GRAN SASSO SCIENCE INSTITUTE
Scuola Universitaria Superiore

G S
S I

SISSA

Cycle XXXI - A. A. 2015–2018

First simultaneous measurement of low-energy pp-chain solar neutrinos and prospects for CNO neutrino detection in Borexino

PhD THESIS
Submitted on July 11th 2019

Candidate:

Daniele Guffanti

Advisors:

Dr. Matteo Agostini
Technische Universität München

Dr. Nicola Rossi
INFN Laboratori Nazionali del Gran Sasso

L'AQUILA, ITALY, JULY 2019

It does not do harm to the mystery to know a little about it.

— Richard Feynman

ABSTRACT

This thesis presents the first simultaneous measurement of low energy neutrinos produced in the Sun through the chain of reactions started from the proton–proton fusion (*pp* chain) and detected by the Borexino experiment.

The Borexino experiment uses about 300 tons of liquid scintillator as target material. Solar neutrinos interact with the electrons of the scintillator and the energy released in the interaction is converted into scintillation light and detected by photo-multipliers. The measurement presented in this thesis is based on a multivariate analysis that exploits the energy, position and pulse shape distribution of the events in order to constrain the background and signal rates. To perform the analysis, new software tools and procedures have been developed, benchmarked and optimized in the framework of this thesis work. The sensitivity and systematic uncertainties have been studied using ensembles of pseudo data produced with Monte Carlo techniques considering all physical processes involved in the neutrino interaction, in the detector response, and in the read-out electronics.

The results presented in this thesis improves previous Borexino findings. The interaction rate of neutrinos produced in the *pp* reaction which drives the energy production in the Sun has been measured to be $134 \pm 10(\text{stat})^{+6}_{-10}(\text{sys})$ counts per day per 100 ton of scintillator (cpd/100 t). The rate of neutrinos produced by the electron capture on ^{7}Be has been constrained to $48.3 \pm 1.1^{+0.4}_{-0.7}$ cpd/100 t that is a factor 2 more accurate than the current prediction of the models that describe the structure and evolution of the Sun. The presence of neutrinos from the *pep* reaction has been established with a 5σ significance, its rate being $2.43 \pm 0.36^{+0.15}_{-0.22}$ cpd/100 t ($2.65 \pm 0.36^{+0.15}_{-0.24}$ cpd/100 t) assuming an high (low) fraction of metals in the Sun.

Under mild assumptions on the nuclear physics, an upper limit on the rate of neutrinos produced in the cycle of nuclear reactions catalyzed by C, N, and O (CNO cycle) has been constrained below 8.1 cpd/100 t at 95% C.L. The sensitivity to a CNO signal is primarily limited by ^{210}Bi and *pep* events that have an almost indistinguishable energy distribution. To improve the sensitivity an independent analysis has been designed to constrain the rate of ^{210}Bi and *pep*, breaking the correlation with the CNO signal. The requirements on the constraints have been defined through an extensive sensitivity study based which shows that under plausible assumptions on the actual CNO neutrino flux, an uncertainty of 1.75 cpd/100 t in the determination of the ^{210}Bi background would result in a 3σ median significance.

In this connection, a strategy for estimating the ^{210}Bi background level based on the measurement of the ^{210}Po daughter has been developed, tested and applied to real data, showing that the precision requested is achievable, and it is currently under consideration within the Borexino Collaboration.

The measurements presented in this thesis along with an independent estimate of ^{8}B neutrinos also based on Borexino data have been used to perform

frequentist and Bayesian hypothesis testing between the predictions of the two most popular solar models. The results show an excellent agreement with the temperature profile expected assuming that the Sun has a high fraction of metals, that is favoured with respect to the predictions based on a lower metal content with odds 5 : 1 or, equivalently, Bayes factor 4.9. Assuming the fluxes predicted by the solar models it is possible to compute the electron neutrino survival probability and study the properties of neutrino oscillation. Borexino measurements establish the existence of matter effect at 98.2% C.L.

The results presented in this thesis are the most accurate estimates of all the low-solar neutrino fluxes to date and provide the foundation of the future measurement of CNO that can be achieved thanks to innovative methods proposed to constrain the background.

CONTENTS

INTRODUCTION

Solar neutrinos have been driving major scientific discoveries for the last 50 years, both on the fundamental properties of neutrinos and for the understanding of the Sun. For instance, neutrino flavour oscillations, which proved that neutrinos are not massless as expected from the Standard Model of Particle Physics, were first experienced and later discovered in solar neutrino experiments. Solar neutrinos provided also the ultimate proof of the nuclear origin of the Sun power. The Sun produces energy by fusing four protons into a helium nucleus through a series of nuclear processes taking place in the Sun core: $\approx 99\%$ of the Sun energy is generated in a chain of reactions starting with the fusion of two protons (*pp*-chain), while the remaining $\approx 1\%$ is produced in a loop of nuclear reactions catalysed by the presence of carbon, nitrogen and oxygen (CNO cycle). Thanks to their small cross section, neutrinos produced in these processes can escape the Sun interior and reach the Earth as direct messenger of the nuclear reactions happening in the Sun core.

The Borexino experiment at the INFN Laboratori Nazionali del Gran Sasso (LNGS) has been the first experiment able to perform low-energy solar neutrino spectroscopy. Measuring the energy released by solar neutrinos allows to separate the contributions of neutrinos produced in the different nuclear reactions, thus probing the Sun engine with an unprecedented detail. Borexino started the data taking in 2007. Neutrinos are detected via their elastic scattering with the electrons of 278 tons of liquid scintillator. The key of the success of Borexino is the impressive low-background level which allows to be sensitive to event rates as low as a few counts per day per 100 ton of liquid scintillator. In the first phase of the experiment (2007–2010) Borexino measured the solar neutrinos produced in the ^{7}Be, *pep*, ^{8}B reactions of the *pp*-chain. A purification campaign that further reduced the background level was performed between 2010-2011, proceeding the second phase of the experiment whose data are analysed in this thesis.

This thesis work presents the most accurate determination of low-energy neutrinos emitted in the *pp*-chain based on Borexino Phase II data along with the prospects for the first detection of the CNO neutrinos in Borexino. The dissertation is structured as follows. Chapter 1 introduces the main concepts of neutrino phenomenology and solar physics required to interpret solar neutrino measurements and reviews the main experimental achievements, including the results of Borexino prior to this thesis work. The main features of the Borexino detector are described in Chap. 2 along with the procedures for the reductions of backgrounds. Chapter 3 describes the development of a multivariate analysis of Borexino data which exploits the energy, position and topology of the events to constrain the background and neutrino interaction

rates. The sensitivity of the analysis is covered in Chap. 4 where the expected statistical uncertainties systematics due to the fit model are discussed. Chapter 5 presents the measurement of the low-energy solar neutrinos interaction rate performed analysing Borexino Phase II data. The results have been combined with an independent measurement of ^{8}B neutrinos (also performed by Borexino) to test the effect of matter on neutrino oscillations and predictions of different solar models.

The second part of the thesis is focused on the search for neutrinos emitted in the CNO cycle, that has never been detected yet. The sensitivity of the analysis to CNO neutrinos is limited by the presence of ^{210}Bi and *pep* which can mimic its the signal. A strategy for the first detection of CNO neutrinos is outlined in Chap. 6. A detailed sensitivity study shows the impact of background constraints on the sensitivity to the CNO neutrino signal, and different approaches to achieve an independent assessment of the background rates are discussed. Chapter 7 details the progresses in the determination of the ^{210}Bi background through the measurement of the ^{210}Po daughter that is supposed to be in equilibrium but whose spatial density is influenced by convective motion in the liquid scintillator due to instabilities in the temperature gradient.

The results presented in this thesis contributed to a comprehensive measurement of the *pp*-chain solar neutrinos performed by Borexino which provides the most accurate determination of the low-energy neutrino fluxes to date. The strategy discussed to achieve the first detection of CNO neutrinos set the basis for the next, ambitious goal of the experiment.

CHAPTER 1

SOLAR NEUTRINOS

The study of solar neutrinos has been one of the driving forces of the progress of neutrino and solar physics in the last 50 years. This fascinating field stands at the crossroad between particle physics, nuclear physics and astrophysics and its development required remarkable joint efforts by these three communities. This chapter introduces the main concepts of neutrino and solar physics needed to interpret solar neutrino measurements. Section 1.1, after a short introduction of the neutrinos within the framework of the Standard Model of Particle Physics, presents the basic phenomenology of neutrino oscillation, both in vaccum and in matter. The working principles of the Sun, encapsulated in the so-called Standard Solar Model, are described in Sec. 1.2 along with the processes producing solar neutrinos; the predictions of the Standard Solar Model that were tested experimentally and gave rise in the recent years to the solar abundances controversy are also discussed. The propagation of solar neutrinos from the Sun core to the Earth is described in Sec. 1.3, while Sec. 1.4 briefly reviews the main experimental achievements of solar neutrino physics to date. Finally, prospects for the new measurements of solar neutrinos are discussed in Sec. 1.5.

1.1 INTRODUCTION TO NEUTRINO OSCILLATION

This Section summarizes the fundamental concepts of neutrino phenomenology that are needed to describe the behaviour of the neutrinos produced in the Sun. After a short introduction on the role of neutrinos in the current framework of elementary particle physics (Sec. 1.1.1), Sec. 1.1.2 presents the theory of neutrino oscillations in vacuum and Sec. 1.1.3 describes the effect on the flavour transition of the matter potential experienced by neutrinos.

1.1.1 *Neutrinos in the Standard Model of Particle Physics*

In contrast to most of the particles discovered in the first, chaotic years at the dawn of particle physics, neutrinos were theoretically postulated before their first experimental observation. The observation of the continuous β spectrum made by Meitner and Hann in 1911 and later confirmed by Chadwick in 1914

seemed to suggest that contrarily to the α and γ decays, energy was not conserved in the β decay process [1–3].

In order to save the energy conservation law, in 1930 Wolfgang Pauli suggested, in his famous letter addressed to the participants of a meeting in Tubingen, "a desperate remedy": the existence of a new, extremely light and hardly detectable neutral particle with spin $1/2$ that is also emitted in the β decay and accounts for the apparent missing energy.

The subsequent progress due to Fermi, the four-fermion β decay Hamiltonian introduced in 1934 [4, 5], turned the initial guess of Pauli into a predictive theory, which became more and more acceptable to the community and that culminated with the first observation of anti-neutrinos at the Savannah River reactor by Cowan and Reines in 1953 [6, 7].

Since then, the understanding of particle physics has made impressive progress. Thanks to many fundamental experiments and to the theoretical work of Glashow, Salam, Iliadis, Higgs, Weinberg, Feynman, Gell-Mann and many others, a comprehensive theory embedding the entire knowledge of particle physics has been developed. The strong, weak and electromagnetic forces are described by so-called Standard Model (SM) of Particle Physics, which has proved itself as an incredible successful theory being able to explain almost all the experimental results obtained so far and to predict a variety of phenomena.

In the Standard Model, twelve spin-$1/2$ elementary particles are present [8]: six quarks which are sensitive to the strong interaction, and six leptons that are not. Both quarks and leptons are paired in three *generations* or *families*; in the leptonic sector, each family is composed by a charged lepton (e^-, μ, τ) and a neutrino with the corresponding flavour (ν_e, ν_μ, ν_τ).

Neutrinos, being electrically neutral, are subjected only to the weak interaction. Already in 1956 Wu showed that the parity is violated in weak processes where charge is exchanged [9], indicating that only (anti)particles in the (right)left-handed helicity state can interact. This result was verified in 1958 by Goldhaber who demonstrated that neutrinos are produced mostly left-handed [10], leading to the idea that right-handed neutrinos are not present in nature.

In the SM, the weak interaction is described jointly with the electromagnetic force [11, 12] by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry group. The $\mathrm{SU}(2)_L$ denotes the symmetry under weak isospin (T) transformations, where the index L reminds that the group acts on the left-handed chiral components of the fermion field. On the other hand, $\mathrm{U}(1)_Y$ describes the hypercharge, that allows us to write the electric charge generator Q as $Q = T_3 + Y/2$, where Y and T_3 are two diagonal matrices, the former proportional to unity and the latter being the third component of the weak isospin operator. The left-handed leptons are arranged in isospin doublets of the form (ν_ℓ, ℓ) which have the same hypercharge, while right-handed components are singlets under $\mathrm{SU}(2)$. Requiring the theory to be invariant under local gauge transformation leads to the introduction of four massless bosonic fields: an isospin triplet of vector bosons associated to the generators of $\mathrm{SU}(2)$ and one additional vector boson for the $\mathrm{U}(1)$ group that is

an isospin singlet. Once properly arranged, these terms represent the photon and the carriers of the weak force, the charged $W^\pm$ and the electrically neutral Z^0. The very large mass of the weak interaction bosons ($m_W = 80.4\,\text{GeV/c}^2$, $m_Z = 91.2\,\text{GeV/c}^2$ [8]) is generated by the spontaneous symmetry breaking induced by the Higgs field and results in a very short interaction range that historically motivated the name *weak* referred to this force.

The Higgs field, along with the weak vector bosons, also gives mass to all the SM fermions through the Yukawa coupling of the right handed fermion with its left-handed doublet and the Higgs field, and since right-handed neutrinos are not foreseen, neutrinos are left massless in the SM.

Neutrinos interact with the other particles of the SM through the exchange of $W^\pm$ and Z^0 bosons. The interactions mediated by the $W^\pm$ are called charged-current (CC) interactions and couple neutrinos to their isospin partners. The process is described by the interaction of the charged leptonic weak current $j_{W,L}$ with the W field in the Lagrangian

$$\mathscr{L}^{\text{CC}} = -\frac{g}{2\sqrt{2}} j^\mu_{W,L} W_\mu + \text{h.c.} \quad \text{with } j^\mu_{W,L} = 2 \sum_{\alpha}^{e,\mu,\tau} \overline{\nu_{\alpha L}} \gamma^\mu \ell_{\alpha L} \qquad (1.1)$$

where $\nu_{\alpha L}$ ($\ell_{\alpha L}$) is the left-handed fermion field describing the α neutrino (charged lepton), g is a coupling constant, γ^μ represents the Dirac matrices and h.c. stands for the hermitian conjugate. The charge-current interactions are responsible for the β decay and were directly observed in the neutrino sector contextually to the neutrino discovery in 1956.

The neutral-current (NC) process is mediated by the Z^0 boson and is described by the interaction of the weak leptonic neutral current $j_{Z,L}$ with the Z_μ field in the Lagrangian

$$\mathscr{L}^{\text{NC}} = -\frac{g}{2\cos\theta_W} j^\mu_{Z,L} Z_\mu$$
$$\text{with} \quad j^\mu_{Z,L} = \sum_{\alpha}^{e,\mu,\tau} \left(g^\nu_L \overline{\nu_{\alpha L}} \gamma^\mu \nu_{\alpha L} + g^\ell_L \overline{\ell_{\alpha L}} \gamma^\mu \ell_{\alpha L} + g^\ell_R \overline{\ell_{\alpha R}} \gamma^\mu \ell_{\alpha R} \right) \quad (1.2)$$

where the definition of the neutral current reflects the fact that the Z^0 is defined as a superposition of two fields introduced imposing the local gauge invariance under SU(2) and U(1). Therefore, the weak neutral current involves also right handed fermions with a strength given by the coefficients $g^{\nu/\ell}_{L/R}$ that is proportional to the mixing among the two original fields represented by the Weinberg angle θ_W. The neutral current interactions were observed for the first time by the Gargamelle experiment at CERN in 1973 [13].

1.1.2 *Neutrino Oscillation in vacuum*

Soon after the discovery of anti-neutrinos from nuclear reactors in 1956, physicists tried to measure neutrinos from a variety of other sources. In particular,

experiments measuring neutrinos produced by nuclear reactions in the Sun or generated in cosmic ray showers in the atmosphere gave puzzling results, often resulting in a flux of neutrinos substantially smaller than the expectations. In fact, contrarily to what was predicted by the SM, neutrinos do have a small mass and the eigenstates of the weak interaction are not the same mass eigenstates involved in the propagation, making the probability of the neutrino to appear in a flavour different that its original one to oscillate while travelling. This phenomenon was hypothesized already in the late 1950s by Pontecorvo in analogy with the K^0–$\bar{K}^0$ oscillation in the quark sector [14, 15], but was demonstrated experimentally only between 1998 and 2001.

In this framework it is possible to express the flavour eigenstates $|\nu_\alpha\rangle$ as a combination of Hamiltonian eigenstates $|\nu_k\rangle$, which can be distinguished by the Greek and Latin label denoting flavour and mass eigenstates respectively. Each Hamiltonian eigenstate is characterized by a specific mass m_k

$$\mathscr{H}|\nu_k\rangle = E_i|\nu_k\rangle \quad \text{where } E_i = \sqrt{\boldsymbol{p}^2 + m_k^2} \tag{1.3}$$

The flavour state $|\nu_\alpha\rangle$ is described by

$$|\nu_\alpha\rangle = \sum_{i=0}^{3} U_{\alpha i}^* |\nu_i\rangle \tag{1.4}$$

where the mixing matrix U is a 3×3 unitary matrix[1] named after Pontecorvo, Maki, Nakagwa and Sakata (PMNS).

Following from Eqs. (1.4) and (1.3), the time evolution of the flavour state $|\nu_\alpha(t)\rangle$ is

$$|\nu_\alpha(t)\rangle = \sum_k U_{\alpha k}^* \mathrm{e}^{-iE_k t} |\nu_k\rangle \tag{1.5}$$

Exploiting the unitarity of U, it is possible to express a Hamiltonian state as a combination of flavour states

$$|\nu_k\rangle = \sum_\alpha U_{\alpha k} |\nu_\alpha\rangle \tag{1.6}$$

Thus, substituting Eq. (1.6) in Eq. (1.5), the evolution of the flavour state $|\nu_\alpha\rangle$ becomes

$$|\nu_\alpha(t)\rangle = \sum_\beta^{e,\mu,\tau} \left(\sum_{k=1}^{3} U_{\alpha k}^* \mathrm{e}^{-iE_k t} U_{\beta k} \right) |\nu_\beta\rangle \tag{1.7}$$

Equation (1.7) shows that if the mixing matrix is not diagonal, a pure flavour state $|\nu_\alpha\rangle$ at $t = 0$ becomes a superposition of different flavour states evolving

---

1 The dimension of U derives from the fact that to describe three independent flavour states, at least three independent mass states must exist. However, if more non-interacting neutrino flavours exist, then more mass eigenstates should be considered and the 3×3 submatrix describing the mixing of the three interacting neutrinos would not unitary.

for $t > 0$. The probability of measuring a flavour $|\nu_\beta\rangle$ after a time t is given by the squared modulus of the transition amplitude $|\langle\nu_\beta|\nu_\alpha(t)\rangle|^2$, which is

$$P_{\nu_\alpha\to\nu_\beta}(t) = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 = \sum_{k,j} U_{\alpha k}^* U_{\beta k} U_{\alpha j} U_{\beta j}^* e^{-i(E_k - E_j)t} \qquad (1.8)$$

In the ultra-relativistic case $E_k \approx E + m_k^2/2E$ where $E = |\boldsymbol{p}|$ and the propagation time t can be approximated by the travelled distance L, leading to

$$\begin{aligned}
P_{\nu_\alpha\to\nu_\beta}(E, L) &= \sum_{k,j} U_{\alpha k}^* U_{\beta k} U_{\alpha j} U_{\beta j}^* \exp\left(-i\frac{\Delta m_{kj}^2 L}{2E}\right) \\
&= \delta_{\alpha\beta} - 4\sum_{k>j} \Re\left[U_{\alpha k}^* U_{\beta k} U_{\alpha j} U_{\beta j}^*\right] \sin^2\left(\frac{\Delta m_{jk}^2 L}{4E}\right) \qquad (1.9) \\
&\quad + 2\sum_{k>j} \Im\left[U_{\alpha k}^* U_{\beta k} U_{\alpha j} U_{\beta j}^*\right] \sin\left(\frac{\Delta m_{jk}^2 L}{2E}\right)
\end{aligned}$$

where Δm_{kj}^2 denotes the squared–mass difference $m_k^2 - m_j^2$ and the natural units are substituted in the argument of the sin function.

Neutrino oscillations were proven jointly by the measurement of solar and atmospheric neutrinos performed respectively by the SNO and Kamiokande experiments, that will be briefly described in Sec. 1.4. Since then, this phenomenon was studied in many different conditions using a variety of neutrino sources, obtaining a very accurate measurement of both the mixing parameters and the mass-squared difference of the Hamiltonian eigenstate, but not the absolute values of the neutrino masses which still is one of the main open questions in neutrino physics.

The many different experiments that studied neutrino oscillations in the last twenty years are justified by the fact that the oscillation parameters do not allow a single experiment to be sensitive enough to *all* the oscillation parameters. The squared–mass difference responsible for the oscillation of solar neutrinos is indeed $\Delta m_{12}^2 = 7.53 \pm 0.18 \times 10^{-5}\,\text{eV}^2$, while the ones causing the oscillation of atmospheric neutrinos are $\Delta m_{31}^2 \approx \Delta m_{32}^2 = 2.44 \pm 0.06 \times 10^{-3}\,\text{eV}^2$. In order for an experiment to be sensitive to oscillation, it should consider neutrinos with energy E and baseline L such that $\Delta m^2 L/4E \approx 1$. In order to estimate the L/E ratio that verifies this condition, the phase factor in Eq. (1.9) can be conveniently expressed replacing the natural units

$$\frac{\Delta m_{kj}^2 L}{4E} = 1.27\frac{\Delta m_{kj}^2[\text{eV}^2]L[\text{m}]}{E[\text{MeV}]} \qquad (1.10)$$

Thus, when an experiment is sensitive to the "atmospheric" mass splitting, the L/E ratio is not large enough to give a substantial effect when combined with the "solar" mass splitting ($\Delta m_{sol}^2 L/4E \ll 1$). Conversely, when the experiment L/E is set to be sensitive to the small solar mass splitting, the oscillation phase is $\Delta m_{\text{atm}}^2 L/4E \gg 1$ and therefore is averaged to $\langle\sin^2 x\rangle = 1/2$.

In many cases it is possible to simplify the three neutrino scenario and describe the oscillation with a simpler effective model considering two neutrinos only. In this framework, the mixing matrix U can be expressed as a rotation depending on one mixing angle only

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \text{with } 0 \leq \theta \leq \frac{\pi}{2} \tag{1.11}$$

and the transition probability in Eq. (1.9) becomes

$$P_{\nu_\alpha \to \nu_\beta}(E, L) = \delta_{\alpha\beta} - (2\delta_{\alpha\beta} - 1)\sin^2 2\theta \sin^2\left(\frac{\Delta m_{kj}^2 L}{4E}\right) \tag{1.12}$$

It is possible to show that a generic unitary matrix can be described by a similar parametrization. In the general case of n families and considering neutrinos as Dirac particles like the rest of the Standard Model fermions[2], a unitary matrix is described by $1/2 n(n-1)$ mixing angles and $1/2(n-1)(n-2)$ phase factors. Hence, in the three neutrino framework of the Standard Model, the PMNS matrix can be expressed as

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \times \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.13}$$

with c_{ij} and s_{ij} indicating $\cos\theta_{ij}$ and $\sin\theta_{ij}$ respectively, while δ is a phase factor that, if different from 0, breaks the CP symmetry.

With an analogous treatment, the results derived in this Section can be obtained also for the anti-neutrino case. The only difference is that to express the anti-neutrino flavour state as a superposition of anti-neutrino mass eigenstates, the weights are given by the complex conjugate of the ones used in Eq. (1.4) ($U_{\alpha k}$), which causes the change of the sign of the term depending on the imaginary part of the mixing matrix product in Eq. (1.9).

1.1.3 *Neutrino oscillation in matter*

In the previous section it was shown that the evolution of the flavour state is determined by the system Hamiltonian. In case neutrinos propagate in ordinary matter, the interaction potential affects ν_e differently from $\nu_{\mu,\tau}$ since the charge-current interaction with the electrons of the medium involves only ν_e. This difference in the interaction potential reflects into the Hamiltonian and therefore impacts on the neutrino flavour oscillation.

---

2 If neutrinos and anti-neutrinos were the same particle just with different helicity, then they would be Majorana particles. Among the deep implications of this fact, two additional phase factors appear in the mixing matrix.

The ν_e–e^- charge-current interaction is described by the Lagrangian

$$\mathscr{L}^{\mathrm{CC}} = -\frac{g^2}{8} j^\dagger_{W,L;\mu} G^{\mu\nu}_{(W)} j^\nu_{W,L} \tag{1.14}$$

where $G^{\mu\nu}_{(W)}$ is the W boson propagator that is given by

$$G^{\mu\nu}_{(W)}(q) = \frac{-g^{\mu\nu} + \frac{q^\mu q^\nu}{m_W^2}}{q^2 - m_W^2 + \mathrm{i}\varepsilon} \tag{1.15}$$

with q being the gauge boson momentum. In the energy range typical for the neutrinos produced in nuclear reactions, $q \ll m_W^2$, thus the gauge boson propagator can be approximated by $G^{\mu\nu}))(W) \approx \mathrm{i}g^{\mu\nu}/m_W^2$. Substituting in Eq. (1.14) one gets

$$\mathscr{L}^{\mathrm{CC}} = -\frac{G_F}{\sqrt{2}} j^\dagger_{W,L;\mu} j^\mu_{W,L} \tag{1.16}$$

where G_F is the Fermi constant that is defined as $g^2/(8m_W^2)$ and is $1.6637(1) \times 10^{-5}\,\mathrm{GeV}^{-2}$. In order to compute the charge-current potential of neutrinos propagating in an electron gas, one should evaluate the average over the distribution of the electrons in the medium rest frame of the Hamiltonian density derived from Eq. (1.16)

$$\mathscr{H}^{(CC)}(x) = \frac{G_F}{\sqrt{2}} \overline{\nu_e}(x)\gamma^\mu(1-\gamma^5)\nu_e(x)\langle \overline{e}(x)\gamma_\mu(1-\gamma^5)e(x)\rangle \tag{1.17}$$

where the term in $\langle \dots \rangle$ indicates the average over the electron momentum and spin states. It can be shown that the only non-vanishing term is

$$\mathscr{H}^{(CC)}(x) = V_{CC}(x)\overline{\nu_{eL}}(x)\gamma^0 \nu_{eL}(x) \tag{1.18}$$

where $V_{CC}(x) = \sqrt{2}G_F n_e(x)$ is the charge-current potential and $n_e(x)$ is the electron density in x. With a similar approach one can compute the potential of the neutral-current interaction, but since it equally affects all the flavours it will result in a coherent phase factor that does not affect the oscillation.

The total Hamiltonian governing the propagation is then

$$\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_I \tag{1.19}$$

where $\mathscr{H}_0$ is the vacuum Hamiltonian in Eq. 1.3 and $\mathscr{H}_I$ includes the effect of the interaction with matter

$$\mathscr{H}_0 \ket{\nu_\alpha} = V_\alpha \ket{\nu_\alpha} \quad \text{with} \quad V_\alpha = V_{CC}\delta_{\alpha e} \tag{1.20}$$

The evolution of a neutrino state with initial flavour α is described by the Schrödinger equation

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\ket{\nu_\alpha(t)} = \mathscr{H}\ket{\nu_\alpha(t)} \tag{1.21}$$

from which one can show that the transition amplitude $\langle \nu_\beta | \nu_\alpha(t) \rangle = \psi_{\alpha\beta}(t)$ evolves as

$$i\frac{d}{dx}\psi_{\alpha\beta} = \sum_\eta \left(\sum_k U_{\beta k} \frac{\Delta m_{k1}^2}{2E} U_{\eta k}^* + \delta_{\beta e}\delta_{\eta e}V_{CC} \right)\psi_{\alpha\eta}(x) \qquad (1.22)$$

where U is the same mixing matrix introduced in Eq. (1.4) and E is the neutrino energy. In deriving Eq. (1.22) the ultra-relativistic approximation $t \sim x$ was used and a term generating a phase common to all the flavours was neglected.

Considering the simplest case of mixing among two neutrino components ν_e, ν_μ and ν_1, ν_2, Eq. (1.22) can be conveniently expressed in matrix form as

$$i\frac{d}{dt}\begin{pmatrix} \psi_{ee} \\ \psi_{e\mu} \end{pmatrix} = \frac{\Delta m_M^2}{4E}\begin{pmatrix} -\cos 2\theta_M & \sin 2\theta_M \\ \sin 2\theta_M & \cos 2\theta_M \end{pmatrix}\begin{pmatrix} \psi_{ee} \\ \psi_{e\mu} \end{pmatrix} \qquad (1.23)$$

with Δm_M^2 and $\sin 2\theta_M$ defined as

$$\Delta m_M^2 = \sqrt{(\Delta m^2 \cos 2\theta - 2EV_{CC})^2 + (\Delta m^2 \sin 2\theta)^2}$$
$$\text{and} \quad \sin 2\theta_M = \frac{\Delta m^2 \sin 2\theta}{\Delta m_M^2} \quad (1.24)$$

that represent the effective mass splitting and mixing angle in matter[3]. The Hamiltonian matrix in Eq. 1.23 is diagonalized by the transformation

$$U_M = \begin{pmatrix} \cos\theta_M & \sin\theta_M \\ -\sin\theta_M & \cos\theta_M \end{pmatrix} \qquad (1.25)$$

which transforms the flavour states into the Hamiltonian eigenstates. Applying U_M to the transition amplitudes in Eq. 1.23, one finds that the matrix describing the evolution has off-diagonal terms equal to $i4Ed\theta_M/dx$ which allows the transition between mass eigenstates. However, if the electron density n_e is constant or varies smoothly (like in the case of the Sun interior), the so-called *crossing* probability P_C, *i.e.* the probability of oscillation between different mass eigenstates, is negligible for energies below $10\,\text{GeV}$.

One can notice that if the electron density varies along the propagation, when the condition

$$\Delta m^2 \cos\theta = 2E\sqrt{2}G_F n_e(x) \qquad (1.26)$$

is met, the effective mixing angle θ_M becomes $\pi/4$ causing maximal mixing between the flavour eigenstates. It should be noticed that such resonance can occur only if the vacuum mixing angle θ is $< \pi/4$ because otherwise $\cos 2\theta < 0$ and the interaction potential is always positive for ν_e. This mechanism is

---

3 Note that in the vacuum limit $n_e \to 0$ the effective mass splitting and mixing angle tend to the corresponding values in vacuum.

known as the MSW effect, from the names of Mikheev, Smirnow and Wolfenstein that first hypothesized this phenomenon [16, 17].

