PC+PPO

mixture (*pulse shape discrimination*). The emission times of produced light $S(t)$ depend on details of the charged particle energy loss: they can be phenomenologically described with a sum of four exponential functions with different time constants τ_i,

$$S(t) = \sum_{i=1}^{4} \frac{w_i}{\tau_i} \mathrm{e}^{-t/\tau_i} \tag{2.21}$$

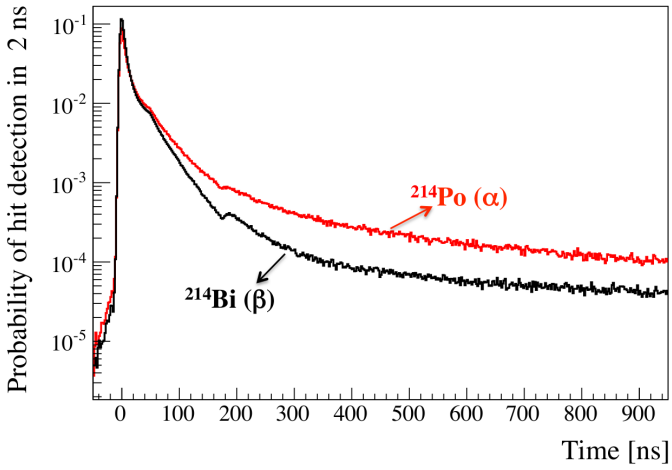


Figure 2.11: α/β identification. Examples of α and β events scintillation time profiles: ^{214}Po for α events, in red line, and ^{214}Bi for β events, in black line.

where w_i is the fraction of the total light emitted associated to each exponential function. Both the τ_i and w_i values are different for α and β particles, resulting in $S(t)$ profiles with different features, which are used to identify or discriminate which particle is associated to the scintillation event. As for example, the quickest exponential accounts the most of the emitted light but with different weights: 90% for β events and 65% for α ones. Therefore, the time distribution of light generated by α particles shows a more relevant tail than the one the β event (see Fig. 2.11). This feature allows to construct specific estimators of the fraction of detected hits for the tails. In this way, α background (i.e. ^{210}Po) can be isolated or statistically subtracted.

Several variables based on the time profile shape have been computed and tested for the α background rejection. The α-β discrimination variable (MLP) commonly used in Borexino is based on the neural network training, whose optimization is performed through a Multi Layer Perceptron algorithm. The estimator training is fulfilled through two ^{214}Bi – ^{214}Po coincidences reference samples (Fig. 2.12), returning a single number output for each event.

The different processes involved in the energy deposition for β^+ and β^- decays can be exploited to disentangle these two classes of events. In approximately 53% of cases, the emitted positrons form an orthopositronium state which has a non negligible lifetime of ≈ 3.1 ns. Moreover, the three 511 keV γ emitted as a consequence of the e^+e^- annihilation deposite their energy in different spots, such that the event cannot be considered pointlike. As a consequence, the time profile of photons emitted in e^+ events is different with respect to the one of e^- ones. The para-positronium time decay (≈ 0.12 ns) is instead completely negligible with respect to the scintillation time scale. Moreover, the e^+ energy deposition is not pointlike, as for β^- events, but slightly disperse because

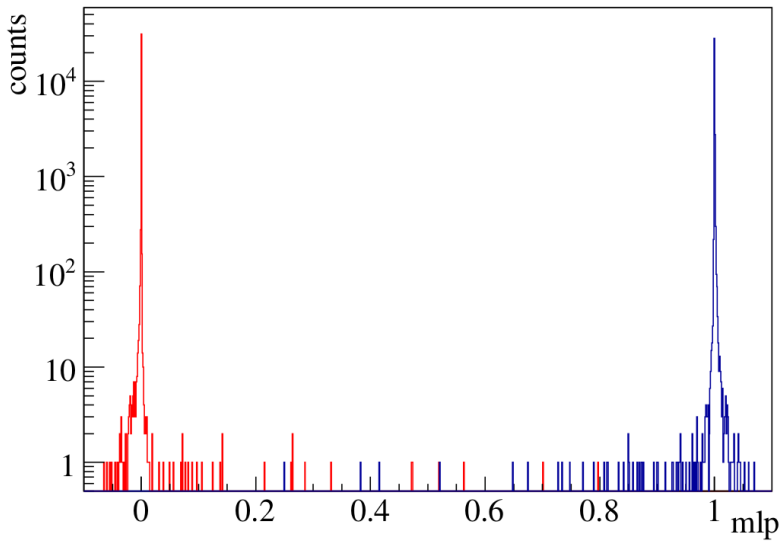


Figure 2.12: α/β identification. Distribution of MLP variable for α(red) and β (blue) events obtained by tagging a ^{214}Bi - ^{214}Po time coincidences sample [144].

of the 511 keV back-to-back γ annihilation.

These features result globally in a worsening of position reconstruction performance: the energy deposit of e^+ exhibits a larger time and spatial spread than the corresponding one generated by e^-. A pulse-shape discrimination parameter is built, named PS-$\mathcal{L}_{\mathrm{PR}} = \max_{\vec{r},\vec{t}} \mathcal{L}_{\mathrm{pos}}(\vec{r},\vec{t})$, i.e. the maximized value of the likelihood function employed in the position reconstruction (Eq. 2.20). This identification is crucial to isolate the cosmogenic ^{11}C, the most abundant e^+ event producer in Borexino and at the same time, an extremely delicate background for the measurement of pep-ν and CNO-ν (see Sec. 2.3.1). An example of the PS-$\mathcal{L}_{\mathrm{PR}}$ parameter distributions for e^+ and e^- samples is shown in Fig. 2.13.

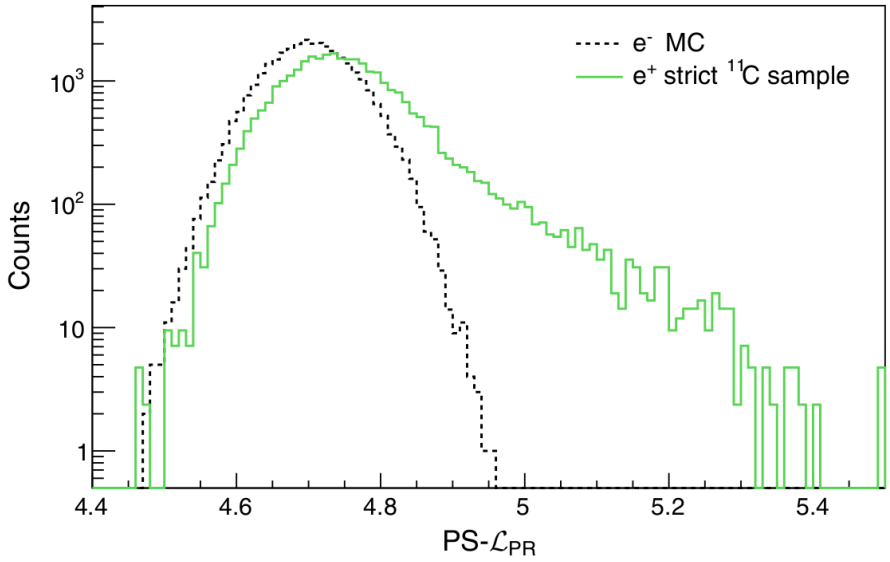


Figure 2.13: e^+/e^- identification. Comparison of the distributions of the PS-$\mathcal{L}_{\mathrm{PR}}$ variable for MC generated e^+ events (green, continuous line) and for e^- events selected from the data (black, dashed line). The latter category of events is a high-purity ^{11}C sample, obtained with the optimized TFC method (see Sec. 4.1.3), using very strict selections on the energy and on the time correlation with the neutron and muon tracks.

2.5 Borexino Phase-II Results

Thanks to its unique design and radiopurity levels, the Borexino detector achieved unprecedented results about solar neutrino physics during its ten year long data-taking. During the Phase-I (May 2007 - May 2010), Borexino detected for the first time the flux of the ^{7}Be-ν [62], which was the primary scientific goal of the experiment. Then, it ruled out any significant day-night asymmetry of ^{7}Be-ν interaction rate [64], and achieved the first direct observation of the pep-ν [138] and pp-ν [137]. The tightest upper limit so far on the CNO-ν flux has been also set [65]. Going to the higher energy range, Borexino have measured for the first time the ^{8}B-ν with an energy threshold 3 MeV. All the associated neutrino fluxes are found in agreement with the SSM+MSW predictions.

The Phase-II data taking (December 2011 - May 2016), which followed a purification campaign, allowed to perform the complete spectroscopy of solar neutrinos from the pp-chain, as described in Ref.[39]. The analysis has been splitted into two separate energy region characterized by different backgrounds: a low-energy region (LER) $0.19 - 2.93$ MeV, to measure the pp-ν, ^{7}Be-ν and pep-ν interaction rates, and a high-energy region (HER) $3.2 - 16$ MeV, dedicated to the ^{8}B neutrinos.

The most relevant results concerning the low energy analysis, which led to the simultaneous measurement of the interaction rates of pp-ν, ^{7}Be-ν, pep-ν with a global fit in an extended energy range $(0.19 - 2.93)$ MeV are extensively reviewed in Sec. 2.5.1 [39, 63]. All the pp-chain flux precisions have been improved with respect to the Phase-I analysis, reaching 2.7% for ^{7}Be-ν, 9.5% for pp-ν. The pep-ν discovery is claimed, rejecting at $> 5\sigma$ significance the hypothesis of absence of the pep reaction in the Sun.

Only an upper limit for the CNO-ν rate has been set by means of this analysis, reinforcing the one obtained from the Borexino Phase-I data. For what concerns the higher energy region, the ^{8}B-ν flux determination doubles the precision with respect to the previous Borexino result ($\approx 8\%$) and is found in agreement with the high-precision measurement performed by SuperKamiokande $\Phi(^8\text{B-}\nu) = 2.345 \pm 0.014\,(\text{stat}) \pm 0.036\,(\text{syst}) \cdot 10^6\,\text{cm}^{-2}\,\text{s}^{-1}$ [52].

As will be illustrated in Sec. 2.5.3, these measurements can be used either to test the MSW-LMA paradigm assuming SSM flux predictions or, in a complementary way, to probe our understanding of solar physics assuming the validity of the neutrino oscillation mechanism. When the flavor oscillation paradigm is assumed, the solar fusion hypotesis can be probed quantifying the relative intensity of the two primary terminations of the pp-chain, while the solar metallicity LZ-SSM or HZ-SSM scenarios can be tested analyzing the $\Phi(^7\text{Be-}\nu)$ and $\Phi(^8\text{B-}\nu)$ fluxes. Comparing the solar luminosity from neutrino fluxes with the very well-known value extracted by photon solar spectrum, it is also possible to prove the thermodynamic equilibrium over a 10^5 years timescale. On the other side, assuming the SSM flux predictions, the flavor oscillation physics in the solar matter is investigated. The vacuum-LMA hypotesis is disfavoured at 98.2% C.L., and results are in excellent agreement with the expectations from the MSW-LMA paradigm. Borexino is thus the first detector able to investigate the oscillation framework in the full solar neutrino energy range. More details on this analysis will be given in the next paragraphs, highlighting in particular those aspects in common to the CNO-ν analysis, which is the main subject of this thesis.

Species settings - Multivariate fit for pp-chain configuration			
Component	Status	Energy spectra	Event type
pp-ν	Free (Linked to pp/pep ratio)	All	β^-
pep-ν	Free (Linked to pp/pep ratio)	All	β^-
^{7}Be-ν	Free	All	β^-
CNO-ν	Fixed to HZ-SSM / Free	All	β^-
^{8}B-ν	Fixed to HER value	All	β^-
^{14}C	Penalty	All	β^-
Pile-up	Penalty	All	β^-
^{85}Kr	Free	All	β^-
^{210}Bi	Free	All	β^-
^{210}Po	Free	All, with separate rates	α
^{11}C	Free	All, with separate rates	β^+
^{10}C	Free	TFC-tagged only	β^+
^{6}He	Free	TFC-tagged only	β^-
^{214}Bi	Free	All	γ
^{206}Tl	Free	All	γ
^{40}K	Free	All	γ

Other fit settings - MC fit		
Fiducial volume		$r < 2.8\,\mathrm{m}$, $-1.8\,\mathrm{m} < z < 2.2\,\mathrm{m}$
Fiducial mass		71.3 t
Exposure		$905.5\,\mathrm{days} \times 100\,\mathrm{t}$
Variables		N_h^n, r^3, PS-$\mathcal{L}_{\mathrm{PR}}$
Ranges	N_h	$85 \leq N_h^n \leq 900$
	r^3	$290 \leq N_h^n \leq 937$
	PS-$\mathcal{L}_{\mathrm{PR}}$	$409 \leq N_h^n \leq 645$

Table 2.3: Phase-II LER multivariate fit: neutrino and background species. List of the species involved in the multivariate fit (top table) and other fit configurations (bottom table, Monte Carlo fit example). The CNO-ν is fixed according to HZ-SSM or LZ-SSM predictions; only in the case of determination of its upper limit it is left free and the pp-ν and pep-ν rates are linked through the pp/pep flux ratio. A similar procedure is applied for the pep-ν rate. The *separate rates* notation marks that the species rate is considered as two different parameters for the two TFC-tagged and TFC-subtracted energy spectra (^{210}Po and ^{11}C).

2.5.1 Simultaneous determination of low energy neutrinos interaction rates

The simultaneous determination of the pp-ν, pep-ν and ^{7}Be-ν interaction rates has been performed for the first time analyzing the Borexino Phase-II dataset, for a total livetime of 1291.5 days. Data spans the time period between 11 December 2011 and 17 July 2016, corresponding to an exposure of 905.5 days $\times$ 100 t, which amounts to approximately 1.6 times the Borexino Phase-I one. The Phase-II data benefits from the reduction of the contaminants of many background isotopes thanks to the extensive purification campaign carried out in 2010-2011, leading to the following background levels: ^{238}U $< 9.4 \cdot 10^{-20}$ g/g (95% C.L.), ^{232}Th $< 5.7 \cdot 10^{-19}$ g/g (95% C.L.), ^{85}Kr and ^{210}Bi reduced by factors ≈ 4.6 and ≈ 2.3 respectively.

In order to determine the solar neutrino rates, the main issue is given by the discrimination of the neutrino signal from the background. The main flow of the solar neutrino analysis in Borexino is shortly summarized hereafter, while it is extensively described in the context of the search for CNO-ν in Chapter 4. We remind that neutrino-induced events in liquid scintillator are intrinsically indistinguishable on an event-by-event basis from the most of backgrounds due to β or γ decays. The *standard cuts* are a set of selection conditions for background reduction, optimized for the solar neutrino low energy analysis. These are applied as the first step for the analysis, aiming to remove the taggable backgrounds: muons and muon-induced events, noise events, radioactive decays from delayed coincidences. The analysis is restricted to a software-defined central portion of the scintillator, named Fiducial Volume (FV), in which the signal over noise ratio is maximized. The suppression of external background as much as possible is crucial, since its rate is exponentially increasing going from the center towards the Inner Vessel. Events are selected in the FV defined by the conditions $r < 2.8$ m, -1.8 m $< z < 2.2$ m. After these two selection steps, the cosmogenic ^{11}C events still represent the major background in the 1.1 MeV $\lesssim E \lesssim$ 1.8 MeV region (see Sec. 2.3.3). The main technique for the ^{11}C tagging is named Three-Fold coincidence (TFC, see Sec. 4.1.3), and exploits the peculiar time and spatial features of its decay products: the ^{11}C positron decay, the parent muon, and the signal from capture of the free neutron. Another tagging technique to identify the ^{11}C background is given by the e^+/e^- pulse shape discrimination variable, as described in Sec. 2.4.1. The standard cuts are also the basis of my analysis to extract the CNO-ν signal; more details on them will be given in Chapter 4.

Even after applying the standard cuts, several backgrounds are still present and it is necessary to apply a multivariate fit to extract the rates of signal and background from the data. The expected distributions of the physical quantities of interest (energy estimators, radial position of events, e^+/e^- discriminator variable) are built according to analytical models or Monte Carlo simulations, and then are fitted against selected data. This such dual approach enhances the robustness of the analysis results, allowing also to evaluate the systematic uncertainty associated to the fitting method. The multivariate fit will be used in this thesis analysis to extract the CNO-ν signal and will be described in more detail in Chapter 4.

The multivariate fit is performed within a $(0.19 - 2.93)$ MeV energy window, the so-called LER (Low Energy Region). Most of the neutrino signals and backgrounds rates are left free to vary in the fits, with some noticeable exceptions (see Tab. 2.3). The ^{8}B-ν spectrum covers only marginally the LER, therefore the related interaction rate does not significantly impact on the analysis, and it is fixed to the independent HER result $R(^{8}\text{B-}\nu) = 0.46$ cpd/100t (see Sec. 2.5.2). The proper determination of the ^{14}C event rate is crucial for the multivariate fit success, since it dominates the

Solar ν	R_ν (Borexino Phase-II) [cpd/100ton]	R_ν (GS98) - HZ [cpd/100ton]	R_ν (AGS09) - LZ [cpd/100ton]
pp	$134 \pm 10^{+6}_{-10}$	131.1 ± 1.4	132.2 ± 1.4
^{7}Be	$48.3 \pm 1.1^{+0.4}_{-0.7}$	47.9 ± 2.8	43.7 ± 2.5
pep (HZ)	$2.43 \pm 0.36^{+0.15}_{-0.22}$	2.74 ± 0.04	2.78 ± 0.04
pep (LZ)	$2.65 \pm 0.36^{+0.15}_{-0.24}$	2.74 ± 0.04	2.78 ± 0.04
CNO	< 8.1 (95% C.L.)	4.92 ± 0.78	3.52 ± 0.52

Table 2.4: Phase-II multivariate fit results: best estimates for neutrino interaction rates. First column reports the pp-ν, ^{7}Be-ν, pep-ν and CNO-ν solar neutrinos interaction rate, averaged on the multivariate analytical and Monte Carlo fits. The first error is the statistical uncertainty derived by profiling the likelihood under Wilks approximation, while the second error is the systematic uncertainty. R(pep - ν) depends on the choice of solar metallicity scenario (LZ or HZ) to which the CNO-ν rate is set. The second and the third columns reports the expected interaction rate in Borexino according to the LZ-SSM and HZ-SSM scenarios and assuming the validity of MSW flavor oscillations.

lower part of the energy spectrum: possible bias for its rate may crucially affect other components (^{85}Kr, pp-ν, ^{210}Po). The ^{14}C rate can be measured with good precision by an indipendent analysis based on double cluster events [137]. The rate obtained from this analysis is 40 ± 1 Bq/100t and is used as a pull-term constraint in the multivariate fit.

For what concerns pep-ν, the tight correlation of its spectral shape with the one of CNO-ν prevent the multivariate fit to disentangle the two rates (see Fig. 2.16). Therefore, to extract pep-ν rate it is necessary to constrain the one of CNO-ν in the fit, in particular to the SSM predictions assuming the validity of MSW flavor oscillations. This topic will be addressed more in detail in Chapter 3, in the context of CNO-ν determination. The results of the fit for pep-ν rate when the CNO-ν rate is fixed according to the HZ-SSM or to the LZ-SSM scenarios differ by 0.22 cpd/100t and are therefore quoted separately in Table 2.4.

The interaction rates of solar neutrinos and the background species, obtained by averaging the results of the analytical and Monte Carlo approaches of Phase-II analysis, are reported in Tab. 2.4 and Tab. 2.5 respectively.

The four distributions fitted in Phase-II Monte Carlo multivariate fit are shown in Fig. 2.14 (TFC tagged and TFC subtracted energy spectra) and in Fig. 2.15 (PS $- \mathcal{L}_{\rm PR}$ and radial histograms). Neutrino rates are consistent with respect to the Borexino Phase-I results, showing improved precisions. In particular, $R(^7\text{Be} - \nu)$ is measured with an unprecedented precision of 2.7%.

Thanks to the improved R(pep - ν) precision, the hypotesis of absence of pep-ν is ruled out at more than 5σ C.L. The marginalized $\Delta\chi^2$/NDF profile for the pep-ν is shown in top plot of Fig. 2.17. An upper limit on the CNO-ν flux is obtained fixing the pp/pep ν rate to the predictions of the SSM: R(CNO - ν) < 8.1 cpd/100t at 95% C.L. This limit is compatible with the one set by Borexino Phase-I data ($R^{\text{Bx - Phase - I}}$(CNO - ν) < 7.9 cpd/100t at 95% C.L.), but relying on weaker assumptions on the pep-ν rate [9]. The marginalized $\Delta\chi^2$/NDF profile for the CNO-ν interaction rates is shown in bottom plot of Fig. 2.17.

[9] The Borexino Phase-I analysis assumed the pep-ν to be fixed to the SSM predictions.

Background	Rate (Borexino Phase-II) [cpd/100ton]
^{14}C [Bq/100t]	40.0 ± 2.0
^{85}Kr	6.8 ± 1.8
^{210}Bi	17.5 ± 1.9
^{11}C	26.8 ± 2.0
^{210}Po	260.0 ± 3.0
Ext. ^{40}K	1.0 ± 0.6
Ext. ^{214}Bi	1.9 ± 0.3
Ext. ^{208}Tl	3.3 ± 0.1

Table 2.5: Phase-II multivariate fit results: best estimates for background interaction rates (^{14}C, ^{85}Kr, ^{210}Bi , ^{11}C, ^{210}Po and external backgrounds) averaged on the analytical and Monte Carlo multivariate fits. The uncertainties are reported adding in quadrature the statistical and the systematic ones.

2.5.2 Measurement of ^{8}B-ν interaction rate

The determination of the ^{8}B-ν flux, whose spectrum extends up to ≈ 15 MeV, requires a separate High Energy Region (HER) analysis, and a different fitting strategy with respect to the one performed in the low energy region. The most recent results [39, 145] are based on the combined Phase-I and Phase-II dataset analysis, spanning from January 2008 to December 2016, totalling a $1.5\,\mathrm{kt}\cdot\mathrm{y}$ exposure. This amounts to ~ 11.5 times the statistics used in the previous analysis [136].

The reason for a separate analysis is due to the different backgrounds affecting the higher energy region of Borexino spectrum. The selected HER is subdivided into two independent regions, named HER-I ($3.2\,\mathrm{MeV} < E < 5.7\,\mathrm{MeV}$) and HER-II ($5.7\,\mathrm{MeV} < E < 16\,\mathrm{MeV}$). The standard data selection cuts described in Sec. 4.1 allow to discard most of the background events for both the HERs.

In Borexino, the energy deposited by natural radioactivity never exceeds $Q(^{208}\mathrm{Tl}) = 5$ MeV: in fact the α-decaying isotopes, as the ^{210}Po , suffer from the scintillation quenching which shifts the event energy by a factor 10-12, moving them to the LER region. The other background events surviving the selections for both HERs are mainly the cosmogenic ^{11}Be isotope, which is subtracted through a separate selection technique, and γ from neutron capture reactions.

Afterwards, a radial fit is performed to disentangle signal from background, separately for events belonging to the HER-I and HER-II energy regions. As mentioned previously, the HER-II is the less populated region, for what concerns both signal and background, and the radial model is based on two components only: the ^{8}B-ν profile, which is uniformly distributed in the scintillator volume, and the external neutron capture components, which is exponentially decreasing approaching going from the IV to the center. The HER-I model instead includes three additional radial components due to the presence of the ^{208}Tl background (bulk, dissolved in the scintillator; surface, intrinsic to the nylon vessel; emanation, due to the diffusion from the IV).

The radial fits of the selected events in the HER-I and HER-II analysis are shown in Fig. 2.18 [145]. The extracted ^{8}B solar neutrino interaction rate is $R(^8\mathrm{B}\text{-}\nu) = 0.220 \pm^{+0.015}_{-1.016}\,(\mathrm{stat.})\,^{+0.006}_{-0.006}\,(\mathrm{stat.})\,\mathrm{cpd/100t}$, improving the precision to $\approx 8\%$, more than a factor two with respect to the previous Borexino result

[136]. The fundamental ingredients for this improvement are given by the lower ^{208}Tl contamination in the purified scintillator, the refinement of the detector background model, the accuracy of the Monte Carlo simulations, which allowed to extend the analysis to a larger portion of scintillator.

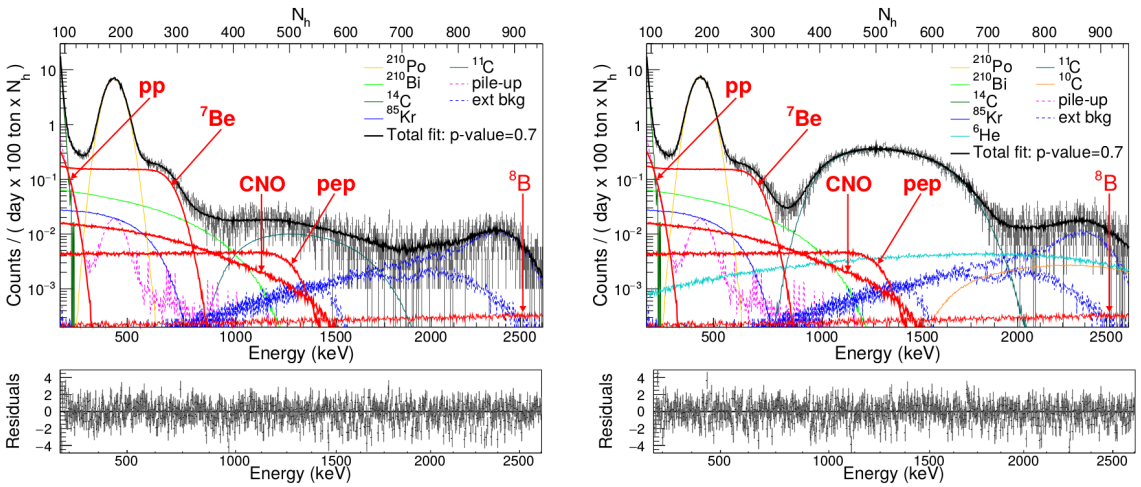


Figure 2.14: Phase-II multivariate fit: energy spectra. Monte Carlo fit results for the TFC-subtracted (left) and the TFC-tagged (right) energy spectra, with residuals shown in respective bottom panels. The sum of the individual components from the fit (black lines) is superimposed on the data (grey points). The analysis has been performed using N_h^n as energy estimator.

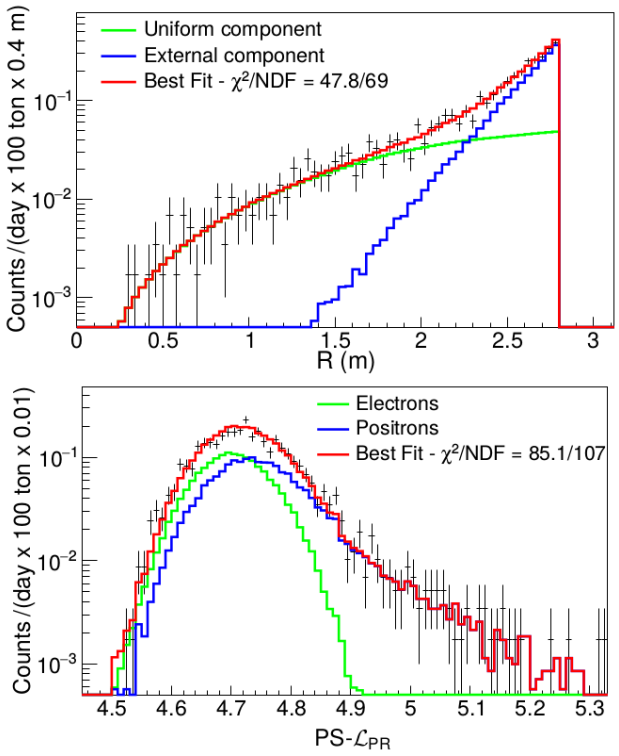


Figure 2.15: Phase-II multivariate fit: β^+/β^- pulse-shape variable and radial histograms. Radial (top) and the PS $-$ $\mathcal{L}_{\mathrm{PR}}$ (bottom) distributions of the events (black crosses) to be fitted in the multivariate analysis from the TFC-subtracted spectrum, in the energy intervals $N_h^n > 290$ and $409 < N_h^n < 645$, respectively.

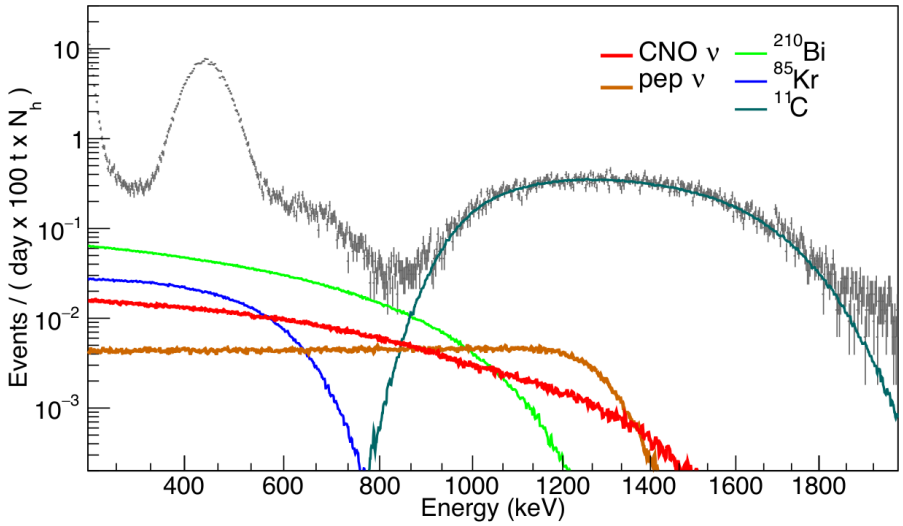


Figure 2.16: ^{210}Bi **- CNO - pep spectral shape degeneracy**. The Borexino Phase-II energy spectrum after selection cuts is superimposed to the expected profiles for some of the LER analysis species of interest. The strong similarity between the ^{210}Bi - CNO-ν, and pep-ν spectral shapes can be noticed.

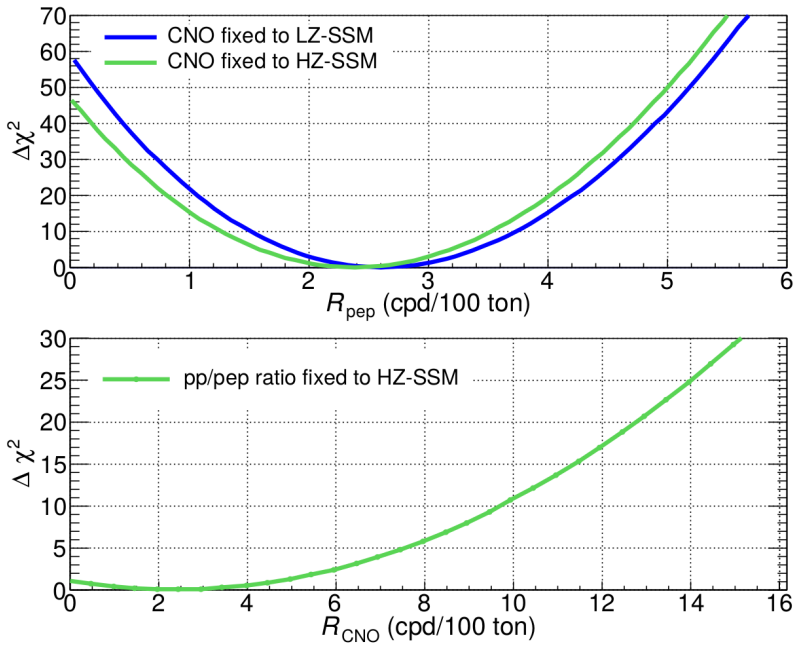


Figure 2.17: Borexino Phase-II results: marginalized $\Delta\chi^2$/NDF profiles for the pep-ν (top plot) and CNO-ν (bottom plot) interaction rates (bottom plot).

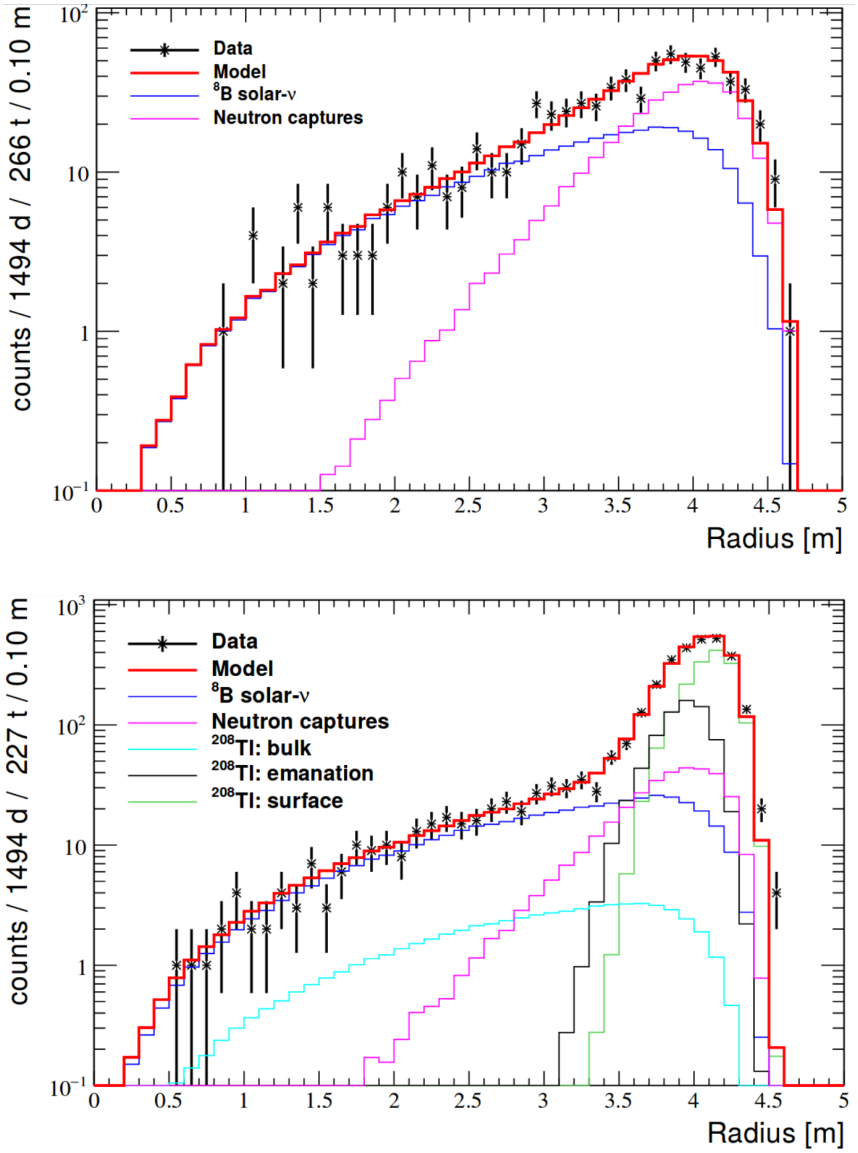


Figure 2.18: Phase-II HER analysis: radial fits. Top panel: fit of the events radial distribution in HER-I window (3.2 MeV $< E <$ 5.7 MeV). Bottom panel: fit of the events radial distribution in the HER-II window (5.7 MeV $< E <$ 16 MeV).

2.5.3 Implications on solar models and on flavor oscillation scenario

The comprehensive Borexino measurements of the pp-chain ν fluxes cover simultaneously, for the first time, all the solar neutrino energy spectrum of interest with the same dataset and analysis procedure. These measurements can be used either to get implications on solar physics or on the flavor oscillation framework: either to test the MSW-LMA paradigm, assuming SSM flux predictions or, conversely, to probe our understanding of solar physics assuming the validity of the neutrino oscillation mechanism.

Testing the Standard Solar Model

The Borexino Phase-II neutrino rates measurements can be employed to test the Standard Solar Models predictions. The validity of the oscillation framework needs to be assumed to extract the total flux of neutrinos: Borexino detects all the neutrino flavors, but the cross section for ν_μ and ν_τ detection is ≈ 6 times lower with respect to the ν_e one. We now review the main implications of the previously described measurements on some solar physics aspects.

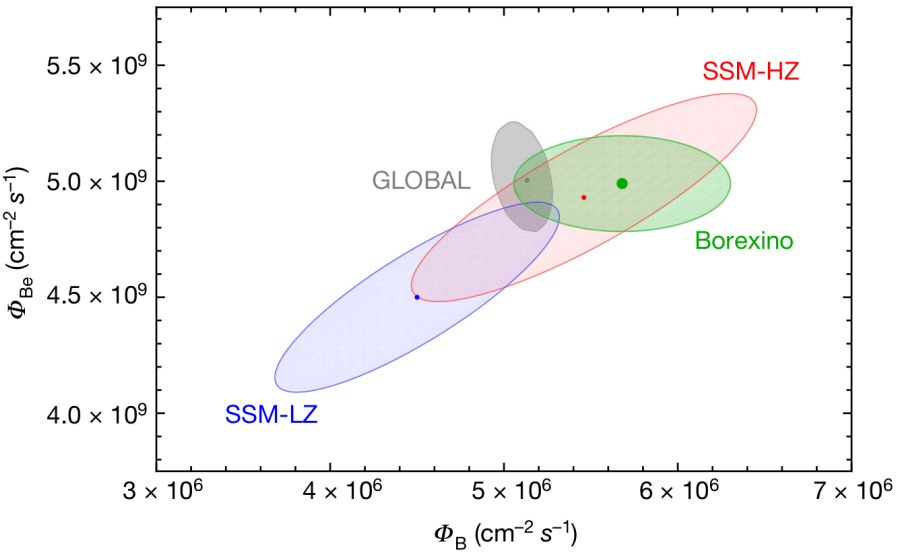


Figure 2.19: Phase-II results implications on solar metallicity: $\Phi(^7\text{Be}) - \Phi(^8\text{B})$ plane. Borexino contours for $\Phi(^7\text{Be})$ and $\Phi(^8\text{B})$ fluxes (green point and area), along with the theoretical SSM-HZ and SSM-LZ predictions (red and blue ellipses, respectively). The Borexino results are compatible with both metallicity scenarios, even if dedicated frequentists and Bayesian statistical analysis shows a mild preference for HZ. Combining Borexino data with solar and KamLAND data in a global fit (grey ellipse), this tendence to HZ is weakened.

As discussed in Sec. 1.5, the precise measurements of the solar neutrino fluxes is the keystone to solve the solar metallicity problem. The two metallicity scenarios provide almost equivalent fluxes for the first stages of the *pp* chain fluxes, while more significative differences show up in the latest steps (see Tab. 1.1: ^{7}Be: $\approx 8\%$ HZ-LZ flux difference, ^{8}B: $\approx 17\%$ HZ-LZ flux difference). The Borexino Phase-II contours for $\Phi(^7\text{Be}\text{-}\nu)$ and $\Phi(^8\text{B}\text{-}\nu)$ fluxes can be analyzed in a dedicated plane (see Fig. 2.19) to probe the solar metallicity problem, comparing the experimental results with the

HZ-SSM and LZ-SSM flux predictions. In the full picture, the theoretical model suffers from the larger uncertainty, mainly due to the astrophysical factors for the $^3\text{He} + {}^4\text{He}$ and $p + {}^7\text{Be}$ reactions, and to the solar opacity values.

As shown in Fig. 2.19, the Borexino $\Phi({}^7\text{Be}\text{-}\nu)$ and $\Phi({}^8\text{B}\text{-}\nu)$ results (green ellipse) are compatible with both HZ and LZ scenarios; dedicated frequentists and Bayesian statistical analysis shows a mild preference for HZ [39]. Combining Borexino data with solar and KamLAND data in a global fit (grey ellipse), this tendence to HZ scenario is weakened. To shed light on the still unsolved solar metallicity problem, a precise measurement for CNO-ν flux is needed: it would be the ideal probe thanks to the larger difference of the HZ-SSM and LZ-SSM flux predictions $\approx 28\%$ (see Tab. 1.1).

Solar neutrinos have the unique feature to provide real-time information on the core of the Sun. The measured neutrino fluxes can be used to extract the total power generated by nuclear reactions in the Sun core, since the amount of released energy for each nuclear reaction is well known from nuclear physics. The *luminosity constraint* connects the neutrino fluxes to the total solar luminosity at the Earth surface [146, 147]:

$$L^{\nu}_{\odot} = 4\pi\, d^2_{\text{ES}} \sum_i \alpha_i\, \Phi_i \qquad i = \text{pp, pep, } {}^7\text{Be, } {}^8\text{B, CNO} \tag{2.22}$$

where Φ_i and α_i are respectively the i-th neutrino flux and the thermal energy released, while $d_{\text{ES}} = 1\,\text{AU}$ is the average Earth-Sun distance. The α_i numerical values, available at Ref. [147] are determined by the differences between nuclear masses and are independent on details of the solar model at the 10^{-4} level. The Eq. (2.22) sum runs over all the neutrino fluxes, whose associated fusion reactions could in principle contribute significantly to the energy budget of the sun. Under basic assumptions on the equilibrium for some of the nuclear reactions, the total solar luminosity is $L^{\nu}_{\odot} = 3.89^{+0.35}_{-0.42} \cdot 10^{33}\,\text{erg/s}$, which is compatible with the more accurate value obtained in the photon channel: $L^{\gamma}_{\odot} = (3.846 \pm 0.015) \cdot 10^{33}\,\text{erg/s}$. Consequently, the nuclear source of the Sun powering is proven once again. Moreover, a photon takes on average $\sim 10^5$ years to reach the photosphere from the generation point in the Sun core (see Sec. 1.5), it is possible to claim that the Sun is found in thermodynamical equilibrium over a such timescale.

The simultaneous measurement of pp-ν and ${}^7\text{Be}$-ν fluxes allows Borexino to probe the solar fusion hypotesis in a second way. The following ratio $R_{\text{I/II}}$ quantifies the relative intensity of the two primary terminations of the pp-chain (*pp*-II and *pp*-I[10]) [148]:

$$R_{\text{I/II}} = \frac{R({}^3\text{He} + {}^4\text{He})}{R({}^3\text{He} + {}^3\text{He})} = \frac{2\Phi({}^7\text{Be}\text{-}\nu)}{\Phi(\text{pp}\text{-}\nu) - \Phi({}^7\text{Be}\text{-}\nu)} \tag{2.23}$$

The experimental value from Phase-II, $R^{\text{BxII}}_{\text{I/II}} = 0.178^{+0.027}_{-0.023}$, is compatible with the SSM predictions: $R^{\text{LZ}\text{-}\text{SSM}}_{\text{I/II}} = 0.161 \pm 0.010$ and $R^{\text{HZ}\text{-}\text{SSM}}_{\text{I/II}} = 0.180 \pm 0.010$. On the other side, the high experimental 14% uncertainty on $R^{\text{BxII}}_{\text{I/II}}$ prevents this test to discriminate between the two metallicity scenarios.

Solar matter effects in flavor oscillation

Once the validity of the SSM flux predictions is assumed, information about the flavor oscillation framework can be inferred. From the measured interaction rates of pp-ν, ${}^7\text{Be}$-ν, ${}^8\text{B}$-ν, one can

[10] pp-III termination can be neglected because of the extremely low B.R. $\sim 10^{-5}$. Note also that, since the ${}^8\text{B}$-ν contribution is very low in the LER, it has been neglected at the denominator of Eq. (2.23).

determine the survival oscillation probability Eq. 1.70 both for the flavor-dominated regime and for the matter-dominated regime[11].

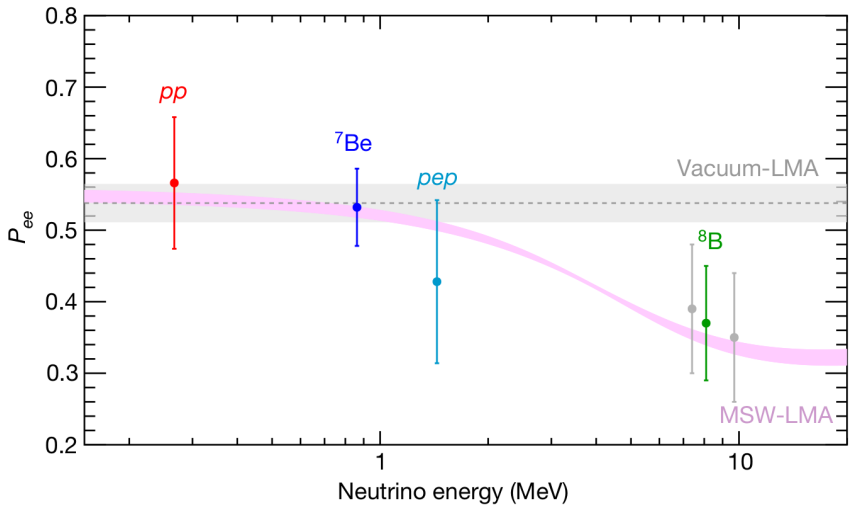


Figure 2.20: Phase-II results implications on flavor oscillations: ν_e survival probability P_{ee} as a function of neutrino energy. Borexino Phase-II results are shown for pp-ν (red point), ^{7}Be-ν (blue point), pep-ν (cyan point) and ^{8}B-ν (green point for the full HER range, and grey points for the separate HER-I and HER-II sub-ranges), assuming the validity of HZ-SSM predictions for solar neutrino fluxes. The grey band is the $\pm1\sigma$ prediction for the vacuum-LMA oscillation scenario, while the pink band represents the $\pm1\sigma$ prediction of MSW-LMA scenario. Borexino results disfavour the vacuum-LMA hypotesis at 98.2% C.L., while they are in excellent agreement with the expectations from the MSW-LMA paradigm [39].

The full scenario is reported in Fig. 2.20, where the ν_e survival probability P_{ee} is shown as a function of the neutrino energy for all the solar neutrino reactions measured by Borexino (pp-ν in red, ^{7}Be-ν in blue, pep-ν in cyan, ^{8}B-ν HER/HER-I/HER-II in green/grey). The grey horizontal band shows the $\pm1\sigma$ prediction for the vacuum-LMA scenario, where matter oscillation effects are not present and P_{ee} does not depend on neutrino energy E, but only on the θ_{12}, θ_{13} and Δm^2_{12} oscillation parameters[12]:

$$P_{ee}^{\text{vacuum}} = \cos^4\theta_{13}\left(1 - \frac{1}{2}\sin^2 2\theta_{12}\right) + \sin^4\theta_{13} \tag{2.24}$$

The pink band represents the $\pm1\sigma$ prediction of MSW-LMA scenario, where the matter effect plays an important role for the higher energies of the (matter-dominated) regime:

$$P_{ee}^{\text{MSW-LMA}}(E) = \frac{1}{2}\cos^4\theta_{13}\left(1 + \cos 2\theta_{12}^M(E)\,\cos 2\theta_{12}\right) \tag{2.25}$$

where $\theta_{12}^M(E)$ absorbs the energy dependence and the matter effects, and is defined following the description of Sec. 1.7.2. Borexino provides the most precise measurement of P_{ee} in the LER, where

[11] The Standard Model $\nu_e - e^-$ cross-section is also assumed [4].

[12] Oscillation parameters θ_{12}, θ_{13} and Δm^2_{12} are extracted from the global neutrino oscillation analysis found at Ref. [149].

flavour conversion is vacuum-dominated. In the higher energy window, where the solar matter effects take over, the Borexino results are in agreement with the high-precision measurements performed by SuperKamiokande [52] and SNO [57]. Overall, Borexino results disfavour the vacuum-LMA hypotesis at 98.2% C.L., while they are in excellent agreement with the expectations from the MSW-LMA paradigm [39].

2.6 The SOX project and other physics fields goals

Besides solar physics, the unique features of Borexino make it competitive also for measurements not related to solar neutrino physics. Some of the analysis and results obtained during the data taking are given by the discovery of geo-neutrinos [12, 13, 14], the search for gravitational wave events follow-up [150, 151], the constraining of neutrino magnetic moment [152], and the search for rare events beyond the Standard Model [153, 154] and Non-Standard neutrino interactions [155].

Also a search for light sterile neutrinos with Borexino was originally planned. The *SOX* experiment [156] (*Short distance neutrino Oscillations with BoreXino*) aimed to probe the LSND and MiniBooNE anomalies (see Sec. 1.9.1) by placing an intense artificial ^{144}Ce-^{144}Pr $\overline{\nu}_e$ source of ~ 1.5 PBq activity close to Borexino. The starting material for the source was planned to be composed of 2.8 ton of spent nuclear fuel from Kola Nuclear Power Plant in Russia. The arrival of the source at LNGS and the following data taking were expected to start in the beginning of 2017-2018, lasting for $1 - 1.5$ year(s). Unfortunately, severe issues in the production of the artificial source definitely determined the shut-off of the SOX project [157]. I have worked on the SOX project throughout my first PhD year, focusing on the evaluation of the experiment sensitivity studies for the sterile neutrino hypotesis; results are reported in Chapter 8.

2.7 Conclusions

The Borexino experiment has provided the most complete measurement of all the solar neutrino fluxes from the pp-chain nuclear fusion reactions. The measured neutrino interaction rates from the analysis of Phase-II dataset are compatible with the previous Borexino Phase-I results and, in all cases, their precision is improved. The pep-ν absence hypothesis is rejected, for the first time, at $> 5\sigma$ confidence level. The implications of these neutrino fluxes measurements have been reviewed both in terms of solar physics field, for what concerns the solar fusion and the solar metallicity problem, and in terms of the flavor oscillation framework, extracting the ν_e survival probability in a wide energy range. The missing tile of the solar neutrino framework consists in the search for CNO-ν evidence: the results of the Borexino Phase-III dataset analysis, fully dedicated to the CNO-ν hunt, will be described in Chapter 7.

Chapter 3

Strategy for the search of CNO neutrinos with Borexino

Borexino has successfully determined the fluxes of solar neutrinos emitted by all the fusion reactions belonging to the so-called pp chain [39, 63, 136, 62, 138, 137]. The only missing elements in this otherwise complete experimental picture of the interior of the Sun are the CNO neutrinos. As discussed in Chapter 1, the CNO cycle is believed to be the main engine of very massive stars, while it is expected to contribute only to $\approx 1\%$ of the Sun luminosity. Observing neutrinos from the CNO cycle fusion reactions would have therefore a striking importance in astrophysics, since it would provide the first experimental proof of the existence of this important source of energy in the core of the stars. So far, only limits on the CNO neutrino flux have been experimentally obtained: the most stringent upper limit on the CNO-ν flux is $\Phi_{\mathrm{CNO}} < 7.7 \cdot 10^8\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at 95% C.L. and has been set by Borexino itself [65].

The search for a CNO-ν signal with Borexino represents the main subject of this PhD thesis. It is worth to recall that Borexino plays a unique role in the CNO neutrinos hunt, thanks to its unprecedented radiopurity levels. Indeed, extensive sensitivity studies which will be reported in details in this Chapter show that Borexino has the chance to tackle CNO neutrinos provided that the interaction rate of ^{210}Bi isotope, the most troublesome background, is determined independently in the scintillator and constrained in the multivariate analysis. The indications stemming from these sensitivity studies have driven the strategy for the CNO-ν search adopted in this thesis, which is based on four main pillars:

1. An effective *data selection*, aiming to enhance significantly the signal over background ratio.
2. A *multivariate fit* to disentangle the neutrino signal from the residual background after the selection cuts, based on either an analytical or Monte Carlo description of the detector response.
3. An accurate *Monte Carlo simulation framework* to describe the detector response function.
4. A technique to tag the ^{210}Bi events, which represent the most important background for the CNO-ν analysis, in order to constraint the associated interaction rate in the multivariate fit.

This Chapter is organized as follows. The Borexino sensitivity to CNO-ν signal without and with the application of independent constraints to the ^{210}Bi background and to pep-ν interaction rates

is described respectively in Sec. 3.1 and Sec. 3.2. The sensitivity studies that I will present are the result of the work of many people within the Borexino analysis team; I have taken part in the rate-only studies, which are discussed in the first part of Sec. 3.2. The strategy for the CNO-ν measurement followed within this PhD thesis is outlined in Sec. 3.3: the key elements necessary for the CNO-ν rate determination will be shortly reviewed. More details on each step of this path will be thoroughly discussed in the following Chapters.

3.1 Borexino sensitivity to CNO-ν signal

The preliminary step towards the detection of a CNO-ν signal is to study via Monte Carlo simulations the sensitivity of Borexino under realistic conditions: this allows to define which are the elements playing the most relevant roles and therefore to elaborate a strategy for the analysis.

The sensitivity to CNO-ν is driven by two main factors:

- **Low signal/background ratio for the CNO-ν interaction rate in Borexino.** According to the SSM predictions [25], the CNO-ν reactions are expected to play a secondary role in the solar reaction framework: this translates into a faint neutrino flux, and therefore into a very low counting rate in Borexino. The expected CNO-ν interaction rate in Borexino, according to the high metallicity and low metallicity scenarios, is $R^{\mathrm{HZ}}(\mathrm{CNO}\text{-}\nu) = 4.92 \pm 0.78\,\mathrm{cpd/100t}$ or $R^{\mathrm{LZ}}(\mathrm{CNO}\text{-}\nu) = 3.56 \pm 0.52\,\mathrm{cpd/100t}$ respectively. In the CNO-ν energy region of interest, the leading contributions in Borexino are instead provided by ^{7}Be-ν, ^{210}Bi and ^{11}C, as it can be seen from the full multivariate fit shown in Fig. 2.14.

  The expected CNO-ν signal/background ratio for several energy ranges is shown in Fig. 3.2; the signal is given by the CNO-ν rate, while the background is represented by all the other species present in the spectrum. The s/b ratio is reported as a function of the lower and the upper energy boundaries of the selected energy range, which are respectively the horizontal and vertical axes. The injected rates are given by the HZ-SSM predictions for neutrinos, and by the Phase-II results for radioactive backgrounds.

  Even considering the most favourable energy window $220 < N_p^{\mathrm{dt1-n}} < 418$ ($0.5\,\mathrm{MeV} \lesssim E \lesssim 1.1\,\mathrm{MeV}$ of the recoiled e^-), the CNO-ν contribution is marginal, since the signal/background ratio is s/b ≈ 0.28. Assuming instead a wider $160 < N_p^{\mathrm{dt1\text{-}n}} < 550$ energy window[1], which covers most of the CNO-ν energy spectrum ($0.37\,\mathrm{MeV} \lesssim E \lesssim 1.4\,\mathrm{MeV}$ of the recoiled e^-), one gets s/b < 0.1.

- **CNO-ν, ^{210}Bi and pep-ν spectral shape degeneracy.** Fig. 3.1 shows the simulated energy spectra of e^- scattered by CNO-ν, pep-ν and produced in the ^{210}Bi decay. The CNO-ν events energy shape in Borexino does not show any prominent structure but, on the contrary, a smooth featureless profile, similar to the one of ^{210}Bi background and pep-ν [2]. As will be described later on, this shape similarity prevents the multivariate fit to distinguish and isolate separately the interaction rates of the three contributions and, consequently, it drastically reduces the sensitivity to a CNO-ν signal.

[1] We remind that the $N_p^{\mathrm{dt1\text{-}n}}$ energy estimator counts the number of fired PMTs within 230 ns from the cluster beginning, normalized to a fixed number of 2000 working PMTs.

[2] Clearly, the overall energy spectrum is composed by many other contributions overlapped (pp-ν, ^{210}Po , ^{7}Be-ν, ^{85}Kr, ^{11}C, and external backgrounds; see e.g. the multivariate fit shown in Fig. 2.14), which are not reported in Fig. 3.1 for clarity, since they play a secondary role in the CNO-ν sensitivity subject.

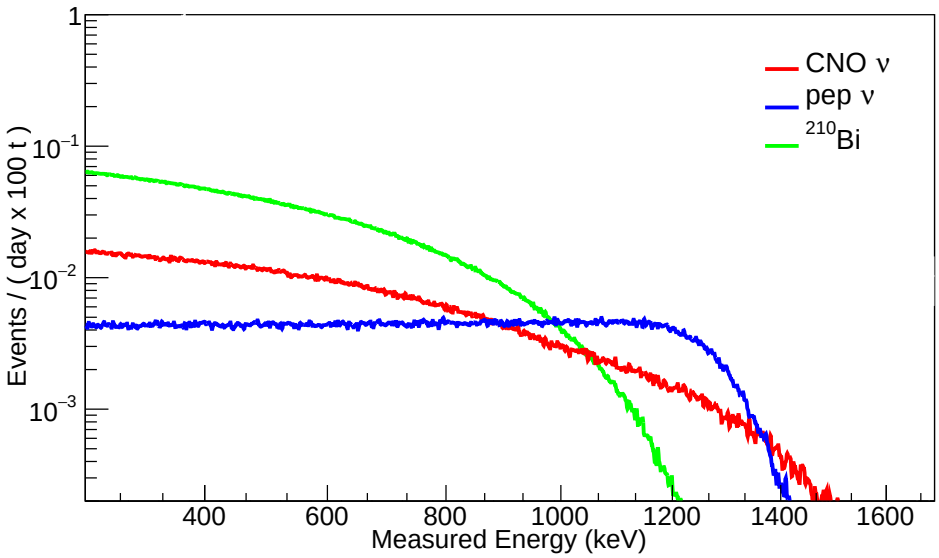


Figure 3.1: Simulated energy spectra of e^- scattered by CNO-ν, by pep-ν and produced in the ^{210}Bi decay (respectively in red line, blue line and green line). Injected rates for CNO-ν and pep-ν are matched to the HZ-SSM predictions, while the ^{210}Bi one is obtained from Phase-II analysis results [63, 39].

To evaluate in a more quantitative way the impact of the most relevant backgrounds on the Borexino sensitivity to CNO-ν, a thorough Monte Carlo study has been performed[3]: 10^4 datasets with an exposure of 900 days $\times$ 100t, close to the one we expect to reach in the dataset dedicated to the CNO-ν analysis, have been simulated. The contribution of CNO-ν signal and of the relevant backgrounds (^{210}Bi , pep-ν, ^{7}Be-ν, ^{11}C, and so on) has been included following the realistic relative proportions. Then the simulated data undergo the standard selection, to reduce the background on an event by event basis, and are then fitted according to the same multivariate analysis techniques that are used for the analysis of real data (described in Sec. 4.1 and Sec. 4.2 respectively).

The results of the multivariate fit for these 10^4 Monte Carlo pseudo-experiments are shown in Fig. 3.3 for CNO-ν, pep-ν and ^{210}Bi rates, as red histograms. The injected rates correspond to the HZ-SSM predictions for neutrinos, and to Phase-II result for ^{210}Bi background. The non-diagonal frames report the species correlation plots; the injected rates are marked by vertical and horizontal black lines. A strong correlation among the species is clearly visible from the non-diagonal plots, inducing a severe bias in the extracted rates for the three species. In particular, the reconstructed CNO-ν rate shows a very broad and biased distribution, with an high peak for the null value. Also the ^{210}Bi and pep-ν output distributions are significantly biased, being not centered on the injected values and strongly asymmetric. In other words, from a statistical point of view the analysis has no sensitivity to a CNO-ν signal, but only to the sum of the three CNO-ν, pep-ν and ^{210}Bi interaction rates.

The main conclusion of this study is that the CNO-ν, pep-ν and ^{210}Bi spectral shape similarity is the most critical factor preventing a straightforward CNO-ν detection in Borexino. The only way to break this correlation consists in the determination of the ^{210}Bi background and of the pep-ν rate

[3] See Chapter 5 for Monte Carlo simulation code details.

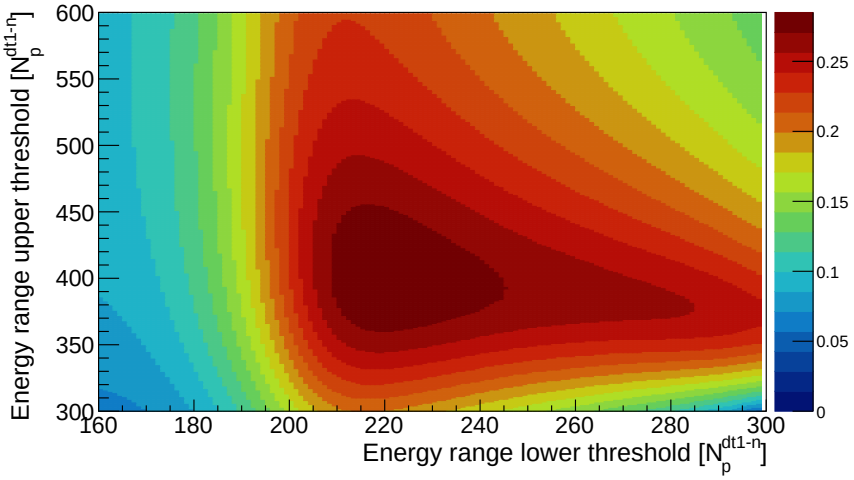


Figure 3.2: Expected CNO-ν signal/background ratio in color scale scanning several reconstructed energy ranges in the $N_p^{\mathrm{dt1-n}}$ variable. Horizontal and vertical axes represent the lower and upper boundarie of the energy range respectively. Signal is given by CNO-ν rate only, while background is given by all the other possible neutrino and background contributions. All the injected rates have been selected from Phase-II results [63, 39], with the exception of neutrinos which follow HZ-SSM prediction; in particular, $R^{\mathrm{HZ-SSM}}(\mathrm{CNO}\text{-}\nu) = 4.92 \pm 0.78\,\mathrm{cpd/100t}$.

independently from the fit, to constrain them in the multivariate analysis itself. The pep-ν rate can be safely constrained based on the Standard Solar Model predictions for the neutrino fluxes [25]. Along with the pp-ν reaction, it starts the pp chain sequence; its neutrino flux, and therefore the related interaction rate in Borexino, are estimated at the 1% level; accounting also for the error on the flavor oscillation parameters, the expected rate in Borexino is $R^{\mathrm{HZ-SSM}}(\mathrm{pep}\text{-}\nu) = 2.74 \pm 0.04\,\mathrm{cpd/100t}$ or $R^{\mathrm{LZ-SSM}}(\mathrm{pep}\text{-}\nu) = 2.78 \pm 0.04\,\mathrm{cpd/100t}$, depending on the metallicity hypothesis. On the contrary, the ^{210}Bi background rate is unknown and must be evaluated independently: the technique to tag the ^{210}Bi rate by means of its daughter isotope, ^{210}Po , will be discussed in detail in Chapter 6. In the following, we will show the impact of a constraint on ^{210}Bi with different precisions on the CNO-ν sensitivity.

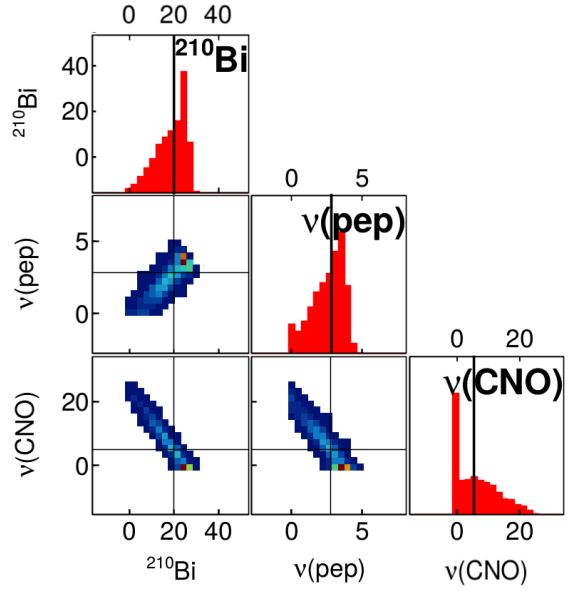


Figure 3.3: Results of the multivariate fits of 10^4 Monte Carlo pseudo-experiments without independent constraints to pep-ν and ^{210}Bi interaction rates. Red histograms represent the output interaction rates for the three components, while the non-diagonal frames are the species correlation plots. The injected rates for each dataset correspond to the HZ-SSM predictions for neutrinos and to Phase-II result for ^{210}Bi (17.5 cpd/100t), and are marked by vertical and horizontal black lines. The study included all the background and neutrino species, but only those components that mostly influence the sensitivity to CNO-ν are displayed. The strong correlations among these species forbid to measure the CNO-ν and pep-ν and ^{210}Bi decay rate in an unbiased way.

3.2 Borexino sensitivity to CNO-ν signal in case of pep-ν and ^{210}Bi constraints

In the previous section we have shown that Borexino has little or no sensitivity to CNO-ν signal due to the strong correlation between CNO-ν, ^{210}Bi and pep-ν spectral shapes. A possible way out to break this degeneracy, and therefore to recover this sensitivity, is to set independent constraints on ^{210}Bi and pep-ν interaction rates in the multivariate fits.

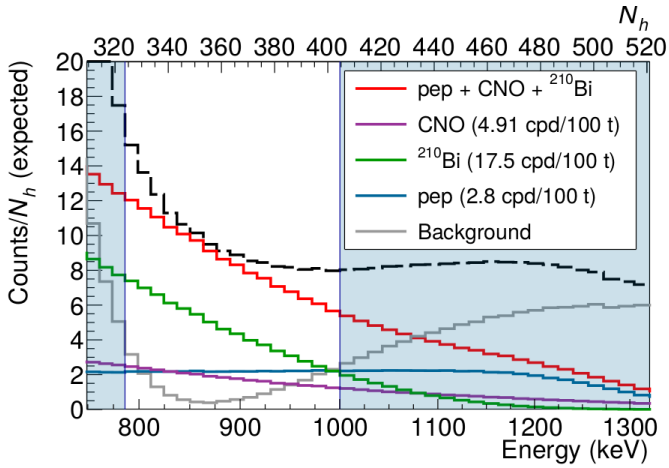


Figure 3.4: Expected Borexino spectrum assuming the rates predicted according to the the HZ-SSM for CNO-ν (purple profile) and pep-ν (blue profile), and injecting a rate of 17.5 cpd/100t for ^{210}Bi background (green profile). These three components are used for the simplified counting sensitivity analysis described in the text.

The impact of these constraints on Borexino CNO-ν sensitivity can be preliminary shown with a simple counting analysis of simulated events, in which the spectral shape information of signal and backgrounds is not considered. This simplified model allows to grasp roughly the required uncertainties on the constraints of background rates to achieve a CNO-ν evidence. The selected energy range, expressed in number of normalized hits, is $322 < N_h^n < 408$ (790 keV $\lesssim E \lesssim$ 1.0 MeV of the recoiled e^-), optimized with respect to the expected CNO-ν abundance, as shown in Fig. 3.4. This energy window is mainly populated by ^{210}Bi , CNO-ν, pep-ν and ^{7}Be-ν events; since this latter one plays a marginal role for the CNO-ν sensitivity and is also known with high precision, it will be neglected. For a given exposure $\mathcal{E}$, the total number of detected events $N_{\rm tot}$ is computed as:

$$N_{\rm tot} = \mathcal{E}\,(\varepsilon_{\rm CNO} R_{\rm CNO} + \varepsilon_{\rm pep} R_{\rm pep} + \varepsilon_{\rm Bi} R_{\rm Bi}) \tag{3.1}$$

where R_i is the interaction rate of the i-th component (i = CNO, pep, ^{210}Bi) and ε_i is the fraction of events for the i-th component into the energy window of interest. Supposing that ^{210}Bi and pep-ν rates are independently constrained, with measured values $\overline{R}_{\rm Bi} \pm \overline{\sigma}_{\rm Bi}$ and $\overline{R}_{\rm pep} \pm \overline{\sigma}_{\rm pep}$, the statistical uncertainty on CNO-ν $\sigma_{\rm CNO}$ can be isolated:

$$R_{\rm CNO} = \frac{N_{\rm tot}}{\mathcal{E}\cdot\varepsilon_{\rm CNO}} - \frac{\varepsilon_{\rm Bi}}{\varepsilon_{\rm CNO}}\overline{R}_{\rm Bi} - \frac{\varepsilon_{\rm pep}}{\varepsilon_{\rm CNO}}\overline{R}_{\rm pep} \tag{3.2}$$

$$\sigma_{\mathrm{CNO}} = \frac{1}{\mathcal{E} \cdot \varepsilon_{\mathrm{CNO}}} \sigma_{N_{\mathrm{tot}}} \oplus \frac{\varepsilon_{\mathrm{Bi}}}{\varepsilon_{\mathrm{CNO}}} \overline{\sigma}_{\mathrm{Bi}} \oplus \frac{\varepsilon_{\mathrm{pep}}}{\varepsilon_{\mathrm{CNO}}} \overline{\sigma}_{\mathrm{pep}} \tag{3.3}$$

where the $\oplus$ notation refers to the quadrature sum ($a \oplus b = \sqrt{a^2 + b^2}$). The colormap in the left panel of Fig. 3.5 shows σ_{CNO} as a function of $\overline{\sigma}_{\mathrm{pep}}$ and $\overline{\sigma}_{\mathrm{Bi}}$, according to Eq. (3.3). $\overline{\sigma}_{\mathrm{pep}}$ is varied between 0 cpd/100t and 0.4 cpd/100t ($\approx 14\%$), while $\overline{\sigma}_{\mathrm{Bi}}$ is varied between 0 cpd/100t and 5 cpd/100t ($\sim (30 - 40)\%$). A 3 years livetime data taking is assumed. The right panel of Fig. 3.5 shows σ_{CNO} as a function of $\overline{\sigma}_{\mathrm{Bi}}$, extracted projecting the left panel plot to two choices of $\overline{\sigma}_{\mathrm{pep}} = 0.28$ cpd/100t and $\overline{\sigma}_{\mathrm{pep}} = 0.028$ cpd/100t, corresponding respectively to relative uncertainties $\overline{\sigma}_{\mathrm{pep}}/\overline{R}_{\mathrm{pep}}$ of 10% (top panel) and 1% (bottom panel).

The pep-ν uncertainty is relevant only for the low $\overline{\sigma}_{\mathrm{Bi}}$ regime ($\overline{\sigma}_{\mathrm{Bi}} \lesssim 1.5$ cpd/100t), i.e. when ^{210}Bi rate is well-known, and is almost negligible otherwise. This behaviour can be observed from the iso-σ_{CNO} contours (white solid lines in left panel of Fig. 3.5) and from $\sigma_{\mathrm{CNO}}(\overline{\sigma}_{\mathrm{Bi}})$ trends (right panel of Fig. 3.5). The dependence of the Borexino sensitivity as a function of livetime is shown in Fig. 3.6 for several $\overline{\sigma}_{\mathrm{Bi}}$ values. For the CNO-ν analysis dataset, a total livetime of 3 years is reasonably expected; additional exposure would not improve significantly the sensitivity to CNO-ν signal. A lower $\sigma_{N_{\mathrm{tot}}}$, indeed, does not impact significantly on σ_{CNO} since the first term of Eq. (3.3) contributes only marginally to the overall budget for large exposures.

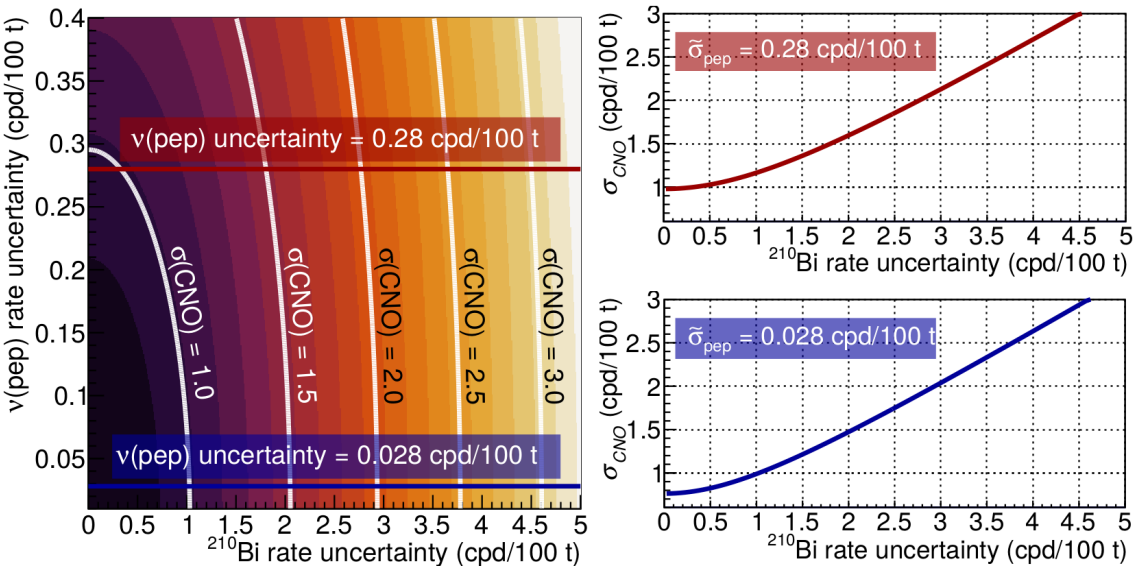


Figure 3.5: Borexino sensitivity to CNO-ν signal assuming pep-ν and ^{210}Bi independent constraints, from a simplified counting analysis. Left panel: expected uncertainty for the CNO-ν neutrino interaction rate σ_{CNO} as a function of the error in the determination of the ^{210}Bi and pep-ν interaction rates ($\overline{\sigma}_{\mathrm{Bi}}$ and $\overline{\sigma}_{\mathrm{pep}}$), according to the sensitivity counting analysis Eq. (3.3). Right panel: expected σ_{CNO} as a function of $\overline{\sigma}_{\mathrm{Bi}}$, projecting the left panel plot for fixed pep-ν constraint uncertainties $\overline{\sigma}_{\mathrm{pep}} = 0.28\,\mathrm{cpd/100t}$ and $\overline{\sigma}_{\mathrm{pep}} = 0.028\,\mathrm{cpd/100t}$, corresponding respectively to relative uncertainties $\overline{\sigma}_{\mathrm{pep}}/\overline{R}_{\mathrm{pep}}$ of 10% and 1%.

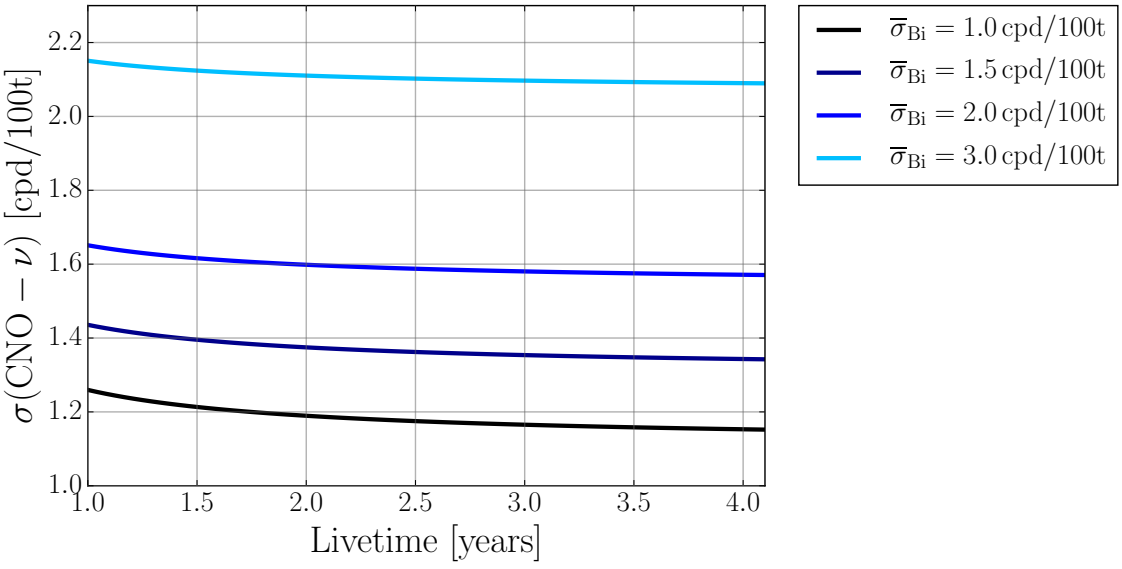


Figure 3.6: Borexino sensitivity to CNO-ν: dependence on data taking livetime. Expected σ_{CNO} as a function of livetime, in years, for several ^{210}Bi strength constraints: ($\overline{\sigma}_{\mathrm{Bi}} = 1.0\,\mathrm{cpd/100t}$, $\overline{\sigma}_{\mathrm{Bi}} = 1.5\,\mathrm{cpd/100t}$, $\overline{\sigma}_{\mathrm{Bi}} = 2.0\,\mathrm{cpd/100t}$, $\overline{\sigma}_{\mathrm{Bi}} = 3.0\,\mathrm{cpd/100t}$). A $\approx 1.4\%$ constraint has been assumed for pep-ν, i.e. $\overline{\sigma}_{\mathrm{pep}} = 0.04\,\mathrm{cpd/100t}$, according to the HZ-SSM and MSW predictions. The Fiducial Volume defined for the CNO-ν analysis has been considered.