The transformation (1.25) has exactly the same form of the mixing matrix in Eq. (1.11), thus, if the electron density is constant, the oscillation probability for a ν_e travelling in matter within the two-flavour model has the same structure of the result obtained in vacuum (Eq. 1.12):

$$P_{\nu_e \to \nu_\mu}(x) = \sin^2 2\theta_M \sin^2\left(\frac{\Delta m_M^2 x}{4E}\right) \tag{1.27}$$

But if the electron density varies adiabatically along the neutrino path, the *survival probability* for electron neutrinos, *i.e.* $P_{\nu_e \to \nu_e}$, becomes

$$P_{\nu_e \to \nu_e}^{\text{adiabatic}}(x) = \frac{1}{2} + \frac{1}{2}\cos 2\theta_M^{(i)}\cos 2\theta_M^{(f)} +$$
$$\frac{1}{2}\sin 2\theta_M^{(i)}\sin 2\theta_M^{(f)}\cos\left(\int_0^x \frac{\Delta m_M^2(x')}{2E}\mathrm{d}x'\right) \tag{1.28}$$

where the (i) and (f) superscripts are used to indicate the effective mixing angle at the creation and detection point.

1.2 THE STANDARD SOLAR MODEL

The most accurate description of the Sun is the one provided by the Standard Solar Model (SSM), which embodies the entire knowledge of stellar and nuclear physics to reproduce the Sun evolution to match its present day properties [18]. The development of the Standard Solar Model was started by J. Bahcall in the early 1960s [19], motivated by the need of a calculation of the solar neutrino fluxes by R. Davis [20] that a few years later performed the first measurement of neutrinos emitted by the Sun [21]. Since the first generation of models many SSMs [22–34] were developed exploiting the improved description of the physical processes and more accurate observational constraints, still maintaining a crucial importance for solar neutrino experiments.

In this Section, the main aspects of the SSM are discussed. Sec. 1.2.1 presents the fundamental concepts underlying the model, while the nuclear reactions that power the Sun and generates solar neutrinos are described in Sec. 1.2.2. The main predictions of the SSM are presented in Sec. 1.2.3 together with the so-called *solar metallicity puzzle*, which is currently one of the main problems in the field of solar physics and significantly affects the predictions of the solar neutrino fluxes.

1.2.1 *Fundamental concepts*

The evolution of the Sun described by the SSM is defined by the considered physical processes as well as by the boundary conditions the evolving star is required to match in order to reproduce the present day status of the Sun.

In the SSM framework [18], the Sun produces its energy only through the fusion of four protons into a nucleus of ^{4}He via the series of nuclear processes which will be described in Sec. 1.2.2 and that are responsible for the production of solar neutrinos. Energy is created in the Sun core and it is transported by radiative processes until it reaches the outer part of the Sun ($\approx 0.71 R_\odot$) where convective motions occur. In the innermost region of the Sun the radiative energy transport is described calculating the plasma opacity which depends on its temperature, density and chemical composition. Since the Sun is stable, the SSM also assumes that the radiative and particle pressure exactly compensates the gravitational force establishing a situation of hydrostatic equilibrium.

In order to describe the Sun, the evolution of a chemically homogeneous gas cloud with mass $M_\odot$ is computed for a time $\tau_\odot$ equal to the age of the Sun, during which the mass loss is considered negligible. The SSM requires the evolved star to match the current solar luminosity $L_\odot$, radius $R_\odot$ and metal to hydrogen abundance $(Z/X)_\odot$ observed in the photosphere [35]. In order to satisfy these conditions the initial fraction of helium and metals (Y_i, Z_i), together with the parameter α_{MLT} from the mixing length theory, are adjusted. Oversimplifying the scenario, one can say that $L_\odot$ is related to the initial helium concentration Y_i, the α_{MLT} parameter is related to the Sun radius $R_\odot$, and the initial metal abundance Z_i and the current photosphere metal fraction are expected to be equal once the effect of diffusion is taken into account. This is however a crude approximation, since the above quantities are correlated with each other [35].

Once the parameters Y_i, Z_i and α_{MLT} are calibrated, the SSM describes the evolution of the Sun from its formation up to now (and possibly beyond). It is worth to stress at this point that the SSM is of extremely importantce not only for solar physics but for astrophysics in general, since it is the benchmark for every stellar evolution model that relies on the evaluation of α_{MLT} in the Sun.

1.2.2 *Energy production in the Sun*

The idea that the Sun is powered by nuclear reactions dates back to the early 1920s [36, 37] but it was only after the discovery of the tunnel effect by Gamow in 1928 [38] that the concept was developed to provide an accurate description of the energy production in stars. In the following years many authors explained the solar power with the conversion of four protons into one Helium nucleus, which releases an energy of 26.7 MeV [39–41]. This conversion can take place in two ways, both suggested in the late 1930s. The first, suggested by H. Bethe in [42], is the proton–proton (*pp*) fusion chain, which is a series of nuclear processes started by the combination of two protons into a deuteron. The second mechanism, itroduced independently by Weitzsäcker [43, 44] and Bethe [45], is the so-called CNO cycle where the reaction is catalysed by the presence of ^{12}C. These two processes are described in Sec. 1.2.2.1 and 1.2.2.2 with particular attention on the production of neutrinos.

In principle, the rate per unit volume of a reaction between nuclei A and B is [18]

$$R_{AB} = \frac{n(A)n(B)}{1 + \delta_{AB}} \langle \sigma v \rangle_{AB} \tag{1.29}$$

with n indicating the local density of the two nuclei, δ_{AB} avoids double counting in case of identical nuclei and $\langle \sigma v \rangle$ represents the product of the process cross section σ and the two nuclei relative velocity v averaged over the thermal distribution of the velocities.

The A and B nuclei in order to interact should overcome the repulsive Coulomb barrier, but the thermal energy in a stellar environment is usually much smaller. The reactions take place thanks to the tunnel effect which allows to cross the potential barrier with a probability depending on the width of the barrier at the energies under consideration.

The cross sections of the nuclear processes happening in the Sun are a fundamental ingredient of the Standard Solar Model. The measurement of these cross sections is not trivial since at such low energies the reaction rate is extremely small. This difficulty is partially overcome by performing the measurement in underground laboratories where the cosmic ray background is strongly suppressed [46]. When it is not possible to directly probe the energy range of interest, the measured cross sections are extrapolated or, when possible, *ab initio* calculations are performed. A comprehensive review of the cross sections used in the most recent Standard Solar Model can be found in [47].

1.2.2.1   *The pp chain*

The sequence of reactions of the *pp* chain is represented in Fig. 1.1. In the first step of the chain, two protons are combined to create a deuteron through the *pp* reaction

$$\boldsymbol{pp} \qquad p + p \rightarrow {}^2\text{H} + e^+ + \nu_e \qquad Q = 0.42\,\text{MeV} \tag{1.30}$$

which is the basis of the entire *pp* chain. The process is mediated by the weak interaction therefore its rate is significantly smaller than the one of reactions mediated by the strong interaction. Hence, the rate of the *pp* reaction basically determines the rate of all the following processes in the *pp* chain. For the same reason, the process cross section is too small to be measured in laboratory conditions, but it can be accurately computed from the theory of weak interactions [47].

The reaction releases an energy of $0.42\,\text{MeV}$ that is distributed among the products, leading to a continuous neutrino spectrum with maximum energy equal to the reaction Q value.

A deuteron can also be formed in the similar *pep* reaction

$$\boldsymbol{pep} \qquad p + e^- + p \rightarrow {}^2\text{H} + \nu_e \qquad Q = 1.44\,\text{MeV} \tag{1.31}$$

$$p + p \rightarrow {}^2\text{H} + e^+ + \boxed{\nu_e} \qquad p + e^- + p \rightarrow {}^2\text{H} + \boxed{\nu_e}$$

99.76%   0.24%

$${}^2\text{H} + p \rightarrow {}^3\text{He} + \gamma$$

84.6%   15.4% $2.5 \times 10^{-5}\%$

$${}^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He} + 2p \qquad {}^3\text{He} + {}^4\text{He} \rightarrow {}^7\text{Be} + \gamma \qquad \boxed{\text{hep } \nu}$$

99.89%   0.11%

$${}^7\text{Be} + e^- \rightarrow {}^7\text{Li} + \boxed{\nu_e} \qquad {}^7\text{Be} + p \rightarrow {}^8\text{B} + \gamma$$

$${}^7\text{Li} + p \rightarrow 2\text{He}^4 \qquad {}^8\text{B} \rightarrow {}^8\text{Be}^* + e^+ + \boxed{\nu_e}$$

$${}^8\text{Be} \rightarrow 2\,{}^4\text{He}$$

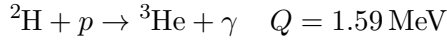**ppI**   **ppII**   **ppIII**   **hep**
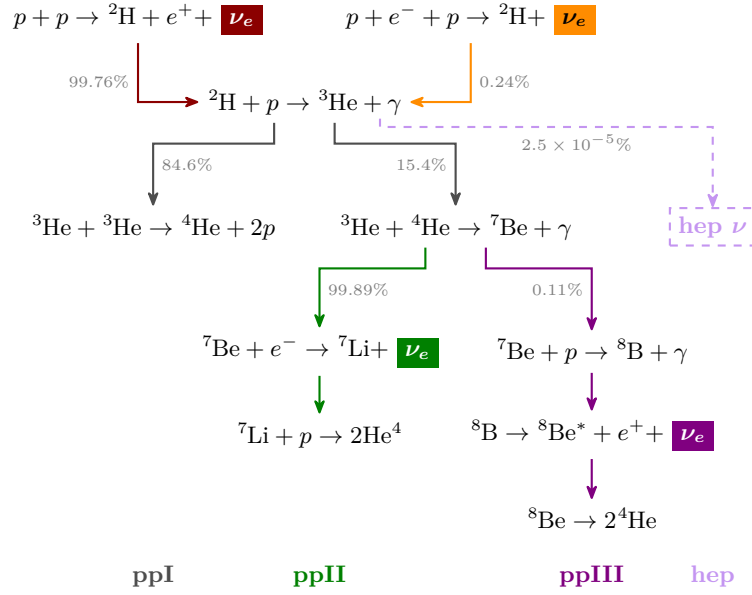
Figure 1.1: Sequence of nuclear processes in the proton–proton chain. The fraction close to the arrows represent the branching ratio of the process in the SSM. The different neutrinos emitted are highlighted with coloured boxes.

which differently from the *pp* reaction produces mono energetic neutrinos with an energy of $1.44\,\text{MeV}$. This process is ≈ 400 times less likely than the *pp* reaction, and their ratio is practically fixed by nuclear physics while the Sun temperature conditions only play minor a role.

Once the deuteron is produced, ${}^3\text{He}$ is formed by the absorption of a proton.

$${}^2\text{H} + p \rightarrow {}^3\text{He} + \gamma \quad Q = 1.59\,\text{MeV}$$

At this point, the *pp* chain can proceed in four different ways leading to four different branches of the chain depending on the interactions of ${}^3\text{He}$. The branches of the *pp* chain are named ppI, ppII, ppIII and hep and are briefly described in the following paragraphs.

The ppI branch In the first case, ${}^3\text{He}$ interacts with another ${}^3\text{He}$ nucleus in the Sun, forming a nucleus of ${}^4\text{He}$ and generating two free protons without emitting any neutrinos.

The ppII branch The ${}^3\text{He}$ nucleus can interact with a ${}^4\text{He}$ nucleus to produce a ${}^7\text{Be}$ nucleus. Beryllium-7 decays via electronic capture with the reaction

$$\mathbf{{}^7Be} \quad {}^7\text{Be} + e^- \rightarrow {}^7\text{Li} + \nu_e \quad Q = 0.862\,\text{MeV}$$
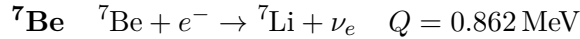
In $\approx 90\%$ of the cases ${}^7\text{Be}$ decays directly to the ground state of ${}^7\text{Li}$ emitting a mono energetic neutrino with $E_\nu = Q$, while in the remaining 10% it decays to the first excited state, emitting a $0.384\,\text{MeV}$ mono energetic neutrino followed by a $0.477\,\text{MeV}$ γ from the de-excitation of ${}^7\text{Li}^*$. The chain is then terminated with the production of two ${}^4\text{He}$ nuclei with the reaction ${}^7\text{Li}(p, \alpha) \uparrow 4He$.

The ppIII branch The first step of the ppIII termination of the *pp* chain is the same described for ppII, but in this case the ^{7}Be nucleus absorbs a proton producing a ^{8}B. This reaction is much less likely than the electron capture process, but it still is extremely relevant for neutrino physics, because ^{8}B is unstable and decays in

$$\mathbf{^8B} \quad ^8\text{B} \rightarrow \text{}^8\text{Be}^* + e^+ + \nu_e \quad Q = 15.5\,\text{MeV}$$
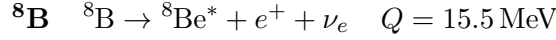
thus producing a neutrino with a continuous energy spectrum extended up to 15 MeV. The chain is then closed by the decay of the excited Beryllium-8 nucleus in two Helium nuclei.

The hep branch Another possibility is the interaction of the ^{3}He nucleus with a proton, giving

$$\mathbf{hep} \quad ^3\text{He} + p \rightarrow \text{}^4\text{He} + e^+ + \nu_e \quad Q = 19.8\,\text{MeV}$$
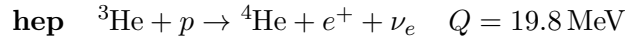
The process produces the *pp* chain highest energy neutrinos, but it is mediated by the weak interaction and thus it is much less probable than the other terminations of the chain.

1.2.2.2   *The* CNO *cycle*

The conversion of four protons into a ^{4}He nucleus that provides energy to the stars can be achieved also through a loop of nuclear reactions where protons are captured by Carbon, Nitrogen and Oxygen that act like catalysts. In principle, four different cycles of nuclear reactions involving different isotopes are possible in stellar environments, but in the case of the Sun, almost all the energy produced by the fusion catalysed by the CNO cycle is due to the CNO-I cycle sketched in Fig. 1.2a with a small contribution of the CNO-II, while the other cycles impact is negligible.

Involving heavier elements than the ones considered in the *pp* chain, the Coulomb barrier influencing the (p, α) and (p, γ) reactions is higher and induces a strong dependence on the local temperature of the star. At the typical temperature for Sun-like stars the energy production rate of the CNO cycle is $\propto T^{16.7}$, while the one of the *pp* chain is roughly $\propto T^4$ [48]. Therefore, if an adequate amount of CNO catalysts are present, the CNO cycle is the dominant production mechanism for high temperature and the *pp* chain is the dominant source of energy at low temperature as shown in Fig. 1.2b.

The CNO-I cycle The first CNO cycle is depicted on the left side of Fig. 1.2a. The crucial element for the CNO cycle is the presence of ^{12}C, which is the

(a) The CNO cycle I-II

(b) *pp*-chain and CNO energy production vs. *T*
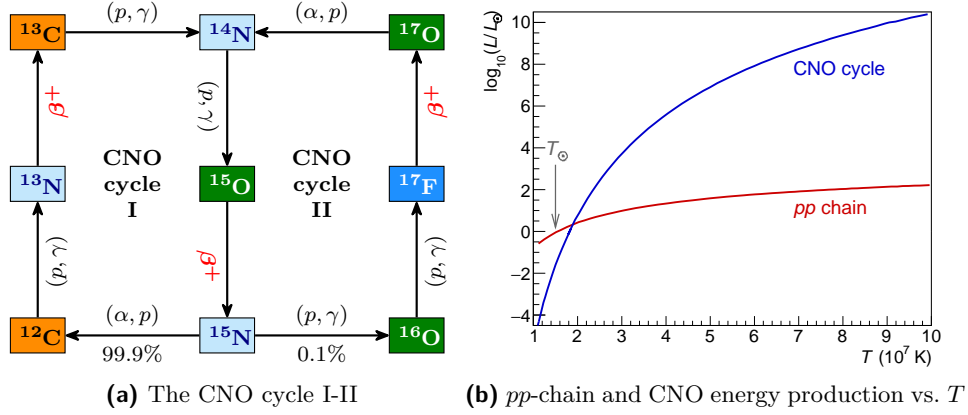
Figure 1.2: (a) Sequence of nuclear reactions of the CNO I-II bi-cycle. Neutrinos are emitted in the β^+ decays. **(b)** Stellar energy production rate as a function of the core temperature for the *pp*-chain and CNO cycle. The temperature of the Sun core $T_\odot$ is shown for reference. Figure elaborated from [47].

most abundant element heavier than He present in stars. The reactions loop is the following:

$$^{12}\text{C} + p \rightarrow {}^{13}\text{N} + \gamma \quad Q = 1.95\,\text{MeV} \tag{1.32}$$

$$^{13}\text{N} \rightarrow {}^{13}\text{C} + e^+ + \nu_e \quad Q = 1.20\,\text{MeV} \tag{1.33}$$

$$^{13}\text{C} + p \rightarrow {}^{14}\text{N} + \gamma \quad Q = 7.54\,\text{MeV} \tag{1.34}$$

$$^{14}\text{N} + p \rightarrow {}^{15}\text{O} + \gamma \quad Q = 7.35\,\text{MeV} \tag{1.35}$$

$$^{15}\text{O} \rightarrow {}^{15}\text{N} + e^+ + \nu_e \quad Q = 1.73\,\text{MeV} \tag{1.36}$$

$$^{15}\text{N} + p \rightarrow {}^{12}\text{C} + {}^4\text{He} \quad Q = 4.96\,\text{MeV} \tag{1.37}$$
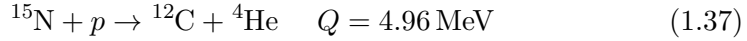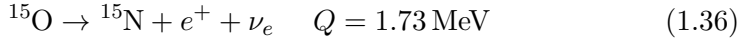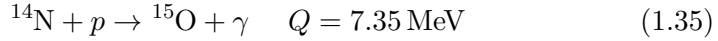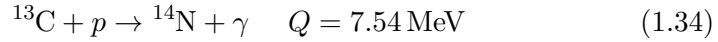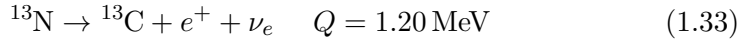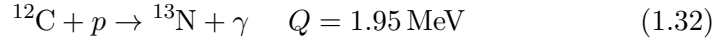
As for the *pp* chain, the rate of the CNO cycle reactions is determined by the slowest reaction of the sequence, which is $^{14}\text{N}(p,\gamma)^{15}\text{O}$. This reaction was measured by the LUNA collaboration in [49, 50], yielding a cross section smaller than what initially believed and thus reducing the impact of the CNO energy production in the Standard Solar Model.

Neutrinos in the CNO-I cycle are emitted in the β^+ decay of ^{13}N and ^{15}O; they have a continuous spectrum with endpoint equal to 1.20 MeV and 1.73 MeV respectively. The same nuclei undergoing β^+ decay can also decay via electron capture, although with much smaller probability. These processes cause the emission of mono-energetic neutrinos (ecCNO) with an energy 1.022 MeV higher than the Q value of the corresponding β decay, but their expected flux for the Sun is so small that only $> 20\,\text{kton}$ experiments can have a chance to observe them and are therefore not considered in this thesis [51, 52]. Since the experimental sensitivity of current experiments does not allow to measure individually the ^{15}O and the ^{13}N neutrinos, all the neutrinos emitted

in the CNO cycle are grouped in a single effective component referred to as CNO neutrinos.

The CNO-II cycle The second CNO cycle originates from the ^{15}N produced in CNO-I and undergoing a rarer (p, γ) reaction in place of the (p, α) process that would have closed the cycle.

Since the initial (p, γ) reaction is quite unlikely, the CNO-II cycle accounts for only $\approx 2\%$ of the total energy produced in the CNO mechanism. In this cycle (right of Fig. 1.2a) neutrinos are produced via the β^+ decay of ^{17}F with an endpoint of 1.74 MeV along with the ^{15}O neutrinos already introduced in CNO-I.

1.2.3 *Predictions of the SSM and the Solar Metallicity puzzle*

Once the parameters of the SSM have been calibrated to match the present conditions of the Sun, the model provides a complete description of our star.

One of the most important predictions of the SSM is the flux of neutrinos produced in the nuclear reactions described in Sec. 1.2.2 that take place in the Sun core. The spectrum of solar neutrinos is shown in Fig. 1.3.
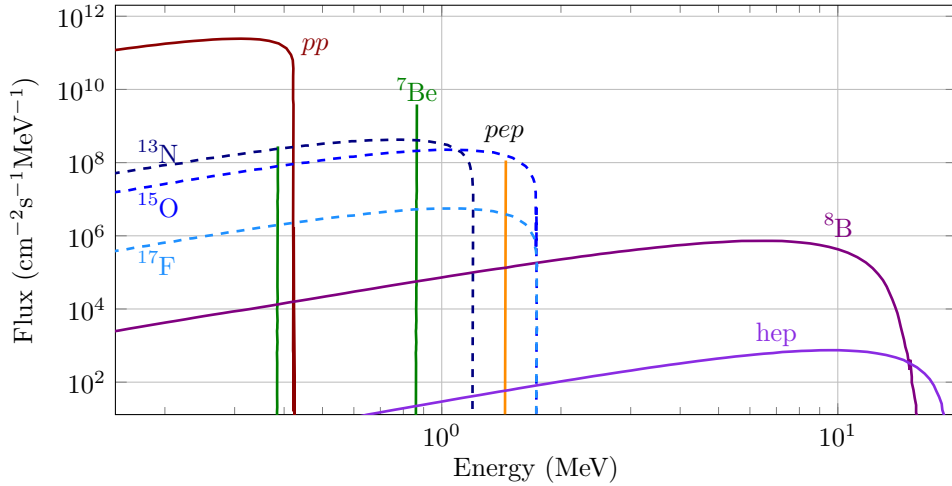
Figure 1.3: Solar neutrino spectrum expected from the SSM [33].

Another significant result of the SSM is the description of the Sun interior that can be probed by helioseismology. Just like seismology on Earth, helioseismology studies the propagation of pressure waves in the Sun. This phenomenon was observed already in 1962 [53] as a periodic shift in the absorption line on the Sun photosphere and in the last two decades these observations have provided the most stringent constraints to the inner structure of our star. Thousands of resonant modes were identified with an angular degree ranging from 0 to several hundreds, which allows to investigate the interior structure with great accuracy [35].

As explained in Sec. 1.2.1, the SSM is built matching the present condition of the Sun and therefore strongly depends on the measured properties of our star. In particular, the fraction of heavy elements[4] is a crucial parameters for calibrating the model. The abundances of chemical elements in the Sun is determined mostly by atomic spectroscopy measurements in the Sun photosphere, which relies on a modelling of the temperature and density stratification to link the intensity of the absorption lines with the element abundance. A three-dimensional radiation hydrodynamic description of the photosphere and line formation at non-local thermodynamic equilibrium were introduced in the model of the photosphere starting from [54], and leading to lower the estimation of CNO metals of the order of 30-40%.

Another way to determine the metal abundance in the Sun is by using the composition of meteorites as a reference for refractory elements. The fraction of refractories inferred by spectroscopic measurements is in general in good agreement with meteorites abundances apart from the AGSS09 catalogue [55] (which is obtained using the novel three-dimensional modelling of the photosphere) where small discrepancies were initially found for Fe, Ca and Mg. These deviations were reduced in [56, 57], and therefore the AGSS09 abundances are used taking meteorites values for refractory elements, leading to the AGSS09met set [34].

In [34] the Standard Solar Model is built assuming two different sets of metal abundances: the GS98 [58] catalogue was used as a reference for the "old" photospheric analysis method, while AGSS09met was chosen to represent the newest measurements. The SSM obtained assuming the GS98 abundances is referred to as the High-metallicity Standard Solar Model (HZ SSM), while the one relying on the AGSS09met is named Low–metallicty Standard Solar Model (LZ SSM) because of the lower content of metals.

A different fraction of metals in the Sun alters the opacity, which impacts on the seismic behaviour of the Sun itself. The sound speed profile predicted by the LZ SSM is in strong tension with the results obtained by helioseismology as shown in Fig. 1.4 and this large discrepancy, that was reported for the first time in [30], gave rise to the so-called *solar metallicity puzzle*.

The failure of the LZ SSM in describing the sound speed profile cannot be interpreted as a clear sign of a wrong composition of the Sun since in [59] it was shown that a wrong modelling of the opacity can in principle account for such deviations even maintaining the AGSS09met composition. It is extremely challenging to measure the opacity at conditions close to the ones of the Sun interior, therefore opacity is usually computed starting from the radiative properties of the atoms. The difference between the values of the two most widely used collections (OP [60] and OPAL [61]) is as large as 3% and cannot explain the measured discrepancy. In a recent experiment, the Iron opacity was measured at temperatures not far from the ones typical of stars [62] finding a value larger than the expected. Iron has a crucial role in the Sun since it accounts for $\approx 25\%$ of the total opacity at the conductive–convective boundary, but

---

4 Here "hevay elements" refers generically to the elements heavier than He.
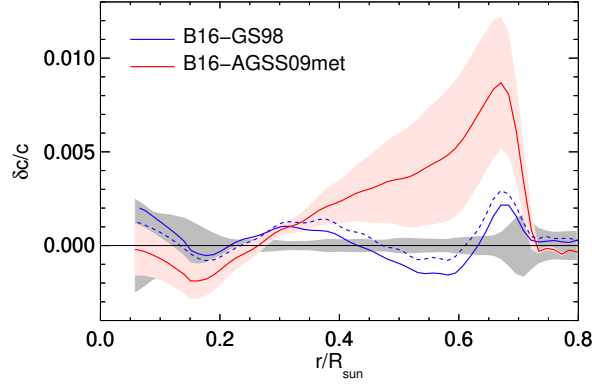
Figure 1.4: Fractional sound speed difference between the HZ/LZ SSM and helioseismological observations. The grey area reflects the uncertainty in the extraction of the sound speed from the data, while the red band around the LZ SSM prediction represents the model uncertainty. An analogue uncertainty should be considered for the HZ SSM, but it was not shown for the sake of clarity. The prediction of the HZ SSM in [33] is shown for comparison as blue dashed line. Figure from [34].

even taking into account this discrepancy in the model uncertainty [34] the LZ SSM still results hardly compatible with the observed data.

The difference in the opacity between the HZ and LZ SSM does not affect only the seismic behaviour of the Sun, but also the temperature profile that is in turn correlated with the solar neutrino production rates. The predicted solar neutrino fluxes for the HZ and LZ SSM are reported in Tab. 1.1 and offer an additional way to probe the SSM predictions as it will be shown in Sec. 5.2.2. In this framework CNO neutrinos are particularly interesting because their flux is correlated to the chemical abundance of C, N and O along with the solar temperature. The importance of CNO neutrinos for a solution of the metallicity puzzle is detailed in Chap. 6.

Finally, it is worth to mention a recent attempt to build a SSM in a data-driven way using as an *input* of the model the solar neutrino data and results from helioseismology [63] in order to determine the solar composition and to reconstruct the solar opacity profile from the available data. Even with this construction the LZ SSM results disfavoured respect to the HZ SSM, but the significance of the discrepancy is reduced.

1.3   PROPAGATION OF SOLAR NEUTRINOS

Thanks to their extremely small cross section, neutrinos produced by the nuclear processes in the Sun can escape the dense stellar environment which causes a photon to take more than 10^4 years to diffuse from the Sun core to the photosphere. For this reason, solar neutrinos can be considered direct messengers of nuclear processes in the Sun.

Solar ν	HZ SSM Flux cm^{-2}s^{-1}	LZ SSM Flux cm^{-2}s^{-1}
pp	$5.98(1 \pm 0.006) \times 10^{10}$	$6.03(1 \pm 0.005) \times 10^{10}$
^{7}Be	$4.93(1 \pm 0.06) \times 10^{9}$	$4.50(1 \pm 0.06) \times 10^{9}$
pep	$1.44(1 \pm 0.009) \times 10^{8}$	$1.46(1 \pm 0.009) \times 10^{8}$
^{8}B	$5.46(1 \pm 0.12) \times 10^{6}$	$4.50(1 \pm 0.12) \times 10^{6}$
hep	$7.98(1 \pm 0.30) \times 10^{3}$	$8.25(1 \pm 0.30) \times 10^{3}$
CNO	$4.88(1 \pm 0.11) \times 10^{8}$	$3.51(1 \pm 0.10) \times 10^{8}$

Table 1.1: Expected neutrino fluxes from the SSM when assuming the HZ (GS98, [58]) and LZ (AGSS09met, [34, 55]) composition model. The ^{13}N, ^{15}O and ^{17}F neutrino fluxes are summed into a single CNO component because current experimental sensitivity does not allow to study the different contributions individually. Values taken from [34].

However, the neutrino oscillations introduced in Sec. 1.1.2 and 1.1.3 should be considered when describing the propagation of solar neutrinos because detectors in general are not equally sensitive to all neutrino flavours. Indeed, it was through the measurement of flavour-dependent interactions of solar neutrinos that neutrino oscillations were observed (unconsciously) for the first time [21].

Solar neutrinos are produced in the Sun core in the electron flavour, described as a superposition of the eigenstates of the Hamiltonian in matter with weights given by $\cos\theta_M^{(i)}$ and $\sin\theta_M^{(i)}$ where the superscript (i) indicates that the effective mixing is evaluated at the production point. The survival probability in the two neutrino framework can be evaluated from Eq. (1.28), noticing that since the distance between the production and detection point is huge, the phase of the cos function is $\gg 2\pi$ for the energy resolution of the detector, and therefore the term is averaged to zero. In the full three–neutrino formalism, the P_{ee} then becomes [64]

$$P_{\nu_e \rightarrow \nu_e}(x) = \sin^4\theta_{13} + \cos^4\theta_{13}\left[\tfrac{1}{2} + \tfrac{1}{2}(1 - P_C)\cos 2\theta_{12}^{\mathrm{M(i)}}\cos 2\theta_{12}\right] \quad (1.38)$$

where P_C is the crossing probability (negligible below $10\,\mathrm{GeV}$ for adiabatic variations of the electron density like the ones in the Sun) and the effective mixing at the detection point $\theta_{12}^{\mathrm{M(f)}}$ is set to θ_{12} because most of the propagation happens in vacuum.

It is possible to consider a few limit cases for Eq. (1.38). Figure 1.5a reports the electron density as a function of the Sun radius computed by the SSM in [30]. The values of the effective mixing angle for such values of electron density are shown in Fig. 1.5b for a neutrino energy of $0.2\,\mathrm{MeV}$ (typical for pp neutrinos), $0.83\,\mathrm{MeV}$ (^{7}Be neutrinos) and $7.0\,\mathrm{MeV}$ (^{8}B neutrinos) assuming the Large Mixing Angle (LMA) solution for θ_{12} as indicated by the KamLAND

experiment. The matter effect is very small for the low energy pp and ^{7}Be neutrinos, $\theta_{12}^{\mathrm{M(i)}} \approx \theta_{12}$ and thus $P_{ee} \approx 1 - \frac{1}{2}\sin^2 2\theta_{12}$. On the other hand, for high energies ($E \gg 2\,\mathrm{MeV}$) the Sun core electron density is such that $\theta_{12}^{(\mathrm{M})} \approx \pi/2$ and thus $P_{ee} \approx \sin^2\theta_{12}$ independently on the neutrino energy.

A measurement of solar neutrinos covering a wide energy range offers the possibility to test both the vacuum approximation at low energy and the matter effect enhancing the oscillation for $E \gtrsim 3\,\mathrm{MeV}$.

(a) e^- density n_e **(b)** Effective mixing angle
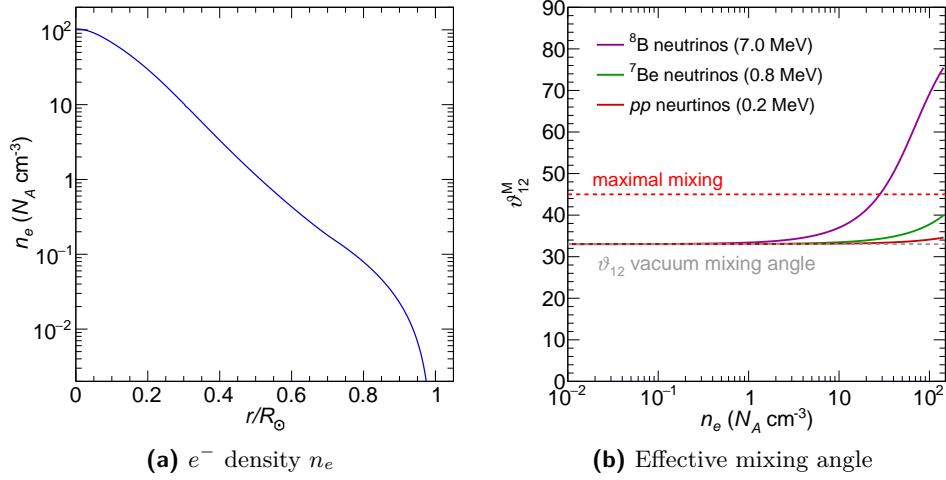
Figure 1.5: (a) Profile of the electron density in the Sun as a function of the radius (data from `http://www.sns.ias.edu/~jnb/SNdata/Export/BS2005/`, [30]). **(b)** Effective mixing angle in matter as a function of the electron density for the typical energies of solar neutrinos computed from Eq. (1.24).

1.4 MEASUREMENTS OF SOLAR NEUTRINOS

As already mentioned in this Chapter, the experimental measurement of solar neutrinos was a key element for the discovery of neutrino oscillations.

The main difficulty in the detection of solar neutrinos is due to the small interaction cross section of these particles, which requires large volume experiments with high sensitivity and low background. In particular, the suppression of the cosmic ray background is a key element for almost every neutrino experiment since, even with advanced active veto techniques, the high rate of cosmic rays makes a measurement of neutrinos on the surface extremely challenging. For this reason, all the experiments that measured solar neutrinos in the years were installed underground, with hundreds of meters of rock shielding the detectors from cosmic rays.

In this section the main results of this more than 50-year long investigation are summarized, grouping the experiments according to their working principle. Section 1.4.1 is devoted to the radiochemical experiments while Sec. 1.4.2

describes the results obtained with Cherenkov detectors. The main results of the Borexino experiment obtained before this thesis work are reported in Sec. 1.4.3 which presents liquid scintillator experiments.

1.4.1 *Radiochemical experiments*

Radiochemical experiments detect neutrinos through the inverse β decay reaction induced on some target material and exploits chemical processes to identify the transmuted nuclei. This kind of experiment can only provide an integrated measurement of the solar neutrino flux above a certain energy threshold that proved to be extremely significant to test the predictions of the SSM, but did not allow to separate the contributions of the different components of the solar neutrino spectrum.

This detection technique was used in the first pioneering solar neutrino experiment installed in the Homestake Gold Mine in South Dakota, USA. The Homestake Solar Neutrino Detector consisted in a steel tank containing 615 ton of tetrachloroethylene (C_2Cl_4) and detected solar neutrinos with the reaction $^{37}Cl(\nu_e, e^-)^{37}Ar$ which has an energy threshold of 0.814 MeV. Unstable Argon-37 atoms ($\tau_{^{37}Ar} \approx 35$ days) were extracted from the tank using chemical methods every two months and detected in small proportional counters where they decayed back to ^{37}Cl emitting an Auger electron. The production rate of ^{37}Ar from solar neutrinos was only 0.5 atoms per day, mostly due to ^8B neutrinos thanks to the increase of the cross section with energy, which corresponds to an average number of ≈ 16 ^{37}Ar detected atoms over 133 ton of Cl every two months. The very low number of atoms required an extremely careful extraction procedure as well as very rigid requirements on possible leaks from the apparatus. Already in its first year of activity, the experiment measured a rate lower than the expectations of the SSM [21] rising the so-called *solar neutrino puzzle* that was later solved with the discovery of neutrino oscillations. The experiment continued its activity until 1994, measuring an average neutrino capture rate of $2.56 \pm 0.16(stat) \pm 0.16(sys)$ SNU[5] [65], to be compared to the 9.3 ± 1.3 SNU predicted by the SSM.

The Homestake Chlorine experiment was followed by two experiments using the inverse β decay of ^{71}Ga, *i.e.* $^{71}Ga(\nu_e, e^-)^{71}Ge$ This reaction has an energy threshold of 0.233 MeV and thus the measurement is sensitive to all the neutrinos emitted by the Sun, including the abundant and low-energy *pp* neutrinos that marks the reaction driving the energy production in the Sun.

The GALLium EXperiment (GALLEX) ran in the INFN Gran Sasso National Laboratory in Italy from 1991 to 1997 [66] and after a short interruption continued the data taking until 2003 under the name of GNO (Gallium Neutrino Observatory). The detector consisted in a 101 ton of gallium chloride ($GaCl_3 - HCl$) solution for a total amount of 30.3 ton of Ga. The com-

---

5 SNU stands for Solar Neutrino Unit, a unit of measurement conveniently introduced for radiochemical experiments which corresponds to 10^{-36} captures per target nucleus per second.

bined results of GALLEX/GNO [67] gave a capture rate of $69.3 \pm 4.1(\text{stat}) \pm 3.6(\text{sys})$ SNU that is roughly one half of the SSM expectation.

The Soviet-American Gallium Experiment (SAGE) also exploited the ^{71}Ga inverse β decay, but using 50 ton of (molten) metallic Gallium. The experiment, located in the underground laboratories excavated under Mt. Andrychi in Russia, began its activity in 1989 and is still running. The latest published results [68] report a capture rate of $65.4^{+3.1}_{-3.0}(\text{stat})^{+2.6}_{-2.8}(sys)$ SNU, in good agreement with GALLEX/GNO.

1.4.2 *Water Cherenkov detectors*

Water Cherenkov detectors allow the measurement of neutrinos by observing the track of ultra-relativistic charged leptons produced by the neutrino interaction inside the detector. When these charged particles travel faster than light in a medium, the emitted Cherenkov photons can be detected and can be used to reconstruct the track direction that can be used to separate the tracks pointing back to the Sun from the isotropic background. These detectors often use water as the active material resulting in a active volume of several kilo-tons at relatively low cost, while the main drawback is the relatively high threshold (of some MeV) which allows to measure only the ^{8}B neutrinos. It is worth to notice that while radiochemical experiments exploit purely charged-current interactions, the elastic scattering of neutrinos on electrons involves all the neutrino flavours, although the cross section of ν_e is approximately six times larger than the one of $\nu_{\mu,\tau}$.

Among water Cherenkov detectors, a primary role is occupied by the Kamiokande experiment (1983–1995) [69, 70] and its successor Super-Kamiokande [71–74] (1996–running), which are hosted at the Kamioka mine in Japan and have provided the most accurate measurement of the ^{8}B neutrino flux continuously lowering the detection threshold. The latest published results are obtained with a remarkable 3.49 MeV threshold and report a ^{8}B neutrino flux of $(2.308 \pm 0.020(\text{stat}) \pm 0.040(\text{sys})) \times 10^6\,\text{cm}^{-2}\text{s}^{-1}$, to be compared with the expectations from the SSM in Tab. 1.1.

The SNO (Sudbury Neutrino Observatory) experiment [75] is another water Cherenkov neutrino detector located in a mine near Sudbury in Canada. The detector consisted of 1 kton of isotopically pure heavy water (D_2O) contained inside a spherical, 6 m radius, acrylic vessel. Differently from Kamiokande and Super-Kamiokande, the experiment is designed to be sensitive to both charged (CC) and neutral current (NC) neutrino interactions separately through the following detection channels

$$\text{CC:}\quad \nu_e + d \to e^- + p + p$$
$$\text{NC:}\quad \nu_\alpha + d \to \nu_\alpha + p + n$$
$$\text{ES:}\quad \nu_\alpha + e^- \to \nu_\alpha + e^-$$

where $\alpha = e, \mu, \tau$. After one year of data taking, the measured CC interaction rate showed a significant deficit with respect to elastic scattering measurements (ES) performed both by Super-Kamiokande and SNO itself, suggesting the presence of a non-ν_e active neutrino component in the solar neutrino flux [76]. This hypothesis was unambiguously proven with the measurement of the NC interaction rate [77–79] which, being equally sensitive to all the neutrino flavours, gave results consistent with the expectation of the SSM thus demonstrating the reliability of the SSM calculations and, independently from any model assumption, the existence of neutrino oscillation.

1.4.3   *Borexino and other liquid scintillator experiments*

After the discovery of neutrino oscillation with the accurate measurement of ^{8}B neutrinos performed by SNO and Super-Kamiokande, large liquid scintillator experiments were deployed with the goal of measuring the low-energy components of the solar neutrino spectrum.

This kind of detectors are able to measure the recoil energy of the electrons produced in neutrinos elastic scattering with a very low energy threshold. The biggest challenge for this detection technique comes from the undistinguishable background due to radioactive contaminations, that in standard conditions would exceed the solar neutrino signal by orders of magnitude.

The main experiment exploiting liquid scintillators for the measurement of solar neutrinos is Borexino, which is the topic of this thesis. The Borexino experiment, that will be described in Chap. 2, is running since 2007 at the INFN Gran Sasso National Laboratories in Italy and consists of a 270 ton ultrapure liquid scintillator active volume readout by 2212 photo-multipliers.

During the first phase of the experiment (2007–2010) it performed the first direct measurement of the mono-energetic ^{7}Be solar neutrinos [80] thanks to a fit of the reconstructed energy spectrum which separates the neutrino from the other radioactive spectral components. The accuracy of the results was later improved [81, 82] finding a flux of $(3.10 \pm 0.15) \times 10^9 \mathrm{cm}^{-2}\mathrm{s}^{-1}$, consistent with the oscillation scenario. The MSW mechanism was tested looking for variations in the ^{7}Be neutrinos interaction rate between night and day due to the possible effect of the interaction potential that neutrinos experience crossing the Earth. The measured asymmetry is consistent with the MSW-LMA solution which predicts a negligible effect [83].

Thanks to the development of a refined data analysis method for the rejection of the cosmogenic background [84], Borexino was also able to provide a first evidence of *pep* neutrinos measuring an interaction rate of 3.10±0.6(stat)± 0.3(sys) counts per day per 100 ton along with a strong limit to the CNO neutrino interaction rate [84].

Besides the investigation of the low-energy component of the solar neutrino spectrum, the liquid scintillator technique exploited by Borexino allowed to measure the ^{8}B neutrino flux with the lowest energy threshold of 3 MeV, find-

ing a flux of $(2.4 \pm 0.4) \times 10^6 \, \mathrm{cm^{-2} s^{-1}}$ [85] in good agreement with the result of SNO and Super-Kamiokande which benefit from a much larger detector mass.

After an additional purification campaign in 2010–2011, Borexino began its Phase II, whose data are analysed in this thesis. It is worth to mention that already before the end of Phase II, Borexino was able to mark an outstanding achievement by measuring the interaction rate of the *pp* neutrinos which, once corrected for the survival probability, results in a flux of $(6.6 \pm 0.7) \times 10^{10} \, \mathrm{cm^{-2} s^{-1}}$, in good agreement with the predictions of the SSM [86].

Another neutrino experiment which uses liquid scintillator is KamLAND. Located in the same experimental cave that hosted Kamiokande in the Kamioka mine in Japan, it was built with the goal of measuring electron-antineutrinos from nuclear reactors but it could also provide a measurement of the of the ^{7}Be neutrino flux with $\approx 15\%$ uncertainty [87] and of ^{8}B [88] neutrinos, although not as accurate as the one performed in the near Super-Kamiokande experiment.

1.5 OUTLOOKS

The study of solar neutrinos played a pivotal role in the history of particle physics with the discovery of neutrino oscillations. These outstanding results crowned 50 years of joint efforts of the neutrino and astrophysical community, but solar neutrino physics is still far from running out of interesting questions.

The solar abundances puzzle briefly presented in Sec. 1.2.3 is still unsolved more than ten years after it was outlined for the first time. In this period of time, remarkable experimental progresses made it possible to test the Standard Solar Model with unprecedented accuracy using the measurement of solar neutrino fluxes.

The Borexino experiment is currently the most sensitive solar neutrino detector and already during its Phase I opened the era of precision physics for the low-energy ^{7}Be solar neutrinos. After the purification campaign of 2011, Borexino Phase II could profit from the lowest radioactive background ever achieved in liquid scintillators and the results presented in this thesis combined with the prospects for the detection of CNO neutrinos confirmed the potential of this experiment to give new substantial contributions to the field.

Beside Borexino, many new experiments are expected to start operations in the next future. The SNO+ experiment is a 780 ton liquid scintillator experiment that exploits the SNO detector facilities in SNOLab, Ontario, Canada. Short after the commissioning, the detector was filled filled with ultrapure water in order to perform a short test of the apparatus that resulted in a measurement of the ^{8}B neutrino flux [89]. The experiment is about to start its liquid scintillator phase and although the main scientific goal is the search for the neutrinoless double-β decay in ^{130}Te, if radioactive contaminations in the scintillator will be low enough it will be possible to profit from the very large detector mass and improve Borexino measurments of solar neutrino fluxes [90].

Another experiment with the potential for contributing to solar neutrino is the Juno experiment, located in China. Juno [91] is a liquid scintillator experiment with an active volume of 20 kton that was developed with the primary goal of determine the neutrino mass hierarchy with a measurement of anti-neutrinos produced by two neighbour nuclear plants [92]. The data taking is planned to start in 2021, and the excellent design energy resolution (3% at 1 MeV) and huge detector mass make it possible to measure the solar neutrino fluxes with great accuracy. Like for all liquid scintillator experiment, the final sensitivity of Juno will depend on the successful purification of the scintillator. If a background of the level of Borexino is achieved, Juno can measure both low energy solar neutrinos and ^{8}B with an energy threshold as low as 2 MeV [93]. This energy region is particularly interesting because it offers the possibility to probe the transition region of the survival probability between the vacuum–dominated and the matter enhanced oscillation region, where possible effects of new physics may appear [94].

The measurement of high energy solar neutrinos will be improved by the forthcoming Hyper-Kamiokande experiment. The Hyper-Kamiokande experiment [95] is planned for 2023 and will consist of two water Čerenkov detectors with a mass of 258 ton each that will pursue a broad physics programme. Concerning solar neutrinos, Hyper-Kamiokande will improve the measurement of ^{8}B neutrinos and thanks to its unprecedented exposure it may be capable to measure for the first time the very low hep neutrino flux.

On a longer time scale, new techniques will likely be considered for the measurement of solar neutrinos. The next generation of double-phase TPCs for Dark Matter direct search can in principle be an excellent solar neutrino detector. Such experiments are designed to have a large mass, good energy resolution and extremely low background which are the same requirement of neutrino experiments. Sensitivity studies performed assuming liquid Xenon [96] and liquid Argon [97] as target show that, given the cosmic rays attenuation provided by currently available underground facilities, measurements of pp and CNO neutrinos with a $\approx 1\%$ and 15% accuracy are achievable. It is also worth to mention some recent progress in the development of water based liquid scintillators [98], which can in principle allow to build very large detectors [99, 100] with the possibility separate in time the prompt Cherenkov signal and the scintillation light [101]. Reconstructing the Cherenkov and scintillation signal separately on a event-by-event basis would be a major breakthrough because it will allow to exploit both the directionality of Cherenkov light and the good energy resolution provided by the scintillator [102].

THE BOREXINO EXPERIMENT

Borexino is a large, low background, liquid scintillator experiment primarily devoted to the measurement of low energy solar neutrinos. The detector was initially conceived as a prototype for BOREX, a multi kton Boron-doped liquid scintillator experiment capable of detecting ^{8}B neutrinos via both charged and neutral current interactions [103]. The possibility to reduce the internal background of the liquid scintillator to an extremely low level opened the way for the direct detection of the low energy ^{7}Be neutrinos that were previously measured only by radiochemical experiments and changed the goal of the experiment itself. Borexino successfully measured the ^{7}Be [80] neutrino flux and the outstanding background suppression allowed to go way beyond the initial goal of the experiment measuring also ^{8}B [85], *pep* [84] and *pp* [86] neutrinos and standing out as the most sensitive detector for low energy solar neutrinos.

This chapter introduces the key features of the Borexino experiment that are exploited by the analysis methods developed in this thesis. After presenting the principles of the detection of solar neutrinos in Borexino in Sec. 2.1, Sec. 2.2 describes the detector, Sec. 2.3 outlines the concepts of the data acquisition system and Sec. 2.4 illustrates the energy and position event reconstruction. The main sources of background are discussed in Sec. 2.5 and the data selection cuts used to suppress it are detailed in Sec. 2.6 along with additional discrimination techniques based on the specific pattern of the detected scintillation photons.

2.1 DETECTION PRINCIPLES

Solar neutrinos of all flavours are detected in Borexino through elastic scattering (ES) with electrons of the liquid scintillators

$$\nu_\alpha + e^- \rightarrow \nu_\alpha + e^- \qquad (\alpha = e, \mu, \tau) \qquad (2.1)$$

which transfer part of the neutrino energy to the electron. The electron recoils and deposits energy in the scintillator causing the excitation of the molecules that de-excite to the ground state emitting photons that are eventually detected by photo-multipliers (PMTs).

Figure 2.1: Neutrino–electron elastic scattering cross section as a function of the neutrino energy for ν_e (solid red line) and $\nu_{\mu,\tau}$ (dashed blue line). The larger cross section of ν_e is due to the additional interaction channel allowed by the charged-current interaction.
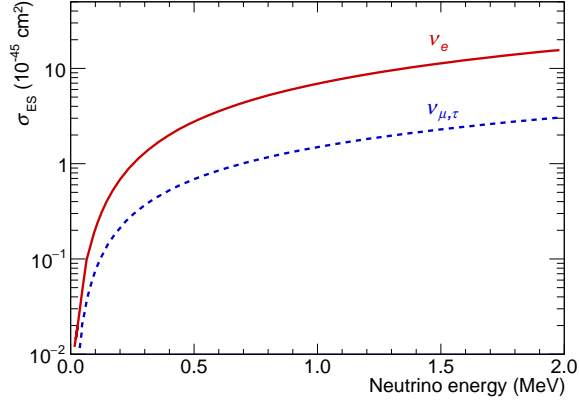
Neutrinos can interact with electrons both via the charged and neutral current interaction (Sec. 1.1.1). At energies of the order of $1\,\mathrm{MeV}$ the process cross section at leading order is computed averaging on the electron spin initial state and summing over the possible spin states in the final configuration. The resulting differential cross section for the electron final kinetic energy T_e is

$$\frac{\mathrm{d}\sigma_\alpha}{\mathrm{d}T_e}(E_\nu, T_e) = \frac{\sigma_0}{m_e}\left[g_1^2 + g_2^2\left(1 - \frac{T_e}{E_\nu}\right)^2 - g_1 g_2 \frac{m_e T_e}{E_\nu^2}\right] \qquad (2.2)$$

where m_e is the electron mass, $\sigma_0 = 2G_F^2 m^2 \pi^{-1} \approx 88.06 \times 10^{-46}\,\mathrm{cm}^2$, while the g_1 and g_2 coefficients are

$$g_1 = \begin{cases} \frac{1}{2} + \sin^2\theta_W & \text{for } \alpha = \nu_e \\ -\frac{1}{2} + \sin^2\theta_W & \text{for } \alpha = \nu_{\mu,\tau} \end{cases} \quad \text{and} \quad g_2 = \begin{cases} \sin^2\theta_W & \text{for } \alpha = \nu_e \\ \sin^2\theta_W & \text{for } \alpha = \nu_{\mu,\tau} \end{cases}$$
$$(2.3)$$

and describe the differences between the ν_e and $\nu_{\mu,\tau}$ due to the fact that while the first can interact both via CC and NC the latter are only interested by the NC process. Radiative corrections have been computed in [104], but their contribution is non negligible only for the highest energy part of the ^{8}B neutrino spectrum. As a consequence of the additional interaction channel, the integral elastic scattering cross section of ν_e is 4–5 times larger than the one of $\nu_{\mu,\tau}$, as shown in Fig. 2.1.

The interaction rate in Borexino for the component ν of the solar neutrino spectrum is given by [82]

$$R_\nu = N_e \Phi_\nu \int \frac{\mathrm{d}\lambda}{\mathrm{d}E_\nu}\mathrm{d}E_\nu \int \left\{ \frac{\mathrm{d}\sigma_e}{\mathrm{d}T_e}(E_\nu, T_e)P_{ee}(E_\nu) + \right.$$
$$\left. \frac{\mathrm{d}\sigma_\mu}{\mathrm{d}T_e}(E_\nu, T_e)\left[1 - P_{ee}(E_\nu)\right] \right\}\mathrm{d}T_e \quad (2.4)$$
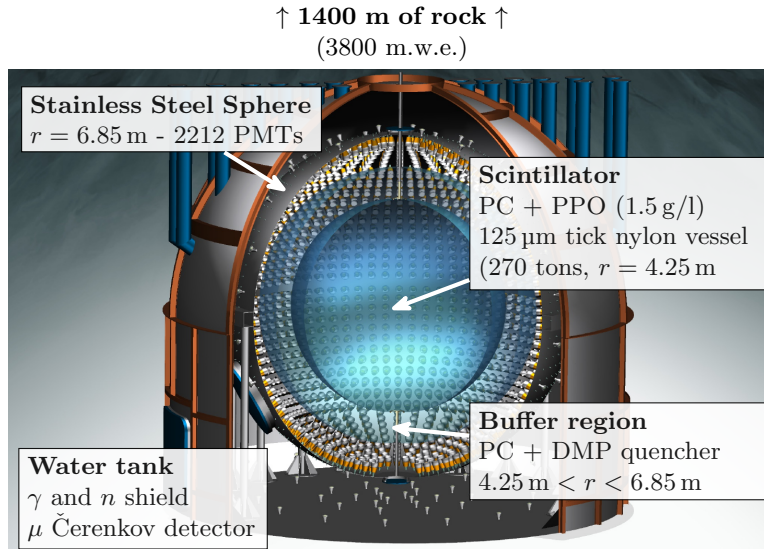
where N_e is the number of target electrons, Φ is the neutrino flux, $\mathrm{d}\lambda/\mathrm{d}E_\nu$ is the neutrino energy spectrum and P_{ee} is the survival probability introduced in Sec. 1.3. The resulting interaction rate is of the order of a few tens of counts

per day per 100 ton of scintillator (cpd/100 t), equivalent to 10^{-9} Bq/kg, to be compared with the typical ~ 10 Bq/kg rate of electrons and γ-rays produced by radioactive decays in standard materials that would leave a signal in the scintillator indistinguishable from the one due to neutrinos on a event–by–event basis. The huge disproportion between the expected signal and the typical background rate was the biggest challenge of the Borexino experiment and drove the design of the entire apparatus and experimental procedures. The accomplished reduction of 9–10 orders of magnitude of the background rate makes Borexino the most *radiopure* environment on Earth and is the cornerstone of all of its scientific achievements.

2.2 DETECTOR DESCRIPTION

The Borexino experiment is installed at the underground INFN Gran Sasso National Laboratories in Assegi (AQ), Italy, where an overburden of 1400 m of rocks (3800 m.w.e.) shields it from cosmic radiation.

Figure 2.2 shows a cutaway of the Borexino detector that is built aiming to progressively reduce the background towards the detector centre where the detector active volume is located.

↑ 1400 m of rock ↑
(3800 m.w.e.)

Stainless Steel Sphere
$r = 6.85$ m - 2212 PMTs

Scintillator
PC + PPO (1.5 g/l)
125 µm tick nylon vessel
(270 tons, $r = 4.25$ m

Buffer region
PC + DMP quencher
4.25 m $< r < 6.85$ m

Water tank
γ and n shield
μ Čerenkov detector

Figure 2.2: Cutaway of the Borexino detector.

Neutrinos are detected in Borexino through their interaction with 278 t of ultra pure liquid scintillator (LS), consisting in a mixture of pseudocumene (PC, 1,2,4-trimethylbenzene) with 1.5 g/l of 2,5-diphenyloxazole (PPO) acting as scintillating solute. Electrons scattered by neutrinos as well as other charged particles lose energy in the scintillator mostly interacting with the solvent and the excitation energy is transferred to the PPO molecules via non-radiative

processes. PPO de-excites onto the first excited singlet state that is the only level that de-excites with the emission of optical photons.

The amount of light produced by a particle interacting in the scintillator depends on the amount of energy deposited per unit length ($\mathrm{d}E/\mathrm{d}x$), that is characteristic of the type of particle. Particles with a large $\mathrm{d}E/\mathrm{d}x$ such as αs generate a high ionization density that favours the energy dissipation through non-radiative processes which result in a strong quenching of the scintillation light. The scintillator cocktail was optimized to ensure a large scintillation yield, a high transparency and a short decay time that are mandatory to provide a good energy and position resolution as well as the possibility to discriminate between the interaction of αs and electrons/γ-rays based on the different time development of the scintillation signal (Sec. 2.6.4.1).

The liquid scintillator occupies the innermost part of the detector and it is contained in a 4.25 m radius, 125 μm thick nylon Inner Vessel (IV) that delimits the detector Active Volume (AV). Scintillation photons are detected by 2212 8 in diameter photo-multipliers (PMTs) installed on the inner surface of a 6.85 m radius Stainless Steel Sphere (SSS) that encloses the entire inner detector. Among the 2212 PMTs, 1828 are equipped with light concentrators designed to enhance the light collection while shielding the photons that are not coming from the detector active volume.

The buffer volume between the SSS and the IV is filled with 1040 t of non-scintillating liquid consisting of PC with a small quantity of DMP (dimethylphtalate) that quenches the PC scintillation and provides an additional shielding from external γ-rays and neutrons coming from the SSS, from the PMTs glass and from the light concentrators. This buffer layer is divided in two parts by a 5.5 m, 125 μm nylon membrane to prevent gas emanation into the IV.

The SSS is contained into a 18 m diameter, 16.9 m high domed Water Tank (WT) filled with ultra-pure water that shields γ-rays and neutrons from the rock of the surrounding environment. At the same time, the Water Tank acts like an outer detector and provides an active veto for the residual cosmic muons. Indeed, despite the flux reduction of 6 orders of magnitude provided by the LNGS overburden, ≈ 4000 muons crosses the detector every day, much more than the number of expected neutrino interactions. For this reason, the water tank was coated with highly refractive Tyvek foils and equipped with 208 photo-multipliers in order to measure the Cherenkov radiation produced by crossing muons.

2.3 DATA ACQUISITION

The optical photons produced by the liquid scintillator after the energy deposition of particles interacting in the detector active volume are detected by 2212 PMTs mounted on the inner surface of the stainless steel sphere. In the energy window of interest for low-energy neutrino spectroscopy (0.1–3 MeV), approximatively 500 photo-electrons (p.e.) per MeV are expected to be recorded by

the detector, meaning that Borexino PMTs operate mostly in single photo-electron regime.

To reconstruct the energy deposited by an event in the detector, the charge collected by each PMT as well as the arrival time of the scintillation photons must be measured. To do so, each PMT is AC coupled to an analogue front-end followed by a digital board which digitizes the time and charge information of every photo-electron. The front-end feeds the digital board with a fast timing signal consisting in the amplified PMT signal itself. The timing signal is sent to a dual threshold discriminator which fires only when the higher threshold ($\approx 25\%$ of the average signal) is crossed, but retains the time information of the lower threshold crossing to filter the PMT dark noise. The discriminator output triggers two FADC (Flash Analogue-to-Digital Converter), one reconstructing the event timing and the other recording the collected charge that is provided by a dedicated gateless integrator [105] implemented on the front-end board which holds the charge plateau for 80 ns before discharging with a time constant of 500 ns. The logic board samples the integrated signal at its baseline and at its peak (after 80 ns) and the collected charge is computed by taking the difference between the two values. In case two photo-electrons are recorded separated by a time interval shorter than 80 ns they are reconstructed as a single "hit" with the time corresponding to the first p.e. and charge given by the sum of two integrated signals. The charge and time information of each PMT are stored in internal memory buffers, and to prevent data jamming during the writing and sampling operations a dead time of 140 ns is introduced every time the discriminator fires.