Median discovery power as a function of independent constraint strenghts

A more rigorous analysis to establish the discovery power of Borexino as a function of the ^{210}Bi constraint strength [158] has been performed exploiting the 10^4 Monte Carlo pseudo-datasets described in Sec. 3.1; the multivariate fit has been now performed in the $132 < N_h^n < 937$ energy window, constraining ^{210}Bi and pep-ν rate with different precisions. The ^{210}Bi measurement information is implemented in the multivariate fit as an additional gaussian pull-term to the overall likelihood to be minimized (see Sec. 4.2).

In order to properly determine the discovery power to the CNO-ν signal, a frequentist hypothesis test has been performed using a profiled likelihood ratio q_0 as a test statistics [159]. The two considered hypotheses are:

- Null hypothesis H_0: only the background is present, i.e. $R(\text{CNO-}\nu) = 0\,\text{cpd/100t}$;
- Alternative hypothesis H_1: the CNO-ν signal is also present, i.e. $R(\text{CNO-}\nu) > 0\,\text{cpd/100t}$.

Since we want to perform a signal discovery analysis, the hypothesis to be tested is the null one H_0. The *median discovery power* of this measurement is defined as the median p-value obtained testing H_0 under the assumption of signal H_1. The employed test statistics q_0 is defined as follows:

$$q_0 = -2\left[\ln \mathcal{L}\left(0, \hat{\theta}\right) - \ln \mathcal{L}\left(\hat{\mu}, \hat{\theta}\right)\right] \tag{3.4}$$

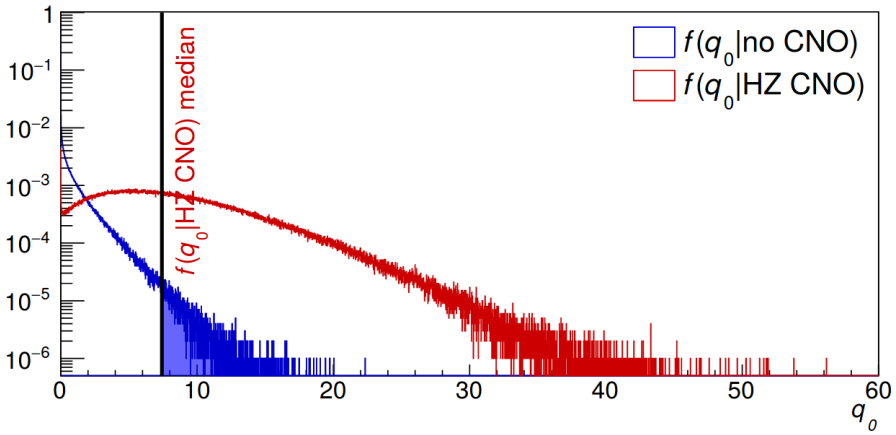


Figure 3.7: Example of the q_0 test statistic distributions obtained with toy MC experiments when the CNO-ν signal (HZ-SSM) is injected or absent in the data (red and blue distributions, respectively). The discovery power is the p-value corresponding to the median of $f(q_0|\text{HZ CNO})$ with respect to $f(q_0|\text{no CNO})$. In this case, ^{210}Bi rate has been constrained at 14% while pep-ν rate at 1%.

The first and the second term of Eq. (3.4) r.h.s are respectively the minimized negative log-likelihood (NLL) assuming $R(\text{CNO-}\nu) = 0\,\text{cpd/100t}$ and the absolute minimum of NLL leaving $R(\text{CNO-}\nu)$ as a free parameter. q_0 quantifies how well a model which assumes the absence of a CNO-ν signal describes the data: small values of q_0 indicates a good agreement, while large values of q_0 indicate a bad compatibility.

This test requires to obtain the q_0 PDFs for the background-only hypothesis $f(q_0|$ no CNO) and in case of the presence of a CNO-ν signal $f(q_0|$ HZ/LZ CNO), following Eq. (3.4). Both of them have been extracted considering two sets of 10^4 toy MC datasets: one set is created without injecting CNO-ν signal, while in the second ensemble the CNO-ν signal is injected according to the HZ-SSM predictions. In both cases, to compute q_0, each dataset has been fitted twice, with the same multivariate fit procedure that will be applied on the real data (see Sec. 4.2): once treating $R(\text{CNO-}\nu)$ as a free parameter, and the second time fixing it to zero. As the final step, the discovery power is computed as the p-value of $f(q_0|$ no CNO) corresponding to the median value of $f(q_0|$ HZ/LZ CNO), as shown in Fig. 3.7.

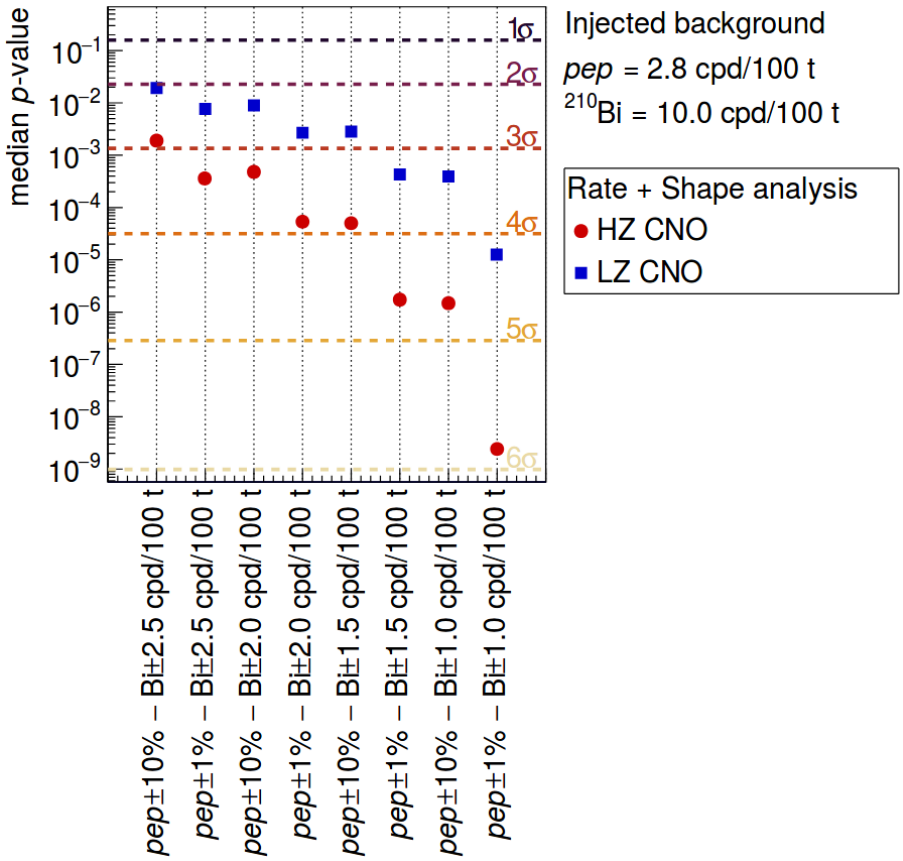


Figure 3.8: Median discovery power of a CNO neutrino signal, for the two metallicity scenarios HZ-SSM (blue points) and LZ-SSM (red squares), for different uncertainties on the ^{210}Bi and pep-ν constraint rates. The injected background rates are $R(\text{pep-}\nu) = 2.8\,\text{cpd/100t}$ and $R(^{210}\text{Bi}) = 10\,\text{cpd/100t}$.

The median discovery power to a CNO-ν signal resulting for different ^{210}Bi and pep-ν rates constraint strengths is shown in Fig. 3.8; results for the HZ-SSM and LZ-SSM scenarios are reported in blue circles or red squares respectively. The injected background rates are $R(\text{pep-}\nu) = 2.8\,\text{cpd/100t}$ and $R(^{210}\text{Bi}) = 10\,\text{cpd/100t}$. More severe constraints imply an higher sensitivity to CNO-ν signal, as expected; the ^{210}Bi statistical uncertainty plays the dominant role in the overall budget, and the pep-ν constraint strength gets more relevant for more precise ^{210}Bi determinations. This study

shows that, for the HZ scenario, an evidence for a CNO-ν signal can be claimed if $\overline{\sigma}_{\mathrm{Bi}} \lesssim 2\,\mathrm{cpd/100t}$ and, simultaneously, $\overline{\sigma}_{\mathrm{pep}} \sim 1\%$, with a median significance at $\sim 4\sigma$ level. In case of LZ scenario, the same constraint strength would lead to a $\sim (2.5-3)\sigma$ significance level.

Median discovery power for a ^{210}Bi upper limit

As we will see in Sec. 6.2, the most precise proposed method to determine the ^{210}Bi rate analyzing the ^{210}Po spatial distribution is not able to get a proper ^{210}Bi rate measurement, but only a ^{210}Bi upper limit. This is due to the fact that this method does not fully disentangle the ^{210}Po rate contribution in equilibrium with ^{210}Bi , because of the presence of an additional unsupported ^{210}Po term. An upper limit can be set since it has been proven that this out-of-equilibrium *migration term* induces only positive bias to the overall ^{210}Po term.

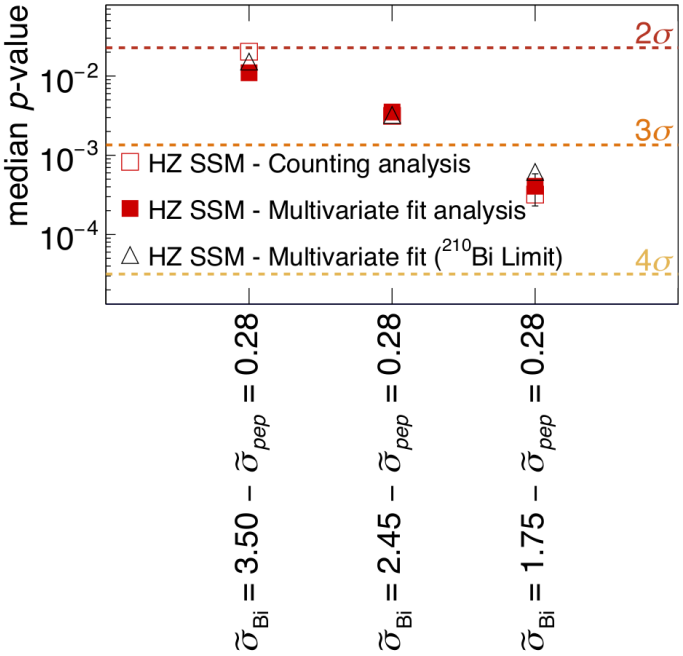
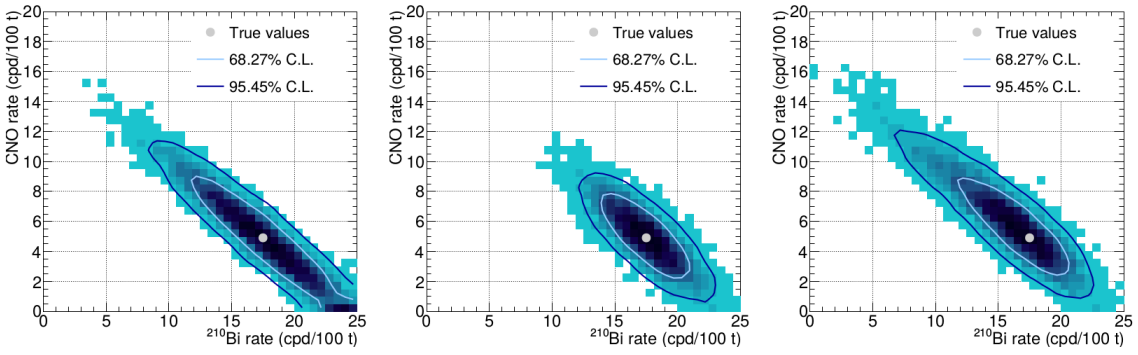


Figure 3.9: Median discovery power of a CNO neutrino signal in case of a ^{210}Bi upper limit as a function of ^{210}Bi and pep-ν constraint strengths, for a simple counting analysis (white squares), for a full analysis with ^{210}Bi gaussian penalty (red squares), and for a full analysis with ^{210}Bi upper limit (black triangles). The injected background rates are $R(\mathrm{pep}\text{-}\nu) = 2.8\,\mathrm{cpd/100t}$ and $R(^{210}\mathrm{Bi}) = 10\,\mathrm{cpd/100t}$.

Fig. 3.9 shows the median discovery power to a CNO-ν signal as a function of ^{210}Bi and pep-ν constraint strengths, for a simple counting analysis (white squares), for a full analysis with ^{210}Bi gaussian penalty (red squares), and for a full analysis with ^{210}Bi upper limit (black triangles).

The median discovery power for the ^{210}Bi upper limit case is found in agreement with the one obtained with a symmetric gaussian penalty: the CNO-ν sensitivity therefore is mostly unchanged. Analyzing the ^{210}Bi - CNO-ν contours in both cases (see Fig. 3.10), we can notice that, even if the sensitivity to the CNO-ν absence hypothesis is not worsened, the upper limit of the CNO-ν rate will not be stringent, since a lower limit on the ^{210}Bi background is not provided. The presence



(a) ^{210}Bi rate free (b) ^{210}Bi rate constrained (c) ^{210}Bi upper limit

Figure 3.10: Distribution of the ^{210}Bi and CNO-ν best fit estimates obtained fitting thousands of simulated datasets applying a 10% constraint on the pep-ν rate and leaving the ^{210}Bi free (a), constraining it with a 14% precision (b), and imposing an upper limit (c). Injected rates for CNO-ν and pep-ν are matched to the HZ-SSM predictions, while the ^{210}Bi one is obtained from Phase-II analysis results [63, 39] (17.5 cpd/100t).

of an asymmetric CNO-ν in this latter case will be verified in the Phase-III analysis described in Chapter 7. In conclusion, an upper limit on the ^{210}Bi rate will be sufficient to claim for an evidence of CNO-ν, but it will not allow to obtain a precise measurement of the CNO interaction rate that is needed to definitely settle the solar metallicity problem.

3.3 Strategy for the CNO-ν determination

The sensitivity studies discussed in the previous paragraph drive the strategy for the CNO-ν analysis. I summarize hereafter the crucial steps on which this analysis is based. In the following chapters, the details of these steps will be thoroughly described.

1. **Phase-III dataset.** The most important conclusion of the sensivity studies is that, in order to gain sensitivity to the CNO-ν signal, the ^{210}Bi contamination in the scintillation must be constrained precisely and independently from the multivariate analysis. For reasons that will be clear in Chapter 6, this is possible by means of the ^{210}Bi - ^{210}Po tag only if the detector is thermally stable. This condition has been reached only from the second half of 2016. For this reason, the CNO-ν analysis must be performed on a dedicated dataset starting from July 2016, which has been named Phase-III.

2. **Data selection.** Neutrino-induced events are intrinsically indistinguishable on an event-by-event basis from the most of backgrounds due to β or γ decays. Many selection cuts can be applied to partially remove several categories of backgrounds on an event-by event basis: muons and muon-induced events, noise events, radioactive decays from delayed coincidences. The fiducial volume (FV) cut selects events in a central scintillator region, highly reducing the external background rate. The cosmogenic events (^{11}C, ^{10}C, ^{6}He) can be identified via dedicated tagging techniques exploiting the decay peculiarities of their products. This data selection analysis has been originally designed for Phase-II analysis; throughout this thesis it will be applied to the Phase-III dataset defined above. My work on this topic will be discussed in Chapter 4.

3. **Multivariate analysis.** Even after the application of the selection cuts, background is still present. To disentangle the neutrino signal components from the background ones, a multivariate analysis is performed. It involves the simultaneous fit to the distributions of three physical quantities of interest: the event reconstructed energy, the radial and the β^+/β^- pulse-shape parameter distributions. The reference distributions for these three quantities are built using either analytical models or Monte Carlo simulations. Then, the model is fitted against data to extract the interaction rates of each background and neutrino signal. The multivariate analysis has been originally designed for the Phase-II dataset, and has been then improved and optimized for the Phase-III one. My work on this topics will be reviewed in Chapter 4.

4. **Monte Carlo simulations** are crucial to obtain an accurate description of the detector response, on which the multivariate fit is based. The Borexino Geant4-based Monte Carlo code was tuned in 2009 thanks to detector calibrations, and found to be in excellent agreement with data at the $\sim 1\%$ level until 2016. After this date, the simulations showed some issues, especially in reproducing properly the energy response of the detector. This is a problem notably for the CNO-ν analysis, which is based on data starting from July 2016 (Phase-III). One of the main task of my work was therefore to solve this issue in order to have Monte Carlo simulations to be used for the CNO-ν analysis, in particular for the multivariate fit. This topic will be reviewed in Chapter 5.

5. ^{210}Bi **independent constraint.** The studies described in Sec. 3.1 and Sec. 3.2 show that Borexino can be sensitive to a CNO-ν signal only if the rate of the ^{210}Bi background is

constrained in the multivariate fit. In order to do so, we exploit the fact that the ^{210}Bi rate can be determined independently by measuring the rate of its daughter nuclide, the ^{210}Po . This strategy is complicated by the non-straightforward time and spatial ^{210}Po distribution in the scintillator. Also, it relies on the assumption that the ^{210}Bi is uniform and stable in time. My work on this topic will be reviewed in Chapter 6.

The results of these steps will be combined in Chapter 7 to give a first result on CNO-ν determination. The first indication of a CNO-ν signal will be illustrated there, including a first preliminary study of the related systematic effects.

Chapter 4

Data selection and multivariate fit for the CNO neutrinos analysis

The success of the Borexino CNO-ν rate determination relies on accurate strategies for the discrimination of the neutrino signal from the background events. Indeed, neutrino-induced events are intrinsically indistinguishable on an event-by-event basis from the most of the β or γ backgrounds. A set of event-by-event selection cuts aiming to remove the taggable background (radioactive decays from delayed coincidences, muons and muon-induced events, cosmogenic ^{11}C events) is the first step for the CNO neutrinos analysis.

Unfortunately, even after the selection cuts, residual background is still present on data. For this reason, in order to disentangle the neutrino signal from it, a multivariate fit technique must be applied; the reference distributions of the quantities of interest (energy estimators, radial position, e^+/e^- discriminator variable) are built according to analytical models or Monte Carlo simulations, and then are fitted against data.

This Chapter is organized as follows. Sec. 4.1 reviews the data selection strategy for the CNO neutrinos analysis, discussing the basic quality cuts (Sec. 4.1.1), the Fiducial Volume definition (Sec. 4.1.2), the techniques for the cosmogenic ^{11}C identification (Sec. 4.1.3). The Phase-III-TMP dataset, which will be analyzed in this thesis for the CNO-ν determination, is defined in Sec. 4.1.4. The effect of selection cuts on the Borexino Phase-III-TMP energy spectrum, along with my work on the *stability plots*, is shown in Sec. 4.1.5. The multivariate fit strategy is introduced in Sec. 4.2, along with the analytical and Monte Carlo approaches for the construction of reference PDFs (Sec. 4.2.1 and Sec. 4.2.2 respectively). As a preliminary crosscheck, Sec. 4.3 reports the results of multivariate fit I have obtained analyzing the Phase-III-TMP dataset, for both the fit approaches and with CNO-ν rate fixed, in the same configuration fit used for the pp chain neutrino analysis [39]. Conclusions are outlined in Sec. 4.4.

4.1 Data selection for the CNO-ν analysis

In the following, the data selection flow employed for the CNO-ν analysis is described in detail. This sequence of selection cuts has been originally designed and optimized for Phase-II analysis, to extract the pp chain neutrinos fluxes [63]; in principle, no further modifications are necessary

for Phase-III-TMP dataset, which is fully dedicated to the CNO-ν analysis (see Sec. 4.1.4). I have carefully analyzed the impact of these selection cuts on the new data, studying the selection stability as a function of time, and finding no evidences for anomalies or issues (see Sec. 4.1.5).

4.1.1 Standard cuts

The *standard cuts* are a set of selection conditions for background reduction, optimized for the solar neutrino low energy analyses. They aim to identify and exclude the taggable backgrounds, due to muons and muon-induced events, radioactive decays from delayed coincidences, noise events [65, 63, 137].

- **Muon and muon-daughter removal.** Muons are the only relevant charged component of cosmic radiation penetrating the rock shielding over the LNGS halls. Thanks to the Gran Sasso mountain shielding, the muon surface flux is reduced to $1.2\,\mathrm{m}^{-2}\mathrm{h}^{-1}$, with a suppression factor $\sim 10^{-6}$ with respect to the flux at the Earth surface. Muon events are excluded from the analysis dataset thanks to the combined information coming from the Inner Detector and the Outer Detector (described in Sec. 2.4), achieving a $> 99.9\%$ tagging efficiency (see Sec. 2.3.3). When a muon is identified, a 300 ms veto is applied over the data taking, in order to exclude the cosmogenic unstable isotopes produced by muons.

- **Single energy deposit.** Only single cluster events are accepted. We recall that clusters are the group of photons belonging to the same scintillator event, as described in Sec. 2.4.1. In case the clustering algorithm identifies 0-cluster events or multiple clusters, the event is discarded.

- **No coincidence events.** Delayed coincidences separated by a time interval longer than the acquisition gate width are detected as two independent events. All events reconstructed with mutual distance smaller than 1.5 m and taking place within a 2 ms time window are rejected. Monte Carlo dedicated tests showed that this cut removes the ^{214}Bi $-^{214}$ Po decay delayed coincidences with high efficiency.

- **Charge estimator quality control.** The reconstructed charge estimator is compared to the number of fired PMTs. In case of unreasonable charge excess or deficit, the event is discarded. This cut allows to drop a large fraction of events induced by the electronic noise.

- **Additional quality controls for noise suppression.** One of the common sources of noise events during data taking is the electronic noise from a single electronics rack. If the total fraction of hits that are recorded on the most active crate exceeds the 75% of all hits, the event is dropped.

The efficiency of this sequence of selection cuts has been evaluated extensively both on Monte Carlo simulations and analyzing the radioactive-source calibration data. The fraction of neutrino events discarded by these cuts is found to be at the $\lesssim 10^{-4}$ level, well below the required precision level for the solar neutrino analyses and therefore negligible [63].

4.1.2 Fiducial Volume

One of the data selection main goal is the suppression of the external background γ events, which are mainly due to PMTs radioactive contamination: indeed, the external background rate is exponentially increasing going from the center towards the Inner Vessel. For this reason, all the neutrino analyses are restricted within a software-defined central portion of the scintillator, named *fiducial volume* (FV), in which the signal over background ratio is maximized. The FV definition is possible thanks to the accurate performance of the Borexino position reconstruction ($\approx 12\,\mathrm{cm}$ coordinate uncertainty at the detector center and at 1 MeV energy), which allows to locate an event inside or outside a defined FV with relatively high precision. Note that a systematic error on the FV definition would introduce a bias on the target mass and therefore on the final neutrino interaction rates. Dedicated studies based on the detector calibration runs show that, with the current position reconstruction performance, the systematic error on the fiducial mass is $\lesssim 1\%$.

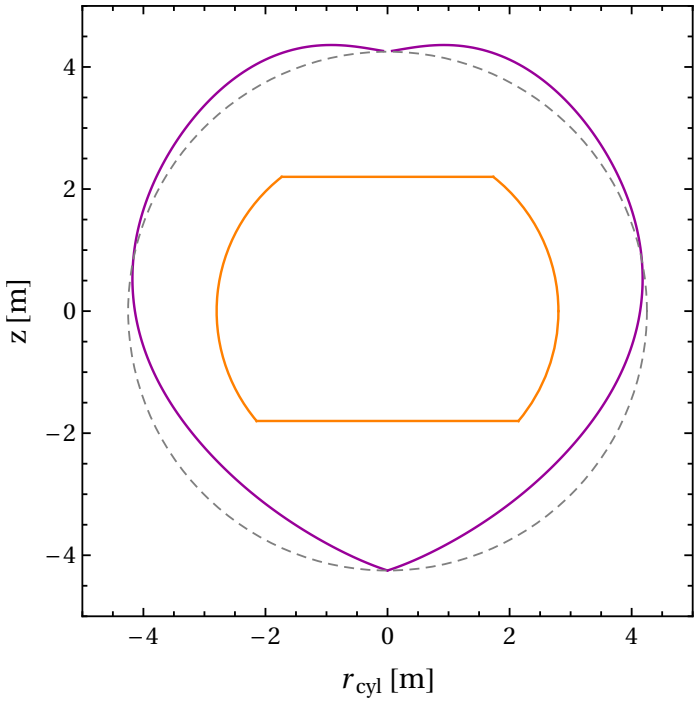


Figure 4.1: Fiducial Volume for the CNO-ν analysis in the $r_{\mathrm{cyl}} - z$ plane (orange solid line), comparing to a recent Inner Vessel shape (first week of September 2019, in purple solid line) and to a spherical ideal Inner Vessel shape (grey dashed line). $r_{\mathrm{cyl}} = \sqrt{x^2 + y^2}$ is the cylindric radius.

The FV definition is strictly connected to the data analysis to be performed. For what concerns the CNO-ν analysis, we use the same FV selection used in the Phase-II analysis, which follows the conditions $r < 2.8\,\mathrm{m}$, $-1.8\,\mathrm{m} < z < 2.2\,\mathrm{m}$ [39] (see the $r_{\mathrm{cyl}} - z$ plane sketch in Fig. 4.1). The defined FV contains a 71.3 ton scintillator fiducial mass. The z cut is asymmetric because of the non-spherical IV shape, which is more deformed in the South region; the polar regions of the sphere are excluded to suppress as much as possible the external background coming from the end-cap mechanical supports located at the top and bottom of the IV [65].

4.1.3 ^{11}C suppression

The selection cuts described above have no effect on the cosmogenic isotope ^{11}C, which still dominates the $1.1\,\text{MeV} \lesssim E \lesssim 1.8\,\text{MeV}$ region of the Borexino energy spectrum. ^{11}C is produced in-situ by muons, undergoing a β^+ decay with a $\tau_{^{11}\text{C}} = 29.4\,\text{min}$ lifetime. The ^{11}C rate in Borexino is $\approx 28\,\text{cpd/100t}$, to be compared with the CNO-ν and pep-ν signals which are expected to be $\sim 5\,\text{cpd/100t}$ and $2.7\,\text{cpd/100t}$ respectively. If not identified through specific techniques, its presence spoils unavoidably the pep-ν and CNO-ν rate measurements.

As discussed in Sec. 2.3.3, ^{11}C events are continuously produced by muons through spallation processes on the scintillator ^{12}C (Eq. (4.1)), therefore the interaction rate in Borexino cannot be reduced by purification techniques. Its production is accompanied by a prompt neutron, which thermalizes and then is captured on hydrogen[1]; the newly generated $^2\text{H}^*$ then de-excites to the fundamental state emitting a delayed 2230 keV γ event. Summarizing:

$$\mu + {}^{12}\text{C} \rightarrow {}^{11}\text{C} + n \implies \begin{cases} {}^{11}\text{C} \rightarrow {}^{11}\text{B} + e^+ + \nu_e & \tau_{^{11}\text{C}} = 29.4\,\text{min} \\ n + {}^1\text{H} \rightarrow {}^2\text{H}^* \rightarrow {}^2\text{H} + \gamma & \tau_n = 256\,\mu\text{s} \end{cases} \tag{4.1}$$

The ^{11}C candidate events can be identified in two complementary ways. The main tagging technique is named Three-Fold coincidence (TFC), and exploits the time and space concurrence of muon and neutron events in order to exclude the subsequent ^{11}C decays [65, 160]. The e^+/e^- pulse-shape discrimination variable, described later on, instead relies on the peculiar time features of the electron and positron classes of events.

Three-Fold Coincidence

The Three-Fold Coincidence is an algorithm for the cosmogenic event reduction based on the key idea of vetoing space-time regions of the detector after μ-n coincidences, such that the following β^+ decay of ^{11}C is excluded. In most cases, a cylindrical veto along the parent muon track with 80 cm radius is applied (see left panel of Fig. 4.2), for a two-hours time window. For the rarer case of high neutron multiplicity from the ^{11}C spallation, a two hours veto is applied over the whole FV [65]. The geometry and the time window for the region to be vetoed have been optimized, and are the result of a balance between the ^{11}C rejection power and the remaining exposure.

The TFC algorithm has been recently improved by the collaboration [39], implementing a likelihood $\mathcal{L}_{\text{TFC}}$ that an event is identified as a ^{11}C, evaluating the distance in space and time from the parent muon, the distance from the neutron, the neutron multiplicity, and the muon loss energy density $\text{d}E/\text{d}x$. This upgrade guarantees a ^{11}C tagging efficiency of $(92 \pm 4)\%$, keeping the $(64.28 \pm 0.01)\%$ of the total exposure.

An example of $\log(\mathcal{L}_{\text{TFC}})$ distribution as a function of the reconstructed energy, expressed in the $N_p^{\text{dt1-n}}$ estimator[2], is reported in right panel of Fig. 4.2. The green-dashed horizontal line represents the $\mathcal{L}_{\text{TFC}}$ threshold used for the selection. The high selection performance of the TFC can be also noticed by eye: the large majority of the ^{11}C events, which are included in the $350 \lesssim N_p^{\text{dt1-n}} \lesssim 700$ energy range, lies above this threshold.

[1] Neutron capture by ^{12}C produces a 4.95 MeV γ but the probability is small ($\approx 1\%$) compared to the dominant hydrogen capture case (($\approx 99\%$).

[2] We remind that the $N_p^{\text{dt1-n}}$ energy estimator counts the number of fired PMTs within 230 ns from the cluster beginning, normalized to a fixed number of 2000 working PMTs.

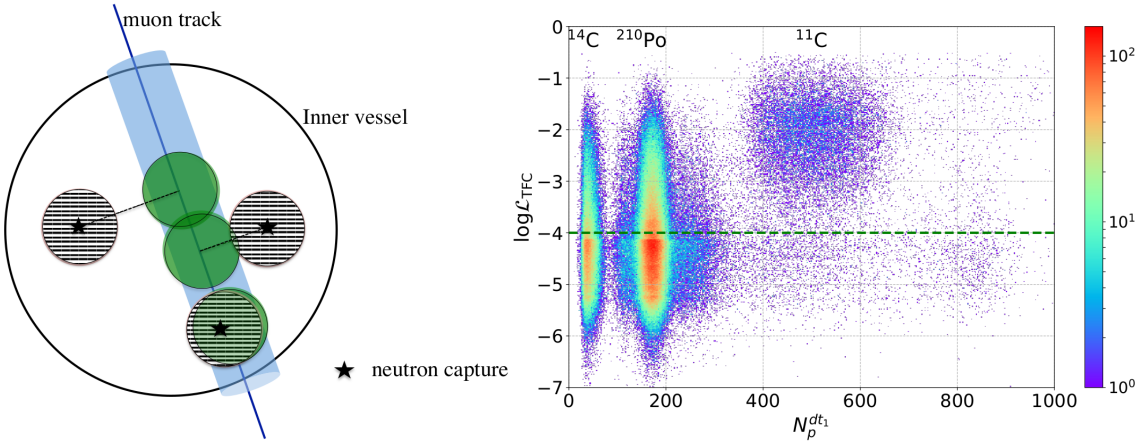


Figure 4.2: Three-Fold Coincidence technique. The spatial regions vetoed in the TFC method are shown in left panel: a cylinder around the muon track (blue) and some examples of spheres centered around the point where the de-excitation γ following the neutron capture is reconstructed (areas with horizontal lines around the stars) and their projections along the muon track (green areas) [65]. Right panel: example of $\log(\mathcal{L}_{\mathrm{TFC}})$ distribution as a function of the reconstructed energy, expressed in $N_p^{\mathrm{dt1-n}}$ estimator. The green-dashed horizontal line represents the $\mathcal{L}_{\mathrm{TFC}}$ threshold used for the selection, above/below which the events are assigned to the TFC-tagged/subtracted energy spectrum. It is clearly visible that the majority of the events of the ^{11}C energy decay spectrum lies above this threshold.

Based on the $\mathcal{L}_{\mathrm{TFC}}$ threshold, the events are assigned to two samples: one depleted (TFC-subtracted) and one enriched (TFC-tagged) in ^{11}C. These two sets are simultaneously fitted in the multivariate analysis procedure (see Sec. 4.2), in order to exploit all the data information. The impact of TFC algorithm on the reconstructed energy spectrum can be seen in Fig. 4.5.

e^+/e^- discrimination

The different processes involved in the energy deposition for e^+ and e^- are exploited to disentangle the β^+ and β^- classes of events (see also Sec. 2.4.1). In approximately 53% of cases, the emitted positrons form an orthopositronium state which decays in a non-negligible lifetime of ≈ 3.1 ns. Moreover, the γ emitted as a consequence of the e^+e^- annihilation deposit their energy in different spots, such that the event cannot be considered pointlike. As a consequence, the time profile of photons emitted in e^+ events is different with respect to the one of e^- ones.

We have studied dedicated estimators able to discriminate between positron-like and electron-like events. The pulse shape discriminator parameter used in this thesis (PS-$\mathcal{L}_{\mathrm{PR}}$) is based on the maximum likelihood value in output to the position reconstruction algorithm PS-$\mathcal{L}_{\mathrm{PR}} = \max_{(\vec{r},t)} \mathcal{L}_{\mathrm{pos}}(\vec{r},t)$ (Eq. 2.20) [39]. An example of the PS-$\mathcal{L}_{\mathrm{PR}}$ parameter distributions for e^+ and e^- samples is shown in Fig. 4.3.

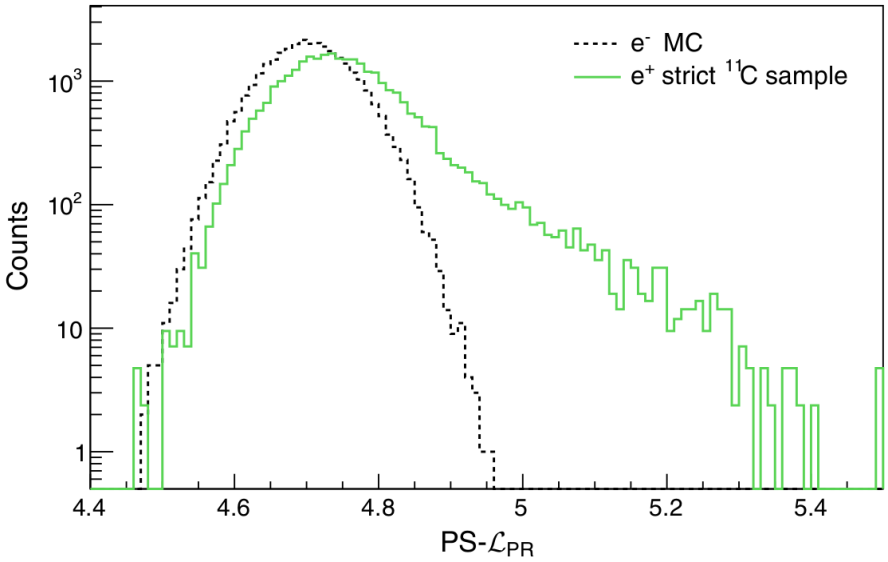


Figure 4.3: e^+/e^- identification. Comparison of the distributions of the PS-$\mathcal{L}_{\mathrm{PR}}$ variable for Monte Carlo generated e^- events (green, continuous line) and for e^+ events selected from the data (black, dashed line). The latter category of events is a high-purity ^{11}C sample, obtained with the optimized TFC method (see Sec. 4.1.3), using very strict selections on the energy and on the time correlation with the neutron and muon tracks.

4.1.4 The Phase-III-TMP dataset

As already outlined in Chapter 3, the CNO-ν analysis must be performed in the so-called Phase-III dataset, which started in July 2016 and is still ongoing. The reasons for this choice are connected to the possibility of determining the ^{210}Bi rate via the so-called ^{210}Bi -^{210}Po tag (see Chapter 6), to constrain it in the multivariate fit. In particular, this method requires the uniformity of ^{210}Bi distribution within the FV and in the same time that no external source of ^{210}Po is present.

- As it will be shown later in Sec. 6.3, after the scintillation purification, which took place in 2010, the ^{210}Bi contamination was not uniform within the FV: the cleanest scintillator portion was found at the FV top. It took several years for the complete mixing to take place.
- The proposed strategy to determine the ^{210}Bi rate relies on the hypothesis that the contamination of ^{210}Po coming from the IV into the FV is minimized. Scintillator convective motions can detach ^{210}Po from the IV and bring it inside the FV: in order to attenuate them, being strongly correlated with the detector temperature, Borexino has been thermally insulated starting from 2015. Only in 2016 the effects of the insulation started to be evident and therefore only the dataset starting from July 2016 is considered.

The Phase-III dataset is still open and Borexino is still taking data. The work of this thesis is based on a subset of the Phase-III dataset, and will be referred to as *Phase-III-TMP*, which ranges from 17 July 2016 to 02 June 2019. Of course, as more data will be available, they will be included in the final analysis which will be officially released. The Phase-III-TMP cumulative exposure trend is shown in Fig. 4.4, reaching a final 644.24 days $\times$ 100 ton value; the energy spectrum in

$85 \leq N_p^{\mathrm{dt1-n}} \leq 900$ range ($0.19\,\mathrm{MeV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$) is reported in Fig. 4.5 and will be described in the next Sec. 4.1.5.

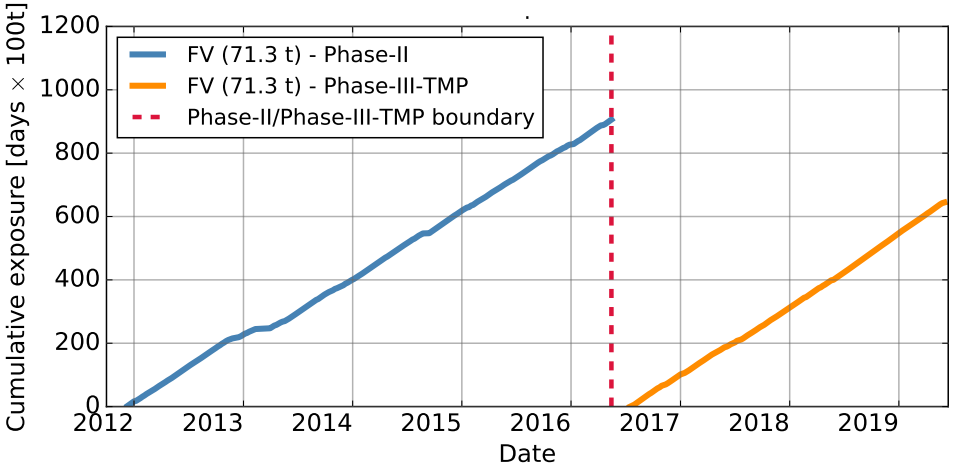


Figure 4.4: Cumulative exposure in days $\times$ 100 t collected by Borexino during the Phase-II and Phase-III-TMP data-taking periods, in the FV defined for the CNO-ν analysis. The red dashed line marks the separation between the two Phases.

4.1.5 Effects of selection cuts on the Borexino Phase-III-TMP data

The effect of the basic cuts, of FV selection, and of ^{11}C events identification by means of the TFC algorithm on the Borexino energy spectrum is shown in Fig. 4.5. The raw energy spectrum from the Phase-III-TMP dataset is displayed as a blue histogram, in the normalized $N_p^{\mathrm{dt1-n}}$ reconstructed energy variable. The standard cuts effect is shown in the violet histogram: the muon and muon-daughter cuts remove the higher energy part of the raw spectrum, while the consequences of the other quality cuts described in Sec. 4.1.1 are not visible by eye. The FV selection (red histogram) allows to increase drastically the signal over background ratio, removing a large part of the external background and highlighting some spectral features. The application of the TFC algorithm allows to discard a relevant portion of ^{11}C (orange histogram, TFC-subtracted spectrum). Some of the contributions populating the Borexino energy spectrum can be identified even by eye. The ^{14}C dominates the low energy region up to $N_p^{\mathrm{dt1-n}} \approx 120$; the ^{210}Po peak, which represents the most abundant background after ^{14}C, culminates at $N_p^{\mathrm{dt1-n}} \approx 180$. The ^{7}Be-$\nu$ Compton-like shoulder is evident by eye in the $280 \lesssim N_p^{\mathrm{dt1-n}} \lesssim 350$ range, while the ^{11}C dominates the $380 \lesssim N_p^{\mathrm{dt1-n}} \lesssim 700$ region.

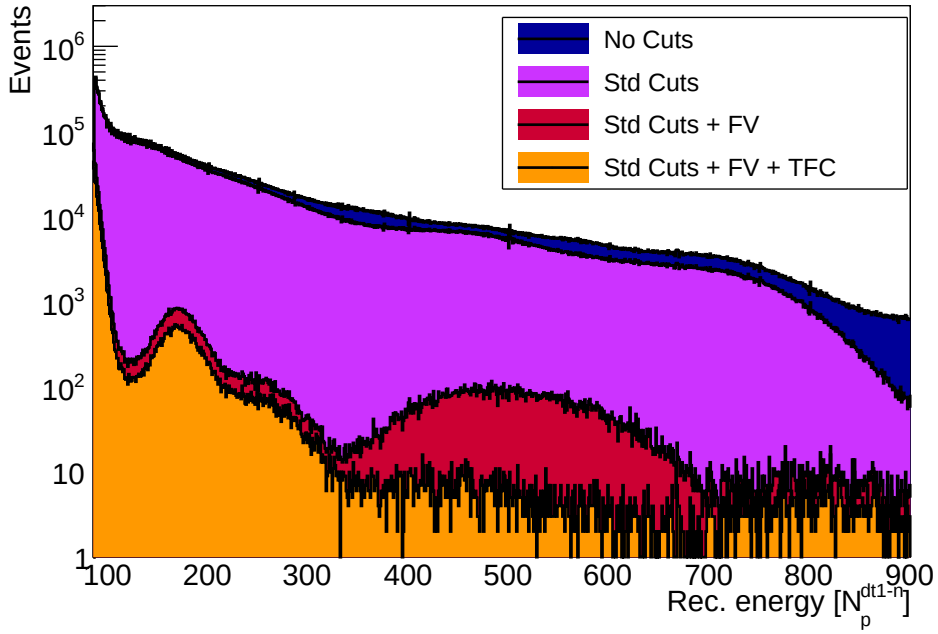


Figure 4.5: Effects of standard cuts, FV selection, and TFC algorithm application, on the Borexino raw energy spectrum. The raw energy spectrum from the Phase-III-TMP dataset, in the normalized $N_p^{\mathrm{dt1-n}}$ variable, is displayed as a blue histogram. The effect of standard cuts is shown in the violet histogram: the muon and muon-daughter cuts remove the higher energy part of the raw spectrum, while the consequences of the other basic cuts described in Sec. 4.1.1 are not visible by eye. The FV selection (red histogram) allows to increase drastically the signal over background ratio, highlighting some spectral features. The application of the TFC algorithm, described in Sec. 4.1.3, allows to discard a relevant contribution of ^{11}C (orange histogram, TFC-subtracted spectrum). The ^{14}C dominates the low energy region up to $N_p^{\mathrm{dt1-n}} \approx 120$; the ^{210}Po peak, which is the most abundant background in Borexino after the ^{14}C, culminates around $N_p^{\mathrm{dt1-n}} \approx 180$. The ^{7}Be-$\nu$ Compton-like shoulder ($280 \lesssim N_p^{\mathrm{dt1-n}} \approx 350$) is evident by eye, while the ^{11}C spectrum covers the $380 \lesssim N_p^{\mathrm{dt1-n}} \lesssim 700$ region.

The rate evolution in Phase-III-TMP for two selected energy windows in the FV are shown in Fig. 4.6 (*stability plots*), with the aim of monitoring and checking the detector stability. Top panel shows the rate in the $85 \leq N_p^{\mathrm{dt1-n}} \leq 100$ energy window ($0.19\,\mathrm{MeV} \lesssim E \lesssim 0.24\,\mathrm{MeV}$), which exhibits a linear increase in time: this is due to the PMTs loss, which causes a worsening in the energy resolution. Therefore, the ^{14}C energy distribution, which peaks at $N_p^{\mathrm{dt1-n}} \approx 40$, is broadened in time and a larger and larger fraction of events is moving towards higher energies, leaking into the $85 \leq N_p^{\mathrm{dt1-n}} \leq 100$ window. As we will see in Sec. 4.3.1 and Sec. 4.3.2, $N_p^{\mathrm{dt1-n}} = 85$ corresponds to the lower energy threshold for the multivariate fit range. The bottom panel shows the rate in the $100 \leq N_p^{\mathrm{dt1-n}} \leq 140$ energy window ($0.24\,\mathrm{MeV} \lesssim E \lesssim 0.33\,\mathrm{MeV}$), which is found quite stable in time within the uncertainty bars. A second order resolution effect is present due to the ^{14}C leakage, as for the $85 \leq N_p^{\mathrm{dt1-n}} \leq 100$ window, inducing a neglectable events increase. Overall, no particular issues has been found in the Phase-III-TMP dataset which will be used for the CNO-ν analysis.

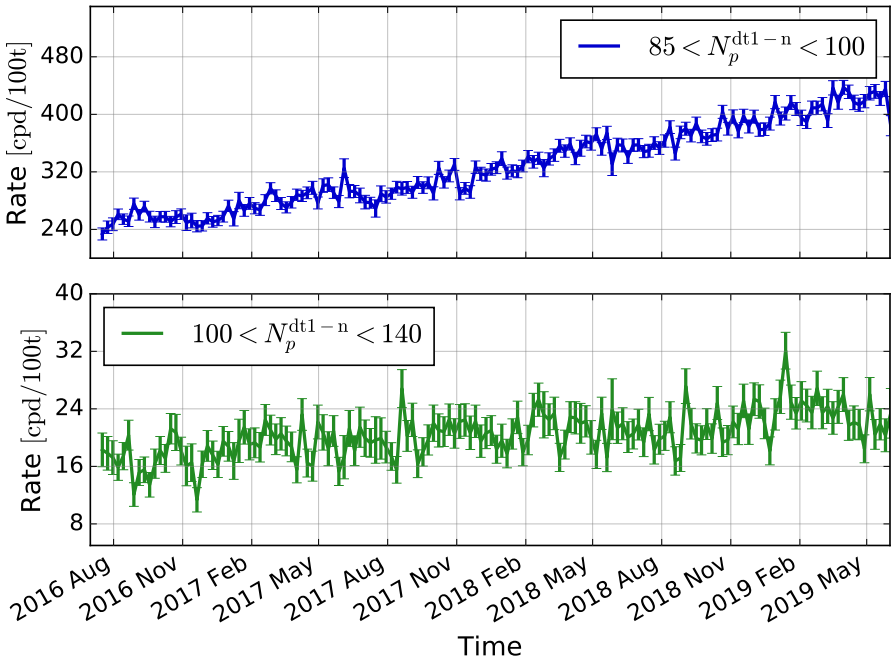


Figure 4.6: Rate evolution in Phase-III-TMP for three selected energy windows (*stability plots*) in the FV: $85 \leq N_p^{\mathrm{dt1-n}} \leq 100$, i.e. $0.19\,\mathrm{MeV} \lesssim E \lesssim 0.24\,\mathrm{MeV}$ (top panel) and $100 \leq N_p^{\mathrm{dt1-n}} \leq 140$, i.e. $0.24\,\mathrm{MeV} \lesssim E \lesssim 0.33\,\mathrm{MeV}$ (bottom panel).

4.2 Multivariate fit for the CNO-ν analysis

In the following, the multivariate analysis technique adopted in the CNO-ν analysis to disentangle the signal and the backgrounds will be discussed in details. The technique, based on the frequentist statistics, relies on the simultaneous fit of the distributions of three physical quantities: the energy of the scattered electrons, the radial position and e^+/e^- discrimination estimator. These distributions are built on an event-by-event basis for those events passing the selection cuts (see Sec. 4.1) and are fitted as the sum of signal and backgrounds components (named also *species*). The interaction rate for each of the species is a free parameter of the fit.

The energy spectrum is the most powerful key to separate the signal and the background, especially in case of distinctive features. For example, as seen in the Fig. 4.5 spectrum the electron scattered by the mono-energetic ^{7}Be-ν displays a characteristic Compton-like shoulder; ^{210}Po background, which undergoes an α decay, is easily recognizable by its gaussian shape. On the contrary, the identification of species with featureless energy profiles, as the CNO-ν or ^{210}Bi , is way less straightforward. The radial distribution contributes to distinguish events originating from interactions that take place uniformly in the scintillator volume, which display a characteristic $\sim r^2$ behaviour, from the external backgrounds, which show a decreasing exponential trend approaching to the detector center. The e^+/e^- pulse-shape estimator is useful to disentangle the ^{11}C background, which decays β^+, from all the other β^--like events lying in the same energy window. This separation is critical and delicate especially for the determination of CNO-ν and pep-ν interaction rates.

The reference distributions corresponding to each signal and background components (*probability*

density function, or *PDF*) are obtained both analytically and via Monte Carlo simulations. The fit is performed by using both the PDFs approaches, providing an important cross-check of systematics related to the detector response function modeling, as will be described in Sec. 4.2.1 and Sec. 4.2.2 respectively.

Data are fitted through the maximization of a multi-dimensional binned likelihood function[3]. The overall likelihood function $\mathcal{L}(\vec{k}|\vec{\theta})$ is built as the product of four independent terms: the reconstructed energy distribution for the TFC-tagged dataset ($\mathcal{L}^{\mathrm{TFC}}_{\mathrm{tag}}$), the reconstructed energy distribution for the TFC-subtracted datasets ($\mathcal{L}^{\mathrm{TFC}}_{\mathrm{sub}}$), the radial distribution ($\mathcal{L}_{\mathrm{Rad}}$) and the pulse shape e^+/e^- parameter distribution ($L_{\mathrm{PS}-\mathcal{L}_{\mathrm{PR}}}$) [63]:

$$\mathcal{L}(\vec{k}|\vec{\theta}) = \mathcal{L}^{\mathrm{TFC}}_{\mathrm{sub}}(\vec{k}|\vec{\theta}) \cdot \mathcal{L}^{\mathrm{TFC}}_{\mathrm{tag}}(\vec{k}|\vec{\theta}) \cdot \mathcal{L}_{\mathrm{Rad}}(\vec{k}|\vec{\theta}) \cdot \mathcal{L}_{\mathrm{PS}-\mathcal{L}_{\mathrm{PR}}}(\vec{k}|\vec{\theta}) \tag{4.2}$$

The symbol $\vec{\theta}$ indicates the set of the arguments with respect to which the function is maximized (e.g. the signal and background interaction rates) and $\vec{k}$ generically indicates the set of the experimental data used to evaluate the likelihood. Each of the four terms represent a standard Poisson binned likelihood function:

$$\mathcal{L}_i(\vec{k}|\vec{\theta}) = \prod_{i=1}^{N} \frac{\lambda_i^{k_i}(\vec{\theta})}{k_i!} \mathrm{e}^{-\lambda_i(\vec{\theta})} \tag{4.3}$$

where the product is calculated over the N bins of the histogram associated to the selected distributions. The two TFC-subtracted and TFC-tagged spectra are fitted keeping the rates of the most part of the components in common, apart from the cosmogenic backgrounds ^{11}C itself, ^{6}He and ^{10}C (see Tab. 2.2), and also ^{210}Po , which may be not distributed in a perfectly homogeneous way throughout the FV.

Additional constraints on some of the fit parameters can be applied in the form of multiplicative Gaussian *pull terms* to the overall likelihood Eq. 4.2 (*penalty method*):

$$\mathcal{L}(\vec{k}|\vec{\theta}) \rightarrow \mathcal{L}(\vec{k}|\vec{\theta}) \cdot \exp\left\{ \frac{1}{2} \left(\frac{x - x_m}{\sigma} \right)^2 \right\} \tag{4.4}$$

where $x_m \pm \sigma$ is the independently estimated value of the selected parameter. This procedure represents an intermediate way between fixing and freeing of a parameter, allowing to include in the multivariate fit the information coming from an independent measurement.

Another possibility for the additional constraint consists in setting an *upper limit* for a given interaction rate. This is the relevant case for the ^{210}Bi background, since the analysis in Chapter 6 will show that only an upper limit for ^{210}Bi rate can be inferred independently on the multivariate fit. In this case, the constraint is applied by means of an asymmetric half-gaussian distribution, obtained considering Eq. (4.4) distribution only for $x > x_m$. This will be a crucial element for the CNO-ν analysis (see Chapter 7), in which a constraint to ^{210}Bi is needed to gain sensitivity on the CNO-ν signal.

The multivariate analysis is performed by means of the statistical framework analysis named *GooStats*, developed, validated and benchmarked internally to the Borexino collaboration [161]. The

[3]The employment of a likelihood instead of a simpler χ^2-based fit allows to treat properly those energy spectrum bins with low statistics and that are consequently found in the Poisson regime.

use of Graphical Processing Units (GPU) have been developed and applied to the fit optimization, critically reducing the computation time by about two orders of magnitude if compared with the previous standard CPU-based algorithms. Without this improvement, the feasibility of the analysis would have been threatened. The key point of the GPU method is that the computations of the expected events for each multidimensional bin can be treated independently: the thousand of cores inside the GPU can perform the calculations separately, highly decreasing the fitting time. This advantage is especially critical for the analytical fit, as it will be shown later on, because of the larger number of free parameters involved which naturally slows down the computation.

We perform the multivariate analysis following two approaches. In the analytical approach, the reference shapes for all the signals and backgrounds to be used in the fit are obtained by an analytical model. In the Monte Carlo approach, the reference shapes are extracted by means of accurate Monte Carlo simulations. In the following two paragraphs, I will describe the main ideas beyond each one of the two approaches, comparing advantages and limitations. It is important to note that the two methods are, to some extent, independent, and that the associated systematic uncertainties are different. This double possibility for the multivariate fit provides an important crosscheck for the analysis results.

4.2.1 Analytical fit

In this approach, the *energy response function* $f(N_p; E)$ used to convert the energy released in the scintillator E into the reconstructed energy estimator N_p, is derived analytically. In principle, the analytical energy response function depends on a complex chain of parameters: the position where the interaction takes place, the particle type, the scintillator and material properties, the detector geometry, the electronics response features, the detector evolution and instabilities in time. In a more practical way, since it is impossible to obtain an explicit analytical dependence taking into account all these parameters, the adopted models make use of approximations and integration over several of the listed parameters.

To illustrate the analytical model, we select the energy estimator N_p as reference. The estimator used in the analytical fit in this thesis, $N_p^{\mathrm{dt1-n}}$, is mainly a slight variant of N_p (see Sec. 2.4.1) [4]. The expected reconstructed spectrum $\dfrac{\mathrm{d}N}{\mathrm{d}E}(N_p)$ for a given species can be obtained by convoluting the correspondent theoretical energy spectra $\dfrac{\mathrm{d}N}{\mathrm{d}E}(E)$ with the energy response function $f(N_p; E)$:

$$\frac{\mathrm{d}N}{\mathrm{d}E}(N_p) = \int_{\mathrm{E_{min}}}^{\mathrm{E_{max}}} f(N_p; E)\ \frac{\mathrm{d}N}{\mathrm{d}E}(E)\ \mathrm{d}E \tag{4.5}$$

where $\dfrac{\mathrm{d}N}{\mathrm{d}E}(E)$ is well-known for all the involved species, since it only depends on the interaction cross section and on the event kinematics. The shape of the response function is generally characterized by its central moments, mainly the mean and variance, and is modeled as a scaled Poisson function:

$$f(N_p; E) = \frac{m^{[N_p s(E)]}}{[N_p(E)\ s]!}\mathrm{e}^{-m} \qquad m = \frac{[N_p(E)]^2}{\sigma_p^2(E)} \qquad s = \frac{\sigma_p^2(E)}{N_p(E)} \tag{4.6}$$

[4] Extensions for other estimators, defined in Sec. 2.4.1, follow similar descriptions and can be found at Ref. [65]

where the parameters m and s depend on the mean value $N_p(E)$ and variance $\sigma_p^2(E)$. The relation between the mean number of collected photoelectrons $N_{pe}(E)$ and the event energy E, is given by the sum of scintillation and Čerenkov contributions:

$$N_{pe}(E) = Y_0 \cdot [Q(E) \cdot E + f_{\rm Ch} \cdot F_{\rm Ch}(E)] \tag{4.7}$$

where Y_0 [p.e./MeV] is the photoelectron yield at the detector center, $Q(E)$ is the quenching factor, $F_{\rm Ch}(E)$ describes the Čerenkov light dependence on energy for electrons, $f_{\rm Ch}$ weights the relative importance of scintillation and Čerenkov light emissions. The mean number of $N_p(E)$ can be now derived as [65]:

$$N_p(E) = N_{\rm tot}\left[1 - {\rm e}^{-\mu}\left(1 + p_t\mu\right)\right]\left(1 - g_C\mu\right) \qquad \mu = \frac{N_{pe}(E)}{N_{\rm tot}} \tag{4.8}$$

g_C is named geometrical correction factor, which takes into account in an effective way the impact of non-homogeneities in light collection to the N_p mean value; it is usually computed analyzing pure e^- Monte Carlo samples for several energies, simulated throughout the whole dataset of interest. p_t is the fraction of a single photoelectron charge spectrum below the electronics threshold, and it is calculated from early calibration measurements.

The N_p variance instead ($\sigma_p^2(E)$) can be modeled separately for α and β events:

$$\sigma^2_{p\alpha}(E) = f_{\rm eq}\left[1 - (1 + v_1)p_1\right] N_p(E) + v_T^{\alpha} N_p^3(E) \tag{4.9}$$

$$\sigma^2_{p\beta}(E) = f_{\rm eq}\left[1 - (1 + v_1)p_1\right] N_p(E) + v_0^{\alpha} N_p^3(E) \tag{4.10}$$

$$+ v_T^q \left(\mu\frac{p_0}{p_1}\right)^2 N_p^2(E) + v_N\, N_p(E) \tag{4.11}$$

where the introduced parameters are:

- $f_{\rm eq}$: equalization factor, i.e. the ratio between number of active PMTs and the $N_{\rm tot} = 2000$ PMTs reference configuration number;
- v_1: relative variance of PMT triggering probability;
- $p_1 = 1 - {\rm e}^{-\mu}$: probability that a specific PMT is fired;
- $p_0 = e^{-\mu}$: probability that a specific PMT is not fired;
- v_T^{α}: width of ^{210}Po peak;
- v_T^0: related to the spatial non-uniformity of the number of fired PMTs;
- v_T^q: related to the non-uniformity of light collection;
- v_N: intrinsic resolution parameter for β events.

Most of the parameters listed above are determined from calibrations and are fixed in the fit thanks to calibration runs or independent lab measurements (quenching-related parameters, k_β, $f_{\rm Ch}$, p_t, g_C,

v_1, v_T^q). However, some of them are left free to vary in the fit procedure: the light yield parameter Y_0, which drives the energy scale, and the energy resolution parameters v_T^0, v_N and v_T^α.

The reference shapes built with the analytical approach, in the $N_p^{\mathrm{dt1-n}}$ reconstructed energy variable, are shown for neutrinos and backgrounds in Fig. 4.7. The four neutrino shapes (pp-ν, pep-ν, ^{7}Be-ν, CNO-ν) are shown in light blue lines. To build these shapes, the detector response parameters described in Sec. 4.2.1 have been fixed to Phase-II analysis values. The lower energy threshold set for the histograms is $N_p^{\mathrm{dt1-n}} = 85$ ($E \approx 0.19\,\mathrm{MeV}$), matching the lower energy threshold for the Phase-III-TMP analytical fit.

The reference shape for pile-up events must be computed separately. We remind that pile-up events are the result of two uncorrelated events that fall in the same trigger acquisition window so close in time that they cannot be identified separately, and are therefore detected as a single event. In the analytical approach, the pile-up contribution is calculated for each of the species involved in the fit. The most relevant components of the pile-up spectrum are:

- ^{14}C overlapped with another ^{14}C;
- ^{14}C overlapped with ^{210}Po ;
- ^{14}C overlapped with external γ;

where the external γ are events without enough energy to induce the trigger condition because they lost most of their energy in the buffer. Note that ^{14}C is always involved in the pile-up process because of its dominant interaction rate in Borexino ($\approx 40\,\mathrm{Bq/100t}$), if compared to all the other species. In particular, the ^{14}C-^{14}C couple dominates the pile-up spectrum in the $160\,\mathrm{keV} \lesssim E \lesssim 265\,\mathrm{keV}$ range.

Since it is impossible to build the pile-up reference shape in a full analytical way, two possible semi-empirical data-driven methods are employed [137, 63]:

- *Convolution method.* In this approach, the pile-up spectrum is given by the mathematical convolution of spectral shape of each species with the spectrum of random events. The typical spectrum for random events is obtained from a dedicated trigger, and is acquired during regular data taking with a 0.5 Hz frequency [39].
- *Synthethic method.* The algorithm for the synthetic pileup production considers the hits recorded within a time interval of duration dti (230 ns for dt1 variables) in the second half of the acquisition gate[5] and superimposes them with the primary cluster, which caused the trigger to occur and is placed at the beginning of the DAQ gate. These synthetic events are then processed with the standard reconstruction code. The pileup spectrum is defined as the set of synthetic events whose energy estimator is increased by at least E_{min}, with respect to the corresponding original event, after the overlapping of the additional hits [162, 39]. An example of synthetic pile-up spectrum can be seen in Fig. 4.8.

In summary, in the analytical fit, in which the reference shapes are built by means of an analytical model, the rates of each signal and background components are left free to vary in the fit, together

[5] It is worth to remind that for every triggered event, we register all the hits in a time window of 16.5 μs, which is much longer than the cluster duration (typically around 1 μs).

with the light yield and the detector resolution parameters. The possibility to adapt the energy response makes the fit able to finely adjust on data, especially in case of physical changes of the detector response. On the other hand, due to the larger number of free parameters and to the possible correlations between them, this approach shows less stability with respect to the Monte Carlo one.

4.2.2 Monte Carlo fit

In the Monte Carlo approach, the energy reference shapes are built by the optimized and tuned simulation code developed by the Borexino collaboration, described in Chapter 5. For each signal and background species, events are generated through massive simulations following the detector evolution in time. The simulated events then undergo the same selection cuts applied to the real dataset, and are processed by the Echidna reconstruction code. The reference PDFs obtained by these simulations are then properly normalized.

All the species, with the exceptions of γ external background and ^{14}C, are simulated in a spherical volume with 3.75 m radius. The chosen number of generated events is proportional to the run effective livetime and in the same time is set to largely exceed the number of detected events on data: 50 times higher for external backgrounds, 200 times higher for the all the other species except for ^{14}C. In this way, the statistical uncertainty of the Monte Carlo simulations (and PDFs) can be neglected. The ^{14}C case must be treated separately because of its extremely high statistics, $40\,\mathrm{Bq/100t} = 3.46 \cdot 10^5\,\mathrm{cpd/100t}$. Therefore, since simulations matching the real ^{14}C abundance are unfeasible, the commonly simulated ^{14}C PDFs are built with $\sim 1/3$ of the real data statistics, resulting in a worse statistical error with respect to data. This criticity is handled considering a scaled poisson distribution in the likelihood computation for very low energy bins [39, 163]: the data statistics is scaled to the Monte Carlo one, with an artificial increasing of the errors, implying a weakened importance of the likelihood computation for this energy region.

The simulation of γ external background due to PMT radioactivity requires a different strategy, since the tracking from PMTs until the inner FV is extremely time consuming. A dedicate *importance sampling algorithm* have been employed, with the aim to optimize the computational efficiency of the simulation. More details about the events generation can be found in Ref. [162, 163].

Except for the external background, the only event type which may show significant non-homogeneities in the scintillator is ^{210}Po ; in Chapter 6 the spatial distribution and time stability of ^{210}Po will be extensively discussed. In order to take into account second order effects in the detector response due to non-homogeneities, ^{210}Po is simulated following its real spatial distribution and time evolution extracted by data. We remind that ^{210}Po is the only relevant α-decaying isotope and that its decay events can be finely isolated by means of a α/β discrimination variable cut (see Sec. 2.4.1).

The simulation of pile-up is performed superimposing Monte Carlo events belonging to different categories: two events are overlapped, introducing an artificial time correlation [162]. The time correlation is typically created by sampling a random delay between the events in a selectable range (typically equal to a few μs). The output created in this way is then processed through the standard electronics simulation and reconstruction code. The pile-up spectrum is simulated only for the three most relevant contributions, given by the combination of $^{14}\mathrm{C} - {}^{14}\mathrm{C}$ events, $^{14}\mathrm{C} - {}^{210}\mathrm{Po}$ events or ^{14}C−external background events. An example of Monte Carlo pile-up spectrum can be seen in Fig. 4.8.

In conclusion, the Monte Carlo fit employs reference PDFs produced by the Borexino Monte Carlo code that will be described in detail in Chapter 5, along with my work on its upgrade. The only parameters left free to vary are the neutrino and background rates, since the detector response (energy scale, resolution effects, non-uniformities and non-linearities) is implicitely taken into account by the Monte Carlo simulations. This makes the Monte Carlo fit more stable then the analytical one, thanks to the lower number of free parameters, but in the same time less flexible to adjust to non-predicted variations of the detector response.

Exceptions to the complete fixing of the detector response parameters are given by the α (^{210}Po) and ^{11}C events energy scales: these two parameters are left free to vary such that fine-tunings at the $\sim 1\%$ level can take place. In the first case, the fine scale adjustment is motivated by the large statistics in ^{210}Po in real dataset, higher than the one of calibration runs available for the tuning of the α response parameters [6]. The natural consequence is that it is not possible to achieve a 0.5% uncertainty (corresponding to 1 hit over 200) for the ^{210}Po energy peak, which is the most crucial Borexino background after ^{14}C. For what concerns ^{11}C, it is the only β^+ event present in the Borexino spectrum: the slightly different energy deposition processes involved in the formation of ortho-positronium justifies this small energy scale adjustments.

[6] Especially for the earlier years of Phase-II.

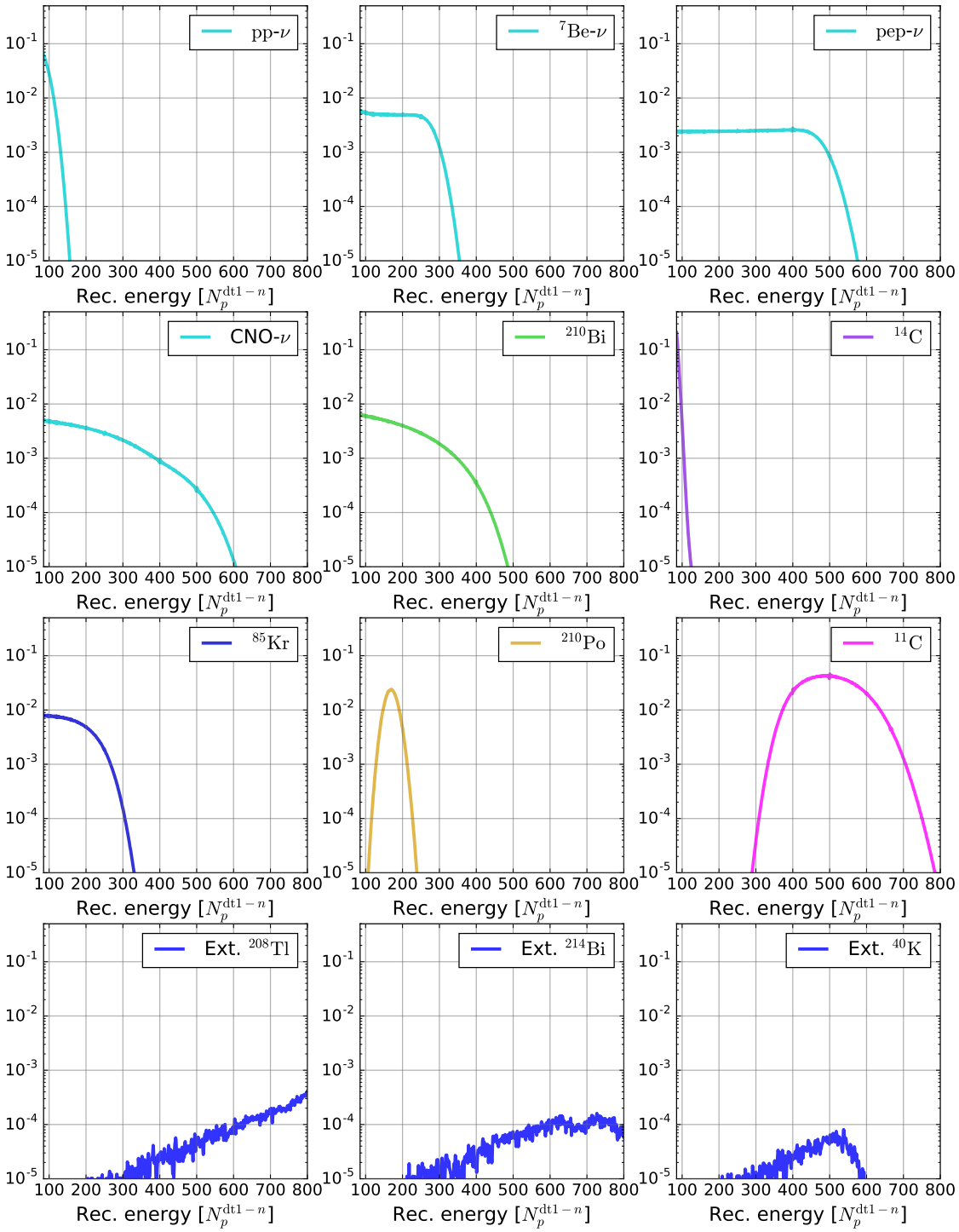


Figure 4.7: Examples of energy reference shapes built with the analytical approach in the $N_p^{\mathrm{dt1-n}}$ reconstructed energy variable. The four neutrino shapes (pp-ν, pep-ν, ^{7}Be-ν, CNO-ν) are shown in light blue lines, while the background species in other colors lines. In this case, the detector response parameters described in Sec. 4.2.1 have been fixed to the Phase-II analysis values. In the analytical fit, these detector response parameters are left free to vary.

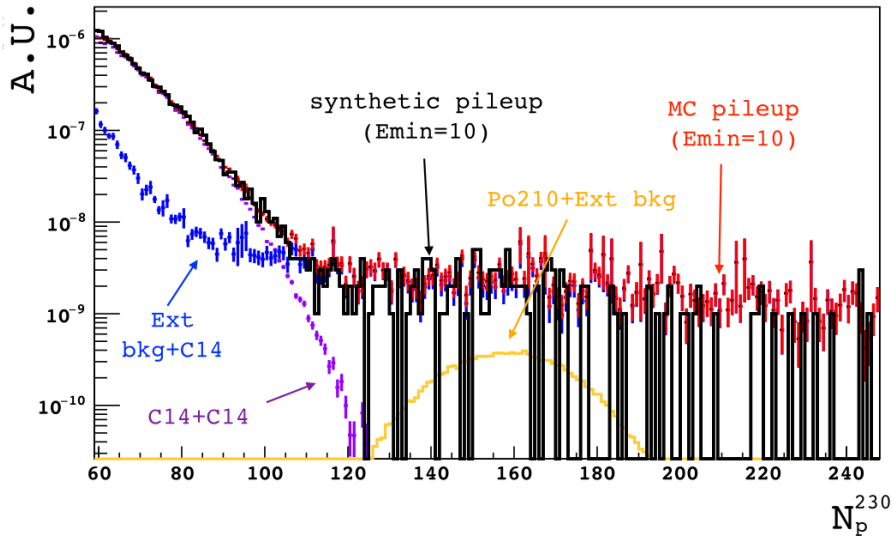


Figure 4.8: Examples of **synthetic** (black histogram) and **Monte Carlo pile-up spectra**.

4.3 Multivariate fit for Phase-III-TMP dataset with fixed CNO-ν rate

In the following, the results of Phase-III-TMP dataset multivariate analysis are shown in the so called *pp chain configuration*, for the determination of interaction rates of pp chain neutrinos. This fit configuration has been extensively used for the Phase-II analysis, leading to the simultaneous determination of all the pp chain solar neutrino rates [39, 63]. Throughout this Chapter, it will be mainly used as a preliminary crosscheck for Phase-III-TMP dataset before performing the fit in the *CNO configuration*, dedicated to the search of CNO-ν signal, which will be employed in the final Chapter 7.

In the pp chain configuration, all the interaction rates from pp chain neutrinos are left free to vary in the fit (with the exception of ^{8}B-ν, which is fixed to HER Phase-II analysis result), while the CNO-ν rate is fixed according to the HZ-SSM predictions $R^{\mathrm{HZ-SSM}}(\mathrm{CNO\text{-}}\nu) = 4.92 \pm 0.78\,\mathrm{cpd/100t}$. The ^{14}C background is constrained by a gaussian 40 ± 1 Bq pull term, according to Eq. (4.4).

4.3.1 Analytical fit in Phase-III-TMP with CNO-ν rate fixed

The results of the analytical multivariate fit, performed in the *pp chain* configuration on Phase-III-TMP dataset, are shown in Fig. 4.9. The fit is performed over the $85 \leq N_h^{\mathrm{n}} \leq 950$ reconstructed energy range ($0.19\,\mathrm{MeV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-), including also the lowest energy part of the spectrum, dominated by ^{14}C and pp-ν events. Energy spectra for the TFC-subtracted (left) and the TFC-tagged (right), along with residuals (bottom panels) are shown in Fig. 4.9. The sum of the individual components from the fit (black lines) is superimposed on the data (grey points).

The interaction rates for the species are reported in Tab. 4.1. The results of the fit are good, as quantified by a low $\chi^2/\mathrm{NDF} = 1.21$, in the same line with the one obtained for the Phase-II analysis published in Ref. [39]. The residuals don't show particular structures nor anomalies. Most of the rates extracted by the fit are consistent with the ones obtained in the previous Phase-II analysis.

However, there is an issue related to the pp-ν rate, which is the most relevant contribution at low energy left free to vary. The resulting rate $R^{\mathrm{Bx\text{-}III}}(\mathrm{pp\text{-}}\nu) = 174.2 \pm 12.7\,\mathrm{cpd/100t}$ is significantly discrepant with respect both to the SSM predictions and to the Borexino Phase-II best estimate (respectively, $R^{\mathrm{Bx\text{-}II}}(\mathrm{pp\text{-}}\nu) = 134.6 \pm 10^{+6}_{-10}\,\mathrm{cpd/100t}$ [39] and $R^{\mathrm{SSM\text{-}HZ}}(\mathrm{pp\text{-}}\nu) = 132.2 \pm 1.4\,\mathrm{cpd/100t}$). This discrepancy is an indication that something is going on in the low energy region of the spectrum, for reasons to be understood.