The experiment trigger must be able to recognize scintillation events, which are characterized by the quasi-simultaneous occurrence of many PMTs hits [106]. Given the large number of PMTs and the relatively high dark count rate, the trigger was implemented with a purely digital logic which fires only when a programmable number of PMTs detected a hit within a short time window. The typical number of fired PMTs required to trigger the detector is 20, corresponding to an energy threshold of ≈ 50 keV. The trigger time window is defined to be 100 ns, its minimum length being limited by the arrival time spread of the optical photons, which take at most 50 ns to cross the whole inner detector. In addition to the trigger for physics events, the Borexino Trigger Board also implements many other trigger types that are used for monitoring purposes.

The Borexino data acquisition was optimized for the detection of low energy solar neutrinos, but in principle it can be used also for the reconstruction of high energy events by correcting for the channels saturation. However, the possibility of detecting high energy ($E > 10$ MeV) neutrinos, especially from transient astrophysical events, motivated the implementation of a separate system based on the idea that with high energy events the individual record of all the 2212 PMTs is not needed. In this system the PMTs are grouped by solid angle sectors in 98 units, and the signal is recorded by a 400 MHz, 8-bit waveform digitizer.

2.4   EVENT RECONSTRUCTION

The information recorded by the main acquisition system described in the previous section are the time and the collected charge measured by each fired PMT. These two information constitute a "hit". Hits from fired PMTs are acquired within a 16.5 μs gate. Given the dark count rate of the order of a few hundreds counts per second per PMT, an average of 16 noise hits are expected in the acquisition window. A dedicated reconstruction algorithm, called `echidna`, analyses the hit time distribution in each acquisition gate in order to separate scintillation events from noise. Unphysical "bad" hits are removed by a decoding algorithm, while hits recognized as belonging to the same scintillation signal are grouped into different "clusters".

If two events in the scintillator are separated by a very short time (a few hundreds of ns), the clustering algorithm may fail in recognizing the two events as separate clusters and a single cluster is reconstructed. These random coincidences are limited to the lower energy part of the spectrum and are described in Sec. 3.3.3.

Clusters are the fundamental objects of Borexino data and are used to estimate energy, position and every other possible property to be related to an event. The description of the energy end position estimators is presented in Sec. 2.4.1 and 2.4.2 respectively, while other features such as the characteristic hit time distribution used to discriminate neutrino events from background are discussed in Sec. 2.6.4.

2.4.1   *Energy reconstruction*

Borexino defines tree main different energy estimators named N_p, N_h and $N_{p.e.}$ which represents the number of hit PMTs, the number of hits and the number of photo-electrons reconstructed for each cluster. In order to account for the variation of the number of active PMTs due to failures or temporary disabling, the estimators are normalized to 2000 working channels with the relation

$$N_{p,h,p.e} = N_{p,h,p.e}^m(t) \cdot f_{eq}(t) \tag{2.5}$$

where the superscript m indicates the measured value of the estimator at the time t and $f_{eq}(t)$ is an equalization factor defined as $f_{eq}(t) = 2000/N'(t)$ with $N'(t)$ being the number of working channels at the time t.

The different definition of energy observables leads to different features of the energy response: while $N_{p.e.}$ is linear over a wide energy range since it accounts for the detection of multiple p.e. N_p and N_h show a pronounced non-linearity but are less affected by dark noise that is critical at low energies.

In order to facilitate the computation of the study of random coincidences of multiple events, two additional variable with a "fixed duration" were derived from N_p. The $N_p^{dt_1}$ and $N_p^{dt_2}$ observables, defined as the number of fired PMTs in the first 230 ns and 400 ns after the beginning of the cluster, were introduced

for the first measurement of *pp* neutrinos; on one hand they benefit from a lower dark noise due to the short time window, but on the other they are more affected by the loss of linearity with respect to N_p.

2.4.2 *Position reconstruction*

The reconstruction of an interaction vertex in Borexino is based on the recorded hits pattern which is defined by the detection time of a photon (t_i) and by the position of the fired PMT ($\boldsymbol{r}_i$). To obtain the most likely position of the event, the position reconstruction algorithm maximizes the likelihood $\mathcal{L}_{\mathrm{PR}}(\boldsymbol{r}_0, t_0 | (\boldsymbol{r}_i, t_i))$ that the event occurred at time t_0 in the position $\boldsymbol{r}_0$ given the hits pattern. In the computation of the likelihood, the algorithm subtracts from each detection time t_i a position-dependent time-of-flight T_{flight}^{j} computed as

$$T_{\mathrm{flight}}^{j} = |\boldsymbol{r}_0 - \boldsymbol{r}_j| \cdot \frac{n_{\mathrm{eff}}}{c} \qquad (2.6)$$

where n_{eff} is the "effective refractive index". The effective refraction index defines an effective speed for the optical photons, taking into account in a single parameter the different group velocity, absorption and re-emission probability and detection efficiency of photons with different wavelengths. Given the complexity of the processes involved, it is hard to compute ab initio the effective refractive index and its value is obtained from calibration measurements.

The likelihood function is maximized comparing the hit time distribution (corrected for the time-of-flight) to the reference PDFs of the hit detection probability as a function of the time elapsed from the emission of the scintillation light. The PDFs are shown in Fig. 2.3a and their shape depends on the amount of charge associated to each hit.

The position reconstruction algorithm was tested during the calibration campaign when the whole detector was mapped with calibration sources of different energies, whose exact position was determined by CCD cameras. Figure 2.3b reports the spatial resolution in the x, y and z coordinates as a function of the energy for calibration sources positioned at the detector centre. For the x and y coordinates the spatial resolution goes from $15\,\mathrm{cm}$ at low energies ($150\,N_{p.e.} \sim 300\,\mathrm{keV}$) to $9\,\mathrm{cm}$ at $1\,\mathrm{MeV}$ ($\sim 500\,N_{p.e.}$). The vertical coordinate z is reconstructed with a slightly worse precision because the PMT coverage in z has a larger granularity [82].

2.5 BACKGROUND IN BOREXINO

The key of the outstanding results of Borexino in the field of solar neutrinos is the achievement of extreme low background levels. In this section the main components of the background are presented along with the experimental procedures implemented to reduce their impact. Section 2.5.1 discusses the background due to residual cosmic muons and muon-induced cosmogenic isotopes; the background due to radioactive decays occurring outside the de-
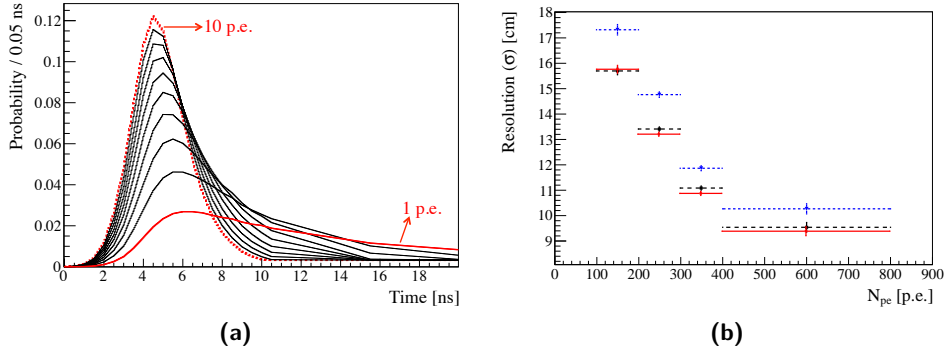
Figure 2.3: (a) Hit time detection PDFs as a function of time. The different curves are for increasing values of the hit charge collected by the PMT. **(b)** Spatial resolution of the position reconstruction algorithm for the x (solid red line), y (dashed black) and z (dotted blue) coordinate as a function of the energy for events produced by calibration sources placed in the detector centre. Figures from [82].

tector active volume, such as in the detector and PMTs structure and in the inner vessel nylon, is presented in Sec. 2.5.2; Sec. 2.5.3 finally presents the internal background due to the decay of radioactive isotopes that pollute the liquid scintillator.

2.5.1   *Muons and Cosmogenic background*

The flux of cosmic muons in Gran Sasso Laboratories is attenuated by 6 orders of magnitude thanks to the 3800 m w e overburden, resulting in $1.2 \, \mu m^{-2} h^{-1}$. Muons crossing the detector can be tagged with good accuracy by both the outer detector, which measures the Cherenkov light produced by the muon, and by the inner detector that can distinguish track-like from point-like interactions in the liquid scintillator.

   High energy muons crossing the detector can produce a number of unstable isotopes both via direct interaction or through the generation of hadronic showers. The main cosmogenic isotopes in Borexino are reported in Tab. 2.1. Most of them have a very short lifetime and their contribution is suppressed by applying a 300 ms veto to the entire detector after each muon. The only three components of the cosmogenic background with a relatively large production rate and a lifetime long enough to survive the muon veto and give a non negligible contribution to the energy spectrum are ^{6}He, ^{10}C and ^{11}C.

   The dominant cosmogenic background in Borexino is ^{11}C, which has a lifetime $\tau = 29.4$ min and undergoes β^+ decay with a Q-value of 960 keV. The total energy released in the scintillator is given by the sum of the energy deposited by the positron and the two 511 keV annihilation γ-rays, and is between 1020 and 1980 keV. This energy region is critical for the measurement of *pep* and CNO neutrinos (Sec. 1.2.2.1 and Sec. 1.2.2.2), especially since their rate is only $\approx \, ^1/_{10}$ of the ^{11}C background one. To obtain a measurement of

Isotope	Lifetime	Energy (MeV)	Decay
n	255 μs	2.23	capture γ on ^{1}H
^{12}N	15.9 ms	17.30	β^+
^{13}B	25.0 ms	13.40	β^-, γ
^{12}B	29.1 ms	13.40	β^-
^{8}He	171.7 ms	10.70	β^-, γ, n
^{9}C	182.5 ms	16.50	β^+
^{9}Li	257.2 ms	13.60	β^-, γ, n
^{8}B	1.11 s	18.00	β^+, α
^{6}He	1.16 s	3.51	β^-
^{8}Li	1.21 s	16.00	β^-, α
^{11}Be	19.9 s	11.50	β^-
^{10}C	27.8 s	3.65	β^+, γ
^{11}C	29.4 min	0.96	β^+
^{7}Be	76.9 d	0.48	EC γ

Table 2.1: Cosmogenic isotopes in Borexino ordered for increasing lifetime. The components that give a non negligible contribution to the Borexino spectrum are highlighted. In case of β decays the quoted energy is the decay Q value. From [82].

the *pep* and CNO neutrino ^{11}C events are tagged as described in Sec. 2.6.3 exploiting their correlation in space and time with a muon track and with the detection of at least one neutron that is often produced in the same process.

2.5.2 *External and Surface background*

The external background is due to radioactive decays taking place in the material that contains and surrounds the scintillator. The main contributions in Borexino come from the PMT glass, the light collectors and from the steel of the detector structure. This background is strongly suppressed thanks to the shielding provided by the non-scintillating buffer surrounding the detector active volume and only relatively high energy γ-rays from ^{40}K, ^{208}Tl and ^{214}Bi can reach the detector active volume.

These eternal γ-rays are progressively shielded by the liquid scintillator, therefore they can be further reduced selecting events inside a Fiducial Volume (FV) that considers only the innermost part of the active volume. The impact of external γ-rays is small for the measurement of ^{7}Be neutrinos, but becomes relevant for the *pep* and CNO neutrinos. For the solar neutrino analysis it is possible to constrain the background due to external γ-rays by exploiting the peculiar spatial distribution of their interactions which decreases exponentially towards the detector centre as discussed in Sec. 3.3.2.

The nylon of the inner vessel was produced meeting very high requirements in radiopurity (^{226}Ra activity < 21 mBq/kg, [106, 107]) but is more contaminated than the liquid scintillator. Radioactive decays on the nylon surface can be reconstructed at some distance from the inner vessel inside the detector active volume, but the definition of a fiducial volume for the analysis removes these events. The rate of γ–rays produced in the nylon and interacting inside the fiducial volume is estimated to be $\ll 1$*cpd* and therefore negligible compared to other sources of background. The relatively high surface contaminations are used to reconstruct the shape of the inner vessel, that is monitored on a weekly basis.

Background		Requirement	Results	
Type	Source		Phase I	Phase II
^{14}C	intrinsic PC	$< 10^{-18}$ g/g	$\sim 2 \times 10^{-18}$ g/g	$\sim 2 \times 10^{-18}$ g/g
^{238}U	dust, metals	$< 10^{-16}$ g/g	$\sim 1.6 \pm 0.1 \times 10^{-17}$ g/g	$< 9.5 \times 10^{-20}$ g/g
^{232}Th	dust, metals	$< 10^{-16}$ g/g	$\sim 5 \pm 1 \times 10^{-18}$ g/g	$< 7.2 \times 10^{-19}$ g/g
^{40}K	dust, PPO	$< 10^{-18}$ g/g	< 0.42 cpd/100 t	not seen
^{210}Po	^{222}Rn	< 1 cpd/t	~ 1 cpd/t	< 1 cpd/t
^{222}Rn	mat. emanation	< 10 cpd/100 t	< 1 cpd/100 t	< 0.1 cpd/100 t
^{85}Kr	air, nuclear test	< 1 cpd/100 t	30 ± 5 cpd/100 t	~ 5 cpd/100 t
^{39}Ar	air, cosmogenic	< 1 cpd/100 t	$\ll {}^{85}$Kr	$\ll {}^{85}$Kr
^{210}Bi	^{222}Rn	-	40 cpd/100 t	$\lesssim 20$ cpd/100 t

Table 2.2: Main sources of internal background in Borexino. The results achieved in Phase I and Phase II are reported along with the maximum allowed activity and the main reduction strategy.

Environmental γ–rays and neutrons produced in the walls of the Hall C are shielded by the water filling the outer detector and their rate inside the active volume is negligible.

2.5.3 *Internal background*

The internal background is due to the decays of radioactive isotopes contaminating the liquid scintillator. Table 2.2 reports the tolerable background level for different contaminations of the liquid scintillator along with the strategy implemented to reduce them. Purification and handling procedures [108–110] were developed and tested at the Borexino Counting Test Facility (CTF) [111] and are the key for the success of the Phase I of Borexino. Between 2010 and 2011, before beginning the Phase II of the experiment whose results are presented in this thesis, an additional purification campaign further reduced the Borexino internal background, confirming Borexino as the most sensitive low-energy neutrino observatory to date. In the following of this section the main sources of internal background are presented.

^{14}C

Carbon-14 (β^-, Q-value $= 156$ keV, $\tau = 8270$ yr) is chemically identical to ^{12}C copiously present in organic scintillator and therefore cannot be removed with chemical processes. In order to reduce the ^{14}C/^{12}C fraction from the typical value of 10^{-12} g/g, the Borexino scintillator was derived from petroleum from deep underground [82, 112] obtaining a concentration of $\approx 2 \cdot 10^{-18}$ g/g.

Despite the impressive reduction of the ^{14}C concentration, ^{14}C is by far the most abundant background in Borexino with an average rate of the order of 3.5×10^6 cpd/100 t. The Borexino trigger threshold (Sec. 2.3) corresponding to ≈ 50 keV reduces the trigger rate to $\lesssim 30$ Hz. However, the high ^{14}C rate makes the probability of reconstructing two decays in a single cluster non-negligible. The random coincidence of two ^{14}C is a crucial background at low

energy for the measurement of *pp* neutrinos and will be addressed in Sec. 3.3.3 and Sec. 4.2.

^{85}Kr

Krypton-85 (β^- (99.53% B.R.), Q-value $= 687\,\text{keV}$, $\tau = 15.4\,\text{yr}$) is present in the atmosphere with an average activity of $\approx 1\,\text{Bq/m}^3$ which requires extreme care to avoid any contact with air in the scintillator handling and detector filling operations. Its contamination can be efficiently reduced by stripping the liquid scintillator with N_2 gas [113].

Krypton-85 can also decay onto the $^{85\text{m}}$Rb, although with B.R. $\approx 0.43\%$, emitting an electron with $173\,\text{keV}$ maximum energy followed by a $514\,\text{keV}$ γ-ray from the $^{85\text{m}}$Rb de-excitation with a mean lifetime of $2\,\mu\text{s}$. The coincidence in space and time of these two events can be used to obtain an independent measurement of the ^{85}Kr rate inside the scintillator. Using this method, a rate of $\approx 30\,\text{cpd/100\,t}$ was measured in Phase I [82] but after the 2010-2011 purification campaign the ^{85}Kr concentration is expected to be reduced by a factor ~ 5.

^{39}Ar

Argon-39 (β^-, Q-value $= 565\,\text{keV}$, $\tau = 386.6\,\text{yr}$), similarly to ^{85}Kr, is present in the atmosphere with an average activity of $1.43\,\text{Bq/m}^3$ and a leak tight experimental environment is required in order to avoid contaminations. Unlike ^{85}Kr, ^{39}Ar has no secondary decay channels that can be exploited to obtain an independent measurement of its rate, therefore great efforts were made to ensure a very low ^{39}Ar contamination. The contamination in the liquid scintillator was reduced by N_2 gas stripping with contamination of Ar in nitrogen as low as $\approx 0.005\,\text{ppm}$. With such pure gas, the expected ^{39}Ar rate is less than $0.02\,\text{cpd/100\,t}$ and therefore negligible [82].

^{238}U *chain*

Isotopes belonging to the ^{238}U chain contaminate the liquid scintillator as micron and sub-micron particulate (dust), but their concentration is strongly reduced by the purification procedures adopted by Borexino, in particular by the scintllator filtering, distillation and "water extraction" [113].

The level of the contaminations due to the unstable isotopes belonging to the ^{238}U chain in secular equilibrium can be estimated exploiting the fast sequence of ^{214}Bi and ^{214}Po decays

$$^{214}\text{Bi} \rightarrow {}^{214}\text{Po} + e^- + \overline{\nu}_e \qquad (Q = 3.27\,\text{MeV})$$
$$^{214}\text{Po} \rightarrow {}^{210}\text{Po} + \alpha \qquad (E_\alpha = 7.7\,\text{MeV})$$

where the lifetime of ^{214}Po is $238\,\mu\text{s}$. An increase of the ^{214}Bi–^{214}Po "fast coincidence" rate followed by an exponential decrease with the mean lifetime

$\tau = 5.5$ days of ^{222}Rn is observed after operations on the detector (filling, insertion of calibration sources, ...), but no persistent contamination is measured. This indicates that the increase in the tagged ^{214}Bi–^{214}Po coincidences is due to the ^{222}Rn reaching the detector active volume.

The "baseline" rate of ^{214}Bi–^{214}Po fast coincidences allows to constrain the ^{238}U concentration to be lower than 9.5×10^{-20} g/g during Phase II, improving the already impressive background level achieved in Phase I.

^{210}Bi

Bismuth-210 (β^-, Q-value $= 1160$ keV, $\tau = 7.23$ days) is one of the main background in the Borexino spectrum since it affects the region of interest for ^{7}Be, *pep* and CNO neutrinos. In particular, the ^{210}Bi background is critical for the sensitivity to CNO neutrinos since its energy spectrum is almost identical to the energy distribution expected from CNO neutrino events. The presence of ^{210}Bi in the liquid scintillator comes from a contamination of the ^{210}Pb parent nucleus ($\tau_{^{210}\text{Pb}} = 32.2$ yr) that decays into ^{210}Bi with a Q-value of 65 keV too small to be detected. A method for obtaining an independent measurement of the ^{210}Bi rate based on the ^{210}Bi equilibrium with the parent ^{210}Pb is discussed in Sec. 6.3.2.

^{210}Po

Polonium-210 (α, $E_\alpha = 5.3$ MeV, $\tau = 199.6$ days) is the daughter of ^{210}Bi but during Phase I as well as at the beginning of Phase II a large out-of-equilibrium contamination made it the second largest source of background after ^{14}C.

The mono-energetic α particle ionization in the liquid scintillator is strongly quenched and the reconstructed energy falls in the region of the ^{7}Be neutrinos. The α decay of ^{210}Po can be efficiently tagged exploiting the characteristic time development of the scintillation light (Sec. 2.6.4.1).

^{232}Th *chain*

Like the isotopes of the ^{238}U decay chain, isotopes descending from ^{232}Th that contaminates the scintillator as dust are efficiently removed by the purification procedures [113]. The residual contamination of ^{232}Th daughters is measured by the delayed coincidence of the ^{212}Bi–^{212}Po decays

$$^{212}\text{Bi} \rightarrow {}^{212}\text{Po} + e^- + \overline{\nu}_e \qquad (Q = 2.3 \text{ MeV})$$
$$^{212}\text{Po} \rightarrow {}^{208}\text{Pb} + \alpha \qquad (E_\alpha = 9.0 \text{ MeV})$$
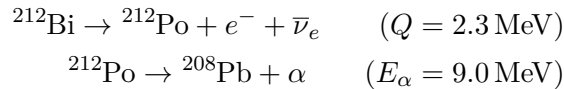
The rate of ^{212}Bi–^{212}Po fast coincidences can vary after some operations on the detector (filling, calibration), but no persistence variation was recorded and the ^{212}Bi–^{212}Po recovered its initial values in a few days. The concentration of the ^{232}Th chain isotopes during Phase II is estimated from the asymptotic ^{212}Bi–

^{212}Po rate and results in an upper limit at 7.2×10^{-19} g/g, which improves the Phase I limit of one order of magnitude.

^{40}K

Potassium-40 (β^- (BR 89%), Q-value= 1310 keV, $\tau = 1.85$ Gyr) is a primordial nuclide with a natural abundance of 0.012%. Large traces of ^{40}K were found in commercially available PPO to the extent that its rate would have been thousands of times larger than the solar neutrino interactions one. Potassium is removed from the liquid scintillator by filtering, distillation and water extraction [113]. It is possible to measure the content of ^{40}K exploiting its secondary decay channel into ^{40}Ar* that in turn decays emitting a 1460 keV γ-ray. The analysis of Phase I data found an upper limit of 0.4 cpd/100 t (95% C.L.) on the ^{40}K and Phase II confirmed no evidence of ^{40}K.

2.6 BACKGROUND REDUCTION TECHNIQUES

After being processed by the reconstruction code, runs acquired in the same week are grouped in a single pre-filtered file (DST). In order to reduced the background events, only events satisfying a series of conditions discussed in this section are selected for the analysis.

Cosmogenic background events, noise and ^{214}Bi–^{214}Po coincidences are removed by applying the data selection cuts described in Sec. 2.6.1. External and surface background is reduced selecting the events reconstructed inside a Fiducial Volume which can be defined differently depending on the goal of the analysis as discussed in Sec. 2.6.2. The impact of data selection cuts is clearly visible from Fig. 2.4, where structures in the spectrum like the peak due to the quenched α decay of ^{210}Po and the "bump" due to the decay of ^{11}C become clearly visible. In order to reduce the cosmogenic background due to long-living cosmogenic isotopes, a veto based on the threefold coincidence of an event with the passage of a muon and the detection of a neutron capture is implemented and discussed in Sec. 2.6.3.

Another possibility to disciminate solar neutrinos from α or β^+ background events relies on the different time development of the emission of the scintillation light and is discussed in Sec. 2.6.4.

2.6.1 *Data selection*

In order to remove muons and short-living cosmogenic background, fast coincidences and noise events, the following cuts are applied:

1. *Muons and muon daughters*: muons are tagged either by the outer detector or by the characteristic track-like signal reconstructed in the inner detector. As soon as a muon is identified the entire detector is vetoed for 300 ms, during which most of the cosmogenic unstable isotopes produced by muon-induced showers decay.
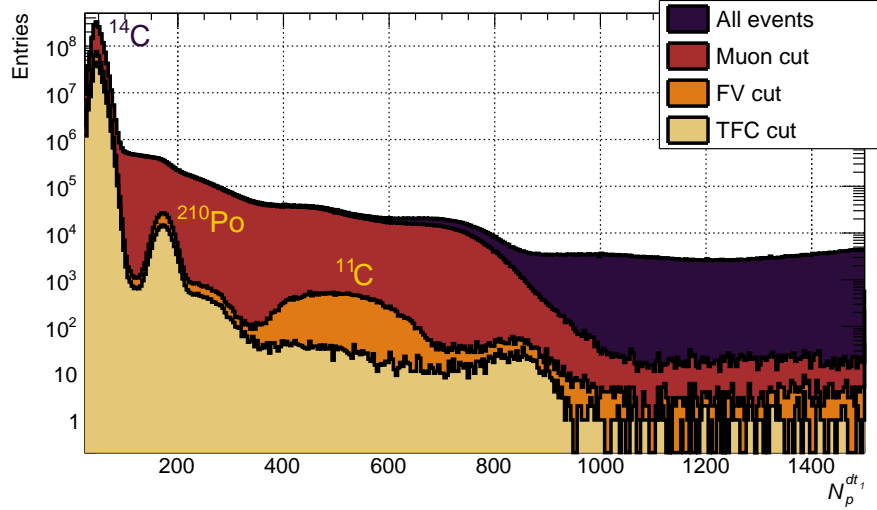
Figure 2.4: Effects of the data selection cuts on Borexino Phase II spectrum. Starting from the total spectrum, cuts removing cosmic muons and short-living cosmogenic isotopes (Sec. 2.6.1), external background (Sec. 2.6.2), and the long-living cosmogenic ^{11}C (Sec. 2.6.3) are applied in series.

2. *Single Clusters*: only events where a single cluster is identified by the reconstruction algorithm are accepted. If the reconstruction algorithm failed to identify at least one point-like scintillation event or reconstructs more than one cluster in the acquisition gate, the event is rejected.

3. *Fast coincidences*: in order to tag and remove ^{214}Bi–^{210}Po delayed coincidences, events reconstructed with a mutual distance smaller than 1.5 m within a 2 μs time window are dropped.

4. *Start time*: clusters with a strange position of the first hit time are accounted as faulty trigger and discarded.

5. *Reconstructed charge control*: the total collected charge is compared to the number of fired PMTs. If the charge is too large or too small given the PMTs hit the event is dropped.

6. *Crate Fraction*: if more than 75% of the fired PMTs are readout and powered by the same crate the event is attributed to crate noise and discarded.

The total dead time introduced by these cuts is dominated by the 300 ms veto applied to remove muons and cosmogenic background. The fraction of good events removed by these cuts was determined from calibrations and simulations and resulted to be much less than 1% [114].
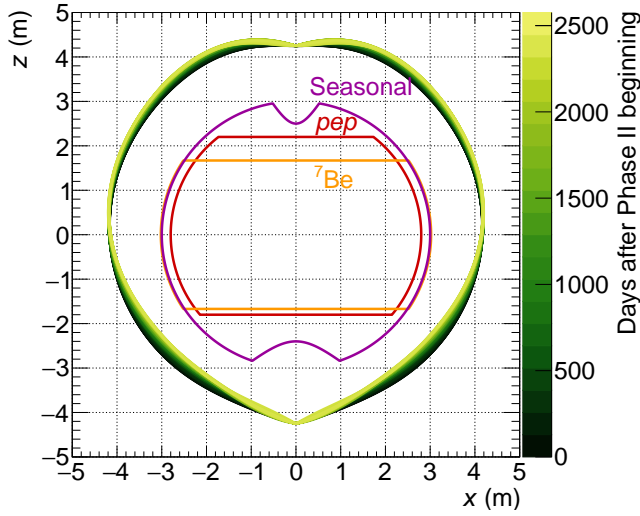
Figure 2.5: Definition of the FV selection in the x–z plane for different solar neutrino analyses compared to the vessel shape evolution throughout Phase II.

2.6.2 *Fiducial volume cuts*

The fiducial volume cut is defined with the goal of minimizing the external and surface background while selecting the largest possible mass of scintillator where the detector energy response is uniform[1] The optimal trade-off between these requirements led to slightly different definitions of the fiducial volume in Borexino solar neutrino analyses depending on the physics goal of the measurement.

Figure 2.5 shows the different definitions of the fiducial volume used for the ^{7}Be neutrino precision measurement, for the analysis of *pep* and CNO neutrinos and for the measurement of the seasonal modulation of the ^{7}Be neutrino flux. The cut on the vertical coordinate z that is applied in different ways for all the fiducial volume choices is motivated by the presence of the end-caps at the top and bottom of the inner vessel. The end-caps are an additional source of external background and their presence also affects the light collection efficiency at the top and bottom of the detector.

^{7}Be FIDUCIAL VOLUME   The fiducial volume for the measurement of ^{7}Be was defined applying a radial cut of $r < 3.021$ m and a vertical cut of $|z| < 1.67$ m, resulting in a total mass of 75.46 tons.

pep FIDUCIAL VOLUME   The fiducial volume chosen for the *pep* and CNO neutrino analysis during Phase I [84] is defined with a radial cut of $r < 2.8$ m and an asymmetric cut in the vertical coordinate ($z > -1.8$ m, $z < 2.2$ m) motivated by the presence of external background generated in the end-caps and by the requirement of a uniform light collection efficiency [115]. The hard radial cut was motivated by the stronger impact of the external background

---

1 The requirement of a uniform energy response of the detector in the selected volume is important to ensure a good energy resolution.

on the *pep* neutrino analysis with respect to the ^{7}Be one. This definition of the fiducial volume corresponds to a mass of 71.3 ton and is the one that is chosen for all the results presented in this thesis where not stated otherwise.

SEASONAL MODULATION FIDUCIAL VOLUME   A dedicated analysis was developed to measure the seasonal modulation of the ^{7}Be flux due to the Earth orbit eccentricity. This particular measurement is less affected by the external background and by the loss of resolution compared to the measurements of the neutrino interaction rate, and therefore the fiducial volume was extended in the vertical direction applying two parabolic cuts in correspondence of the end-caps [116]. The corresponding mass is 98.6 tons.

2.6.3   *Threefold coincidence method*

The background due to cosmogenic ^{11}C decays is one of the main backgrounds in Borexino and is clearly visible in the events spectrum shown in Fig. 2.4. Cosmic muons reaching the detector have an average energy of 320 GeV [117] and can produce ^{11}C either by spallation (*i.e.*, through a direct interaction mediated by a virtual photon)

$$^{12}\text{C} + \mu \rightarrow \mu + {}^{11}\text{C} + n + \text{secondaries} \tag{2.7}$$

and through the interaction of ^{12}C with products of muon-induced showers. The production channels of ^{11}C were studied in [118], showing that in 95% of the cases at leas one neutron is present in the final state. It is then possible to tag ^{11}C exploiting their correlation in space and time with the passage of a muon and the delayed γ-ray(s) from the capture of one or more neutrons by hydrogen. This space-time correlation is often referred to as threefold coincidence (TFC).

The first implementation of the TFC algorithm established a veto in the region of the detectors interested by the passage of a muon track and by the neutron captures [82, 84]. A more recent version of the TFC [119] computes for each event the probability of being a comogenic background event. The likelihood function is built taking into account the relative timing and distance of the event from the muon track and from the neutron capture vertex, the neutron multiplicity and the muon energy deposit along the track. Figure 2.6 shows the likelihood ($\mathcal{L}_{\text{TFC}}$) as a function of the reconstructed energy for Borexino Phase II data. Events with $\mathcal{L}_{\text{TFC}}$ below are certain threshold are collected in the TFC-subctracted dataset (shown in Fig. 2.4), while remaining events constitutes the TFC-tagged dataset. The $\mathcal{L}_{\text{TFC}}$ threshold used to define the TFC-subtracted and -tagged datasets is chosen to maximize the TFC efficiency cut while preserving the exposure of the TFC-subtracted dataset. In the analysis presented in this thesis, the ^{11}C-tagging efficiency is $(92 \pm 4)\%$ and the TFC-subtracted datasets retains $(64.28 \pm 0.01)\%$ of the total exposure [119].
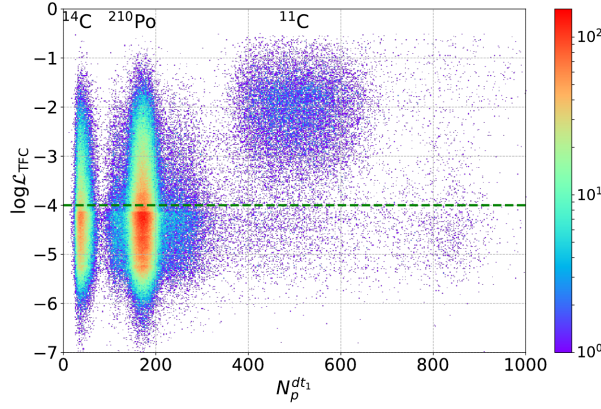
Figure 2.6: Distribution of $\mathcal{L}_{\text{TFC}}$ as a function of the $N_p^{dt_1}$ energy estimator. ^{11}C events are clearly separated from the rest of the events. The green dashed line indicates the threshold used to define the TFC-subtracted and -tagged datasets. Figure from [119].

In addition to the selection based on the value of $\mathcal{L}_{\text{TFC}}$, the TFC algorithm vetoes the whole detector for 2 hours at the beginning of each DST file since DSTs are independent on each other and the information on muons are not shared from a DST to the next. Similarly, another full volume veto is applied at the beginning of each run since a muon could be lost in the time between the two runs. The duration of the veto is $[600 + 3600(1 - \exp\{-3\Delta t/\tau\})]$ s, where Δt is the time interval between the two runs and τ is the neutron capture lifetime (≈ 0.25 ms) [114].

2.6.4 *Pulse shape discrimination*

The emission time of scintillation light depends on the details of the energy loss of a particle inside the liquid scintillator. The resulting hit time distribution, or *pulse shape*, can then be used to discriminate e^--like events, like the ones expected from the interaction of solar neutrinos, from background due to α or β^+ decays as described in the following sections.

2.6.4.1 α/β *discrimination*

The time development of a scintillation event, just like the amount of emitted light, depends on the $\mathrm{d}E/\mathrm{d}x$ and thus on the particle type.

Scintillation photons are emitted by the de-excitation of the first excited singles state, but highly ionizing α particles create along their track a region with high ionization density, with a relatively large number of PPO molecules in the first triplet excited state (T_{10}). This state cannot de-excite emitting light nor can pass to the singlet state: the only way to decay is by molecular interaction with another molecule in the T_{10} state which brings one of the molecules at the ground states and the other to the first excited singlet state that promptly de-excites emitting a *delayed* photon. The molecular interaction yield scales as the squared of the concentration of molecules in the T_{10} state, therefore the fraction of light emitted with a delay will be larger for the particle with larger $\mathrm{d}E/\mathrm{d}x$ [120].

The different hit time distribution of α particles with respect to β^- events is shown in Fig. 2.7a where ^{214}Bi–^{214}Po events selected through the fast coincidence method (Sec. 2.6.1) are used as a reference for β^- and α respectively.

(a)   **(b)**
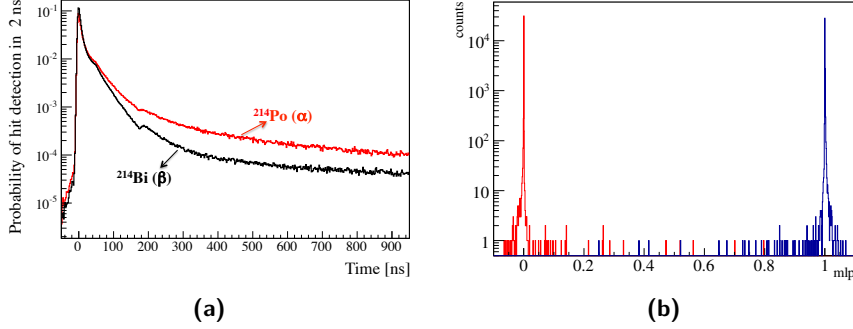
Figure 2.7: (a) Hit time distribution for α and β^- events selected with the ^{214}Bi–^{214}Po fast coincidence. Figure from [82]. **(b)** Distribution of the MLP classifier for a α (red) and β^- (blue) training sample obtained selecting ^{214}Bi–^{214}Po fast coincidences. Figure from [116].

During Phase I a linear discrimination technique based on the Gatti parameter was used to separate α and β events [82]. More recently, a new approach based on a neural network was developed [116]. The new method exploits a MultiLayer Perceptron (MLP) algorithm that is implemented in the `TMVA` package of the `ROOT` analysis framework [121].

The algorithm classifies an event as belonging to one or the other of the considered classes on the base of a vector of features selected by the user, such as different tail-to-total, the mean time and the number of hits inside the cluster [114]. The neural network is *trained* using part of the ^{214}Bi–^{214}Po events as reference samples for the α and β^- classes. After the training, the discrimination efficiency is evaluated on a test sample of events independent from the one used for the training: the algorithm assigns to each tested event a value close to 0 or 1 depending on the most likely class. Figure 2.7b displays the distribution of the MLP output for a ^{214}Bi–^{214}Po independent test sample and shows the power of this discrimination method.

The MLP classifier allows to tag α decays in the scintillator, like the ones due to ^{210}Po, with great efficiency but the dependence of the MLP performance for different events and detector conditions still requires some studies. For instance, one should consider that the neural network was trained using ^{214}Po α events, which have a higher energy compared to those of ^{210}Po. A possible energy dependence of the MLP efficiency is therefore not excluded. Similarly, the loss of PMTs during Phase II can slightly affect the MLP performance: for this reason studies with a specific dataset defined to minimize the changes of the detector response [93] are foreseen in the near future.
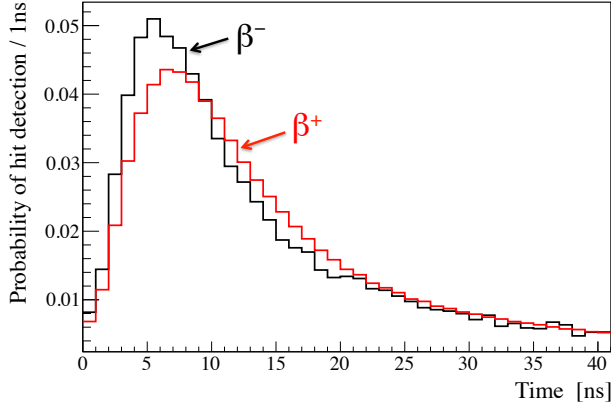
Figure 2.8: Hit time distribution for β^+ and β^- events. Figure from [82].

2.6.4.2 e^+/e^- pulse shape discrimination

The interaction of positrons inside the liquid scintillator is characterized by a specific time evolution and topology of the energy deposit that can be exploited to separate the neutrino signal from the β^+ decays background due to ^{11}C that survived the TFC cut (Sec. 2.6.3).

A positron loses its energy primarily via ionization and in 95% of the cases it forms a e^+–e^- bound state ("positronium") with an electron in the scintillator before annihilating with the emission of two 511 keV γ-rays [82]. Depending on the relative orientation of the electron and positron spin, the positronium has a different mean lifetime, being 125 ps for the singlet state (para-positronium) and 140 ns for triplet states (ortho-positronium). The formation probability of para- and ortho-positronium in matter, as well as their mean lifetime, differs from the values computed in vacuum because of the interactions with the surrounding medium. For the Borexino scintillator, the formation probability of ortho-positronium is 51.2% [122]. The corresponding mean lifetime is reduced to 3.12 ns [122], that is compatible with the fastest component of the scintillation light. The delayed annihilation of positronium therefore produces a visible effect in the hit time distribution of β^+ events with respect to β^- interactions: Fig. 2.8 shows the difference between the hit time distribution for a sample of β^+ events obtained from a clean sample of ^{11}C selected applying a strict TFC cut and for ^{214}Bi β^- decays selected thanks to the ^{214}Bi–^{214}Po fast coincidence.

Another signature of β^+ decays comes from their decay topology. Indeed, the average path of the two annihilation γ-rays in the liquid scintillator is in the 10–20 cm range, which is larger than the typical spatial resolution of the position reconstruction.

During Phase I the output of a Boosted Decision Tree (BDT) neural network was introduced as a pulse shape parameter (PS-BDT) [82, 84, 115]. The neural network considers both the features of the hit time profile and the event topology and was trained using a sample of ^{11}C events selected with a strict TFC cut and ^{214}Bi β^- decays in the energy region interested by ^{11}C.

The distribution of the PS-BDT parameter for a set of events independent from the training dataset is shown in Fig. 2.9a.

(a) PS-BDT   **(b)** PS-$\mathcal{L}_{\mathrm{PR}}$
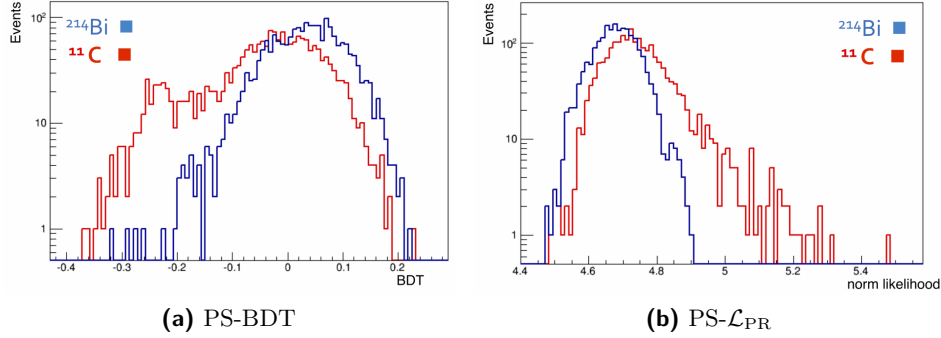
Figure 2.9: Test of pulse shape discrimination on a sample of strictly selected ^{11}C and ^{214}Bi events using the PS-BDT **(a)** and the PS-$\mathcal{L}_{\mathrm{PR}}$ parameters. Figure from [114].

A new pulse shape discrimination parameter based on the likelihood of the position reconstruction $\mathcal{L}_{\mathrm{PR}}$ (Sec. 2.4.2) was introduced for the analysis of Borexino Phase II data [114]. The "diffused" energy deposit of a β^+ event (prompt localized ionization, followed by two delayed γ-rays) creates multiple local maxima in the likelihood function and causes a general worsening of the position reconstruction that results in smaller values of the maximum likelihood. The value of the maximized likelihood depends on the number of hits in the cluster. This energy dependence is almost completely removed by normalizing the likelihood for the number of fired PMTs [114]. The distribution of the normalized likelihood pulse shape parameter[2] PS-$\mathcal{L}_{\mathrm{PR}}$ for the same test sample used for PS-BDT is shown in Fig. 2.9b.

The choice of the PS-$\mathcal{L}_{\mathrm{PR}}$ offers approximatively the same discrimination power of PS-BDT, but also brings some advantages. In particular, the new pulse shape variable is easier to simulate and is less dependent on the events energy compared to PS-BDT [114].

2.7 CONCLUSIONS

The achievement of an extremely low background level has been crucial for the success of the Borexino experiment, proven by the outstanding scientific results obtained since the very beginning of the data taking in 2007.

Despite the impressive reduction, the residual background surviving the data selection cuts described in Sec. 2.6 has a rate compatible with the one expected from solar neutrinos. Figure 2.10 shows the expected contributions

---

2 Since the position reconstruction algorithm finds the maximum likelihood minimizing the negative logarithm of $\mathcal{L}_{\mathrm{PR}}$, the PS-$\mathcal{L}_{\mathrm{PR}}$ variable is computed from the normalized minimum value of $-\log\mathcal{L}_{\mathrm{PR}}$. This means that high values of PS-$\mathcal{L}_{\mathrm{PR}}$ indicate a bad position reconstruction accuracy.
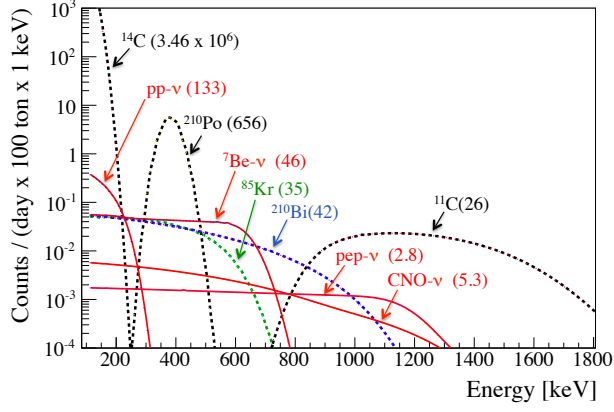
Figure 2.10: Expected contributions to the Borexino energy spectrum during Phase I. The numbers in parenthesis are the expected event rate in cpd/100 t. Figure from [82].

of both background events and electrons scattered by solar neutrinos to the Borexino energy spectrum during Phase I. In order to separate background and neutrino events and to constrain their rate, a fit is performed on the data.

The fit methods and procedures are described in the next Chapter, that presents the development of a new fit tool that has been used to obtain the results presented in this thesis.

The unique radiopurity of the Borexino detector allowed to pursue other scientific goals besides the measurement of solar neutrinos summarized in Sec. 1.4.3. In particular, the large distance from active nuclear plants makes Borexino an ideal candidate to measure the flux of electron anti-neutrinos produced by the decay of radioactive isotopes in the Earth interior. The measurements of geo-neutrinos in Borexino [123–125] confirmed the presence of the geoneutrino signal with a 5.9σ significance and constrained the radiogenic heat production from ^{238}U and ^{232}Th inside the Earth in between 23 and 36TW.

CHAPTER 3

TOOLS AND PROCEDURES FOR A MULTIVARIATE FIT

Despite the outstanding radiopurity of the Borexino experiment, the relatively low event rate and the presence of background make the extraction of the solar neutrino signal challenging. Signal and background events are often indistinguishable on an event–by–event base and their discrimination relies primarily on a fit of the energy spectrum.

The Borexino analysis methods developed before 2017 (briefly described in Sec. 3.1) were successfully used to obtain relevant physics results, but showed some limitations when including additional variables to the analysis to improve the sensitivity to the neutrino signal. This motivated the development of a new fitting tool, designed to have a full multivariate approach, be more flexible and improve the computational performance. In this chapter, all the steps of the development of the fitting tool are reported, starting from the conceptual structure and its benchmarking against the previous analysis described in Sec. 3.2. Thanks to the new tool a new analysis procedure became possible and has been tested. A major improvement with respect to the previous analyses concerns the inclusion of other variables in addition to the energy observables that now are treated as additional dimensions. The construction of the multi-dimensional models needed for this kind of analysis is summarized in Sec. 3.3, while in Sec. 3.4 an optimized choice of the binning is presented. The details of the fit configuration that will be used on data are briefly discussed in Sec. 3.5.

3.1 PRIOR STATE OF THE ART

What makes the Borexino experiment unique in the landscape of neutrino experiments is its capability to measure the signal of low energy neutrinos and perform a spectral analysis extended to very low energy. As for any spectral analysis, a crucial ingredient is the accurate knowledge of the expected signal, of the relevant sources of background and of the detector response modelling.

Each of the energy observables introduced in the previous chapter is characterized by a specific response function that describes its probability distribution function when energy is deposited in the detector. The energy response depends on several factors such as the position inside the detector, the type

of interacting particle, the scintillator properties as well as the status of the detector and of the readout electronics at the time of the event.

The analysis of Borexino adopts two independent models of the detector response: the first is an analytical model of the response function that provides the mean and the variance of the energy estimator as a function of the deposited energy; the second model is based on the Monte Carlo simulations of the detector behaviour.

The so-called *analytical* response function [82] is derived taking into account both the properties of the scintillator, such as the light yield, and of the detector, *e.g.* the spatial uniformity of the light collection and the number of active PMTs. The theoretical spectrum of each component is convolved with the response function. Rates and some parameter of the detector response, such as the light yield, are left free in the fit while the rest of the response function parameters are fixed after they have been tuned with Monte Carlo and calibration data. In this approach the free parameters of the response function are determined by the data themselves during the maximization of the likelihood, which makes the analysis less prone to systematic uncertainties concerning the detector response modelling. On the other hand, additional nuisance parameters contribute in enlarging the statistical uncertainty on the measured rates.

Another way to obtain the energy response is by performing a detailed simulation of the whole experiment. The Borexino Monte Carlo simulations [126] is a custom package based on Geant4 [127] that carefully reproduces the detector geometry and all the physical processes following the interaction of a particle. Events are generated in the detector according to their expected spatial distribution and photons emitted after the energy deposit in the liquid scintillator are tracked individually until they are absorbed or reach the PMT photocathode. The PMT detection probability is defined according to the characteristic quantum efficiency at each photon wavelength. The full electronic chain is also simulated and the output of the simulation is processed using the same reconstruction code used for real data. Great attention is given to the simulation of the optical properties of the materials that have been carefully tuned with calibration data [128] and benchmark measurement, making it possible for the simulation to achieve a sub-percent accuracy [126]. An additional crucial requirement for a reliable simulation is the possibility to take into account the detector evolution in time. In particular the deformation of the inner vessel, which impacts locally on the overall optical absorption length and on the external background contamination, is traced on a weekly basis. The characteristic dark rate of each PMT, its effective quantum efficiency and its gain, that are later used in the processing of the events, are also taken from the data every week.

In this fit approach, the model of each spectral component is obtained simulating ≈ 200 times the number of events in the data. The output of the simulation consists in files resembling the raw data that are processed through

the very same reconstruction software. In the spectral analysis the only free parameters are the rates that act like scaling factors of the models.

In the previous analyses performed on the Borexino spectrum, the fit range and the dataset were optimized according to the solar neutrino component that was under consideration [80, 84, 86]. An ambitious goal of the Borexino analysis pursued in the last years was to perform a simultaneous analysis of *all* the low energy components of the solar neutrino spectrum. As pointed out in Sec. 2.6.3, the ^{11}C background has a crucial impact on the sensitivity to the $\nu(pep)$ and $\nu(\text{CNO})$ signal. Carbon-11 rate can be suppressed by the TFC cut, but this reduces the available statistics for the ^{7}Be neutrino measurement. In order to maximise the sensitivity to both ^{7}Be and pep/CNO neutrinos, the TFC tagged and subtracted datasets are simultaneously analysed maximizing a likelihood given by the product of the likelihoods of the two datasets

$$\mathcal{L}_{\text{tot}}(\boldsymbol{\theta}) = \mathcal{L}_{\text{TFC}}^{\text{sub}}(\boldsymbol{\theta}) \cdot \mathcal{L}_{\text{TFC}}^{\text{tag}}(\boldsymbol{\theta}) \tag{3.1}$$

where $\mathcal{L}_{\text{TFC}}^{\text{sub,tag}}(\boldsymbol{\theta})$ is a binned Poissonian likelihood

$$\mathcal{L}_{\text{TFC}}^{\text{sub,tag}}(\boldsymbol{\theta}) = \prod_{i=0}^{N_{\text{bins}}} \frac{[\lambda_i(\boldsymbol{\theta})]^{k_i}}{k_i!} e^{-\lambda_i^{k_i}(\boldsymbol{\theta})} \tag{3.2}$$

with the index i spanning the bins of the spectrum and k_i and $\lambda_i(\boldsymbol{\theta})$ being the number of entries and the expected number of entries in the i-th bin respectively. The parameters entering the likelihood are indicated by $\boldsymbol{\theta}$ and are partly shared by the TFC-subctracted/tagged terms. The rates of the neutrino components are common for the two datasets along with part of the nuisance parameters like the rate of some background contaminations. Other contributions to the background (like ^{210}Po and the cosmogenic ^{11}C, ^{10}C and ^{6}He) differently populate the TFC-tagged and TFC-subctracted datasets and their rates are kept independent (see Sec. 3.5).

The sensitivity of the analysis can be improved including additional information that helps constraining some components of the Borexino spectrum. The external background due to γ rays generated in the stainless steel sphere and in the PMT glass are more and more attenuated as they penetrate towards the detector centre, showing therefore a characteristic radial distribution. Adding the radial coordinate of each event can hence help increasing the accuracy of the analysis. Similarly, the pulse shape parameter introduced in Sec. 2.6.4.2 can improve the precision in the determination of the ^{11}C background that is critical for the measurement of *pep* and CNO neutrinos. The first analysis of this kind was already performed during Phase I [82, 84] but the lack of computing resources prevented the implementation of a full multivariate approach and the radial and Pulse Shape information were included in the fit by adding to the likelihood in Eq. (3.1) two additional multiplicative terms

$$\mathcal{L}_{\text{tot}}(\boldsymbol{\theta}) = \mathcal{L}_{\text{TFC}}^{sub}(\boldsymbol{\theta}) \cdot \mathcal{L}_{\text{TFC}}^{tag}(\boldsymbol{\theta}) \cdot \mathcal{L}_{\text{rad}}(\boldsymbol{\theta}) \cdot \mathcal{L}_{PS-\mathcal{L}_{pr}}(\boldsymbol{\theta}) \tag{3.3}$$

In such procedure the distribution of the events radial position and of the Pulse Shape parameter is integrated over a given energy range and treated as independent data sets. This strategy introduces correlations between the number of counts in different histograms since the same event falls both into the energy spectrum *and* in the pulse shape and radial projections; to solve this issue the normalization of the model for the radial/pulse shape distribution was set to be equal to the total number of entries in the projected histograms, leading to a likelihood term of the form

$$\mathcal{L}_{\mathrm{rad/PS}\text{-}\mathcal{L}_{\mathrm{pr}}}(\boldsymbol{\theta}) = \prod_{s=1}^{P} \prod_{j=1}^{M} \frac{[a_s \lambda_{js}(\boldsymbol{\theta})]^{k_{js}}}{k_{js}!} e^{-a_s \lambda_{js}(\boldsymbol{\theta})} \tag{3.4}$$

where the index s runs over the number of considered energy windows, while j spans over the bins of the projected histograms. Similarly to Eq. (3.1) $\lambda_{js}(\boldsymbol{\theta})$ and k_{js} indicate the expected and the observed number of entries of the j-th bin of the s-th projection. The a_s coefficients set the normalization of the model as described before, being defined as

$$N_s = \sum_{j=1}^{M} a_s \lambda_{js}(\boldsymbol{\theta}) \tag{3.5}$$

This fit procedure was extensively discussed in [82, 115] and validated against Monte Carlo generated pseudo-datasets.

This analysis strategy is not fully satisfying for two reasons. First, the standard multivariate fit does not take into account properly the correlations between the analysis variables that can be significant in case the detector energy response changes as a function of the radius. Secondly, the construction of the likelihood in Eq. (3.3) leads to a rigid structure of the analysis that does not allow to easily test different configurations or add new variables. These limitations motivated an overall review of the multivariate fit strategy and the implementation of a new analysis tool.

3.2   DEVELOPMENT AND BENCHMARKING OF A NEW FITTING TOOL

To overcome the limitations of the previous analysis method, a new fitting tool, named `bx-stats`, was developed with the intention to have a general, flexible tool for frequentist multivariate analysis. The new fit structure and its main features are described in Sec. 3.2.1 while the results of a benchmarking against the previous fitting tool are presented in Sec. 3.2.2.

3.2.1   *Structure of the `bx-stats` fitting tool*

The `bx-stats` software was built as an interface to the `m-stats` software package[1], a `ROOT`-based frequentist analysis tool designed to perform maximum likelihood fits of multidimensional histograms. In this framework, a multivariate analysis of energy, radius and pulse shape can be performed avoiding the "projections" onto one-dimensional histograms used in the previous analysis. The likelihood is the product of a Poisson function for each bin:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i_0}^{N_0} \prod_{i_1}^{N_1} \ldots \prod_{i_n}^{N_n} \frac{\lambda_{\boldsymbol{i}}^{k_{\boldsymbol{i}}}(\boldsymbol{\theta})}{k_{\boldsymbol{i}}!} e^{-\lambda_{\boldsymbol{i}}(\boldsymbol{\theta})} \tag{3.6}$$

with $k_{\boldsymbol{i}}$ and $\lambda_{\boldsymbol{i}}(\boldsymbol{\theta})$ denoting respectively the actual and the expected number of counts in the $\boldsymbol{i} = \{i_0, i_1, \ldots i_n\}$ bin.

In the configuration of the analysis the user can define several *datasets*, each one characterized by its own exposure, variable ranges and fit model. Indeed, for each dataset all the signal and background components must be listed specifying their name: the free parameters of the analysis are the rates of the single components, that are treated as one in case a component with the same name is present in multiple datasets. The different datasets are combined in the minimisation of the Poisson Negative Log-Likelihood (NLL) in an analogue way as shown for the TFC-subctracted/tagged datasets in Eq. (3.1)

$$-\log \mathcal{L}_{\text{tot}}(\boldsymbol{\theta}) = \sum_{n=1}^{N_D} \sum_{\boldsymbol{i}} \left(-k_{\boldsymbol{i}}^{(n)} \lambda_{\boldsymbol{i}}^{(n)}(\boldsymbol{\theta}) + \lambda_{\boldsymbol{i}}^{(n)}(\boldsymbol{\theta}) + \log \Gamma(k_{\boldsymbol{i}}^{(n)} + 1) \right) \tag{3.7}$$

where n denotes the specific datasets out of the N_D considered, the vector $\boldsymbol{i}$ indicates the bin of the multi-dimensional histogram corresponding to the n-th dataset and $\Gamma(x)$ is the Gamma function.

A common situation in a physics analysis where many parameters are involved is to rely on some other "control" measurements to constrain some of the nuisance parameters. The correct way to take into account the additional information in a frequentist analysis approach consists in including in the analysis an additional dataset for each of the control measurements. This is typically approximated adding a Gauss function as a multiplicative term to the likelihood for each of the constrained parameters θ_i. Therefore, if one of the fit parameters is already known from the results of an independent estimation, it is possible to include this information simply by adding to the NLL in Eq. (3.7) a penalty term of the form

$$-\ln \mathcal{L}_{\text{penalty}}(\theta_i; \mu_i, \sigma_i) = \frac{1}{2}\ln(2\pi) + \ln \sigma + \frac{1}{2}\left(\frac{\theta_i - \mu_i}{\sigma_i}\right)^2 \tag{3.8}$$

---

1 M. Agostini, `https://github.com/mmatteo/m-stats`

where μ_i and σ_i are the mean and standard deviation obtained from the controlled measurement of the parameter θ_i.

3.2.2 *Benchmarking of the new fitting tool*

The performance of `bx-stats` was benchmarked comparing its results against those of the previous fitting tool when the same dataset is analysed and using the exact same settings by the two softwares. In particular, the MINUIT settings were set identically including the minimisation algorithm, its tolerance and the maximum number of calls (Tab. 3.1). The PDFs were obtained for each component using the same simulations. The comparison was performed using N_h as energy observable and considering the *pep* fiducial volume cut (Sec. 2.6.2). The energy range was 120–1000 N_h and the binning was 1 N_h.

The simultaneous fit of multiple datasets with the new fitting tool was validated performing simultaneous (non multivariate) fits of many pseudo-datasets consisting of TFC-subctracted and tagged energy spectra. The validation of the multivariate fit approach was tested in a similar way including the radial distribution in addition to the energy spectrum.

SIMULTANEOUS FIT OF MULTIPLE DATASETS   The simultaneous fit on the TFC-subctracted/tagged datasets was validated by drawing a comparison between the results of the two fit methods when the same dataset is analysed.

The pseudo-data were produced using `bx-stats` and stored to disk. These pseudo-datasets, consisting in a TFC-subctracted and tagged spectrum, are sampled injecting the rates reported in Tab. 3.2 and assuming both a realistic exposure (corresponding to 4 years of data taking) and an unrealistically large one (100 times larger). The fit parameters are initialized in the same way in the two softwares, sharing the same starting value and fit initial step size.

Method	Max. Calls	Tolerance
SIMPLEX	1×10^4	0.1
MINIMIZE	1×10^8	0.1
Standard Minuit Settings		
MINIMIZE	2×10^3	0.1

Table 3.1: Settings of the minimization method used in the validation, unless otherwise specified. The Standard Minuit Settings usually used in the old approach are also shown for comparison.

To test the fit stability we studied the cases of "far" and "close" starting values of the fit parameters with respect to the actual injected rate. From here to the end of this section, every reference to the "close" and "far" starting points is to be referred to the rates reported in Tab. 3.2, where the "far" starting values were defined to be 0.1 cpd/100 t for all the components while the "close" ones correspond to 80% of the true injected rate.

Species	Injected rate (cpd/100 t)	Close start. value (cpd/100 t)	Far start. value (cpd/100 t)
$\nu(^7\mathrm{Be})$	48.0	44.0	0.1
$\nu(pep)$	2.8	2.3	0.1
$\nu(\mathrm{CNO})$	5.0	4.0	0.1
$^{210}\mathrm{Bi}$	20.0	16.0	0.1
$^{210}\mathrm{Po}$	250.0	200.0	0.1
$^{85}\mathrm{Kr}$	5.0	4.0	0.1
$^{11}\mathrm{C}$ sub.	2.8	2.2	0.1
$^{11}\mathrm{C}$ cmp.	65.8	48.0	0.1
Ext. bkg. $^{214}\mathrm{Bi}$	1.0	0.8	0.1
Ext. bkg. $^{208}\mathrm{Tl}$	2.0	1.6	0.1
Ext. bkg. $^{40}\mathrm{K}$	1.0	0.8	0.1

Table 3.2: Summary of the injected rates together with the corresponding starting points both for the values close to the injected ones and far from them.