Sources for this low energy anomaly are not fully understood yet, and the analysis team is strongly focusing on it. I have investigated some of the possible explanations:

- **Detector response function.** A possible reason of the anomalous result described above could be researched in the loss of resolution due to the decreasing number of working PMTs as a function of time. I recall in fact that Borexino is losing working PMTs since the beginning of its data taking in 2007, with a nearly constant rate of ~ 70 PMTs/year. If the analytical model does not reproduce correctly the impact of the PMT loss on the detector response as a function of time, this may introduce a bias in the output of the fit. In order to test this hypothesis, I tried to vary some of the response function parameters that were fixed in the fit, as explained in Sec. 4.2.1, like g_C (which takes into account the uniformity of the

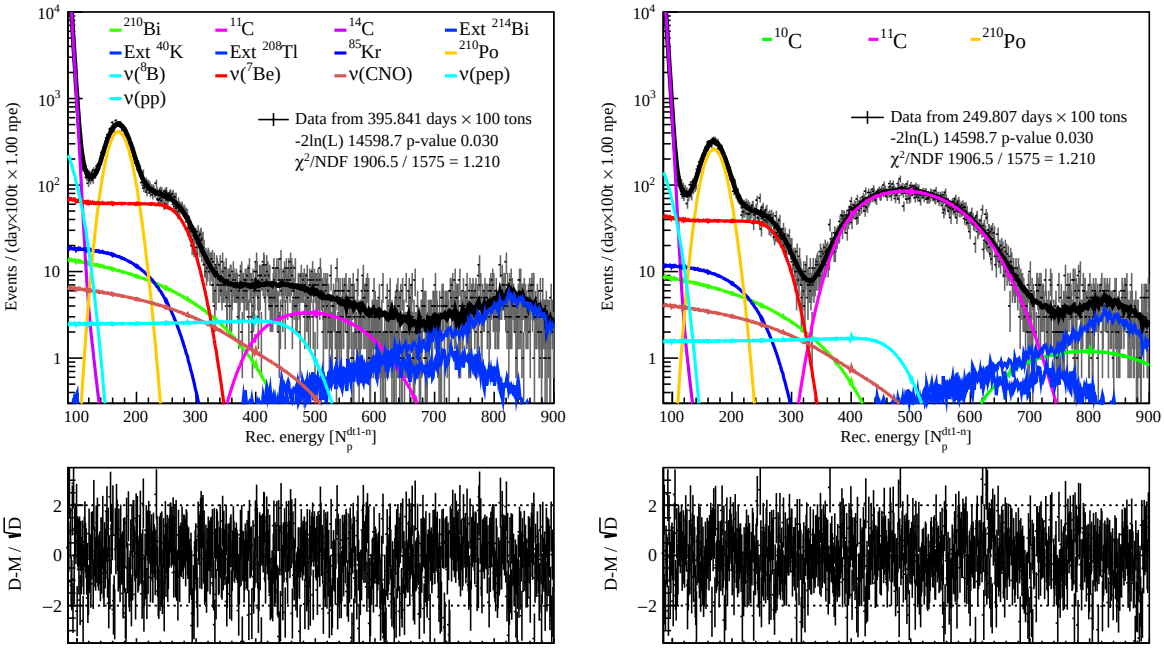


Figure 4.9: Analytical fit for Phase-III-TMP dataset with CNO-ν rate fixed: TFC-subtracted (left) and TFC-tagged (right) energy spectra (top panels), with residuals (bottom panels). The sum of the individual components from the fit (black lines) is superimposed on the data (grey points). The analysis has been performed within the $85 \leq N_p^{\mathrm{dt1-n}} \leq 900$ energy window ($0.19\,\mathrm{MeV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-). The residuals are calculated in every bin as the difference between the data counts and the fit result, divided by the square root of the data counts. The interaction rates are reported in Tab. 4.1.

response function) or f_{Ch} (which is related to the Čerenkov light contribution to the total number of photons). The result of the fit are robust with respect to small variations of these parameters. Note that I varied these parameters within the limit allowed by the model and by the calibrations data. In conclusion, the pp-ν rate anomaly is not solved by small fine-tunings of the energy response function.

- **Pile-up modeling approach.** I investigated the possibility that the anomaly could be due to the pile-up convolution method which, for some unknown reasons (possibly again related to the loss of active PMTs in time), could be not performing correctly in the latest years. In order to check this hypothesis, I performed the fit using the synthethic pile-up method, instead of convolution one. I recall that this is a second method to treat pile-up, in which its reference PDF is built from data themselves (see Sec. 4.2.1). The results of the fit performed with synthetic pile-up are shown in Tab. 4.2; they are consistent with those obtained with convolution method, and pp-ν rate is still significantly high.

- **Identification of a critical Phase-III-TMP subperiod.** In order to understand the origin of the problem, I have tried to perform the fit on three independent sub-periods of Phase-III: Phase-III-A (2016 July - 2017 July), Phase-III-B (2017 July - 2018 July) and Phase-III-C (2018 July - 2019 June). The results are shown in Fig. 4.10 for the most important

Analytical fit results [cpd/100t], $85 \leq N_p^{\mathrm{dt1-n}} \leq 900$ **range pp chain configuration**	
pp-ν	174.2 ± 12.7
^{7}Be-ν	47.4 ± 1.1
CNO-ν	4.92 [fixed]
pep-ν	3.18 ± 0.33
^{8}B-ν	0.46 [fixed]
^{14}C	39.95 ± 0.21 [constrained]
^{210}Bi	9.0 ± 1.3
^{11}C	27.2 ± 0.3
^{85}Kr	10.2 ± 1.5
^{210}Po	45.82 ± 0.53
^{214}Bi (ext.bkg.)	2.80 ± 0.40
^{40}K (ext. bkg.)	0.15 [fixed]
^{208}Tl (ext. bkg.)	4.51 ± 0.21
χ^2/NDF	1.21

Table 4.1: Analytical fit in pp configuration for Phase-III-TMP dataset: results. The analytical multivariate fit is performed in the $85 \leq N_{\mathrm{p}}^{\mathrm{dt1-n}} \leq 900$ reconstructed energy range ($0.19\,\mathrm{MeV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-). The pp chain configuration is adopted, i.e. the neutrinos rates from pp chain are left free to vary in the fit (with the exception of ^{8}B-ν, which is fixed to HER Phase-II analysis result), while CNO-ν rate is fixed according to the HZ-SSM predictions $R^{\mathrm{HZ-SSM}}(\mathrm{CNO\text{-}}\nu) = 4.92 \pm 0.78\,\mathrm{cpd/100t}$. The ^{14}C background is constrained by a gaussian pull term $40 \pm 1\,\mathrm{Bq}$. The results of the fit for neutrino and background rates (in cpd/100t units, except for ^{14}C in Bq/100t), and the χ^2/NDF, are reported.

species: pp-ν, ^{7}Be-ν, ^{85}Kr, ^{210}Bi . The main conclusion of this study is that the anomaly is present only in the last year (Phase-III-C) where $R(\mathrm{pp\text{-}}\nu) = 223 \pm 36\,\mathrm{cpd/100t}$, while the other two years the situation seems to be under control. Note also that in Phase-III-C the high pp-ν rate seems to be correlated with anomalous ^{7}Be-ν and ^{85}Kr results.

More detailed studies show that the problematic data seems to be confined in time to the second semester of 2018. More work is needed to pinpoint the cause of the problem. Meanwhile, I will perform the fit for the CNO-ν extraction only in the higher $140 \leq N_p^{\mathrm{dt1-n}} \leq 900$ renergy range ($330\,\mathrm{keV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-), which seems unaffected by this issues (see Chapter 7). We will show that Borexino is anyway sensitive to a CNO-ν signal, even in this fit *restricted* energy range.

4.3.2 Monte Carlo based fit in Phase-III-TMP for the pp chain analysis

As discussed in Sec. 4.2.2, the Monte Carlo fit is an alternative way to perform the multivariate analysis to extract the solar neutrino rates. As will be discussed in details in Chapter 5, the Monte Carlo fit was showing some problems in fitting data after 2016, making it impossible to analyze the Phase-III-TMP dataset with this method. After a long work of debugging, upgrade and re-tuning in which I have been heavily involved, the problem was fixed and the Monte Carlo can now fully

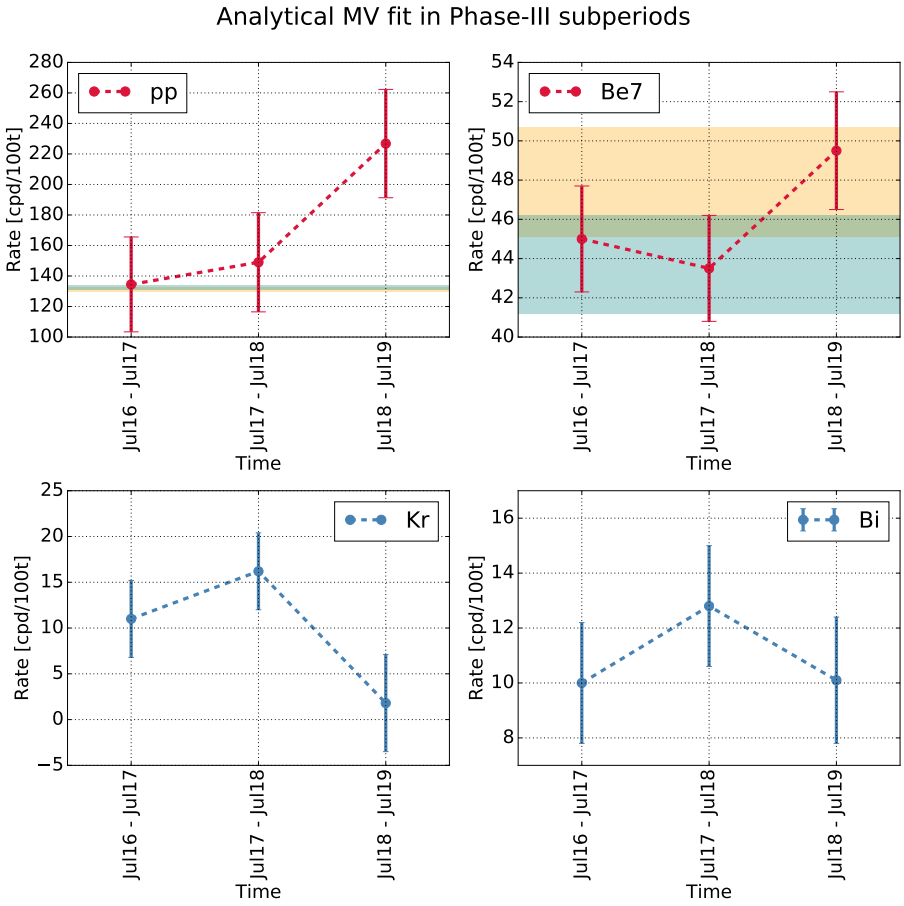


Figure 4.10: Analytical multivariate fit results for Phase-III-TMP yearly sub-periods with CNO-ν rate fixed.. Yearly profiles for pp-ν, ^{7}Be-ν, ^{85}Kr and ^{210}Bi best estimates. CNO-ν and pep-ν rates are fixed to HZ-SSM predictions. The blue and yellow horizontal bands correspond respectively to the 1σ LZ-SSM and HZ-SSM predictions for the pp-ν and ^{7}Be-ν interaction rates.

reproduce data after 2016, as will be discussed thoroughly in Chapter 5.

Here I show the first results of the Monte Carlo based fit in the pp chain configuration for Phase-III-TMP dataset. The fit is performed over the $140 \leq N_h^{\mathrm{n}} \leq 950$ reconstructed energy range, that corresponds to $330\,\mathrm{keV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-. I had to leave out the lowest energy part of the spectrum, because the pile-up Monte Carlo reference PDFs, whose contribution is crucial for this energy range, are currently being simulated after the major Monte Carlo upgrade.

The results of the fit are shown in Fig. 4.11 and reported in Tab. 4.3. We can see that the $\chi^2/\mathrm{NDF} = 1.30$ and associated residuals are good, and that the neutrino and background rates in output from the fit are consistent with the previously published results of Phase-II [39].

Analytical fit results [cpd/100t], $85 \leq N_p^{\mathrm{dt1-n}} \leq 900$ range pp chain configuration, synthetic pile-up	
pp-ν	166.5 ± 20.0
^{7}Be-ν	47.4 ± 1.1
CNO-ν	4.92 [fixed]
pep-ν	3.19 ± 0.34
^{8}B-ν	0.46 [fixed]
^{14}C	39.94 ± 0.22 [constrained]
^{210}Bi	9.0 ± 1.3
^{11}C	27.3 ± 0.3
^{85}Kr	10.3 ± 1.5
^{210}Po	45.79 ± 0.54
^{214}Bi (ext.bkg.)	2.80 ± 0.40
^{40}K (ext. bkg.)	0.15 [fixed]
^{208}Tl (ext. bkg.)	4.51 ± 0.21
χ^2/NDF	1.21

Table 4.2: Analytical fit for Phase-III-TMP dataset with synthethic-pile-up and with CNO-ν rate fixed: results. The analytical multivariate fit is performed in the $85 \leq N_{\mathrm{p}}^{\mathrm{dt1-n}} \leq 900$ reconstructed energy range ($0.19\,\mathrm{MeV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-), in the pp chain configuration and employing the synthethic pile-up method. The results of the fit for neutrino and background rates (in cpd/100t units, except for ^{14}C in Bq/100t), and χ^2/NDF, are reported.

4.4 Conclusions

The strategy for the background reduction in the solar neutrino analysis in Borexino has been described, along with the multivariate fit procedure according to the two complementary approaches to obtain the reference shapes (analytical and Monte Carlo approaches). These techniques have been designed by the collaboration for the determination of the pp-chain neutrino fluxes (Sec. 2.5.1), and will be extensively used in the next chapters for the determination of CNO neutrinos on the Phase-III-TMP dataset (Chapter 7).

As a preliminary crosscheck, both the analytical and Monte Carlo based multivariate fits have been performed on the Phase-III-TMP in the pp chain configuration (i.e. pp chain neutrinos rates are left free to vary, while CNO-ν is constrained according to HZ-SSM predictions). In both cases, results are compatible with respect to the Phase-II analysis ones.

The fit results in pp chain configuration shown in this Chapter can be considered as preliminary (and successful) crosschecks, confirming the robustness of the two fit approaches on the Phase-III-TMP dataset. Therefore, given the upgrade of Monte Carlo simulations and the ^{210}Bi rate independent determination, that will be described in the following Chapter 5 and Chapter 6, it will be possible to proceed with the CNO-ν determination analysis in Chapter 7.

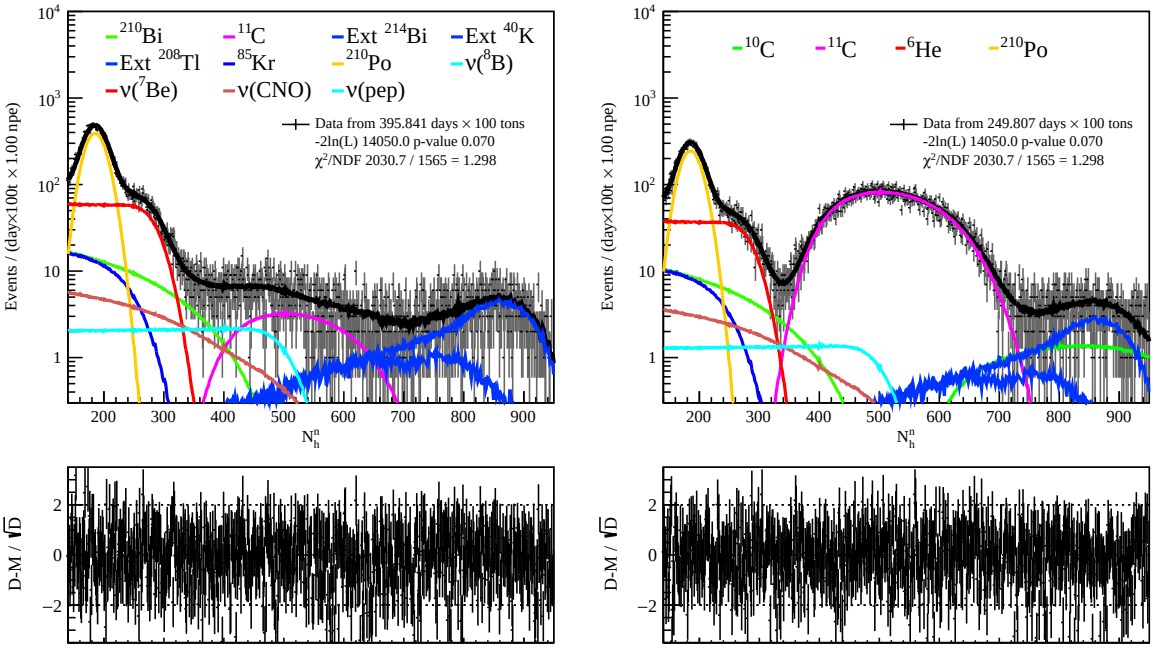


Figure 4.11: Monte Carlo fit for Phase-III-TMP dataset with CNO-ν rate fixed: TFC-subtracted (left) and the TFC-tagged (right) energy spectra (top panels), with residuals (bottom panels). The sum of the individual components from the fit (black lines) is superimposed on the data (grey points). The analysis has been performed within the $140 \leq N_h^n \leq 950$ energy window ($330\,\mathrm{keV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$ of the scattered e^-). The residuals are calculated in every bin as the difference between the data counts and the fit result, divided by the square root of the data counts. The interaction rates are reported in Tab. 4.3.

MC fit results [cpd/100t], $140 \leq N_h^n \leq 950$ **range pp chain configuration**	
^{7}Be-ν	46.8 ± 1.4
CNO-ν	4.92 [fixed]
pep-ν	2.72 ± 0.29
^{8}B-ν	0.46 [fixed]
^{210}Bi	13.2 ± 2.3
^{11}C	27.0 ± 0.3
^{85}Kr	10.0 ± 1.8
^{210}Po	51.44 ± 0.5
^{214}Bi (ext.bkg.)	3.62 ± 0.45
^{40}K (ext. bkg.)	0.15 [fixed]
^{208}Tl (ext. bkg.)	5.08 ± 0.22
χ^2/NDF	1.30

Table 4.3: Monte Carlo fit for Phase-III-TMP dataset with CNO-ν rate fixed: results. The Monte Carlo multivariate fit is performed in the $140 \leq N_h^{\rm n} \leq 950$ reconstructed energy range ($330\,\text{keV} \lesssim E \lesssim 2.9\,\text{MeV}$ of the scattered e^-). The pp chain configuration is adopted: pep-ν neutrinos rates from pp chain are left free to vary in the fit, ^{8}B-ν is fixed to HER Phase-II analysis result, while CNO-ν rate is fixed according to the HZ-SSM predictions $R^{\rm HZ-SSM}(\text{CNO - }\nu) = 4.92 \pm 0.78\,\text{cpd/100t}$; pp-$\nu$ is not included due to the restricted energy range. The best estimates for neutrino and background rates (in cpd/100t units), and the χ^2/NDF, are reported.

Chapter 5

The Monte Carlo simulation of the Borexino detector for the CNO neutrinos analysis

An accurate and validated simulation is essential for any particle experiment and Borexino makes no exception. As discussed in the previous Chapter, the multivariate fit to extract the CNO neutrinos signal is crucially based on the comprehension of the Borexino response function, which can be obtained through a detailed Monte Carlo simulation [39, 136, 138, 162].

The Borexino Monte Carlo code [162] is a simulation package, based on Geant4, which includes all the processes occurring in the Borexino detector, from the energy deposition, to the light emission, propagation and detection, up to the signal processing. The simulation needs in input a large set of parameters, mainly related to the different material properties. Many of these parameters, like the scintillator features, have been measured before the Borexino construction in small scale laboratory tests. However, even those parameters needed to be adjusted and fine-tuned in the realistic experimental conditions. This was done thanks to an extensive campaign of calibrations with radioactive sources, which was carried out in 2009 and allowed to define the values of all input parameters [65, 164]. After this tuning, the Monte Carlo simulation was able to reproduce the relevant physical quantities characterizing the event, such as the reconstructed position, the energy scale, and in general, the detector response as a function of position at the $\approx 1\%$ level over a $r < 3\,\mathrm{m}$ Fiducial Volume and in a wide $0.1\,\mathrm{MeV} \lesssim E \lesssim 3\,\mathrm{MeV}$ energy rande.

The excellent accuracy reached by the Monte Carlo simulations was one of the key elements which led to the numerous Borexino publications on the so-called Phase-II dataset (Dec 2011 - May 2016) including the comprehensive measurement of pp-chain solar neutrino fluxes, discussed in Sec. 2.5.1 and Sec. 2.5.2 [39]. However, after 2016 the Monte Carlo simulations failed to reproduce data with the required precision. This showed up as inconsistencies in the energy scale when trying to fit data using the signal and background reference shapes obtained from simulations. Together with the Borexino analysis team I devoted a significant effort to solve this issue, which would have prevented us to perform the Monte Carlo fit on the Phase-III dataset, dedicated to CNO-ν analysis which started in July 2016. The problem was found to be related to the photomultiplier quantum efficiencies (QE) which are given as input to the simulations and are calculated from data on a weekly basis. After reviewing and improving the code which calculates the QE, the Monte Carlo input parameters have been re-tuned and the reference energy shapes for signal and backgrounds have been produced again. This has solved the problem and now it is possible to perform the Monte

Carlo fit also on Phase-III data.

This Chapter is organized as follows. Sec. 5.1 reviews the Borexino simulation chain to generate a physical event: the particle interaction and energy loss in the scintillator (Sec. 5.1.1), the photon emission and propagation (Sec. 5.1.2), the electronic readout and trigger condition (Sec. 5.1.3). Sec 5.2 illustrates the Borexino calibrations, used to tune the Monte Carlo parameters. The problem of Monte Carlo simulations after 2016 is addressed in Sec. 5.3. The quantum efficiency topic, along with their role in explaining the previously mentioned puzzle, is illustrated in Sec. 5.4. My work on the quantum efficiency validation and on the Monte Carlo tuning is extensively described in Sec. 5.5 and Sec. 5.6 respectively. The simulations of the Phase-III-TMP reference PDFs for the Monte Carlo multivariate fit are outlined in Sec. 5.7; the fit results are shortly described in Sec. 5.8. Conclusions will be outlined in Sec. 5.9.

5.1 Simulation chain basic steps

The Monte Carlo simulation for Borexino events consists in a chain of numerical codes performing the following steps [162]

1. **Event generation:** classes for the generation of solar neutrinos, calibration sources, radioactive decays, anti-neutrinos.
2. **Energy loss:** for each particle, the energy loss in the materials is simulated.
3. **Scintillation photon production and tracking:** the number and the wavelength of scintillation or Čerenkov photons are simulated according to the particle type, the energy and the material in which the energy is deposited. Then, they are individually tracked during their path, taking into account the possible interactions with the detector materials. The tracking is completed when photons reach a PMT or are absorbed.
4. **PMT signal generation and processing:** if a PMT is fired, the code simulates the photoelectron generation, according to the PMT quantum efficiency, and the processing of the pulse signal following the electronic chain. The trigger condition is also checked.

The first three steps are handled by *g4bx2*, a code for simulation and tracking mainly based on the Geant4 platform, optimized according to the Borexino geometry and involved processes; it includes classes for handling geometry, light tracking, detector response, up to the photon absorption or interaction with a photocatode. The simulation of the electronic chain, from the photoelectron generation up to the trigger condition, is performed by *bxelec*, a sofware developed by the Borexino collaboration. The result of the whole simulation chain is a raw data file with same structure as the one produced by the Borexino data acquisition system; at last, it is analyzed by Echidna, the Borexino event reconstruction code which also is used for the reconstruction of real data. The Borexino Monte Carlo simulation steps are now reviewed with more detail.

5.1.1 Event generation

The physics of particle interactions and propagation inside the detector is managed by the Geant4 platform, both for primary and secondary particles ($e^{\pm}$, γ, α, ν_e, $\overline{\nu}_e$, μ, n, ions, and so on) [162].

Radioactive decays are handled by the Geant4 code which, for most of the radionuclides, provides the isotope daughter spectra and the decay branching ratios. The only exceptions are given by ^{14}C and ^{210}Bi backgrounds: dedicated generators have been developed by the Borexino collaboration, allowing to vary the β spectrum shape parameters if needed. In these two cases, in fact, the shape factors are known with large errors and we needed to check how this uncertainty propagates to the multivariate fit results. Solar neutrinos are simulated via a dedicated generator, which takes into account their energy and the related flavor oscillation probability, as well as the elastic $\nu_e - e^-$ cross sections, including radiative corrections [162].

5.1.2 Photons generation and light tracking

The Borexino scintillator is made of pseudocumene (PC) mixed to 1.5 g/l of PPO as solute; the energy deposited after an interaction is firstly transferred to PC and then to PPO mainly through non-radiative processes. The de-excitation of π electrons is responsible for the emission of scintillation photons. The scintillation light is simulated following the PC+PPO emission spectrum shown in Fig. 5.2 by the black points, which peaks at ≈ 380 nm wavelength.

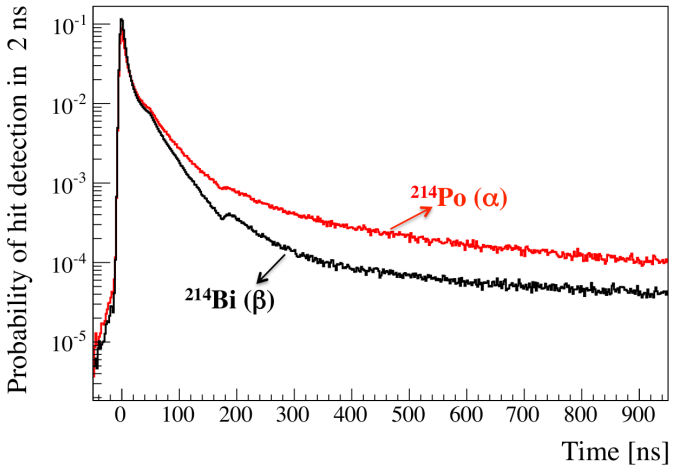


Figure 5.1: Examples of α and β events **scintillation time profiles**: ^{214}Po for α events, in red line, and ^{214}Bi for β events, in black line.

The scintillation light yield, which represents the number of emitted photons per MeV of deposited energy, is one of the most important input parameters of the Monte Carlo simulations, and has been fine-tuned on calibration data resulting in $Y_0 = 1.36 \cdot 10^4$ photons/MeV. The time distribution of the emitted photons is described as the sum of four exponential profiles with different time constant τ_i; examples of the full scintillation time profiles are shown in Fig. 5.1.

The relative weights of each exponential curve, w_i, along with the corresponding τ_i, have been measured for pseudocumene in small scale laboratory tests, but have been fine-tuned analyzing the calibration data. The w_i and τ_i values after tuning are reported in Tab. 5.1, along with the other main quantities given as input to the Monte Carlo simulations [162, 163]; note that these quantities are different for α and β events. This difference is at the basis of the pulse-shape discrimination technique. The simulation takes also into account the Čerenkov light: the Čerenkov spectrum extends up to the far ultraviolet region, which in principle is not detectable by the PMTs, but

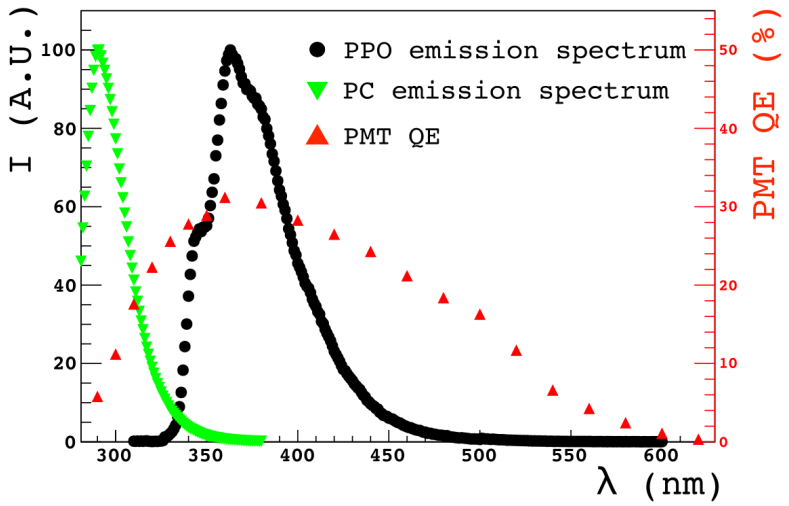


Figure 5.2: PC and PC+PPO emission wavelength spectra, respectively in green and black markers. The measured PMT quantum efficiency as a function of wavelength is shown by the red triangles. The PPO solute acts as a wavelength shifter, matching the final emission spectrum towards the QE peak region and therefore optimizing the light collection.

Čerenkov photons can be absorbed by the scintillator and re-emitted at lower wavelengths. However their contribution is small in Borexino, amounting to $\approx 1\%$ of the scintillator photons emitted at 1 MeV.

The light tracking is performed on a photon by photon basis. All the possible interactions with the single components of the detector are taken into account, including the Rayleigh scattering, the absorption and the re-emissions due to PPO and DMP molecules. The optical interactions with the detector materials and material interfaces (nylon vessel) and surfaces (reflections on light concentrators, SSS and PMT photocathodes) are effectively accounted for through a set of parameters for each component of interest [162].

5.1.3 Photoelectron signal and readout electronic

The electronic simulation code reproduces the electronic chain and the trigger system response, based on the information of the PMT pulse times. Each one of the 2212 Borexino PMTs has slightly different light detection efficiency, due to intrinsic differences of photocathodes. To guarantee an accurate reproduction of the reconstructed energy, the PMT photocathode parameters cannot be considered equal a priori. The PMTs quantum efficiencies (QE) need to be given as input to the simulation chain as a function of time, to account for possible variation of the PMTs performance, and need to be calculated by data on periodic basis. This is a crucial issue which has proven to be the cause of the problem with Monte Carlo after 2016, as will be explained in Sec. 5.4.

Parameter	β events		α events	
	Meas.	Tuned	Meas.	Tuned
τ_1 [ns]	3.95	3.7	4.15	3.9
τ_2 [ns]	23.56	24	19.90	26
τ_3 [ns]	78.86	60	99.91	110
τ_4 [ns]	546.4	600	617.2	630
w_1	0.933	0.889	0.679	0.674
w_2	0.024	0.055	0.144	0.146
w_3	0.022	0.027	0.102	0.103
w_4	0.021	0.029	0.075	0.077
Y_0 [photons/MeV]	$1.10 \cdot 10^4$		$1.36 \cdot 10^4$	
$k\,[\mathrm{cm} \cdot \mathrm{MeV}^{-1}]$	0.0109	0.01098	0.012	0.01055

Table 5.1: Monte Carlo simulations: input parameters for the reproduction of the scintillator response [162, 163]. The *measured* columns refer to the values determined by small scale tests in laboratory, while the *tuned* columns indicate the values after the Monte Carlo tuning. The four τ_i represent the time constant for each of the four exponential time distributions of the scintillator emitted photons, with relative weight w_i. Y_0 is the scintillator light yield, expressed in scintillation photons for 1 MeV of deposited energy. k_α and k_β are the quenching Birks parameters, describing the non linearity in energy deposition.

5.2 Detector calibrations and Monte Carlo tuning

Calibrations have been exploited to achieve a better comprehension of the detector response, namely to determine the energy scale, to describe non-linearity effects, and to study the detector response dependence on position, energy, trigger efficiency [164]. On the Monte Carlo side, they are a fundamental pawn to tune the simulations input parameters, to study extensively the agreement with data in order to validate the simulation chain.

5.2.1 Borexino calibration campaigns

Four Borexino calibration campaigns have been carried out between October 2008 and July 2009[1]: α, β, γ and neutron sources were deployed in about 300 positions inside the detector active region. The largest part of the calibration points is located in positions with radial distance from the detector center r lower than 3 m, covering the FVs of interest for the solar neutrino analyses. Overall, the four internal calibration campaigns lasted for 54 days with a duty cycle of 65%. The sources used during the calibration campaigns are reported in Table 5.2.

Even though the Borexino internal calibration system was designed and built following very strict requirements on radiopurity, the calibration campaigns always carry a non-zero risk of contamination. For this reason, internal calibrations have been performed more than one year after the beginning of Borexino data acquisition, when a significant sample of good data had been collected. Comparing data acquired before and after each calibration campaign, it has been possible to search for any

[1] October 2008 (I), January 2009 (II), June 2009 (III), and July 2009 (IV).

Source	Type	E [MeV]	Position	Motivation
^{57}Co	γ	0.122	IV volume	Energy scale
^{139}Ce	γ	0.165	IV volume	Energy scale
^{203}Hg	γ	0.279	IV volume	Energy scale
^{85}Sr	γ	0.514	z axis + $R < 3$ m sphere	Energy scale + FV
^{54}Mn	γ	0.834	along z axis	Energy scale
^{65}Zn	γ	1.115	along z axis	Energy scale
^{60}Co	γ	1.173, 1.332	along z axis	Energy scale
^{40}K	γ	1.460	along z axis	Energy scale
^{222}Rn – ^{14}C	γ, β	0 – 3.20	IV volume	FV + uniformity
	α	5.5, 6.0, 7.4	IV volume	FV + uniformity
^{241}Am – ^{9}Be	n	0 – 9	sphere $R < 4$ m	Energy scale + FV
394 nm laser	photons	-	center	PMT equalization

Table 5.2: Radioactive sources used during the Borexino internal calibration campaigns [164]. The radionuclides, the emitted particle types and the decay energies are shown in the first three columns. The fourth column indicates the positions where the sources have been deployed within the scintillator. The main purposes for the individual source measurements are summarized in the fifth column; the "FV" notation indicates that sources have been used to study the systematics on the position reconstruction, in particular within the $r < 3$ m region.

hints of any calibration-induced contamination. This analysis showed that the global internal calibration program left no detectable evidence of impurities [164].

In the following, the radioactive sources used for calibrating the detector are discussed in more details:

- **^{222}Rn–^{14}C source.** 182 calibration points have been scanned with a ^{222}Rn – ^{14}C compound source which provides α, β , and γ events across a large energy window (0 MeV $< E <$ 3.2 MeV). The ^{222}Rn chain consists of three α emitters: ^{222}Rn, ^{218}Po and ^{214}Po, with energies of 5.5 MeV, 6.0 MeV and 7.4 MeV, respectively. Due to the high quenching factor (≈ 12), α particles produce an equivalent energy around 400 keV (see Sec. 2.4.1). Moreover, the ^{222}Rn – ^{14}C source provides two β/γ emitters ^{214}Pb and ^{214}Bi with endpoints of 1.0 MeV and 3.2 MeV, respectively. The ^{222}Rn – ^{14}C sources used in the calibrations were prepared with an initial activity of $\sim$ 100 Bq.

- **γ sources**. Mono-energetic γ sources contributed crucially to the knowledge of the detector energy response, determining the absolute energy scale and the non-linearities. Several sources have been employed, as reported in Tab. 5.2, spanning a wide energy range which ranges from a 0.122 MeV minimum (^{57}Co) to a 2.5 MeV maximum (^{60}Co). The activity of each of the γ sources employed was approximately 2 Bq [164].

- **^{241}Am – ^{9}Be neutron source.** This is a fundamental source for what concerns the anti-neutrinos analysis, since it allows to have available energy calibration points up to 10 MeV and to scan the neutron capture efficiency in the detector volume. In particular, this set of simulations would have been extremely useful for the SOX project (see Chapter 8), based

on $\overline{\nu}_e$ detection. A ^{241}Am $-$ ^{9}Be neutron source, with a 10 Bq activity, was deployed in thirty positions. Neutrons are produced via two main reactions, $^{9}\mathrm{Be} + \alpha \rightarrow n + {}^{12}\mathrm{C}$ and $^{9}\mathrm{Be} + \alpha \rightarrow n + {}^{12}\mathrm{C}^{*}$, in which neutron energies reach 11 MeV and 6.5 MeV, respectively. The second reaction also produces one or two γ rays with a total energy of 4.4 MeV from the $^{12}\mathrm{C}^{*}$ de-excitation. These photons, together with the recoil protons from neutron scattering in the medium, are responsible for a prompt scintillator signal. Afterwards, neutrons thermalize in the scintillator in a $\tau \approx 256\,\mu$m mean time and are captured either on protons or ^{12}C nuclei, emitting characteristic capture rays [2].

- **Laser source.** During the third internal off z-axis calibration campaign, a dedicated 394 nm pulsed laser system has been used to provide light to a diffuser ball located at the center of the detector. The primary goal of the laser source calibration was to check the PMT time synchronization independently from the main equalization system (see Sec. 2.4.1). The standard PMT equalization is performed routinely by means of a multiplexed optical fiber system, which delivers the laser light directly and simultaneously on each photocathode [165]. During the laser source calibration runs, the laser and the detector were synchronously triggered at 50 Hz at several different laser intensities.

5.2.2 Monte Carlo tuning strategy

The Monte Carlo tuning for Borexino is an intricated and delicate procedure, whose primary goal is the optimization of the simulation code accuracy to reproduce the collected data. The strategy adopted in the tuning is sketched in Fig. 5.3 and is briefly summarized hereafter, while the previously shown Table 5.1 lists several of the tuned parameters for the detector response reproduction.

The tuning procedure is based on an iterative comparison of Monte Carlo simulations with 2008-2009 calibration data [162].

The first step is the optimization of the agreement with the measured uniformity of the detector energy response, which is first of all driven by photon attenuation lengths and material reflectivities; a dedicated tuning was performed thanks to the $^{14}\mathrm{C} - {}^{222}$Rn calibration runs, analyzing the ^{214}Po α energy peak scale. The attenuation lengths for PC, PPO, DMP and nylon have been tuned through multiplicative factors which scale the values previously measured in lab on small scale samples. The light collection uniformity is also dependent on the concentrator reflectivities, which are tuned through the optimization of ratio of the fired PMTs with and without light concentrators ("cone-no cone ratio") and of the fired PMTs far (> 4 m) and near (< 4 m) from the source ("far-near ratio"). These two variables are not dependent on the energy scale and linearity.

The second step consists in the study of the time distribution of the collected light. The time constants, the absorption-reemission delays, and the electronic response need to be tuned. α and β low energy source data (mainly the α-decaying ^{210}Po and the β-decaying ^{85}Sr isotope) have been analyzed, selecting runs with the sources deployed at the detector center: in this way, geometrical effects due to anisotropy of photon propagation, and therefore of light collection, are removed. The tuned value of attenuation lengths and reflectivities in the previous step are given as inputs.

The final and most important step is the tuning of the energy scale, analyzing the γ calibration runs: this depends on the absolute scintillation light yield Y_0, which drives the energy scale, on

[2] In addition, neutrons can be captured on iron, nickel and chromium nuclei of the stainless steel source insertion arm, resulting in the emission of de-excitation γs with energies up to 9.3 MeV.

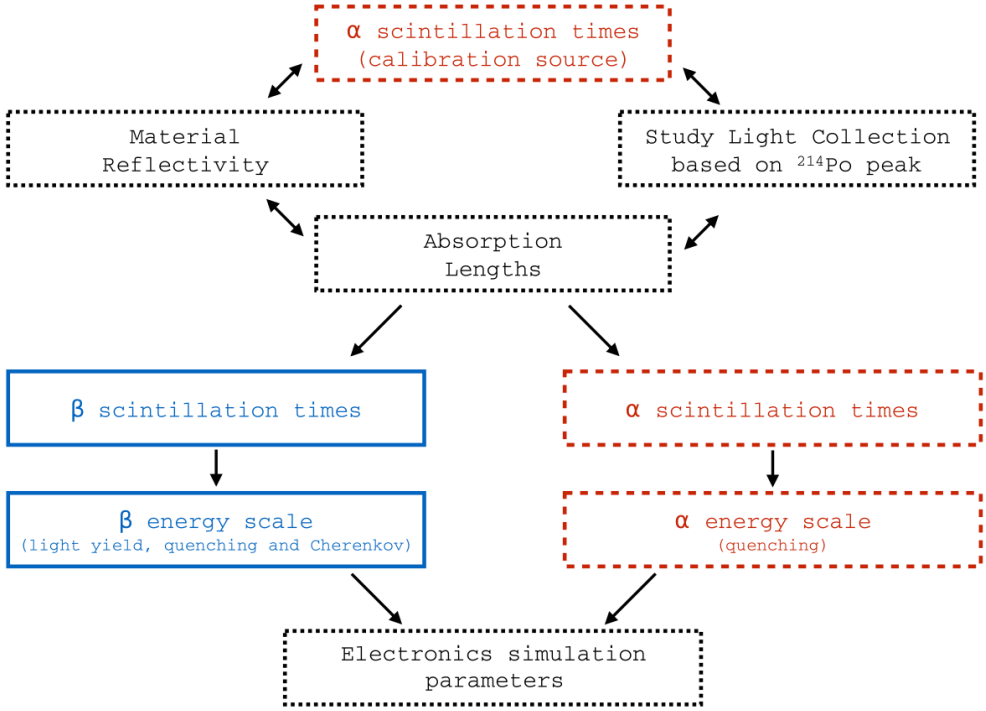


Figure 5.3: Scheme of the Monte Carlo tuning strategy. Continuous boxes (blue) are relative to β particles, while dashed ones (red) refer to α particles. Finer dashed boxes (black) indicate sets of response parameters not depending on the particle type.

the Birks quenching parameter k_β, which describes energy non linearities, and marginally to the yield of the Čerenkov light. β and α events energy scales are firstly commonly tuned, since these two events categories share most of these parameters. Once these values are fixed by the previous iteration, the α energy scale is separately refined through the Birks quenching parameter k_α.

As hinted in the previous lines, the correlations between the large amount of parameters are one of the main issues to address during the tuning. Once the tuning has been completed according to the procedure described above, the simulation code is validated furtherly and tested on independent datasets, as the ^{241}Am-^{9}Be calibration source data, the external background calibration data, and the ^{210}Po events selected over the 2012-2013 years.

5.2.3 Monte Carlo performances

After the tuning described above, Monte Carlo simulations match data very well: some exemplary plots will be shown throughout this paragraph. The hit time distributions agreement is a key step, since they critically affect all the time-related quantities, including the reconstructed position and the pulse-shape discriminator variables. An example of data-Monte Carlo comparison and agreement for reconstructed hit times, for a γ calibration example run, can be seen in Fig. 5.4.

A precise energy scale matching is as well crucial for a properly working simulation. The reconstructed energy spectra of γ de-excitations from eight different calibration sources are reported in Fig. 5.5, in terms of the N_h^{n} energy variable. This energy estimator counts the number of fired PMTs

within the data acquisition time window, normalized to a fixed number of 2000 working PMTs (as described in Sec. 2.4.1). The energy peaks, from left to right, belong to the ^{57}Co, ^{139}Ce, ^{203}Hg, ^{85}Sr, ^{54}Mn, ^{65}Zn, ^{40}K and ^{60}Co sources (see Tab. 5.2). These provide eight energy calibration points covering most of the energy region of interest for Borexino, i.e., from as low as 122 keV (^{57}Co source) up to 2.5 MeV (sum of the two γ rays from ^{60}Co source). Extracting the energy peak value, it is possible to obtain the energy scale conversion. This was essential to fine-tune the Monte Carlo code, which should reproduce the scintillator non-linear response over this relatively large range of energies. The energy position and sigma of each γ source gaussian-like peak is reproduced by Monte Carlo at the 0.8% level.

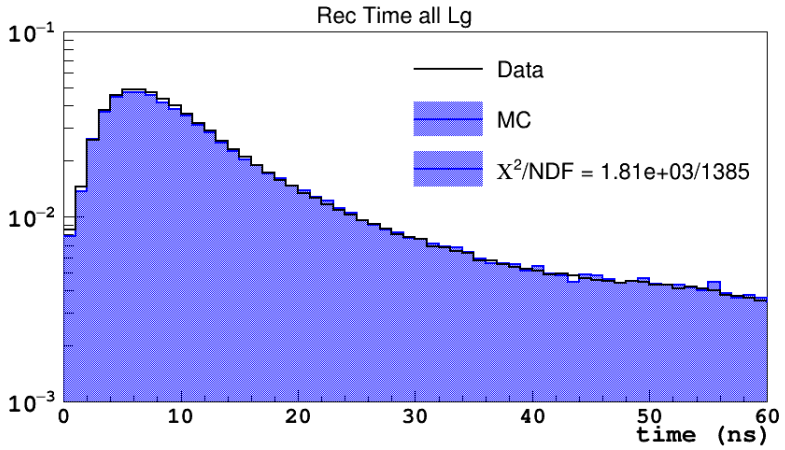


Figure 5.4: Data-Monte Carlo comparison and agreement for **reconstructed hit times**, i.e. the hit time distribution corrected for the time of flight in ns, for a γ calibration example run.

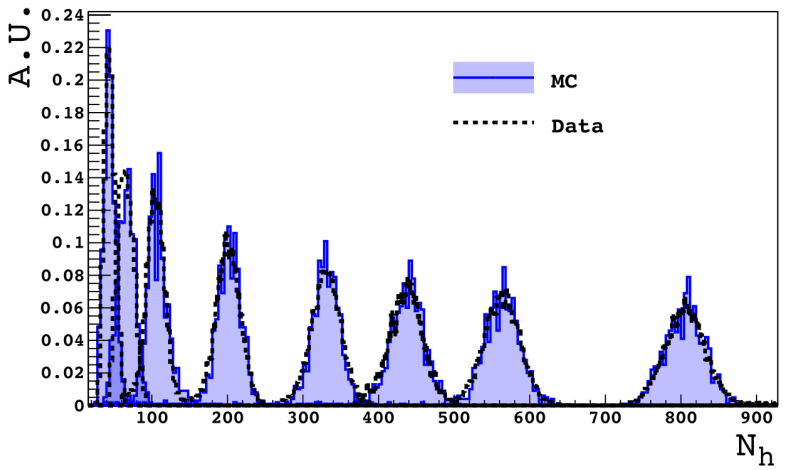


Figure 5.5: Energy spectra of γ isotopes from eight different calibration sources expressed in terms of the N_h^n energy variable, along with the Monte Carlo predicted spectrum. The peaks from left to right belong to the ^{57}Co, ^{139}Ce, ^{203}Hg, ^{85}Sr, ^{54}Mn, ^{65}Zn, ^{40}K and ^{60}Co sources.

For what concerns the energy response as a function of position, we know that many factors can contribute to a non uniform response: the light absorption, which depends on the travelled

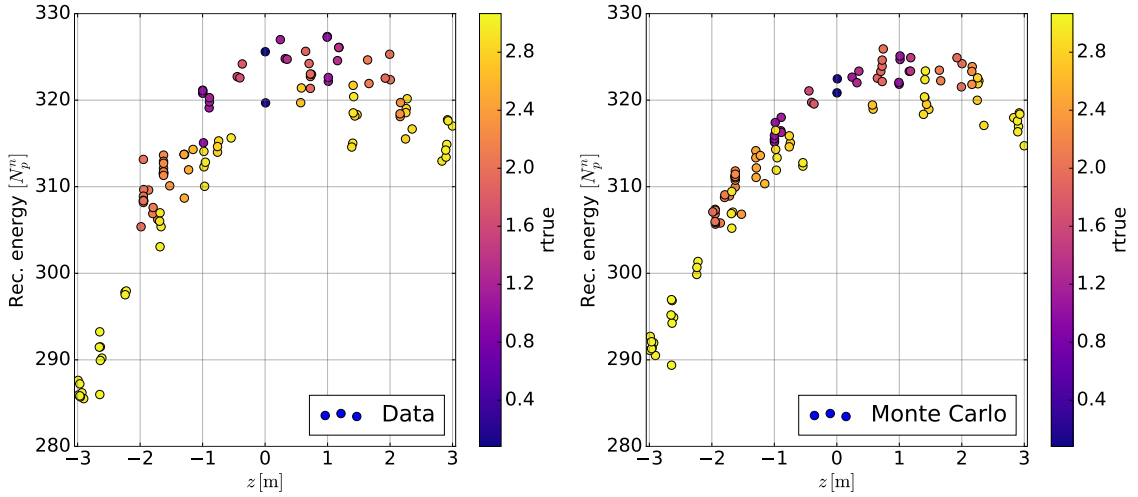


Figure 5.6: Energy response uniformity in space: data (left) and Monte Carlo simulations (right). The reconstructed energies in N_p^n of ^{214}Po α peak vs the z-coordinate, exctracted from the ^{14}C-^{222}Rn calibration sources, are shown. The points at fixed z value correspond to different x and y coordinates; the color scale indicates the reconstructed radius in meters.

path, the different response of individual electronics channels, the non-uniform distribution of non-working electronic chains. Fig. 5.6 shows the z-coordinate dependency of the reconstructed energy, expressed in the normalized NPMTs estimator N_p^{n}. Left and right panels display respectively the real $^{14}\mathrm{C} - {}^{222}\mathrm{Rn}$ source calibration data and the related Monte Carlo simulations. Each point on the plot is associated to one of the calibration runs. The points at fixed z position reflect different x and y coordinates: the color scale is indeed related to the event reconstructed radius. The light collection is maximized at the center of the detector, while it decreases approaching to the poles. The North-South trend asymmetry is due to the higher concentration of broken PMTs in the bottom hemisphere with respect to the upper one, especially close to the south pole. The differences of reconstructed energy at $z = \pm 3\,\mathrm{m}$ with respect to the the center are

$$\frac{\Delta N_p^n(z = -3\,\mathrm{m})}{N_p^n(z = 0\,\mathrm{m})} \approx -12\% \qquad \frac{\Delta N_p^n(z = 3\,\mathrm{m})}{N_p^n(z = 0\,\mathrm{m})} \approx -3.5\% \tag{5.1}$$

This large non-uniformity is well reproduced by Monte Carlo simulations (right panel of Fig. 5.6); differences between data and Monte Carlo reconstructed energy are found to be within 1%.

In conclusion, the Monte Carlo tuned on calibrations is able of reproducing well the energy scale, its dependences on position and the photon time distributions which are crucial elements for the analysis.

5.3 Motivations for the re-tuning of Monte Carlo code for Phase-III data

As discussed in the previous paragraph, the Borexino Monte Carlo code tuned on 2009 calibration data has been able to reproduce the detector response at the 1% level in the FV. The detector response has proven to be stable for years at the % level, despite the natural degradation and the consequent death of hundreds of PMTs. From 2016 on, the Monte Carlo energy scale fails to match the data one; in particular, Monte Carlo runs simulated in the 2016 detector conditions show that the ^{210}Po energy peak is found $\approx 1.5\%$ higher than data. In addition, the ^{7}Be-ν Compton edge is clearly shifted to inconsistently larger energies.

The origin of this disagreement could be in principle due to a physical effect, e.g. a natural worsening of the detector light yield in time. Indeed, the stability of the scintillator light yield cannot be assumed a priori: many scintillator-based detectors are reported to suffer from non negligible degradations (of the order of some percent) in few years from the beginning of data taking. Since the calibration campaigns date back to 2009, an independent check of the current absolute energy scale is not possible.

^{210}Po can be employed as a standard candle to monitor independently the relative light yield time evolution on data, thanks to its high rate and to the fact that its energy distribution is gaussian, whose position is easily identifiable. The ratio of ^{210}Po energy peaks for different years is proportional to the light yield variations. Fig. 5.7 shows the relative variations of light yield parameter Y_0 with respect to its 2012 value Y_0^{2012}, both for ^{210}Po data (green line) and Monte Carlo simulations (red line). The beginning of Phase-III-TMP dataset is marked by the orange dashed line. This study shows that the energy scale calculated on data, and therefore the light yield, is stable or slightly decreasing in time, by less than 1% over the whole Phase-II and Phase-III-TMP periods; Monte Carlo simulations instead show an increase of light yield by almost 2%. This energy scale mismatch between data and Monte Carlo could be the source of the Monte Carlo fail after 2016.

As a further crosscheck, the light yield evolution in time for Monte Carlo datasets can be also evaluated monitoring the average reconstructed energy for ^{14}C data events, calculated over a 1 m or 2 m radius sphere in the geometrically normalized hits estimator [3] $50 \leq N_h^{\mathrm{ng}} \leq 100$ range ($0.10\,\mathrm{MeV} \lesssim E \lesssim 0.21\,\mathrm{MeV}$); this is shown in Fig. 5.8, respectively in blue or orange lines. This ratio is directly linked to the light yield variations in time: from 2012 on, the simulations show a smooth increase of $\approx +0.6\,N_h^{\mathrm{n}}$, equivalent to a $\approx +1\%$ shift for the energy scale. This plot confirms the results already shown for the ^{210}Po events (Fig. 5.7): the Monte Carlo artificially simulates an increase of light yield which is not present on data.

In the following, I will show that this anomalous light yield behaviour in time is related to the PMT quantum efficiency set (QE), given as input to the Monte Carlo simulations. This led to the review and the improvement of quantum efficiency computation algorithm, and to the consequent production of new Monte Carlo simulations without this issue.

[3] The geometrical normalization corrects the selected non-normalized energy estimator to 2000 reference working channels (as the standard normalization does) and, moreover, takes into account geometrical acceptance effects of the energy reconstruction, which are position dependent.

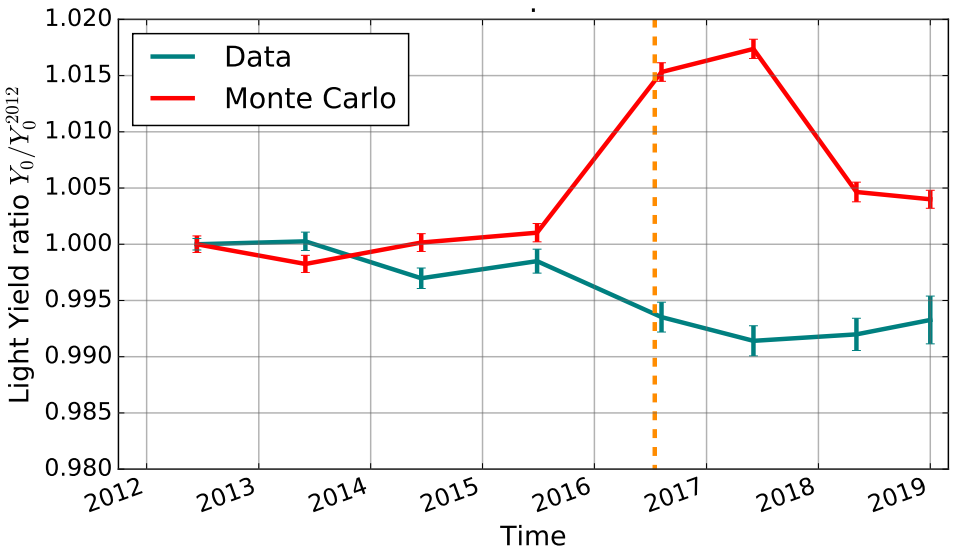


Figure 5.7: Evolution of Borexino energy scale in time, by analyzing ^{210}Po events: relative variations of light yield parameter Y_0 with respect to the 2012 value Y_0^{2012} both for data (green trend) and Monte Carlo (red trend). ^{210}Po events are selected in the innermost region of the scintillator. The beginning of Phase-III-TMP dataset is marked by the orange dashed line.

5.4 Effective quantum efficiency

The effective quantum efficiencies (QE) are a set of time dependent parameters (one parameter for each of the 2212 photomultiplier tubes) given as input to the Monte Carlo simulations. They take into account in an effective way the response of each PMT and the electronic channel coupled to it. In particular, they include the following effects:

- *Intrinsic quantum efficiency of the photocatode*, i.e., the probability that a photoelectron (p.e.) is emitted upon arrival of a photon. This depends on each photocatode details (imperfections, uneveness, and so on);
- *Electronic channel properties*, such as the PMT gain, threshold differences and the electronic noise.

Both the effects are not fixed in time a-priori and their variations may significantly impact on the energy response scale and homogeneity. Therefore, the QE set must be calculated on a periodic basis.

Historically, the calculation of QE makes use of ^{14}C events, because of their extremely high statistics ($R(^{14}\mathrm{C}) \approx 40\,\mathrm{Bq/100t}$), their uniformity in the scintillator volume, and their low energy (which guarantees the PMT to work in single-photo electron regime). On a weekly basis, ^{14}C events located inside a $r < 2\,\mathrm{m}$ spherical volume are selected, such that the solid angle between the events and all the PMTs can be assumed to be equal. The QE of each PMT is calculated as the total number of hits recorded by that PMT normalized to the total number of events and to the total number of hits.

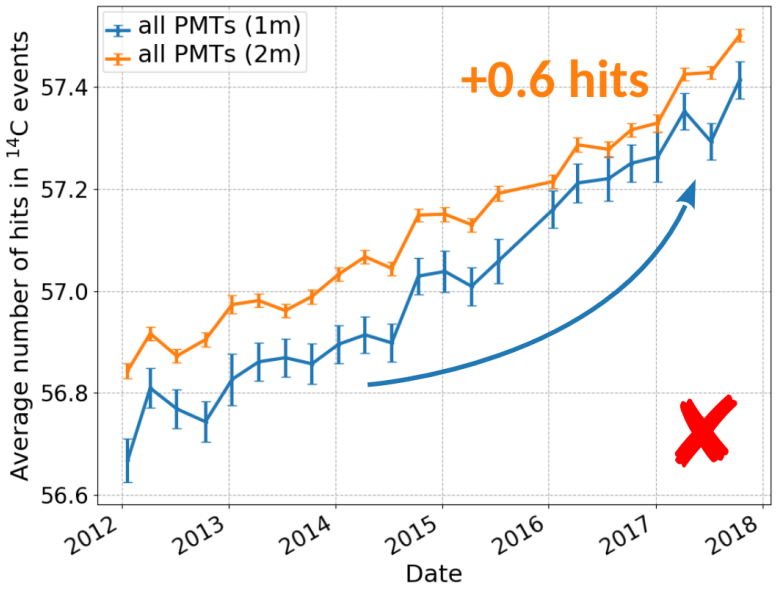


Figure 5.8: Monte Carlo simulations: reconstructed energy (N_h^{n}) for ^{14}C events simulated in different years, calculated over a 1 m or 2 m radius sphere (respectively in blue or orange lines). This plot confirms the results already shown for the ^{210}Po events (Fig. 5.7): the Monte Carlo artificially simulates an increase of light yield which is not present on data.

To be more precise, for a given sample of N_{ev} events included in a time window t of r runs, the QE(p) for the p-th PMT is computed comparing the hits firing the single PMT with respect to the overall hits:

$$N_h(p,r) = \sum_{\mathrm{ev}=1}^{N_{\mathrm{ev}}} N_h(p,\mathrm{ev}) \qquad N_{\mathrm{ev}}(p) = \sum_{r|p\,\mathrm{enabled}}^{r\in t} N_{\mathrm{ev}}(r) \tag{5.2}$$

where $N_h(p,\mathrm{ev})$ is the number of hits firing the p-th PMT for a given ev-th event, and $N_{\mathrm{ev}}(p)$ is the sum of all events among all runs $r \in t$ where the PMT was enabled. Note that 1828 PMTs over the total 2212 ones are equipped with *light concentrators* (also named *cones*), installed to improve their collection efficiency for scintillation photons coming from the center, resulting in an enhanced QE value with respect to the no-concentrator (*no cones*) PMTs.

To get a more accurate evaluation, the proper $N_h(p,\mathrm{ev})$ should be free from dark noise hits, which at this point are subtracted to N_h to get the *clean hits* N_h^{clean}

$$\mathrm{QE}(p) = \frac{N_h^{\mathrm{clean}}(p)}{N_{\mathrm{ev}}(p)} \tag{5.3}$$

Thanks to its extremely long mean life ($\tau_{^{14}\mathrm{C}} = 8270\,\mathrm{y}$) the ^{14}C rate is stable at a high level of precision, therefore an ideal QE trend is expected to be constant in time. Note that we are not interested in the absolute value of QE, but in the relative evolution in time.

Investigating in detail the standard approach for the QE computation, it was found that it suffers from two weaknesses:

1. The selection energy window for ^{14}C is fixed in time. Since the number of live PMTs decreases in time, this energy window corresponds to higher energy events as time goes by. This implies a higher number of collected photons, which turns into an apparent increase of QE.

2. The average value of the QE distribution, for a specific time window, is forced to be stable in time. This assumption generates a bias in the QE calculation procedure: in principle it not valid, since the average PMT performance could change in time. For example, $\langle QE \rangle$ could decrease because of the ageing of the PMT performances, reflecting on worse detection performances. As a consequence, this overall scaling may induce unphysical values $\mathrm{QE}(p) > 1$ for some of the PMTs in order to match the required $\langle \mathrm{QE} \rangle$; to avoid this, the algorithm forces it to the maximum allowed value $\mathrm{QE}(p) = 1$.

These two effects stand out especially when the number of alive PMTs is significantly decreased with respect to the initial status, explaining why the energy scale inconsistency showed up only from 2016 on.

5.4.1 QE upgrade

The algorithm to calculate the QE has been upgraded to solve the two weaknesses previously mentioned, as follows:

1. In order to avoid the first issue, the ^{14}C event selection is performed using only a subset of PMTs, named *B900 PMTs set*. This set is formed by PMTs which have survived the entire Phase-III-TMP until now. Furthermore, they have been selected among the ones with lower dark noise. In this case the number of active PMTs included in the selection is constant in time, and so the energy of the selected events is.

2. The average value of the QE distribution is no more scaled to a fixed value in time.

The average number of hits for ^{14}C events as a function of time calculated considering only fired B900 PMTs is shown in Fig. 5.9, within a 1 m or 2 m radius sphere (respectively in blue or orange lines). The rate is almost stable in time, displaying maximum variations at the $\approx 0.3\%$ level for $r < 1\,\mathrm{m}$ in six years of data taking.

The new procedure to compute QE is now described. ^{14}C events are selected according to the geometrically normalized hits variable $50 \leq N_h^{\mathrm{ng}} \leq 100$ ($0.10\,\mathrm{MeV} \lesssim E \lesssim 0.21\,\mathrm{MeV}$), on a weekly basis. The notations $N_h(p, \mathrm{ev})$ or $N_h(p, r)$ refer, from now on, to the number of geometrically normalized hits N_h^{ng} for a given PMT p for an event or a run respectively. The B apex refers to the variables calculated over the B900 PMTs set.

$$\text{B900 PMTs:} \qquad N_h^B(p, r) = \sum_{ev=1}^{N_{\mathrm{ev}}} N_h^B(p, \mathrm{ev}) \qquad t^B_{\mathrm{cluster}}(r) = \sum_{\mathrm{ev}=1}^{N_{\mathrm{ev}}} t_{\mathrm{cluster}}(\mathrm{ev}) \tag{5.4}$$

$$\text{All PMTs:} \qquad N_h(p, r) = \sum_{ev=1}^{N_{\mathrm{ev}}} N_h(p, ev) \qquad t_{\mathrm{cluster}}(r) = \sum_{\mathrm{ev}=1}^{N_{\mathrm{ev}}} t^B_{\mathrm{cluster}}(\mathrm{ev}) \tag{5.5}$$

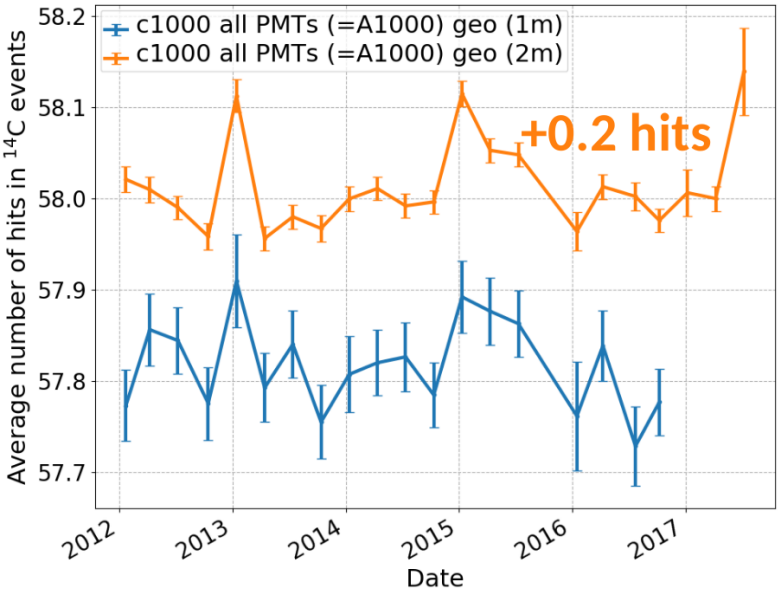


Figure 5.9: Monte Carlo simulations: energy in N_h^{n} for ^{14}C events simulated in different years, calculated for the B900 PMTs set over a 1 m or 2 m radius sphere (respectively in blue or orange lines). The events rate is almost stable in time, displaying maximum variations at the $\approx 0.3\%$ level for $r < 1$ m in six years of data taking.

For each enabled PMT, the total number of dark noise events[4] $N^{\mathrm{dark}}(r)$ is selected; the sum of hits for each run, assuming a gate length $t_{\mathrm{gate}}(r)$ for that run, is:

$$N_h^{\mathrm{dark},B}(p,r) = \sum_{ev=1}^{N_{\mathrm{ev}}^{\mathrm{dark}}} N_h^B(p,\mathrm{ev}) \tag{5.6}$$

Summing over the runs where the PMT was enabled, one gets the number of events for a given PMT (as in Eq. 5.2):

$$N_{\mathrm{ev}}^B(p) = \sum_{r|p\,\mathrm{enabled}}^{r\in t} N_{\mathrm{ev}}^B(r) \qquad N_{\mathrm{ev}}(p) = \sum_{r|p\,\mathrm{enabled}}^{r\in t} N_{\mathrm{ev}}(r) \tag{5.7}$$

$$N_h^B(p) = \sum_{r|h\,\mathrm{enabled}}^{r\in t} N_h^B(p,r) \qquad N_h(p) = \sum_{r|p\,\mathrm{enabled}}^{r\in t} N_h(p,r) \tag{5.8}$$

The dark rate and the dark hits can be calculated in each PMT for the resulting time window as

$$R^{\mathrm{dark}}(p) = \frac{N_h^{\mathrm{dark}}(p)}{t_{\mathrm{dark}}(p)} = \frac{\sum_r N_h^{\mathrm{dark}}(p,r)}{\sum_r N^{\mathrm{dark}}(r)\cdot t_{\mathrm{gate}}(h)} \tag{5.9}$$

[4]Dark noise events are selected by means of 0-cluster condition and by a dedicated random trigger.

The *clean hits* are free from the dark noise contributions and are the proper quantity for the direct calculation of QE, obtained from the subtractions $N_h^{B,\text{clean}}(p) = N_h^B(p) - N_h^{\text{dark}}(p)$ and $N_h^{\text{clean}}(p) = N_h(p) - N_h^{\text{dark}}(p)$:

$$\text{QE}(p) = \frac{N_h^{B,\text{clean}}(p)}{N_{\text{ev}}(p)} \tag{5.10}$$

This new QE Eq. (5.10) is biased by the user choice of best PMTs. The QE adjusted values can be obtained correcting as follows:

$$\text{QE}^{\text{corr}}(h) = \begin{cases} \text{QE}(p) & \text{if } p \in \text{B900} \\ \text{QE}(p) \cdot \frac{R_{\text{biased}}}{R_{\text{true}}} & \text{if } p \notin \text{B900} \end{cases} \tag{5.11}$$

$$R_{\text{biased}} = \frac{\overline{\text{QE}^B(p)},\ p \in \text{B900}}{\overline{\text{QE}^B(p)},\ p \notin \text{B900}} \qquad R_{\text{true}} = \frac{\overline{\text{QE}(p)},\ p \in \text{B900}}{\overline{\text{QE}(p)},\ p \notin \text{B900}} \tag{5.12}$$

For a PMT with concentrator, its QE(p) value is then multiplied by a scaling factor calculated via Monte Carlo simulations. Finally, the obtained QE distribution is scaled to be less than 1 (*relative QE*).

In order to verify that the new QE set improves the Monte Carlo energy scale reproduction with respect to the previous one, the same plot already shown in Fig. 5.7 is reproduced in Fig. 5.10, including the results obtained with the new QE. The evolution in time of the ^{210}Po peak energy obtained from simulations with the *NewQE* (blue trend), matches much better the data (green trend), at the $\approx 0.5\%$ level, with respect to simulations performed with the old set of QE, named from now on *StdQE* (red trend).

5.5 Validation of the new QE on calibration data

I have performed an important cross-check of the newly produced QE by analyzing the calibration data taken in 2009. I exploited the data collected during the several calibration runs when the source has been deployed in the center of the detector; in particular, in the following I will show results for the ^{85}Sr source, a 551 keV γ emitter. The main results and conclusions shown in this section have been checked also for several γ calibration runs, deployed in different position within the calibrations $r < 3$ m FV.

The main goal of this study is to validate the new approach for the QE calculations, described in the previous Sec. 5.4. In order to do this, I compared the following set of data:

1. *Data*: ^{85}Sr γ source calibration data.
2. *MC Std*: simulated ^{85}Sr data with the Monte Carlo using the StdQE (as computed in Sec. 5.4).
3. *MC New*: simulated ^{85}Sr data with the Monte Carlo using the NewQE.

For each set of data I calculated the QE for every PMT according to the following procedure: for each PMT, for the events in a $r < 1$ m radius sphere, the total number of hits collected by a PMT is

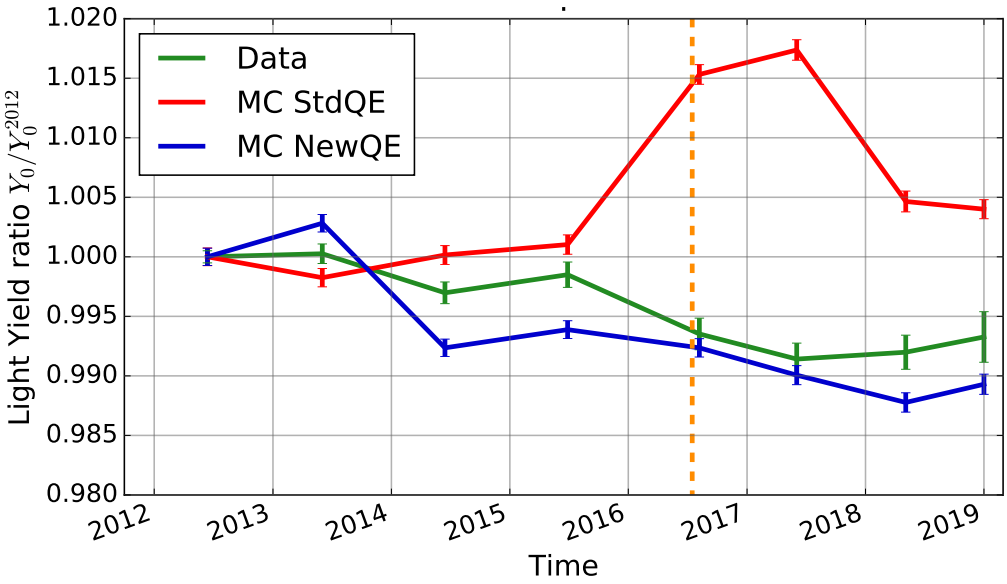


Figure 5.10: Evolution of Borexino energy scale in time, by analyzing ^{210}Po events: relative variations of light yield parameter Y_0 with respect to the 2012 value Y_0^{2012} for data (green) Monte Carlo simulations assuming StdQE set (red), Monte Carlo simulations assuming NewQE set (blue). The beginning of Phase-III-TMP dataset is marked by the orange dashed line.

divided by the total number of detected events in the run. The three QE distributions obtained in this way (named *DataQE*, *MC StdQE*, *MC NewQE*), can be seen in Fig. 5.11, both for PMTs with light concentrator (red histograms) and for PMTs without light concentrators (blue histograms). It is clear that the distribution obtained with the MC NewQE matches better the data one with respect to StdQE one. This can also be seen from the difference plot (bottom right histogram of Fig. 5.11).

This improvement can be seen even more clearly by the correlation plots of MC StdQE vs DataQE and MC NewQE vs DataQE, shown in left and right panel respectively of Fig. 5.12. The correlation coefficients, $R^2_{\mathrm{Data-Std}} = 0.761$ and $R^2_{\mathrm{Data-New}} = 0.895$ respectively, show that the NewQE sets follows better the DataQE profile.

The angular (θ, φ) NewQE map for PMTs with concentrators is shown in Fig. 5.13. According to a spherical coordinate system, $-\pi < \theta < \pi$ is the polar angle and $0 \leq \phi \leq \pi$ is the azimuthal angle. This map has been compared to the similar DataQE one, shown in Fig. 5.14, and are found to be consistent: no significant differences or anomalies show up. A spot of high-performing PMTs can be detected in the $(\theta, \phi) \in [250°, 350°] \times [-50°, 0°]$ angular region, easily identifiable in Fig. 5.13 and Fig. 5.14 by the grouped blue-scale circles. Note that, integrating on the two North and South hemispheres, the latter one shows a larger average QE value: this implies that the southern PMTs are more performant with respect to the North ones, showing a more efficient light collection. This is confirmed both by data and Monte Carlo simulations.

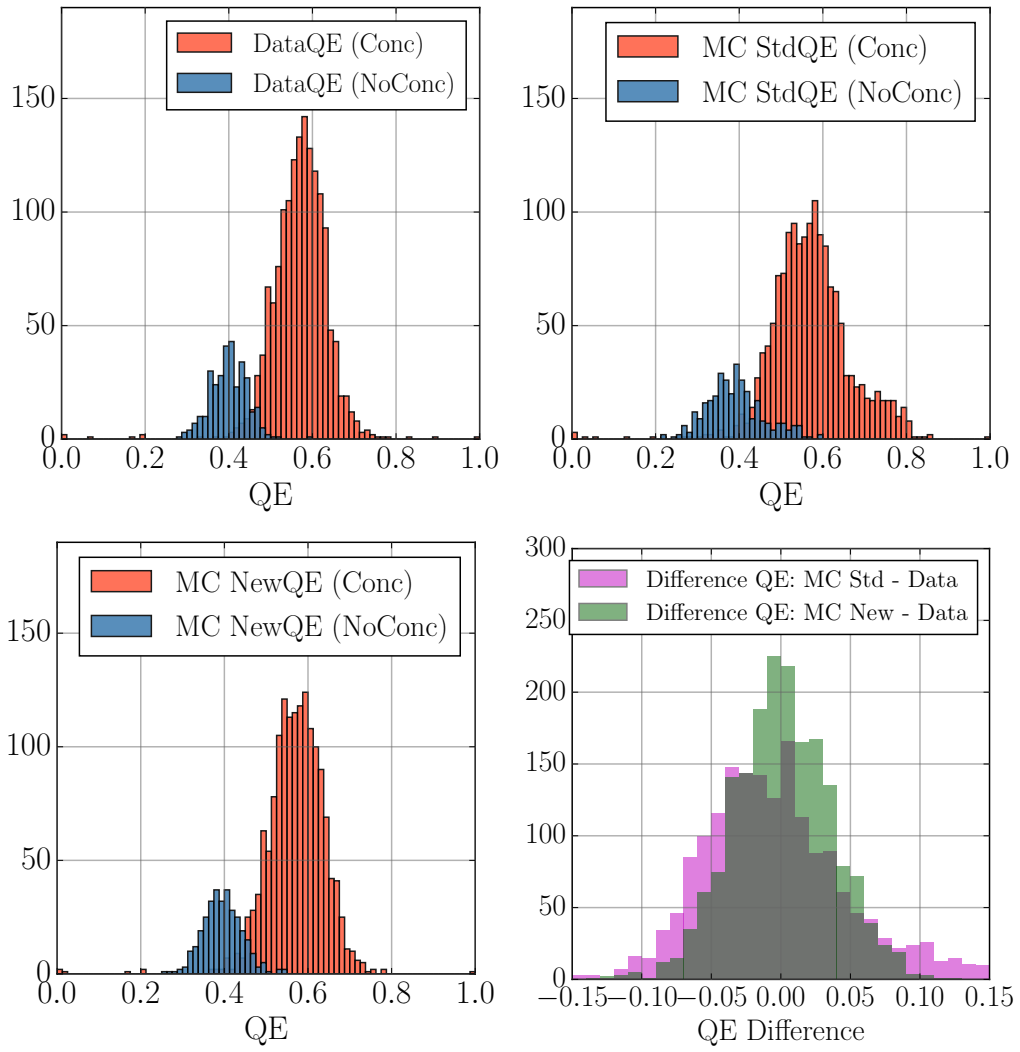


Figure 5.11: QE distributions calculated from three set of data: ^{85}Sr γ calibration data (upper left plot, *DataQE*), ^{85}Sr γ simulated data assuming StdQE (upper right plot, *MC StdQE* set), ^{85}Sr γ simulated data assuming NewQE (lower left plot, *MC NewQE* set). Concentrators/No concentrators PMTs QE are respectively shown in red or blue histograms. The lower right plot shows the diffferences between QE calculated from the previous distributions: pink histogram for MC StdQE and MC DataQE difference, green histogram for NewQE and DataQE difference.