In the case of huge statistics and close starting values the two fit methods give exactly the same results, as one can see from the clear correlation between the results returned by the fits for a subset of the fit parameters shown in Fig. 3.1a. The maximum likelihood estimator is unbiased for both software as the mean value of each component (reported in the bottom right boxes in Fig. 3.1a) is equal to the injected one (shown in Tab. 3.2). Setting a far starting value for all the fit parameters doesn't change the outcome of `bx-stats`. Conversely, the change of the starting point has an impact for the standard fit, that in some cases converges to a local minimum. This behaviour can be recognized from the distribution of the reconstructed rates of $\nu(\mathrm{CNO})$, $\nu(pep)$ and $^{210}\mathrm{Bi}$ since, as it will be extensively described in the next chapter, these components are strongly correlated and are very sensitive to fit instabilities. Figure 3.1b shows that while `bx-stats` correctly reconstructs the rate of CNO neutrinos and $^{210}\mathrm{Bi}$, in approximatively 4% of the cases the previous analysis returns a null CNO rate and, as an effect of the existing correlation, very high rate of $^{210}\mathrm{Bi}$.

Assuming a nominal statistics of 1500 days, the situation changes quite drastically as both the fit methods are not able to resolve the above mentioned interplay among the $\nu(\mathrm{CNO})$, $\nu(pep)$ and $^{210}\mathrm{Bi}$ rates. When using close starting values for the fit parameters, the two methods give the same results as in the previous comparison. When setting the starting values far from the actual injected rate the standard fit appears to be more prone to converge to a local minimum, showing an excess of instances where the CNO rate rails to zero with respect to the `bx-stats` results. The better stability of `bx-stats` can be explained by the absence of numerical approximations that were introduced in the previous fit in order to have a common interface for both the Monte Carlo and the Analytical fit method.

(a) Close starting values

(b) Far starting values
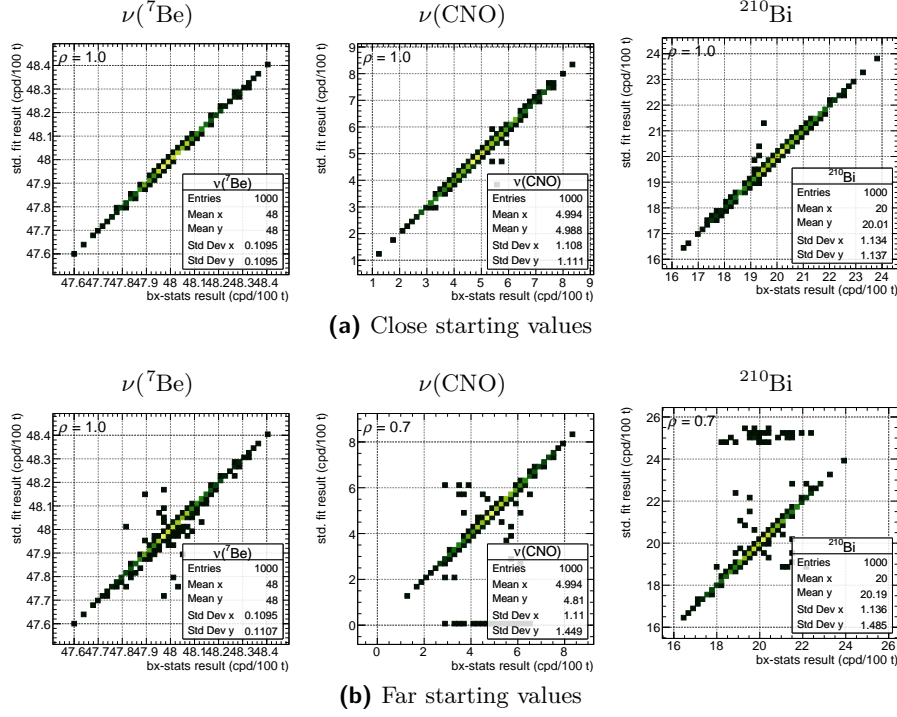
Figure 3.1: Correlation between the best estimate obtained using the standard fit tool and `bx-stats` on the same datasets generated assuming a large sample imitating a ≈ 400 years long data taking and setting close (top) and far (bottom) starting values for the fit parameters. In the top left corner of each panel the correlation coefficient is reported.

SIMULTANEOUS FIT OF n-DIMENSIONAL DATASETS   One of the most interesting features of `bx-stats` is the fact that for each spectral component it can handle multi-dimensional PDF with an arbitrary number of dimensions. The fact that `bx-stats` performs the multivariate fit using n-dimensional histograms and PDFs does not allow a direct comparison with the existing fitter that uses an approximated likelihood. Nonetheless, the closest possible configuration for the two methods was reproduced in order to study the impact of such an approximation.

In the previous multivariate fit approach only two radial distributions are considered: one for the uniform "bulk" components and one for the external background. Neutrino events, as well as cosmogenic and intrinsic contamination[2] backgrounds are uniformly distributed in the volume, and their reference PDF was built using a toy-MC technique sampling events inside the *pep* fiducial volume without assuming any energy dependence. The same method could not be applied to the external background since the spatial distribution of external γ interactions cannot be easily modelled and therefore had to be

---

2 This is not the case for ^{210}Po contamination, as it will be shown in Chap. 7. However ^{210}Po events are out of the region where the contribution of external background is relevant, and therefore they are excluded from the radial analysis.
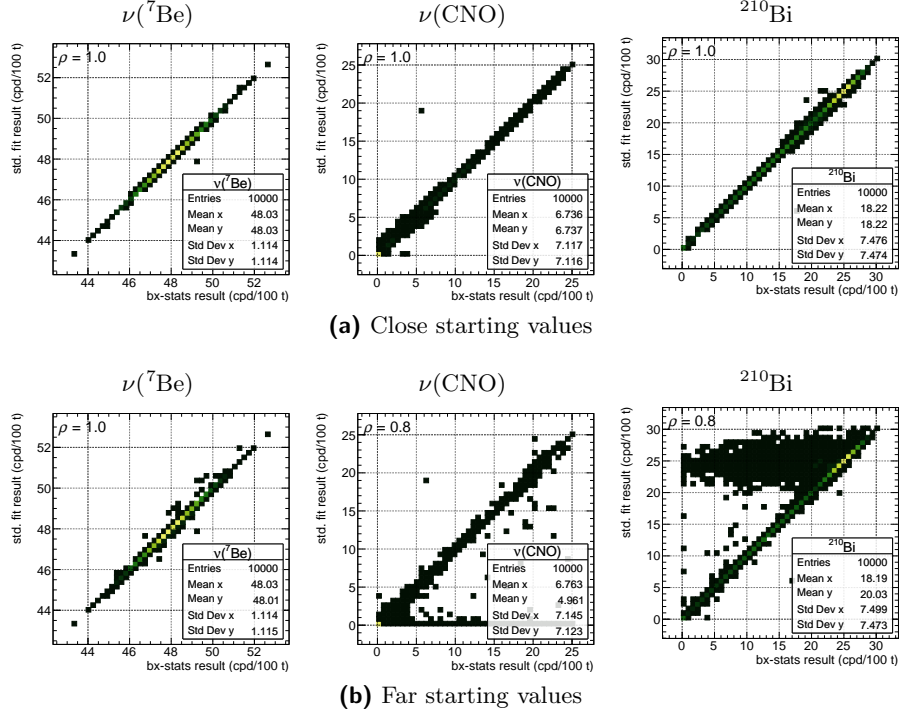
Figure 3.2: Correlation between the maximum likelihood estimator obtained using the standard fit tool and `bx-stats` on the same datasets generated assuming an nominal exposure of 1500 days of data taking and setting close (top) and far (bottom) starting values for the fit parameters. In the top left corner of each panel the correlation coefficient is reported.

built using a Monte Carlo simulation; for the purpose of this test the radial distribution of ^{208}Tl was assumed for all the external background components since this was the component for which the largest Monte Carlo production was available. To make a fair comparison between the two fit methods, the multivariate PDFs used by `bx-stats` were modified to impose the same radial distribution to all the bulk components removing any dependence on the deposited energy, and the same was done for external background PDFs.

The radial fit was performed between 500 and 900 N_h with both the fitters, while the complementary fit on the ^{11}C sub./tag. energy spectra was performed with the same settings used in the previous tests.

The scatter plot in fig. 3.3 shows that the results obtained by the previous fit and `bx-stats` are compatible for each realization even for the external background component, despite the fact that they are treated in a different way in the two analysis methods.

The comparison between the results of the previous fit and `bx-stats` was a crucial point in the validation of the latter. No relevant discrepancies were found in the tests presented in the previous sections that covered every possible configuration of the fit procedure.
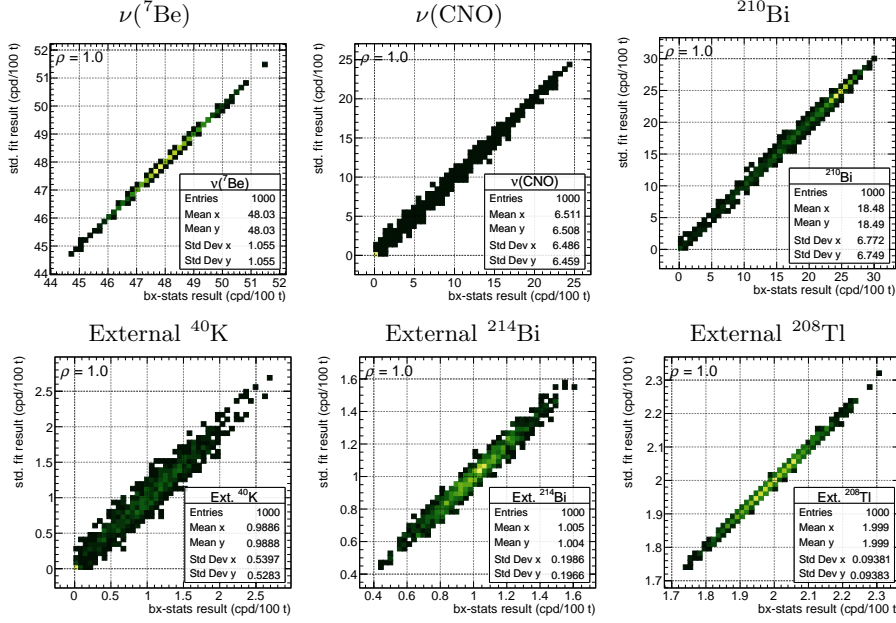
Figure 3.3: Correlation between the best estimate obtained by the previous fitting tool and `bx-stats` when performing a multivariate analysis considering the energy and radial distributions on the same MC-generated datasets. The datasets were created assuming a live time of 1500 days and setting the starting value of the fit parameters close to the actual injected rate.

The new fitting tool then reproduces the results provided by the standard fit, but also brings some significant advantages:

- *Increased stability:* Thanks to the fewer numerical approximations, `bx-stats` is more stable with respect to the previous fit, converging to the global minimum even with far starting values.

- *Speed:* The common interface of the analytical and Monte Carlo fit in the previous fitting tool requires many numerical conversions. On the contrary, `bx-stats` directly implements the histogram fitting, resulting much faster than the previous fit. The time needed for a simultaneous fit is reduced from a few minutes to approximatively one second and from more than two hours to a few minutes for the multivariate fit (in energy and radius). A test shows that the net improvement per call is close to a factor 40 (fig. 3.4). This improvement in the performance opened the way to a number of studies that have never been considered because of the unsustainable required computational time, including a detailed evaluation of the systematic uncertainties.

- *Flexible and multivariate:* thanks to an accessible interface, `bx-stats` makes it easier to perform a combined analysis on multiple datasets, and the full multivariate design allows to flexibly add new variables for

a multivariate fit without concerns about the unusual approach used in the previous multivariate analysis.
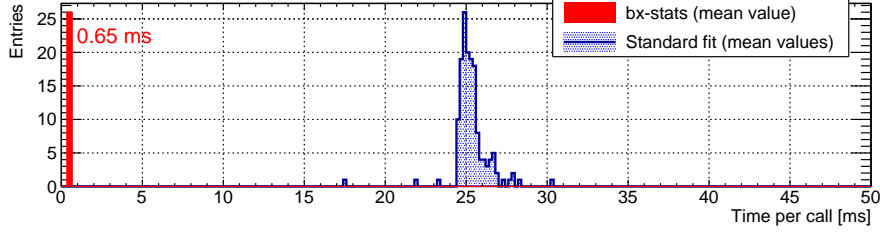
Figure 3.4: Average time per call for `bx-stats` and for the Standard fit. This result was obtained generating 500 pseudo-datasets and fitting them with both the methods, setting for both of them a maximum number of calls too small to reach the convergence.

3.3 MULTIVARIATE PDFS PRODUCTION

The PDF for each of the components considered in the fit model are produced using a comprehensive simulation of the Borexino detector following the same approach used in [114] and summarized in this section. The procedure adopted for the production of most of the components of the fit model is described in Sec. 3.3.1. Some components require a dedicated treatment: the simulation of the external background and other non uniform components is discussed in Sec. 3.3.2, while the production of the abundant low–energy ^{14}C and the connected pile-up events are presented in Sec. 3.3.3. The modelling of the pulse shape parameter, motivated by the increased dimensionality of the PDFs, is discussed in Sec. 3.3.4.

3.3.1 *Standard spectral components*

The PDF of any spectral components with energy higher than the ^{14}C and uniformly distributed inside the detector are produced according to the following procedure.

The decay of background isotopes and the interactions of neutrinos in the scintillator are simulated with Monte Carlo methods which cover also the propagation of the scintillation photons and reproduce the full electronics and reconstruction chain as described in Sec. 3.1.

To take into account the trigger and software reconstruction efficiency of each component, the PDFs are normalized using the ratio between the number of reconstructed events and the total number of events simulated in the considered fiducial volume.

3.3.2 *Non uniform components*

The background due to γ-rays produced by the decays of ^{40}K, ^{214}Bi and ^{208}Tl in the PMTs and in the stainless steel sphere has a radial dependence since the γ-rays flux is attenuated exponentially as a function of the distance from the production site. These isotopes are generated at their production site outside the inner vessel and propagated to the fiducial volume applying an importance sampling algorithm to improve the computational efficiency of the simulation [114, 126]. A Monte Carlo simulation of the γ interaction allows to take into account the energy dependence of the cross section, that results in a different radial profile for the different components of the external background (Fig. 3.5).
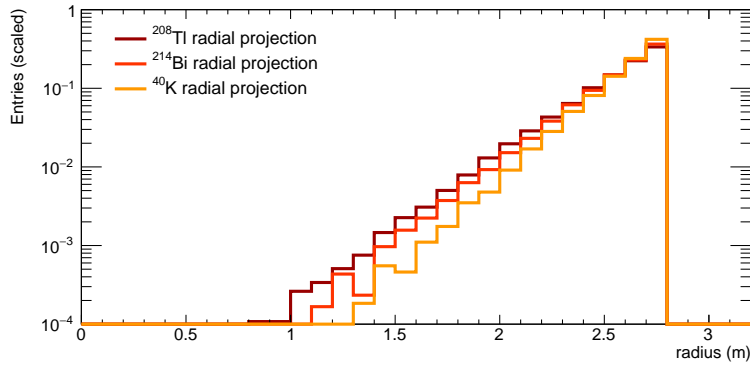
Figure 3.5: Radial distribution of external background events generated with a Monte Carlo simulation.

Another background component that is not uniformly distributed in the detector is ^{210}Po. As it was described in Sec. 2.6.4.1 α events are identified with very high efficiency using a pulse shape discrimination method based on the output of a MLP algorithm. In this way it was noted already in 2013 that the concentration of ^{210}Po is not constant in the detector volume, but has strong local non-uniformities that evolve with time. This behaviour will be discussed in Chap. 7. The position inside the detector affects the energy response function and for this reason ^{210}Po was simulated reproducing the actual distribution of the events in space and time [114].

3.3.3 *Low energy components*

The Carbon-14 present in the liquid scintillator, with its endpoint at 156 keV, is the most relevant background for the measurement of the *pp* neutrino rate. The low energy deposited in the scintillator by these events impacts on the accuracy of the position reconstruction algorithm as only a few hundreds scintillation photons are emitted. To correctly model the ^{14}C events reconstruction in the analysis volume, decays in the whole scintillator volume must be generated.

It is by far the largest radioactive contamination of the scintillator, being its rate independently evaluated to be $40 \pm 1 \, \mathrm{Bq}/100 \, \mathrm{t}$ [86]; the huge ^{14}C rate makes hard to simulate even a fraction of the actual number of events accumulated in over 4 years of data taking. The number of simulated events is indeed only $\approx 1/15$ of the one in the data [114], thus introducing in the PDF statistical fluctuations much larger than the ones present in the data. In the previous MC fit strategy this issue was taken into account by artificially increasing the uncertainty of the data to make it compatible with the one of the MC PDFs for those bins in which the MC PDFs were underpopulated, suppressing the impact of this part of the spectrum in the total likelihood [114].

In this thesis a different approach was adopted: with the assistance of Xue-Feng Ding, the ^{14}C spectral shape was taken from a fit of the ^{14}C spectrum with an analytical model using an independent dataset from the one of neutrino analysis. The ^{14}C spectrum was obtained selecting those events that happened in the $16 \, \mu s$ that followed a trigger started by a previous event. These "second-cluster" events are not affected by the energy threshold needed to trigger the acquisition and therefore allow to reconstruct a wider portion of the ^{14}C spectrum (Fig. 3.6a). The spectrum was fitted using the *analytical* approach introduced in Sec. 3.1 and the resulting shape was used in place of the MC generated events to build the ^{14}C PDF. A comparison of the relevant high-energy tail of ^{14}C described by the low-population original MC PDF and the one arising from this "hybrid" approach is shown in Fig. 3.6b.

(a)   **(b)**
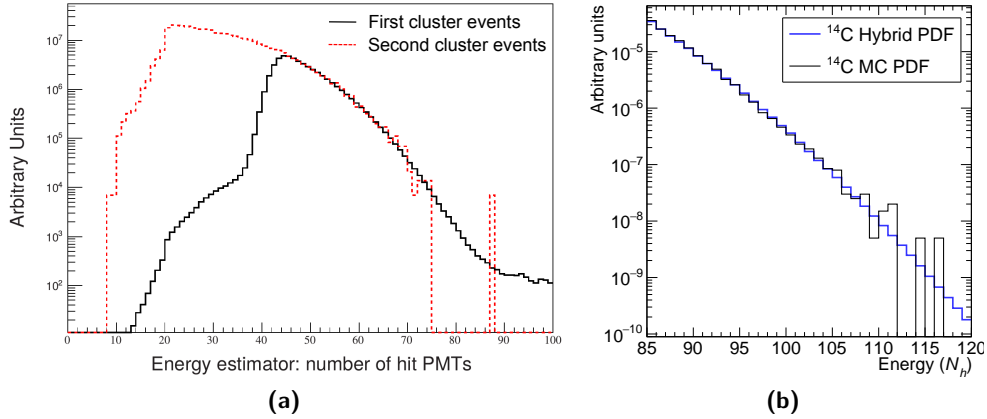
Figure 3.6: (a) Comparison of the spectrum obtained with self-triggering events (black) and by selecting events happening within $16 \, \mu s$ (red). Figure from [86]. **(b)** Comparison between the low-populated MC ^{14}C PDF and the one obtained from the fit of second-cluster events with an analytical model of the detector response.

The rate of ^{14}C is indeed so high that it is possible that a ^{14}C event can happen during the formation of the signal of a previous event, so close in time that the two events cannot be reconstructed separately. The random coincidences of multiple events in the detector, here referred to as pile-up,

strongly affects the low energy part of the spectrum whose understanding is crucial for the determination of the *pp* neutrino rate.

The pile-up energy distribution was studied in detail in [86] where two main strategies were adopted: in the first the PDF of each spectral component was convolved with the spectrum of random triggers events, while in the second the so called "synthetic" pile-up spectrum was built by artificially overlapping hits recorded after the end of a trigger window to the ones at the start of the cluster starting time.

More recently, after major improvements of the Monte Carlo simulation, it became possible to simulate also the pile-up spectrum. As described in [114, 126], pile-up components are produced generating two events correlated in time and then processed as standard events simulating the electronic chain and using the same reconstruction algorithm used for data. This procedure has the benefit to reproduce the physical process generating the pile-up, including those cases where none of the two events could trigger the detector individually and that are not reproduced by the synthetic method [126]. The components identified as relevant for pile-up were [126]:

- ^{14}C coincidences with degraded γ-rays

- ^{14}C–^{14}C coincidences

- ^{210}Po coincidences with degraded γ-rays

where degraded γ-rays here refer to those photons produced by decays in the PMTs and in the supporting structure that, conversely to those described in the previous section, already lost most of their energy in the buffer via Compton scattering and reach the FV with an energy below the trigger threshold.

Their relative weights were obtained by fitting the synthetic pile-up spectrum and their contribution to the total pile-up PDF is shown in Fig. 3.7.
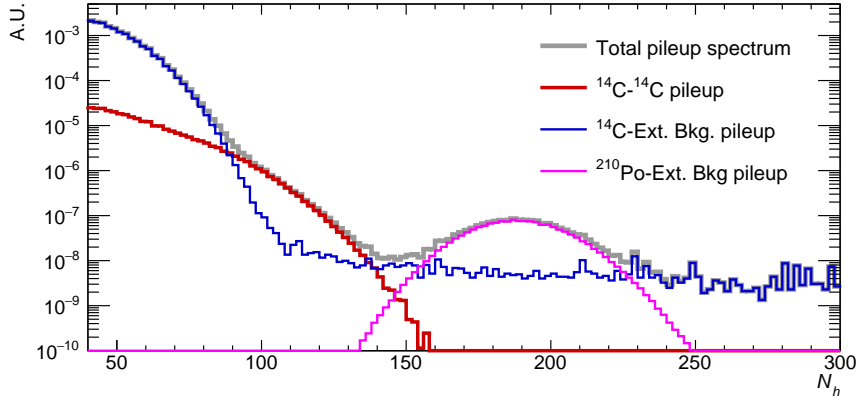
Figure 3.7: Monte Carlo PDF of random coincidence events, with the different contributions highlighted.

3.3.4   *Pulse Shape Parameter modelling*

The increased dimensionality of the PDFs introduced in this fit method has the effect to reduce the number of entries per bin. For a limited number of simulated events, the uncertainty due to statistical fluctuations in the PDFs becomes relevant. Since additional Monte Carlo simulations would have been unpractical, in order to reduce the complexity of the statistical treatment [129–131] the pulse shape variable is described using a heuristic model. This decision was supported by the fact that, as discussed in Sec. 2.6.4.2, the pulse shape variable does not show a strong energy dependence and that it was not possible to test with adequate accuracy its dependence on the events position without a dedicated calibration campaign.

The PDF of the pulse shape parameter was obtained fitting the distribution of the pulse shape variable of simulated events in the ^{11}C energy range with a heuristic model: while the electron-like components show a distribution compatible with a Gaussian, eternal background is described with the convolution of a Gaussian with an exponential,

$$f_\gamma^{\mathrm{PS}}(x) = \tfrac{1}{2} \exp\left\{\tfrac{1}{2}\lambda \cdot (2\mu + \lambda\sigma^2 - 2x)\right\} \cdot \mathrm{erfc}\left[\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right] \tag{3.9}$$

and the same model is used to model the pulse shape distribution for positrons, with the addition of second Gaussian to take into account the "bump" at $\mathcal{L}_{\mathrm{PS}} \approx 5.2$

$$f_{e^+}^{\mathrm{PS}}(x) = (1 - \epsilon) \cdot \frac{1}{2} \exp\left\{\tfrac{1}{2}\lambda \cdot (2\mu + \lambda\sigma^2 - 2x)\right\} \cdot \mathrm{erfc}\left[\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right] +$$
$$+ \begin{cases} 0 & \text{for } x < \mu - \sigma \\ \mathrm{Gauss}(x, \mu_2, \sigma_2) & \text{for } x \geq \mu - \sigma \end{cases} \tag{3.10}$$

The distributions of the pulse shape parameter in the energy range of interest for ^{11}C are shown in Fig. 3.8a with their analytic model. The Monte Carlo simulation reproduces with sufficient accuracy the shape of the distribution. The distribution of the pulse shape variable obtained from a subset of ^{11}C events was compared to the one obtain from MC generated events and a small shift was compensated in the model. The comparison between the pulse shape distribution of real ^{11}C events and the one described by the heuristic model is shown in Fig. 3.8b.

To reproduce the model in a MV PDF the procedure was the following: let's indicate the content of the i-th energy, j-th radial and k-th pulse shape bin as n_{ijk}. First the total number of entries for a given energy–radius bin is computed

$$N_{ij} = \sum_k n_{ijk} \tag{3.11}$$
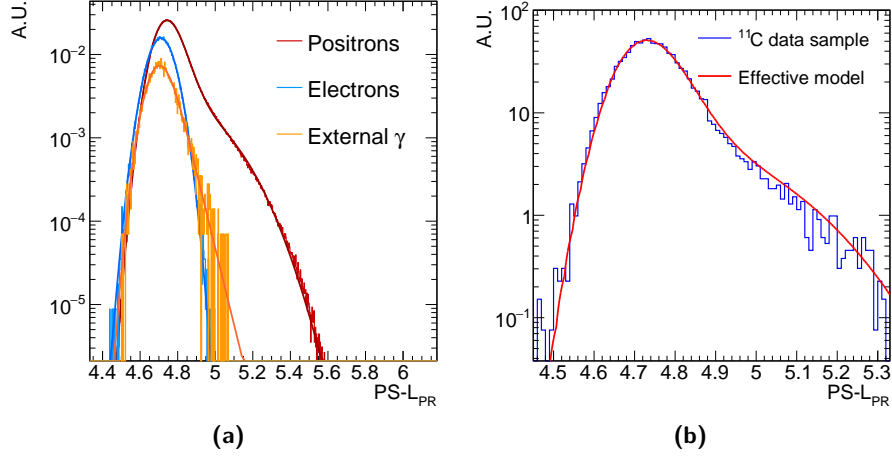
Figure 3.8: (a) Distribution of the pulse shape parameter in the ^{11}C energy range for positrons, electrons and external background in the *pep* FV obtained with a Monte Carlo simulation. The different scaling is set to better appreciate differences in the distributions. **(b)** Comparison between the $\mathcal{L}_{PR}$ distribution for a subset of ^{11}C events and the model derived from the MC distribution and corrected for a small shift observed.

then the bin content of the new "hybrid" PDF is defined as

$$\tilde{n}_{ijk} = N_{ij} \cdot \int_{\Delta_k} f_{PS}(x)\mathrm{d}x \tag{3.12}$$

where $f_{PS}(x)$ is the analytical PDF of the pulse shape variable and the integral is performed over the k-th bin. This method damps significantly the statistical fluctuations since in each bin the relative statistical uncertainty is not related to the number of entries in that bin but is $\sqrt{N_{ij}}/N_{ij}$.

3.4 BINNING OPTIMIZATION

The local event rate in Borexino varies significantly throughout the energy spectrum, going from ≈ 0.1 ev /day/100 t/N_h in the region of $\nu(pp)$ to ≈ 0.005 ev /day/100 t/N_h at the energy of interest for the search of the $\nu(pep)/\nu(\text{CNO})$ signal.

Similarly the events radial distribution grows as r^3 for uniformly distributed events, further increasing the inequality in the datasets bin content. The binning of the dataset was therefore optimized with the purpose of reducing statistical fluctuations providing only physical information to the analysis and mitigating the difference in the bin content between the highly- and low-populated bins in the process.

For what concerns the radial dimension, the natural choice was to adopt as a spatial variable r^3 in place or r. This option has the advantage of hav-

ing approximatively the same number of entries per bin for bulk components using a uniform binning. In addition, the outer bins —uniform in r^3— gets narrower with the increasing radius in the r variable, which is needed for an efficient discrimination of the external background. From studies performed on simulated datasets it emerged that even considering only five $5\,\mathrm{m}^3$-large bins (Fig. 3.9) the precision in the determination of the external background is already excellent and does not improve significantly when adding more bins.
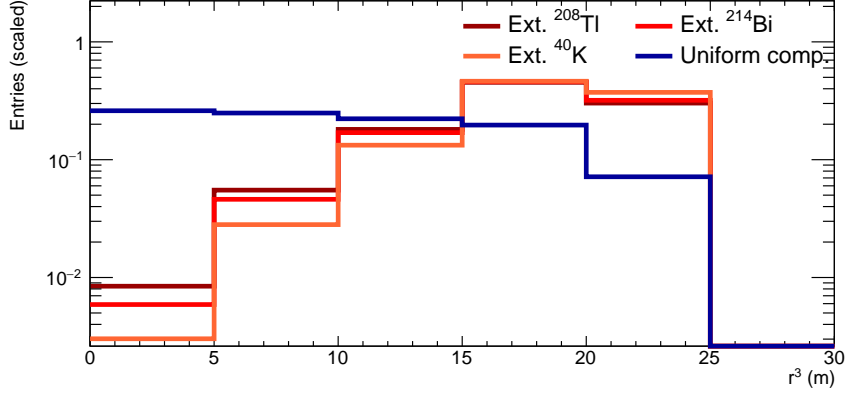
Figure 3.9: Distribution of r^3 as reconstructed by the MC simulation for the external background and a uniform component.

In the energy domain, for the N_h energy observables, the binning was chosen according to the detector resolution. Point-like, monochromatic e^- sources were simulated in the detector and the reconstructed N_h distribution width was studied (Fig. 3.10) as a function of N_h using a heuristic model of the form

$$\sigma(N_h) = a \cdot \sqrt{N_h} + b \cdot N_h \tag{3.13}$$

Using Eq. (3.13) as a model of the energy resolution it is possible to derive a new energy variable N_h' such that an increase of $1\,N_h'$ is equivalent to increasing N_h by $1\sigma(N_h)$. This variable can be defined simply imposing

$$\frac{\mathrm{d}N_h'}{1} = \frac{\mathrm{d}N_h}{\sigma(N_h)} \tag{3.14}$$

which results in

$$N_h' = \frac{2}{b} \log\left(1 + \frac{b}{a}\sqrt{N_h}\right) \tag{3.15}$$

Adopting the transformed variable N_h' uniformly binned is equivalent to use variable-size bins that scales as the resolution of the detector.

To be sure not to lose any detail of the spectral information and in order to have at least four bins in which the $\nu(pp)$ contribution dominates, the bin width in the energy domain was set to $\Delta = \frac{1}{4}N_h'$, corresponding to $\frac{1}{4}\sigma$, but the accuracy in the parameter reconstruction in the part of the spectrum with $N_h > 120$ is unaltered if $\frac{1}{2}\sigma$ bins are considered.
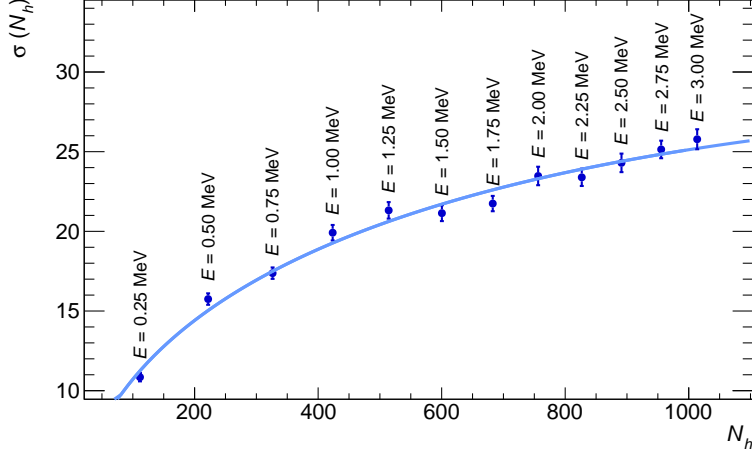
Figure 3.10: Width of the N_h distribution for point-like, monochromatic e^- sources simulated at $x = 2\,\mathrm{m}$ at different energies. The trend is fitted with the model in Eq. 3.13.

The N_h' energy variable is adopted in most of the results presented in this thesis. However, to allow an easier comparison with previous Borexino results the N_h energy observable will be quoted. Similarly, energy spectra will be shown using the equivalent variable-size bins in the N_h in place of the uniformly binned N_h' variable.

As a byproduct of the binning optimization the total number of bins was significantly reduced, contributing to further improve the computational performance of the analysis.

3.5 FIT CONFIGURATION

The configuration of the fit model that it is used to fit the data is slightly more complex than the one that has been used to benchmark the fit tool in Sec. 3.2.2. In addition to the components listed in Tab. 3.2, the background due to ^{14}C and random coincidences introduced in Sec. 3.3.3 are considered in the fit model.

The rate of most of the components is kept correlated in both the TFC-subtracted and -tagged datasets, but there are however some exceptions. The first, natural, case is ^{11}C for the reason that the TFC algorithm is implemented with the specific goal of suppressing its contribution in the TFC-subtracted spectrum. Similarly ^{10}C and ^{6}He, both of cosmogenic origin, are considered only in the TFC-tagged dataset since a time fit of events after muons shows that their contribution after the TFC cut is negligible.

Another component which has different rate in the two datasets is ^{210}Po. This is due to the fact that at the beginning of Phase II the data taking was unstable, with the TFC algorithm vetoing the entire detector for some hours after every run interruption (see Sec. 2.6.3). During the same period of time

the ^{210}Po contamination was higher in the detector, and as the concentration of ^{210}Po decreased due to natural decay, the duty cycle of the experiment sensibly improved, leaving a slightly higher content of ^{210}Po events in the TFC-tagged dataset, while all the other component –being their rate constant in time— are not affected by the variation trend of the duty cycle.

Neutrinos from ^{8}B are not considered in the fit since their rate in the 0.2–2.9 MeV energy range is of the order of 0.5 cpd/100 t and therefore negligible compared to other backgrounds.

3.6 CONCLUSIONS

The tools and procedures for a multivariate analysis of Borexino data were developed, benchmarked and optimized. The analysis presented in this chapter exploits the events energy, position and their pulse shape to constrain the solar neutrino signals and their backgrounds. The limitations of the previous multivariate analysis method have been overcome implementing the fit of multidimensional histograms which results in a more stable, fast and accurate fitting strategy.

The PDFs for signal and background components are created using large ensembles of simulated data. In order to reduce the statistical uncertainty associated to the PDFs caused by the finite number of simulated events, the distribution of the pulse shape variable was decoupled from the energy and position observables and described by a heuristic model.

The binning of the PDFs was optimized with the goal of reducing the impact of statistical fluctuations in the data while not losing any physical information. For this reason the third power of the radial coordinate was chosen as spatial variable. Similarly, a transformed energy observable derived from the number of reconstructed hits was introduced to reproduce a variable size binning scaling with energy proportionally to the standard deviation of the N_h estimator.

CHAPTER 4

SENSITIVITY TO LOW ENERGY SOLAR NEUTRINOS

This Chapter presents the sensitivity of the analysis method to the solar neutrino signal. A qualitative discussion of the main signatures of solar neutrinos in the Borexino data is presented in Sec. 4.1. Section 4.2 shows how similarities in the events distributions of some fit components generates correlations among the reconstructed rates. The parameters correlations affect the potential of the analysis for some components whose determination relies on additional information included in the analysis in the form of external constraints. The analysis sensitivity is presented in Sec. 4.3 reporting the expected statistical uncertainty of the reconstructed solar neutrino rates, while systematic uncertainties due to the fit model are discussed in Sec. 4.4. Finally, in Sec. 4.5 the study of systematics is used to determine the fit range which minimizes the overall uncertainty on the neutrino rate.

4.1 SIGNATURES OF SOLAR NEUTRINOS IN BOREXINO DATA

The interaction of solar neutrinos in the detector produces a specific signal in the data that allows to separate it from the background. In Sec. 4.1.1 the main features of the solar neutrino energy spectrum that are exploited by the analysis are presented. A MV analysis like the one described in the previous Chapter results sensitive to some second order effects in the data that were never considered previously. In particular, Sec. 4.1.2 describes the spatial dependence of the energy response whose precise modelling is one of the most important advantages brought by a MV analysis.

4.1.1 *Features of the solar neutrino signal in the energy spectrum*

The spectrum of neutrinos produced by the nuclear processes in the Sun core covers a wide energy range extending up to 18 MeV. The overwhelming majority of neutrinos, however, are emitted with an energy lower than 1.5 MeV. The interaction of these low-energy neutrinos with electrons in the active volume of the detector causes the electrons recoil that is eventually detected. The energy spectrum of recoiling electrons is continuous even for mono-energetic neutrinos sources such as ^{7}Be decays. The reconstructed energy distribution depends on the energy distribution of the incoming neutrinos and the contributions of

different sources can be extracted by fitting the energy spectrum. This task is particularly challenging because of radioactive contaminations in the liquid scintillator which create a background that mimics the neutrino signal.

The expected spectrum of Borexino for a 4 year exposure, assuming the solar neutrino rates predicted by the Standard Solar Model and a preliminary estimation of background levels, is shown in Fig. 4.1 after removing part of the ^{11}C events thanks to the TFC cut.
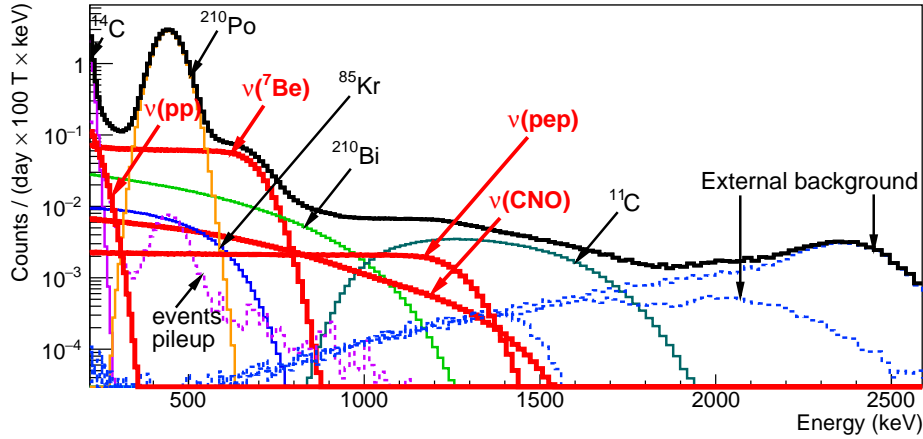
Figure 4.1: Expected energy spectrum after the TFC cut which removes most of ^{11}C events. Solar neutrino components are highlighted with bold red lines.

The very low-energy contribution of *pp* neutrinos (endpoint of 261 keV) is relevant only in a very narrow energy window right above the ^{14}C endpoint. The region of the spectrum where the analysis is sensitive to the *pp* neutrino signal can also be observed in Fig. 4.2, where the expected ratio between the signal and background fluctuations is shown. The contribution of ^{14}C–^{14}C pile-up gives a signal extremely similar to the one of *pp* neutrinos and therefore it is crucial to study it carefully.

Mono-energetic ^{7}Be neutrinos are the dominant component of the Borexino spectrum between 300 and 700 keV. Beryllium-7 decays via electron capture to the ground state of ^{7}Li with a branching ratio of 89.56%, emitting a mono-energetic neutrino of 862 keV. In the remaining 10.44% of the cases it decays to the first excited state of ^{7}Li generating a mono-energetic 384 keV neutrino followed by a 478 keV γ-ray due to the de-excitation to the ground state. Mono-energetic neutrinos generate to an electron recoil spectrum with a characteristic kink that is visible in Fig. 4.1 at 665 keV.

At higher energies, the solar neutrino signal is due to CNO and *pep* neutrinos. These two components are the weakest species of the solar neutrino spectrum that can be detected in such analysis. Together with ^{210}Bi, they dominate the energy spectrum between the $\nu(^7$Be) shoulder and the residual ^{11}C. External background dominates the higher energy part of the spectrum, affecting also part of the region of interest for *pep* and CNO neutrinos.
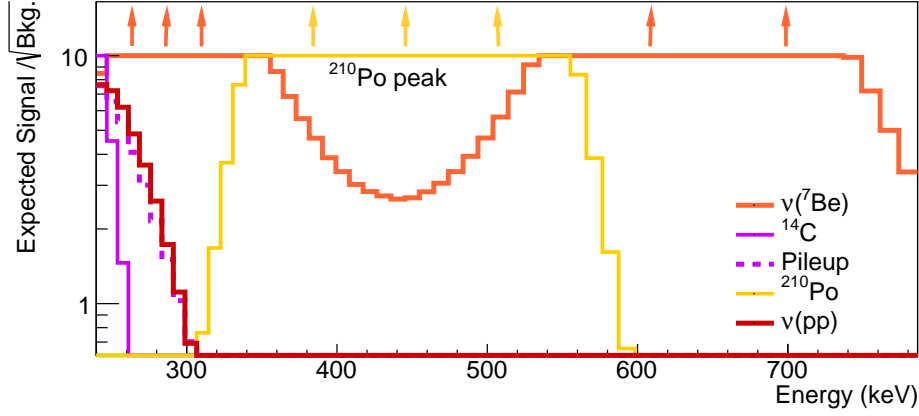
Figure 4.2: Signal strength of low energy fit components, computed as the ratio between the expected signal and overall background fluctuations assuming an exposure of ≈ 4 years.

The radial distribution plays an important role to distinguish among external background and uniform components: Fig. 4.3 represents the expected signal strength, defined as in Fig. 4.2 but in the energy–radius space. By looking at the figure it is therefore possible to identify the most important regions for the reconstruction of a given component, such as the distinctive radial pattern of the external background. Also the ^{11}C signal strength shows a peculiar structure in the energy–radius space due to the spatial dependence of the energy response that is exploited by the new multivariate approach.

4.1.2 *Spatial dependence of the energy response*

As mentioned before, one of the main advantages of a multivariate fit is that it takes into account the correlations among the variables used in the analysis.

For instance, Fig. 4.4 shows the reconstructed hits (N_h) and the charge energy observables ($N_{p.e.}$) as a function of the true energy for point-like, mono-energetic electrons simulated on the detector equatorial plane at different distances from the detector centre. The N_h observable exhibits a strong non-linearity in the energy response that depends on the radial position and gets more pronounced at large radii. The $N_{p.e.}$ variable instead appears to have a better linearity, but events generated far from the centre are reconstructed at higher $N_{p.e.}$ with respect to those simulated in the innermost part. This is due to the specific definition of the energy observable: if two or more photons originating from the same event are detected by one PMT in a very short time window, they will be reconstructed as a single hit; on the other hand, the integrated charge collected by the PMTs in case of multiple hits yields a larger signal that accounts for the real number of photoelectrons.

The spatial dependence of the energy response is mostly due to geometrical effects: the probability of multiple hits on the same PMT depends on the solid

(a) ^{210}Bi   **(b)** ν(CNO)   **(c)** ν(pep)

(d) ^{11}C   **(e)** External ^{214}Bi   **(f)** External ^{208}Tl
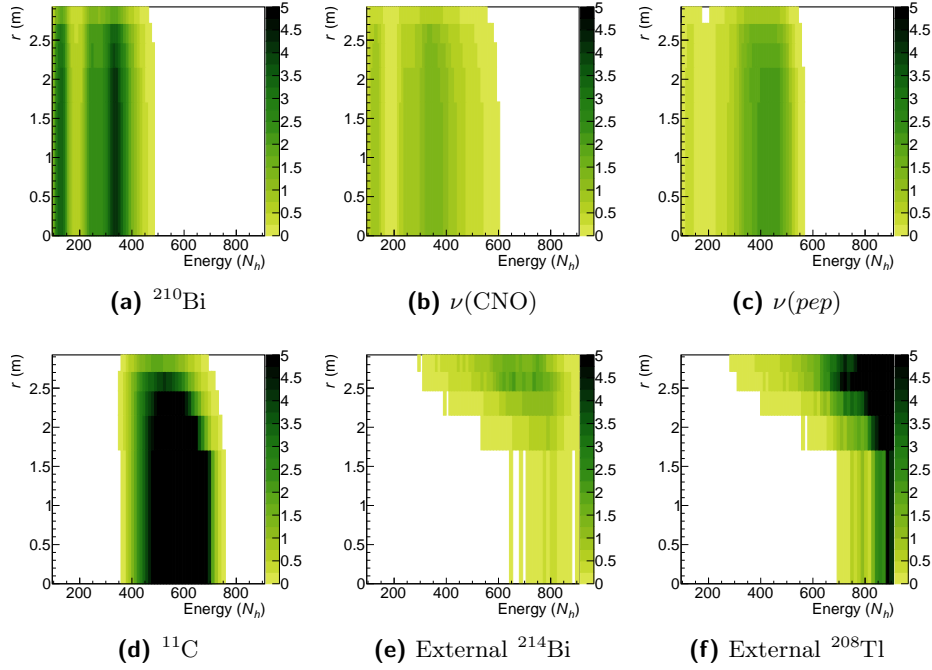
Figure 4.3: Expected signal strength of some of the components of the Borexino spectrum as a function of energy and radius. The signal strength is defined as the ratio between the expected signal and the fluctuations of the background.

angle covered by the PMTs while the number of absorbed photons depends on the distance between the interaction point and the PMTs. For this reason, events with the same energy are reconstructed with different $N_h/N_{p.e.}$ depending on their position inside the detector. Figures 4.4a, 4.4c shows that for the typical radii considered in the analysis ($r < 3.1$ m in fiducial volume cuts) the impact on the N_h variable is stronger at high energy, while it is less relevant for $E \lesssim 0.8$ MeV.

The "compression" of the energy scale observed at large radii in the N_h observable shifts the ^{11}C peak (Fig. 4.5a) affecting its radial distribution as well. Figure 4.5b shows the radial distribution of real events due to ^{11}C selected with a strict TFC cut compared with a sample of simulated events (that include the effect of N_h saturation) and with the distribution expected assuming a uniform sampling in the selected fiducial volume.

A correct modelling of the spatial dependence of the energy response is a fundamental ingredient for a multivariate analysis and the Monte Carlo simulation is expected to reproduce correctly these effects [126, 128]. If the influence of the interaction position on the energy response is not considered, the rates reconstructed by the analysis will be biased. Figure 4.6 shows the distribution of the reconstructed parameters obtained analysing an ensemble of pseudo–experiments. The datasets were generated including the spatial dependence of the detector response in N_h and then fitted considering the correct model
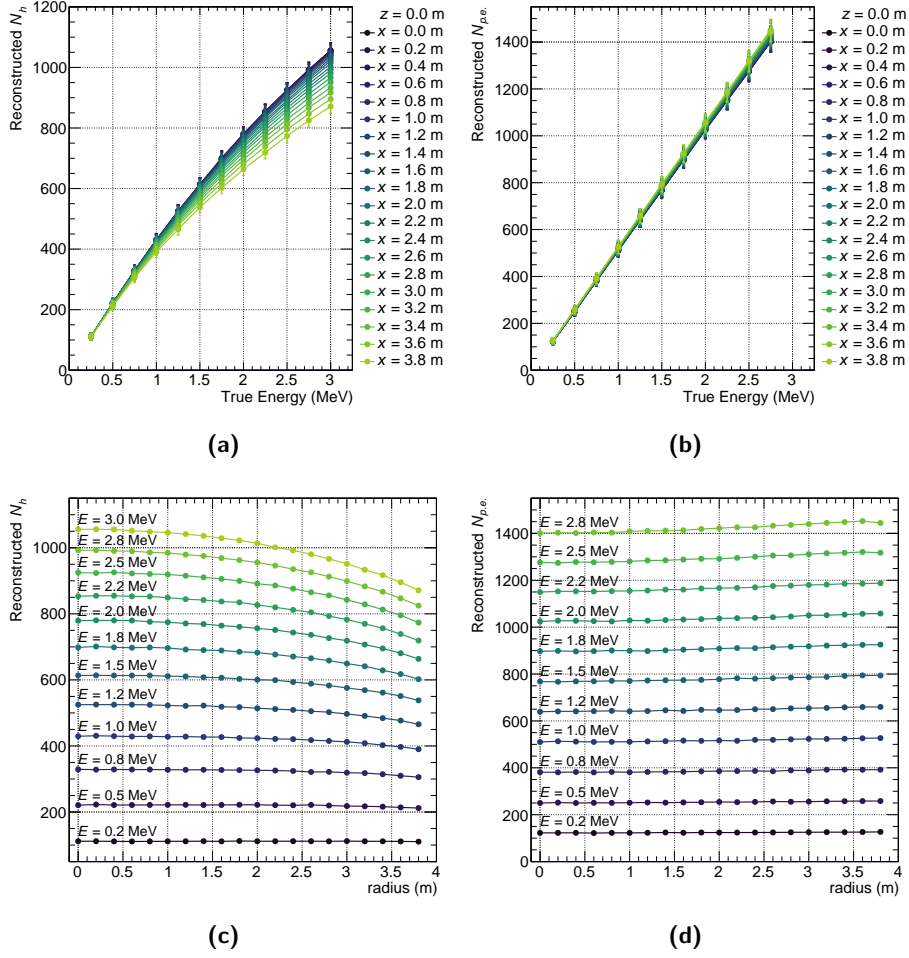
Figure 4.4: Reconstructed energy as a function of the position for point-like monochromatic e^- sources: Figures **(a)**, **(b)** shows N_h and $N_{p.e.}$ respectively as a function of the true energy, while the reconstructed energy at different radii for a given simulated energy is shown in **(c)**, **(d)**.

of the energy response and a simplified one where all the external background sources had the ^{208}Tl radial distribution while the bulk components radial distribution was obtained with a toy-MC sampling uniformly the *pep* FV.

4.2 PARAMETER CORRELATIONS AND EXTERNAL CONSTRAINTS

The features of the spectral shapes helping the fit that has been described in the previous section gives a qualitative hint of what components are within reach of the analysis. In particular, similarities in the events distributions of different fit components introduce correlations in the analysis that can be observed in the distribution of the reconstructed rates from an ensemble of pseudo-experiments. Such correlations reduces the sensitivity of the analy-

(a)

(b)
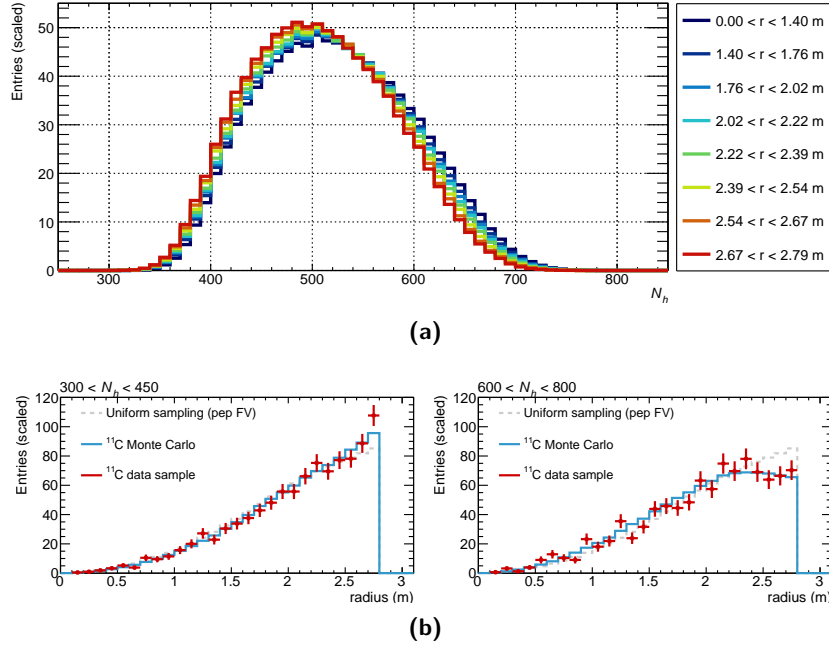
Figure 4.5: Effect of the position dependence of the energy response on ^{11}C. **(a)** ^{11}C obtained from a MC simulation at different radii within the *pep* FV. **(b)** Comparison between the radial distribution of of a clean data sample of ^{11}C events selected with a strict TFC cut and the predictions from a simulation for two different energy range. The expected distribution assuming no spatial effect on the energy reconstruction is also shown for reference.

sis and require independent estimates of the background rate which act like external constraints and break the above mentioned correlations.

Similarly to what was done for the validation of the fit procedure described in Sec. 3.2, thousands of pseudo-datasets have been generated assuming certain rates of the fit components and the exposure of the real dataset. These Monte Carlo generated datasets are analysed as real data and the distributions of the reconstructed rates, when the Wilck's theorem is valid, are Gaussian centred on the injected value and with the σ equal to the statistical uncertainty. Thus σ becomes increasingly large if the fit is not sensitive to a parameter.

Figure 4.7 shows the distributions of the reconstructed rates of some of the fit components (red histograms) as well as their correlations for 10^4 simultaneous multivariate fit of TFC-subctracted/tagged pseudo-datasets. A correlation is present between the reconstructed $\nu(\text{CNO})$ rate and the one of $\nu(pep)$ and ^{210}Bi. This effect is due to the similarity between the energy distribution of CNO neutrino events and the ^{210}Bi β decay spectrum (Fig. 4.1), combined to the fact that $\nu(\text{CNO})$, $\nu(pep)$ and ^{210}Bi components are effectively reconstructed based on the data in the same regions of the energy spectrum (Fig. 4.3). The analysis cannot resolve individually the single contributions but it is sensitive to the sum of the three components. The sensitivity to the CNO neutrino rate is indeed so low that the fit results can reach up to

(a) ^{11}C-sub.   **(b)** Ext. ^{40}K background
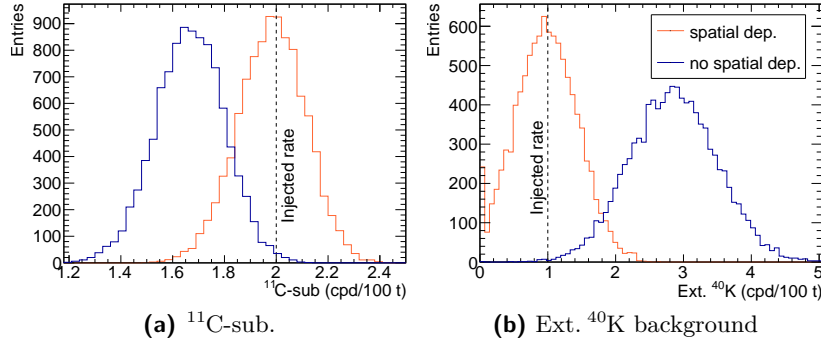
Figure 4.6: Distribution of the best estimate obtained from pseudo-experiments including the position dependence of the energy response in the generation of the dataset. Pseudo-data are fitted using two sets of PDFs that do and do not take into account this effect.

20 cpd/100 t when a rate of 4.91cpd/100 t is injected, with a non negligible fraction of the realization in which the best estimate of the CNO rate is zero. The physical boundary (the rate of CNO neutrinos cannot be negative) breaks the Wilck's theorem, making the distribution of the CNO reconstructed rate not Gaussian. This effect is propagated via the correlations in the fit to the ^{210}Bi and $\nu(pep)$ rate, deforming the best estimate distributions.

It is worth to notice that also the low energy part of the spectrum impacts on the ^{210}Bi reconstruction. Indeed in the narrow energy range between the ^{14}C dominated region and the ^{210}Po peak it is possible to see from Fig. 4.3 that ^{210}Bi accounts for a non negligible part of the signal. In the same region ^{85}Kr is present, and since there is no clear spectral feature to disentangle the two components a negative correlation appears and propagates to $\nu(pep)$ and $\nu(\text{CNO})$.

The correlation between $\nu(\text{CNO})$, $\nu(pep)$ and ^{210}Bi can be broken by constraining one of the parameters involved. In principle the best option would be to obtain an independent measurement of the ^{210}Bi background rate to improve the analysis sensitivity to the *pep* and CNO neutrinos: this is indeed crucial for the search of CNO neutrinos as it will be extensively described in Chap. 6, but at present such a constraint is not available.

The expectations from the Standard Solar Models can however be used to constrain the rate of CNO neutrinos. In the SSM, the CNO flux is predicted with an accuracy of $\approx 10\%$, but the central value depends significantly on the set of spectroscopic measurements used to evaluate the heavy elements abundance in the Sun (Sec. 1.2.3). The CNO rate has been constrained using both the prediction of the High and Low Metallicity Standard Solar Model applying a multiplicative Gauss penalty to the likelihood as indicated in Sec. 3.2. The parameter correlations for the CNO rate constrained to the High-Metallicity prediction are shown in Fig. 4.8. The $\nu(pep)$ and ^{210}Bi reconstructed rates are now described by Gauss distributions.
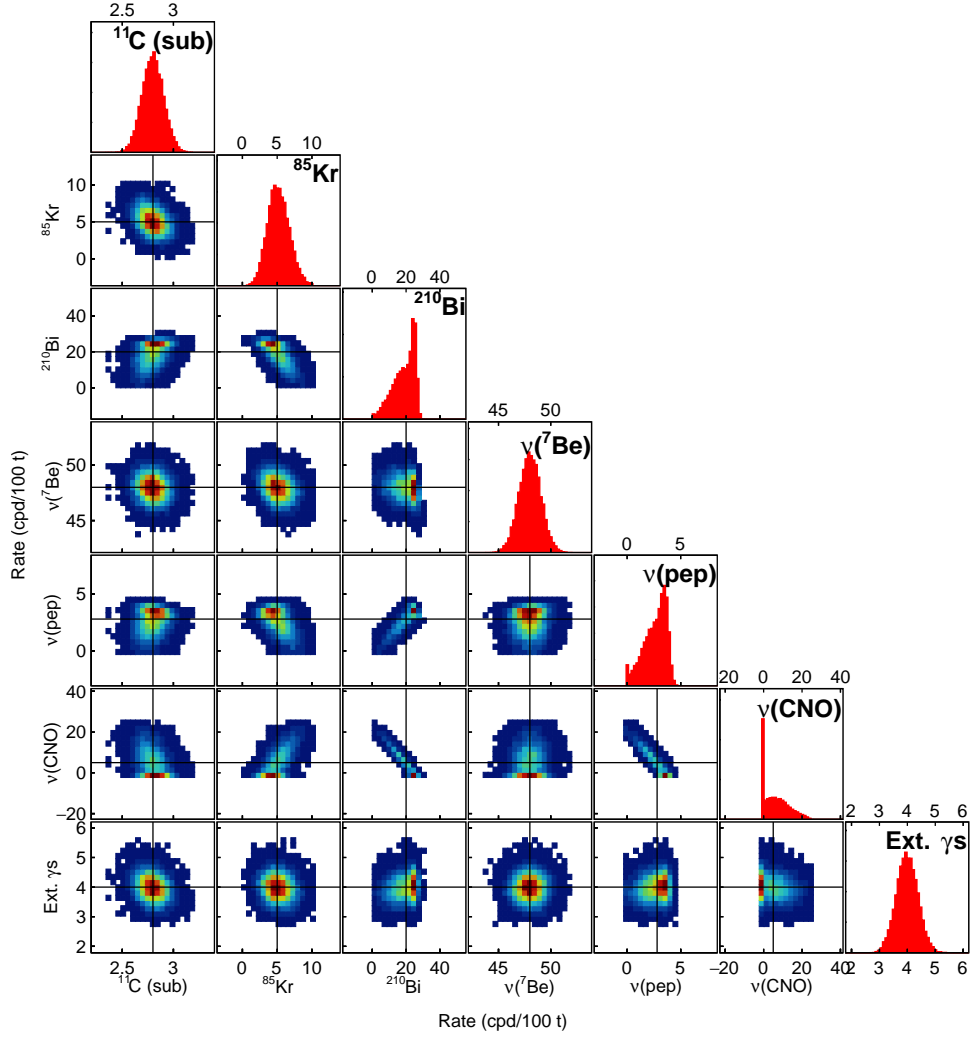
Figure 4.7: Distribution of the best estimate obtained from $\approx 2 \times 10^4$ pseudo–datasets.