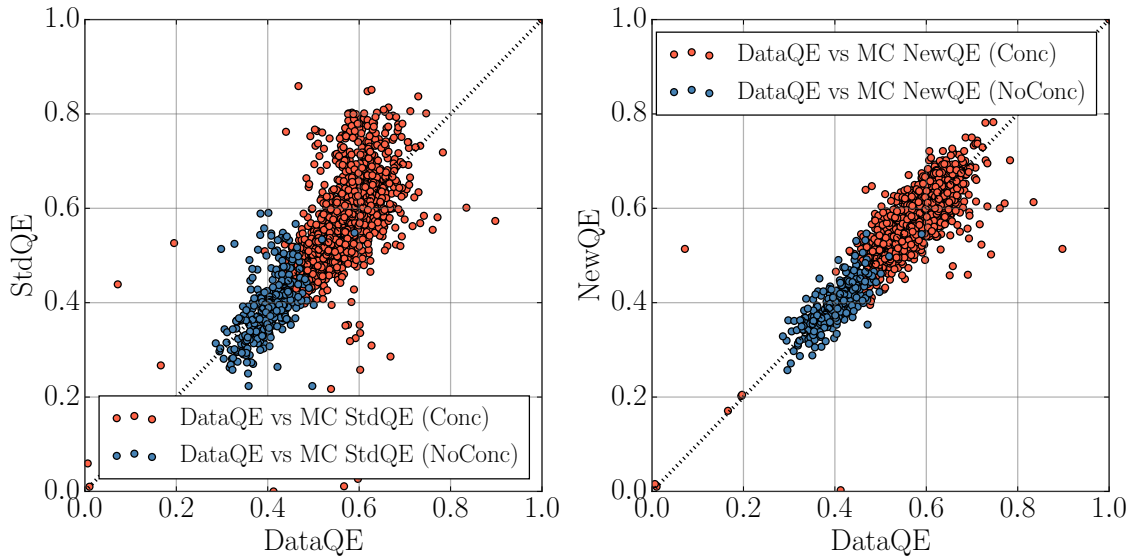


Figure 5.12: Correlation plots of MC StdQE vs DataQE (left panel) and MC NewQE vs DataQE (right panel). The correlation coefficients are respectively $R^2_{\text{Data-Std}} = 0.761$ and $R^2_{\text{Data-New}} = 0.895$. The black dashed diagonal line is just a reference, showing a perfectly correlated trend ($R^2 = 1$).

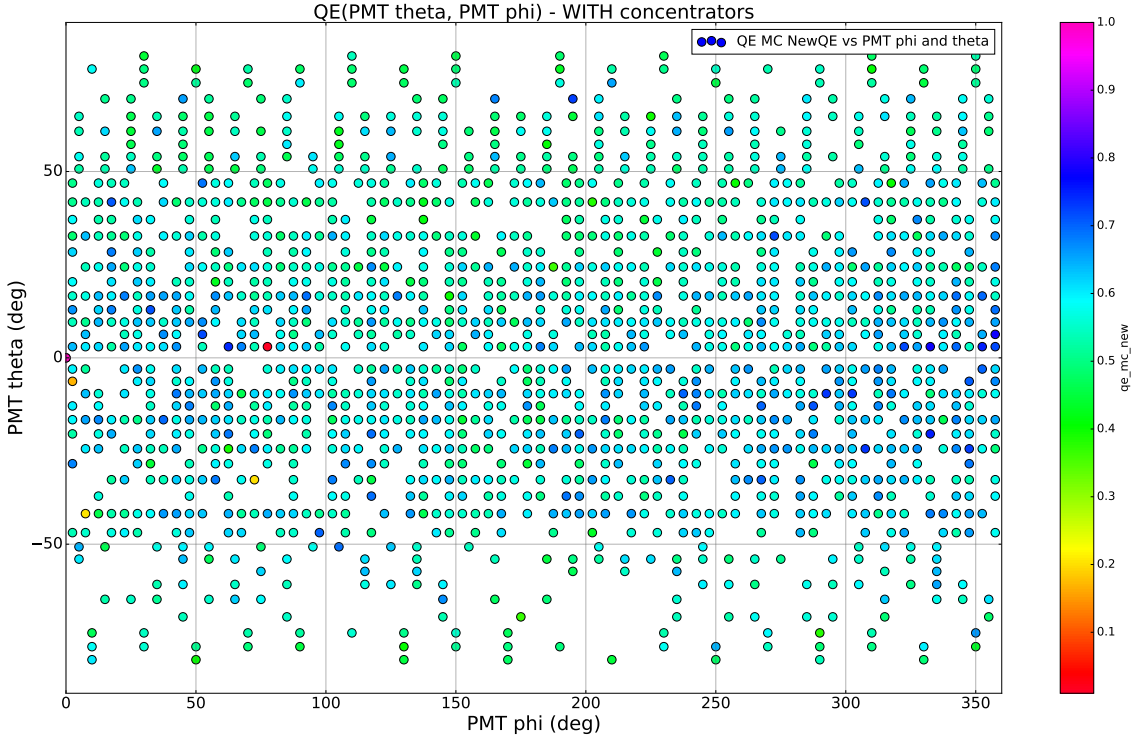


Figure 5.13: Monte Carlo simulations: **angular** (θ, φ) **map of the MC NewQE dataset** for PMTs with light concentrators.

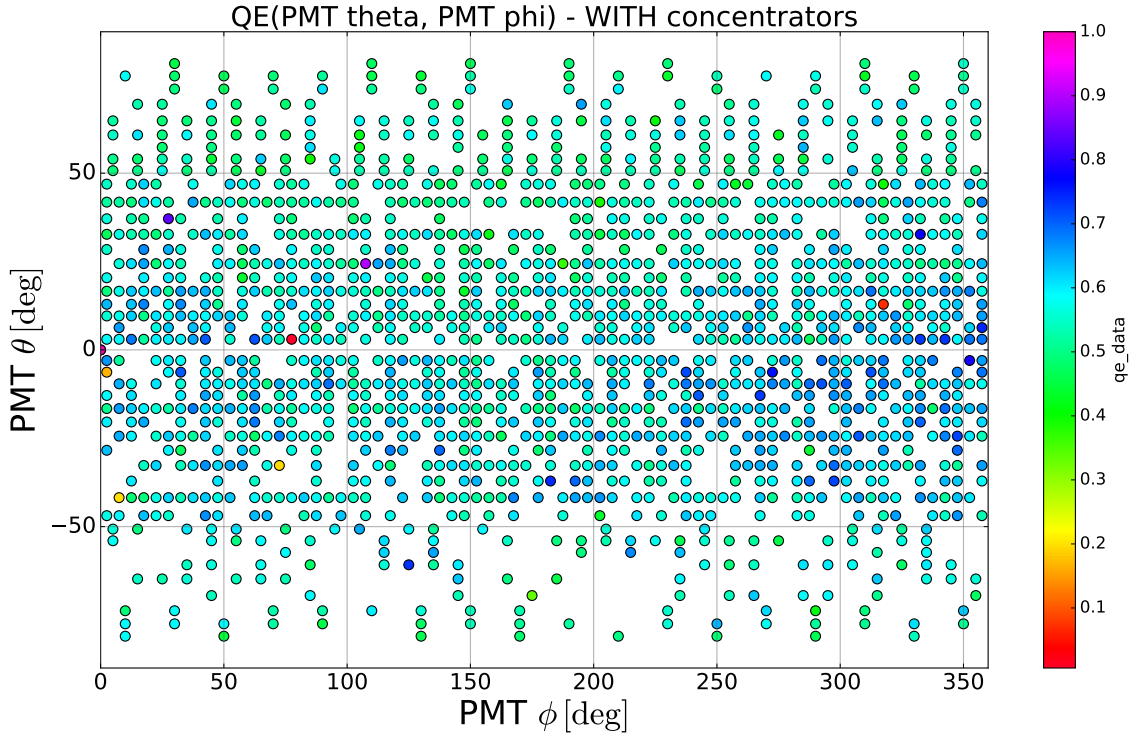


Figure 5.14: Angular (θ, φ) map of DataQE for PMTs with light concentrators.

5.6 Monte Carlo tuning after the QE upgrade

In principle, the upgrade on the QE calculations should impact only on the evolution of the QE in time, giving a better agreement between Monte Carlo simulations and data after 2016. The agreement between Monte Carlo and calibration data taken in 2009 should remain unchanged with respect to the simulation performed with the old QE. However, since the QE impact on the uniformity of the response function, which in turn could affect other observables, we cannot a priori exclude that the tuning of the Monte Carlo input parameters should be re-done once Monte Carlo uses the new QE set.

In order to study this, I have performed a very careful comparison of all the relevant observables obtained on γ calibration data and on Monte Carlo simulations with the NewQE and the StdQE sets. Some examples of this comparison are shown in Fig. 5.15, Fig. 5.16, Fig. 5.17 and Fig. 5.18 for the γ ^{65}Zn source, where I plot the distributions of two energy estimators (N_p^{n} and N_h^{n}), the photon time distributions, the reconstructed position, the cone/no-cone ratio, for data (black), Monte Carlo simulations with StdQE (red) and Monte Carlo simulations with NewQE (blue).

For each one of the comparison, a χ^2 estimator is built to quantify the agreement between data and Monte Carlo, for both StdQE and NewQE. This is done for each of the 40 physical quantities of interest and for all the 14 γ calibration runs used in the tuning. The results are condensed in Fig. 5.19, where the distributions of all the χ^2/NDF are plotted for Monte Carlo with StdQE and NewQE (top and bottom panels, respectively). These two distributions are similar, with a 1.38 median value for StdQE and 1.31 for NewQE, showing that the latter ones perform similarly or slightly better with respect to the StdQE ones. In conclusion, we don't have found any compelling motivation to perform a further global tuning of Monte Carlo after the QE update.

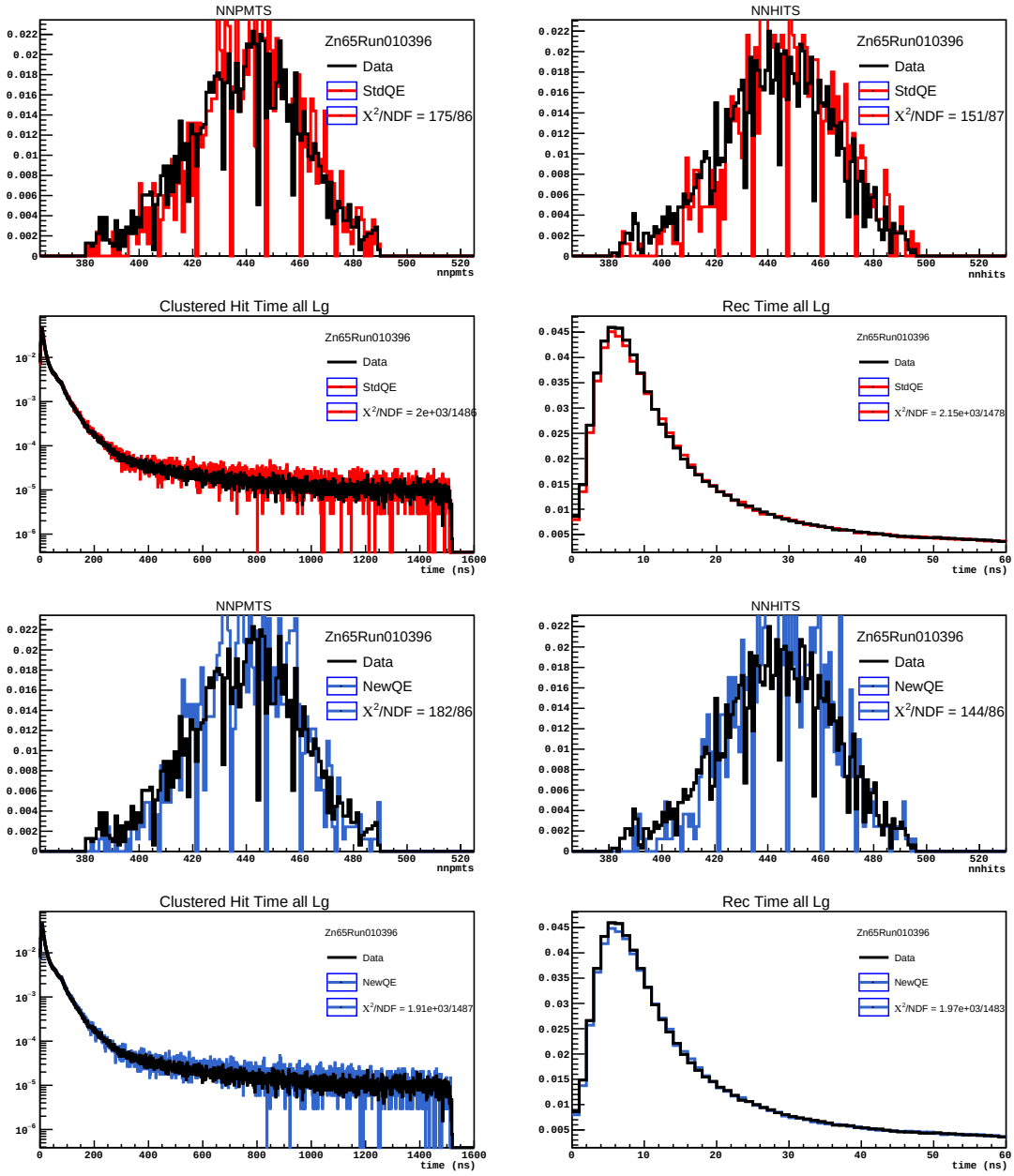


Figure 5.15: Data-MC agreement for basic energy and position variables for ^{65}Zn calibration events in the center, for StdQE (red histograms) and NewQE (blue histograms) simulations. For each quadruplet, the following variables are shown: N_p^{n} (top left panel), N_h^{n} (top right panel), cluster hit time (bottom left panel), reconstructed time (bottom right panel).

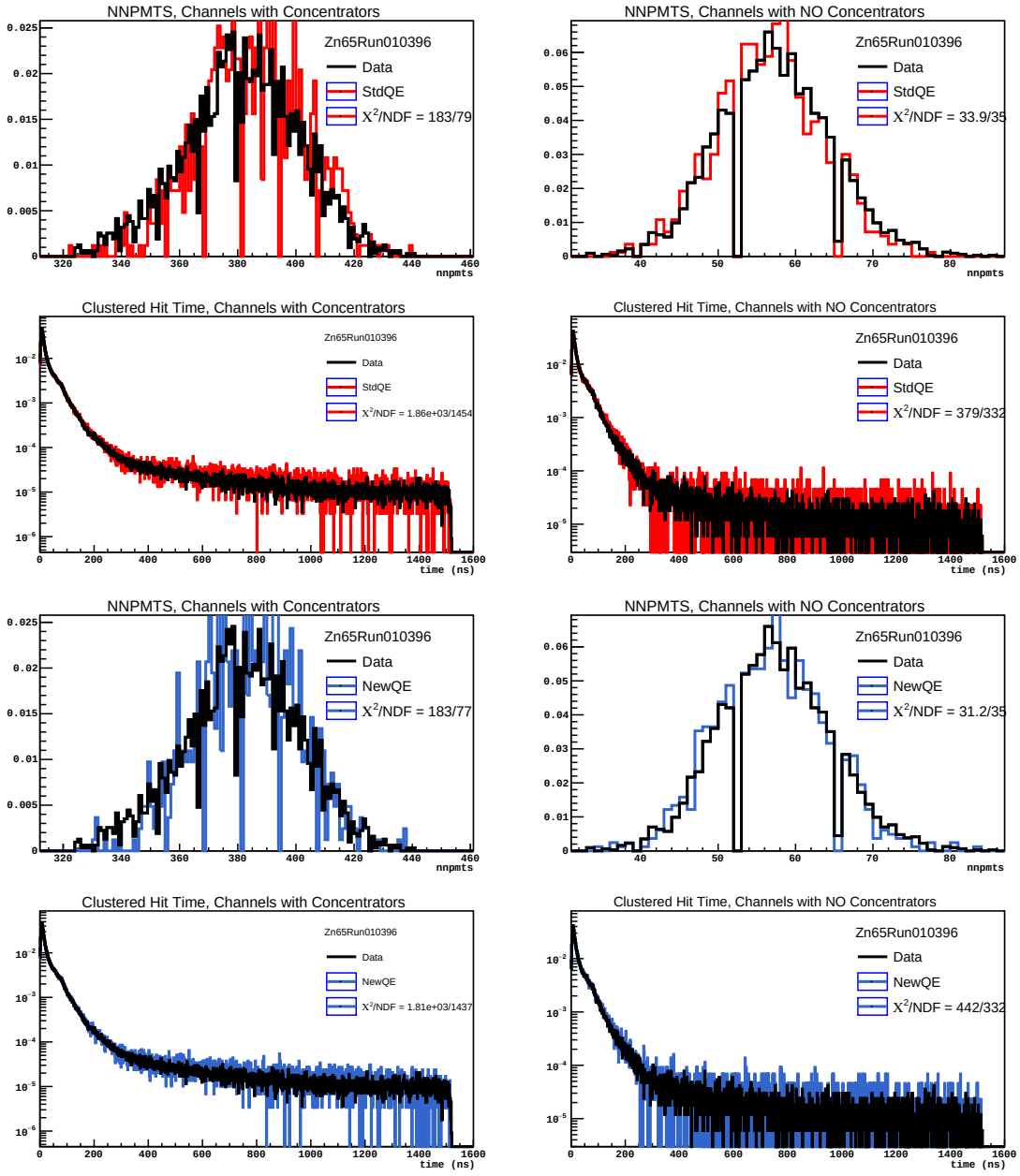


Figure 5.16: Data-MC agreement for variables for "cone PMTs" or "no cone PMTs" subsets (i.e. variables calculated considering only PMTs with / without the light cocentrators), for ^{65}Zn calibration events in the center, for StdQE (red histograms) and NewQE (blue histograms) simulations. For each quadruplet, the following variables are shown: N_p^{n} for cone PMTs (top left panel), N_p^{n} for no cone PMTs (top right panel), cluster hit time for cone PMTs (bottom left panel), cluster hit time for no cone PMTs (bottom right panel).

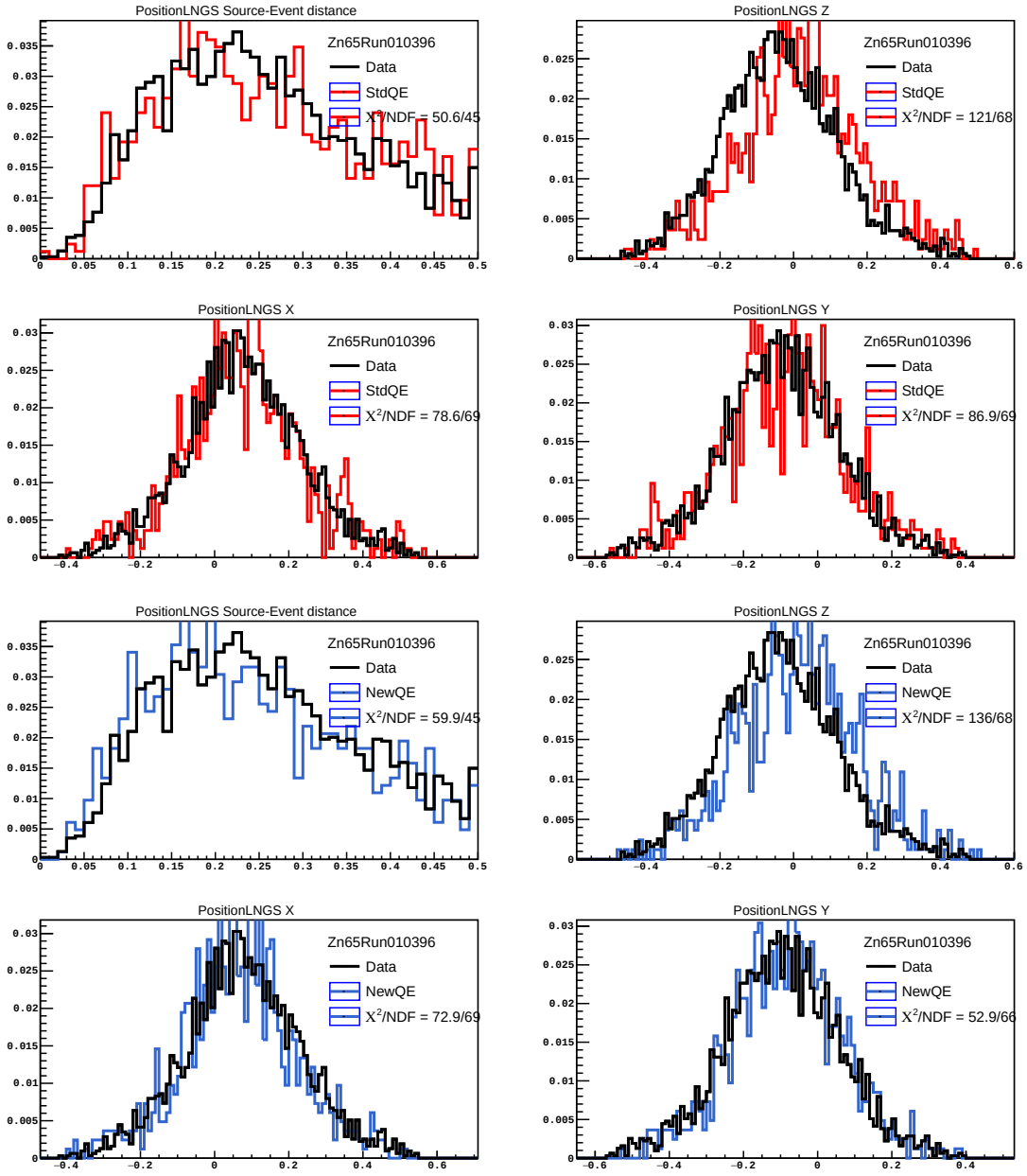


Figure 5.17: Data-MC agreement for position variables for ^{65}Zn calibration events in the center, for StdQE (red histograms) and NewQE (blue histograms) simulations. For each quadruplet, the following variables are shown: reconstructed radial distance r (top left panel), x reconstructed coordinate (bottom left panel), y reconstructed coordinate (top right panel), z reconstructed coordinate (bottom right panel).

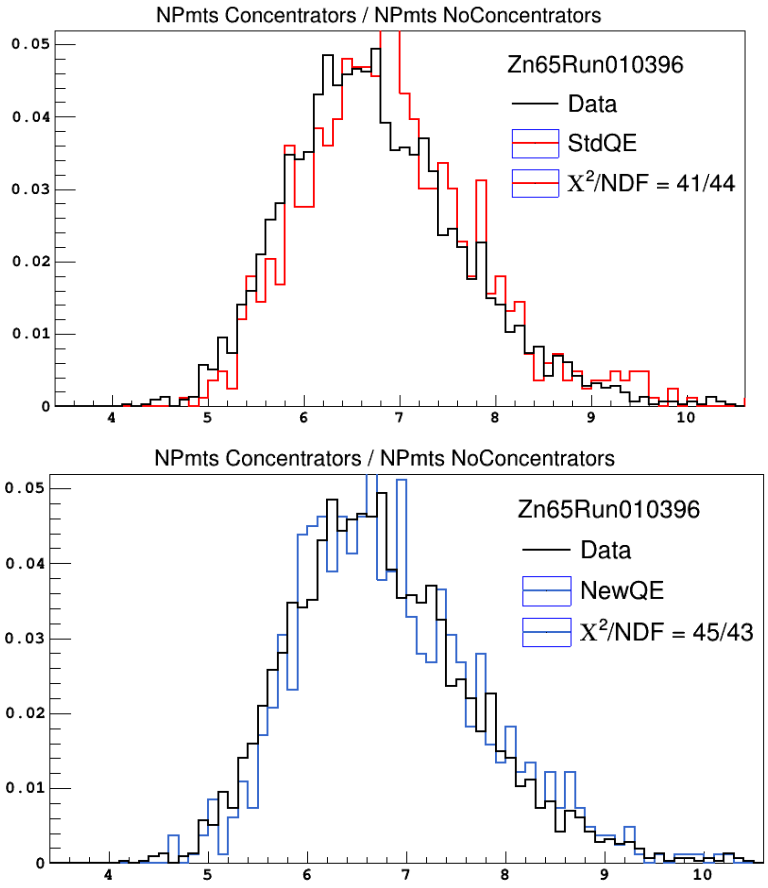


Figure 5.18: Data-MC agreement for the ratio between number of photons collected by PMTs with light concentrators on the number of photons collected by PMTs without light concentrators. The distributions are shown for ^{65}Zn calibration events in the center, for data (black histogram), MC StdQE simulations (red histogram) and MC NewQE simulations (blue histogram).

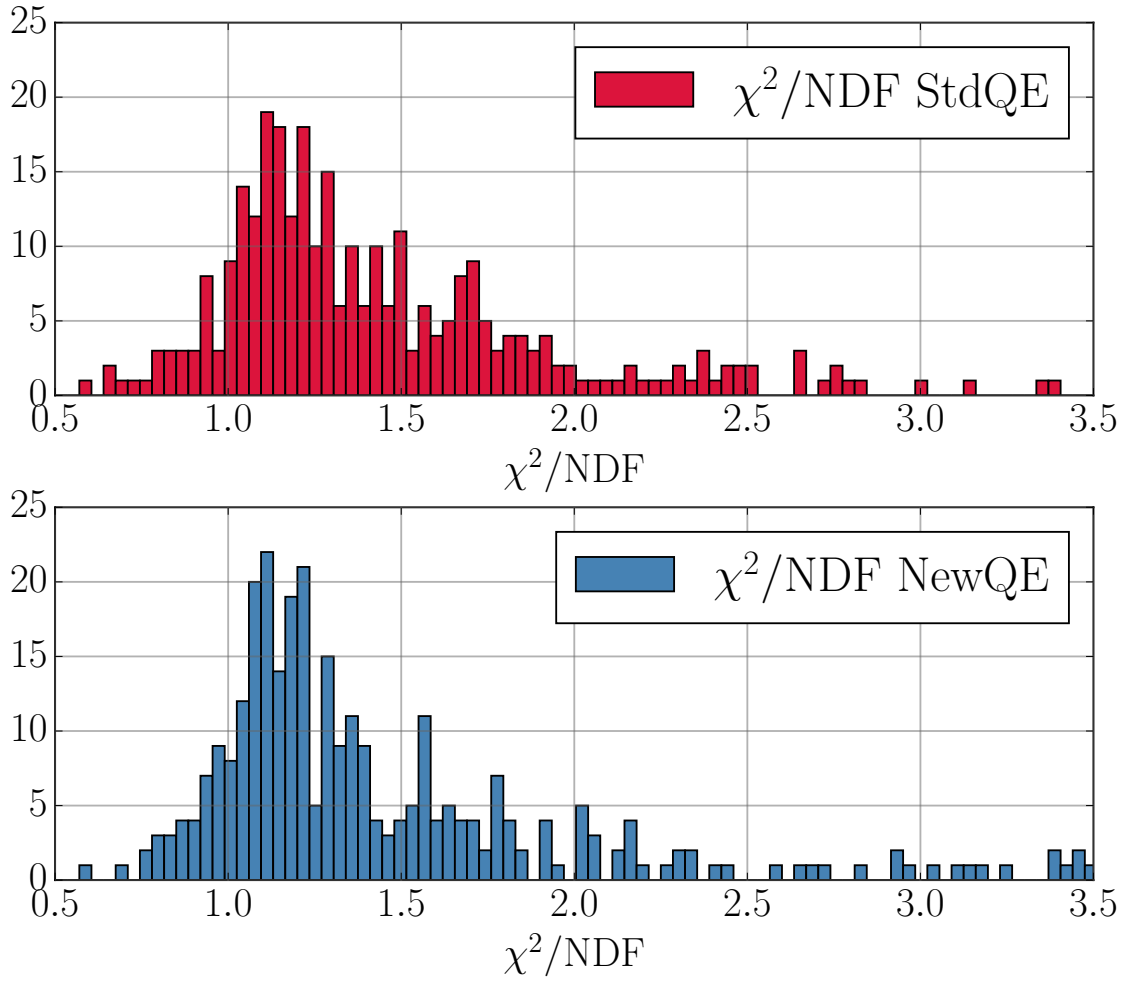


Figure 5.19: Data-MC χ^2/NDF agreements for the several physical quantities describing the detector response: StdQE histogram (top panel), NewQE histogram (central panel). The median values of the two distributions are respectively 1.38 and 1.31, showing that the NewQE perform similarly or slightly better with respect to the StdQE ones.

.

5.6.1 Change of the Geant4 version

Together with the QE update, the collaboration decided to switch to the version 10.5 of Geant4, with respect to the version 10.2 which was employed so far. The main expected changes are related to the quenching factors update, as shown in Fig. 5.20, which in the Borexino range of interest ($10\,\text{keV} \lesssim E \lesssim 10\,\text{MeV}$) affects mainly the α energy scale (left panel). The β quenching factors (right panel) are instead mostly unchanged in the Borexino energy window of interest. After having demonstrated that the other Monte Carlo parameters are unaffected by the Geant4 version upgrade, I finely tuned the Birks factor for α particles, k_α, which drives the α energy scale (see Sec. 5.1). ^{210}Po events have been selected over three months of data to get a reasonable statistics ($\gtrsim 10^4$ events), and into a very strict central volume ($r < 1\,\text{m}$) to limit as much as possible the position dependent effects for energy reconstruction (see Fig. 5.6). The performed α tuning allowed to restore back the energy scale, slightly improving the data-MC agreement with respect to the one obtained in the previous tuning (see Fig. 5.21).

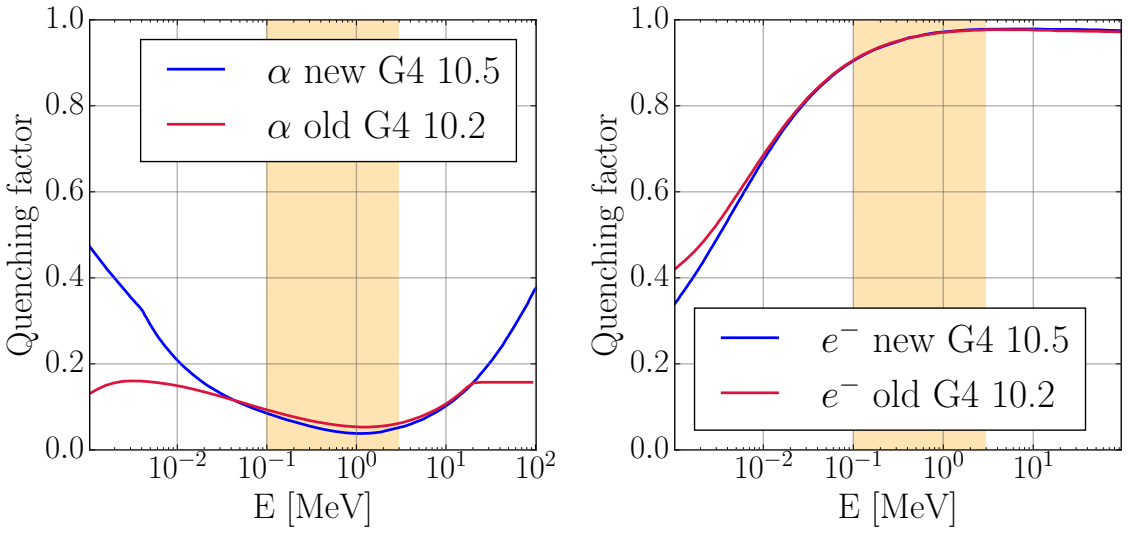


Figure 5.20: Quenching factors for α and e^- events simulations (respectively, left and right panels), for the Geant4 10.2 and 10.5 versions. Orange bands cover the energy region of interest for solar neutrino detection with Borexino.

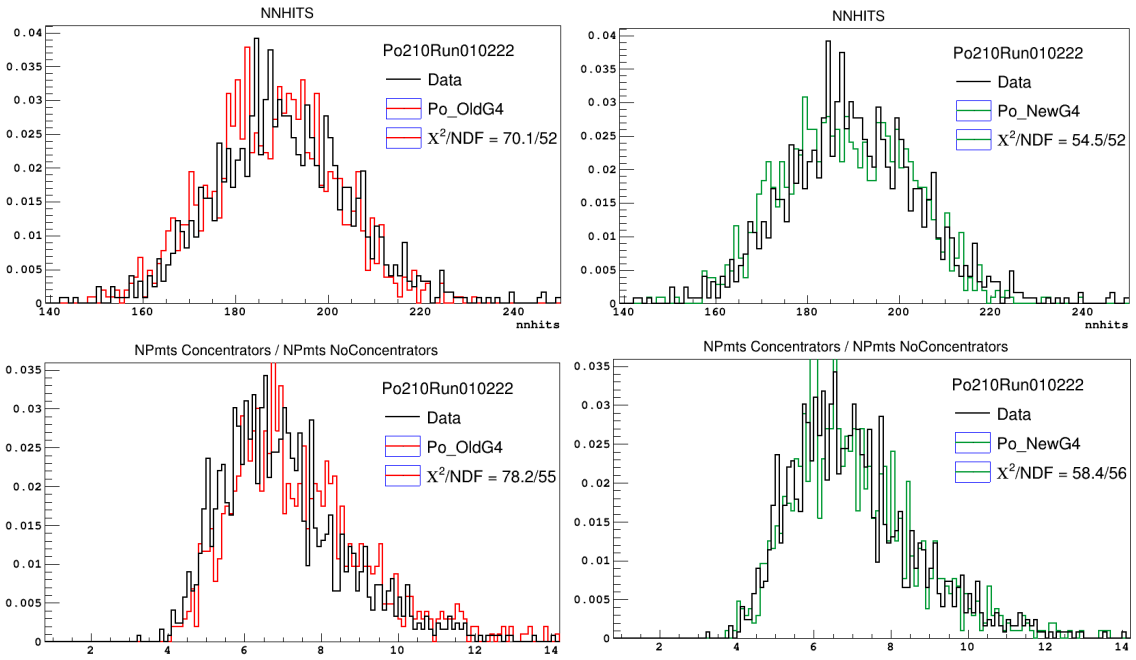


Figure 5.21: ^{210}Po **energy peak value for data** (black histogram), old 10.2 Geant4 version simulations (red histogram) and updated 10.5 Geant4 version (green histograms) simulations: N_h^n spectrum (upper plots) and N_p^n concentrators/no-concentrators ratio (lower plots). Only new QE has been employed. After the new tuning stage, the data-MC agreement, quantified by the χ^2/NDF value, is slightly improved for both the distributions.

5.7 Simulations of the Phase-III-TMP PDFs for the Monte Carlo multivariate fit

Once the QE upgrade and the following tuning stages have been completed, I performed a massive simulation of all the signal and background PDFs relative to Phase-III-TMP dataset. This is needed in order to perform the multivariate fit to extract the CNO-ν signal[5]. The technical steps to create the Monte Carlo based PDFs are summarized hearafter and are schematized in Fig. 5.22.

- **Events generation and reconstruction.** For each species of interest, events are simulated on a run-by-run basis, according to the real detector conditions (number of live PMTs, trigger threshold, and so on). The chosen number of generated events is proportional to the run effective livetime and in the same time is set to largely exceed the number of detected events on data: 200 times higher for the all species except for external background (only 50 times higher, being very time consuming) and for ^{14}C. As discussed in Sec. 4.2.2, ^{14}C rate is extremely high ($\approx 40\,$Bq/100t) and it is therefore not possible to simulate it with the same statistics; so the number of generated events is reduced by a factor 3. Therefore, except for ^{14}C, the statistical uncertainty associated to the Monte Carlo simulations can be neglected. All the simulated data are then processed with the reconstruction code Echidna, exactly as for real data.
- **Data quality filtering.** Once reconstructed, the events undergo the standard selection cut sequence described in Sec. 4.1. The basic quality cuts and the FV selection are applied. The runs belonging to the Phase-III-TMP dataset are then merged in a single file.
- **Normalization.** The obtained histograms need to be normalized according to the ratio between generated and reconstructed events. This step implicitely takes into account also the trigger threshold effect, which is particularly important for species whose spectrum lies in the low energy region.
- **Final PDF file.** For each signal and background component, and for the entire Phase-III-TMP, the final output files contain the energy PDFs for several normalized estimators: N_h^{n}, N_p^{n}, N_{pe}^{n} and the associated dt1 and dt2 variants (see Sec. 2.4.1). These files are given as inputs to the multivariate fitter.

As an example, Fig. 5.23 shows the energy PDFs for for CNO-ν (top) and ^{7}Be-ν (bottom) species, built with the new Monte Carlo (blue profile) and compared with the ones simulated with the old Monte Carlo (red profile): the effect of the QE upgrade is clearly visible as a shift towards lower energy values.

5.8 Monte Carlo fit in Phase-III-TMP with CNO-ν rate fixed

After completing my work on the upgrade of the Monte Carlo simulations, I have performed a first test of the newly produced PDFs by fitting Phase-III-TMP data in the so-called *pp chain configuration*, i.e., keeping the CNO-ν interaction rate fixed to the Standard Solar Model prediction,

[5] the other alternative to produce the same PDFs consists in using an analytical model for the energy response description, as described in Sec. 4.2.1

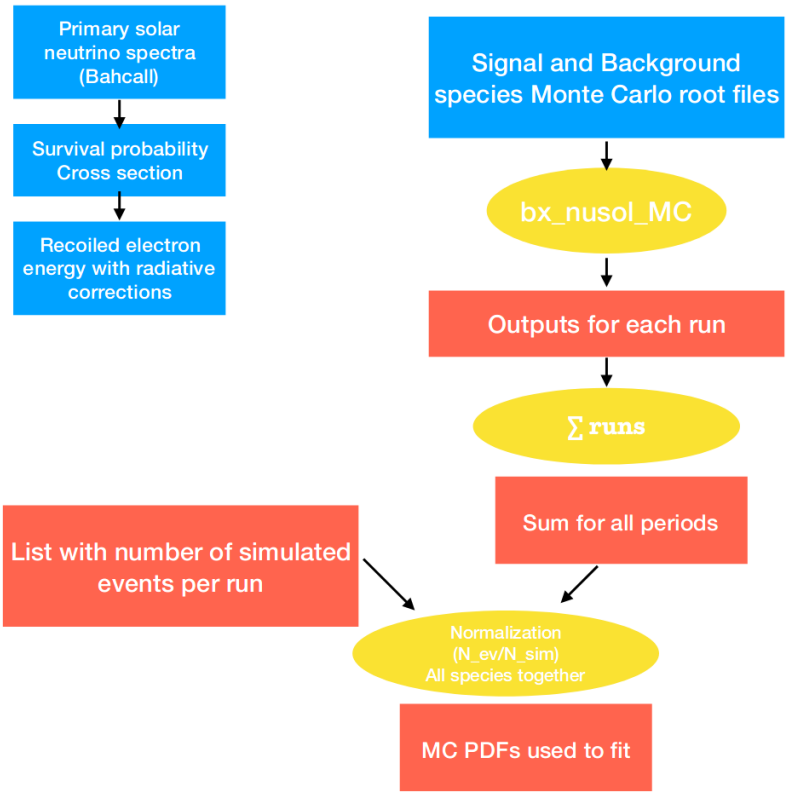


Figure 5.22: Sequence of the Monte Carlo PDFs construction steps.

in the high metallicity scenario (HZ-SSM), and leaving free the interaction rates of the other neutrino species. This study is mainly meant to verify that, after the upgrade of the QE set, the Monte Carlo is able of fitting data even after 2016, with results consistent with the Phase-II one. The results of this test have been anticipated in Sec. 4.3.2 and shown in Tab. 4.3 and Fig. 4.11 , proving that the new Monte Carlo reference shapes reproduce well data (χ^2/NDF = 1.3, similar to the one obtained for Phase-II analysis) and the neutrino results are consistent with the one obtained by Borexino in Phase-II and published in Ref. [39].

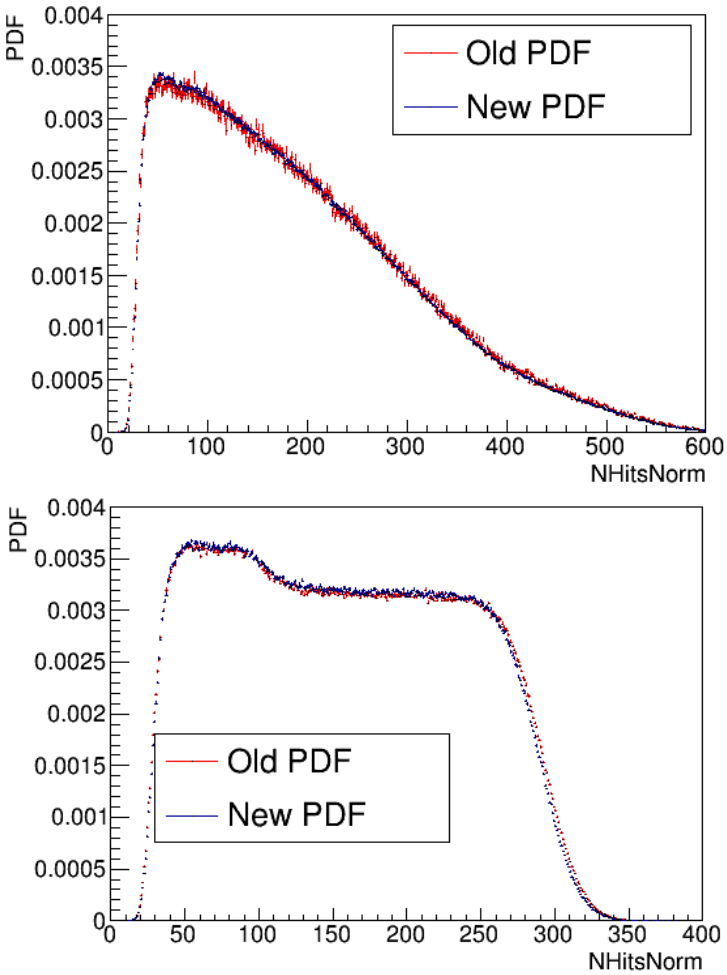


Figure 5.23: Exemplary energy PDFs for CNO-ν (top) and ^{7}Be-ν (bottom), built with the new Monte Carlo (blue profile) and compared with the ones simulated with the old Monte Carlo (red profile): the effect of the QE upgrade is clearly visible as a shift towards lower energy values.

5.9 Conclusions

This Chapter has been fully devoted to the Borexino Monte Carlo code and in particular to its upgrade, which is a crucial step for the CNO-ν analysis. Together with the analysis team, I have spent a significant effort in understanding the issue which was affecting the previous version of the code and which prevented the multivariate analysis fit of data after 2016. The problem was found to be related to the quantum efficiencies given in input to the simulations and was solved adopting an upgraded alghoritm to calculate them. I have tested the new QE inputs and verified that it is adequate for our purposes. By means of the calibration data re-analysis, I have also verified that the upgraded Monte Carlo does not require new global tuning. Finally, I have produced all the signal and background reference shapes needed for the multivariate fit, and I performed the first fit of Phase-III-TMP data in the so-called pp chain configuration, that is, fixing the CNO-ν rate. This test shows that the fit performs well and gives results compatible to the one published in Ref. [39].

We can conclude that the Monte Carlo is ready to perform the analysis to extract the CNO-ν signal which is the main goal of this thesis and will be described in details in Chapter 7.

Chapter 6

Determination of the ^{210}Bi background rate for the CNO neutrinos analysis

In Chapter 3 we have shown, by means of dedicated sensitivity studies, that Borexino has the chance to isolate a CNO-ν signal with relevant significance if a constraint of the interaction rate on the most important background, the ^{210}Bi isotope, is set in the multivariate analysis (see Sec. 3.2). This is needed because of the similarity of ^{210}Bi and CNO-ν spectral shapes, which induces strong correlations in the fit reducing the sensitivity to CNO neutrinos. As seen from the sensitivity studies, the ^{210}Bi interaction rate must be determined with a precision better than 2 cpd/100t in order to reach a median sensitivity on CNO-ν of 3σ or more. In this Chapter I will discuss the method to determine the ^{210}Bi rate and the current results obtained with it. The ^{210}Bi rate constraint will be used in Chapter 7 to perform the multivariate fit and extract the CNO-ν signal.

The independent constraint on ^{210}Bi can be obtained by tagging its daughter isotope ^{210}Po , which decays α and can therefore be easily identified via α/β discrimination techniques. If the chain $^{210}\text{Pb} \rightarrow {}^{210}\text{Bi} \rightarrow {}^{210}\text{Po}$ is in secular equilibrium, the activity of ^{210}Po is *de facto* the same as the ^{210}Bi one (apart for negligible corrections due to the finite mean lifetime of ^{210}Pb), making the measurement straightforward.

Unfortunately, this strategy is complicated by additional ^{210}Po contributions out of equilibrium, which are present in the detector. In particular, the most annoying one is due to the ^{210}Po contamination present on the Inner Vessel which can be detached from the nylon and carried inside the innermost region of the scintillator (the Fiducial Volume in which the analysis is done) by convective currents. It has been observed that these convective currents are triggered by variations in the temperature profile of the scintillator. In order to minimize this effect, the detector has been thermally insulated starting from 2015. In spite of this, ^{210}Po motions due to convection have not stopped completely and a dedicated study of the spatial profile of the ^{210}Po distribution is needed to extract the supported term which is directly related to ^{210}Bi .

This chapter is structured as follows. Sec. 6.1 illustrates the ^{210}Bi -^{210}Po link, the key relation to extract an independent ^{210}Bi rate; in particular, Sec. 6.1.1 reports the ^{210}Po distributions in space and time in the Borexino data, along with the impact of temperature evolution and detector thermal stability. The strategy to extract the ^{210}Po supported rate is illustrated in Sec. 6.2. A qualitative method (*radial analysis*) based on an approximated model of the expected spatial distribution of ^{210}Po will be discussed in Sec. 6.2.1. The results on the ^{210}Po supported rate which I will use in the

CNO-ν analysis come from a more quantitative method (*clean bubble analysis*) which is reported in Sec. 6.2.2. I have validated this method with a toy Monte Carlo study, described in Sec. 6.2.3. Sec. 6.3 reports my studies on the time stability and homogeneity of ^{210}Bi decay rate, along with the final ^{210}Bi rate result (Sec. 6.3.2. Conclusions are outlined in Sec. 6.4.

6.1 The ^{210}Bi - ^{210}Po link

The strategy adopted in this thesis to constrain the ^{210}Bi rate is based on the fact that this isotope is part of a decay chain starting with ^{210}Pb:

$$^{210}\mathrm{Pb} \xrightarrow[Q=63.5\,\mathrm{keV}]{\tau\approx 32\,\mathrm{y}} {}^{210}\mathrm{Bi} \xrightarrow[Q=1160\,\mathrm{keV}]{\tau=7.23\,\mathrm{d}} {}^{210}\mathrm{Po} \xrightarrow[Q=5.3\,\mathrm{MeV}]{\tau=199.1\,\mathrm{d}} {}^{206}\mathrm{Pb} \tag{6.1}$$

If there are no external sources of ^{210}Po and ^{210}Bi , the time evolution of the three isotope activity is shown in Fig. 6.1, where we assumed (as an example) an initial condition of $R(^{210}\mathrm{Pb}, t=0) = 20\,\mathrm{cpd/100t}$. Because of its small mean lifetime ($\tau_{\mathrm{Bi}} = 7.23\,\mathrm{d}$) compared to the one of ^{210}Pb ($\tau_{\mathrm{Pb}} = 32\,\mathrm{y}$), the ^{210}Bi activity goes almost immediately in equilibrium with the ^{210}Pb one, while ^{210}Po reaches equilibrium after $\sim 2\,\mathrm{y}$. After the equilibrium is reached, the following relation between the rates of the three isotopes holds:

$$\tau_{\mathrm{Pb}} \gg \tau_{\mathrm{Bi}} \qquad \tau_{\mathrm{Pb}} \gg \tau_{\mathrm{Po}} \overset{t/\tau_{\mathrm{Po}} \gg 1}{\Longrightarrow} \qquad \frac{R_{\mathrm{Po}}}{R_{\mathrm{Bi}}} = \frac{R_{\mathrm{Po}}}{R_{\mathrm{Pb}}} = \frac{\tau_{\mathrm{Pb}}}{\tau_{\mathrm{Pb}} - \tau_{\mathrm{Po}}} \approx 1.015 \tag{6.2}$$

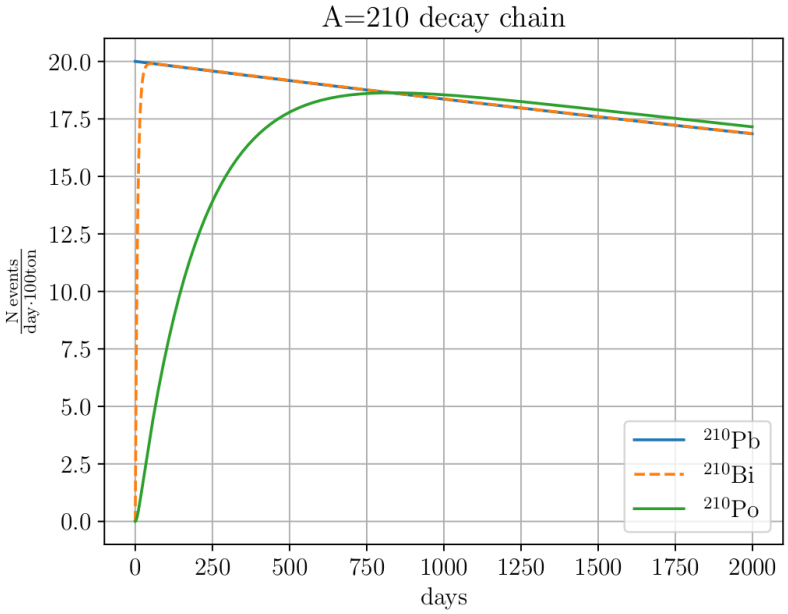


Figure 6.1: Time evolution of ^{210}Pb $\rightarrow$ ^{210}Bi $\rightarrow$ ^{210}Po chain decay species in absence of ^{210}Po source terms, assuming a reference ^{210}Pb rate $R_{\mathrm{Pb}}(t=0) = 20\,\mathrm{cpd/100t}$ (^{210}Pb in blue solid line, ^{210}Bi in orange dashed line, ^{210}Po in green solid line).

Consequently, at equilibrium the rate of ^{210}Bi is equal to the one of ^{210}Po with a very good approximation.

This gives us an hint on how to proceed: ^{210}Po decays emitting a mono-energetic α of 5.3 MeV, which is quenched in the scintillator to an equivalent energy of approximately 400 keV. The signature of this class of events in Borexino is therefore a very clear gaussian-like peak, which is easily recognizable with respect to events originated in β decays. Furthermore, α events can be distinguished from β ones by means of pulse-shape discrimination techniques, which exploit the different time distributions of photons emitted by α and β particles in the scintillator (see Sec 2.4.1). For all these reasons, measuring the ^{210}Po rate is relatively straightforward in Borexino. If the condition of secular equilibrium holds, this measurement provides the direct and independent estimation of the ^{210}Bi rate that we were looking for.

Unfortunately, the possibility to connect ^{210}Bi and ^{210}Po decay is complicated by two additional ^{210}Po contributions out of Eq. (6.1) chain equilibrium:

1. **Unsupported ^{210}Po rate from the scintillator purification phase.** A ^{210}Po contamination has been accidentally inserted in the scintillator during the Water Extraction purification phase, which took place 2010. Therefore, at the beginning of Phase-II, in 2012, a high ^{210}Po rate of ~ 1400 cpd/100t was present in the Fiducial Volume. Fortunately, this contamination was not supported by ^{210}Pb and therefore has decayed exponentially with the ^{210}Po mean lifetime. For the Phase-III analysis it contributes negligibly to the total ^{210}Po rate, since almost 6 years (corresponding to $\sim 8\tau_{\mathrm{Po}}$) have elapsed after the ^{210}Po insertion;

2. ^{210}Po **source term.** We know that ^{210}Pb (and therefore the full chain of Eq. (6.1)) is present on the nylon Inner Vessel which contains the Borexino scintillator. There is no experimental evidence that ^{210}Pb itself has detached from the IV nylon surface and has entered in the scintillator reaching the FV (see Sec. 6.3). Even if ^{210}Bi from the vessel would enter in the scintillator, it is short-lived and, not being supported by ^{210}Pb, it would decay before reaching the FV. This is not the case for ^{210}Po , which has a relatively long lifetime ($\tau_{\mathrm{Po}} \approx 200$ d). We have evidence that convective motions induced by changes in the temperature profile of the scintillator can detach ^{210}Po from the nylon and bring it into the FV. This spoils the secular equilibrium assumption on which the ^{210}Bi -^{210}Po tag in Eq. (6.2) is based.

Due to the presence of the ^{210}Po source term described in point 2), the secular equilibrium assumption does not hold anymore and the evolution in time of ^{210}Po , ^{210}Bi and ^{210}Pb activity is perturbed by the presence of a source term $S(\vec{x}, t)$:

$$\frac{\mathrm{d}n_i(t)}{\mathrm{d}t} = -\frac{n_i(t)}{\tau_i} + \frac{n_{i-1}(t)}{\tau_{i-1}} + \underbrace{S_i(\vec{x}, t)}_{\text{source term}} \qquad i = {}^{210}\mathrm{Pb}, {}^{210}\mathrm{Bi}, {}^{210}\mathrm{Po} \tag{6.3}$$

where n_i are the number of decays for each isotope. We will see throughout this Chapter that the source term $S(\vec{x}, t)$ has a complex dependence on time and position and that it is the main obstacle to the ^{210}Po supported rate determination.

6.1.1 ^{210}Po evolution in time and thermal insulation of the Borexino detector

The evolution in time of the ^{210}Po counting rate from 2012 until now is shown in Fig. 6.2. To obtain this plot I selected events in a relatively narrow window around the ^{210}Po energy peak ($150 < N_{pe}^{ng} < 300$, where N_{pe}^{ng} is the geometrically normalized charge N_{pe} estimator[1]). Furthermore, I applied the basic quality cuts described in Sec. 4.1 and an additional cut in the α/β discrimination variable, MLP < 0.3, to select α events (see Sec. 2.4.1). An example of the α events spectrum can be seen in Fig. 6.3. The α tagging efficiency of the MLP variable and of its variations in time have been separately estimated by Monte Carlo simulations and are shown in Fig. 6.4; the related corrections have been applied to the selected data.

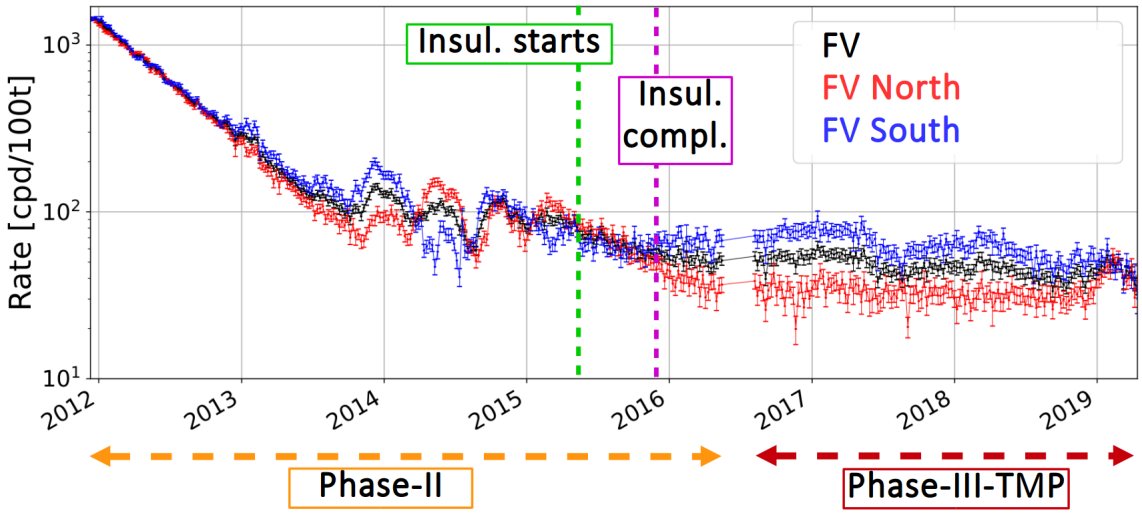


Figure 6.2: ^{210}Po counting analysis: ^{210}Po decay rate evolution from the beginning of Phase-II in the whole FV (black profile) and in its northern and southern hemispheres (respectively, red and blue profiles). The two vertical lines flag the steps towards the detector thermal insulation and stabilization of convective motions: thermal insulation beginning, in May 2015 (green line), insulation completed in December 2015 (purple line).

The black line in Fig. 6.2 shows the evolution in time of the ^{210}Po rate in the FV used for the CNO analysis[2], while the blue and red curves refer to events in its South and North hemisphere respectively.

Until 2015, the rate was dominated by the ^{210}Po unsupported term described in the previous Sec. 6.1, which started out at around ~ 1400 cpd/100t in 2012 and decayed away following the ^{210}Po lifetime. This large rate has hidden second order effects which became evident when the unsupported term decayed away. For example, starting from the end of 2014 it is possible to see strange upward fluctuations of the ^{210}Po rate with an asymmetric North/South distribution. These fluctuations were found to be highly correlated to external temperature changes (due to the season or to other effects). As discussed in the previous Sec. 6.1, the explanation for this behaviour has

[1] The geometrical normalization corrects the energy estimator to 2000 reference working channels (as the standard normalization does) and, moreover, takes into account geometrical acceptance effects of the energy reconstruction, which are position dependent.

[2] Defined by the $r < 2.8$ m and -1.8 m $< z < 2.2$ m conditions, as described in Sec. 4.1.2.

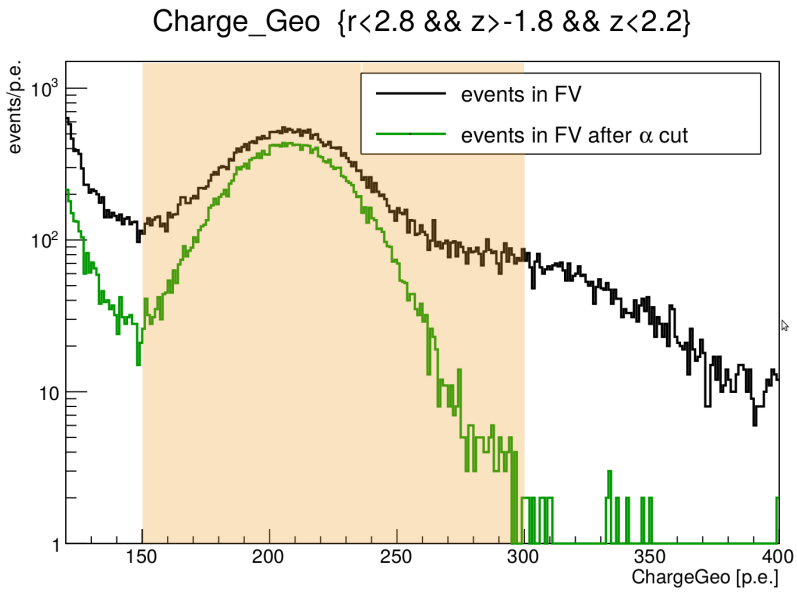


Figure 6.3: Energy spectrum in N_{pe}^{ng} variable, in FV, for Phase-III-TMP dataset before and after the MLP < 0.3 selection cut used to isolate α events (respectively in black and green). The ^{210}Po gaussian-like distribution is clearly visible by eye. The α events selected for the ^{210}Po time evolution are those included in the $150 < N_{pe}^{\mathrm{ng}} < 300$ energy window, marked as the shaded orange region.

to be searched in the presence of ^{210}Po on the Inner Vessel, which is brought inside the FV by convective motions triggered by temperature variations.

In order to damp this effect, in 2015 the Borexino collaboration has decided to thermally insulate the detector with a 20 cm thick layer of mineral wool (see Fig. 6.5). The insulation was completed at the end of 2015. In addition, in 2016 a Temperature Active Control System (TACS) has been installed on the top ring of the Water Tank. It provides heat to compensate for the seasonal cooling in Fall and Winter time and it is useful to further decouple the upper part of the detector from the Hall C temperature variations. A TACS upgrade has been realized in the end of 2018, mounting two supplementary pipe rings at smaller latitudes.

In order to keep the detector temperature under monitoring, a system of temperature probes has been mounted (Latitudinal Temperature Probe System, LTPS). It consists of 54 probes distributed inside the Stainless Steel Sphere, on the Water Tank outer wall surface and on the Water Tank upper dome (see Fig. 6.6). This system provides an accurate profile of the temperature in the Borexino detector as a function of position and time. Fig. 6.7 shows the temperature profile obtained with LTPS as a function of time: the different colors correspond to probes located at different height (blue is the bottom, red is the top). We can see that after the insulation, the stability of the temperature is much improved. Note that there is always a temperature gradient in the vertical axis (the top regions are warmer than the bottom ones): this is actually good because it helps keeping the fluid stationary, as confirmed also by thermal simulations (see Sec. 6.1.2).

The temperature stabilization has been effective in damping the ^{210}Po variations due to convective motion. This can be seen looking again at the evolution in time of the ^{210}Po rate depicted in Fig. 6.2. After the completion of the thermal insulation (marked with a blue vertical line) the fluctuations of

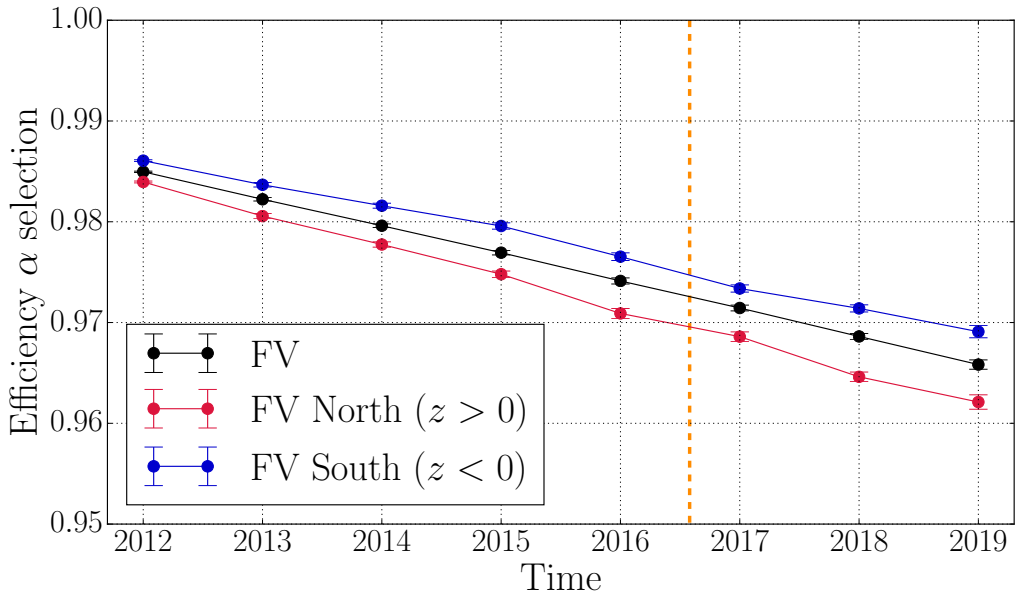


Figure 6.4: Time evolution of the MLP efficiency for α events tagging in the FV, in the range $150 < N_{pe}^{\text{ng}} < 300$. The orange dashed line marks the beginning of Phase-III-TMP dataset. Estimations have been carried out exploiting high statistics Monte Carlo simulations.

the ^{210}Po rate are less evident. Unfortunately, some residual convective effects are still present. This can be also seen from Fig. 6.8, which shows the content of ^{210}Po as a function of time and vertical coordinate. The ^{210}Po rate is calculated considering 35 iso-volumic layers defined in a $r < 3\,\text{m}$ sphere, cylindrically simmetric with respect to the z-axis (named also *z-slices*). The ^{210}Po rate is shown in logarithmic color scale, and ranges from 10 cpd/100t (dark blue) to 300 cpd/100t (dark red). Fig. 6.8 shows that before insulation there were big injections of ^{210}Po moving upwards in Fall when the temperature of Hall C was decreasing and downwards in Spring when the temperature of Hall C was increasing. This effect has been effectively damped by insulation, but some residual seasonal effects are still clearly visible. However, it is possible to see by eye an innermost *clean bubble* just above the equator where the presence of convective motions seems to be null (or at least very small). This region is relatively stable in the last two years. Based on this observation, a more rigorous method has been developed to define this *clean bubble* and determine the supported ^{210}Po term (and consequently the ^{210}Bi rate). This will be described in details in Section 6.3.1.

Figure 6.5: Photography of the Borexino detector, showing the external layer of mineral wool applied to improve the thermal insulation.

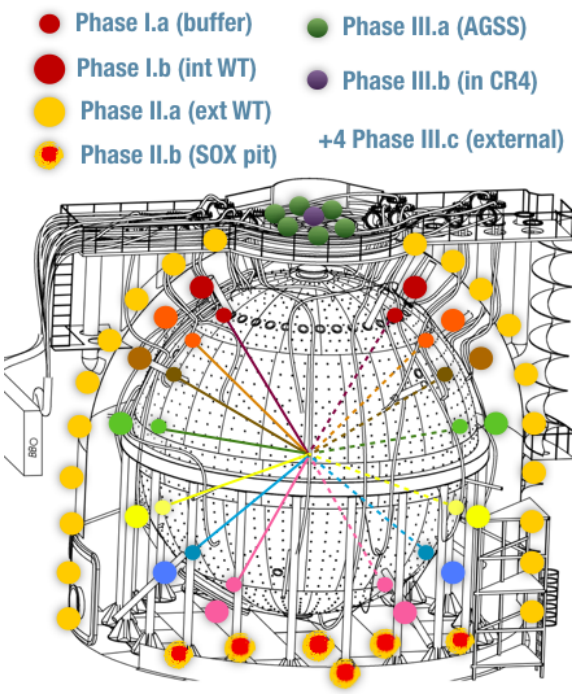


Figure 6.6: The **Latitudinal Temperature Probe System** in Borexino (LTPS): position of the 54 temperature probes [166].

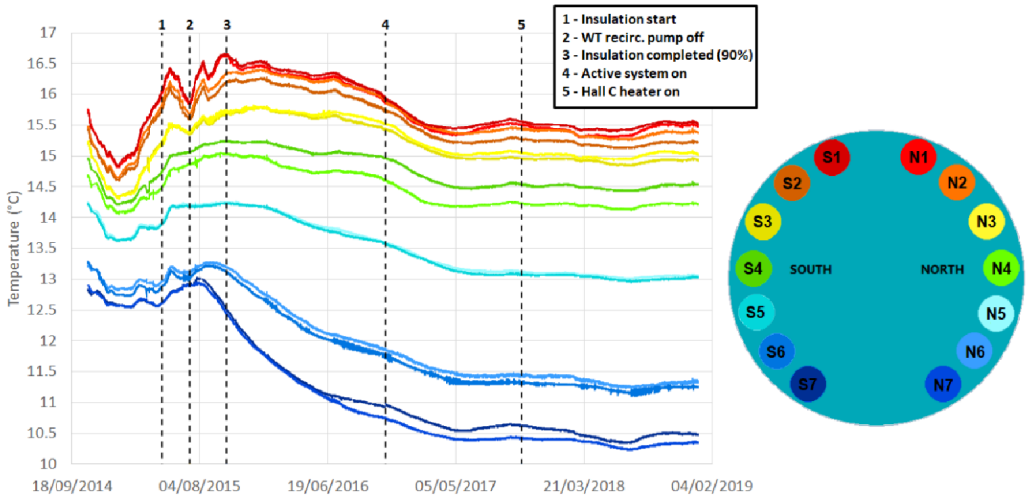


Figure 6.7: Temperature evolution in the detector buffer during and after the insulation. Different colors are associated to the vertical positions of the probes (blue are on the botton, red on the top, see also the sketch on the right). Note that S and N indicate the geographic South and North, respectively.

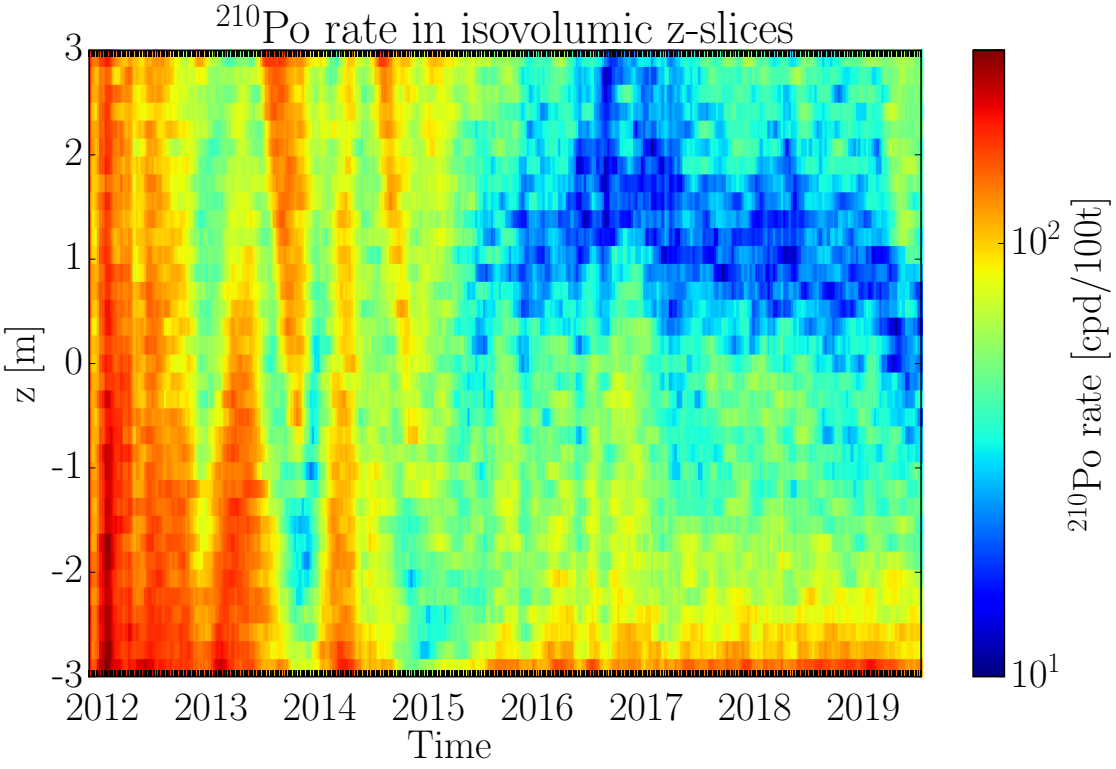


Figure 6.8: Map of ^{210}Po rate evolution in time from 2012 on, calculated in 35 iso-volumic layers, cylindrically, symmetric with respect to the z axis and defined in a 3 m sphere. ^{210}Po rate is shown in logarithmic color scale, which ranges from 10 cpd/100t (dark blue) to 300 cpd/100t (dark red), as a function of time (horizontal axis) and of the z coordinate of the layers (vertical coordinate). Oscillatory trends, visible up to half of 2015, are generated by the seasonal temperature variations of the scintillator.

6.1.2 Fluid dynamics simulations

Numerical simulations of the Borexino detector thermal and fluid-dynamics may help to achieve the proper understanding of the ^{210}Po evolution in time and pattern in space, and also suggest effective counter-measures to minimize convection inside the detector. A system of two-dimensional conductive models has been implemented into CFD simulations (Conductive Computational Fluid Dynamics) to predict the heat exchange processes of all the detector components [166].

The temperature instability impact in terms of convective motions has also been studied through CFD simulations. A 2D model has been employed, including as input profile temperatures the output of the LTPS system. The simulations confirm qualitatively the behaviour observed on data, i.e. that the stabilization of the vertical temperature gradient reduces the fluid vertical movements. However, even with a stable vertical gradient, some very slow horizontal flows with typical speed scale $\sim 10^{-7}$ m/s are still present due to small inhomogeneities in the temperature boundary conditions. Those movements could be responsible for the residual convection that we are seeing on data.

A simplified transport model was developed, aiming to simulate the ^{210}Po migration inside the inner vessel for insulated periods. Both ^{210}Po and ^{210}Bi are assumed to be emitted from the IV surface; a correlation between the minimum ^{210}Po rate and the minimum of horizontal gradient is found. Recently, the numerical analysis has been extended by improving the two-dimensional model accuracy as well as the validation framework. Fig. 6.9 shows estimations for ^{210}Po rate in the scintillator for the period October 2018 from numerical fluid dynamic simulations, displayed as color scale (from 35 cpd/100t for dark blue regions to 90 cpd/100t for dark red regions) in the x-z coordinates plane. Calculations have been carried out assuming either a spherical IV shape (left panel), or a realistically deformed IV shape (right panel). Results for the two cases share many features: a ^{210}Po z-gradient is clearly visible, as a local minimum region 40 cpd/100t slightly above the equator. The φ symmetry of the ^{210}Po rate distribution seems to be preserved.

Fig. 6.10 shows exemplary ^{210}Po rate trends along the z-coordinate (measured taking the IV bottom point as $z = 0$ m) during April 2018 - October 2018 period [166]: numerical simulations (dotted lines) and real data (solid lines). The main features of ^{210}Po evolution in time are qualitatively reproduced; however, discrepancies show up especially in the southern hemisphere, where convections are expected to be more critical.

In conclusion, more work is needed before the simulations achieve the required accuracy to be predictive. However, this is not crucial for the determination of the ^{210}Po supported term: indeed, our method described in Sec. 6.2 does not require a precise knowledge of the details of the transport model: it looks for a minimum in the ^{210}Po spatial distributions with very mild underlying assumptions.

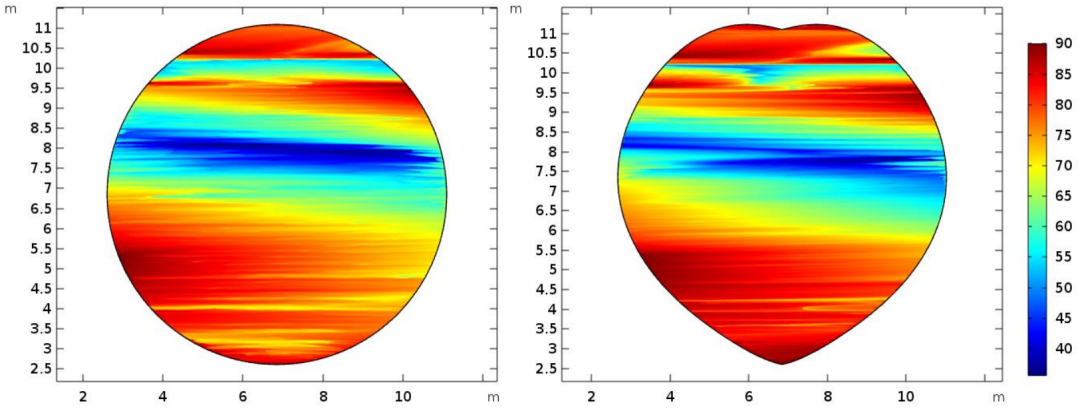


Figure 6.9: Fluid dynamic simulations: estimations for ^{210}Po rate in the scintillator for the period October 2018 from numerical fluid dynamic simulations, displayed as color scale (from 35 cpd/100t for dark blue regions to 90 cpd/100t for dark red regions) in the x-z coordinates plane. Calculations have been carried out assuming either a spherical IV shape (left panel), or a realistically deformed IV shape (right panel).

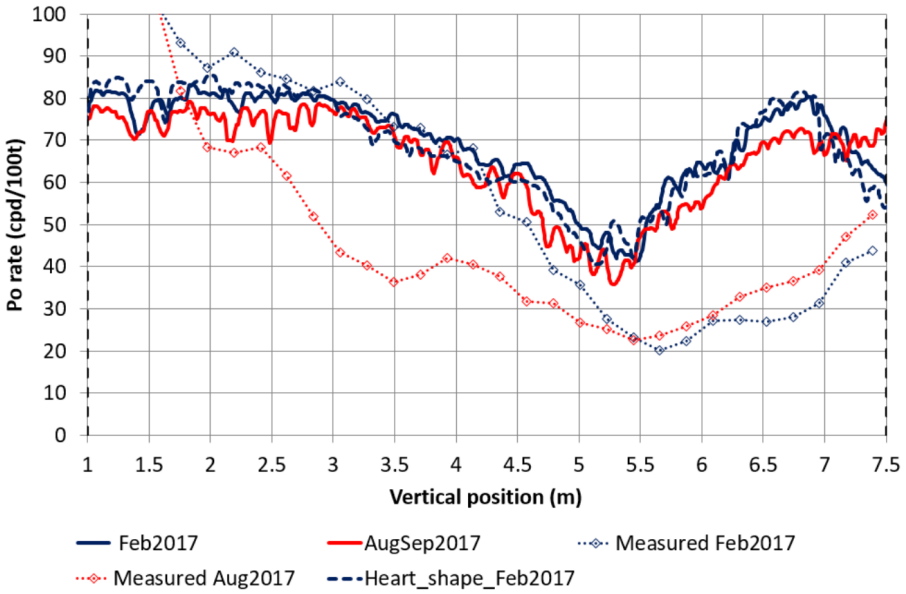


Figure 6.10: Fluid dynamic simulations: exemplary ^{210}Po rate trends along the z-coordinate during April 2018 - October 2018 period [166]. $z = 0$ m corresponds to the bottom of the stainless steel sphere. Blue and red correspond to two different periods in time (Feb 2017 and Aug-Sep 2017). Continuous lines are simulation, while dotted lines are data.

6.2 Determination of the ^{210}Bi interaction rate from the supported ^{210}Po term

In Sec. 6.1 we have seen that the thermal insulation of Borexino was successful in reducing the convective motions which caused the migration of ^{210}Po from the Inner Vessel into the FV. Unfortunately, the motions have not completely stopped and the time/space pattern of the ^{210}Po rate is still complicated to interpret (see Fig. 6.8). Further attempts to completely stop convection by increasing even more the temperature control are being pursued, like for example, installing a system to stabilize the experimental Hall C temperature. In the next months we will see whether these attempts will be successful. However, we can observe that even now we have a clean bubble, corresponding to approximately 20 t of scintillator, where convective currents seems to be reduced (blue region around $z \sim (0.5 - 1)$ m in Fig. 6.8, which appeared starting from 2016).

The success of this procedure is critically dependent on the capability to:

1. Define the clean bubble spatially and determine the ^{210}Po rate minimum in it fitting the ^{210}Po distribution, in an unbiased way;
2. Prove that the minimum of the ^{210}Po rate in the clean bubble can be extrapolated to the entire FV; this requires to prove that ^{210}Bi is uniform throughout the whole FV and has been so for the entire set of data used in the analysis (Phase-III-TMP);

For what concerns point 1), in Sec. 6.2.1 I will describe my results obtained using a simplified model which assumes that the ^{210}Po spatial distribution depends only from the distance from the detector center (radial-only dependence). This is only an approximation since the ^{210}Po contribution due to residual convective motions depends instead on the distance from the IV, which is not spherical, as described in Sec. 2.1. For this reason, the results of this *radial analysis*, can be taken only as qualitative.

A more quantitative analysis, the so-called *clean bubble* analysis, is discussed in Section 6.2.2. The result obtained from this second method will be used to set an upper limit to the ^{210}Bi rate in the multivariate fits shown in Chapter 7. To verify possible systematic errors associated to this procedure, I have performed its validation by means of a toy Monte Carlo technique, showing that the minimum of ^{210}Po rate obtained can be only positively biased. This analysis, described in Sec. 6.2.3, confirms that the method is suitable for the extraction of a ^{210}Po upper limit.

For what concerns point 2) I performed several studies which confirm that, within errors, ^{210}Bi is spatially uniform in the FV, as will be described in Section 6.3. The uncertainty on this statement is quantified and included in the systematic error associated to the ^{210}Bi rate.

6.2.1 ^{210}Po radial analysis method

A possible approach to determine the supported ^{210}Po rate R^s_{Po} assumes that the spatial distribution of ^{210}Po events present in the scintillator can be modeled along the radial direction. In the simplest case, I have adopted an analytical model describing the ^{210}Po radial distribution as the sum of three independent contributions:

- **Vessel term**, related to the ^{210}Po attached to the IV surface;

- **Convection term**, describing in an effective way the ^{210}Po which, once detached from the IV, moves towards the innermost scintillation region due to fluid convection (see Sec. 6.1). It is expected to decrease exponentially going towards the central part of the scintillator.
- **Supported term**, directly related to the ^{210}Bi background rate: it is uniformly distributed in the scintillator volume.

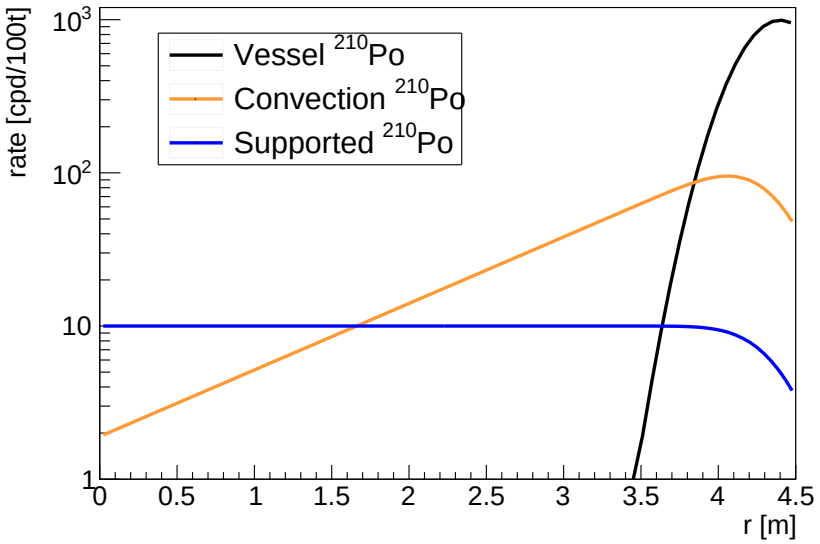


Figure 6.11: ^{210}Po radial analysis method: analytical model. The analytical shapes of the three contributions to the ^{210}Po rate are shown in different colors: vessel term (black), convection term (orange), supported (blue).

The three analytical shapes of the contributions listed above are shown in Fig. 6.11.

This method has the advantage to describe globally the ^{210}Po distribution including all terms (not only the supported one) and is therefore less prone to biases induced by the selection of an *ad-hoc* region free from convection. On the other hand, it relies on approximations which do not take into account the complexity of the ^{210}Po rate distribution. For this reason, this method may be affected by large systematics, difficult to evaluate, and we are using it to have only a qualitative estimate of the supported ^{210}Po term.

Event selection

The analysis is performed on the Phase-III-TMP dataset (17 July 2016 - 02 June 2019). Data are selected through the standard quality cuts. In addition, the α candidate events are selected with the pulse-shape discrimination variable selection MLP < 0.3. Corrections for the MLP efficiency, shown in Fig. 6.12, are applied as a function of the events radius.

As shown in Sec. 5.2.1, for a fixed energy release in the scintillator, the Borexino reconstructed energy decreases when events position approaches the vessel, due to the geometric acceptance decrease and dead PMT asymmetry (see Fig. 5.2). Since we are interested in almost nearly the

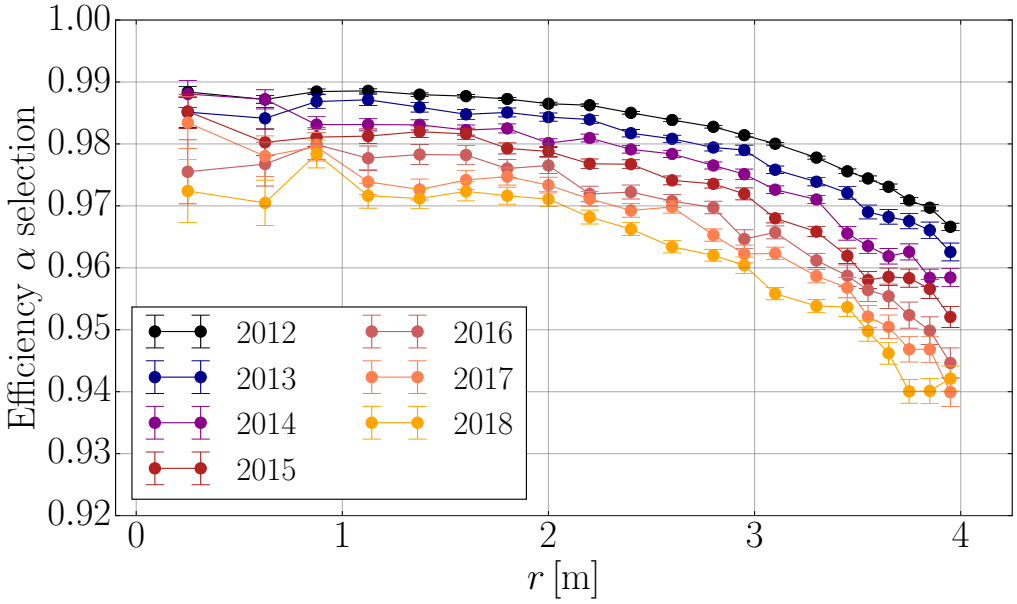


Figure 6.12: MLP variable efficiency for α events as a function of reconstructed event radius. This evaluation has been carried out analyzing ^{210}Po high-statistics Monte Carlo samples.

whole IV volume, the chosen energy window for this analysis is a wide $140 \leq N_{pe}^{\text{ng}} \leq 270$ range, which safely include all the ^{210}Po events. The reconstructed energy vs radius ranges are shown in Fig. 6.13 in red dashed lines in the $N_{pe}^{\text{ng}} - r$ plane. The ^{210}Po gaussian peak value for several radii has been estimated independently and is shown in black points; the associated $\pm 1\sigma$, $\pm 2\sigma$ and $\pm 3\sigma$ reconstructed energy intervals are displayed in different shades of blue.

The radial distribution of α candidate events, after the efficiency correction, before and after the MLP corrections, is shown in Fig. 6.14.

Fit method and results

As summarized previously, the employed model for the ^{210}Po events radial distribution consists of three terms: vessel, convection and supported. The analytical description of the rate of α tagged events as a function of reconstructed radius r, $R_{\text{Po}}(r)$, is given by:

$$R_{\text{Po}}(r) = [R^{\text{v}}_{\text{Po}}(r) + R^{\text{conv}}_{\text{Po}}(r) + R^{s}_{\text{Po}}] * \text{Gauss}(0, \sigma_L) = \tag{6.4}$$

$$= \left[\underbrace{R^{\text{v}}_{\text{Po}}\delta(r_{\text{IV}})}_{\text{Vessel}} + \underbrace{R^{\text{conv}}_{\text{Po}} \text{e}^{(r-r_{\text{IV}})/\lambda_c}}_{\text{Convection}} + \underbrace{R^{s}_{\text{Po}}}_{\text{Supported}} \right] * \text{Gauss}(0, \sigma_L) = \tag{6.5}$$

$$= R_{\text{Po}}(r_{\text{IV}}; R^{\text{v}}_{\text{Po}}, R^{\text{diff}}_{\text{Po}}, R^{s}_{\text{Po}}, \lambda_c, \sigma_L) \tag{6.6}$$

where r_{IV} is the average IV radius, λ_c is the average effective convection length, R^{v}_{Po} is the vessel ^{210}Po term, $R^{\text{conv}}_{\text{Po}} = R^{\text{conv}}_{\text{Po}}(r = r_{\text{IV}})$, R^{s}_{Po} is the supported ^{210}Po term. While $R^{\text{v}}_{\text{Po}}(r)$ is localized around the IV radius, assuming a gaussian shape due to the finite position resolution, $R^{\text{conv}}_{\text{Po}}(r)$

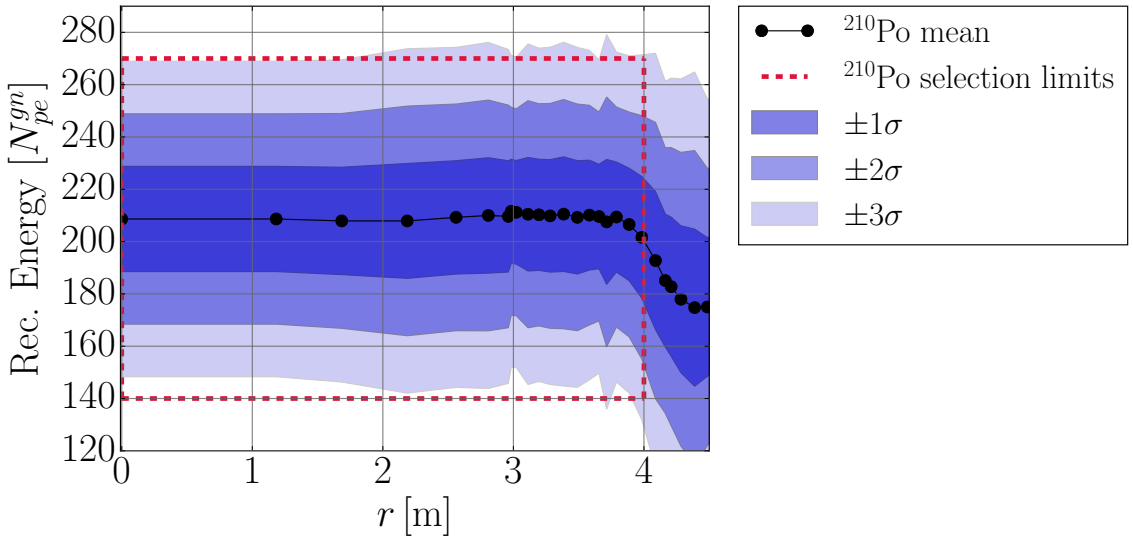


Figure 6.13: ^{210}Po radial analysis method: reconstructed energy vs radius. The ^{210}Po gaussian peak value for several radii has been estimated independently and is shown in black points; the associated $\pm1\sigma$, $\pm2\sigma$ and $\pm3\sigma$ reconstructed energy intervals are displayed in different shades of blue. The dashed red rectangle shows the region selected by the cuts ($140 < N_{pe}^{\mathrm{ng}} < 270$) and ($r < 4\,\mathrm{m}$).

rapidly decrease towards the detector center and $R^{\mathrm{s}}_{\mathrm{Po}}(r)$ is constant throughout the scintillator. The position reconstruction smearing is modeled as a Gaussian with σ_L resolution parameter.

A delicate assumption of this model is connected to the IV shape, which is known to be asymmetric with respect to the azimuthal θ angle. Therefore, the fraction of ^{210}Po detaching from IV and diffusing into the scintillator may be slightly θ-asymmetric as well, at least in the portions of scintillator extremely close to the IV. The IV shape symmetry with respect to the φ polar angle can be safely assumed. Performing the radial analysis it is implicitely assumed that the angular θ and φ ^{210}Po rate dependences, whose exact modeling would be unworkable, can be absorbed in a simpler and *effective* radial distribution ($R_{\mathrm{Po}}(r, \theta\, \varphi) = R_{\mathrm{Po}}(r)$). It follows that the four parameters λ, r_{IV}, $R^{\mathrm{conv}}_{\mathrm{Po}}$ and $R^{\mathrm{v}}_{\mathrm{Po}}$ are implicitely *averaged* over the angular coordinates. To limit the impact of a non-perfectly-spherical vessel, the FV for this analysis is selected such that the average distance from the vessel is at least $\approx (50 - 60)\,\mathrm{cm}$.

The fit is performed separately in the North and in the South hemisphere, considering large spherical FVs: $r < 3.8\,\mathrm{m}$ and $z > 0\,\mathrm{m}$ for North hemisphere, while $r < 3.4\,\mathrm{m}$ and $z < 0\,\mathrm{m}$ for South one. These maximum radial values for the analysis FV have been determined by means of a χ^2/NDF optimization of the fit scanning several FVs, maximizing the data-model agreement; in any case, results are proven to be stable against the radial fit range.

The vessel term plays a crucial role for the success of the analysis being the term most sensitive to σ_L. Unfortunately, since we don't fit over the whole volume, only the lower tail of its gaussian distribution contributes and the fit struggles to determine r_{IV}. For this reason, r_{IV} is determined independently, from the spatial distribution of the β contamination on the vessel, and is fixed in the fit.

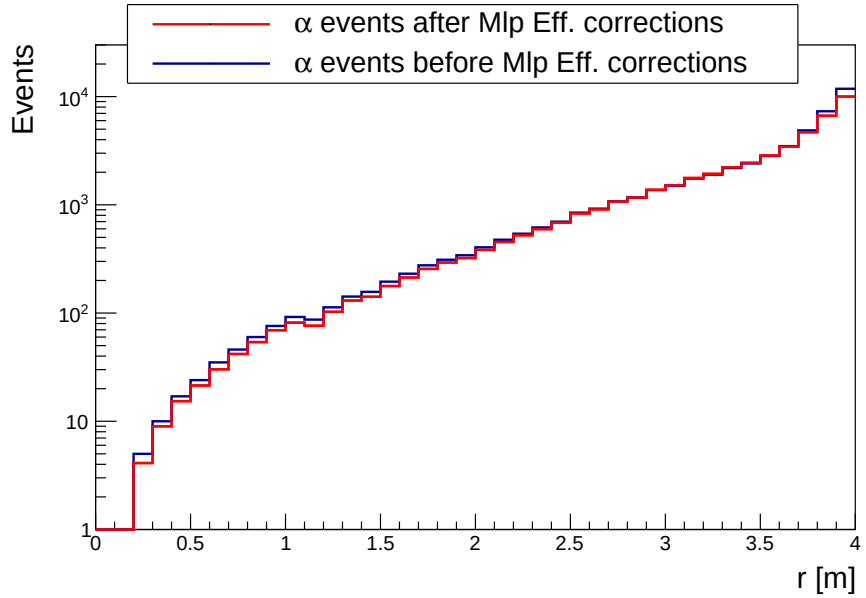


Figure 6.14: ^{210}Po radial analysis method: radial distribution of α candidate events after basic cuts (quality cuts, energy, MLP, hemisphere), before and after the MLP efficiency corrections, respectively in blue and red profiles.

The fit to the Phase-III-TMP dataset with the radial model described by Eq. (6.6) is shown in Fig. 6.15, for the FV North hemisphere. Results for both North and South hemispheres are reported in Tab. 6.1; they are found consistent and with a good data-agreement, quantified by $\chi^2/\mathrm{NDF}_{\mathrm{North}} = 0.98$ and $\chi^2/\mathrm{NDF}_{\mathrm{South}} = 0.80$ respectively. Note that the convection lengths found for the two hemispheres ($\lambda_c^{\mathrm{North}} = 1.44 \pm 0.12$ and $\lambda_c^{\mathrm{South}} = 1.53 \pm 0.25$) exceed of seven times the expected diffusion-only lengths for a liquid scintillator ($\approx 20\,\mathrm{cm}$); this restates the importance of convective motions within the scintillator. The combined North and South hemisphere analysis return an overall ^{210}Po supported rate $R^s_{\mathrm{Po}} = 10.8 \pm 2.3\,\mathrm{cpd/100t}$, with a $\approx 21\%$ relative uncertainty.

I have verified that the result of this method is stable against details of the data selection (energy range, MLP selection and efficiency) and of the fit (radial fit range). I have varied the most relevant parameters of interest one by one within reasonable ranges; the overall variation of the result concerning the ^{210}Po supported rate change by less than $0.5\,\mathrm{cpd/100t}$, as shown in Tab. 6.2. These variations could be considered as an estimation of the systematic uncertainties associated to the radial analysis itself. However, as already mentioned at the beginning, the results of this method should be taken only as qualitative, since they are based on strong approximations concerning the convective ^{210}Po spatial distribution.

6.2.2 ^{210}Po clean bubble method

This method is based on the search for the *cleanest bubble* in the scintillator, i.e., a region where the convective motions are negligible or marginal, and the ^{210}Po count rate is mostly due to the supported term. A rough localization of the bubble can be performed even by eye, looking at the

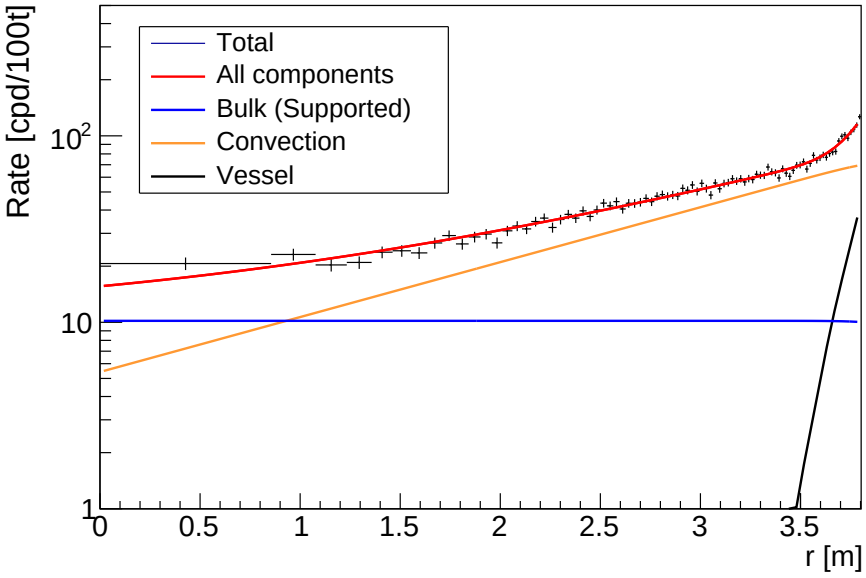


Figure 6.15: ^{210}Po radial analysis method: radial fit for the Phase-III-TMP dataset according to Eq. (6.6) in the North hemisphere. The result of the fit is shown in red. The three contributions considered in the fit are shown in black (vessel term), orange (convective term) and blue (supported term).

map of the ^{210}Po rate as a function of time (Fig. 6.8): a clean blue spot around $z \sim 1\,\mathrm{m}$ with a low ^{210}Po counting rate ($\sim (10-15)\,\mathrm{cpd/100t}$) is evident starting from 2016. A more detailed and rigorous analysis based on a spatial fit of the ^{210}Po distribution will be described in the following.

Only data from 2017 are considered so that the contribution from unsupported ^{210}Po term is negligible: the dataset ranges from Jan 2017 to Dec 2018, for a total livetime of 625.5 days. In this assumption, the minimum ^{210}Po rate inside the *clean bubble* R^b_{Po} can be described as the sum of supported term R^s_{Po} and the convection term R^c_{Po}:

$$R^b_{\mathrm{Po}} = R^s_{\mathrm{Po}} + R^c_{\mathrm{Po}} \tag{6.7}$$

Consequently, thanks to the ^{210}Bi -^{210}Po secular equilibrium relation, an upper limit for ^{210}Bi rate can be set as

$$R_{\mathrm{Bi}} = R^s_{\mathrm{Po}} \leq R^b_{\mathrm{Po}} = R^{s\text{-rec.}}_{\mathrm{Po}} \tag{6.8}$$

where the equal sign holds in case the clean bubble region is free from any convective motion, i.e. in case of pure ^{210}Bi -^{210}Po secular equilibrium. In the following method, R^b_{Po} is extracted, but the two ^{210}Po contributions R^s_{Po} and R^c_{Po} cannot be disentangled. Therefore, according to Eq. (6.8), only an upper limit for the ^{210}Bi rate will be set.

Unlike the radial analysis described in Sec. 6.2.1, the clean bubble method does not make any assumption on the spatial distribution of the supported and convective ^{210}Po terms. Instead, it

	^{210}Po radial analysis, Phase-III-TMP	
	North	South
R^s_{Po} [cpd/100t]	10.6 ± 2.4	12.7 ± 7.6
λ_c [m]	1.44 ± 0.12	1.53 ± 0.25
σ_L [m]	0.22 ± 0.01	0.27 ± 0.01
χ^2/NDF	0.98	0.80

Table 6.1: ^{210}Po radial analysis method: results for the Phase-III-TMP dataset, in the North and South hemisphere. The fit is performed in the $r_{\mathrm{North}} < 3.8\,\mathrm{m}$ and $r_{\mathrm{South}} < 3.5\,\mathrm{m}$ regions for North and South hemispheres respectively.

Parameter	Uncertainty [cpd/100t]	%
Radial fit range	±0.09	±0.7%
Energy range cut	±0.345	±1.7%
MLP threshold cut	±0.23	±1.9%
MLP efficiency	±0.046	±0.4%
Total	±0.43	±3.3%

Table 6.2: ^{210}Po radial analysis: stability against fit parameters involved in the data selection or in the fit method.

searches for a region which can be fitted by a spheroidal shape:

$$R_{\mathrm{Po}}(x,y,z) = R^b_{\mathrm{Po}}\left[1 + \frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{c^2} + \frac{(z-z_0)^2}{b^2}\right] \tag{6.9}$$

This function is simply the generalization of the 1D parabola and describes the local Taylor expansion around a point of minimum (x_0, y_0, z_0). The analysis procedure can be summarized in three steps:

1. **3D fit.** The fitting can be simplified by means of a preliminary step. A 3D ^{210}Po rate fit is performed on the full data set (2017-2018) in the range $|x| < 3\,\mathrm{m}$, $|y| < 3\,\mathrm{m}$, $-1\,\mathrm{m} < z < 3\,\mathrm{m}$, assuming Eq. (6.9) as analytical function. From this first fit we get a rough estimate of x_0, y_0, z_0, a, b and c. We obtain that x_0 and y_0 are compatible with 0 while $z_0 \sim 0.8\,\mathrm{m}$. We also find that a is approximately equal to c: therefore it is possible to assume a φ cylindrical symmetry and approximate Eq. (6.9) to

$$R_{\mathrm{Po}}(\rho,z) = R^b_{\mathrm{Po}}\left[1 + \frac{\rho^2}{a^2} + \frac{(z-z_0)^2}{b^2}\right] \tag{6.10}$$

   where $\rho = \sqrt{x^2+y^2}$ is the cylindrical radius.

2. **Not-aligned fit.** From the α counting analysis, it is known that the ^{210}Po clean bubble center z_0 is not perfectly stable in time because of temperature changes, which induce ^{210}Po

motions. For this reason, the fit is then performed on a monthly basis on a 70 t fiducial mass, using the expression Eq. (6.10) as a function of ρ^2 and z, to extract the $z_0(t)$ trend in time (*not aligned fit*).

3. **Aligned fit.** As a final step, the 2D fit is performed after having *aligned* data at the center of the *bubble* to the previously obtained $z_0(t)$ trend, in a 20 t fiducial mass. This is equivalent to perform the fit with respect to ρ^2 and $z - z_0(t)$ variables. The results of the fit are shown in Fig. 6.16; the extracted ^{210}Po minimum rate is $R^b_{\mathrm{Po}} = 11.8 \pm 0.8$ cpd/100t, including the statistical uncertainty only. The projections of the 2D histogram to the ρ^2 and $z - z_0$ axes are shown in Fig. 6.17.

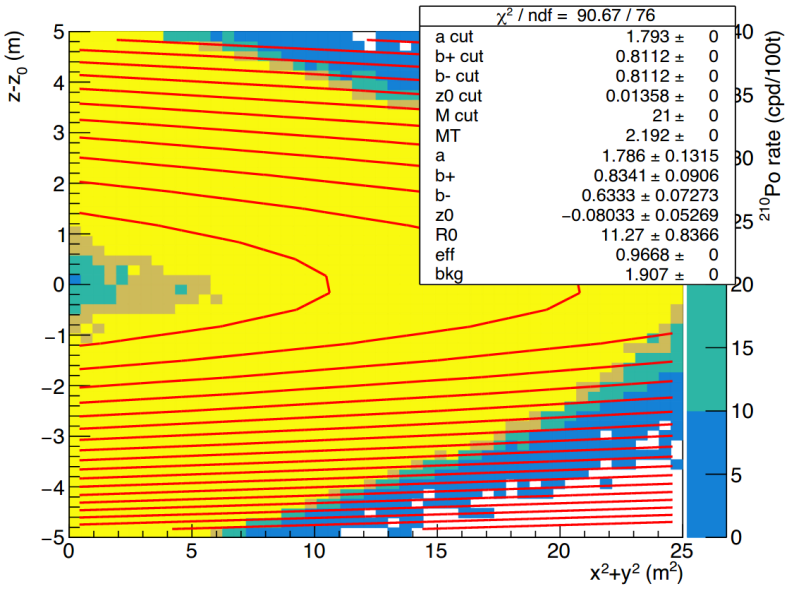


Figure 6.16: ^{210}Po **clean bubble method:** 2D fit of ^{210}Po rate as a function of ρ^2 and $z - z_0(t)$, where z_0 is the center of the fit bubble, in 2017 Jan - 2018 Dec period (625.5 days), according to Eq. (6.10).

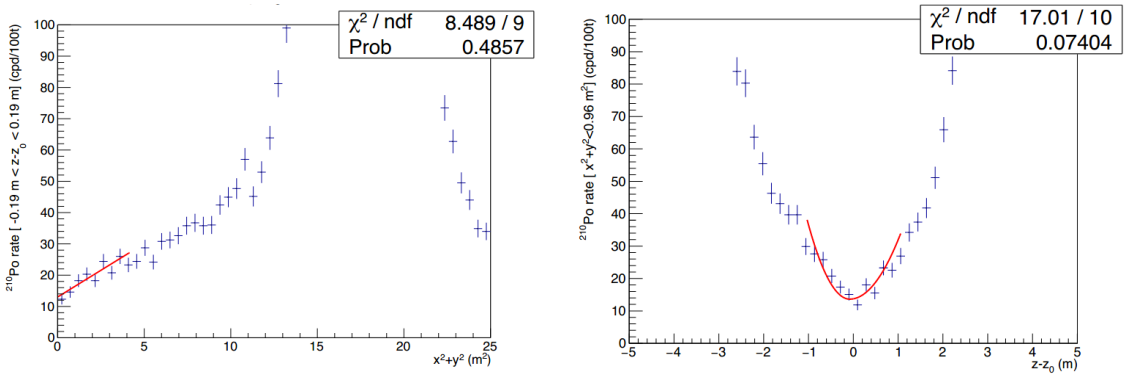


Figure 6.17: ^{210}Po **clean bubble method:** projections of Fig. 6.16 2D fit of ^{210}Po rate on the ρ^2 and $z - z_0(t)$ axes, in 2017 Jan - 2018 Dec period (625.5 days), according to Eq. (6.10).

It is worth to underline that the effect of the thermal stability operations could bring significant advantages in the extraction of a ^{210}Bi constraint. Indeed, in the best case, the *clean bubble* size could be enlarged in the next future thanks to the detector thermalization, generating a ^{210}Po rate *plateau*: this would be a clear sign that the ^{210}Po rate has reached a physical minimum. If this plateau will be sufficiently large, we could directly extract $R^s_{\rm Po}$ and therefore $R_{\rm Bi}$ with a significant precision, instead of an upper limit as discussed previously. However, until now we don't have a clean evidence of a ^{210}Po rate plateau (as seen in right panel of Fig. 6.17, which shows a ^{210}Po minimum) and therefore we will use only the bubble analysis result as an upper limit on ^{210}Bi rate.

Systematic studies

In the following, we discuss the systematic errors associated to the previously described *clean bubble* method.

1. **Clean bubble size**. I have evaluated the error associated to the choice of the clean bubble size, by varying it from a minimum of 10 t to a maximum of 35 t: the contribution to the error is estimated to be ± 0.42 cpd/100t.
2. **Choice of the analytical function for the 2D fit**. The impact of the choice of the fitting function on the results can be analyzed choosing alternatively a smoother *pan-like shape* or a sharper *V-shape* instead of the standard paraboloid one (Eq. (6.10)). The projections of the 2D fits over the $z - z_0(t)$-direction, selecting these two fitting functions, are shown respectively in the left and in the right panel of Fig. 6.18. The $R^b_{\rm Po}$ differences with respect to the standard case are quoted as the systematics errors, and have been evaluated as $\pm^{+0.00}_{-2.38}$ cpd/100t. The negative error is not included in the overall systematic uncertainty for the $R(^{210}\text{Bi})$ upper limit.
3. **Homogeneity of ^{210}Bi rate.** Beyond the systematic errors discussed so far, we still have to evaluate the uncertainty associated to the extrapolation of the *clean bubble* result to the entire FV used for the CNO-ν analysis. This is the main subject discussed in the next Sec. 6.3.

As shown also in detail in Sec. 6.3.2, we anticipate the final estimate of the ^{210}Bi content $R(^{210}\text{Bi}) = 11.8^{+1.6}_{-2.8}$ cpd/100t, which will be converted in a ^{210}Bi rate upper limit. The quoted error has been obtained summing in quadrature the previously mentioned systematic and statistical errors.

6.2.3 Toy MC for the ^{210}Po clean bubble method validation

Besides the systematic studies shown until now, we have to check whether the the 2D paraboloid (Eq. (6.9)) fitting procedure introduces biases to the extracted $R^b_{\rm Po}$ value, and therefore on the $R(^{210}\text{Bi})$ result. I have evaluated this possibility by means of dedicated Monte Carlo studies: I have simulated thousands of pseudo-datasets in which the ^{210}Po spatial distribution is given by the sum of two terms, the supported ^{210}Po plus convection ^{210}Po , and then I have fitted them with the same procedure used for the real data (as described in the previous Sec. 6.2.2).

I followed this procedure:

1. **PDFs generation.** The first step is the generation of two $\rho^2 - z$ 2-dimensional PDFs, one for

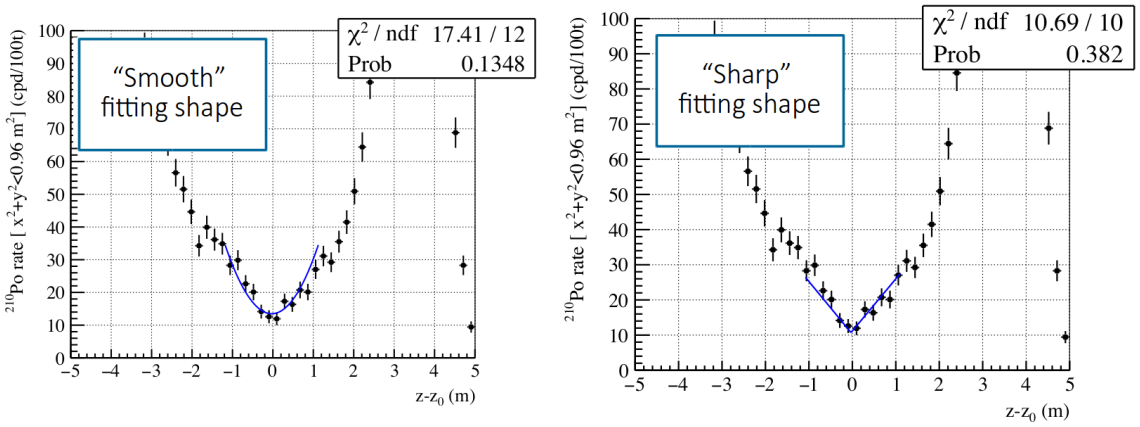


Figure 6.18: ^{210}Po clean bubble method: systematic associated to the choice of the analytical model shape. Projections of 2D fit of ^{210}Po rate on the $z - z_0(t)$ axes, in 2017 Jan - 2018 Dec period (625.5 days), according to analytical shapes slightly different from the standard paraboloid one (Eq. (6.10)).

each ^{210}Po contribution: supported and convection[3]. Events are firstly simulated uniformly in the FV and, on an event-by-event basis, a weight is applied to the event to properly describe the ^{210}Po distribution of the chosen term.

The supported term is generated uniformly throughout the whole IV, therefore assigning a trivial weight which does not depend on the event position $w_{\text{supported}}(x, y, z) = 1$. The convection term is the most critical one, since currently there is no exact analytical dependence to describe it. In this work, the weight is assumed exponentially decreasing as a function of the event distance from the Inner Vessel[4], i.e. $w_{\text{convection}}(x, y, z) = \mathrm{e}^{-d_{\text{IV}}(x,y,z)/\lambda}$. λ is the convection length; I have generated convection ^{210}Po PDFs with different λ assumptions, taking into account that from the radial analysis on data (described in Sec. 6.2.1) it is expected to be included within the 0.8 m − 1.6 m range. The finite position reconstruction resolution is taken into account on an event-by-event basis by means of a gaussian smearing, while the real deformed shape of the IV is properly considered.

Exemplary 2D $\rho^2 - z$ PDFs are shown in Fig. 6.19, for the ^{210}Po supported term (left panel) and the ^{210}Po convection term (right panel). A realistic IV shape, obtained averaging the ones from the 2 years of data-taking considered by the clean bubble analysis, has been assumed; the employed convection length in this example is $\lambda = 1.0$ m.

2. **PDFs sampling.** The second step consists in sampling the previously generated PDFs according to the real expected statistics in the same period of the clean bubble analysis. In this way, a *pseudo-dataset* is generated.

   I have generated pseudo datasets in different conditions: assuming different convection lengths λ and different values of R^s_{Po} and $R^{\text{conv}}_{\text{Po}}$. An example of the pseudo-datasets can be seen in Fig. 6.20, where I have assumed $\lambda = 1.0$ m, $R^{s\text{ - inj.}}_{\text{Po}} = 10$ cpd/100t and, at the ^{210}Po convection

[3] In particular, the ^{210}Po vessel term is not necessary, since the 2D fitting considers only a 20 t region within the FV, where the vessel term is absent.

[4] Several possibilities for the $w_{\text{convection}}(x, y, z)$ modeling have been analyzed; the conclusions of these studies are unchanged with respect to what is shown in the following.

minimum, $R_{\mathrm{Po}}^{s\text{-inj.}}/R_{\mathrm{Po}}^{\mathrm{conv.-inj.(min)}} \approx 2$. In particular, one can note that the generated ^{210}Po distribution reproduces the typical *clean region* found on data, and centered approximately at $z \approx 0.5\,\mathrm{m}$ and $\rho \approx 0\,\mathrm{m}$ (dark blue region). However it is worth to remind again that the goal of this study is not to describe a complete and fully detailed picture of the ^{210}Po behaviour within the FV, but to provide a reasonable ^{210}Po distribution in order to validate the 2D fitting procedure.

3. **Dataset fitting** As a final step, the sampled $\rho^2 - z$ histogram is then fitted with Eq. (6.10) for each of the pseudo-datasets in a 20 t mass region, as for the analysis performed on data.

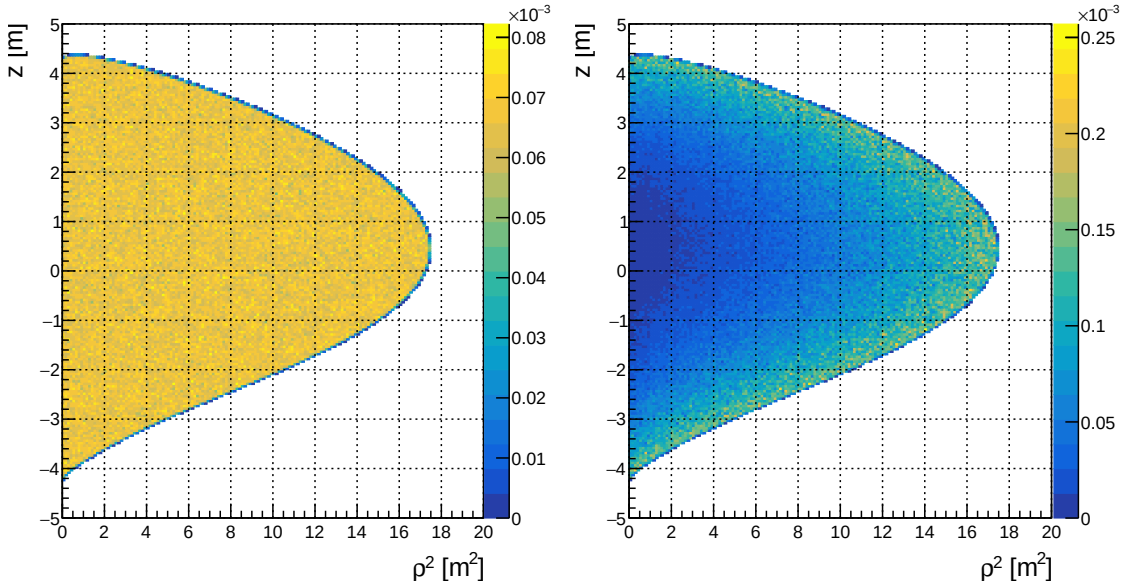


Figure 6.19: Exemplary **toy MC** ^{210}Po **2D PDFs** $\rho^2 - z$ for the ^{210}Po supported term (left panel) and for the ^{210}Po convection term (right panel). The employed convection length is $\lambda = 1.0\,\mathrm{m}$.

In the following, we report an example in the $\lambda = 1.0\,\mathrm{m}$ scenario: results of 10^3 fits of simulated pseudo-experiments with a 2 y statistics are shown in Fig. 6.21. The histograms for the following quantities are shown: injected and reconstructed rates of supported ^{210}Po ($R_{\mathrm{Po}}^{s\text{-inj.}}$ and $R_{\mathrm{Po}}^{s\text{-rec.}}$ respectively); difference of these two ^{210}Po rates ($\Delta R_{\mathrm{Po}}^{s}$); z_0 of clean bubble; χ^2/NDF of the fit.

The most important conclusion is that, in this parameters configuration, the method induces only an overestimation of the ^{210}Po supported term: in 99.8% of cases, the reconstructed ^{210}Po minimum is higher than the related injected value. The average bias, displayed by the mean of the gaussian-like $\Delta R_{\mathrm{Po}}^{s} = R_{\mathrm{Po}}^{s\text{-rec.}} - R_{\mathrm{Po}}^{s\text{-inj.}}$ distribution (upper right plot), is $+1.9\,\mathrm{cpd/100t}$. The data-agreement is good, as quantified by the χ^2/NDF distribution (lower left panel). Note also that the bubble z_0 distribution (lower right panel) peaks around $\sim 0.5\,\mathrm{m}$, which is a qualitatively reasonable result if compared to the convection ^{210}Po distribution shown in Fig. 6.19.

I have performed similar studies for several combinations of the input parameters λ, $R_{\mathrm{Po}}^{s\text{-inj.}}$ and $R_{\mathrm{Po}}^{\mathrm{conv-inj.}}$ in a wide range, and I have verified that the conclusions reported for the previous example still hold; in particular, $R_{\mathrm{Po}}^{s\text{-inj.}}$ can be only overestimated by the method (i.e. $R_{\mathrm{Po}}^{s\text{-rec.}} > R_{\mathrm{Po}}^{s\text{-inj.}}$). As

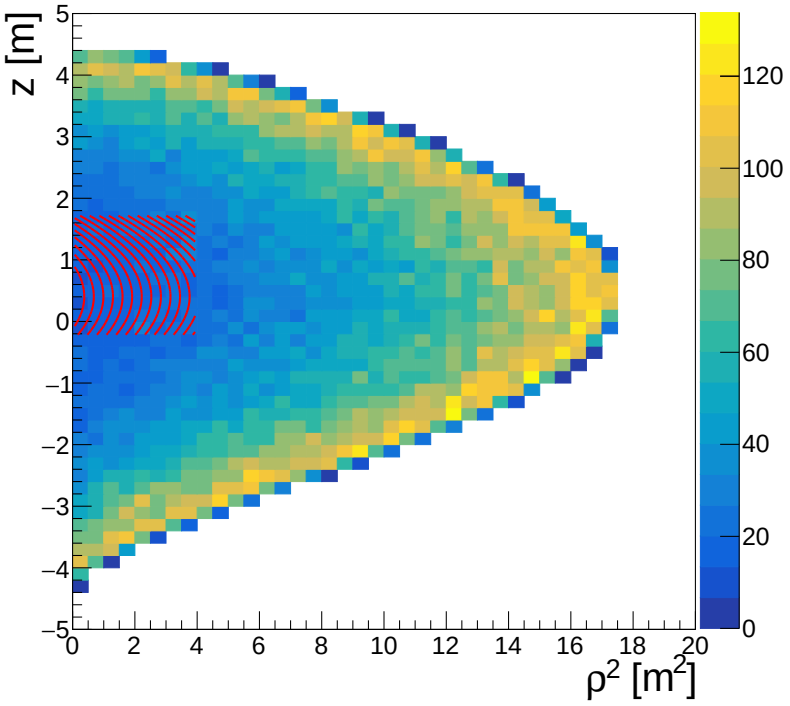


Figure 6.20: Exemplary **toy MC** ^{210}Po **pseudo-dataset fit**: $\rho^2 - z$ histogram of ^{210}Po rate (in color scale), superimposing the fitting according to the clean bubble method (red lines). The input parameters are: $\lambda = 1.0\,\mathrm{m}$, $R_{\mathrm{Po}}^{s\text{-inj.}} = 10\,\mathrm{cpd/100t}$ and, at the ^{210}Po convection minimum, $R_{\mathrm{Po}}^{s\text{-inj.}}/R_{\mathrm{Po}}^{\mathrm{conv.-inj.(min)}} \approx 2$ The fitted region consists of 20 t of scintillator mass.

an example, I report the R_{Po}^{s} distributions scanning the λ parameter, from $\lambda = 0.6\,\mathrm{m}$ to $\lambda = 2.0\,\mathrm{m}$, fixing $R_{\mathrm{Po}}^{s\text{-inj.}} = 10\,\mathrm{cpd/100t}$ and, at the ^{210}Po convection minimum, $R_{\mathrm{Po}}^{s\text{-inj.}}/R_{\mathrm{Po}}^{\mathrm{conv.-inj.}}(\mathrm{min}) \approx 0.5$. In all these cases, it is immediate to note that the $R_{\mathrm{Po}}^{s\text{-rec.}}$ can be only positively biased.

In conclusion, the toy MC study provides confidence that the 2D fit adopted in the *clean bubble method* never underestimates the ^{210}Po supported term, and is therefore suitable for the extraction of an upper limit for the ^{210}Bi rate in the scintillator.

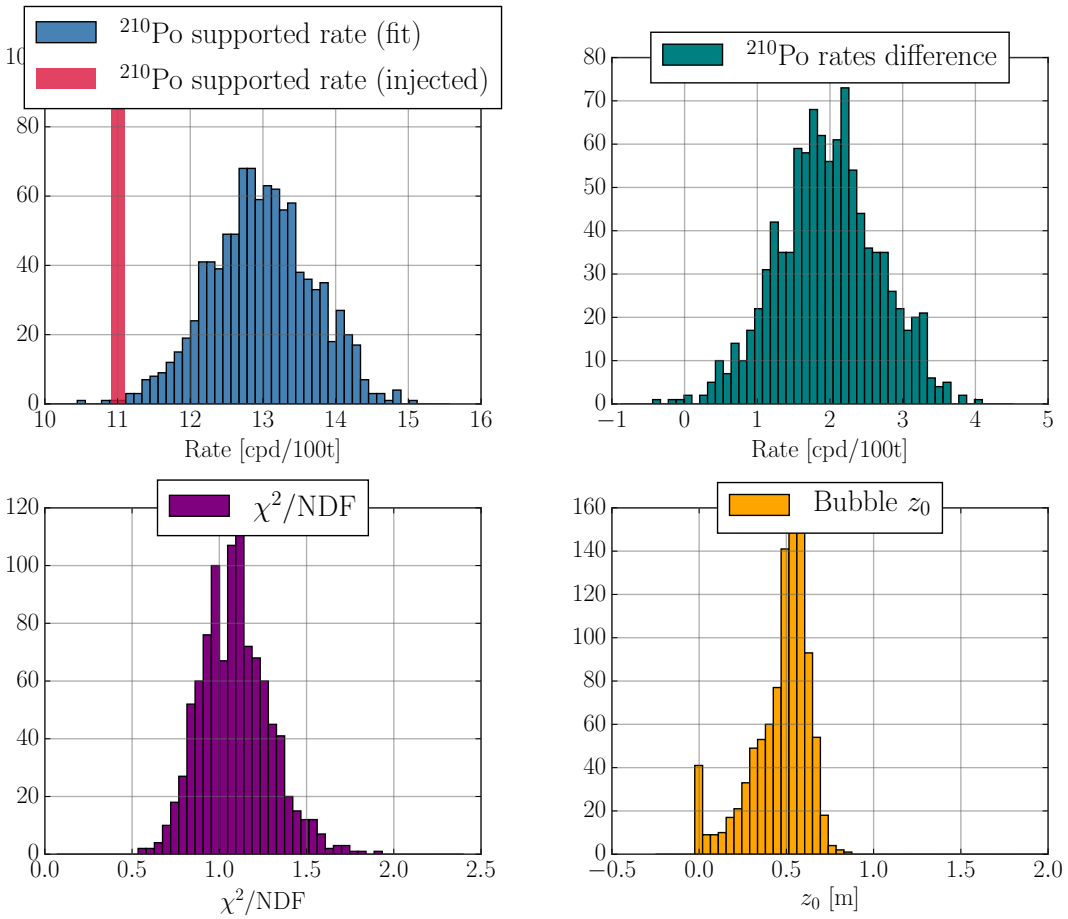


Figure 6.21: Toy MC for the ^{210}Po spatial distribution: results of 10^3 fits of simulated pseudo-experiments with a 2 y statistics. The following distributions are reported: injected and reconstructed rates of supported ^{210}Po (top left panel); difference of these two rates (top right panel); χ^2/NDF of the fit (bottom left panel); z_0 of clean bubble (bottom right panel).

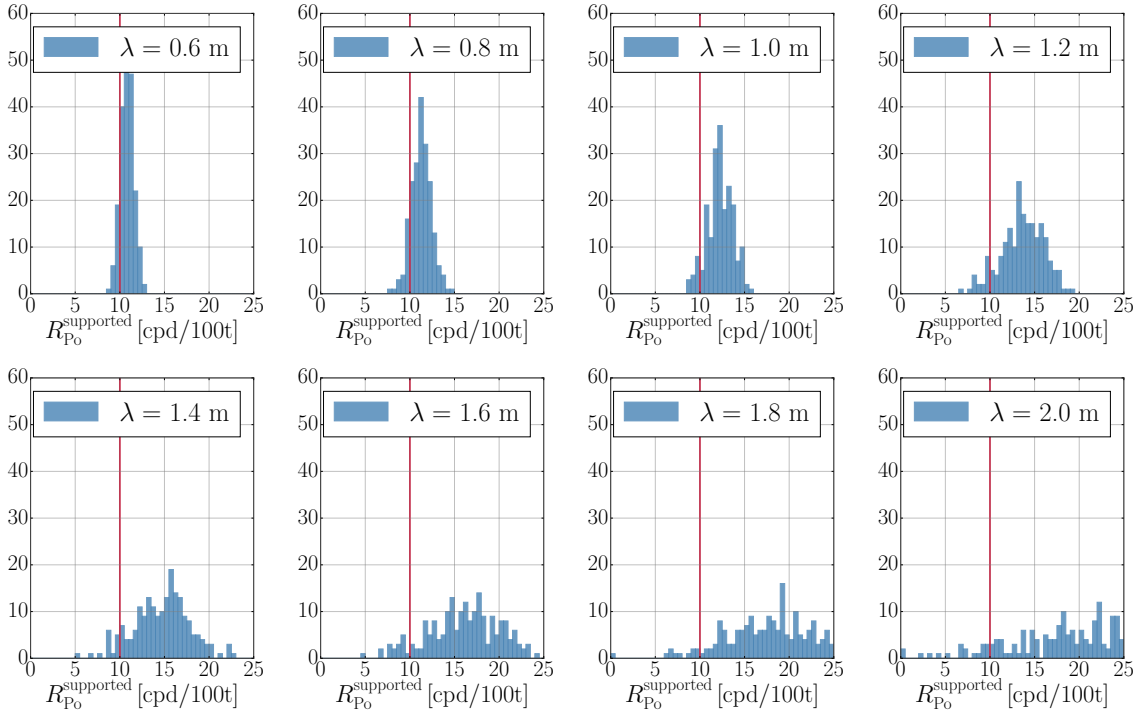


Figure 6.22: Toy MC of ^{210}Po distribution: results of 200 fits of simulated pseudo-experiments with a 2 y statistics varying the λ parameter ($\lambda = 0.6$ m, $\lambda = 0.8$ m, $\lambda = 1.0$ m, $\lambda = 1.2$ m, $\lambda = 1.4$ m, $\lambda = 1.6$ m, $\lambda = 1.8$ m, $\lambda = 2.0$ m). For each inset, the ^{210}Po injected and reconstructed supported rates are reported, respectively as red vertical lines (10 cpd/100t) and as blue histograms.

6.3 Homogeneity and time stability of ^{210}Bi decay rate

In order to apply the result of the clean bubble analysis to the entire FV and in the entire Phase-III-TMP data, we must prove that ^{210}Bi is both uniform and stable in time.

As a preliminary estimation, I have studied the evolution in time and the distribution in space of the β count rate. ^{210}Bi events are a fraction of the β events selected by the basic quality cuts sequence (see Sec. 4.1). In spite of its secondary role in the total counting rate, ^{210}Bi is the only contribution within this energy window which in principle may vary in time and space: the solar neutrino interactions are necessarily stable in time[5] and homogeneously distributed in the detector. Thus, the study of the β total counting analysis variations is directly linked to the determination of ^{210}Bi homogeneity and time stability features.

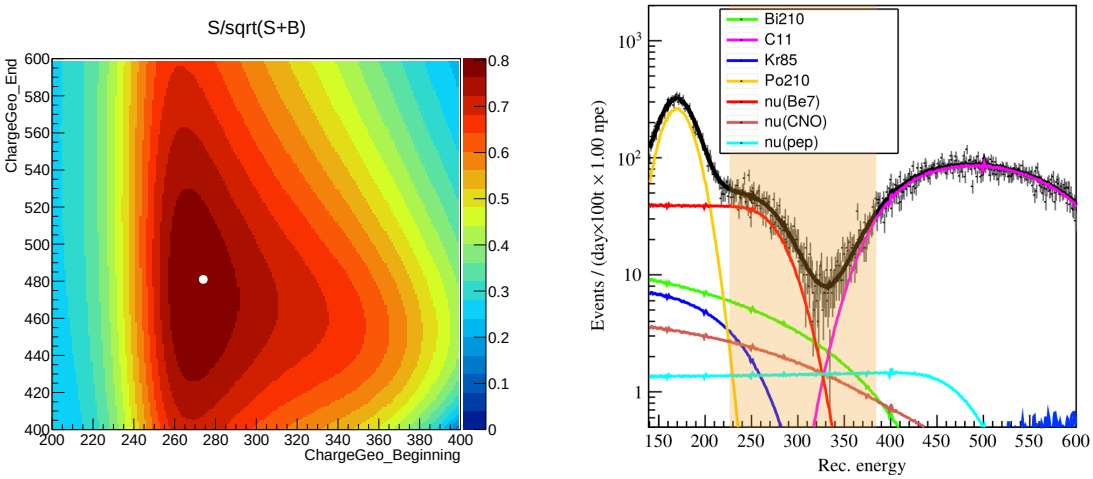


Figure 6.23: β counting analysis: definition of the optimized energy region. Left panel: 2D profile of figure of merit employed for the definition of the optimized energy region, in color scale, as a function of the two window borders. Right panel: collocation of the optimized energy region for the β counting analysis within the Borexino energy spectrum.

In order to decouple this analysis from the complex time-position pattern of the ^{210}Po rate I have selected events in an energy window far from the ^{210}Po peak. Furthermore, I optimized the energy window in order to maximize the contribution of ^{210}Bi with respect to the other species. Figure 6.23 (left panel) shows the 2D profile of the figure of merit ($F = R(^{210}\text{Bi})/\sqrt{R(^{210}\text{Bi}) + R_{\text{all}}}$ as a function of the low and high energy bounds[6]. The best window is found to be $284 \leq N_{pe}^{\text{ng}} \leq 471$ ($0.55\,\text{MeV} \lesssim E \lesssim 0.92\,\text{MeV}$) and is shown in Fig. 6.23 (right panel), superimposed to the Borexino energy distribution and to the simulated spectra of each one of the species. Even in this optimized energy window, the ^{210}Bi events contribute only with $\approx 16\%$ of the total count rate: the other present species are the ^{7}Be-ν ($\approx 61\%$), the CNO-ν ($\approx 8\%$), the ^{11}C ($\approx 7\%$), the pep-ν ($\approx 6\%$) and the ^{85}Kr ($\sim 2\%$).

The time evolution of β event rate in FV and in the optimized energy window is reported in

[5] Except for neutrino flux seasonal modulations, whose amplitude is about 6.7% and with a 1 y period, due to the eccentricity of the Earth's orbit.

[6] The injected rates used for this estimation are, for most of the neutrino and background species, the Phase-II ones; $R(^{210}\text{Bi}) \approx 12\,\text{cpd}/100\text{t}$, as from the clean bubble analysis will show (see Sec. 6.2.2).

Fig. 6.24. In order to damp the fluctuations due to low statistics (in this window the total β count rate is $\sim 12\,\mathrm{cpd/100t}$) I performed a one year moving average: for each time bin considered, the β rate is calculated integrating over the previous and the next six months. This process increases the statistics for each time bin, smoothes the β trend in time, damping the high frequency fluctuations and at the same time averaging the effect of the solar neutrino rate seasonal modulations. As a drawback, it introduces correlations between points distant less than six months from each other.

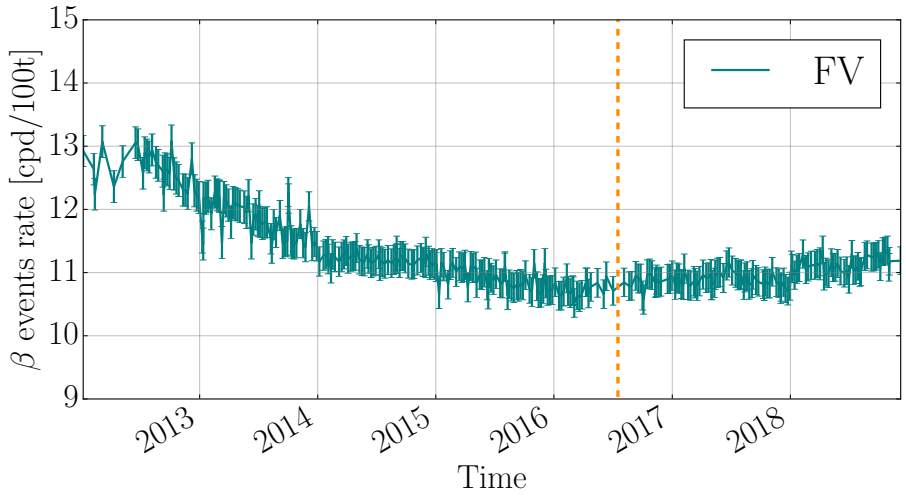


Figure 6.24: β counting analysis in the FV for the $284 \leq N_{pe}^{\mathrm{ng}} \leq 471$ energy window and after a 1 year moving average. The beginning of Phase-III-TMP dataset is marked by the orange dashed line.

Phase-II data show a significant decrease in the event rate going from 2012 to 2016 ($R^{\beta}_{2012} = 12.94 \pm 0.25\,\mathrm{cpd/100t}$ and $R^{\beta}_{2016} = 10.48 \pm 0.25\,\mathrm{cpd/100t}$), while in Phase-III-TMP the rate stabilizes: $R^{\beta}_{2018} = 10.91 \pm 0.25\,\mathrm{cpd/100t}$. This decrease can be explained by the technical details of the purifications carried out in 2010: indeed, the scintillator was extracted from the bottom of the detector, purified and then inserted again in the top hemisphere. This means that, immediately after purifications, the cleanest scintillator in term of ^{210}Pb contamination (and therefore also of ^{210}Bi) was stratified near the top of the Inner Vessel, and some of it was out of the Fiducial Volume considered in the analysis. It took few years for the clean scintillator to completely mix and therefore enter in the FV. This could explain the apparent decrease of ^{210}Bi rate during the Phase-II, and of its subsequent stabilization.

From these considerations, it is clear that we cannot use the full dataset from 2012 until now to perform the CNO-ν analysis. We must select a subset of data where the ^{210}Bi can be considered uniform in space. We decided to consider only data starting from July 2016 (the so-called Phase-III, already introduced in Chapter 3 and in Sec. 4.1.4) with the following motivations:

- We know that ^{210}Pb (and therefore ^{210}Bi and ^{210}Po as well) have a diffusion length of the order of $\lambda \sim 20\,\mathrm{cm}$. This is a typical value for the diffusion length in liquid hydrocarbons, which according to the Fick's laws of diffusion can be evaluated as

$$\lambda = \sqrt{D\tau_{\mathrm{Po}}} \qquad D \sim 10^{-9}\,\mathrm{m^2/s} \tag{6.11}$$

where D is the diffusion coefficient: it can be obtained by the Stokes-Einstein equation which describes the diffusion of spherical particles in a liquid in case of laminar flow. If diffusion were the only mechanism occurring in the scintillator for the ^{210}Pb migration, simulations show that it would have not had time to homogenize in space even after 4-6 years.

- However, we know from the careful analysis of the ^{210}Po distribution (see Fig. 6.8) that, before thermal insulation, convective currents were present in the detector which effectively enlarged the transport length λ and allowed the fluid to mix. From the study of the ^{210}Po movements in 2012-2016 period, we can estimate that the effective λ, resulting from the combination of diffusion and convection, was of the order of several meters.
- Assuming $\lambda = 1\,\mathrm{m}$ (which is a very conservative assumption) we find, from the thermodynamical simulations, that after 4 years from the beginning of Phase-II the fluid should have mixed at least on a mass scale of 20 t.

Following these very conservative estimations we have decided to focus the CNO-ν analysis only on data from July 2016 on (Phase-III). From the previous considerations we are confident that in this period ^{210}Pb (and therefore ^{210}Bi) is uniform on a scale of $\sim 20\,\mathrm{t}$. In the following Sec. 6.3.1 we will evaluate the systematic uncertainty associated to the extrapolation of this hypothesis to the entire volume used in the analysis (71.3 t).

6.3.1 ^{210}Bi spatial homogeneity

The clean bubble analysis described in Sec. 6.2.2 allowed to find the minimum of the ^{210}Po rate distribution in a subset of the FV. This minimum rate will be used to constrain $R(^{210}\mathrm{Bi})$, as an independent measurement, in the multivariate fit for the CNO-ν analysis, in Chapter 7. As hinted in the Sec. 6.3, we must prove the ^{210}Bi rate homogeneity on the full volume used in the CNO-ν analysis.

I have followed two approaches to quantify the uncertainty related to the ^{210}Bi homogeneity:

1. Study of the homogeneity inside and outside the the *clean bubble* region: the β counting rate in a 20 t mass around the ^{210}Po rate minimum is compared with the one the complementary region within the FV.
2. Study of the homogeneity in the z-direction: a set of software-defined *z-slices* is considered, and the homogeneity is evaluated assuming linear variations for the β rate along the z-axis.

In the following, these two methods are reviewed in detail.

^{210}Bi homogeneity inside and outside the *clean bubble* region

In this approach, I have compared the β counting rate in a 20 t mass around the minimum of ^{210}Po rate, and in its complementary region within the FV, throughout the full Phase-III-TMP. I have chosen an energy window $150 \leq N_{pe}^{\mathrm{ng}} \leq 471$ ($0.32\,\mathrm{MeV} \lesssim E \lesssim 0.92\,\mathrm{MeV}$) wider with respect to the one for analysis described in Sec. 6.3, in order to increase the statistics and consequently to lower as much as possible the uncertainty associated to the β extracted rates. This energy range contains a fraction $\varepsilon_{\mathrm{Bi}} = 55.1\%$ of the ^{210}Bi events; since it contains also the ^{210}Po gaussian peak,

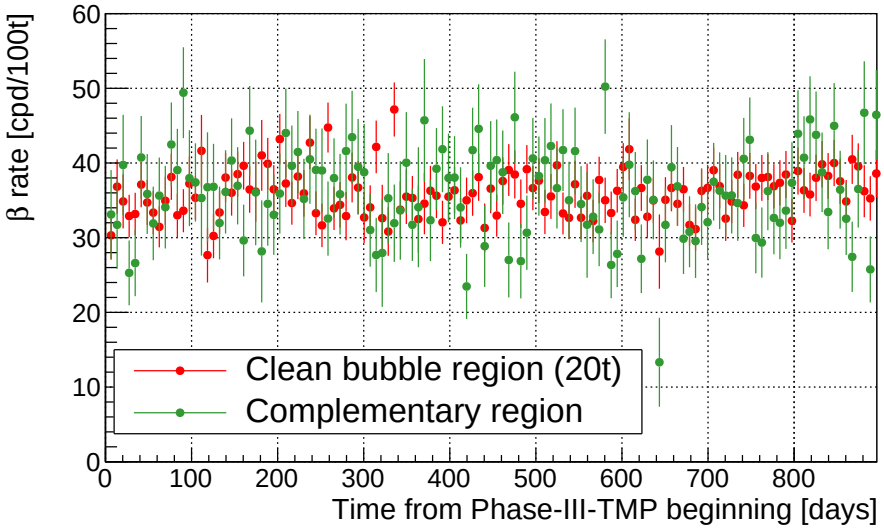


Figure 6.25: β counting analysis: extrapolation of clean bubble analysis result to the FV, time evolution. The evolution of β counting rate in time, in the energy window $150 \leq N_{pe}^{\rm ng} \leq 471$ ($0.32\,\text{MeV} \lesssim E \lesssim 0.92\,\text{MeV}$), is shown for the clean bubble, described in Sec. 6.2.2 (green points), and for the FV (red points).

we must apply a MLP selection cut to remove them, as described in Sec. 6.1.1. Fig. 6.25 shows the β rate trends in time for the clean bubble region (green points) and for the complementary one (red points). Note that the clean bubble trend is more scattered, because of the less available statistics ($\sim 1/3$).

The extracted β rates, integrated over the Phase-III-TMP period, are $R^{\beta}_{\rm Bubble} = 36.23 \pm 0.48\,\text{cpd/100t}$ and $R^{\beta}_{\rm Compl} = 36.04 \pm 0.27\,\text{cpd/100t}$. The β rate difference between the two volumes is $\Delta R^{\beta} = R^{\beta}_{\rm Bubble} - R^{\beta}_{\rm Compl} = -0.19 \pm 0.55\,\text{cpd/100t}$, corresponding to a ^{210}Bi rate difference of

$$\Delta R(^{210}\text{Bi}) = \Delta R^{\beta}/\varepsilon_{\rm Bi} = -0.45 \pm 1.30\,\text{cpd/100t} \tag{6.12}$$

which is consistent with null value.

Homogeneity in the z-direction

As a further crosscheck of the ^{210}Bi uniformity I have also performed a study considering a set of slices symmetric to the z axis, to investigate possible inhomogeneities along the vertical direction. To get sufficiently high statistics, the FV is divided into six 11.86 t slices, and I have considered the same wide energy window $150 \leq N_{pe}^{\rm ng} \leq 471$ ($0.32\,\text{MeV} \lesssim E \lesssim 0.92\,\text{MeV}$) of the previous Section.

The β event rates for each of the six z-slices are reported in Fig. 6.26, for different years in Phase-III (top panels and bottom left panel) and for the full Phase-III-TMP dataset (bottom right panel) in the $150 \leq N_{pe}^{\rm ng} \leq 471$ energy window; it can be seen also by eye that no evident features or variations of the β rate are present along z. As for the previous β rate analysis, the 1 y integration eliminates the neutrinos seasonal modulations.

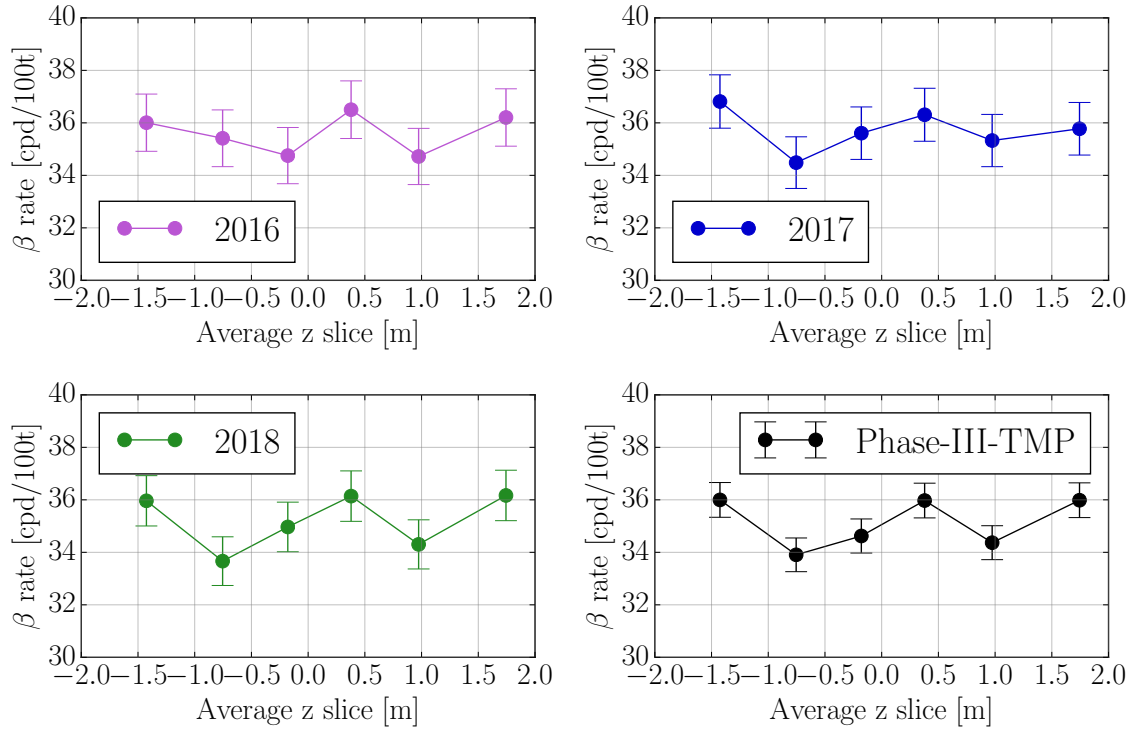


Figure 6.26: β slices counting analysis: β events rate as a function of z for different years in Phase-III (top panels and bottom left panel) and for the full Phase-III-TMP dataset (bottom right panel) in the optimized $284 \leq N_{pe}^{\text{ng}} \leq 471$ energy window.

To quantify the systematic associated to the ^{210}Bi rate homogeneity, we can model linearly the β events rate $R^{\beta}(z)$ and the ^{210}Bi rate as:

$$R^{\beta}(z) = R_0^{\beta} + K_z^{\beta} \cdot z \implies R(^{210}\text{Bi})(z) = R^{\beta}(z)/\varepsilon_{\text{Bi}} \tag{6.13}$$

where K_z^{β} is a *distortion coefficient* which takes into account possible linear variations of the ^{210}Bi rate as a function of z. The ^{210}Bi rate difference within the FV, $\Delta R(^{210}\text{Bi})$, and its uncertainty $\sigma_z(^{210}\text{Bi})$ can be evaluated as:

$$\Delta R(^{210}\text{Bi}) = \frac{K^{\beta}}{\varepsilon_{\text{Bi}}} \cdot \frac{z_{\text{max}} - z_{\text{min}}}{2} \implies \sigma_z(^{210}\text{Bi}) = \frac{\sigma(K_z^{\beta})}{\varepsilon_{\text{Bi}}} \cdot \frac{z_{\text{max}} - z_{\text{min}}}{2} \tag{6.14}$$

where z_{min} and z_{max} represent the z-coordinates boundaries of the FV (respectively, -2.2 m and 1.8 m). The β rates have been linearly fitted according Eq. 6.13: the resulting R_0^{β}, K_z^{β} and $\Delta R(^{210}\text{Bi})$ parameters are reported in Tab. 6.3 for different years of Phase-III-TMP.

K_z^{β} values are compatible with zero for all the time periods analyzed, confirming the β rate homogeneity in a ≈ 10 t mass scale: on the full Phase-III-TMP, one gets $\Delta R(^{210}\text{Bi}) = 0.47 \pm 0.91$ cpd/100t. Note that this error is slightly smaller with respect to what found with the previous method (1.30 cpd/100t). To be conservative, we decided to include the largest error on the systematic budget.

Period	R_0^β [cpd/100t]	K_z^β [cpd/100t m^{-1}]	$\Delta R(^{210}\mathrm{Bi})$
2016	35.6 ± 0.44	0.03 ± 0.42	0.12 ± 1.53
2017	35.7 ± 0.41	-0.09 ± 0.38	-0.32 ± 1.38
2018	35.2 ± 0.39	0.19 ± 0.37	0.69 ± 1.35
Phase-III-TMP	35.1 ± 0.27	0.13 ± 0.25	0.47 ± 0.91

Table 6.3: β slices counting analysis: linear fit parameters (Eq. 6.13) for β events rate in six ~ 11 t isovolumic z-slices defined inside the FV, as a function of years and in the Phase-III-TMP.

6.3.2 Final result on the ^{210}Bi rate

All the systematics associated to the ^{210}Bi independent measurement are reported in Tab. 6.4: the one associated to the ^{210}Po supported rate extraction, as addressed in Sec. 6.2.2, and the one related to the ^{210}Bi homogeneity, as discussed in Sec. 6.3.

To get the final uncertainty for the ^{210}Bi , the Tab. 6.4 ^{210}Bi systematic uncertainty has to be summed in quadrature with the with the 0.8 cpd/100t statistical one. We remind that we are not interested in the negative uncertainty, since in this analysis we are looking for a $R(^{210}\mathrm{Bi})$ upper limit only. Therefore, for the positive uncertainty, we get $\sigma_{\mathrm{Bi}} = \sqrt{\left(\sigma_{\mathrm{Bi}}^{\mathrm{stat.}}\right)^2 + \left(\sigma_{\mathrm{Bi}}^{\mathrm{syst.}}\right)^2} = \sqrt{(0.8\,\mathrm{cpd/100t})^2 + (1.3\,\mathrm{cpd/100t})^2} = 1.6\,\mathrm{cpd/100t}$. Summarizing, from the analysis discussed in this Chapter, I can conclude that my result for the ^{210}Bi rate is $R(^{210}\mathrm{Bi}) = 11.8^{+1.6}_{-2.8}\,\mathrm{cpd/100t}$. This can be converted into an upper limit for ^{210}Bi rate $R(^{210}\mathrm{Bi}) < 14.9\,\mathrm{cpd/100t}$ at 95% C.L. (see Eq. (6.8)).

Systematic source	Systematic error [cpd/100t]
^{210}Bi rate spatial homogeneity	± 1.30
^{210}Po *clean region* size	± 0.42
^{210}Po analytical 2D fitting shape	$^{+0.0}_{-2.4}$
Total budget	$^{+1.4}_{-2.8}$

Table 6.4: β slices counting analysis: linear fit parameters (Eq. 6.13) for β events rate in six ~ 11 t isovolumic z-slices defined inside the FV, as a function of years and in the Phase-III-TMP.

6.4 Conclusions

In this Chapter I have discussed the strategy to determine the ^{210}Bi rate from the ^{210}Po one. After the completion of the Borexino thermal insulation (second half of 2016), the two isotopes are nearly in secular equilibrium, even though the system is still perturbed by residual convective motions which inject small amount of ^{210}Po out of equilibrium in the FV.

In spite of the complexity of the system, I have shown it is possible to extract the ^{210}Po supported term (and therefore the ^{210}Bi rate) by means of a fit to a 20 t clean bubble of the detector almost free from the convective currents (*clean bubble method*). By studying the uniformity in space and

the stability in time of the β count rate, I have demonstrated that the ^{210}Bi rate extracted in the clean bubble can be extrapolated to the entire FV volume and in the entire Phase-III-TMP period. I have extracted the associated systematic uncertainty according to two different methods, and in order to be conservative, I have quoted the largest one.

The final result for the ^{210}Bi rate is $R(^{210}\text{Bi}) = 11.8^{+1.6}_{-2.8}\,\text{cpd/100t}$, which can be converted into an upper limit $R(^{210}\text{Bi}) < 14.9\,\text{cpd/100t}$ at 95% C.L. This value will be used in Chapter 7 to constrain the ^{210}Bi rate in the multivariate fit, in order to extract the CNO-ν interaction rate.

Chapter 7

First indication of a solar neutrinos signal from the CNO cycle reactions

All the analysis ingredients presented in the previous Chapters will be combined in this last Chapter to obtain the first indication of a CNO-ν signal with Borexino detector.

The multivariate fit, used to disentangle the CNO-ν signal from the backgrounds, will be performed on the Phase-III-TMP dataset (17 July 2016 - 02 June 2019). Following the indications coming from the sensitivity studies discussed in Sec. 3.2, the fit will be performed constraining the pep-ν and ^{210}Bi interaction rates to the values obtained independently: the pep-ν rate will be constrained to the SSM predictions, while the ^{210}Bi rate will be constrained to the upper limit derived from the ^{210}Bi -^{210}Po analysis shown in Chapter 6. Both analytical and Monte Carlo approaches for the multivariate fit will be employed, showing consistent results and leading to a CNO-ν signal with significance in the $\sim 4\sigma$ range.

The chapter is organized as follows: the multivariate fit results obtained for Phase-III-TMP dataset by means of Monte Carlo or analytical multivariate analysis will be presented and discussed in Sec. 7.1.1 and Sec. 7.1.2, respectively. An overview of the systematic errors affecting the analysis is given in Sec. 7.2. The significance of the results on CNO-ν is outlined in Sec. 7.3. The implications of the CNO-ν results in term of solar physics will be outlined in Sec. 7.4. Conclusions and perspectives are discussed in Sec. 7.5.

7.1 The CNO-ν analysis

In this Section I will discuss the results of the analysis performed to extract the CNO-ν rate with Borexino. I have analyzed data from the Phase-III-TMP dataset (2016 July 17 - 2019 June 02) which has a total exposure of 644.24 days $\times$ 100 ton. After applying the selection cuts thoroughly described in Sec. 4.1, I performed the multivariate analysis, i.e. the simultaneous fit of the energy, radial and pulse-shape distributions, with both the Monte Carlo and analytical approaches (see Sec. 4.2). The results given by the two methods are consistent, as described in Sec. 7.1.1 and Sec. 7.1.2, respectively.

7.1.1 Monte Carlo based multivariate fit: results

As discussed in Sec. 4.3.2, the Monte Carlo fit is based on the reference shapes for signal and background obtained by simulations: every species which contribute to the total rate in Borexino is simulated for each run of Phase-III-TMP following the experimental conditions recorded in the database (run duration, trigger threshold, number of working PMTs, and so on).

After solving the issue which was causing a mismatch between the simulations and data for Phase-III-TMP, the Monte Carlo was upgraded (see Sec. 5.5 and Sec. 5.6) and I re-simulated the reference PDFs to be used in the multivariate fit. As shown in Sec. 4.2.2, the improved Monte Carlo is now able to fit the Phase-III-TMP data: the fit in the so-called pp chain configuration, i.e. fixing the CNO-ν interaction rate and leaving free all the other neutrino interaction rates, gives results consisted with the one published on Phase-II data [39].

Note that the production of the Monte Carlo reference shapes is not completed yet: ^{14}C and pile-up PDFs are very CPU-time consuming and they are being finalized only in these days. For this reason, the multivariate fit is currently performed only starting from $N_h^n = 140$ (where N_h^n is the normalized number of hits), corresponding to $\approx 320\,\text{keV}$, in order to cut out the low energy part where these species have their spectra. The pile-up tail can actually extends above 300 keV, but I have estimated its contribution to be small $< 0.2\,\text{cpd/100t}$ and I included it in the systematic error budget (see Sec. 7.2). In summary, the multivariate Monte Carlo fit is performed in an energy range between $140 \leq N_h^n \leq 950$, i.e. $320\,\text{keV} \lesssim E \lesssim 2.9\,\text{MeV}$.

The parameters left free to vary in the fit are the rates of the signal and background species. The detector response (energy scale, resolution and so on) is automatically taken into account in the simulated reference PDFs. The only exceptions are the position of the ^{210}Po gaussian peak and the starting point of the ^{11}C spectrum which are left free to adjust on data (see Sec. 4.3.2).

MC fit results [cpd/100t], $140 \leq N_h^n \leq 950$ **range**	
^{7}Be-ν	47.3 ± 1.4
CNO-ν	$5.4^{+1.2}_{-1.0}$
pep-ν	2.74 ± 0.04 [constrained]
^{210}Bi	12.1 ± 1.5 [upper limit]
^{85}Kr	9.5 ± 2.2
^{210}Po	51.46 ± 0.59
^{11}C	27.4 ± 0.5
^{214}Bi (ext. bkg.)	3.65 ± 0.54
^{40}K (ext. bkg.)	0.0
^{208}Tl (ext. bkg.)	5.08 ± 0.27
χ^2/NDF	1.30

Table 7.1: Monte Carlo based fit for Phase-III-TMP dataset: best estimates. The Monte Carlo multivariate fit is performed in the $140 \leq N_h^n \leq 950$ reconstructed energy range. The CNO configuration is adopted, i.e. pep-ν rate constrained to HZ-SSM, ^{210}Bi rate constrained to independent estimations (upper limit $R(^{210}\text{Bi}) < 14.9\,\text{cpd/100t}$ at 95% C.L.), CNO-ν rate free to vary. The best estimates for neutrino and background rates (in cpd/100t units) and the fit χ^2/NDF are reported. Only statistical uncertainties are quoted.

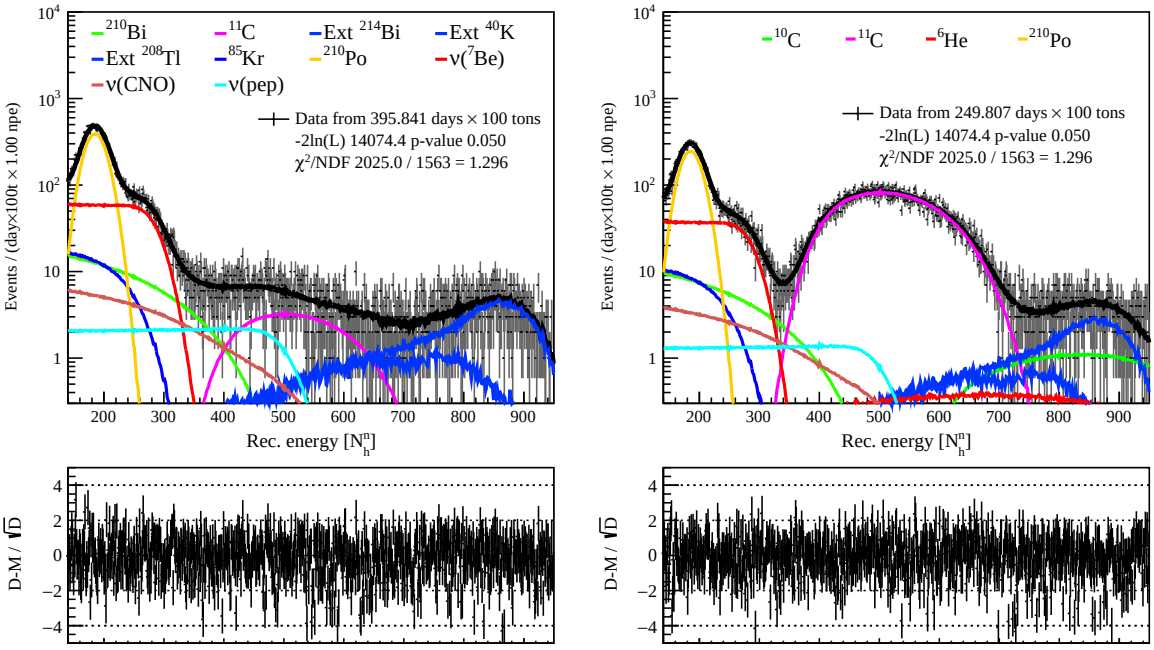


Figure 7.1: Monte Carlo multivariate fit results for Phase-III-TMP dataset in CNO configuration for the TFC-subtracted (left) and the TFC-tagged (right) energy spectra (top panels), with residuals (bottom panels). The sum of the individual components from the fit (black lines) is superimposed on the data (grey points). The analysis has been performed within the $140 \leq N_h^n \leq 950$ energy window, excluding the lower energy fraction of the energy spectrum. The residuals are calculated in every bin as the difference between the data counts and the fit result, divided by the square root of the data counts. The interaction rates are reported in Tab. 7.1.

Following the indications of the sensitivity studies discussed in Chapter 3, in order to break the correlation between CNO-ν, ^{210}Bi and pep-ν spectral shapes, I constrained the interaction rates of ^{210}Bi and pep-ν species to the values obtained independently. For the pep-ν specie, the SSM predictions are exploited, leading to the constraint $R(\text{pep-}\nu) = 2.74 \pm 0.04\,\text{cpd/100t}$. Note that the reported uncertainty of 1.45% accounts for both the theoretical error on the neutrino flux and the one on the oscillation parameters. For what concerns the ^{210}Bi rate the results obtained by means of the ^{210}Po clean bubble method described in Sec. 6.2.2 have been employed, including also the systematic associated to the extrapolation of the clean bubble result to the entire FV used for the CNO-ν analysis (see Sec. 6.3.2): $R(^{210}\text{Bi}) < 14.9\,\text{cpd/100t}$ at 95% C.L.

The Phase-III-TMP Monte Carlo multivariate fit results in this configuration (from now on called *CNO configuration*) are reported in Tab. 7.1: only statistical uncertainties are quoted. Fig. 7.1 displays the TFC-subtracted (left panel) and TFC-tagged (right panel) energy spectra, along with the associated residuals (bottom panels). The data model agreement is good, quantified by $\chi^2/\text{NDF} = 1.30$, and is similar to what was found for the Phase-II multivariate fit. The residual plots don't show any anomalous structure nor biases. The extracted CNO-ν rate is $R^{\text{MC}}(\text{CNO-}\nu) = 5.4^{+1.2}_{-1.0}\,\text{cpd/100t}$ (statistical error only). The other rates in output of the fit (for example the ^{7}Be-ν rate) return results consistent with what was obtained and published for Phase-II [39].

7.1.2 Analytical multivariate fit: results

In the following, the results obtained with the analytical multivariate fit for the Phase-III-TMP dataset will be shown.

In this case the reference shapes for signal and background are obtained by means of an analytical model, thoroughly discussed in Sec. 4.2.1. Some of the parameters which describe the detector response are left free to vary in the fit (light yield, resolution-related parameters, and so on) together with the rates of neutrinos and backgrounds. The large number of free parameters makes the analytical fit more flexible with respect to the Monte Carlo fit. On the other hand, the analytical fit is more prone to correlations between species, induced by the fact that the response parameters can slightly adjust the energy scale and alter the spectral shape of each species.

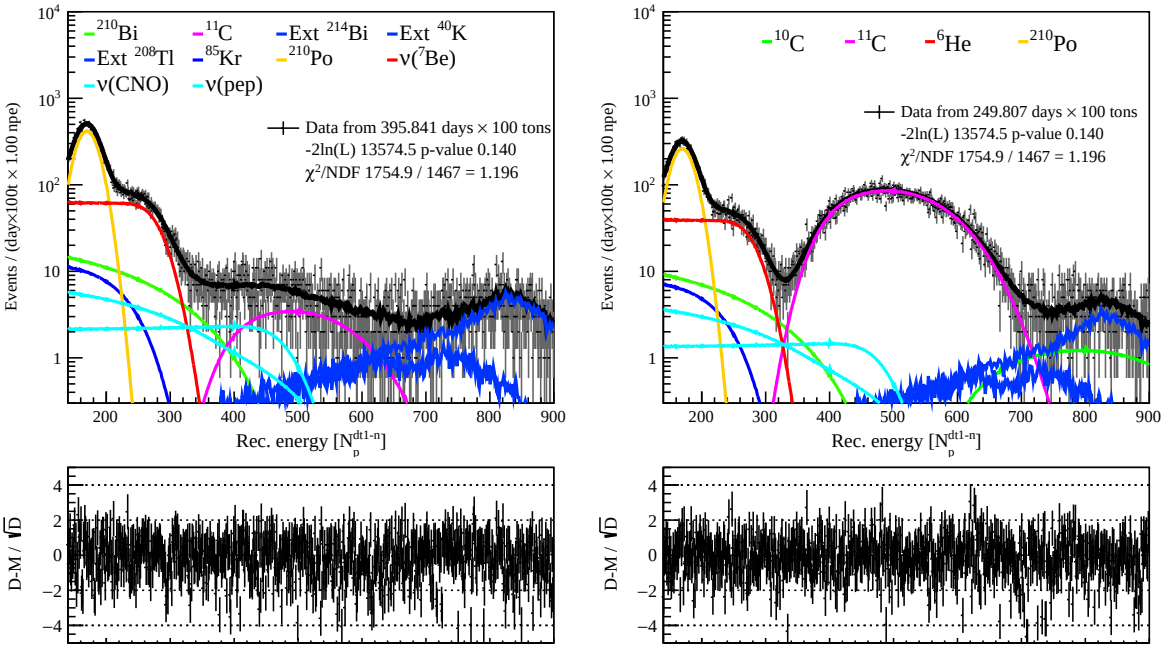


Figure 7.2: Analytical multivariate fit results for Phase-III-TMP dataset in CNO configuration for the TFC-subtracted (left) and the TFC-tagged (right) energy spectra (top panels), with residuals (bottom panels). The sum of the individual components from the fit (black lines) is superimposed on the data (grey points). The analysis has been performed within the $140 \leq N_p^{\mathrm{dt1-n}} \leq 900$ energy window, excluding the lower energy fraction of the energy spectrum. The residuals are calculated in every bin as the difference between the data counts and the fit result, divided by the square root of the data counts. The interaction rates are reported in Tab. 7.2.

The analytical fit is performed in the range $140 \leq N_p^{\mathrm{dt1-n}} \leq 900$ which corresponds to 330 keV $\lesssim E \lesssim$ 2.9 MeV. Like in the Monte Carlo fit discussed in the previous paragraph, also in this case the lower energy portion of the spectrum is not considered. While for Monte Carlo this was due to a practical issue, i.e. the PDFs of the species in the lower energy region are not ready yet, in this case the reason is connected to a problem already mentioned in Sec. 4.3.1. Indeed, in Phase-III-TMP the analytical multivariate fit performed in the so-called pp configuration (i.e. fixing the CNO-ν rate to the SSM predictions) returns value for the pp-ν which are significantly higher than the

Analytical fit results [cpd/100t], $100 \leq N_p^{\mathrm{dt1-n}} \leq 900$ **range**	
^{7}Be-ν	51.2 ± 6.3
CNO-ν	$6.2^{+2.7}_{-2.6}$
pep-ν	2.74 ± 0.04 [constrained]
^{210}Bi	8.1 ± 5.3 [upper limit]
^{85}Kr	1.2 ± 17.4
^{210}Po	47.3 ± 1.9
^{11}C	27.3 ± 0.4
^{214}Bi (ext. bkg.)	2.92 ± 0.52
^{40}K (ext. bkg.)	0.0
^{208}Tl (ext. bkg.)	5.36 ± 0.28
χ^2/NDF	1.20

Table 7.2: Analytical fit for Phase-III-TMP dataset: best estimates The analytical multivariate fit is performed for in the $140 \leq N_p^{\mathrm{dt1-n}} \leq 900$ reconstructed energy range. The CNO configuration is adopted, i.e. pep-ν rate constrained to HZ-SSM, ^{210}Bi rate constrained to independent estimations ($R(^{210}\mathrm{Bi}) < 14.9$ cpd/100t at 95% C.L.), CNO-ν rate free to vary. The best estimates for neutrino and background rates (in cpd/100t) and the fit χ^2/NDF are reported. Only statistical uncertainties are quoted.

Standard Model predictions and not consistent with our previous result published on Phase-II [39]. This effect is evident in the last years of Phase-III-TMP, especially in 2018. The issue is currently under investigation. Some hypotheses on its origin (change in the response function due to loss of PMTs, wrong treatment of pile-up) have been studied (see Section 4.3.1) and have been ruled out. I am currently studying the possibility that this high pp-ν rate is due to some spurious noise affecting data after 2018. In any case, the problem does not seem to affect the higher energy portion of the spectrum. For this reason I decided to perform the fit above $N_p^{\mathrm{dt1-n}} = 140$ (≈ 330 keV). Of course, the range restriction reduces the available information for the fit, implying that some of the species which have a large part of their spectrum at lower energies (like ^{7}Be-ν and ^{85}Kr background) are determined with less precision with respect to the fit which exploits the full energy range. Furthermore, the low energy part of the spectrum (dominated by ^{14}C events) helps constraining the energy scales and reduces the systematics of the analytical fit. So, it is auspicable that in a near future the issue is solved and the fit will be performed in the entire energy range.

Results for the analytical fit for Phase-III-TMP dataset, performed in the CNO configuration, i.e., constraining pep-ν and ^{210}Bi to the value obtained independently, are shown in Fig. 7.2. The best estimates for neutrino and background rates are reported in Tab. 7.2. The extracted CNO-ν rate is $R^{\mathrm{AN}}(\mathrm{CNO}\text{-}\nu) = 6.2^{+2.7}_{-2.6}$ cpd/100t, consistent with the result obtained by the Monte Carlo fit.

7.2 Overview of the systematic errors

In this section, the main sources expected to provide a systematic uncertainty to the CNO-ν measurement are summarized. This part of the analysis is still incomplete and is planned to be more deeply and accurately investigated in the next months; however, the conclusions on the CNO-ν

significance are not expected to change dramatically: indeed, we expect the overall uncertainty of the CNO-ν rates to be dominated by the error in output from the multivariate fit ($\sim 25\%$), mainly due to the error on ^{210}Bi constraint.

Fit approach

The two Monte Carlo and analytical approaches return compatible CNO-ν rates $R^{\mathrm{AN}}(\mathrm{CNO}\text{-}\nu) = 6.2^{+2.7}_{-2.6}$ cpd/100t (analytical) and $R^{\mathrm{MC}}(\mathrm{CNO}\text{-}\nu) = 5.4^{+1.2}_{-1.0}$ cpd/100t (Monte Carlo), showing the consistency of the two methods. Consequently, no systematic uncertainties has been associated to the fit approach.

Correlations between species

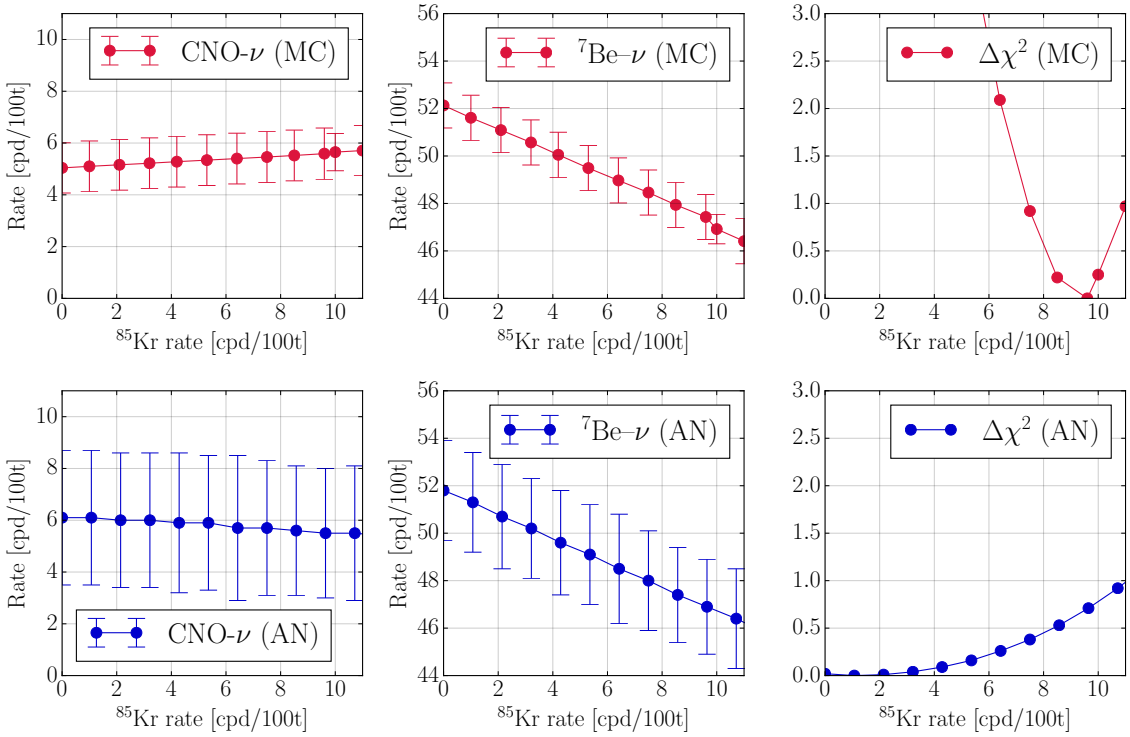


Figure 7.3: CNO-ν analysis: systematics associated to the ^{85}Kr rate. The multivariate fits for both the Monte Carlo (red) and analytical approaches (blue) has been performed according to the configurations illustrated in the previous Sec. 7.1; in addition, the ^{85}Kr has been fixed to several values within reasonable ranges. The variables reported are the CNO-ν rate (left), the ^{7}Be-ν rate (center), the $\Delta\chi^2 = \chi^2 - \chi^2_{\mathrm{min}}$ (right).

Correlations among species involved in the fit may bias the CNO-ν results. For example, dedicated sensitivity studies [63] confirmed by the analysis of Borexino Phase-I and Phase-II data have shown that the ^{7}Be-ν and ^{85}Kr rates are slightly anticorrelated in the multivariate fit, due to the similar recoil e^- spectral shape and to the close end-point energy.

Therefore it is important to study whether a similar correlation ^{85}Kr-CNO-ν is present also in the CNO configuration, and if it could possibly bias the CNO-ν extracted rate. Similarly, the systematics associated to the pep-ν constraint are studied setting a pull term based on the LZ-SSM scenario instead of HZ-SSM one, which is done as default in CNO configuration.

- **Stability for ^{85}Kr background rate variations.** The ^{85}Kr background, which lies in the lowest part of the energy spectrum, has an energy endpoint of862 keV, therefore only a small fraction of its spectrum is found in the fit range. For these reasons, the multivariate fit has little sensitivity to it, and the ^{85}Kr extracted rate can be biased by correlations with other species.

  To study these possible correlations, the fit has been repeated fixing the ^{85}Kr rate to different values, ranging from 0 cpd/100t to 11 cpd/100t. This latter value is defined by an independent analysis for the ^{85}Kr rate determination, which exploits the tagging of ^{85}Kr $-^{85m}$ Rb decays delayed coincidences (see Sec. 2.3.1); the resulting ^{85}Kr rate upper limit [65] is $R(^{85}\text{Kr}) < 10.9$ cpd/100t at 95% C.L for Phase-III-TMP period.

  Results for Monte Carlo and analytical approach are reported in Fig. 7.3, where the top and the bottom rows refer to the Monte Carlo and analytical approach respectively: the reported quantities are the CNO-ν rate (left), the ^{7}Be-ν rate (center), the $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ (right). The extracted CNO-ν rates are found compatible with the results shown in Tab. 7.1 and Tab. 7.2; in conclusion, we do not associate any systematic error to the CNO-ν result for the ^{85}Kr correlation.

- **Stability for pep-ν constraint variations.** A secondary source of systematics is given by the pep-ν pull term; it can be set according to LZ-SSM predictions instead of HZ-SSM ones. Results on CNO-ν rate are compatible, showing negligible variations at the < 0.1 cpd/100t level.

- **^{210}Bi constraint.** Note that the systematics associated to the ^{210}Bi constraint are implicitely included in the ^{210}Bi penalty uncertainty, as explained in Sec. 6.2.2 and Sec. 6.3 consequently, no systematics associated to the ^{210}Bi pull term need to be studied in the multivariate fit. As shown by the sensitivity studies (Sec. 3.2), the ^{210}Bi constraint uncertainty impacts crucially on the precision of the CNO-ν rate.

^{11}C energy scale

We have hinted previously (Sec. 7.1.1 and Sec. 4.3.2) how, in the Monte Carlo fit approach, the parameters left free to vary are the interaction rates (for signal and bakgrounds), while the energy scale parameters for ^{210}Po and ^{11}C backgrounds are free to perform fine tunings. In particular, the ^{11}C species takes on a relevant role for the CNO-ν analysis, since the lowest portion of its spectrum may overlap with the CNO-ν one; therefore, if its shape is not properly reproduced by the Monte Carlo simulations, the CNO-ν rate extracted by the fit may be somehow biased.

The choice to leave ^{11}C energy scale as a free parameter even in the Monte Carlo case is justified by the fact that it is the only e^+-induced species in the Borexino spectrum, and that during the detector calibrations (see Sec. 4.1.1) no β^+ sources were employed. Therefore, energy scale shifts at a $\sim 1\% - 2\%$ level are justified.

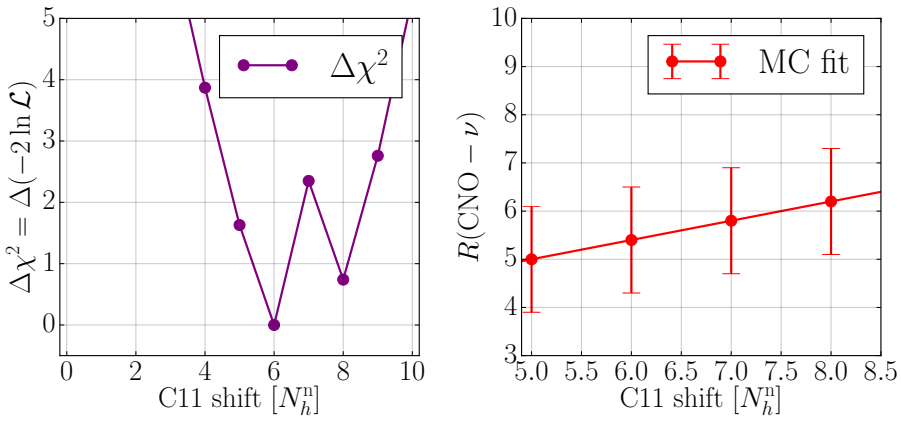


Figure 7.4: CNO-ν analysis: systematics associated to the ^{11}C energy scale. Left panel: $\Delta\chi^2$ profile extracted fitting a highly pure ^{11}C dataset and varying the ^{11}C energy shift parameter. Right panel: $R(\mathrm{CNO}\text{-}\nu)$ as a function of ^{11}C energy shift parameter, obtained fitting the standard Phase-III-TMP dataset with the Monte Carlo approach.

We can evaluate the systematics associated to the ^{11}C energy scale considering a highly pure e^+ (i.e. ^{11}C) sample obtained by data, applying the TFC algorithm on Phase-III-TMP data with very tight selection cuts (see Sec. 4.1.3); the multivariate fit is performed on this dataset varying the ^{11}C energy scale parameter, and extracting the $\Delta\chi^2$ profile (see left panel of Fig. 7.4). Then, the CNO-ν rate distribution is built performing the multivariate fits on the standard Phase-III-TMP dataset, and considering only the fits in a $\pm1\sigma$ region around the $\Delta\chi^2$ minimum. The CNO-ν results, reported in right panel of Fig. 7.4, are fully compatible with the ones reported in Tab. 7.1 and Tab. 7.2: therefore, no systematic uncertainty is associated.

Energy estimator

The multivariate fit is repeated assuming a reconstructed energy estimator different from the usual N_h^n and $N_p^{\mathrm{dt1-n}}$ ones adopted for the Monte Carlo and analytical approaches respectively. Results on the CNO-ν rate using other energy estimators (i.e. $N_p^{\mathrm{dt2-n}}$, another the fixed duration estimator[1] described in Sec. 4.1) are different by less of $\sim 0.1\,\mathrm{cpd/100t}$; therefore, no systematic error is associated to the energy estimator choice.

Fit range

The systematic error associated to the fit range choice can be evaluated scanning the lower energy threshold. The CNO-ν (left panel) and ^{7}Be-ν (right panel) rates as a function of the energy fit starting point are shown in Fig. 7.5; trends associated to analytical and Monte Carlo fits are shown in red and blue lines respectively. Only small variations of the results are observed for the CNO-ν rate, marginal if compared to the statistical uncertainty, proving the CNO configuration solidity; consequently we don't associate any related systematic error. This fit range stability confirms the

[1] We remind that the $N_p^{\mathrm{dt2-n}}$ energy estimator counts the number of fired PMTs within 400 ns from the cluster beginning, normalized to a fixed number of 2000 working PMTs.

robustness of the CNO-ν rate result for both the approaches in the chosen fit energy range.

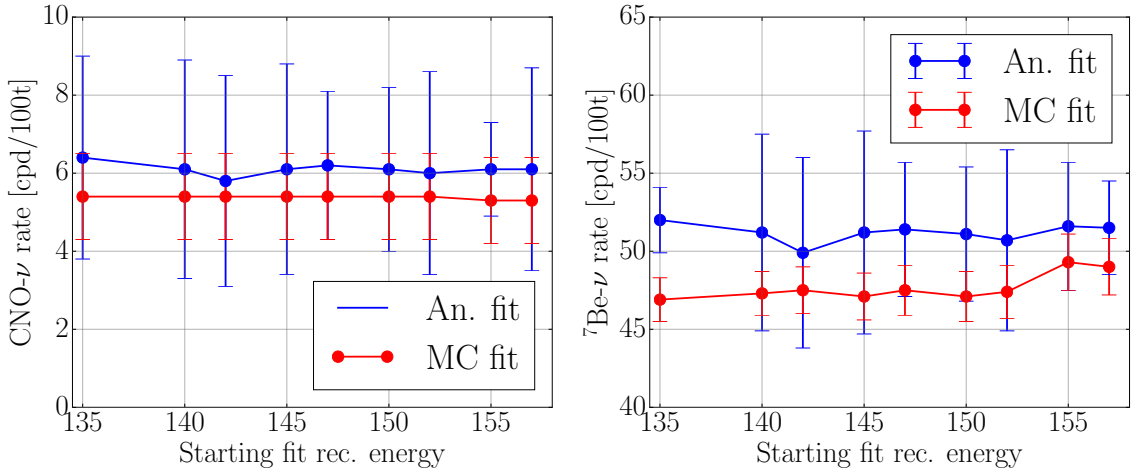


Figure 7.5: Stability against fit energy range. The CNO-ν and ^{7}Be-ν rates are reported as a function of the starting fit energy in $N_p^{\mathrm{dt1-n}}$ or N_h^n (respectively for Monte Carlo or analytical approaches, in red or blue).

Pile-up

While in the analytical approach the pile-up species is automatically included in the fit by means of the convolution method (see Sec. 4.2.1), in the Monte Carlo approach it must be simulated and calculated in the fit separately. As mentioned in Sec. 7.1.1, the pile-up simulations after the Monte Carlo upgrade are still ongoing, therefore pile-up is not included in the Monte Carlo fit. The energy range of the fit $140 \leq N_h^n \leq 950$ still includes a small tail of pile-up events: from independent estimations, I have evaluated its contribution to be $\lesssim 0.20\,\mathrm{cpd/100t}$. This is accounted for as a systematics for the Monte Carlo approach.

Time stability

The multivariate fit is repeated in the three subperiods of Phase-III-TMP dataset: Phase-III-A (July 2016 - July 2017), Phase-III-B (July 2017 - July 2018), Phase-III-C (July 2018 - July 2019). The results are shown in Fig. 7.6, displaying the CNO-ν rate as a function of the dataset time, in the standard CNO configuration, for the Monte Carlo (right panel) and analytical fit approaches. For each of the two panels, the horizontal black lines and the green bars represent the $R(\mathrm{CNO}\text{-}\nu)$ central value and 1σ C.L. obtained analyzing the full Phase-III-TMP period.

In spite of the lower statistics with respect to the full Phase-III-TMP data, the one-year datasets show that Borexino is still sensitive to the CNO-ν signal, displaying compatible results for different years.

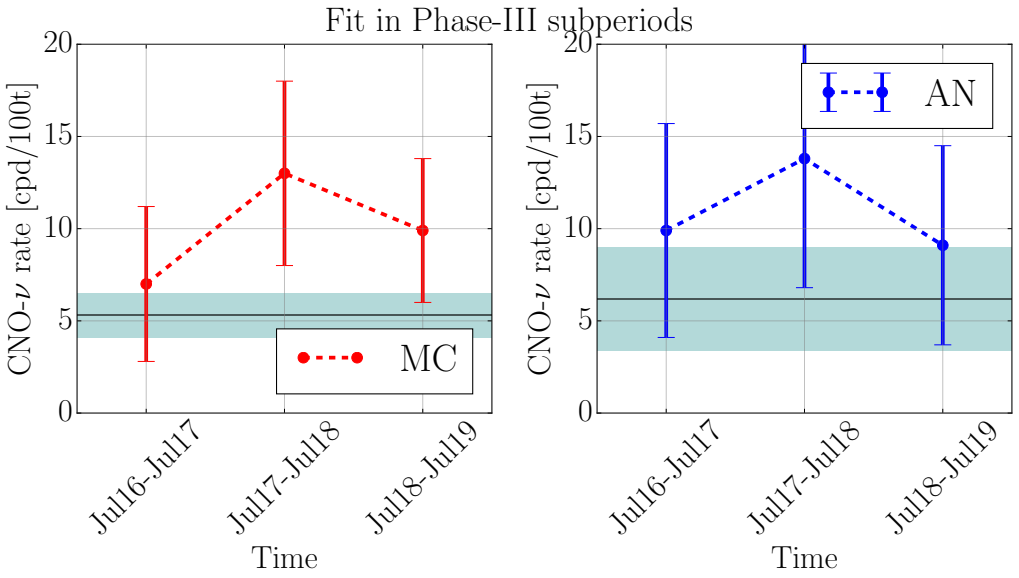


Figure 7.6: Time stability: the Monte Carlo and analytical fits (left and right panels, respectively) are shown in three Phase-III-TMP yearly subperiods, in the standard CNO configuration. For each of the panels, the horizontal black lines and the green bars represent the $R(\text{CNO-}\nu)$ central value and 1σ C.L. obtained analyzing the full Phase-III-TMP period.

Detector-related systematics

A source of systematics could be the non-accurate modeling of the detector response. This has been thoroughly studies for the fit in the pp chain configuration, as published in Ref. [39]. Ensembles of pseudo-experiments have been generated from a family of PDFs based on the full Monte Carlo simulations and fitted using both the Monte Carlo and analytical methods. PDFs including deformations due to possible inaccuracies in the modeling of the detector response (energy scale, uniformity of the energy response, shape of e^+/e^- discriminator variable) and uncertainties in the theoretical energy spectra (^{210}Bi) have to be considered. The magnitude of the deformation was chosen to be within the range allowed by the available calibration data [63].

This study should be repeated for the fit in CNO configuration. We can assume for now that the order of magnitude of the systematic error on CNO-ν rate due to the detector response modeling is similar to the one quoted for the pep-ν species (which has a similar energy range) in the pp chain configuration, i.e. $\approx 0.25\,\text{cpd}/100\text{t}$ [39].

Exposure-related systematics

More marginal sources of systematics, not associated to the fit procedure, are due to the accuracy of fiducial mass determination, of the data taking livetime and of the scintillator density. The fiducial volume uncertainty is the dominant one, and its value corresponds to what calculated in Ref. [65] by means of calibration sources of known positions: $^{+0.6\%}_{-1.1\%}$. The other two sources contribute at $\sim 0.05\%$ level. All these systematics are negligible with respect to the statistical CNO-ν rate uncertainty.

Total systematic uncertainty budget

The systematic uncertainties described in the previous paragraphs for Monte Carlo and analytical fits are reported in the second and third columns of Tab. 7.3. The overall uncertainties is obtained summing in quadrature the quoted errors: $^{+0.33}_{-0.32}$ cpd/100t for the Monte Carlo approach and $^{+0.26}_{-0.25}$ cpd/100t for the analytical one.

	R(CNO - ν) systematics [cpd/100t]	
	Monte Carlo	Analytical
Detector-related	± 0.25	± 0.25
Exposure-related	$^{+0.06}_{-0.03}$	$^{+0.06}_{-0.03}$
Pileup inclusion	± 0.2	-
Overall	$^{+0.33}_{-0.32}$	$^{+0.26}_{-0.25}$

Table 7.3: Relevant sources of systematic uncertainty and their contributions to R(CNO - ν) in cpd/100t, for both the Monte Carlo and analytical approaches (second and third columns, respectively).

7.3 Likelihood profiles for the CNO-ν rate

The systematics study shown in the previous Section indicates that the overall uncertainty on the CNO-ν rate is largely dominated by the statistical error (~ 1.3 cpd/100t) with respect to the systematic one (~ 0.3 cpd/100t). Even though the study of all the possible sources of systematics is not completed yet, we can conclude that the result is robust and provides a convincing indication of the presence of a CNO-ν signal.

Including both the statistical and systematic uncertainties, the best estimates for CNO-ν interaction rates are $R^{\mathrm{MC}}(\mathrm{CNO}\text{-}\nu) = 5.4^{+1.2}_{-1.0}\,(\mathrm{stat.}) \pm 0.3\,(\mathrm{syst.})\,\mathrm{cpd/100t} = 5.4^{+1.3}_{-1.1}\,\mathrm{cpd/100t}$ for the Monte Carlo fit and $R^{\mathrm{AN}}(\mathrm{CNO}\text{-}\nu) = 6.2^{+2.7}_{-2.6}\,(\mathrm{stat.}) \pm 0.3\,(\mathrm{syst.})\,\mathrm{cpd/100t} = 6.2^{+2.7}_{-2.6}\,\mathrm{cpd/100t}$ for the analytical fit.

The confidence interval for this measurement can be precisely extracted by means of the profile likelihoods of CNO-ν interaction rate. These are obtained scanning the R(CNO - ν) parameter space, and are displayed for both analytical and Monte Carlo approaches in Fig. 7.7 (red and blue contours respectively). For both the two approaches, the absence of a CNO-ν signal can be ruled out in the $\sim 4\sigma$ statistical significance range, compatibly with the expectations from sensitivity studies illustrated in Fig. 3.8. In particular, for the analytical and Monte Carlo fits we get 3.6σ and 4.9σ significances respectively. The latter approach is shown to be more stable, thanks to the lower number of free parameters with respect to the analytical approach.

Note that the profile likelihoods are appreciably asymmetric, due to the ^{210}Bi upper limit set in the fit. Indeed, its presence strongly disfavours the lowest CNO-ν values ($R(\mathrm{CNO}\text{-}\nu) \lesssim 5\,\mathrm{cpd/100t}$), while it impacts in a milder way on the higher ones, since a lower limit on the ^{210}Bi background is not provided. Each of the $\Delta\chi^2$ profiles shown can be parametrized with two independent parabolic shapes, each for one side with respect to the minima. The confidence intervals for R(CNO - ν) at 1σ, 2σ and 3σ levels are quoted in Tab. 7.4

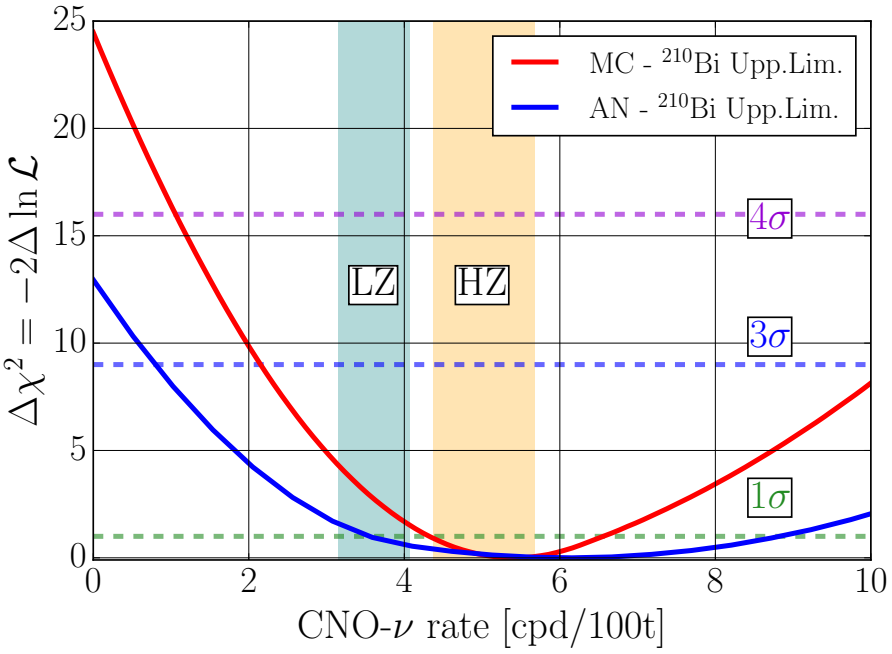


Figure 7.7: Profile likelihoods of the CNO-ν interaction rate both for Monte Carlo and analytical fits (respectively, solid red and solid blue lines). The green and yellow vertical bands show the $\pm 1\sigma$ confidence intervals for the CNO-ν rate in Borexino according to the LZ-SSM and HZ-SSM predictions [24]. The horizontal lines represent three confidence levels for the hypothesis exclusion.

For what concerns solar physics, the result on the CNO-ν flux is found in agreement with the predictions of the SSM both in the high and in the low metallicity hypotheses. Even though our measurements are closer to the HZ-SSM predictions, the associated errors are too large to discriminate between the two hypotheses. This topic is addressed with more detail throughout the next Sec. 7.4.

7.4 Implications on solar physics

7.4.1 Discrimination of metallicity scenario

As discussed in Sec. 1.4, several works (Ref. [31, 32, 33]) have reviewed the spectroscopic determination of the solar photospheric composition. As a consequence of this revision, the input metallicity abundance has been lowered to $Z/X = 0.0165$ (AGS09, *low metallicity* or *LZ* scenario [31]), that is about 28% less than the previous GS98 value (*high metallicity* or *HZ* scenario). Unfortunately, solar models incorporating the LZ hypothesis metallicity as input parameters are not able to describe properly the helioseismology results (as for example, the sound speed profile, as shown in Fig. 1.6); on the other hand, the HZ-SSM predictions are found in better agreement with the same experimental measurements.

The Borexino measurements of the ^{7}Be-ν and ^{8}B-ν fluxes are found in a mild tension against the

	C.I. for $R(\text{CNO-}\nu)$ [cpd/100t]	
	Analytical	Monte Carlo
1σ (68.3% C.L.)	$3.6 - 8.8$	$4.3 - 6.6$
2σ (95.5% C.L.)	$2.1 - 11.3$	$3.2 - 8.3$
3σ (99.7% C.L.)	$0.8 - 12.5$	$2.2 - 10.3$

Table 7.4: Confidence intervals for $R(\text{CNO-}\nu)$ in cpd/100t from the Monte Carlo and analytical Phase-III-TMP multivariate fits reported in Sec. 7.1.1 and Sec. 7.1.2.

most quoted LZ-SSM scenario (rejecting it at 96.6% C.L., i.e. standardized 1.8σ), while are finely consistent with HZ-SSM predictions [39]. Also the CNO-ν Borexino measurement can be used to probe the solar metallicity: the CNO-ν flux indeed depends strongly both on the Sun temperature and on the C+N abundance. The difference in the predictions of CNO-ν flux assuming the two metallicity sets is about 28%.

Considering the HZ-SSM and LZ-SSM as two alternative models, I have performed a frequentist hypotheses test, considering the ^{7}Be-ν and ^{8}B-ν fluxes determined by the Phase-II analysis [39] and the CNO-ν flux result obtained in this thesis by means of the Monte Carlo fit (Sec. 7.1). I define a 3D χ^2 estimator as:

$$\chi^2(\text{SSM}) = \left(\Phi^{\text{SSM}} - \Phi^{\text{Exp}}\right)^{\text{T}} \left(\Sigma^{\text{SSM}} + \Sigma^{\text{Exp}}\right)^{-1} \left(\Phi^{\text{SSM}} - \Phi^{\text{Exp}}\right) \tag{7.1}$$

where Σ^{SSM} and Σ^{Exp} are the covariance matrices of the $(\Phi_{^7\text{Be}}, \Phi_{^8\text{B}}, \Phi_{\text{CNO}})$ SSM fluxes (which is extracted by the more complete Tab. 1.2 [36]) and of the mentioned Borexino measurements respectively[2]:

$$\Sigma^{\text{SSM}}_{\text{HZ}} = \begin{pmatrix} 0.090 & 0.168 & 0.084 \\ 0.168 & 0.436 & 0.221 \\ 0.084 & 0.221 & 0.608 \end{pmatrix} \quad \Sigma^{\text{SSM}}_{\text{LZ}} = \begin{pmatrix} 0.073 & 0.124 & 0.056 \\ 0.123 & 0.292 & 0.124 \\ 0.056 & 0.127 & 0.281 \end{pmatrix} \tag{7.2}$$

$$\Sigma^{\text{Exp}}_{\text{Bx}} = \begin{pmatrix} \sigma^2_{^7\text{Be}} & 0 & 0 \\ 0 & \sigma^2_{^8\text{B}} & 0 \\ 0 & 0 & \sigma^2_{\text{CNO}} \end{pmatrix} \tag{7.3}$$

I employed a test statistics employed t which considers the LZ-SSM and HZ-SSM as two alternative hypothesis:

$$t = -2\log\left[\mathcal{L}(\text{HZ})/\mathcal{L}(\text{LZ})\right] = \chi^2(\text{HZ}) - \chi^2(\text{LZ}) \tag{7.4}$$

The t probability distributions are built by means of a toy Monte Carlo approach: I have randomly generated 10^6 pseudo-datasets, i.e. triplets of $(\Phi_{^7\text{Be}}, \Phi_{^8\text{B}}, \Phi_{\text{CNO}})$, by sampling a 3-dimensional gaussian centered on the HZ-SSM or LZ-SSM predictions, and having a width which includes both theoretical and experimental errors and correlations. The distributions of the three fluxes are shown

[2]The covariance matrix components are $\Sigma_{ij} = \sigma_i \rho_{ij} \sigma_j$, where σ_i is the i-th neutrino flux and ρ_{ij} is the covariance of the i-th and the j-th neutrino fluxes couple.

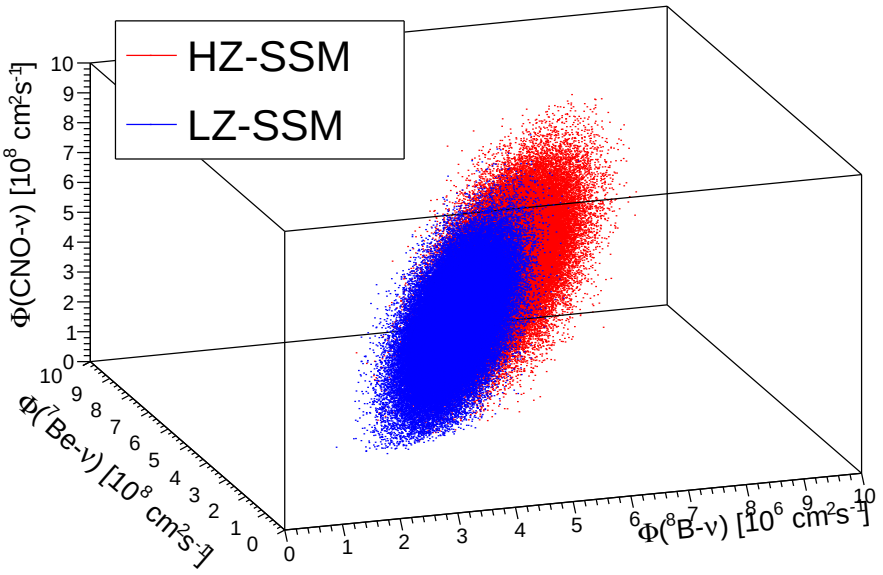


Figure 7.8: Frequentist hypothesis test for the metallicity discrimination: 3D distribution of the $(\Phi_{^7\mathrm{Be}}, \Phi_{^8\mathrm{B}}, \Phi_{\mathrm{CNO}})$ fluxes, obtained by means of pseudo-dataset . The red and the blue distributions are related to the HZ-SSM and LZ-SSM hypotheses respectively.

in Fig. 7.8, in red and blue for HZ-SSM and LZ-SSM hypothesis respectively. For each of the generated triplets, I have evaluated t according to Eq. (7.4).

The distributions of the test statistics t (Eq. (7.4)) are shown in Fig. 7.9 for the two HZ-SSM and LZ-SSM metallicity hypothesis (red or blue respectively). The value of t corresponding to the Borexino results for ^{7}Be-ν, ^{8}B-ν and CNO-ν fluxes, i.e. $t_{\mathrm{Bx}} = t(\Phi^{\mathrm{Bx}}_{^7\mathrm{Be}}, \Phi^{\mathrm{Bx}}_{^8\mathrm{B}}, \Phi^{\mathrm{Bx}}_{\mathrm{CNO}}) = -4.69$, is marked by the black dashed line.

Our measurements show a mild preference for the HZ-SSM scenario; I have evaluated the Borexino exclusion power to the LZ-SSM hypothesis by estimating the blue curve area located at the left of the $f(t|\mathrm{HZ})$ median, i.e. $p_{\mathrm{Bx}} = \int_{-\infty}^{t_{\mathrm{Bx}}} f(t|\mathrm{LZ})\mathrm{d}t = 0.015$. This represents the Borexino exclusion power combining the three fluxes measurement: it disfavors the LZ-SSM metallicity hypothesis at a $1 - 0.015 = 98.5\%$ C.L.. This result can be standardized in terms of gaussian standard deviations as[3] $Z = \Phi^{-1}(1 - p_{\mathrm{Bx}}) = 2.1\sigma$. The Borexino exclusion constraint is slighly stronger than the median sensitivity power (median p-value), which is obtained integrating $f(t|\mathrm{LZ})$ up to the median of the $f(t|\mathrm{HZ})$ distribution, i.e. $p_{\mathrm{median}} = \int_{-\infty}^{t_{\mathrm{median-HZ}}} f(t|\mathrm{LZ})\mathrm{d}t = 0.015$, excluding LZ-SSM at a $1 - 0.034 = 96.5\%$ C.L., corresponding to 1.8σ standard deviations.

The Borexino measurement of the CNO-ν rate does not drastically improve the rejection power based only on the ^{7}Be-ν and ^{8}B-ν fluxes, which was found to be where a 1.8σ (see Ref. [39]). This is mainly due to the large statistical error of the extracted $R(\mathrm{CNO}\text{-}\nu)$.

[3] Where Φ is the Gaussian cumulative distribution function.

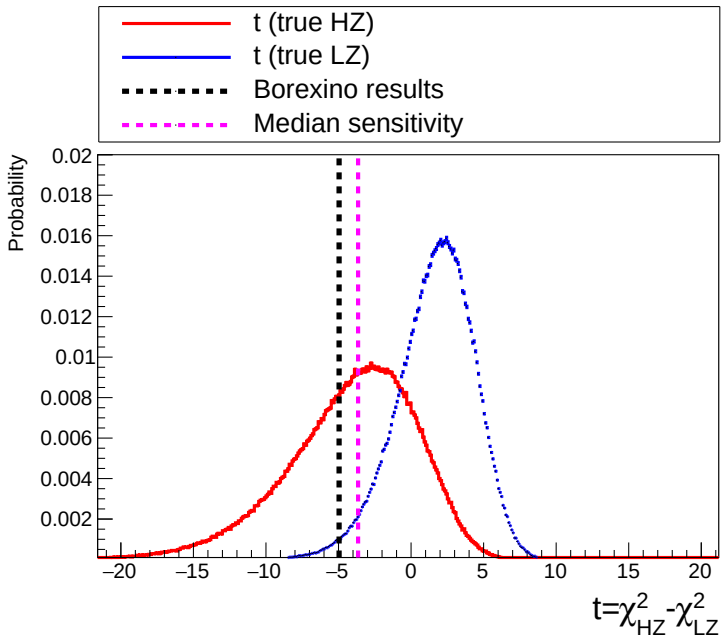


Figure 7.9: Frequentist hypothesis testing for the metallicity discrimination: test statistics t for the HZ-SSM vs LZ-SSM discrimination, combining the information from the fluxes triplet $(\Phi_{^7\mathrm{Be}}, \Phi_{^8\mathrm{B}}, \Phi_{\mathrm{CNO}})$. The red and blue distributions are the Eq. (7.4) test statistics PDFs, evaluated by means of Monte Carlo simulations (HZ-SSM and LZ-SSM in red and blue respectively)s. The black dashed line corresponds to the Borexino results sensitivity power, i.e. $t(\Phi^{\mathrm{Bx}}_{^7\mathrm{Be}}, \Phi^{\mathrm{Bx}}_{^8\mathrm{B}}, \Phi^{\mathrm{Bx}}_{\mathrm{CNO}})$ while the pink dashed line represents the Borexino median discovery power.

Future perspectives

It is interesting to understand whether a more accurate Φ_{CNO} measurement would be able to significantly improve the discrimination power between LZ-SSM and HZ-SSM scenarios. For this reason, I have performed the same frequentist sensitivity study shown above, assuming a more precise CNO-ν determination. I have examined two exemplary cases: $\sigma(\mathrm{CNO}\text{-}\nu) = 0.5\,\mathrm{cpd/100t}$ and the extreme case of $\sigma(\mathrm{CNO}\text{-}\nu) \approx 0.1\,\mathrm{cpd/100t}$.

The t (Eq. (7.4)) probability distributions obtained for the two analyzed cases are shown in Fig. 7.10. This study shows that, if a future Borexino-like experiment will be able of measuring Φ_{CNO} with a 10% precision (i.e. $\sigma(\mathrm{CNO}\text{-}\nu) \approx 0.5\,\mathrm{cpd/100t}$ in Borexino), the resulting median exclusion power (2.2σ) would be only slightly larger than the current one, as displayed in left panel of Fig. 7.10.

Note that the $\sigma(\mathrm{CNO}\text{-}\nu)$ in Borexino is dominated by the error on the ^{210}Bi independent constraint; an uncertainty of $\sigma(\mathrm{CNO}\text{-}\nu) < 0.5\,\mathrm{cpd/100t}$ could be reached only if the convective currents bringing ^{210}Po inside were completely stopped and the clean bubble region expanded to the entire FV.

A further increase of the precision, i.e. 2% (i.e. $\sigma(\mathrm{CNO}\text{-}\nu) \approx 0.1\,\mathrm{cpd/100t}$ in Borexino), would not increase significantly the discrimination, leading to a 2.6σ exclusion power (right panel of Fig. 7.10): this is due to the fact that the theoretical uncertainty would be dominant (0.81 cpd/100t for HZ-SSM and 0.55 cpd/100t for LZ-SSM). No further increase of the discrimination capability would occur until the theoretical error is reduced as well.

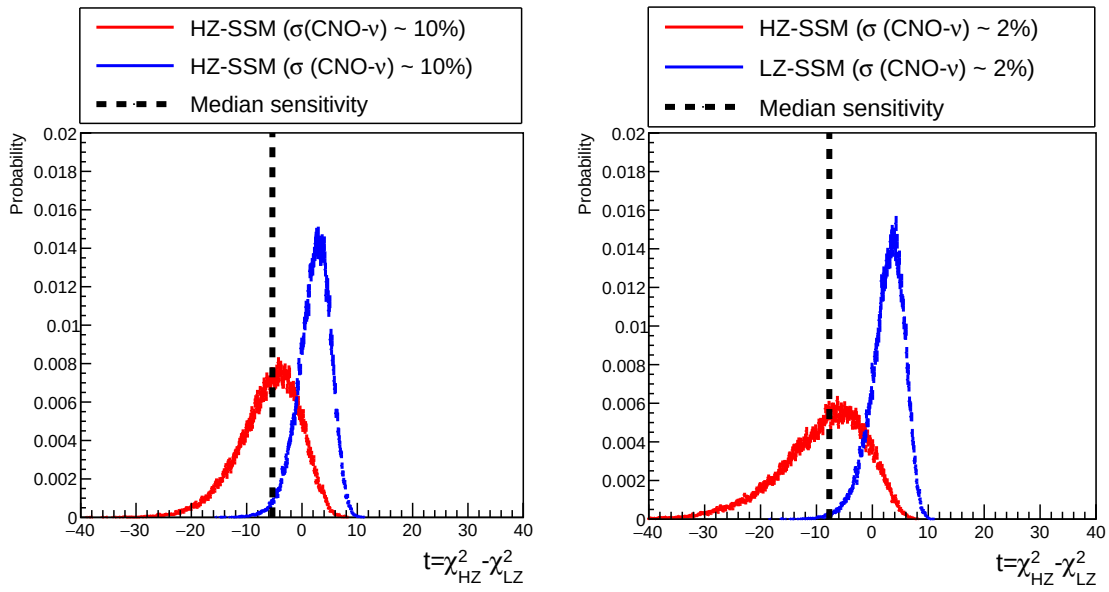


Figure 7.10: Frequentist hypothesis testing for the metallicity discrimination assuming 10% and 2% experimental errors on Φ_{CNO} (left and right panels respectively). The test statistics t for the HZ-SSM vs LZ-SSM discrimination, combining the information from the fluxes triplet $(\Phi_{^7\mathrm{Be}}, \Phi_{^8\mathrm{B}}, \Phi_{\mathrm{CNO}})$, are shown as red and blue distributions, evaluated by means of Monte Carlo simulations (HZ-SSM and LZ-SSM respectively). The black dashed lines correspond to the median exclusion power.

7.4.2 Determination of C+N abundance

In this Section I will show that the measurement of the CNO-ν flux can be used to obtain direct information on the primordial abundance of Carbon and Nitrogen minimizing the uncertainties on many input parameters of the SSM. I will follow the method proposed in Ref. [42, 167], summarizing the basic concept and applying it to the Borexino case as described in Ref. [168].

As discussed in the previous Section, the assumption on metallicity influences significantly the SSM predictions, at least for some of the solar neutrino fluxes. However, the dependence on metallicity is an indirect one: the different metal content affects the solar matter opacity, which in turn has an impact on the temperature profile of the Sun interior, and this ultimately drives the nuclear interaction rates and therefore the neutrino emission rate. This degeneracy between metallicity, opacity and solar temperature makes it difficult to gather information on the input parameters from a measurement of the solar neutrino fluxes.

A similar situation holds for many of the SSM input parameters (those called *environmental parameters*, in Ref.[42]), including age $\tau_{\odot}$, luminosity $\mathcal{L}_{\odot}$, diffusion rate D, together with the already mentioned opacity and metallicity (considering also the heavy elements abundances x_i, where i is an element from O to Fe); the dependence of neutrino fluxes from these parameters collapses into a dependence on the solar core temperature T_c, where T_c is an implicit function of them.

Indeed, it can be shown that for each reaction producing neutrinos in the Sun the main controlling

parameter is T_c. The dependence of the neutrino fluxes from the core temperature is a power-law with a different exponent γ_i for each reaction occuring in the Sun: $\gamma_i < 0$ for the primary pp-ν ($\gamma_{\rm pp} \sim -1$), while it ranges from ~ 10 to ~ 30 for neutrinos produced in the secondary branches of the pp-chain (^{7}Be-ν, ^{8}B-ν, pep-ν) and in the CNO cycle (see Table 7.5).

	pp	^{7}Be	^{8}B	^{13}N	^{15}O
Bahcall and Ulmer 1996 [169]	-1.1	10	24	24.4	27.1
Fuschini and Villante 2019 [170]	-0.8	10.5	23	14.7	19.6

Table 7.5: Values for the power law γ_i coefficients in the $\Phi_i \sim T_c^{\gamma_i}$ relation [169, 170].

The CNO-ν flux (in particular the CN-ν one) has an interesting characteristic with respect to the other fluxes: besides its dependence on the *environmental parameters* which, as in the other cases, collapses into a dependence on T_c, it is also directly dependent on the content of Carbon and Nitrogen which catalize the reaction. It is therefore possible to extract some information on the C and N content in the Sun core by combining the measurement of CNO-ν flux with the one of neutrinos coming from other reactions in the Sun, like for example, ^{8}B neutrinos (whose flux is measured with a $\approx 2\%$ precision by Super-Kamiokande [52]). Both fluxes depends on T_c with different exponent $\gamma_{\rm CN}$ and $\gamma_{^8{\rm B}}$ and, by taking the ratio of their fluxes with a proper weight k, it is possible to nearly cancel out the dependence on T_c (and therefore on the *environmental parameters*) leaving only the dependence on the CN content.

In other words, the ^{8}B-ν flux can be employed as a *thermometer* to measure the Sun core temperature, to be used in the CNO-ν flux measurement to extract the CN content. The weighting factor k is in first approximation related to the ratio $\gamma_{\rm CN}/\gamma_{^8{\rm B}}$, but is determined precisely by minimizing the uncertainty due to the environmental parameters. The procedure to perform this minimization is described in Ref. [168] and leads to $k = 0.735$ assuming the HZ-SSM inputs; it is important to underline that, as expected, the C+N abundance results do not depend from the input metallicity choice.

Following this strategy, the CNO-ν and ^{8}B fluxes weighted ratio is practically independent from T_c and therefore from the environmental parameters, leaving a dependence on x_C and x_N:

$$\begin{aligned}\frac{\Phi_{\rm CN}^{\rm Bx}}{\Phi_{\rm CN}^{\rm SSM}} \Bigg/ \left(\frac{\Phi_{^8{\rm B}}^{\rm SK}}{\Phi_{^8{\rm B}}^{\rm SSM}}\right)^{0.735} &= x_{\rm C}^{0.810} x_{\rm N}^{0.196} D_{\rm C}^{0.183} \times \\ &\times \left[S_{11}^{-0.837} S_{33}^{0.332} S_{34}^{-0.663} S_{17}^{0.735} S_{e7}^{-0.756} S_{114}^{1.003}\right] \times \\ &\times \left[L_\odot^{0.612} k_a^{0.021} k_b^{-0.049} \tau_\odot^{0.289}\right] \times \\ &\times \left[x_{\rm O}^{0.005} x_{\rm Ne}^{-0.004} x_{\rm Mg}^{-0.003} x_{\rm Si}^{0.002} x_{\rm S}^{0.001} x_{\rm Ar}^{0.0001} x_{\rm Fe}^{0.005}\right]\end{aligned} \tag{7.5}$$

where the various terms on the r.h.s. of Eq. (7.5) are the input parameters for solar model construction normalized to the adopted B16-SSM. It is also possible to shown that [170]

$$x_{\rm C}^{\alpha} x_{\rm N}^{1-\alpha} \sim N_{\rm C}^{\alpha} N_{\rm N}^{1-\alpha} \sim N_{\rm C} + N_{\rm N} \tag{7.6}$$

where N_i is the fraction of i-th metal with respect to ^{1}H. Combining Eq. (7.5) with Eq. (7.6), one can now write the proportionality relations between the fluxes and the abundances $x_{\rm C}$ and $x_{\rm N}$,

considering all the other SSM variables as systematic sources:

$$\frac{\Phi_{\rm CNO}^{\rm Bx}}{\Phi_{\rm CNO}^{\rm SSM}} \Bigg/ \left(\frac{\Phi_{^8{\rm B}}^{\rm SK}}{\Phi_{^8{\rm B}}^{\rm SSM}}\right)^{0.735} \sim x_{\rm C}^{0.807} x_{\rm N}^{0.198} = \frac{N_{\rm C}+N_{\rm N}}{N_{\rm C}^{\rm SSM}+N_{\rm N}^{\rm SSM}} \tag{7.7}$$

Inverting Eq. 7.7, it is possible to obtain the sum of C and N abundances $N_{\rm C}+N_{\rm N}$ from the CNO-ν rate measured in this thesis ($R_{\rm CNO}^{\rm Bx}$):

$$\begin{aligned}\frac{N_{\rm C}+N_{\rm N}}{N_{\rm C}^{\rm SSM}+N_{\rm N}^{\rm SSM}} &= \frac{R_{\rm CNO}^{\rm Bx}}{R_{\rm CNO}^{\rm SSM}} \cdot \left(\frac{\Phi_{^8{\rm B}}^{\rm Bx}}{\Phi_{^8{\rm B}}^{\rm SSM}}\right)^{-0.735} \cdot \left[1 \pm \overbrace{2\%}^{^8{\rm B}} \pm \overbrace{0.5\%}^{\rm (env.)} \pm \overbrace{9.3\%}^{\rm (nucl.)} \pm \overbrace{2.7\%}^{\rm (diff)}\right] = \\ &= 1.10 \pm 0.18 \cdot [1 \pm 2\% \pm 0.5\% \pm 9.3\% \pm 2.7\%] = \\ &= 1.10 \pm 0.22\end{aligned} \tag{7.8}$$

It is immediate to note that in Eq. (7.8) the overall uncertainty is dominated by the $R_{\rm CNO}^{\rm Bx}$ error ($\approx 16\%$). Further contributions to the error are obtained by propagating the uncertainties in the environmental, nuclear and diffusion parameters for SSM construction. The dominant error source is given by the S-factors of nuclear reactions, with the largest contributions given by S_{114} (7.3%), S_{34} (3.6%) and S_{17} (3.7%). Note that the error due to the *environmental* inputs of the SSM, included the solar plasma opacity, is relatively small (0.4%), since it has been cancelled out in the weighted ratio of CN and ^{8}B-ν fluxes. From Eq. (7.8), one gets:

$$N_{\rm C+N}^{\rm Bx} = (4.55 \pm 1.33) \cdot 10^{-4} N_{\rm H} \tag{7.9}$$

Metallicity discrimination by using the C+N abundances measurement

The comparison between the C+N abundance in the GS98 (red line) and AGS09 (blue line) catalogues is reported in Fig. 7.11; the Borexino result is shown as the black horizontal bar. As discussed above, the error is dominated by the experimental uncertainty, and being large ($\approx 30\%$) it makes impossible to discriminate between the HZ-SSM and LZ-SSM hypotheses. If the experimental error was reduced to 10%, the theoretical error would start being dominant in the overall error budget (see the green horizontal bar), still limiting the discrimination power.

A discrimination test based on this abundance approach is weaker with respect to the one described in the previous Sec. 7.4.1, since it does not fully exploit the metallicity impact on SSM predictions. However, it is worth to remind the importance of this measurement, which is the first direct determination of the C+N abundance in the Sun core by means of the neutrino fluxes; the advantage of not being dependent on the metallicity inputs and, more generally, on the details of the SSM, guarantees robustness to the measurement.

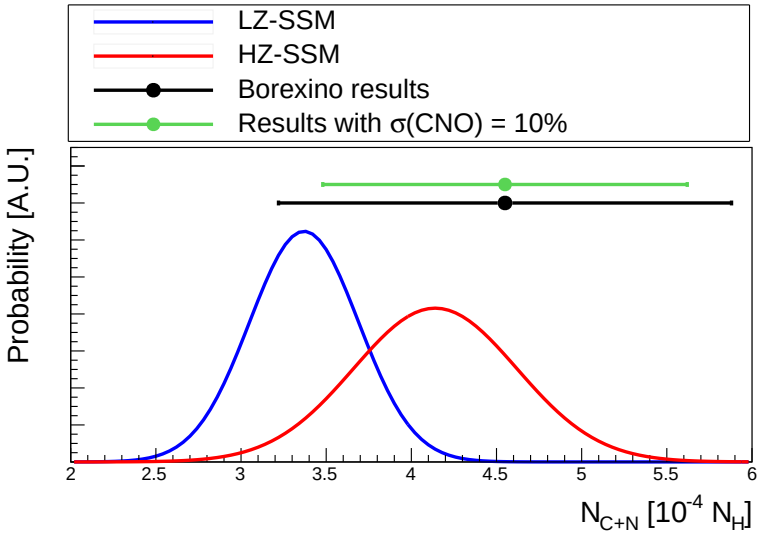


Figure 7.11: Comparison between the C+N abundance reported in the GS98 (red line) and AGS09 (blue line) catalogues; the Borexino result is shown as the green horizontal bar Eq. (7.9)). The abundance determined by a possible next future experiment able to measure the CNO-ν flux measurement with 10% precision is shown as a green horizontal bar.

7.5 Conclusions and perspectives

The analysis of the Borexino Phase-III-TMP dataset described in this Chapter has led to the first indication of the existence of CNO neutrinos. This result has been obtained following both the analytical and Monte Carlo fit approaches to disentangle the CNO-ν signal from the background. A crucial ingredient in the analysis is the constraint to the ^{210}Bi rate in the fit, which was obtained exploiting the ^{210}Bi -^{210}Po link (see Chapter 6).

Including both the statistical and systematic uncertainties, the best estimates for the CNO-ν interaction rate are $R^{\mathrm{MC}}(\mathrm{CNO}\text{-}\nu) = 5.4^{+1.3}_{-1.1}$ cpd/100t for the Monte Carlo fit and $R^{\mathrm{AN}}(\mathrm{CNO}\text{-}\nu) = 6.2^{+2.7}_{-2.6}$ cpd/100t for the analytical fit. The hypothesis of absence of a CNO-ν signal is rejected, by using both the approaches, in the $\sim 4\sigma$ statistical significance range, compatibly with the expectations coming from the sensitivity studies. These results represent the first direct experimental proof of the CNO-ν cycle which, according to the stellar models, is believed to be the primary source of energy for the stars more massive than the Sun.

Furthermore, I have exploited this measurement to infer additional information on solar physics. Combining the CNO-ν interaction rate obtained in this thesis with the previous Borexino ^{7}Be-ν and ^{8}B-ν measurements, I have shown that Borexino alone disfavours the LZ-SSM scenario at 98.5% C.L. (2.2σ). This slightly improves the result obtained by Borexino with the ^{7}Be-ν and ^{8}B-ν results alone. I have also shown that improving the CNO-ν measurement precision would not help much to discriminate between the HZ-SSM and LZ-SSM hypothesis, unless the theoretical error is also reduced.

Exploiting the CNO-ν rate measured in this thesis, I have applied the procedure described in Ref. [42, 167, 168] to extract the sum of Carbon and Nitrogen abundance in the core of the Sun, which is $N^{\mathrm{Bx}}_{\mathrm{C+N}} = (4.55 \pm 1.33) \cdot 10^{-4} N_{\mathrm{H}}$. In spite of the large error, this is the first direct measurement of the C+N primordial abundance and is very robust against details of the Standard Solar Model, since it has been derived minimizing the impact of the so-called *environmental* input parameters.

Chapter 8

SOX: Short distance neutrino Oscillations with BoreXino

The search for sterile neutrinos is still an open question in particle physics. Beyond the three-flavor mixing well-established picture, anomalous disappearances or appearances of neutrino events have been reported in different neutrino oscillation experiments. Well-known examples are LSND and MiniBooNE results, and the evidences of missing $\overline{\nu}_e$ or ν_e in short baseline reactor and radioactive source experiments [98]. All these anomalies could be naturally explained assuming the mixing of the three active flavors with one or two sterile states in the $\Delta m^2_{\text{new}} \approx (0.5 - 2)\,\text{eV}^2$ mass region. However, a tension among the overall results of neutrino oscillation experiments exist, and a part of them do not support the existence of sterile state(s). While earlier CMB data pointed to the confirmation of sterile anomalies, pushing towards the possibility of the existence of more than three neutrino states, recent more accurate data of Planck survey [97] prefer the standard scenario. In a nutshell, further experiments are still needed to clarify the current status, defintely confirming or disproving the sterile neutrino hypotesis.

SOX [156] (Short distance neutrino Oscillations with BoreXino) has been one of the experiments planned to probe the sterile anomalies, and possibly test conclusively the sterile neutrino(s) hypothesis. The SOX data taking was planned to begin in the first half of 2018, lasting $1 - 1.5$ years. The key idea relied on placing an intense artificial ^{144}Ce-^{144}Pr $\overline{\nu}_e$ source close to the Borexino detector, with no alteration to the current detector design. If sterile neutrino in the $\Delta m^2_{\text{new}} \sim \text{eV}^2$ mass range exist, SOX would detect a deficit in the observed flux with respect to the predictions, due to the sterile oscillations. Unfortunately, the SOX project have been stopped before its beginning in February 2018, due to the impossibility of realizing the source with the required characteristics [157]. However, the SOX preparatory work has gone wasted: the preliminary series of sensitivity studies performed, including a wide analysis of the associated systematics, can be used to illustrate the main peculiarities of a short-baseline neutrino experiment for the search of sterile neutrinos, allowing to employ again the acquired expertise in future experiments.

I have been involved in the SOX project during my whole first PhD year. My personal contribution is mainly related to the sensitivity studies in the limit setting configuration, and to the related preparatory work. I have been one of the two main investigators of the SOX sensitivity power, involved in the validation and testing of the SOX sensitivity framework. A large part of my work is related to the study of systematic effects impact on SOX exclusion power, extending the sensitivity framework as necessary. I was also in charge of the Monte Carlo production of $\overline{\nu}_e$ massive simulations

necessary for the events PDF production to be given as input for the sensitivity framework. These simulations have been also largely employed for the study of the detection efficiency dependence on IV distance.

This chapter is organized as follows. The SOX project design and main features are reviewed in Sec. 8.2. The $\overline{\nu}_e$ event selection cuts are discussed in Sec. 8.3. The basic ideas about a short baseline sensitivity analysis are overviewed in Sec. 8.4. An exemplary strategy for the calculation of the total number of $\overline{\nu}_e$ detected during the SOX data taking is discussed in Sec. 8.5. The technical tools and procedures for the sensitivity analysis are described in detail in Sec. 8.6. The analysis of systematic effects is reported in Sec. 8.8 and Sec. 8.9, regarding the source-related and the detector-related systematic effects impact on the SOX exclusion power. Conclusions will be outlined in Sec. 8.10.

8.2 Main features

The SOX project has been a proposed short-baseline neutrino oscillation experiment whose main goal was the probing of sterile anomalies at the $\Delta m^2 \sim E/L \sim 1\,\mathrm{eV}^2$ scale. The Borexino detector is an excellent candidate for hosting a sterile neutrino experiment, thanks to the extremely low background contaminations, the large active mass, the high position and energy reconstruction precision and the expertise achieved in ten years data taking both for geo-neutrinos and solar neutrino physics. The original data-taking schedule was scheduled to start a $1-1.5\,$ys data taking starting from the first half of 2018.

One of the SOX advantages is the possibility to perform the experiment without changes to the current Borexino layout, avoiding a consistent part of the long sequence of preparatory studies. According to the SOX design, a 100 kCi-150 kCi ^{144}Ce–^{144}Pr $\overline{\nu}_e$ source was planned to be placed beneath the Borexino detector (see Fig. 8.2), thanks to a cubical pit accessible through a small squared tunnel (Fig. 8.1) that was built at the time of construction for the future calibration detectors and neutrino magnetic moment measurements. The distance from the $\overline{\nu}_e$ source to the Borexino detector center was 8.51 m. Since the emitted $\overline{\nu}_e$ from the ^{144}Ce–^{144}Pr ranges between $E \sim (1.8-3.0)\,\mathrm{MeV}$, the experiment is particular sensitive to the $\Delta m^2 \sim E/L \sim (10^{-1}-10)\,\mathrm{eV}^2$ sterile state mass range. Basic estimations show that the number of expected events, in absence of sterile neutrino, is found in the order of $\sim (0.7-1.0)\cdot 10^4$ events (see Sec. 8.5), while the background events play a marginal role.

The Russian spent nuclear fuel reprocessing company, Federal State Unitary Enterprise Mayak Production Association, had been identified to be the only facility able to deliver a PBq scale sealed source of ^{144}Ce, which would have been extracted from Kola Nuclear Power Plant in Russia. After several delays in the spent fuel refinement and in the source production, the Mayak Production Association announced the impossibility to produce the required source, marking unavoidably the end of the SOX experiment.

8.2.1 Source requirements

Electronic anti-neutrinos from ^{144}Ce-^{144}Pr source are efficiently detected in Borexino by means of inverse β decay reaction on protons $\overline{\nu}_e + p \rightarrow e^+ + n$ (see Sec. 2.2). The low radioactivity and the clean tag offered by the space-time coincidence between the prompt e^+ and the subsequent neutron capture make accidental background less than one event per year in the total volume.

Figure 8.1: Picture of the tunnel which leads to the pit under the Borexino detector.

The choice of the ^{144}Ce-^{144}Pr couple as $\overline{\nu}_e$ emitters is due to tight and precise constraints. The $\overline{\nu}_e$ IBD energy threshold is $E_{\text{thr}} = 1.806\,\text{MeV}$, thus a source radioisotope with a high endpoint β decay is required [171]. Since for β decay half-life and endpoint energy are strongly anti-correlated, this requirement leads to look for nuclei with half-lives typically shorter than a day, and prevents the employment of a single radioisotope as antineutrino generator. Looking for a cascading pair of β decaying isotopes, the parent nucleus has to be a rather long-lived isotope (with month or year-scale half-life) and the daughter nucleus a short-lived isotope. The latter one must display a β endpoint energy as high as possible above the IBD threshold, in order to maximize the IBD reaction rate. Several pairs of isotopes meeting these requirements have been identified. Combining these physics-driven requirements with the production feasibility constraint, mainly based on the production rate in nuclear reactor cores, the most suitable $\overline{\nu}_e$ generator for a very short baseline experiment is the ^{144}Ce–^{144}Pr pair. It satisfies all the aforementioned features, since the two half-lives are $\tau_{^{144}\text{Ce}} = 285\,\text{days}$ and $\tau_{^{144}\text{Pr}} = 17\,\text{mins}$, while the high ^{144}Pr Q-value is 2.99 MeV. The simplified decay scheme of ^{144}Ce – ^{144}Pr pair is depicted in Fig. 8.3. The theoretical $\overline{\nu}_e$ ^{144}Ce – ^{144}Pr emitted spectrum, and the predicted $\overline{\nu}_e$ spectra detected by Borexino (obtained folding the differential observable $\overline{\nu}_e$ spectrum from ^{144}Pr with the IBD cross section) without sterile neutrino oscillations are reported respectively in the left and in the right panel of Fig. 8.4.

The ^{144}Pr β-decay is followed, for the 0.7% of decays, by a 2.185 MeV γ ray, which represents an additional source of accidental background. The complete isolation of the $\overline{\nu}_e$ generator from the outside is an indispensable condition for implementation of the project: due to the very high activity of the source, this emission constitutes a major radiation protection concern. A dedicated material shielding is necessary to suppress this background, a massive tungsten-alloy cylinder 19 cm thick.

8.2.2 Calorimetric measurements

The knowledge of the source activity is a crucial parameter for the success of a short-baseline neutrino experiment, since is directly linked to the total number of $\overline{\nu}_e$ expected events; for the SOX

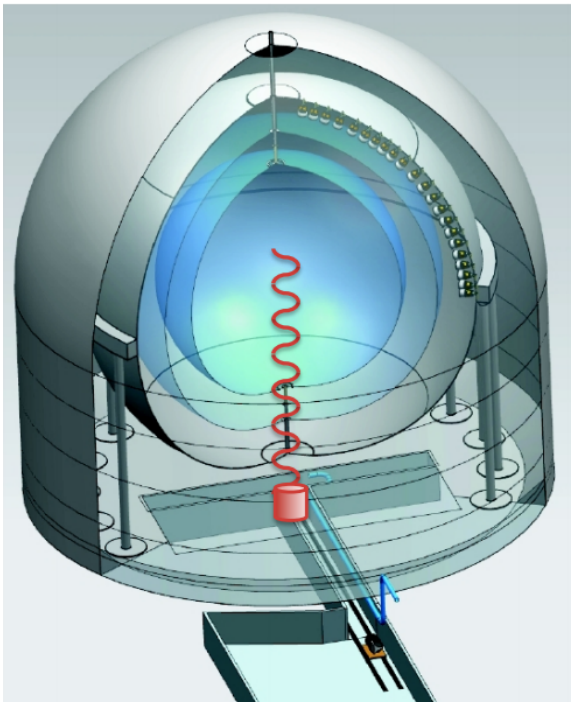


Figure 8.2: Schematic layout of the Borexino detector and of the approximate location of the ^{144}Ce-^{144}Pr $\overline{\nu}_e$ source (represented by the red cylinder under the detector).

experiment, as we will be investigated later (Sec. 8.8), a $\approx 1\%$ precision level is required to achieve the required exclusion power. The highly shielding tungsten-alloy cylinder prevents a direct source activity measurements by β or γ spectroscopic measurements.

Calorimetric measurements offer an attractive alternative to spectroscopic methods. Preliminary analysis has shown that a precision of 1% in the activity might be obtained with a carefully designed and precisely calibrated isothermal calorimeter: the released heat power H is directly linked to the source β decay activity A through the relation

$$A = \frac{H}{\langle S_\beta \rangle} \tag{8.1}$$

where $\langle S_\beta \rangle$ is the mean e^- energy of the β spectrum of ^{144}Ce – ^{144}Pr pair. The calorimeter has been designed to allow the calorimetric measurements both during and after the data taking [172].

Beyond the heat power measurement, Eq. (8.1) shows that great care must be taken in the calculation of the power-to-activity conversion constant, that is given by the β spectrum features. For this reason, a lot of work for both β-spectrum theoretical modelling and spectrometry measurements of ^{144}Ce and ^{144}Pr decays has been carried out throughout the years for the preparation of SOX.

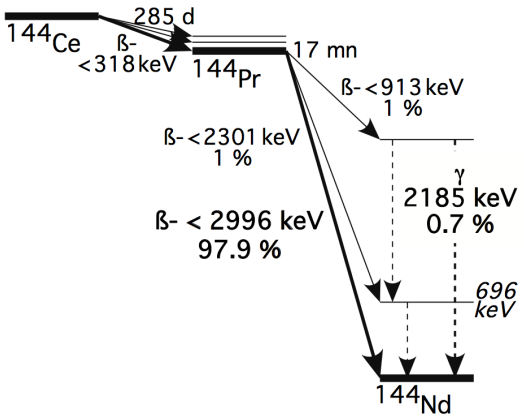


Figure 8.3: Simplified decay scheme of ^{144}Ce–^{144}Pr nuclides pair.

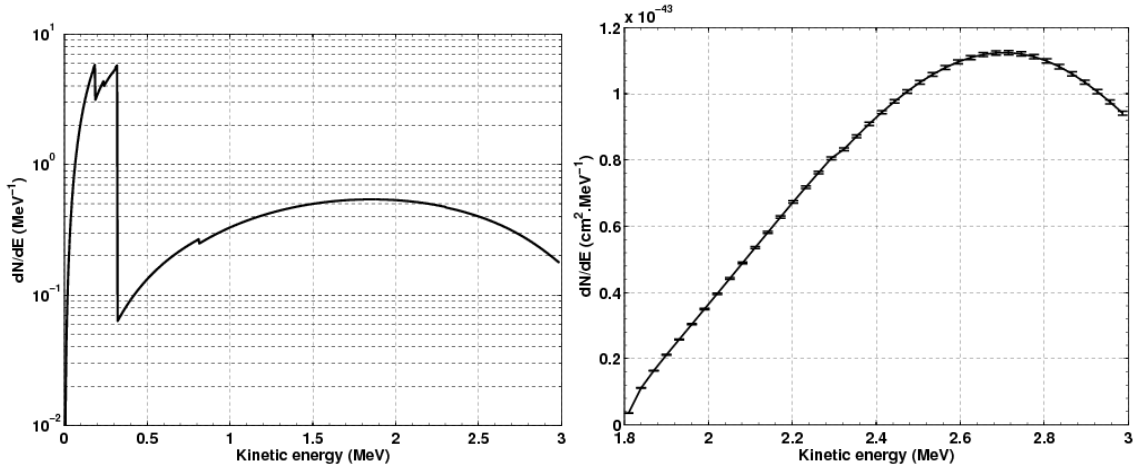


Figure 8.4: Left panel: analytical shape of ^{144}Ce – ^{144}Pr emitted $\overline{\nu}_e$ spectrum. Right panel: prediction of the detected ^{144}Ce – ^{144}Pr $\overline{\nu}_e$ spectrum in Borexino without sterile neutrino oscillations.

8.3 Data selection and backgrounds

Backgrounds faking anti-neutrino interactions can arise from cosmic muons and muon-induced unstable nuclides, from intrinsic contaminations of the scintillator and of the surrounding materials, and from the accidental coincidences of non-correlated events. The starting point for the event selection comes from the expertise acquired by the Borexino collaboration for the geo-neutrino analysis [14]. The list of applied and optimized cuts for the SOX $\overline{\nu}_e$ analysis is the following:

- Veto after muons: a 2 ms veto, i.e. about 8 times the neutron capture time, is applied after all external muons detected by the OD. After the passage of internal muons, able to produce ^{9}Li, ^{8}He, and ^{12}B isotopes, a 300 ms veto for the whole detector is applied (see Sec. 4.1).
- Space correlation between prompt and delayed event. Even if the e^+ and n production is very localized, the distance dR between their detection point is not negligible due to many reasons: the neutron thermalization generates a minimum free path around 30 cm; the γ

interaction is intrinsically not point-like, due to the generated Compton electrons; the position reconstruction (9 cm at 1 MeV significantly smears the reconstruction positions. The toy-MC optimized range for the $\overline{\nu}_e$ selection is $dR < 2$ m.

- Pulse shape discrimination cut for the delayed event, to reject α-like events from radon-correlated background
- Energy cuts: both for prompt and delayed events, the detected energies must match adequately the expected values for e^+ and γ. The minimum threshold for the prompt/delayed charge is 400 p.e. / 500 p.e. (≈ 0.8 MeV / ≈ 1.0 MeV).
- Multiplicity cut: no additional "high-energy" event ($N_{pe} > 400$ p e) event is observed within ± 2 ms around neither the prompt nor the delayed candidate. This cut is meant to suppress the background from undetected muons, as for example neutron-neutron or buffer muon-neutron pairs.

Basic estimations, according to the ^{144}Ce $-^{144}$ Pr source design, show that the expected number of events in case of sterile oscillation absence is found around $\sim (0.7 - 1.0) \cdot 10^4$. Unlike the neutrino analysis, the anti-neutrino one is marginally affected by the presence of backgrounds. Thanks to the extremely low scintillator contaminants and to the clean space-time tagging, the $\overline{\nu}_e$ backgrounds (geo-neutrinos, reactor anti-neutrinos, accidental coincidences) are expected to contribute with $\approx (10 - 15)$ ev/y over the entire scintillator mass [14], corresponding to $\sim (0.11 - 0.18)\%$ of the expected source $\overline{\nu}_e$ events: an harmless value if compared to the 1% uncertainty on the $\overline{\nu}_e$ number of expected events.

For this reason, events can be selected inside a large fiducial volume covering nearly all the scintillator volume, increasing the active mass with respect to solar neutrinos analysis. The FV is usually defined by the scintillator mass located at a distance from IV (d_{IV}) higher than a user-selected value. This choice exclude the scintillator region with the highest detection efficiency uncertainty, which otherwise would reflect heavily on the uncertainty on the expected number of detected events and thus on the exclusion contour power. The delayed tagging efficiency indeed dramatically drops approaching to the IV, due to the possible spill-out of the wandering neutron or of the de-excitation γ. In the same time, these outernmost regions are the less investigated ones by the calibration campaigns, due to priority and technical motivations. This topic will be addressed later on in Sec. 8.9.4.

8.4 Sensitivity studies

The SOX project, as all the other short-baseline neutrino experiments, aims at providing a decisive proof or disproof of the sterile neutrino hypotesis. The main goal of the sensitivity studies is to test the experiment sensitivity to sterile neutrino hypothesis as a function of the mixing parameters. From now on, a single sterile state will be hypotesized, and active-sterile mixing parameters are named $\sin^2 2\theta_{14} = \sin^2 2\theta$ and $\Delta m^2_{14} = \Delta m^2$.

From the statistical point of view, the sensitivity analysis is a hypotesis testing, which can be performed according two different null hypotesis. If SOX would observe clear and unambiguous signatures supporting the existence of the fourth neutrino state, the associated mixing parameters $\sin^2 2\theta$ and Δm^2 could be measured: to claim this evidence, the experiment assumes as null

hypotesis the non-existence of sterile neutrinos, i.e., $H_{00} = \{\sin^2 2\theta = 0\,, \Delta m^2 = 0\}$, and is able to reject it to a sufficiently high confidence level (C.L.). This is named a *discovery analysis*.

The complementary scenario is the *limit setting analysis*. If no evidence of a signal is found, i.e. the discovery analysis fails, SOX data will be used to constrain a part of the mixing parameter $\sin^2 2\theta - \Delta m^2$ space, testing the $H_{XY} = \{\sin^2 2\theta = 0\,, \Delta m^2 = 0\}$ hypotesis instead. The work I have carried out and reported on this thesis focuses entirely on this latter case. The main related goal is the study of the excluded region extension, a key feature of a short baseline experiment, since it represents its hypotesis testing power.

All the short-baseline experiment make use of a binned likelihood function as a starting point for the statistical analysis. We assume that the experiment is able to measure both the reconstructed distance from source L and the reconstructed energy E: the $L \times E$ space is consequently splitted in a $N_E \times N_L$ grid. The 2D likelihood can be written as:

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2) = \prod_{ij}^{N_E \times N_L} \text{Poiss}(N_{ij}^{\text{obs}} | N_{ij}^{\text{exp},s}(\sin^2 2\theta, \Delta m^2) + N_{ij}^{\text{exp},b}) \tag{8.2}$$

$$= \prod_{ij}^{N_E \times N_L} \text{Poiss}(N_{ij}^{\text{obs}} | N^s \cdot \text{PDF}_{ij}^s(\sin^2 2\theta, \Delta m^2) + N^b \text{PDF}_{ij}^b) \tag{8.3}$$

$$N_{ij}^{\text{exp},s} = N^s \cdot \text{PDF}_{ij}^s(\sin^2 2\theta, \Delta m^2) \qquad N_{ij}^{\text{exp},b} = N^b \text{PDF}_{ij}^b \tag{8.4}$$

N_{ij}^{obs} and N_{ij}^{exp} and respectively the numbers of detected and expected (signal or background) events for the ij bin. PDF_{ij}^s and PDF_{ij}^b are the probability density functions for signal or background evaluated in the ij bin; $P(n|\lambda)$ is the Poisson probability for detecting n events given an expectation value λ. In the following, the main ingredients for the calculation of Eq. (8.6) will be introduced, i.e. the calculation of total number of events N^s and the generation of two-dimensional $\text{PDF}(L, E)$.

N^s and N^b are the total expected events for the source $\overline{\nu}_e$ signal and for the background contaminations. They need to be treated as nuisance parameters, since their estimations usually come with a not-negligible statistical uncertainty, and are included in the likelihood construction through a multiplicative gaussian penalty, centered in the expected values N_0^s and N_0^b:

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s, N^b) = \prod_{ij}^{L_{\text{rec}}, E_{\text{rec}}} \text{Poiss}(N_{ij}^{\text{obs}} | N^s \cdot \text{PDF}_{ij}^s(\sin^2 2\theta, \Delta m^2) + N^b \text{PDF}_{ij}^b) \cdot \tag{8.5}$$

$$\cdot \text{Gauss}\left(N^s, N_0^s, \sigma_N\right) \cdot \text{Gauss}\left(N^b, N_0^b, \sigma_{N^b}\right) \tag{8.6}$$

In the most complete case, additional systematic uncertainties can be accounted through extra nuisance parameters, as done for N^s and N^b. These are included to balance for those model parameters not known with sufficient precision. Systematics can be originated by a wide list of sources, ranging from the detector response to the source spectral features. This topic will be addressed extensively in Sec. 8.8 and Sec. 8.9. For what concerns the SOX analysis, since the signal/background ratio estimations are extremely high ($\sim (500 - 1000)$), from now on we will neglect the background contribution ($N^b = N_0^b = 0$)

The Eq. (8.6) likelihood contains simultaneously the combined information on the detector total counting rate (*rate analysis*) and on the L/E oscillometry (*shape analysis*). A pure rate analysis is performed washing out the L/E profile information, considering a single bin in the $L \times E$ space. The pure shape analysis is realized instead removing the gaussian external constraints, losing the information on N^s.

The employed test statistics for the hypotesis testing is a 2D profile likelihood ratio, built according to the Neyman-Pearson lemma[1]

$$t(\sin^2(2\theta), \Delta m^2; N^s) = -2 \ln \frac{\max_N^s \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s)}{\max \mathcal{L}(\sin^2 2\theta, \Delta m^2); N^s)} = \tag{8.7}$$

$$= \min_{N^s} \left(-2 \ln \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s)\right) - \min \left(-2 \ln \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s)\right) \tag{8.8}$$

The first term of Eq. (8.8) minimizes the likelihood with respect to the nuisance parameters N^s, whereas the second term is the absolute minimum. As a consequence of Wilks theorem [173], which will be always assumed in the following, as the sample size is sufficiently large, the test statistic (8.8) becomes asymptotically χ^2-distributed[2]. For the extraction of a defined confidence level, the $\sin^2 2\theta - \Delta m^2$ values are searched for, where the t-values match the CL-quantile of a χ^2 distribution with two degrees of freedom. In the standard sensitivity plots of sterile neutrino experiments, it is usually shown the set of hypotheses that are rejected with a median confidence level of 95% (i.e. median significance of 2σ)

8.5 Total number of expected events

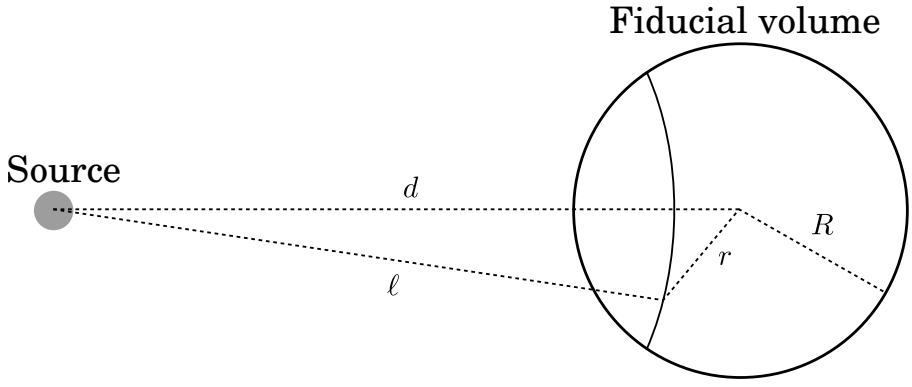


Figure 8.5: Geometric configuration for the $\overline{\nu}_e$ source and the spherical FV.

The estimation of total number of expected events N_0^s is one of the main input parameters of the sensitivity analysis, and can be estimated with sufficient precision following an almost completely analytical approach. First of all, we consider a FV spherical shell of thickness $\mathrm{d}L$ and a given energy

[1] The Neyman-Pearson lemma guarantees that the defined t is most powerful test at significance level α.

[2] The full statistical treatment, which is out of subject for this thesis, construct the test statistics probability distribution on large Monte Carlo datasets, providing the exact coverage.

interval $\mathrm{d}E$. The number of expected $\overline{\nu}_e$ without sterile oscillations in this $L \times E$ infinitesimal bin is given by the product of the $\overline{\nu}_e$ flux $\Phi(L)$, of the targets number $n_{\mathrm{target}}(L)$, of the $\overline{\nu}_e$ cross section $\sigma(E)$, and of the detection efficiency $\mathcal{E}$

$$\frac{\mathrm{d}^2 N_0^s}{\mathrm{d}E\mathrm{d}L} = \Phi(L)\, dn_{\mathrm{target}}(L)\, \mathrm{d}\sigma(E)\, \mathcal{E}(E,\vec{x}) \qquad N_0^s = \int_{L_{\min}}^{L_{\max}} \int_{E_{\mathrm{thr}}}^{Q} \mathrm{d}L\, \mathrm{d}E\, \frac{\mathrm{d}^2 N_0^s}{\mathrm{d}E\mathrm{d}L} \tag{8.9}$$

All the quantities showing up in the previous relation are going to be defined. The flux $\Phi(L)$ at a distance L from the source is connected with the total number of $\overline{\nu}_e$ emitted during the data taking, named $N_{\overline{\nu}}$:

$$\Phi(L) = \frac{N_{\overline{\nu}}}{4\pi L^2} \tag{8.10}$$

Assuming $t_0 = 0$ as the production source time and an exponential decay law for the source activity $A(t)$, with initial value $A(t=0) \equiv A$:

$$N_{\overline{\nu}} = \int_{t_0}^{t_0+t_{\mathrm{exp}}} A(t)\, \mathrm{d}t = \tau A\, \mathrm{e}^{-t_0/\tau} \left(1 - e^{-t_{\mathrm{exp}}/\tau}\right) \qquad A = \frac{H}{\langle S_\beta(E) \rangle} \tag{8.11}$$

where t_0 is the starting data-taking time, t_{exp} is the duration of the data taking and τ is the mean lifetime of the ^{144}Ce radioisotope[3]. According to the signed contract with the Mayak Production Association, the source initial activity range was $A = (100 - 150)\,\mathrm{kCi} = (3.7 - 5.5)\,\mathrm{PBq}$.

Being the source located externally with respect to the detector, only a fraction of emitted $\overline{\nu}_e$ will cross the fiducial volume. This can be taken into account introducing a geometrical factor $S(L)$, that is the spherical surface of radius L crossed by the spherical fiducial volume. In the simplified case of spherical FV (see Fig. 8.5):

$$S(L) = 2\pi L^2 \left(1 - \frac{d^2 - R_0^2 + L^2}{2dL}\right) \tag{8.12}$$

In the most realistic case of a deformed FV which follows the IV evolution in time through the d_{IV} selection cut, the $S(L)$ factor can be computed numerically once the IV shape as a function of (R, θ) couple is provided[4]. Consequently, one gets $n_{\mathrm{target}}(L) = \rho_p\, \mathrm{d}V = \rho_p\, S(L)\, \mathrm{d}L$, where $\rho_p = 5.2976 \cdot 10^{28}\, m^{-3}$ is the numerical proton density of the scintillator, $\mathrm{d}V$ is the shell volume.

From the energy side, the IBD cross-section $\sigma_{\mathrm{IBD}}(E)$ (see the left panel of Fig. 8.4) are:

$$\sigma_{\mathrm{IBD}}(E) = \sigma_0 (E - m_n + m_p) \sqrt{(E - m_n + m_p)^2 - m_e^2} \tag{8.13}$$

where $\sigma_0 = 9.455 \cdot 10^{-48}\,\mathrm{m}^2$, $E_{\mathrm{thr}} = m_n - m_p \approx 1.8\,\mathrm{MeV}$ is the IBD energy threshold, $Q = 2.99\,\mathrm{MeV}$ is the ^{144}Pr emission spectrum Q-value. The total effective cross section $\sigma(E)$ is obtained by multiplying of $\lambda(E)$ with $\sigma_{\mathrm{IBD}}(E)$ (see right panel of Fig. 8.4): $\sigma(E) = \sigma_{\mathrm{IBD}}(E)\, \lambda(E)$

[3] We remind that $\tau_{^{144}\mathrm{Ce}} \gg \tau_{^{144}\mathrm{Pr}}$: the two radioisotope are in secular equilibrium and with excellent approximation their activities coincide.

[4] Since the φ IV shape symmetry is assumed, even in the realistic case the S factor depends again only on L.

The energy spectrum shape S_β, according to β decay theory, can be written as

$$S_\beta(E) = S_\beta^0 \cdot \underbrace{\sqrt{E^2 - m_e^2}\,(E - Q + m_e)^2}_{\text{statistical shape}} \cdot \underbrace{F(E,Z)}_{\text{Fermi function}} \cdot \underbrace{\lambda(W)}_{\text{shape factor}} \qquad \lambda(E) = 1 + aE + \frac{b}{E} + cE^2 \tag{8.14}$$

where Z is the atomic number, m_e is the electron mass, S_β^0 is an overall normalization factor. The purely statistical distribution, extracted through the Fermi golden rule, is modified by the Fermi function $F(Z,E)$, describing the interaction of the e^+ with the Coulomb field of the daughter nucleus, and by a shape factor $\lambda(E)$, accounting for biases due to the so-called allowed approximation (the electron and neutrino wave functions are assumed to be equal to 1) and depends on the forbiddeness of the transition.

It is crucial to underline that, at least for the N_0^s calculations, it is not strictly compulsory to take into account also the position and energy reconstruction resolutions: the smearing redistributes the events in the $L_{\text{rec}} - E_{\text{rec}}$ plane, without lowering significantly N_0^s. Conversely, a N_{ij}^s analytical calculation would be unfeasible because of the complex details of the full detector response, whose accurate reproduction can be performed only by the Monte Carlo simulations.

8.5.1 Tagging efficiency estimations

Following Eq. (8.16), all the terms for N_0^s can be computed analytically with the exception of the detection efficiency $\mathcal{E}(E,\vec{x})$: its spatial dependences cannot be modelled easily, since they are strongly dependent on d_{IV}, and thus on the IV and FV shapes, and on the detector response non-homogeneities. I have contributed to a possible strategy to solve the efficiency issue: the N^s analytical calculation assuming perfect detection efficiency, i.e. $\mathcal{E}(E,\vec{x}) = 1$, can be performed, correcting later on the obtained number with a global efficiency parameter $\mathcal{E}(E,\vec{x}) = \mathcal{E}^{\text{MC}}$ estimated through Monte Carlo studies.

I have generated a 10^6 $\overline{\nu}_e$ events Monte Carlo sample, in which $\overline{\nu}_e$ are uniformly distributed in the IV. The simulation statistics is far exceeding the $\sim 10^4$ expectation, to reduce the statistical uncertainty on the detection efficiency itself. The Monte Carlo sample go through the data selection cuts described in Sec. 8.3, then the global efficiency $\mathcal{E}^{\text{MC}}$ is calculated as

$$\mathcal{E}^{\text{MC}}(E,\vec{x}) = \frac{N_{\text{detected}}(E,\vec{x})}{N_{\text{simulated}}(E,\vec{x})} \implies \mathcal{E}^{\text{MC}} = \frac{N_{\text{detected}}^{\text{total}}}{N_{\text{simulated}}^{\text{total}}} \tag{8.15}$$

Examples of the extracted efficiency values for different event positions are provided in Tab. 8.1. Since the prompt event is extremely localized, the associated tagging efficiency is assumed and crosschecked $\approx 100\%$. The delayed tagging efficiency drop approaching to the IV can be observed in second column data. An example of delayed tagging efficiency trend as a function of d_{IV}, from Monte Carlo simulations, is shown in Fig. 8.6.

The final relation for the N_0^s calculation can now be shown. Combining Eq. (8.5), Eq. (8.10),

Coordinates [m]	Tagging eff.	Cut eff.	Detection eff.
$(0, 0, 0)$	99.49%	100.00%	99.49%
$(0, 0, -1)$	99.50%	100.00%	99.50%
$(0, 0, -2)$	99.51%	100.00%	99.49%
$(0, 0, -3)$	99.40%	98.98%	98.38%
$(0, 0, -4)$	86.34%	76.34%	65.91%

Table 8.1: Delayed event tagging efficiency from Monte Carlo simulations. tagging efficiency, cut efficiency and detection efficiency approaching to the IV on the z-axis.

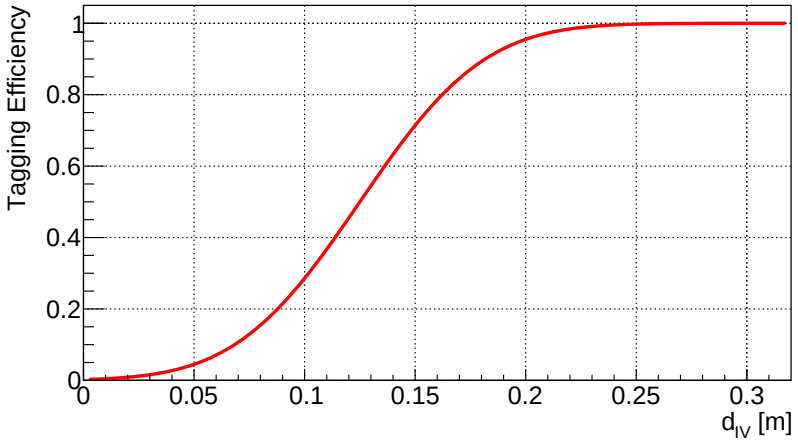


Figure 8.6: Borexino delayed event tagging efficiency as a function of d_{IV}, extracted from Monte Carlo simulations, for scintillator regions very close to the IV ($0\,\mathrm{m} < d_{\mathrm{IV}} \lesssim 0.3\,\mathrm{m}$. Detection efficiency in the detector center is $\approx 99.49\%$.

Eq. (8.12), Eq. (8.13), Eq. (8.15) we get:

$$N_0^s = \mathcal{E}^{\mathrm{MC}} \int_{d-R_0}^{d+R_0} \int_{E_{\mathrm{thr}}}^{Q} \rho_p \, \Phi(L) \, S(L) \, S_\beta(E) \, \sigma_{\mathrm{IBD}}(E) \, \mathrm{d}L \, \mathrm{d}E \tag{8.16}$$

Among the terms appearing in Eq. (8.16), the major sources of bias are given by $\Phi(L)$, through the source activity, and $S_\beta(E)$, through the β shape factor $\lambda(E)$, as it will be investigated in Sec. 8.8.1 and Sec. 8.8.2 respectively. Eq. (8.16) assumes 100% duty cycle and livetime which, if needed, can be easily introduced through overall scaling factors.

8.6 Tools and procedures for the sensitivity analysis

In this section, the main technical steps to perform a sensitivity studies are reviewed. The analysis software flow is shown in Fig. 8.7 and can be summarized as follows:

- **Monte Carlo $\overline{\nu}_e$ simulations** are the starting point for the construction of PDF(L, E) both

for the pseudo dataset and for the model.

- **Processing of the simulated events**: once $\overline{\nu}_e$ events have been simulated, the whole processing chain, made up of offline reconstruction and events filtering, is performed. The output is an Echidna reconstructed RootFile, with the same structure information of a real data one.

- The Echidna RootFiles are then processed, slimmed down and merged to a **light RootFiles**, which includes only the necessary information for the sensitivity studies. These generated light files are the input for the sensitivity analysis framework. The final output file, which includes the basic information on the position and energy distributions, has been used extensively from the collaboration to carry out sensitivity studies.

- **Sensitivity analysis** is performed through a dedicated framework, named `sox-stats`, developed internally to the Borexino/SOX collaboration.

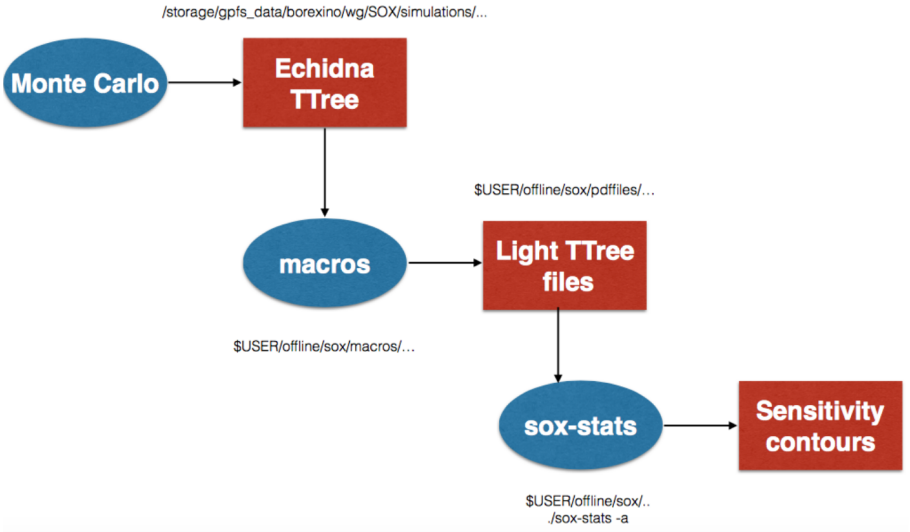


Figure 8.7: Scheme of the analysis software flow.

I was involved in all the steps of the statistical analysis chains: from the process of PDF generation, simulating the high-statistics Monte Carlo datasets, to the building of light input files, to the validation and expansion of the sensitivity framework.

8.6.1 Monte Carlo PDF production

The data sample and each oscillation hypothesis are represented by a 2D histogram with energy (E) and distance (L) axis. The N^{ij} expected events for each $L \times E$ bin cannot be calculated simply integrating numerically Eq. (8.9) over the proper bin limits. The full detector response should be included, and an analytical approach is extremely demanding due to the difficulties of reproducing accurately the position and event reconstructions at large radii. The most efficient way is given by the Monte Carlo simulations.

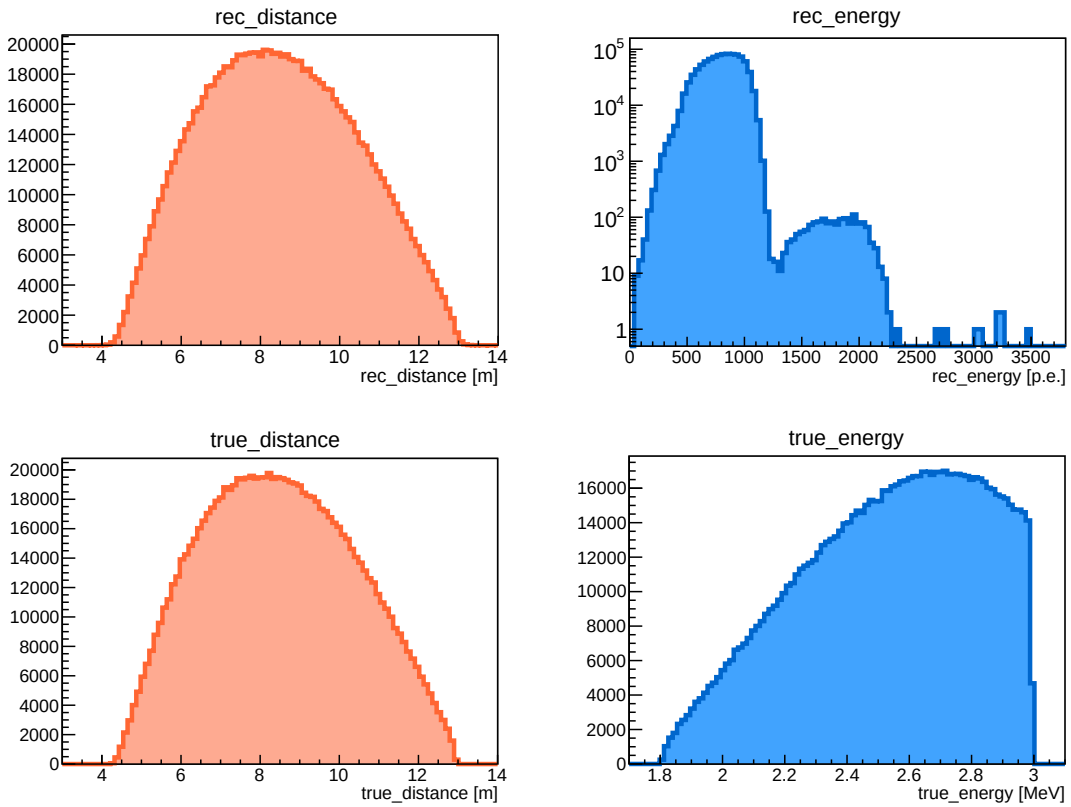


Figure 8.8: Examples of reconstructed and nominal L (orange histograms) and E (blue histograms) histograms of detected $\overline{\nu}_e$ events emitted from a SOX-like pointlike source, respectively on the top and bottom plots. Massive $\approx 10^6$ statistics have been generated, assuming the Borexino detector (live PMTs set, quantum efficiency set, IV shape) according to December 2016 configuration. The reconstructed L histogram is more widely spread with respect to the nominal L one because of the position resolution smearing. The higher energy bump of the reconstructed E histogram is due to the double peak events ($\approx 0.12\%$).

Both nominal and reconstructed 1D L and E PDFs are needed. The nominal variables ones are used to calculate the survival integrated probability (see later), while the reconstructed ones for the number of expected events in a specific bidimensional bin. These PDFs are generated through a Monte Carlo simulation of a SOX-like experiment, in which a number of $\overline{\nu}_e$ N_ν^{MC} are generated from an almost pointlike source located in the pit and are forced to interact inside the active volume. The non-oscillated scenario is assumed: possible sterile oscillations are applied later on, at the sensitivity analysis stage, to save computational time.

I was in charge of the the Monte Carlo massive simulations $\overline{\nu}_e$ samples ($\sim 10^6$ generated events each one) to be given as input of the SOX framework analysis. Examples of nominal and reconstructed L (orange trends) and E (blue trends) histograms of detected $\overline{\nu}_e$ events emitted from a SOX-like pointlike source are shown in Fig. 8.8; the Borexino detector settings (live PMTs set, quantum efficiency set, IV shape) are taken according to December 2016 configuration. The reconstructed distance histogram is more widely spread with respect to the nominal distance one because of the

position resolution smearing. The higher energy bump of the reconstructed energy histogram is due to the double peak events ($\approx 0.12\%$). From this Monte Carlo dataset, the searched PDFs are extracted normalizing the required variables.

8.6.2 sox-stats sensitivity framework

The sensitivity analysis is performed through `sox-stats`, a collection of C++ classes providing interfaces and algorithms for statistical analyses of SOX data. Three main ingredients are needed to perform such an analysis, and each has been implemented as dedicated C++ objects: the generation of the PDFs used to fit the data; the statistical model, including the parameters of the analysis and the definition of the $-2\ln\mathcal{L}$ function; the scanning of the likelihood space and extraction of contours for different probability levels. A scheme of the `sox-stats` framework structure is depicted in Fig. 8.9. I was part of the working team on the `sox-stats` sensitivity framework, playing a role both on the code validation and on its improvement.

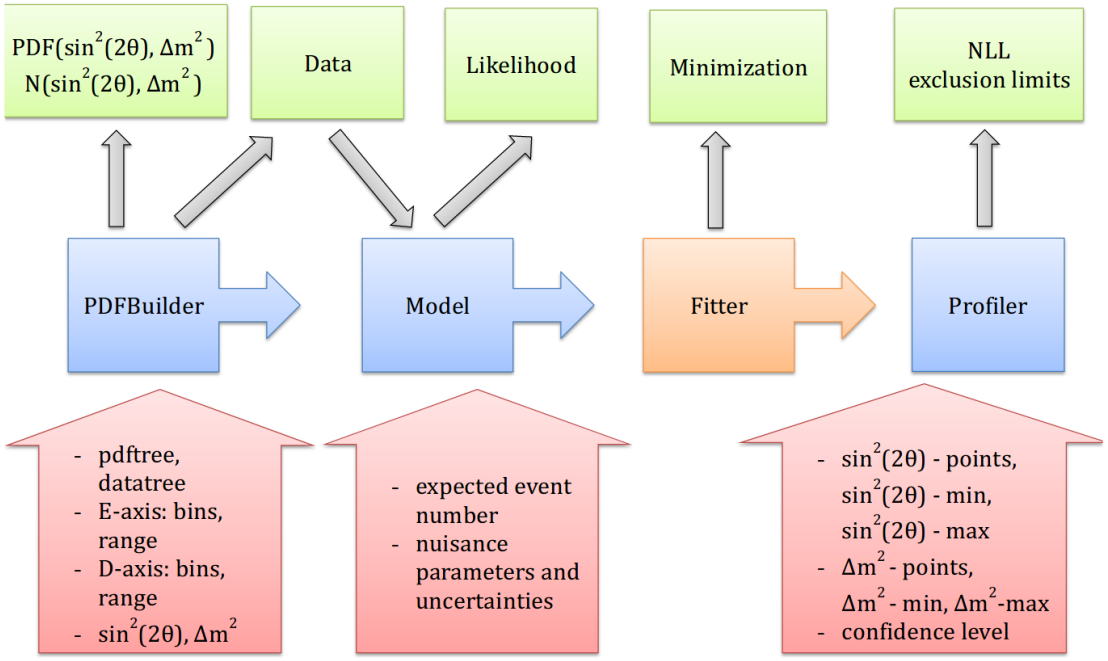


Figure 8.9: Scheme of the `sox-stats` framework structure.

Oscillated PDF generator

Since the limit setting sensitivity scans over the oscillation parameter plane, the PDFs for all the considered hypotesis need to be built. The starting point is given by the high statistic $\overline{\nu}_e$ simulations according to the null hypothesis, described in Sec. 8.6.1, and the associated extracted light tree RootFiles.

By looping over all events of the tree and weighting each event with the corresponding survival probability $P_{\overline{\nu}_e \to \overline{\nu}_e}$, all oscillation hypotheses PDFs can be produced without running the full

simulation again for each point of the parameter plane. Because of the applied $P_{\overline{\nu}_e \to \overline{\nu}_e}$ weight, the histograms cannot be simply normalized to the unity scaling to the total number of generated events, but according to the integrated survival probability $P^{\text{int}}_{\text{surv}}(\sin^2(2\theta), \Delta m^2)$:

$$P^{\text{int}}_{\text{surv}}(\sin^2 2\theta, \Delta m^2) = \frac{1}{N^s_{\text{MC}}} \sum_{k=1}^{N^s_{\text{MC}}} P_{\overline{\nu}_e \to \overline{\nu}_e}(L, E, \sin^2 2\theta, \Delta m^2) \tag{8.17}$$

$P^{\text{int}}_{\text{surv}}(\sin^2 2\theta, \Delta m^2)$ corresponds to the fraction of events expected in SOX data taking with respect to the total number, in case of oscillations with mixing parameters couple $\sin^2 2\theta - \Delta m^2$.

Also the generation of pseudo dataset is performed with Monte Carlo techniques, sampling from the chosen 2D PDF. I implemented the possibility to make use of different PDFs for data and model, allowing to study the impact of a vast records of systematic effects on the SOX exclusion contours.

Two complementary approaches for the PDFs construction can be employed, according to the number of samplings N^s_{MC} during the building stage:

- **Asimov sample:** $N^s_{\text{MC}} \gg N^s$**.** The simulation generates a very high $\overline{\nu}_e$ statistics with respect to the $\sim 10^4$ expected one. When the PDF(L, E) will be generated from the MC sample through the proper normalization, the statistical fluctuations will be almost suppressed, resulting in a smooth and regular profile. The PDF(L, E) median values coincide to the ones obtainable from the averaging of PDFs from multiple experiments, each one with N_0 events. The consequent sensitivity contours extracted from these pdfs represent the median sensitivity values. This sample allows to underline specific physics effects regardless of the statistical fluctuations, which could overlap and hide them.

- **Monte Carlo sample:** $N^s_{\text{MC}} \approx N^s \sim 10^4$. A SOX-expected number of events is generated. In this way, both pdfs and sensitivity contours will be randomly affected by statistical fluctuations, providing the most realistic case. Extracting sensitivity contours many times with Monte Carlo sample, one would obtain the Asimov sample limit and thus the median sensitivity contours. The Monte Carlo sample can be generated also extracting randomly N_0 antineutrino events from the Asimov sample.

Model

In the model the likelihood function, the parameter of interest and the nuisance parameters are defined. N^s is a required input and is defined in the basic model as a nuisance parameter. The basic version of the model likelihood is Eq. (8.6):

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) = \prod_{ij}^{L_{\text{rec}}, E_{\text{rec}}} \text{Poiss}(N^{\text{obs}}_{ij} | N^s_{ij} \cdot \text{PDF}^s_{ij}(\sin^2 2\theta, \Delta m^2)) \cdot \text{Gauss}\,(N^s, N^s_0, \sigma_N) \tag{8.18}$$

Eq. (8.18) can be changed or extended by the user according to the study to be performed, i.e. through gaussian multiplicative terms to account for systematic effects (see Sec. 8.8).

Profiler

A scan over the whole $\sin^2 2\theta - \Delta m^2$ plane is performed and at each point the negative log likelihood (NLL, i.e. $-2\ln\mathcal{L}$) is minimized with respect to the nuisance parameters. After the scan, the NLL is corrected for the absolute minima to obtain the profile likelihood ratio t reported in Eq. (8.8). The Profiler class includes a collection of algorithms for mapping the likelihood function and extracting iso-$(-2\ln\mathcal{L})$ contours of the desired C.L.

8.7 Standard sensitivity plots

A set of standard sensitivity plots have been produced in an expected SOX configuration as a solid benchmark for the future studies.

The full data processing flow described in Sec. 8.3 have been followed. The basic quality cuts (Sec. 8.3) have been applied, and a conservative FV of $d_{\mathrm{IV}} > 0.25\,\mathrm{m}$ is considered. This leads to a detection efficiency of 78.3%. A 99.8% lifetime is estimated, due to the muon veto, and a 99.8% branching ratio of ^{144}Pr main branch, the expected event number varies between 5760 and 8638 for an activity of 100 kCi and 150 kCi, respectively. An uncertainty of 1.5% on the normalization rate (including systematics from source heat power and FV) and an absolute error of $\sigma_b = 0.03$ on the ^{144}Pr spectrum shape factor b are included. The official plot of Fig. 8.10 shows the 95%CL exclusion limits for an activity region of $(100 - 150)\,\mathrm{kCi}$ that is represented by the shaded bands: red / blue / black colors correspond to the rate only / shape only / full analysis. The interpretation of the sensitivity curves features can be divided according to the Δm^2 range, and related to the oscillation lengths:

- Low Δm^2_{14}: $\Delta m^2 \lesssim 0.2\,\mathrm{eV}^2$. The oscillation length $L_{\mathrm{osc}} \sim E/\Delta m^2_{14}$ is wider than the Borexino size. There is no enough space for the oscillations to take place, resulting in a fast power decrease both for rate and shape information.
- Intermediate Δm^2: $0.2\,\mathrm{eV}^2 \lesssim \Delta m^2 \lesssim 2\,\mathrm{eV}^2$. L_{osc} is in the order of the detector size: the entire oscillation profile is sampled and is larger than the experimental resolution, thus the sensitivity is maximized. The shape analysis plays a dominating role in order to cover the experimental anomalies region at 95% confidence level, shown as black closed curves.
- High Δm^2: $\Delta m^2 \gtrsim 2\,\mathrm{eV}^2$. L_{osc} is smaller than the detector resolution on position or energy. The oscillations are thus smeared on an average value, leading to a complete loss of the shape information. The sensitivity is driven by the rate analysis, whose exclusion power is at this point constant for increasing Δm^2.

The total sensitivity is not given by a simple linear combination of the two rate only and shape only contributions, because of their non trivial interplay in the likelihood function. In an optimistic case, the black contour shows that SOX would have been able to exclude the larger part of the anomaly regions, providing a constraining - even if not definitive - claim on the non existence of sterile neutrinos. The signature expected by sterile neutrinos for illustrative Δm^2 values ($0.5\,\mathrm{eV}^2$, $2.0\,\mathrm{eV}^2$, $10.0\,\mathrm{eV}^2$) as a function of L/E is shown in Fig. 8.11.

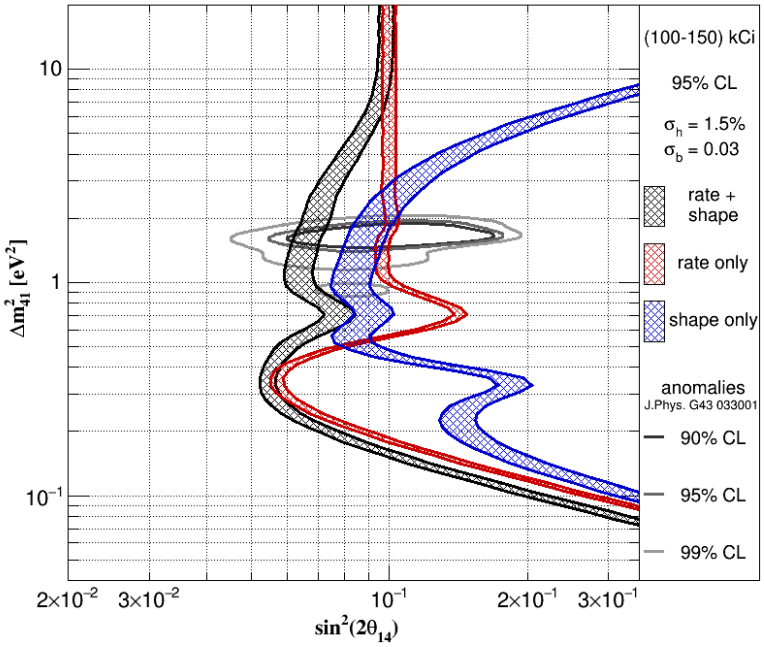


Figure 8.10: Official SOX standard sensitivity plots. Exclusion plot (95% C.L.) in the $\sin^2 2\theta - \Delta m^2$ plane. The rate analysis contour is showed in red, the shape analysis in blue, the full analysis in black. Events are selected through the basic quality cuts and are selected in a $d_{\mathrm{IV}} > 0.25\,\mathrm{m}$ FV. The uncertainty on the source activity is set to be 1.5%. The bands border are related to the lowest expected activity (100 kCi, left borders), and to the highest one (150 kCi, right borders).

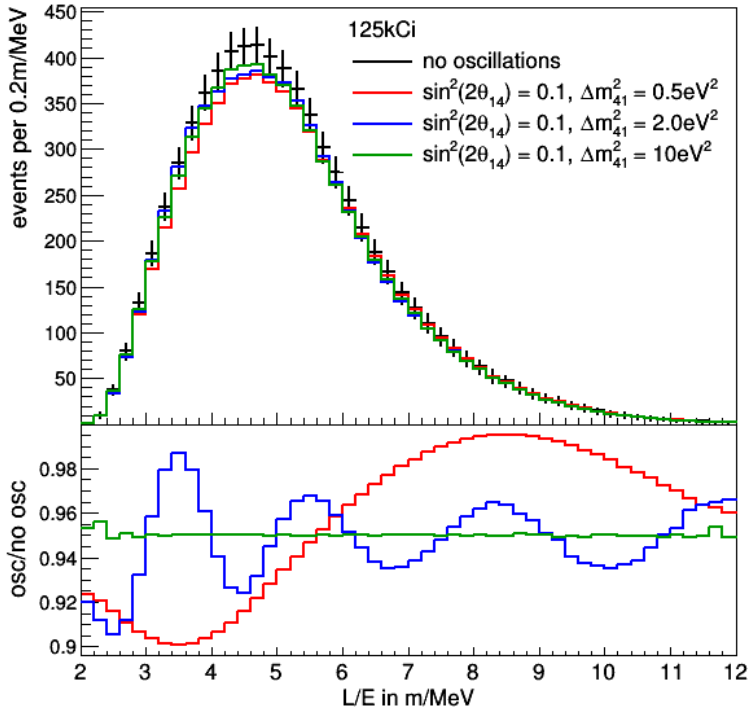


Figure 8.11: Expected SOX signature as a function of L/E: in case of oscillations, with three different values of the oscillation parameters, and in absence of oscillations (top panel). The related ratio of oscillation versus no-oscillation rate for each one of the three cases is shown in bottom panel.

8.8 Source-related systematics

Results about the source-related systematics on the SOX exclusion power are presented in the following. The two main systematics I have investigated are associated to the heat power determination and the β ^{144}Ce emitted spectrum, because of the crucial role they play for the success of the experiment.

8.8.1 Systematics on heat power uncertainty

Calorimetric measurements are crucial for the success of the SOX experiment, because their reliability impact dramatically on the total number of expected events. Since the measured heat power H will be known with not infinite uncertainty, it has to be included in the analysis as a nuisance parameter. This results in a scaling of the total counting rate, without consequences on the events L/E shape. Ad additional gaussian pull-term can be thus included in the likelihood Eq. (8.18):

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) \rightarrow \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s, H) = \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) \cdot \text{Gauss}(H, H_0, \sigma_H) \quad (8.19)$$

The associated systematics can be studied varying the σ_H parameter, and 95% C.L. sensitivity contours are shown in Fig. 8.12, scanning in the $0 \leq \sigma_H \leq +\infty$ range. The gradual degradation to a pure shape analysis can be observed for increasing σ_H values; the sensitivity decreases where no oscillations can be resolved within the detector ($\Delta m^2 \approx 0.3\,\text{eV}^2$ and high $\Delta m^2 > 1\,\text{eV}^2$ range). The contour with $\sigma_H = +\infty$ corresponds to a released H pull term and gives therefore the shape only contours. According to the calorimeter testing, the heat power is measured with a $\sigma_H = 0.5\%$ uncertainty, which holds the experiment exclusion contour covering the allowed regions.

8.8.2 Systematics on ^{144}Ce $-^{144}$ Pr spectrum shape

The β spectrum emitted from the ^{144}Ce $-^{144}$ Pr source is a key input for the antineutrino Monte Carlo generator. The main branch of ^{144}Pr energy distribution, i.e. the spectral shape, defines parameters such as:

- the average heat output per decay of ^{144}Ce $-^{144}$ Pr, which is needed to extract from the planned calorimetric measurements the activity of the source, according to $A = H/\langle S_\beta \rangle$ (Eq. (8.1));
- the expected fraction of e^+ with $E_{e^+} > 1.2\,\text{MeV}$, corresponding to the fraction of $\overline{\nu}_e$ with $E_{\overline{\nu}_e} > E_{\text{thr}} = 1.8\,\text{MeV}$ and, consequently, determines the expected interaction rate inside Borexino.

Consequently, an accurate knowledge of the ^{144}Ce $-^{144}$ Pr is required to avoid biases in rate estimations. Moreover, if the neutrino spectrum shape given as input is sensibly biased, shape oscillations can be mimicked.

The main branch of ^{144}Pr spectrum follows a non unique first forbidden decay that cannnot be fully described by theory. In its simplest form, the spectrum can be described as a function of the shape factor b, which is usually expressed in units of m_e:

$$S_\beta(E, b) = S_\beta(E, b = 0) \cdot \left(1 + \frac{b}{E}\right) \tag{8.20}$$

From the experimental side, the latest independent measures, carried out in 60s, show disagreements up to 10%. The collaboration has been recently performing dedicated measurements of the e^- spectrum achieving a $b_0 = (0.35 \pm 0.03) m_e$ results. To account for the shape factor uncertainty, in the framework analysis b has been added to the list of nuisance parameters through a dedicated Gaussian pull-term:

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) \to \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s, b) = \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) \cdot \text{Gauss}(b, b_0, \sigma_b) \tag{8.21}$$

$$\text{PDF}(E, L, \sin^2 2\theta, \Delta m^2; N^s) \to \text{PDF}(E, L, \sin^2 2\theta, \Delta m^2; N^s, b) \tag{8.22}$$

$$N^s(\sin^2 2\theta, \Delta m^2; N^s) \to N^s(\sin^2 2\theta, \Delta m^2; N^s) = N^s \cdot P^{\text{int}}_{\text{surv}}(\sin^2 2\theta, \Delta m^2) \cdot \text{Sh}(b) \cdot \text{Norm}(b) \cdot \text{En}(b) \tag{8.23}$$

where the three additional terms $\text{Sh}(b)$, $\text{Norm}(b)$ and $\text{En}(b)$ are defined as follows

- $\text{Sh}(b) = \int_{E_{\text{thr}}}^{Q} \frac{N_\nu(E, b)}{N_\nu(E, b_0)}$ accounts for the event number variations due to the shape change;
- $\text{Norm}(b) = \frac{\int N_\nu(E, b_0)}{\int N_\nu(E, b)}$ normalizes the neutrino spectrum;
- $\text{En}(b) = \frac{\langle E(b_0) \rangle}{\langle E(b) \rangle}$ corrects for the spectrum mean energy.

Fig. 8.13 shows sensitivity contours for different σ_b, scanning the $0 \leq \sigma_b < +\infty$ range. The main worsening of sensitivity is due to rate information, due to the presence of the b pull term. Fig. 8.14 shows the interplay of heat and spectral shape uncertainties σ_h and σ_b. The 95% C.L. levels of nouncertainties (black), shape factor released (blue), heat released (red), and shape+heat released (violet) is depicted. From the red and violet curves it can be observed that uncertainty in the spectral shape can mimic spatial oscillations only for $\Delta m^2 < 1\,\text{eV}^2$ range, without affecting sensibly the anomaly $\Delta m^2 \sim 1\,\text{eV}^2$ one.

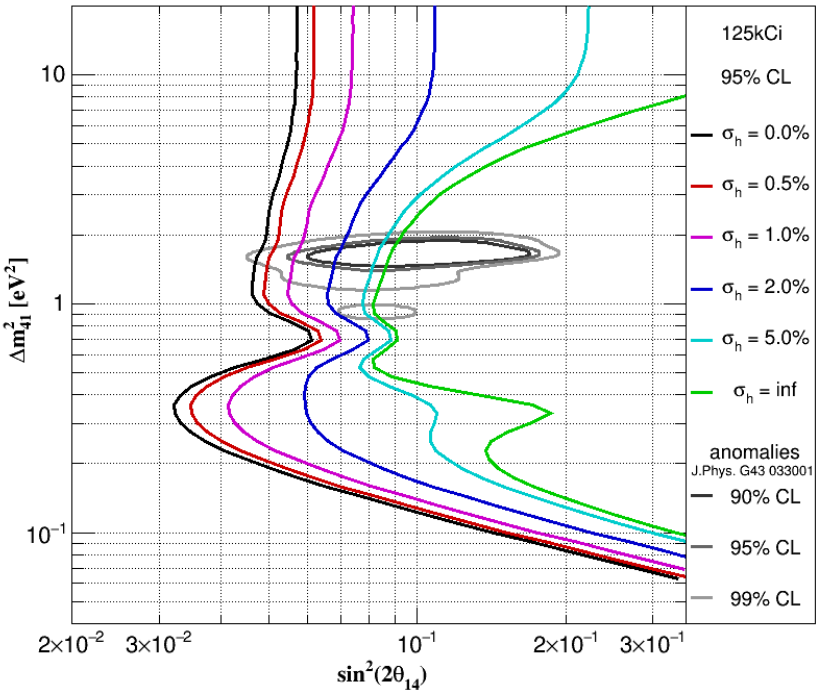


Figure 8.12: Impact of the uncertainty of the calorimetric measurement on the sensitivity. The 95% C.L. exclusion limits are shown for varying uncertainties of the heat power in the $0 \leq \sigma_H \leq +\infty$ range.

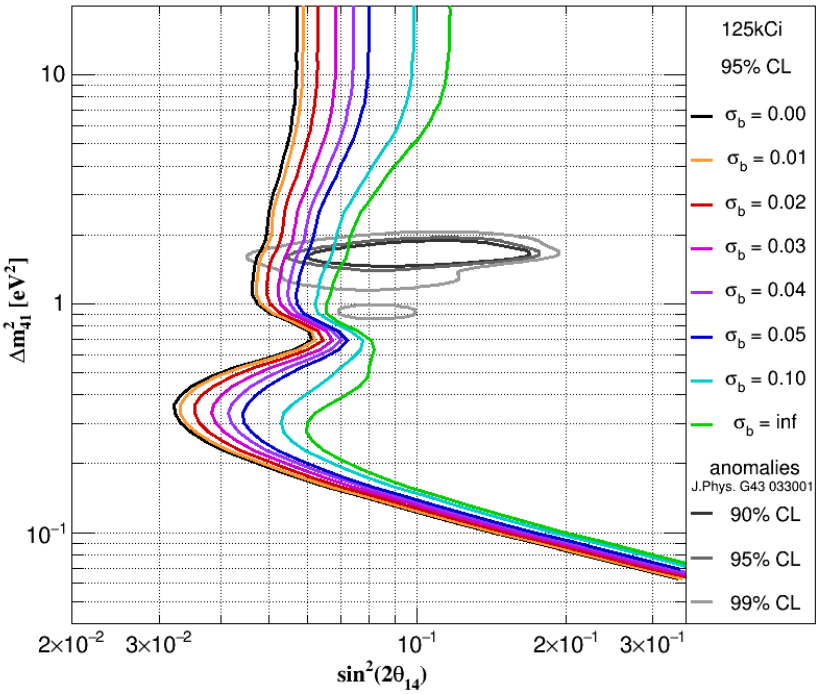


Figure 8.13: Impact on sensitivity of the uncertainty of the b parameter spectral shape of the main branch of ^{144}Pr. The 95% CL levels are shown for different uncertainty values on the shape factor $0 \leq \sigma_b \leq +\infty$ range. The goal of the dedicated spectral measurements is to achieve $\sigma_b < 0.03$.

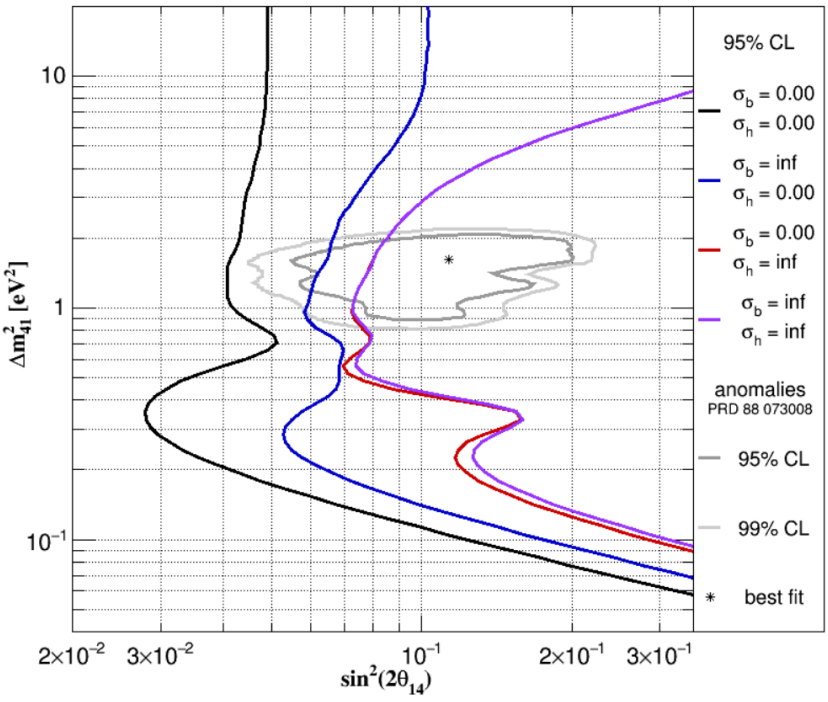


Figure 8.14: Interplay of heat and spectral shape uncertainties σ_h and σ_b. The 95% C.L. levels of nouncertainties (black), shape factor released (blue), heat released (red), and shape+heat released (violet) are shown.

8.9 Detector-related systematics

The knowledge of the detector response, both in energy and in position reconstruction, is naturally integrated in the Monte Carlo simulation code through the tuning based on the 2009 calibration runs. An $^{241}\mathrm{Am} - {}^{9}\mathrm{Be}$ neutron calibration source, dedicated to the anti-neutrino data taking mode, has been employed widely to scan the detector response over a large scintillator volume.

While the energy response scale and its homogeneity in time can be continously monitored through standard candles, which are the ^{210}Po and ^{14}C events, similar independent crosschecks for what concerns the position reconstruction performances are not so straightforward. If biases in the position reconstruction appear for recent data, but did not show up during the 2009 calibrations, these may be not taken into account by Monte Carlo simulations. For instance, a possible source is the deterioration of PMTs performances in time. I have been strongly focused about the topic of position reconstruction bias to sensitivity, since they could alter artificially the events L/E profile and thus the exclusion contours, because of our not sufficient knowledge of the detector response. It is crucial to estimate the type and the size for these bias to mimic oscillations and spoil the sensitivity analysis. My work about the detector-related systematics on the SOX exclusion power is hereafter shown, with a focus on the position reconstruction shift biases.

8.9.1 Impact of position reconstruction resolution

First of all, it is worth to study the position resolution impact on the SOX sensitivity power. Fig. 8.15 shows the change in the shape-only sensitivity contours due to the worsening of position resolution, from the almost ideal $\sigma_L = 1\,\mathrm{cm}$ value to the realistic $\sigma_L = 15\,\mathrm{cm}$ value, to a very worst case scenario $\sigma_L = 25\,\mathrm{cm}$. Rate-only analysis contours are not shown, since they are very marginally affected by σ_L variations, since the number of expected events is nearly unchanged. Studies are performed through a simplified statistical approach, to highlight clearly the main impact of position resolution. Contours differences are mainly found in $\Delta m^2 \gtrsim 1\,\mathrm{eV}^2$ region: since the short wavelength of the oscillations there, the effect of smearing is highly relevant. L/E patterns for this high Δm^2 are more flattened with respect to the low Δm^2 case, where profiles are almost L independent. In the $\Delta m^2 \sim 1\,\mathrm{eV}^2$ mass range of interest for the anomaly range, the effect associated to the σ_L variation is still limited: the expected position resolution worsening due to the decrease of active PMT is not the major concerns for SOX.

8.9.2 Impact of position reconstruction shifts

For what concerns the position resolution biases, we assume that data are affected by some kind of position distortion effects we are not aware of. The standard building of model PDF holds, while the pseudo dataset PDF is artificially altered according to an additional position reconstruction effect. It is convenient to separate the shape and rate analysis to have a wider understanding of the involved effects. We start with the shape information analysis.

- **Random shifts** of x and y coordinate into a $\pm 2\,\mathrm{cm}$ range, which is the reconstruction performance systematic uncertainty band, do not alter the sensitivity power (see Fig. 8.17). Only unpredicted systematic coordinate shifts are potentially worrysome.
- **Shifts of x (or y) reconstructed coordinates in one fixed direction** does not impact

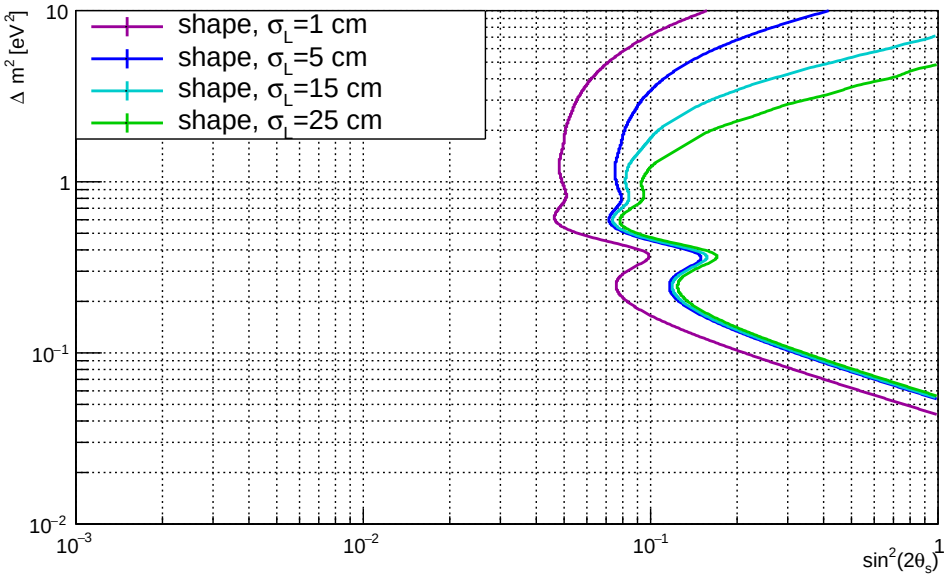


Figure 8.15: Detection resolution σ_L impact on the shape-only sensitivity analysis, performed with a simplified statistical approach. The exclusion contours are shown as a function of vertex detection uncertainty σ_L (purple solid line: 1 cm; blue solid line: 5 cm; light blue solid line: 15 cm; green solid line: 25 cm.)

significantly on the analysis (see left panel of Fig. 8.18). It is possible to justify this behaviour via the L/E ratio variations associated to a Δx shifts. Geometrical considerations about the source position (see Fig. 8.2) allows to calculate the ΔL_x source-detection variation due to a Δx event shift:

$$\Delta L_x = \frac{\mathrm{d}L}{\mathrm{d}x}\Delta x = \frac{x}{L}\Delta x \tag{8.24}$$

which is not negligible only for large x events (i.e. $(x, y, z) = (3, 0, 0)$ and $\Delta x = 5\,\mathrm{cm} \implies \Delta L_x \approx 0.016\,\mathrm{m}$). Since the events are moved in a fixed verse (i.e. ΔL depends on the x coordinate sign), the overall ΔL is approximately negligible. The detector response is reasonably symmetric with respect to x (and y) coordinates, thus also the overall ΔE is unaffected. The combination of these two observation implies almost unchanged L/E ratio and exclusion contours.

- **Shifts of z reconstructed coordinate in one fixed direction**, instead, affects significantly the exclusion plots (see right panel of Fig. 8.18). The ΔL_z source-detection variation due to a Δz event shift can be described as:

$$\Delta L_z = \frac{\mathrm{d}L}{\mathrm{d}z}\Delta z = \frac{z-d}{L}\Delta z \qquad d = 8.51\,\mathrm{m} \tag{8.25}$$

where d is the source-detector center. Unlike the Δx case of Eq. (8.24), ΔL_z is significant for all the detected events thanks to the $z - d$ factor (i.e. $(x, y, z) = (3, 0, 0)$ and $\Delta z = 5\,\mathrm{cm} \implies \Delta L_z \approx 0.031\,\mathrm{m}$). The $\Delta L_z/\Delta L_x$ trend as a function of z, for several x values chosen is plotted in Fig. 8.16: $\Delta L_z/\Delta L_x > 1$ for all the (x, z) couples, pointing at the larger

impact of z shift with respect to the x one. This fact, combined with the asymmetric L and E PDFs with respect to the z coordinate, makes the L/E ratio not invariant with respect to a Δz rigid shift. Consequently, the associated exclusion contours result biased, especially in the intermediate $\Delta m^2 \sim (0.2 - 2)\,\mathrm{eV}^2$ mass region where the shape information plays the major role. A particular z-bias case, whose analysis is motivated by the calibration data results, will be reviewed in the next Sec. 8.9.3.

- **Outward shift of coordinates** (i.e. x-coordinate, in Fig. 8.19). The opposite effect with respect to the previous case is generated: the L/E ratio shows a deficit in the low range, mimicking the oscillation at low Δm^2 and positively biasing the sensitivity contour. The effect is relevant in the $\Delta m^2 \sim 1\,\mathrm{eV}^2$ anomaly region only for $|\Delta x| \gtrsim 5\,\mathrm{cm}$ shifts.

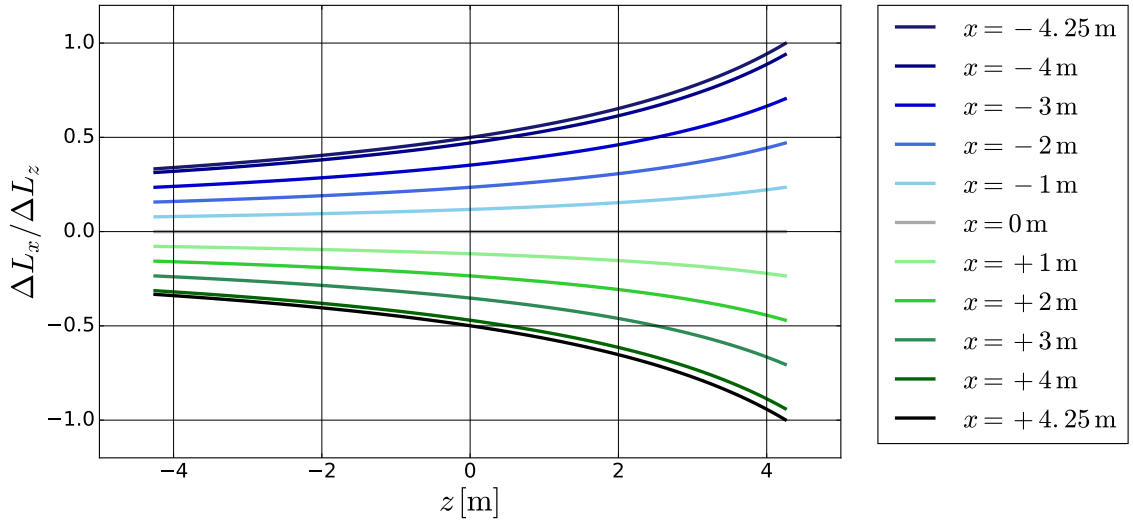


Figure 8.16: Impact of shifts of reconstructed coordinates in one fixed direction: $\Delta L_z/\Delta L_x$ ratio (see Eq. (8.25) and Eq. (8.24)) trend as a function of z, for several x values chosen.

The impact of the previously analyzed effects on the total rate is shown in Fig. 8.20: the ratio between the total number of events in shifted configuration over the unshifted case one, for several coordinate shifts, is reported. Random shifts case is not shown since total rate expected changes is negligible, at the $< 0.05\%$ level. Rigid x shift does not alter significantly the total counting, since the energy response is nearly invariant as a function of small movements of this coordinate. On the opposite case, the z shifts can mimic rate oscillations, since the energy response depend significantly on z coordinate (see Fig. 5.6): the estimated rate variations are 0.7% for a $\Delta z = 2\,\mathrm{cm}$ shift and -1.8% for a $\Delta z = 5\,\mathrm{cm}$ shifts.

This analysis, combined with the following study of z-bias (Sec. 8.9.3), underlines once again the importance of a precise detector knowledge in order to perform a successful sensitivity analysis. Apparently low importance position bias, if still unknown, may possibly reflect into artifacts on the exclusion contours, unvalidating the whole experiment. During the 2009 calibration campaigns, only the well-known z parabolic bias was found, but the detector response is slowly evolving in time, due to the PMTs deterioration: this results in the decreasing of active PMTs and in the same time of their response to scintillation photon (in the form of QE). A new calibration campaign,

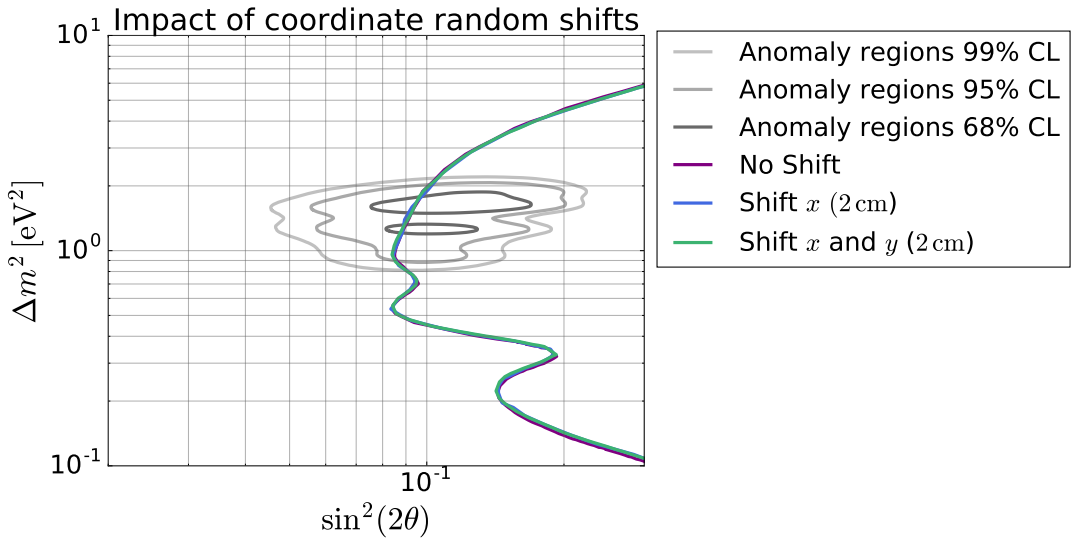


Figure 8.17: Impact of **random shifts** on the SOX exclusion powers. Shape-only exclusion contours in case of data PDFs affected by random 2 cm x (x and y) coordinate shifts, in red solid line (blue solid line).

with a particular focus on the outernmost $R > 3$ m scintillator regions, would be crucial to unearth and identify position reconstruction biases and their sources.

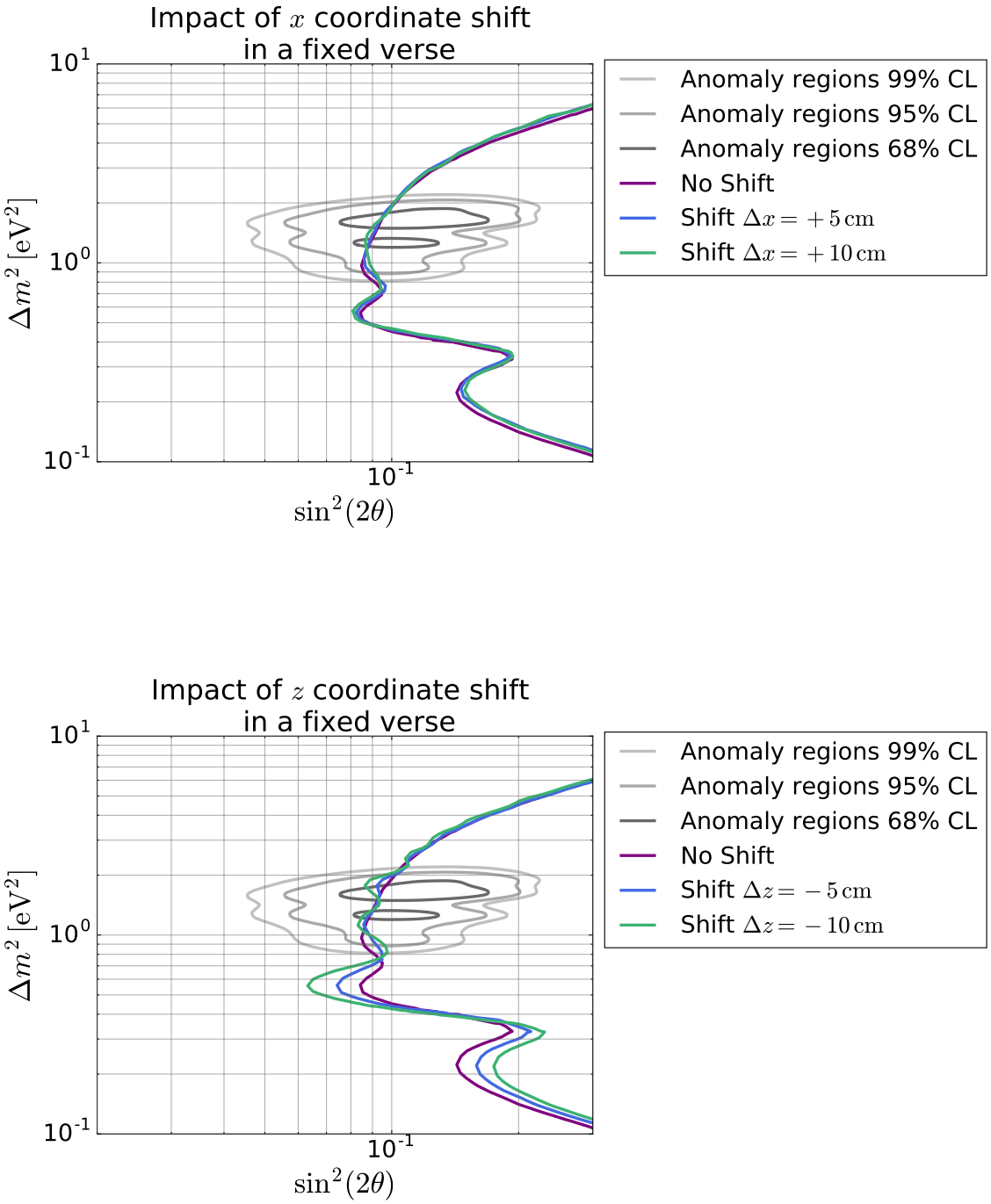


Figure 8.18: Impact of shifts of x or z reconstructed coordinate in one fixed direction on SOX exclusion power. Top (bottom) panel reports the shape-only exclusion contours in case of data PDFs affected by fixed shifts of x (z) coordinate, for $\Delta x = 5\,\mathrm{cm}$ as blue line and $\Delta x = 10\,\mathrm{cm}$ as green line ($\Delta z = -5\,\mathrm{cm}$ as blue line and $\Delta z = -10\,\mathrm{cm}$ as green line), with respect to the non-shifted case (purple line).

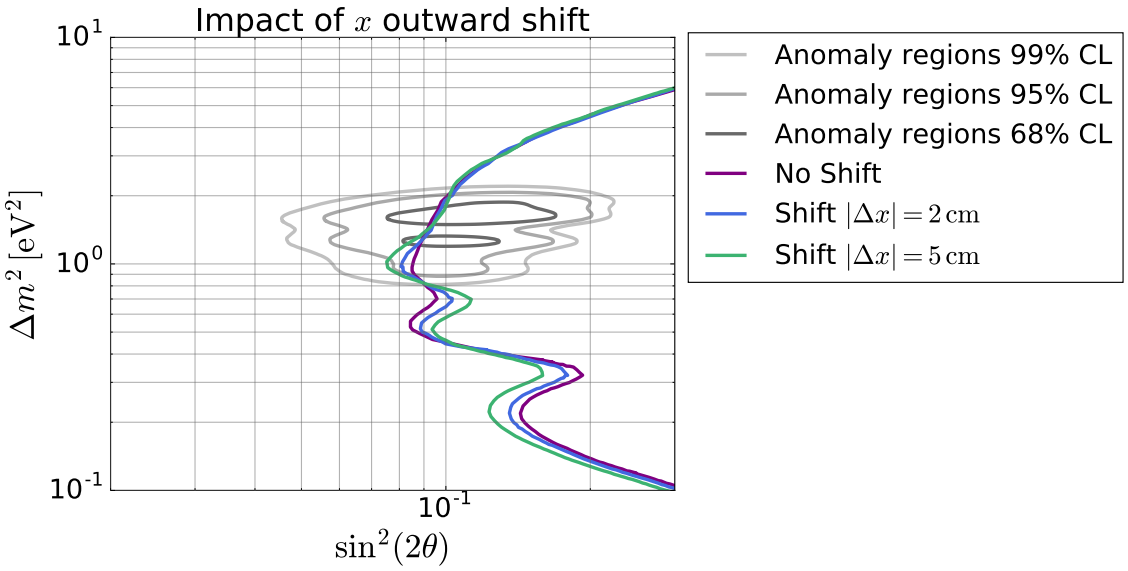


Figure 8.19: Impact of outward x reconstructed coordinate shifts on SOX exclusion power. Shape-only exclusion contours in case of data PDFs affected by x coordinate outward shifts, for $|\Delta x| = 2\,\mathrm{cm}$ as blue line and $|\Delta x| = 5\,\mathrm{cm}$ as green line, with respect to the non-shifted case (purple line).

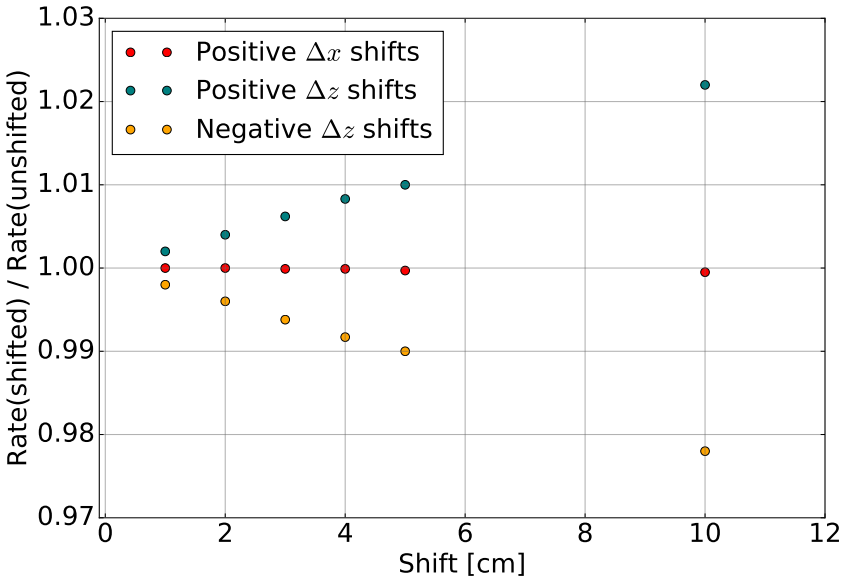


Figure 8.20: Ratio of expected $\overline{\nu}_e$ signal events in case of a coordinate shift applied ($\Delta x > 0$ shift as green points; $\Delta x > 0$ shift as red solid points; $\Delta x > 0$ shift in yellow points) with respect to the unshifted case.

8.9.3 Impact of the z coordinate bias

Calibration campaigns carried out in 2009 (Sec. 5.2.1) highlighted a small position reconstruction bias for the z-coordinate, depending on the event z coordinate itself. The position reconstruction is worst at the detector center, where $\Delta z(|z| < 1\,\mathrm{m}) \approx -3\,\mathrm{cm}$; moving away from it along the z-axis, the position reconstruction slowly gets unbiased. The causes for this effect are still unclear, even if several explanations have been proposed: none of them have provided satisfactory conclusions. For what concerns the solar neutrino analysis, the impact of z-bias to the FV definition is marginal; instead, the SOX analysis can be sensibly influenced by it since all the selected event positions take part in the L PDF building.

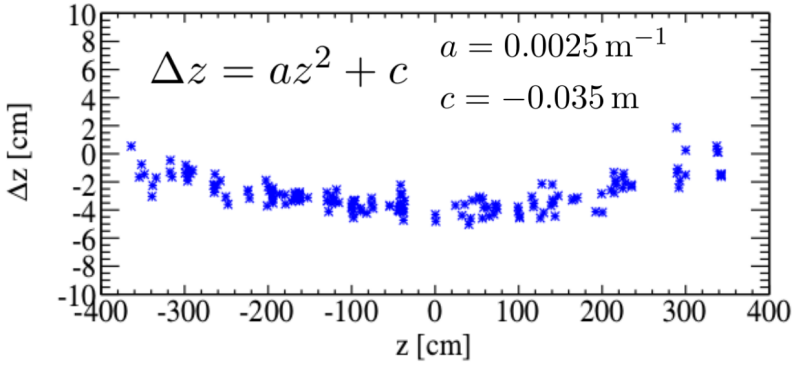


Figure 8.21: z-coordinate reconstructed distortion in Borexino (Δz) plotted as a function of the event z nominal coordinate; the trend is fitted with a parabolic function $\Delta(z) = az^2 + c$. The fitted parameters a_{fit} and c_{fit} values are reported.

Since this effect is not reproduced by the Monte Carlo code, and thus it is not included in the model PDFs, biases of the sensitivity contours may arise if they are left out from the statistical analysis. The impact this z-bias neglection on the SOX exclusion power the can be easily predicted. First of all, the bias trend $\Delta z(z)$ can be extrapolated, in the simplest analytical form, as parabolic and North/South symmetric. A parabolic $\Delta z(z) = az^2 + c$ fit on calibration data is reported in Fig. 8.21, where $a_{\mathrm{fit}} = (2.5 \pm 0.1) \cdot 10^{-3}\,\mathrm{m}^{-1}$ and $c_{\mathrm{fit}} = (-3.49 \pm 0.07) \cdot 10^{-2}\,\mathrm{m}$. We assume that the events dataset suffer from this bias, contrarily to the model PDFs: during the data PDF construction, the z-coordinate of the events is shifted artificially following the $\Delta z(z)$ extracted trend, sampling from a Gaussian distribution which takes into account the parameters uncertainty.

$$z \to z + \mathrm{Gauss}(a_{\mathrm{fit}}, \sigma_a) \cdot z^2 + \mathrm{Gauss}(c_{\mathrm{fit}}, \sigma_c) \tag{8.26}$$

The resulting sensitivity contour, compared to the case without the z-bias included, is shown in Fig. 8.22. One can notice that sensible discrepancies between the two contours exist, causing exclusion power biases in the full $\Delta m^2 \lesssim 3\,\mathrm{eV}^2$ range. The different L PDF included for data and model, due to the z-bias, via the L/E ratio, mimicks oscillations. Differences are observed only for the shape analysis; from the calculation of the total number of expected events, indeed, it can be shown that the total rate in the two cases is almost equivalent. This fact can be observed in the figure because the contours are coincident in the high mass range $\Delta m^2 \gtrsim 3\,\mathrm{eV}^2$, where only the rate information matters (as in Fig. 8.10).

The proper way to deal with this z-bias is to include it in the statistical treatment. Introducing in

the likelihood analysis Eq. (8.18) the uncertainty of the two parameters a and c describing this bias, results showed that a sterile neutrino signature cannot be mimicked.

$$\mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s, a, c) \to \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s, a, c) = \tag{8.27}$$

$$= \mathcal{L}(\sin^2 2\theta, \Delta m^2; N^s) \cdot \text{Gauss}(a, a_0, \sigma_a) \cdot \text{Gauss}(c, c_0, \sigma_c) \tag{8.28}$$

An exemplary profile of the likelihood function for the null hypothesis over the $a - c$ bias parameter plane is depicted in Fig. 8.23. It can be shown that the likelihood exhibits the same minimum for all kind of oscillation hypothesis, corresponding to the couple of injected values $(a_{\text{fit}}, c_{\text{fit}})$: the likelihood fit is able to identify the bias parameters and consequently the sensitivity contours are unchanged with respect to the standard ones. Clearly, from the computational side the inclusion of a and c as additional nuisance parameters stretches sensibly the optimization time.

It is possible to claim that the SOX sensitivity power is not affected by the z-bias if it is known with high precision from calibration data. This latest assumption crucially drives the conclusion of this study: there are no evidence that $\Delta z(z)$ trend, and thus its parabolic parametrization, holds stable in years. A new stage of detector calibrations investigating carefully this bias would be of utmost importance to definitely shed light on this topic.

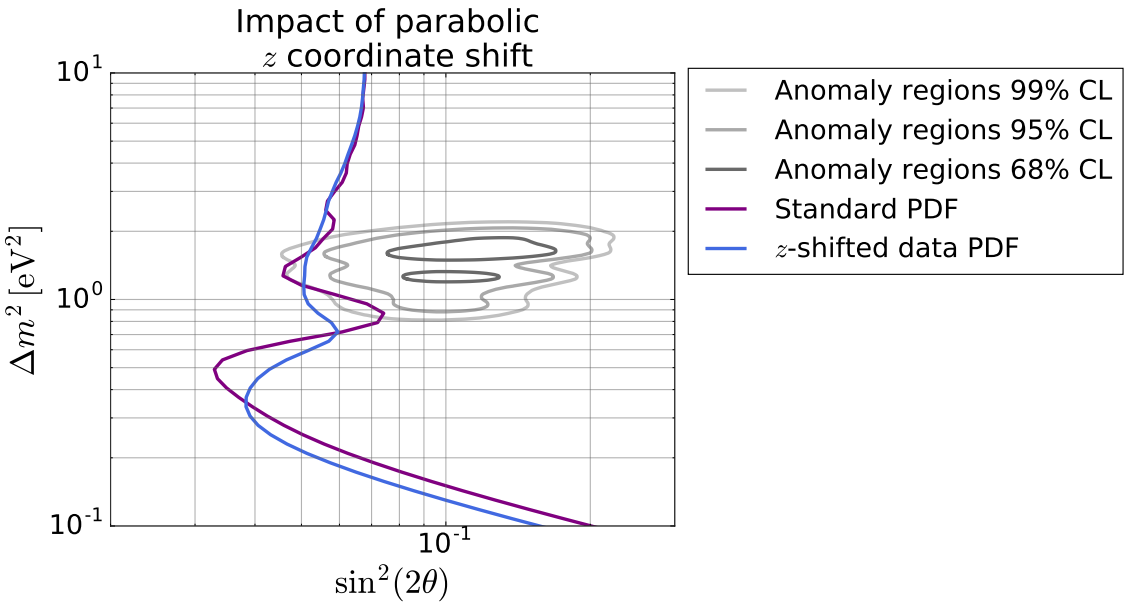


Figure 8.22: Comparison of a 95% C.L. exclusion contour with a z-shifted PDF data one, according to the $\Delta z(z)$ fit shown in Fig. 8.21 (green), compared with a standard configuration one (orange).

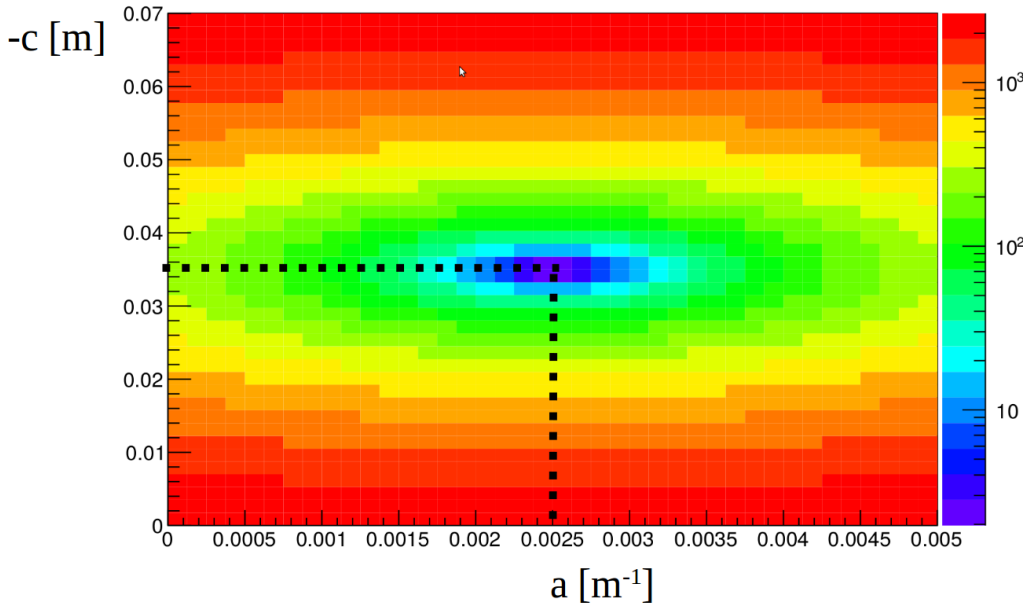


Figure 8.23: Exemplary profile of the likelihood function for the null hypothesis over the 2D bias a and c parameter space. Black dashed lines mark the a and $-c$ values maximizing the likelihood, which are found to correspond to the injected a_{fit} and $-c_{\mathrm{fit}}$ ones.

8.9.4 Influence of vessel uncertainty

To increase the exclusion power, SOX aims to maximize the fiducial volume up to the IV border, to increase the total exposure. This will, in principle, may clash with a series of still poor knowledge of the detector response in the outernmost scintllator regions. One of the most crucial pieces is related to the knowledge of IV shape position. The method currently employed by the Borexino collaboration relies on a radial fit of the large ^{210}Bi contributions which lies on the IV, due to the long-lived ^{210}Pb isotope decaying. The associated statistical errors, extracted directly from the fit, being smaller than 1-2 cm, but inconsistencies with other methods suggest possible systematic errors.

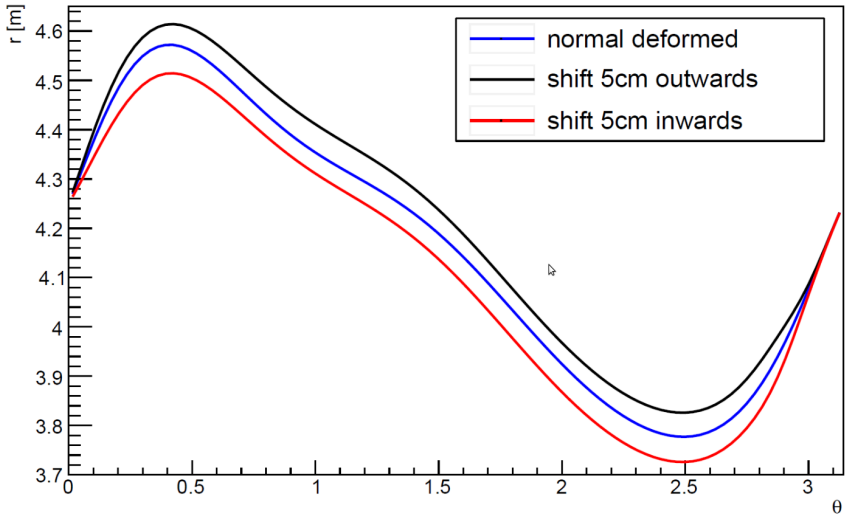


Figure 8.24: The three exemplary IV shapes employed for the study of vessel uncertainty influence on SOX exclusion power. The vessel radius as a function of the azimuthal θ angle, i.e. the $R(\theta)$ trends, are plotted. A real IV shape, from December 2016, determined used the standard algorithm, is shown in blue line. Black and red lines represent IV shapes, respectively shifted 5 cm outwardly or inwardly with respect to the standard (blue) one. The IV is mechanically anchored to the pole endcaps plates: consequently, during the shifting procedure, the $R(\theta)$ at the extreme North and South ($\theta = 0$ and $\theta = \pi$) is forced to the IV design value 4.25 m.

We are interested in studying how the vessel uncertainty impacts on the SOX exclusion power. To be conservative, we assume that the IV shape is misreconstructed with a ± 5 cm systematic error (outwardly or inwardly), not dependent on θ, so that the PDFs for pseudo dataset and model mismatch. Then, sensitivity studies assuming no FV cut are performed. Assuming that the IV is globally mis-determined means that we are considering a the worst case scenario. The three involved IV shapes are shown in Fig. 8.24.

The resulting exclusion contours are shown in Fig. 8.25. The total number of expected events is scaled according to the active volume of the IV configuration, and it is given by 7800 (standard IV), 8300 (outward IV), 7400 (inward IV). "Artifacts" are due to the different IV shapes in data sample and model PDFs: the contours extracted by a biased IV shape are clearly shifted by the mismatch of the rate information. The selection of the events inside a FV selecting events based on their d_{IV} solves the problem of huge rate differences and simultaneously make disappear the artifacts.

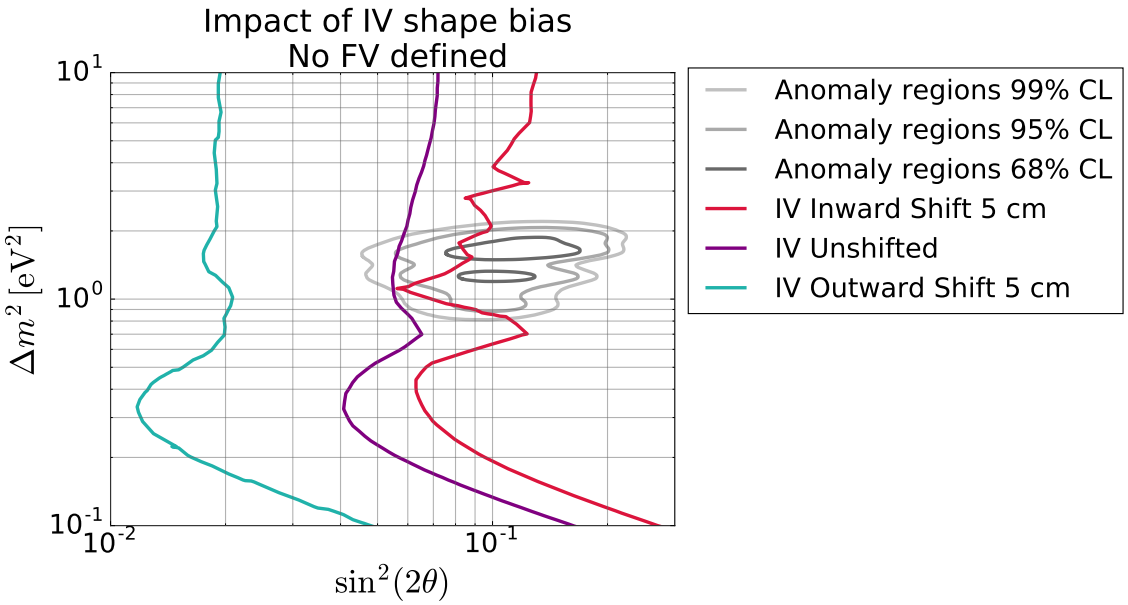


Figure 8.25: SOX exclusion contours (95% C.L.) selecting data in a $d_{\mathrm{IV}} > 0\,\mathrm{cm}$ FV. The IV shape is assumed as standardly deformed for the model, while it is considered differently for the pseudo dataset: standardly deformed (purple contour), 5 cm outwardly deformed (green contour), 5 cm inwardly deformed (red contour)

FVs defined by the $d_{\mathrm{IV}} > 10\,\mathrm{cm}$ or $d_{\mathrm{IV}} > 30\,\mathrm{cm}$, as one can observe in Fig. 8.26 and Fig. 8.27, are sufficient to avoid nearly completely the IV shape issues; simultaneously, in spite of the decreased statistics available, the related sensitivity power allows anyway to cover the 95% C.L. anomaly region.

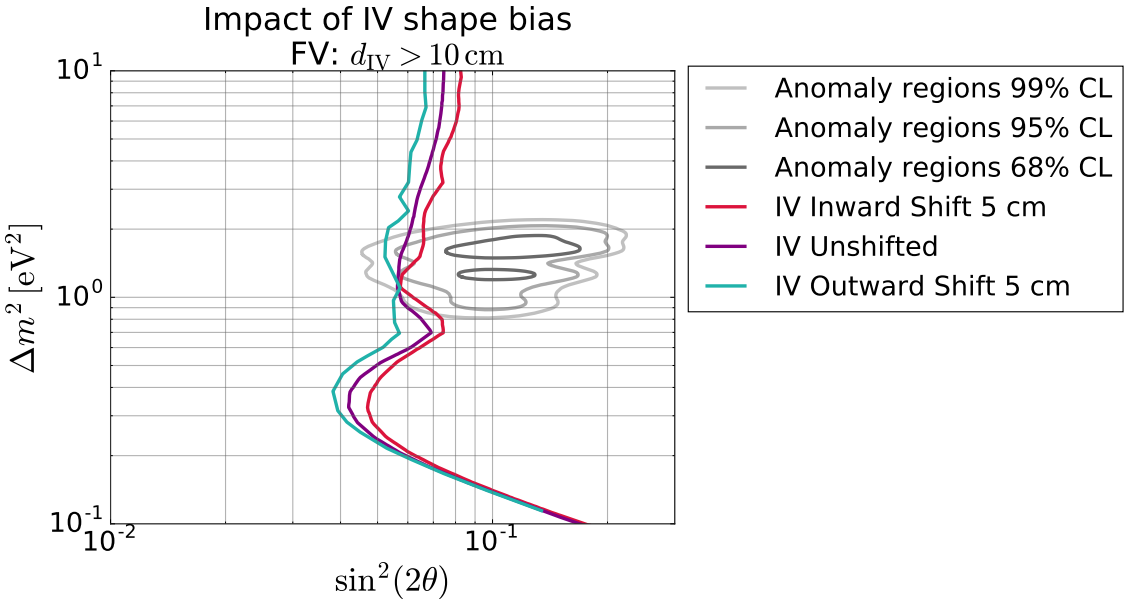


Figure 8.26: SOX exclusion contours (95% C.L.) selecting data in the whole scintillator volume ($d_{\mathrm{IV}} > 10\,\mathrm{cm}$). The IV shape is assumed as standardly deformed for the model, while it is considered differently for the pseudo dataset: standardly deformed (purple contour), 5 cm outwardly deformed (green contour), 5 cm inwardly deformed (red contour).

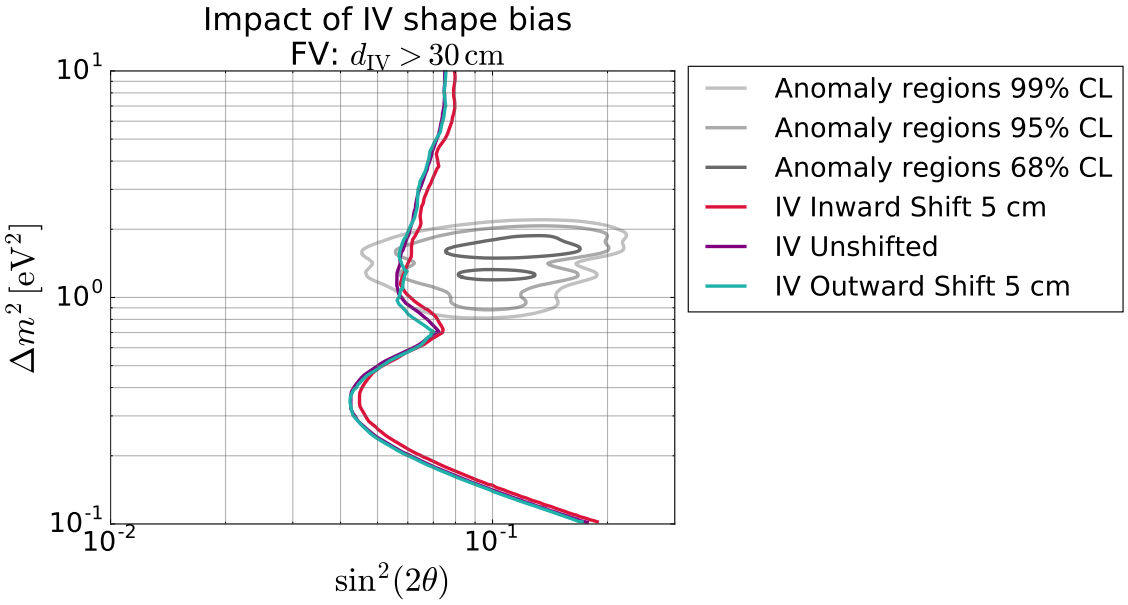


Figure 8.27: SOX exclusion contours (95% C.L.) selecting data in a $d_{\mathrm{IV}} > 30\,\mathrm{cm}$ FV. The IV shape is assumed as standardly deformed for the model, while it is considered differently for the pseudo dataset: standardly deformed (purple contour), 5 cm outwardly deformed (green contour), 5 cm inwardly deformed (red contour).

8.10 Conclusions

The SOX sensitivity studies topic, in case of limit setting, has been extensively reviewed. Results have shown that in a realistic case, the high SOX exclusion power allowed to exclude nearly all the 95% C.L. confidence region in the $\sin^2 2\theta - \Delta m^2$ plane allowed by the sterile anomalies. The most concerning systematic effects, due both to the $\overline{\nu}_e$ source specifications and to the detector response, have been analyzed; with the proper statistical approach, none of them has been found potentially able to spoil the experiment sensitivity power. Despite the SOX project interruption before the data taking beginning, the acquired expertise in the field of sensitivity studies is potentially employable again for next future neutrino experiments searching for sterile neutrinos as primary or secondary physics goal; a possible candidate is the JUNO experiment in China, as discussed in Chapter 9 of Ref. [174].

Chapter 9

Conclusions

The analysis of the Borexino Phase-III-TMP dataset described in this thesis has led to the first indication of the existence of CNO neutrinos. This result has been obtained following both the analytical and Monte Carlo fit approaches to disentangle the CNO-ν signal from the background. A crucial ingredient in the analysis is the constraint to the ^{210}Bi rate in the fit, which was obtained exploiting the ^{210}Bi -^{210}Po link.

Including both the statistical and systematic uncertainties, the best estimates for the CNO-ν interaction rate are $R^{\mathrm{MC}}(\mathrm{CNO}\text{-}\nu) = 5.4^{+1.3}_{-1.1}$ cpd/100t for the Monte Carlo fit and $R^{\mathrm{AN}}(\mathrm{CNO}\text{-}\nu) = 6.2^{+2.7}_{-2.6}$ cpd/100t for the analytical fit. The hypothesis of absence of a CNO-ν signal is rejected, by using both the approaches, with 4.9σ and 3.6σ statistical significance range, compatibly with the expectations coming from the sensitivity studies. These results represent the first direct experimental proof of the CNO cycle reactions which, according to the stellar models, are believed to be the primary source of energy for the stars more massive than the Sun.

Furthermore, I have exploited this measurement to infer additional information on solar physics. Combining the CNO-ν interaction rate obtained in this thesis with the previous Borexino ^{7}Be-ν and ^{8}B-ν measurements, I have shown that Borexino alone disfavours the LZ-SSM scenario at 98.5% C.L. (2.2σ). This slightly improves the result obtained by Borexino with the ^{7}Be-ν and ^{8}B-ν results alone. I have also shown that improving the CNO-ν measurement precision would not help much to discriminate between the HZ-SSM and LZ-SSM hypothesis, unless the theoretical error is also reduced.

Exploiting the CNO-ν rate measured in this thesis, I have applied the procedure described in Ref. [42, 167, 168] to extract the sum of Carbon and Nitrogen abundance in the core of the Sun, which is $N^{\mathrm{Bx}}_{\mathrm{C+N}} = (4.55 \pm 1.33) \cdot 10^{-4} N_{\mathrm{H}}$. In spite of the large error, this is the first direct measurement of the C+N primordial abundance and is very robust against details of the Standard Solar Model, since it has been derived minimizing the impact of the so-called *environmental* input parameters.

Perspectives and outlook

The results on the CNO-ν interaction rate obtained in this thesis are preliminary. We are working on several aspects of the analysis, in order to perform consistency checks, finalize the study of systematic errors and investigate possible improvements.

- *Consistency checks, including the low energy region.* For reasons described in details in the text, I have carried out the analysis in a restricted energy window ($320\,\mathrm{keV} \lesssim E \lesssim 2.9\,\mathrm{MeV}$) with respect to what was done for the pp cycle neutrinos analysis. This restricted range cuts out the contribution of lower energy species, including the one of pp neutrinos. Work to overcome this limitation is in progress and should be completed in the next months. Extending the range of the fit also to the low energy region, both with the analytical and Monte Carlo fits, would provide an important consistency check of the results by simultaneously extracting both the pp chain and the CNO cycle neutrino fluxes.

- *Study of systematic uncertainties.* The study of possible systematic effects needs to be completed, in particular with the addition of a part dedicated to the detector response uncertainty. However, we don't expect this uncertainty to be dominant with respect to the statistical error and therefore the conclusions on the significance of the CNO measurement should not change dramatically with respect to the preliminary result obtained in this thesis.

- *Reduction on the* ^{210}Bi *uncertainty.* The error on the CNO-ν measurement described in this thesis is dominated by the uncertainty on the ^{210}Bi constraint, which is currently $\sim 15\%$. This error is mainly due to the uncertainty on the assumption that the ^{210}Bi is spatially uniform, which is at the basis of the extrapolation of the result on the bubble to the entire fiducial volume.

  On the other hand, the error on the ^{210}Bi rate would be reduced if the clean region enlarges, hopefully including the entire fiducial volume. In order to pursue this goal, we have installed a control system of the Hall C temperature to try to further increase the temperature stability of the detector, in order to reduce the convective motions which prevent the clean bubble to increase in size. In the next months we will see if this is sufficient to reach this goal.

- *Calibration campaigns.* Before the end of the Borexino data-taking, an extensive campaign of calibrations with radioactive sources will be performed. This will allow to carry out further consistency checks, for example, for what concerns the energy response at low energies, where the fit shows the inconsistencies described in the text.

In conclusion, I presented the first indication of a signal of solar neutrinos from the CNO reactions cycle, with a $\sim 4\sigma$ significance, obtained analyzing the Borexino data collected between July 2016 up to June 2019. Several consistency checks are ongoing, while we are still collecting more data. Therefore, the results of my thesis are preliminary, but we do not expect them to change significantly in the final release, which should take place in the next few months.

Bibliography

[1] W. Pauli. Letter to the Tübingen conference. 1930 (cit. on p. 23).

[2] B. T. Cleveland et al. "Measurement of the solar electron neutrino flux with the Homestake chlorine detector". In: *Astrophys. J.* 496 (1998), pp. 505–526 (cit. on pp. 24, 39, 43).

[3] *The Nobel Prize in Physics 2015*. Oct. 2015. URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2015/ (cit. on pp. 24, 44).

[4] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (2018), p. 1898 (cit. on pp. 25, 26, 48, 49, 55, 102).

[5] Christian Spiering. "Towards High-Energy Neutrino Astronomy. A Historical Review". In: *Eur. Phys. J.* H37 (2012), pp. 515–565. arXiv: 1207.4952 [astro-ph.IM] (cit. on pp. 27, 28).

[6] Carlo Giunti and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. 2007. ISBN: 9780198508717 (cit. on pp. 27, 44, 46).

[7] Steven Weinberg. "Universal Neutrino Degeneracy". In: *Phys. Rev.* 128 (3 1962), pp. 1457–1473. URL: https://link.aps.org/doi/10.1103/PhysRev.128.1457 (cit. on p. 27).

[8] E. Baracchini et al. "PTOLEMY: A Proposal for Thermal Relic Detection of Massive Neutrinos and Directional Detection of MeV Dark Matter". In: (2018). arXiv: 1808.01892 [physics.ins-det] (cit. on p. 27).

[9] A. A. Hahn et al. "Anti-neutrino Spectra From ^{241}Pu and ^{239}Pu Thermal Neutron Fission Products". In: *Phys. Lett.* B218 (1989), pp. 365–368 (cit. on p. 27).

[10] F. Von Feilitzsch, A. A. Hahn, and K. Schreckenbach. "Experimental beta spectra from Pu-239 and U-235 thermal neutron fission products and their correlated anti-neutrinos spectra". In: *Phys. Lett.* B118 (1982), pp. 162–166 (cit. on p. 27).

[11] Gianni Fiorentini, Marcello Lissia, and Fabio Mantovani. "Geo-neutrinos and Earth's interior". In: *Phys. Rept.* 453 (2007), pp. 117–172. arXiv: 0707.3203 [physics.geo-ph] (cit. on p. 28).

[12] G. Bellini et al. "Observation of Geo-Neutrinos". In: *Phys. Lett.* B687 (2010), pp. 299–304. arXiv: 1003.0284 [hep-ex] (cit. on pp. 28, 74, 103).

[13] G. Bellini et al. "Measurement of geo-neutrinos from 1353 days of Borexino". In: *Phys. Lett.* B722 (2013), pp. 295–300. arXiv: 1303.2571 [hep-ex] (cit. on pp. 28, 103).

[14] M. Agostini et al. "Spectroscopy of geoneutrinos from 2056 days of Borexino data". In: *Phys. Rev.* D92.3 (2015), p. 031101. arXiv: 1506.04610 [hep-ex] (cit. on pp. 28, 103, 231, 232).

[15] K. Hirata et al. "Observation of a neutrino burst from the supernova SN1987A". In: *Phys. Rev. Lett.* 58 (14 1987), pp. 1490–1493. URL: http://link.aps.org/doi/10.1103/PhysRevLett.58.1490 (cit. on p. 29).

[16] R. M. Bionta et al. "Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud". In: *Phys. Rev. Lett.* 58 (14 1987), pp. 1494–1496. URL: http://link.aps.org/doi/10.1103/PhysRevLett.58.1494 (cit. on p. 29).

[17] E. N. Alekseev et al. "Possible Detection of a Neutrino Signal on 23 February 1987 at the Baksan Underground Scintillation Telescope of the Institute of Nuclear Research". In: *JETP Lett.* 45 (1987). [,739(1987)], pp. 589–592 (cit. on p. 29).

[18] V. S. Berezinsky and G. T. Zatsepin. "Cosmic neutrinos of superhigh energy". In: *Yad. Fiz.* 11 (1970), pp. 200–205 (cit. on p. 29).

[19] J. Ahrens et al. "Sensitivity of the IceCube detector to astrophysical sources of high energy muon neutrinos". In: *Astropart. Phys.* 20 (2004), pp. 507–532. arXiv: astro-ph/0305196 [astro-ph] (cit. on p. 29).

[20] M. G. Aartsen et al. "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector". In: *Science* 342 (2013), p. 1242856. arXiv: 1311.5238 [astro-ph.HE] (cit. on p. 29).

[21] M. G. Aartsen et al. "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert". In: *Science* 361.6398 (2018), pp. 147–151. arXiv: 1807.08794 [astro-ph.HE] (cit. on p. 29).

[22] P. Padovani et al. "Dissecting the region around IceCube-170922A: the blazar TXS 0506+056 as the first cosmic neutrino source". In: *Mon. Not. Roy. Astron. Soc.* 480.1 (2018), pp. 192–203. arXiv: 1807.04461 [astro-ph.HE] (cit. on p. 29).

[23] J. Bahcall. *Neutrino astrophysics*. Cambridge University Press, 1989. ISBN: 0521351138 (cit. on pp. 29, 34).

[24] Aldo Serenelli. "Alive and well: a short review about standard solar models". In: *Eur. Phys. J.* A52.4 (2016), p. 78. arXiv: 1601.07179 [astro-ph.SR] (cit. on pp. 29, 38, 218).

[25] Núria Vinyoles et al. "A new Generation of Standard Solar Models". In: *Astrophys. J.* 835.2 (2017), p. 202. arXiv: 1611.09867 [astro-ph.SR] (cit. on pp. 29, 30, 33, 35, 36, 38, 39, 73, 106, 108).

[26] Aldo M. Serenelli, W. C. Haxton, and Carlos Peña-Garay. "Solar models with accretion. I. Application to the solar abundance problem". In: *Astrophys. J.* 743 (2011), p. 24. arXiv: 1104.1639 [astro-ph.SR] (cit. on pp. 29, 30).

[27] W. C. Haxton, R. G. Hamish Robertson, and Aldo M. Serenelli. "Solar Neutrinos: Status and Prospects". In: *Ann. Rev. Astron. Astrophys.* 51 (2013), pp. 21–61. arXiv: 1208.5723 [astro-ph.SR] (cit. on pp. 29, 35).

[28] Sarbani Basu and H. M. Antia. "Helioseismology and Solar Abundances". In: *Phys. Rept.* 457 (2008), pp. 217–283. arXiv: 0711.4590 [astro-ph] (cit. on pp. 29, 30).

[29] Sarbani Basu. "Global seismology of the Sun". In: *Living Reviews in Solar Physics* 13.1 (2016). ISSN: 1614-4961. URL: http://dx.doi.org/10.1007/s41116-016-0003-4 (cit. on pp. 30, 38).

[30] N. Grevesse and A. J. Sauval. "Standard Solar Composition". In: *Space Sci. Rev.* 85 (1998), pp. 161–174 (cit. on pp. 30, 35).

[31] M. Asplund et al. "The Chemical Composition of the Sun". In: *Annual Review of Astronomy & Astrophysics* 47 (Sept. 2009), pp. 481–522. arXiv: 0909.0948 [astro-ph.SR] (cit. on pp. 30, 35, 218).

[32] Martin Asplund, Nicolas Grevesse, and Jacques Sauval. "The Solar chemical composition". In: *Nucl. Phys.* A777 (2006). [ASP Conf. Ser.336,25(2005)], pp. 1–4. arXiv: astro-ph/0410214 [astro-ph] (cit. on pp. 30, 218).

[33] E. Caffau et al. "Solar Chemical Abundances Determined with a CO5BOLD 3D Model Atmosphere". In: *Solar Physics* 268 (Feb. 2011), pp. 255–269. arXiv: 1003.1190 [astro-ph.SR] (cit. on pp. 30, 218).

[34] Joyce A. Guzik, L. Scott Watson, and Arthur N. Cox. "Can enhanced diffusion improve helioseismic agreement for solar models with revised abundances?" In: *Astrophys. J.* 627 (2005), pp. 1049–1056. arXiv: astro-ph/0502364 [astro-ph] (cit. on p. 30).

[35] H. A. Bethe. "Energy Production in Stars". In: *Phys. Rev.* 55 (5 1939), pp. 434–456. URL: http://link.aps.org/doi/10.1103/PhysRev.55.434 (cit. on p. 31).

[36] URL: http://www.sns.ias.edu/~jnb/ (cit. on pp. 33, 36, 37, 219).

[37] John N. Bahcall. "Solar Neutrinos. I. Theoretical". In: *Phys. Rev. Lett.* 12 (11 1964), pp. 300–302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.12.300 (cit. on p. 31).

[38] John N. Bahcall. "Solar Neutrino Cross Sections and Nuclear Beta Decay". In: *Phys. Rev.* 135 (1B 1964), B137–B146. URL: https://link.aps.org/doi/10.1103/PhysRev.135.B137 (cit. on pp. 32, 33).

[39] M. Agostini et al. "Comprehensive measurement of *pp*-chain solar neutrinos". In: *Nature* 562.7728 (2018), pp. 505–510 (cit. on pp. 34, 43, 67, 91, 95, 101, 102, 103, 105, 107, 108, 116, 119, 121, 122, 123, 131, 132, 136, 139, 143, 172, 174, 208, 209, 211, 216, 219, 220).

[40] S. Fukuda et al. "Solar B-8 and hep neutrino measurements from 1258 days of Super-Kamiokande data". In: *Phys. Rev. Lett.* 86 (2001), pp. 5651–5655. arXiv: hep-ex/0103032 [hep-ex] (cit. on pp. 34, 41).

[41] E. G. Adelberger et al. "Solar fusion cross sections II: the pp chain and CNO cycles". In: *Rev. Mod. Phys.* 83 (2011), p. 195. arXiv: 1004.2318 [nucl-ex] (cit. on pp. 35, 37).

[42] W. C. Haxton and A. M. Serenelli. "CN Cycle Solar Neutrinos and the Sun's Primordial Core Metallicity". In: *The Astrophysical Journal* 687.1 (Nov. 2008), pp. 678–691. URL: https://doi.org/10.1086%2F591787 (cit. on pp. 35, 37, 222, 226, 261).

[43] Jorgen Christensen-Dalsgaard. "Helioseismology". In: *Rev. Mod. Phys.* 74 (2003), pp. 1073–1129. arXiv: astro-ph/0207403 [astro-ph] (cit. on p. 38).

[44] J. Christensen-Dalsgaard et al. "The current state of solar modeling". In: *Science* 272 (1996), pp. 1286–1292 (cit. on p. 38).

[45] S. Basu. "The Solar Metallicity Problem: What is the Solution?" In: *Solar-Stellar Dynamos as Revealed by Helio-and Asteroseismology: GONG 2008/SOHO 21*. Ed. by M. Dikpati et al. Vol. 416. Astronomical Society of the Pacific Conference Series. Dec. 2009, p. 193 (cit. on p. 38).

[46] W. Hampel et al. "GALLEX solar neutrino observations: Results for GALLEX IV". In: *Phys. Lett.* B447 (1999), pp. 127–133 (cit. on p. 40).

[47] M. Altmann et al. "GNO solar neutrino observations: Results for GNO I". In: *Phys. Lett.* B490 (2000), pp. 16–26. arXiv: hep-ex/0006034 [hep-ex] (cit. on p. 40).

[48] M. Altmann et al. "Complete results for five years of GNO solar neutrino observations". In: *Phys. Lett.* B616 (2005), pp. 174–190. arXiv: hep-ex/0504037 [hep-ex] (cit. on p. 40).

[49] Dzh.N. Abdurashitov et al. "The Russian-American gallium experiment (SAGE) Cr neutrino source measurement". In: *Phys. Rev. Lett.* 77 (1996), pp. 4708–4711 (cit. on pp. 40, 55).

[50] D. Vignaud. "Solar neutrinos observed by GALLEX at Gran Sasso". In: *In *Dallas 1992, Proceedings, High energy physics, vol. 1* 1093-1110*. 1992 (cit. on pp. 40, 43).

[51] J. N. Abdurashitov et al. "Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period". In: *Phys. Rev.* C80 (2009), p. 015807. arXiv: 0901.2200 [nucl-ex] (cit. on pp. 40, 55).

[52] K. Abe et al. "Solar Neutrino Measurements in Super-Kamiokande-IV". In: *Phys. Rev.* D94.5 (2016), p. 052010. arXiv: 1606.07538 [hep-ex] (cit. on pp. 41, 43, 91, 103, 223).

[53] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. arXiv: hep-ex/9807003 [hep-ex] (cit. on pp. 41, 43, 55).

[54] G. T. Ewan, W. F. Davidson, and C. G. Hargrove. "The Sudbury Neutrino Observatory - an introduction". In: *Phys. Canada* 48 (1992), pp. 112–115 (cit. on p. 41).

[55] A. B. Balantekin and F. Loreti. "Solar and supernova neutrino physics with Sudbury Neutrino Observatory". In: *Phys. Rev.* D45 (1992), pp. 1059–1065 (cit. on p. 41).

[56] Arthur B. McDonald. "Nobel Lecture: The Sudbury Neutrino Observatory: Observation of flavor change for solar neutrinos". In: *Rev. Mod. Phys.* 88 (3 2016), p. 030502. URL: https://link.aps.org/doi/10.1103/RevModPhys.88.030502 (cit. on pp. 41, 42).

[57] B. Aharmim et al. "Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory". In: *Phys. Rev.* C88 (2013), p. 025501. arXiv: 1109.0763 [nucl-ex] (cit. on pp. 41, 42, 103).

[58] John N. Bahcall and M. H. Pinsonneault. "What Do We (Not) Know Theoretically about Solar Neutrino Fluxes?" In: *Phys. Rev. Lett.* 92 (12 2004), p. 121301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.92.121301 (cit. on p. 42).

[59] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002), p. 011301. arXiv: nucl-ex/0204008 [nucl-ex] (cit. on p. 42).

[60] Q. R. Ahmad et al. "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (2001), p. 071301. arXiv: nucl-ex/0106015 [nucl-ex] (cit. on pp. 42, 43).

[61] G. Alimonti et al. "The Borexino detector at the Laboratori Nazionali del Gran Sasso". In: *Nucl. Instrum. Meth.* A600 (2009), pp. 568–593. arXiv: 0806.2400 [physics.ins-det] (cit. on pp. 43, 67, 68, 71, 81, 82, 83, 84, 86).

[62] A. Ianni. "Neutrino physics with Borexino". In: *Prog. Part. Nucl. Phys.* 66 (2011), pp. 405–411 (cit. on pp. 43, 67, 91, 105).

[63] M. Agostini et al. "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II". In: (2017). arXiv: 1707.09279 [hep-ex] (cit. on pp. 43, 52, 91, 105, 107, 108, 116, 119, 120, 128, 131, 136, 212, 216).

[64] G. Bellini et al. "Absence of day–night asymmetry of 862 keV ^{7}Be solar neutrino rate in Borexino and MSW oscillation parameters". In: *Phys. Lett.* B707 (2012), pp. 22–26. arXiv: 1104.2150 [hep-ex] (cit. on pp. 43, 91).

[65] G. Bellini et al. "Final results of Borexino Phase-I on low energy solar neutrino spectroscopy". In: *Phys. Rev.* D89.11 (2014), p. 112007. arXiv: 1308.0443 [hep-ex] (cit. on pp. 43, 67, 72, 79, 86, 91, 105, 120, 121, 122, 123, 129, 130, 143, 213, 216).

[66] Samoil M. Bilenky and B. Pontecorvo. "The Quark-Lepton Analogy and the Muonic Charge". In: *Yad. Fiz.* 24 (1976). [Sov. J. Nucl. Phys.24,316(1976)], pp. 603–608 (cit. on p. 43).

[67] Samoil M. Bilenky and B. Pontecorvo. "Again on Neutrino Oscillations". In: *Lett. Nuovo Cim.* 17 (1976), p. 569 (cit. on p. 43).

[68] Shalom Eliezer and Arthur R. Swift. "Experimental Consequences of electron Neutrino-Muon-neutrino Mixing in Neutrino Beams". In: *Nucl. Phys.* B105 (1976), pp. 45–51 (cit. on p. 43).

[69] Y. Fukuda et al. "Atmospheric muon-neutrino / electron-neutrino ratio in the multiGeV energy range". In: *Phys. Lett.* B335 (1994), pp. 237–245 (cit. on p. 43).

[70] K. Eguchi et al. "First results from KamLAND: Evidence for reactor anti-neutrino disappearance". In: *Phys. Rev. Lett.* 90 (2003), p. 021802. arXiv: hep-ex/0212021 [hep-ex] (cit. on pp. 43, 55).

[71] J. K. Ahn et al. "Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment". In: *Phys. Rev. Lett.* 108 (2012), p. 191802. arXiv: 1204.0626 [hep-ex] (cit. on pp. 43, 55).

[72] F. P. An et al. "Observation of electron-antineutrino disappearance at Daya Bay". In: *Phys. Rev. Lett.* 108 (2012), p. 171803. arXiv: 1203.1669 [hep-ex] (cit. on pp. 43, 55).

[73] Y. Abe et al. "Indication of Reactor $\overline{\nu}_e$ Disappearance in the Double Chooz Experiment". In: *Phys. Rev. Lett.* 108 (13 2012), p. 131801. URL: http://link.aps.org/doi/10.1103/PhysRevLett.108.131801 (cit. on pp. 43, 58).

[74] D. G. Michael et al. "Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam". In: *Phys. Rev. Lett.* 97 (2006), p. 191801. arXiv: hep-ex/0607088 [hep-ex] (cit. on pp. 43, 55, 65).

[75] K. Abe et al. "Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam". In: *Phys. Rev. Lett.* 107 (2011), p. 041801. arXiv: 1106.2822 [hep-ex] (cit. on pp. 43, 55).

[76] N. Agafonova et al. "Observation of a first ν_τ candidate in the OPERA experiment in the CNGS beam". In: *Phys. Lett.* B691 (2010), pp. 138–145. arXiv: 1006.1623 [hep-ex] (cit. on pp. 43, 63).

[77] M. A. Acero et al. "First Measurement of Neutrino Oscillation Parameters using Neutrinos and Antineutrinos by NOvA". In: *Phys. Rev. Lett.* 123.15 (2019), p. 151803. arXiv: 1906.04907 [hep-ex] (cit. on pp. 43, 55).

[78] M. G. Aartsen et al. "Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data". In: *Phys. Rev.* D91.7 (2015), p. 072004. arXiv: 1410.7227 [hep-ex] (cit. on p. 43).

[79] M. G. Aartsen et al. "Flavor Ratio of Astrophysical Neutrinos above 35 TeV in IceCube". In: *Phys. Rev. Lett.* 114.17 (2015), p. 171102. arXiv: 1502.03376 [astro-ph.HE] (cit. on p. 43).

[80] L. Wolfenstein. "Neutrino Oscillations in Matter". In: *Phys. Rev.* D17 (1978), pp. 2369–2374 (cit. on pp. 43, 51).

[81] S. P. Mikheev and A. Yu. Smirnov. "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos". In: *Sov. J. Nucl. Phys.* 42 (1985). [Yad. Fiz.42,1441(1985)], pp. 913–917 (cit. on pp. 43, 51, 53).

[82] E. Kh. Akhmedov and A. Yu. Smirnov. "Paradoxes of neutrino oscillations". In: *Physics of Atomic Nuclei* 72.8 (2009), pp. 1363–1381. ISSN: 1562-692X. URL: https://doi.org/10.1134/S1063778809080122 (cit. on p. 44).

[83] Alessandro Strumia and Francesco Vissani. "Neutrino masses and mixings and..." In: (2006). arXiv: hep-ph/0606054 [hep-ph] (cit. on p. 44).

[84] C. Giunti and C. W. Kim. "Coherence of neutrino oscillations in the wave packet approach". In: *Phys. Rev.* D58 (1998), p. 017301. arXiv: hep-ph/9711363 [hep-ph] (cit. on p. 44).

[85] S. Nussinov. "Solar neutrinos and neutrino mixing". In: *Phys. Lett.* B63 (1976), pp. 201–203 (cit. on p. 44).

[86] Francesco Capozzi et al. "Global constraints on absolute neutrino masses and their ordering". In: *Phys. Rev.* D95.9 (2017), p. 096014. arXiv: 1703.04471 [hep-ph] (cit. on pp. 46, 48, 73).

[87] C. et al. Alduino. "First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ^{130}Te". In: *Phys. Rev. Lett.* 120 (13 2018), p. 132501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.132501 (cit. on p. 54).

[88] M. et al. Agostini. "Improved Limit on Neutrinoless Double-β Decay of ^{76}Ge from GERDA Phase II". In: *Phys. Rev. Lett.* 120 (13 2018), p. 132503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.132503 (cit. on p. 54).

[89] J. B. Albert et al. "Search for Neutrinoless Double-Beta Decay with the Upgraded EXO-200 Detector". In: *Phys. Rev. Lett.* 120.7 (2018), p. 072701. arXiv: 1707.08707 [hep-ex] (cit. on p. 54).

[90] R. Arnold et al. "Results of the search for neutrinoless double-β decay in ^{100}Mo with the NEMO-3 experiment". In: *Phys. Rev.* D92.7 (2015), p. 072011. arXiv: 1506.05825 [hep-ex] (cit. on p. 54).

[91] A. Gando et al. "Search for Majorana Neutrinos near the Inverted Mass Hierarchy Region with KamLAND-Zen". In: *Phys. Rev. Lett.* 117.8 (2016). [Addendum: Phys. Rev. Lett.117,no.10,109903(2016)], p. 082503. arXiv: 1605.02889 [hep-ex] (cit. on p. 54).

[92] M. Agostini et al. "Probing Majorana neutrinos with double-β decay". In: *Science* 365 (2019), p. 1445. arXiv: 1909.02726 [hep-ex] (cit. on p. 54).

[93] V. N. Aseev et al. "An upper limit on electron antineutrino mass from Troitsk experiment". In: *Phys. Rev.* D84 (2011), p. 112003. arXiv: 1108.5034 [hep-ex] (cit. on p. 54).

[94] Ch. Kraus et al. "Final results from phase II of the Mainz neutrino mass search in tritium beta decay". In: *Eur. Phys. J.* C40 (2005), pp. 447–468. arXiv: hep-ex/0412056 [hep-ex] (cit. on p. 54).

[95] A. Osipowicz et al. "KATRIN: A Next generation tritium beta decay experiment with sub-eV sensitivity for the electron neutrino mass. Letter of intent". In: (2001). arXiv: hep-ex/0109033 [hep-ex] (cit. on p. 54).

[96] M. Aker et al. "An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN". In: (2019). arXiv: 1909.06048 [hep-ex] (cit. on p. 54).

[97] P. A. R. Ade et al. "Planck 2015 results. XIII. Cosmological parameters". In: (2015). arXiv: 1502.01589 [astro-ph.CO] (cit. on pp. 54, 59, 60, 64, 227).

[98] K. N. Abazajian et al. "Light Sterile Neutrinos: A White Paper". In: (2012). arXiv: 1204.5379 [hep-ph] (cit. on pp. 55, 227).

[99] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *JETP* 26 (1968), p. 984 (cit. on p. 55).

[100] P. Anselmann et al. "First results from the Cr-51 neutrino source experiment with the GALLEX detector". In: *Phys. Lett.* B342 (1995), pp. 440–450 (cit. on p. 55).

[101] W. Hampel et al. "Final results of the Cr-51 neutrino source experiments in GALLEX". In: *Phys. Lett.* B420 (1998), pp. 114–126 (cit. on p. 55).

[102] A. Aguilar-Arevalo et al. "Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam". In: *Phys. Rev.* D64 (2001), p. 112007. arXiv: hep-ex/0104049 [hep-ex] (cit. on pp. 56, 57).

[103] A. A. Aguilar-Arevalo et al. "Event Excess in the MiniBooNE Search for $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillations". In: *Phys. Rev. Lett.* 105 (2010), p. 181801. arXiv: 1007.1150 [hep-ex] (cit. on p. 57).

[104] A. A. Aguilar-Arevalo et al. "Significant Excess of ElectronLike Events in the MiniBooNE Short-Baseline Neutrino Experiment". In: *Phys. Rev. Lett.* 121.22 (2018), p. 221801. arXiv: 1805.12028 [hep-ex] (cit. on pp. 58, 63).

[105] R. Acciarri et al. "Design and Construction of the MicroBooNE Detector". In: *JINST* 12.02 (2017), P02017. arXiv: 1612.05824 [physics.ins-det] (cit. on p. 58).

[106] B. Armbruster et al. "Upper limits for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ from muon decay at rest". In: *Phys. Rev. D* 65 (11 2002), p. 112001. URL: https://link.aps.org/doi/10.1103/PhysRevD.65.112001 (cit. on pp. 58, 63).

[107] Ashley Timmons. "The Results of MINOS and the Future with MINOS+". In: (2015). arXiv: 1511.06178 [hep-ex] (cit. on p. 58).

[108] Th. A. Mueller et al. "Improved predictions of reactor antineutrino spectra". In: *Phys. Rev. C* 83 (5 2011), p. 054615. URL: http://link.aps.org/doi/10.1103/PhysRevC.83.054615 (cit. on p. 58).

[109] J. M. Conrad et al. "Sterile Neutrino Fits to Short Baseline Neutrino Oscillation Measurements". In: *Adv. High Energy Phys.* 2013 (2013), p. 163897. arXiv: 1207.4765 [hep-ex] (cit. on p. 59).

[110] C. Giunti and Y. F. Li. "Matter Effects in Active-Sterile Solar Neutrino Oscillations". In: *Phys. Rev.* D80 (2009), p. 113007. arXiv: 0910.5856 [hep-ph] (cit. on p. 60).

[111] Sandhya Choubey. "Signature of sterile species in atmospheric neutrino data at neutrino telescopes". In: *JHEP* 12 (2007), p. 014. arXiv: 0709.1937 [hep-ph] (cit. on p. 60).

[112] Carlo Giunti. "Light Sterile Neutrinos: Status and Perspectives". In: (2015). arXiv: 1512.04758 [hep-ph] (cit. on pp. 61, 64).

[113] M. Antonello et al. "A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam". In: (2015). arXiv: 1503.01520 [physics.ins-det] (cit. on pp. 63, 65).

[114] Johnathon R. Jordan et al. "Severe Constraints on New Physics Explanations of the MiniBooNE Excess". In: (2018). arXiv: 1810.07185 [hep-ph] (cit. on p. 63).

[115] Jiajun Liao, Danny Marfatia, and Kerry Whisnant. "MiniBooNE, MINOS+ and IceCube data imply a baroque neutrino sector". In: (2018). arXiv: 1810.01000 [hep-ph] (cit. on p. 63).

[116] Peter Ballett, Silvia Pascoli, and Mark Ross-Lonergan. "U(1)' mediated decays of heavy sterile neutrinos in MiniBooNE". In: (2018). arXiv: 1808.02915 [hep-ph] (cit. on p. 63).

[117] Enrico Bertuzzo et al. "Dark Neutrino Portal to Explain MiniBooNE excess". In: *Phys. Rev. Lett.* 121.24 (2018), p. 241801. arXiv: 1807.09877 [hep-ph] (cit. on p. 63).

[118] G. Karagiorgi, M. H. Shaevitz, and J. M. Conrad. "Confronting the Short-Baseline Oscillation Anomalies with a Single Sterile Neutrino and Non-Standard Matter Effects". In: (2012). arXiv: 1202.1024 [hep-ph] (cit. on p. 63).

[119] Jiajun Liao and Danny Marfatia. "Impact of nonstandard interactions on sterile neutrino searches at IceCube". In: *Phys. Rev. Lett.* 117.7 (2016), p. 071802. arXiv: 1602.08766 [hep-ph] (cit. on p. 63).

[120] Marcela Carena et al. "Neutrinos in large extra dimensions and short-baseline ν_e appearance". In: *Phys. Rev. D* 96 (9 2017), p. 095014. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.095014 (cit. on p. 63).

[121] Joachim Kopp, Michele Maltoni, and Thomas Schwetz. "Are there sterile neutrinos at the eV scale?" In: *Phys. Rev. Lett.* 107 (2011), p. 091801. arXiv: 1103.4570 [hep-ph] (cit. on p. 64).

[122] Evgeny Akhmedov and Thomas Schwetz. "MiniBooNE and LSND data: Non-standard neutrino interactions in a (3+1) scheme versus (3+2) oscillations". In: *JHEP* 10 (2010), p. 115. arXiv: 1007.4171 [hep-ph] (cit. on p. 64).

[123] Mikhail Danilov. "Searches for sterile neutrinos at very short baseline reactor experiments". In: *4th International Conference on Particle Physics and Astrophysics (ICPPA 2018) Moscow, Russia, October 22-26, 2018*. 2018. arXiv: 1812.04085 [hep-ex] (cit. on p. 64).

[124] I Alekseev et al. "Search for sterile neutrinos at the DANSS experiment". In: *Phys. Lett.* B787 (2018), pp. 56–63. arXiv: 1804.04046 [hep-ex] (cit. on p. 65).

[125] I. Alekseev et al. "DANSS: Detector of the reactor AntiNeutrino based on Solid Scintillator". In: *JINST* 11.11 (2016), P11011. arXiv: 1606.02896 [physics.ins-det] (cit. on p. 65).

[126] Y. J. Ko et al. "Sterile Neutrino Search at the NEOS Experiment". In: *Phys. Rev. Lett.* 118.12 (2017), p. 121802. arXiv: 1610.05134 [hep-ex] (cit. on p. 65).

[127] A. P. Serebrov et al. "NEUTRINO4 experiment: preparations for search for sterile neutrino at 100 MW Reactor SM-3 at 6-12 Meters". In: (2012). arXiv: 1205.2955 [hep-ph] (cit. on p. 65).

[128] J. Ashenfelter et al. "The PROSPECT Physics Program". In: *J. Phys.* G43.11 (2016), p. 113001. arXiv: 1512.02202 [physics.ins-det] (cit. on p. 65).

[129] J. P. Chen et al. "A White Paper on SoLID (Solenoidal Large Intensity Device)". In: (2014). arXiv: 1409.7741 [nucl-ex] (cit. on p. 65).

[130] H. Almazàn et al. "Sterile Neutrino Constraints from the STEREO Experiment with 66 Days of Reactor-On Data". In: *Phys. Rev. Lett.* 121.16 (2018), p. 161801. arXiv: 1806.02096 [hep-ex] (cit. on p. 65).

[131] P. Adamson et al. "Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit". In: (2017). arXiv: 1710.06488 [hep-ex] (cit. on p. 65).

[132] Sandhya Choubey and Dipyaman Pramanik. "Constraints on Sterile Neutrino Oscillations using DUNE Near Detector". In: *Phys. Lett.* B764 (2017), pp. 135–141. arXiv: 1604.04731 [hep-ph] (cit. on p. 66).

[133] A. Bandyopadhyay. "Physics at a future Neutrino Factory and super-beam facility". In: *Rept. Prog. Phys.* 72 (2009). Ed. by S. Choubey et al., p. 106201. arXiv: 0710.4947 [hep-ph] (cit. on p. 66).

[134] Arman Esmaili and Orlando L. G. Peres. "KATRIN Sensitivity to Sterile Neutrino Mass in the Shadow of Lightest Neutrino Mass". In: *Phys. Rev.* D85 (2012), p. 117301. arXiv: 1203.2632 [hep-ph] (cit. on p. 66).

[135] A. I. Belesev et al. "The search for an additional neutrino mass eigenstate in the 2-100 eV region from 'Troitsk nu-mass' data: a detailed analysis". In: *J. Phys.* G41 (2014), p. 015001. arXiv: 1307.5687 [hep-ex] (cit. on p. 66).

[136] G. Bellini et al. "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector". In: *Phys. Rev. D* 82 (3 2010), p. 033006. URL: http://link.aps.org/doi/10.1103/PhysRevD.82.033006 (cit. on pp. 67, 95, 96, 105, 143).

[137] G. Bellini et al. "Neutrinos from the primary proton-proton fusion process in the Sun". In: *Nature* 512.7515 (2014), pp. 383–386 (cit. on pp. 67, 91, 94, 105, 120, 131).

[138] G. Bellini et al. "First Evidence of pep Solar Neutrinos by Direct Detection in Borexino". In: *Phys. Rev. Lett.* 108 (5 2012), p. 051302. URL: http://link.aps.org/doi/10.1103/PhysRevLett.108.051302 (cit. on pp. 67, 91, 105, 143).

[139] G. Alimonti et al. "Science and technology of BOREXINO: A Real time detector for low-energy solar neutrinos". In: *Astropart. Phys.* 16 (2002), pp. 205–234. arXiv: hep-ex/0012030 [hep-ex] (cit. on p. 68).

[140] G.'t Hooft. "Prediction for neutrino-electron cross-sections in Weinberg's model of weak interactions". In: *Physics Letters B* 37.2 (1971), pp. 195 –196. ISSN: 0370-2693. URL: http://www.sciencedirect.com/science/article/pii/0370269371900505 (cit. on p. 72).

[141] C. Arpesella et al. "Radon measurements in the Gran Sasso underground laboratory". In: *Health Physics* 72.4 (1997), pp. 629–632 (cit. on p. 77).

[142] G. Bellini et al. "Muon and Cosmogenic Neutron Detection in Borexino". In: *Journal of Instrumentation* 6 (May 2011), P05005 (cit. on pp. 78, 81).

[143] G. v. Bunau. "J. B. Birks: Photophysics of Aromatic Molecules. Wiley-Interscience, London 1970. 704 Seiten. Preis: 210s". In: *Berichte der Bunsengesellschaft fur physikalische Chemie* 74.12 (1970), pp. 1294–1295. ISSN: 0005-9021. URL: http://dx.doi.org/10.1002/bbpc.19700741223 (cit. on p. 85).

[144] M. Agostini et al. "Seasonal Modulation of the ^{7}Be Solar Neutrino Rate in Borexino". In: *Astropart. Phys.* 92 (2017), pp. 21–29. arXiv: 1701.07970 [hep-ex] (cit. on p. 89).

[145] M. Agostini et al. "Improved measurement of ^{8}B solar neutrinos with 1.5 kt y of Borexino exposure". In: (2017). arXiv: 1709.00756 [hep-ex] (cit. on p. 95).

[146] John N. Bahcall. "The Luminosity constraint on solar neutrino fluxes". In: *Phys. Rev.* C65 (2002), p. 025801. arXiv: hep-ph/0108148 [hep-ph] (cit. on p. 101).

[147] F Vissani. *Luminosity constraint and entangled solar neutrino signals*. Aug. 2018 (cit. on p. 101).

[148] John N. Bahcall. "Solar Models: An Historical Overview". In: *Nuclear Physics B - Proceedings Supplements* 118 (Sept. 2002), pp. 77–86 (cit. on p. 101).

[149] Ivan Esteban et al. "Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity". In: *JHEP* 01 (2017), p. 087. arXiv: 1611.01514 [hep-ph] (cit. on p. 102).

[150] B. P. Abbott et al. "GW150914: The Advanced LIGO Detectors in the Era of First Discoveries". In: *Physical Review Letters* 116.13, 131103 (Apr. 2016), p. 131103. arXiv: 1602.03838 [gr-qc] (cit. on p. 103).

[151] B. P. Abbott et al. "The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914". In: *The Astrophysical Journal Letters* 833.1 (2016), p. L1. URL: http://stacks.iop.org/2041-8205/833/i=1/a=L1 (cit. on p. 103).

[152] M. Agostini et al. "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data". In: *Phys. Rev.* D96.9 (2017), p. 091103. arXiv: 1707.09355 [hep-ex] (cit. on p. 103).

[153] H. O. Back et al. "Search for electron decay mode $e \to \gamma + \nu$ with prototype of Borexino detector". In: *Phys. Lett.* B525 (2002), pp. 29–40 (cit. on p. 103).

[154] H. O. Back et al. "New limits on nucleon decays into invisible channels with the BOREXINO counting test facility". In: *Phys. Lett.* B563 (2003), pp. 23–34. arXiv: hep-ex/0302002 [hep-ex] (cit. on p. 103).

[155] S. K. Agarwalla et al. "Constraints on Non-Standard Neutrino Interactions from Borexino Phase-II". In: (2019). arXiv: 1905.03512 [hep-ph] (cit. on p. 103).

[156] G. Bellini et al. "SOX: Short distance neutrino Oscillations with BoreXino". In: *JHEP* 08 (2013), p. 038. arXiv: 1304.7721 [physics.ins-det] (cit. on pp. 103, 227).

[157] *Official cancellation of SOX project*. Feb. 2018. URL: http://home.infn.it/en/sox/13-esperimenti/2681-sox-short-distance-oscillations-with-borexino-ing (cit. on pp. 103, 227).

[158] D. Guffanti et al. "Sensitivity of Borexino Phase II to CNO neutrinos". Borexino Internal Report. 2018 (cit. on p. 113).

[159] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (2011), p. 1554. URL: https://doi.org/10.1140/epjc/s10052-011-1554-0 (cit. on p. 113).

[160] H. Back et al. "CNO and pep neutrino spectroscopy in Borexino: Measurement of the deep underground production of cosmogenic 11C in organic liquid scintillator". In: *Phys. Rev.* C74 (2006), p. 045805. arXiv: hep-ex/0601035 [hep-ex] (cit. on p. 122).

[161] X.F. Ding. "GooStats: A GPU-based framework for multi-variate analysis in particle physics". In: *Journal of Instrumentation* 13.12 (2018), P12018–P12018. URL: https://doi.org/10.1088%2F1748-0221%2F13%2F12%2Fp12018 (cit. on p. 128).

[162] M. Agostini et al. "The Monte Carlo simulation of the Borexino detector". In: *Astropart. Phys.* 97 (2018), pp. 136–159. arXiv: 1704.02291 [physics.ins-det] (cit. on pp. 131, 132, 143, 144, 145, 146, 147, 149).

[163] Simone Marcocci. "Precision measurement of solar neutrino fluxes with Borexino and prospects for $0\nu\beta\beta$ search with 136Xe-loaded liquid scintillators". PhD thesis. GSSI, Aquila, 2016-10-30. URL: http://borex.lngs.infn.it/Thesis/Simone_Marcocci_PhD_Thesis.pdf (cit. on pp. 132, 145, 147).

[164] H. Back et al. "Borexino calibrations: Hardware, Methods, and Results". In: *JINST* 7 (2012), P10018. arXiv: 1207.4816 [physics.ins-det] (cit. on pp. 143, 147, 148).

[165] B. Caccianiga et al. "A multiplexed optical-fiber system for the PMT calibration of the Borexino experiment". In: *Nucl. Instrum. Meth.* A496 (2003), pp. 353–361 (cit. on p. 149).

[166] David Bravo-Berguño et al. "Fluid-dynamics in the Borexino Neutrino Detector: behavior of a pseudo-stably-stratified, near-equilibrium closed system under asymmetrical, changing boundary conditions". In: (2017). arXiv: 1705.09658 [physics.ins-det] (cit. on pp. 182, 184, 185).

[167] Aldo Serenelli, Carlos Peña Garay, and W. C. Haxton. "Using the standard solar model to constrain solar composition and nuclear reaction S factors". In: *Phys. Rev. D* 87 (4 2013), p. 043001. URL: https://link.aps.org/doi/10.1103/PhysRevD.87.043001 (cit. on pp. 222, 226, 261).

[168] D. Guffanti and F. L. Villante. "Something new inside the Sun: what a future measurement of CNO neutrinos can tell us about the Sun". Internal Note. 2020 (cit. on pp. 222, 223, 226, 261).

[169] John N. Bahcall and Andrew Ulmer. "The Temperature dependence of solar neutrino fluxes". In: *Phys. Rev.* D53 (1996), pp. 4202–4210. arXiv: astro-ph/9602012 [astro-ph] (cit. on p. 223).

[170] F. L. Villante. "Private communication". 2019 (cit. on p. 223).

[171] J. Gaffiot et al. "Experimental Parameters for a Cerium 144 Based Intense Electron Antineutrino Generator Experiment at Very Short Baselines". In: *Phys. Rev.* D91.7 (2015), p. 072005. arXiv: 1411.6694 [physics.ins-det] (cit. on p. 229).

[172] L. Papp et al. "A high precision calorimeter for the SOX experiment". In: *Nucl. Instrum. Meth.* A824 (2016), pp. 699–700 (cit. on p. 230).

[173] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Ann. Math. Statist.* 9.1 (Mar. 1938), pp. 60–62. URL: http://dx.doi.org/10.1214/aoms/1177732360 (cit. on p. 234).

[174] Fengpeng An et al. "Neutrino Physics with JUNO". In: *J. Phys.* G43.3 (2016), p. 030401. arXiv: 1507.05613 [physics.ins-det] (cit. on p. 260).

Acknowledgments

> How wonderful that we have met with a paradox. Now we have some hope of making progress.
>
> Niels Bohr

Non potrei iniziare questa difficile sezione se non con Barbara, per il suo insostituibile supporto, per la fiducia riposta in me e per il tempo che mi ha dedicato in questi anni (senza dimenticare la pazienza in fase di scrittura!). Mi sento in debito con lei e con la sua incredibile, sincera passione per la ricerca.

Il microcosmo Borexiniano è decisamente variegato, e sarebbe impossibile citare esaustivamente tutte le persone con cui ho condiviso una parte di questo percorso. Grazie ad Ale, compagna di ufficio, caffè, produzioni Monte Carlo, pranzi e laureandi. A David, per aver condiviso ufficio, Giappone e conoscenze di fluidodinamica. Un grazie a tutto il gruppo milanese, alias le persone che popolano - o hanno popolato - il terzo piano del LITA: Emanuela, Gianpaolo, Gioacchino, Lino, Davide, Paolo, Augusto, Luca, Andrey e Fede. I feel lucky to be part of the *bxnusol* analysis group, I have really enjoyed working with all of you: thanks to all the Juelich people, tireless and truly welcoming guys, so thanks to Livia, Zara, Mariia, Omer and Sindhu; e grazie ai *gransassini* Nicola, Xuefeng, Chiara. Grazie a Daniele per le utili discussioni di questi mesi e per avermi introdotto nel mondo del SSM. E ovviamente ad Alessio: mi sembra ieri quando mi spiegavi come installare il Monte Carlo su CNAF! E infine a Simone, un esempio da imitare, da sempre.

Vorrei ringraziare inoltre i referee di questo lavoro, Oliviero Cremonesi e Aldo Serenelli, per l'attenzione che hanno dedicato nella lettura e per i consigli ben calibrati, che mi hanno spinto a chiarificare l'esposizione di diverse sezioni critiche.

Via Celoria 16 è ormai diventata per me una sorta di seconda casa, nella quale sono cresciuto anagraficamente e non. Grazie al poliedrico Lorenzo, aka Caccia, dal quale ho sempre da imparare. E a Gigi: tutti gli esercitatori vorrebbero essere affiancati a docenti di tale calibro umano e professionale. Grazie anche a Marco Grassi, per le discussioni su JUNO di questi ultimi mesi e per gli utili suggerimenti. E a Xavi, per un paio di chiacchierate rivelatesi cruciali per la mia crescita.

Al di fuori del mondo accademico, avrei diverse persone da citare, alle quali devo una parte di questi tre anni. Grazie a tutti gli Zippari: anche crescendo, tra part-time e modelle varie, il nostro livello di idiozia continua pericolosamente ad aumentare, com'è giusto che sia. E un grazie a Giorgio, compagno di concerti nonchè una delle persone più leali che io conosca. A Erika e alla sua pazienza stoica. E anche ad E. e S.: senza di voi, non avrei capito tantissime cose. A Francesca, che mi ricorda l'entusiasmo di essere studente. Infine grazie a Virginia, per tante, troppe cose.

Infine, un enorme e fondamentale grazie va alla mia famiglia, e in particolare ai miei genitori. Senza il vostro supporto, nessun passo di questo percorso sarebbe stato possibile.