At low energy, also the measurement of the *pp* neutrino rate is strongly affected by the background due to ^{14}C and multiple events coincidences (*pile-up*, see Sec. 3.3.3). In particular, the similarity between the pile-up and *pp* neutrino events energy distribution in the region of the spectrum where both are relevant (Fig. 4.1,4.2) makes it crucial to measure independently the ^{14}C and pile-up background rate and to constrain them in the analysis. The rate of ^{14}C contaminating the liquid scintillator was determined in [86] by fitting the "second cluster" events spectrum (see Sec. 3.3.3), finding a rate of $40 \pm 1\,\mathrm{Bq}/100\,\mathrm{t}$. With this information, it is then possible to compute the rate of ^{14}C–^{14}C events coincidences considering the mean duration of a cluster. The relative weight of the components of the overall pile-up PDF are fixed [126] and their rate is scaled accordingly, yielding a total pile-up rate of $8.458 \pm 0.081\,\mathrm{Bq}/100\,\mathrm{t}$.
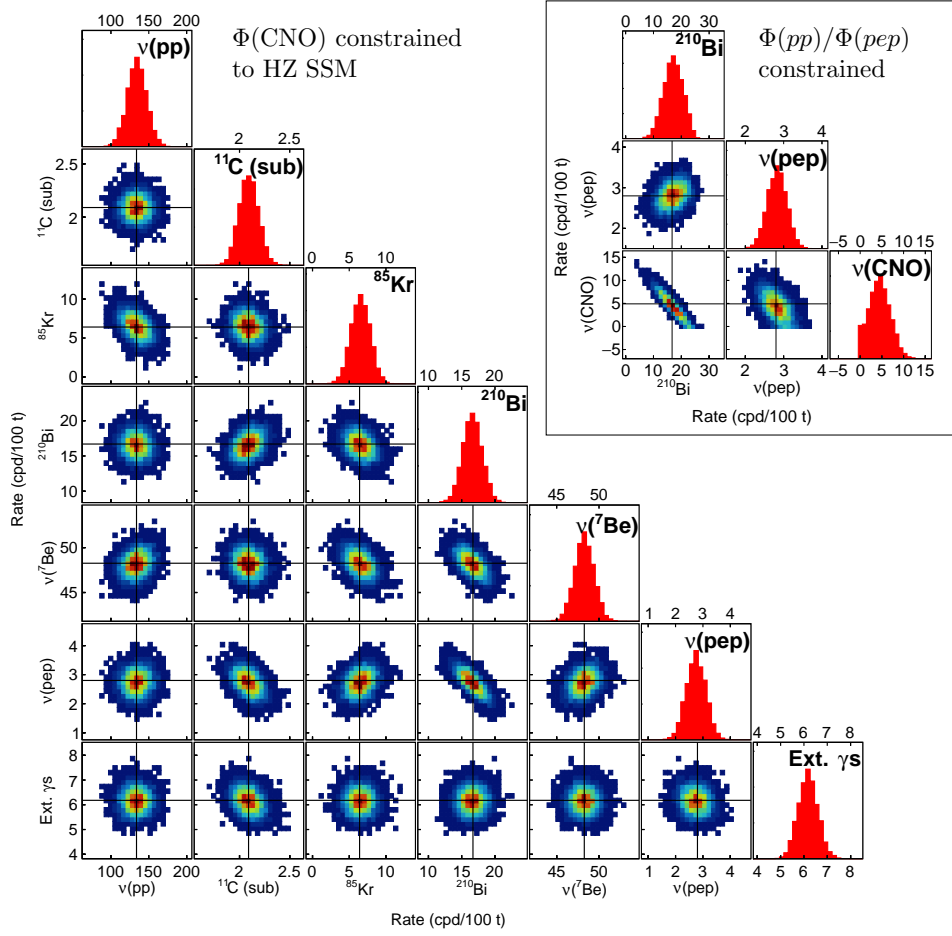
Figure 4.8: *Main figure* Distribution of the reconstructed rates from 10^4 pseudo–datasets when the $\nu(CNO)$ rate is constrained to the value predicted by the HZ Standard Solar Model. *Top right box* Distribution of the $\nu(CNO)$, $\nu(pep)$ and ^{210}Bi rate obtained constraining the ratio between the *pep* and *pp* neutrino rates.

Figure 4.8 shows that the best estimate of *pp* neutrinos rate is affected by the ^{85}Kr reconstructed rate. Indeed, with the background due to ^{14}C and pile-up already constrained in the analysis, a change in the number of events due to ^{85}Kr in the *pp* region induces an opposite variation in the number of $\nu(pp)$ events that results in the anti-correlation of the reconstructed rates.

The possibility to measure with relatively good accuracy the rate of *pp* neutrinos makes it possible to improve the sensitivity to CNO neutrinos. In the same way a constraint on the CNO rate enhances the sensitivity to *pep* neutrinos, an indirect evaluation of the *pep* neutrino rate can be used to obtain a better sensitivity for CNO neutrinos, mitigating the effect of the correlation shown in Fig. 4.7 even if the ^{210}Bi rate is left free in the analysis. To obtain an independent estimation of the *pep* rate it is possible to exploit the relationship between the *pp* and *pep* neutrino fluxes. The nuclear processes originating

the *pp* and *pep* neutrinos (see Sec. 1.2.2.1) are close relatives and take place almost in the same region of the Sun core. As it will be explained in Sec. 6.3.1, these facts make the ratio between the fluxes of *pp* and *pep* neutrinos basically independent from the SSM parameters and accurately determined only by nuclear physics. It is then possible to constrain the ratio between the rate of *pp* and *pep* neutrinos as done in [132]: using this link, the accuracy in the determination of the *pp* neutrino rate is propagated to the *pep* rate, bringing a substantial improvement to the sensitivity to the CNO signal as it is clearly visible from the comparison of the top right box in Fig. 4.8 with Fig. 4.7.

4.3 SENSITIVITY TO LOW ENERGY NEUTRINOS

The expected statistical uncertainty on the maximum likelihood estimators is obtained from the distribution of the reconstructed rates shown in Fig. 4.8.

The independent determination of the rate of ^{14}C, events coincidences (pile-up) and CNO neutrinos described in the previous section are crucial to break the correlations in the analysis. Constraining these backgrounds, the fit is sensitive to the rates of all the components. As a consequence, the reconstructed rates from the fit of pseudo–datasets are distributed according to a Gauss function, which proves that the Wilck's theorem holds and thus allows to consider the standard deviation of the distribution as the expected statistical uncertainty.

The results of this study are summarized in Tab. 4.1. The statistical uncertainty on the *pp* neutrino rate is at the level of 9.5% (12.7 cpd/100, t). The expected accuracy on the ^{7}Be neutrino interaction rate reaches a remarkable 2.4% level (1.13 cpd/100 t), the most precise ever achieved. The *pep* neutrino rate, once the CNO is constrained to the predictions of the HZ SSMs [34], is expected to be measured with a 12.9% precision (0.36 cpd/100 t). Finally, the sensitivity to CNO neutrinos can be improved constraining the ratio between the rate of *pp* and *pep* neutrinos at 47.7 ± 0.8 [119], as predicted by the SSM [34, 132]. Exploiting this link, the 9.5% accuracy on the *pp* rate is propagated to *pep* neutrinos acting as an effective constraint. The degeneracy between the CNO events energy distribution and the ^{210}Bi β-decay spectrum limits the accuracy on the CNO neutrino rate to 51.5% (2.53 cpd/100 t) assuming the flux predicted by the HZ SSM. Further improvement requires an independent assessment of the ^{210}Bi background rate that would break the correlation with the CNO reconstructed rate as discussed in Chap. 6.

4.4 FIT MODEL SYSTEMATIC UNCERTAINTIES

The evaluation of the systematic uncertainties of the fit model is performed by fitting ensembles of pseudo-data generated from a family of models reflecting the uncertainties on the response function and fit model parameters. For each

pseudo-data a model is constructed by randomizing the parameters of the response function within their uncertainties determined from calibration data.

In this study, performed together with my colleague Ding XueFeng [93], four different effects were considered: an uncertainty in the modelling of the N_h non-linearity, the presence of a dependence in the detector response along the vertical axis, a wrong modelling of the PS variable and finally a theoretical uncertainty on the exact shape of the ^{210}Bi β-decay spectrum.

Uncertainty in the non-linearity of the N_h response function can be due to small imprecisions in the tuning of the quenching parameters in the MC simulation as well as to other optical properties like the re-emission efficiency of ultraviolet photons. To generate pseudo-data with a different non-linearity, events are sampled from simulated data that were obtained using the nominal values of the response parameters and N_h is converted into the $N_{p.e.}$ observable. A randomized linear distortion is then applied and the distorted $N_{p.e.}$ is converted back to N_h and used to generate the pseudo-dataset.

The possible variation of the scintillator properties along the vertical co-ordinate is considered after observing some small discrepancies in the ^{210}Po peak position between the data and the simulation. Similarly to the non-linearity case, events were sampled from a set of data simulated according to the nominal response function model; the N_h variable is converted to $N_{p.e.}$ and distorted using a randomized linear model before being converted back in N_h and used to build a pseudo-dataset with a randomized response variation along the vertical direction.

Uncertainties in the modelling in the pulse shape parameter used to enhance the discrimination of electron–positron events (Sec. 3.3.4) were also considered as a possible source of systematics. The pseudo-datasets were generated varying the parameters of the pulse shape PDF (Eqs. 3.9, 3.10) according to the uncertainties returned by the fit on the MC sample used to derive them, with the exception of the μ parameter in Eq. (3.10) for which the uncertainty returned by the fit on a selected sample of ^{11}C β^+ events was used instead.

The theoretical uncertainty on the spectral shape of ^{210}Bi is due to the fact that the β decay to ^{210}Po is a first forbidden non-unique transition and therefore it cannot be modelled with standard nuclear physics. The reference spectrum is taken from the measurement performed by Daniel in [133][1], and possible inaccuracies in the measurements are accounted by applying a linear distortion to the ^{210}Bi that is injected in the pseudo-dataset.

Figure 4.9 shows the distribution of the reconstructed rates obtained fitting pseudo-datasets generated randomizing the nominal values of the model and using the standard PDFs to perform the analysis. The reference distribution obtained from an ensemble of pseudo-data where the nominal model parameters were used (and therefore accounting only for the statistical uncertainties) is shown as a dotted histogram for comparison. The overall impact of the

---

1 The collaboration is undergoing intense efforts to perform a new, more accurate measurement of the ^{210}Bi β^- spectrum.

systematics connected to the fit model, estimated from the broadening of the distribution is reported in Tab. 4.1.
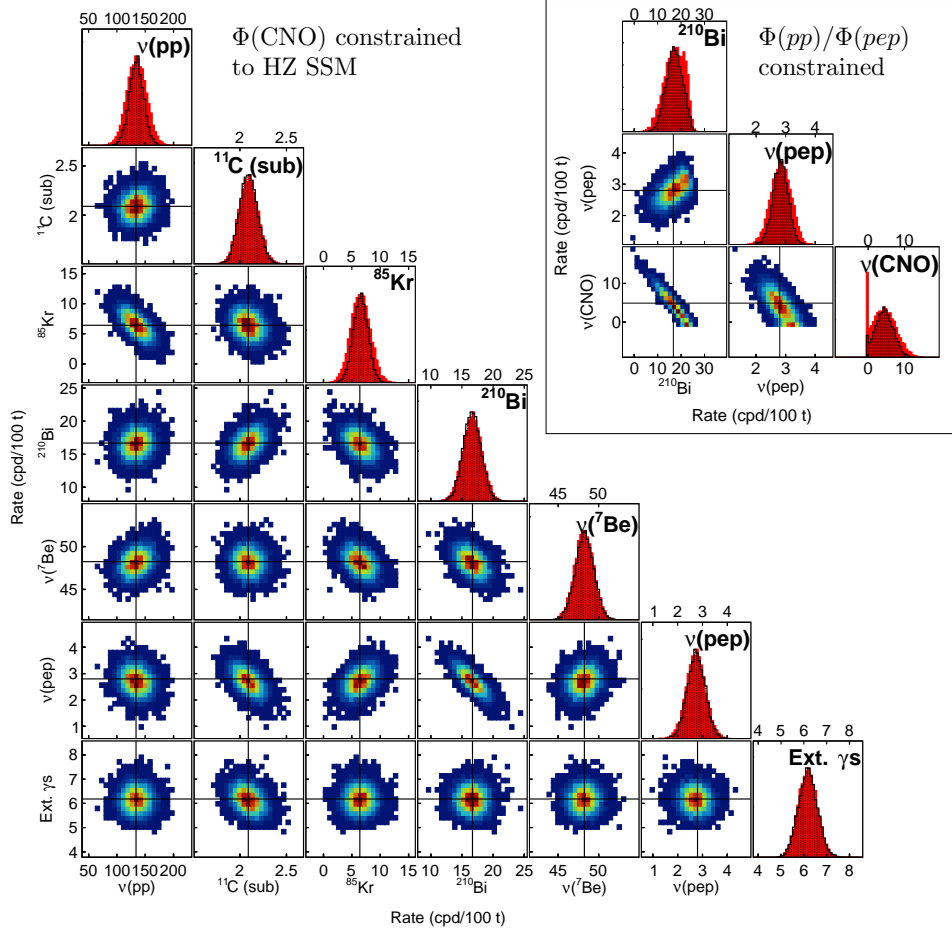
Figure 4.9: Distribution of the reconstructed rates from 10^4 pseudo–experiments generated randomizing within an allowed interval some parameters of the model (see text for explanation). The black dotted histograms show the fit best estimates when the model parameters were fixed at the nominal value. In the main figure the ν(CNO) rate was constrained to the value predicted by the HZ Standard Solar Model, while in the top right box the ratio between the *pep* and *pp* neutrino rates was constrained.

4.5 FIT RANGE DETERMINATION

The fit range has been chosen with the goal of minimizing the total uncertainty of the measurement of *pp* neutrinos. Figure 4.10 shows the trend of the total error on the *pp* neutrino rate as a function of the lower edge of the fit range; the statistical error grows with energy because limiting the spectral information results in larger uncertainties on the *pp* background (leaving ^{14}C and pile-up rate to be determined by the external constraints with the respective accu-

	Injected rate (cpd/100 t)	Stat. uncertainty (cpd/100 t)	Model Systematics (cpd/100 t)
$\nu(\mathrm{pp})$	133.4	$\pm 12.7(\pm 9.5\%)$	$+13.14(+9.8\%)$ $-9.68(-7.2\%)$
$\nu(^7\mathrm{Be})$	48.13	$\pm 1.14(\pm 2.4\%)$	$+0.12(+0.2\%)$ $-0.47(-1.0\%)$
$\nu(\mathrm{pep})$	2.80	$\pm 0.36(\pm 12.9\%)$	$+0.07(+2.4\%)$ $-0.17(-6.0\%)$
$\nu(\mathrm{CNO})^*$	4.91	$\pm 2.53(51.6\%)$	$+2.56(+52.2\%)$ $-2.56(-52.2\%)$
$^{210}\mathrm{Bi}$	16.73	$\pm 1.44(\pm 8.2\%)$	$+0.66(+3.9\%)$ $-0.37(-2.2\%)$
$^{85}\mathrm{Kr}$	6.41	$\pm 1.32(\pm 20.5\%)$	$+1.10(+17.1\%)$ $-1.03(-16.1\%)$
$^{11}\mathrm{C}$ sub	2.10	$\pm 0.10(+4.7\%)$	$+0.04(+2.0\%)$ $-0.03(-1.6\%)$
$^{11}\mathrm{C}$ cmp	71.60	$\pm 0.61(+0.9\%)$	$+0.23(+0.3\%)$ $-0.06(-0.1\%)$
Ext. $^{40}\mathrm{K}$	1.00	$\pm 0.43(+44.4\%)$	$+0.01(+1.0\%)$ $-0.00(-0.0\%)$
Ext. $^{214}\mathrm{Bi}$	1.85	$\pm 0.27(+14.5\%)$	$+0.10(+5.2\%)$ $-0.05(-2.5\%)$
Ext. $^{208}\mathrm{Tl}$	3.35	$\pm 0.14(+4.2\%)$	$+0.03(+0.8\%)$ $-0.04(-1.2\%)$

Table 4.1: Statistical sensitivity and systematic uncertainty for the neutrino and background signal in the analysis. The results are obtained from MC on MC experiments where the CNO rate was constrained assuming the HZ SSM prediction, except for the entry regarding the CNO neutrinos themselves (marked with "*") where the *pep/pp* neutrino flux ratio was constrained instead.

racy) and on the *pp* itself. On the other hand, a systematics uncertainty on the energy scale may move part of the huge number of ^{14}C events above or below the fit range start. Since the ^{14}C and the pile-up rate are constrained by the independent measurement described in Sec. 4.2, the difference in the number of counts is accounted by the *pp* rate and induces a non negligible systematic uncertainty. The lower edge of the fit range minimizing the overall uncertainties on the *pp* neutrino rate has been found to be 106 N_h.
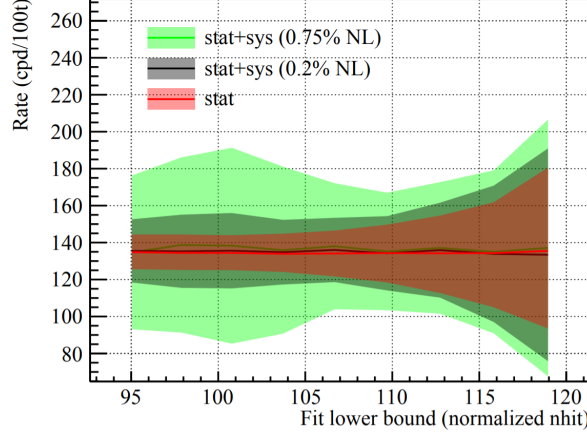
Figure 4.10: Total error on the *pp* rate determination as a function of the fit starting point. The red band shows the statistical uncertainty, while the grey one includes the systematic uncertainties. Finally the green band also includes systematics but assuming a very pessimistic scenario in which larger variations of the response non-linearity are allowed (courtesy of Ding X.F).

4.6   CONCLUSIONS

The sensitivity to low-energy solar neutrinos of the analysis introduced in Chap. 3 relies on specific features in the events distributions. External constraints based on independent measurement of the background as well as on theoretical estimates are crucial to break some correlations in the fit and improve the accuracy of the results.

The sensitivity to the low-energy solar neutrino rates has been studied fitting a large ensemble of pseudo-datasets sampled from simulated data. Similarly, systematic uncertainties have been estimated analysing an ensemble of psuedo-datasets generated from a class of models where the parameters of the response function were randomized within the uncertainties derived by calibration. The overall accuracy on the *pp* and ^{7}Be neutrino rate are 12.6% and 2.5%. The expected uncertainty on the rate of neutrinos from the *pep* reaction when the CNO neutrino rate is constrained according to the predictions of the SSM is 13.9%. On the other hand, when constraining the *pep* rate through the ratio with the *pp* neutrino rate, the statistical uncertainty on the CNO neutrino rate is of the order of 73.3% because of the degeneracy of the events distribution

with the one expected from ^{210}Bi. These estimation of the uncertainties, in particular the one of ^{7}Be neutrinos, shows that a substantial improvement with respect to the previous Borexino results is within reach.

MEASUREMENT OF SOLAR NEUTRINO FLUXES AND IMPACT FOR SOLAR AND NEUTRINO PHYSICS

The analysis described in Chap. 3 and 4 have been applied on a dataset of 1291.51 days of data taking to perform the first simultaneous measurement of pp, ^{7}Be and pep neutrinos [119]. The analysis is summarized in Sec. 5.1 and leads to the most accurate measurement of the low-energy solar neutrinos from the pp chain, along with confirmation of the best upper limit on the interaction rate of CNO neutrinos. These results together with the measurement of ^{8}B neutrinos presented in [134] make of Borexino the first experiment ever capable of providing a comprehensive measurement of all the solar neutrinos emitted throughout the entire pp chain [135] with the only exception of the weak hep neutrino flux. Such a rich set of results has been used to test both fundamental properties of neutrinos and the prediction of the SSMs as discussed in 5.2. On the one hand, assuming the neutrino fluxes predicted by the Standard Solar Model, the electron neutrino survival probability has been studied in a wide energy range spanning from 200 keV (pp neutrinos) to 10 MeV (^{8}B neutrinos), and used to test the MSW-LMA solution of neutrino oscillation (Sec. 5.2.1). On the other hand, assuming the current flavour oscillation framework it is possible to infer from the measured interaction rate the initial neutrino fluxes, that can in turn be used to test the SSM predictions (Sec. 5.2.2). In particular, the measurement of the pp neutrinos is crucial to infer the luminosity of the Sun in the neutrino channel and, combined with the determination of the ^{7}Be flux, it can also be used to compute the relative intensity of the two main terminations of the pp chain. Finally, the measurements of the ^{7}Be and ^{8}B neutrino fluxes have been used to perform an hypothesis test between the HZ and LZ Standard Solar Model.

5.1 MEASUREMENT OF LOW-ENERGY SOLAR ν FLUXES

5.1.1 *Dataset and fit configuration*

The analysis presented in this Chapter is based on data collected from December 14th 2011 to May 21st 2016, equivalent to a total livetime of 1291.5 days (Fig. 5.1).
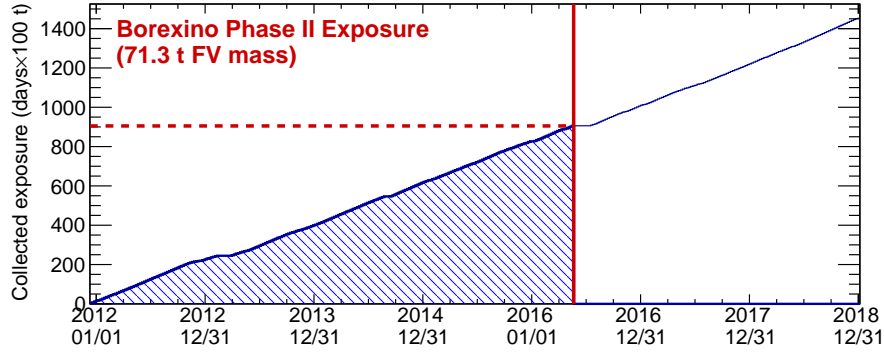
Figure 5.1: Exposure collected by Borexino during Phase II.

Table 5.1: Details of the MV analysis of Borexino Phase II data.

Fiducial volume:	*pep* FV
Exposure:	905.95 days/100 t
Variable:	N_h, r^3, PS-$\mathcal{L}_{PR}$
Range:	107 – 937 N_h
Range r^3:	290 – 937 N_h
Range PS-$\mathcal{L}_{PR}$	409 – 645 N_h (TFC sub only)

The dataset used in this analysis has been selected using the cuts described in Sec. 2.6.1 and the *pep* fiducial volume cut defined in Sec. 2.6.2 (mass 71.3 t), yielding an overall exposure of 905.95 days/100 t, that is $\approx 60\%$ larger than the one collected during the Phase I.

As discussed in Chap. 3 and Chap. 4 the analysis is performed considering the reconstructed number of hits N_h as energy observable, the cubic of the distance from the detector centre r^3 as spatial variable and the pulse shape variable PS-$\mathcal{L}_{PR}$ defined in Sec. 2.6.4.2. The energy range where the fit is performed has been optimized as described in Sec. 4.5 and chosen to be 107–937 N_h. The spatial information is exploited only for reconstructed energy larger than 290 N_h to avoid the non-homogeneous ^{210}Po (Sec. 3.5), while the pulse shape variable is considered only in the ^{11}C region between 409 and 645 N_h. The settings of the multivariate analysis are summarized in Tab. 5.1.

The data are divided into two independent datasets according to the TFC cut (Sec. 2.6.3) and the TFC-subtracted and -tagged datasets are fitted simultaneously considering in the fit model the components listed in Tab. 5.2. The rates of most of the components are *global* parameters, *i.e.* they are kept correlated in both the TFC-subtracted and -tagged datasets, with the exception of ^{210}Po and the ^{11}C, ^{10}C and ^{6}He cosmogenic background (see Sec. 3.5).

The rates of some of the fit components have been constrained in order to break the correlations among the fit parameters as discussed in Sec. 4.2. The rate of ^{14}C and events pile-up has been constrained in all the fit configurations to 40 ± 1 and 8.458 ± 0.081 Bq/100 t respectively on the basis of the independent measurement of the ^{14}C rate (Sec. 3.3.3). The rate of CNO neutrinos has

Component	Constraints	Global	PS type
Neutrino signal			
$\nu(pp)$	Linked‡	Yes	β^-
$\nu(^7\text{Be})$	Free	Yes	β^-
$\nu(pep)$	Linked‡	Yes	β^-
$\nu(\text{CNO})$	Constrained†	Yes	β^-
Background			
^{14}C	Constrained	Yes	β^-
Pileup	Constrained	Yes	β^-
^{85}Kr	Constrained*	Yes	β^-
^{210}Bi	Free	Yes	β^-
^{210}Po	Free	No	$-$
^{11}C	Free	No	β^+
^{10}C	Free	TFC-tag only	β^+
^6He	Free	TFC-tag only	β^-
Ext. ^{40}K	Free	Yes	γ
Ext. ^{214}Bi	Free	Yes	γ
Ext. ^{208}Tl	Free	Yes	γ

Table 5.2: Summary of the fit model components. The constraints applied to the ^{14}C and pile-up rate are described in Sec. 4.3 as well as the one on the CNO neutrino rate and on Φ_{pp}/Φ_{pep} that are applied alternatively when measuring the *pep* and the CNO(‡) rate respectively. An independent constraint on the ^{85}Kr rate (*) is also discussed in Sec. 5.1.2.2 and it has been used to evaluate possible systematics.

been constrained to the values predicted by the HZ and LZ SSMs to break the correlation with the *pep* neutrino rate.

5.1.2   *Systematic uncertainties*

Systematic uncertainties in the analysis can be roughly grouped in two main divisions: the first are the *correlated* systematics that affect in the same way all the fit parameters and are usually related to the exposure, while *uncorrelated* systematics are connected to the fit model and have a different impact on the various parameters of the analysis. The influence of these sources of uncertainties are described in the following sections. In Sec. 5.1.2.1 the uncertainties in the determination of the exposure are discussed, followed by uncertainties related to the fit model, Sec. 5.1.2.2 discuss the impact of an independent measurement of the ^{85}Kr background rate on the fit results while systematics related to the detector response model are presented in Sec. 5.1.2.3.

5.1.2.1   *Exposure uncertainties*

The uncertainties on the exposure arise both from the limited accuracy in the determination of the live time and of the FV mass. The live time before the TFC veto is computed as discussed in [136], where the total uncertainty was estimated to be of the order of $\approx 0.05\%$. The evaluation of the exposure after the TFC veto is less trivial and it is performed using a dedicated simulation. The resulting uncertainty on the dead time is $< 0.5\%$ and impacts only on the determination of the *pep* neutrino rate, but it is practically negligible with respect to the statistical sensitivity and other sources of systematics.

The uncertainties related to the FV mass come mainly from the limited accuracy of the position reconstruction algorithm introduced in Sec. 2.4.2, whose impact in an energy range that includes most of ^{7}Be, *pep* and CNO neutrino events[1] was estimated to be $(-1.1\%, +0.6\%)$ [81, 84, 115, 136] for the *pep* FV. The density of the liquid scintillator is also considered when computing the systematics of the total exposure since variations in the temperature of the detector induce the thermal expansion of the liquid scintillator. Given the current knowledge of the temperature evolution inside the detector, this effect is quantified to impact for $\approx 0.05\%$ of the exposure leading to a negligible contribution in the measurement of the solar neutrino fluxes.

5.1.2.2 ^{85}Kr *independent measurement*

The contamination level of ^{85}Kr in the liquid scintillator in principle can be independently determined exploiting a secondary decay channel into the excited state of ^{85m}Rb. In this case the emission of a β-decay electron with a

---

1 The same error is adopted also at lower energy where most of the signal due to *pp* neutrinos lies. However this is not expected to have a significant impact since this effect is almost negligible when compared to other systematics in the measurement of the *pp* neutrino rate.

Q-value of 173 keV is followed by a 514 keV γ-ray from the de-excitation of ^{85m}Rb, which has a mean lifetime of $\approx 2\,\mu s$. The coincidence in space of the β-decay electron and of the de-excitation γ-ray in a short time window can be used to provide a measurement of the ^{85}Kr rate. Unfortunately, the branching ratio of the decay to the ^{85m}Rb excited state is just 0.43%, and only 3 candidate events were found in the period included in the analysis. In order to consider this information in the analysis, an *ad hoc* penalty term was added to the likelihood of the fit.

5.1.2.3 *Detector response model*

The impact of systematic uncertainties in the fit model has been estimated by performing a thousands of fits on ensembles of pseudo-data generated from a family of models covering the uncertainties on the model parameters. As discussed in Sec. 4.4, three different features of the detector response were randomized at the same time in order to estimate the overall effect including possible interplay between some features. Namely, the randomization of the response model accounted for a wrong modelling of the detector non-linearity for the N_h variable, the presence of a non modelled spatial dependence of the detector response along the vertical axis and uncertainties in the PS-$\mathcal{L}_{PR}$ variable model. The theoretical uncertainty on the ^{210}Bi β-decay spectral shape has been treated in the same way.

5.1.3 *Results*

The data and the result of the fit are shown in Fig. 5.2 and Fig. 5.3. Figure 5.2 shows the projection to energy variable of the TFC-subtracted and -tagged datasets along with the fit result, while Fig. 5.3 shows the projection of the TFC-subtracted dataset to the radial and pulse shape variable. In the lower pad the ratio between the data and the result of the fit indicates that the model describes properly the data.

The goodness of fit however cannot be estimated by a simple χ^2 test since in the multidimensional analysis many bins might be too low populated to allow Gaussian approximation to be valid. The minimum of the negative log-likelihood obtained by the fit (see Sec. 3.2.1) has been chosen as a test statistics and its distribution is derived fitting an ensemble of pseudo-data generated using the same model used for the fit. Figure 5.4 shows the value of the test statistics obtained from the data realization along with the expected distribution. The p-value is ≈ 0.5, thus indicating a good compatibility between the data and the model.

The fit has been performed constraining the CNO rate to the values predicted by both the HZ and LZ SSM. Confidence intervals for the fit parameters have been built using a profile likelihood under Wilk's approximation. Figure 5.5 and Fig. 5.6 reports the profile likelihood for the low energy neutrino components: the rate of CNO neutrinos is determined by the external
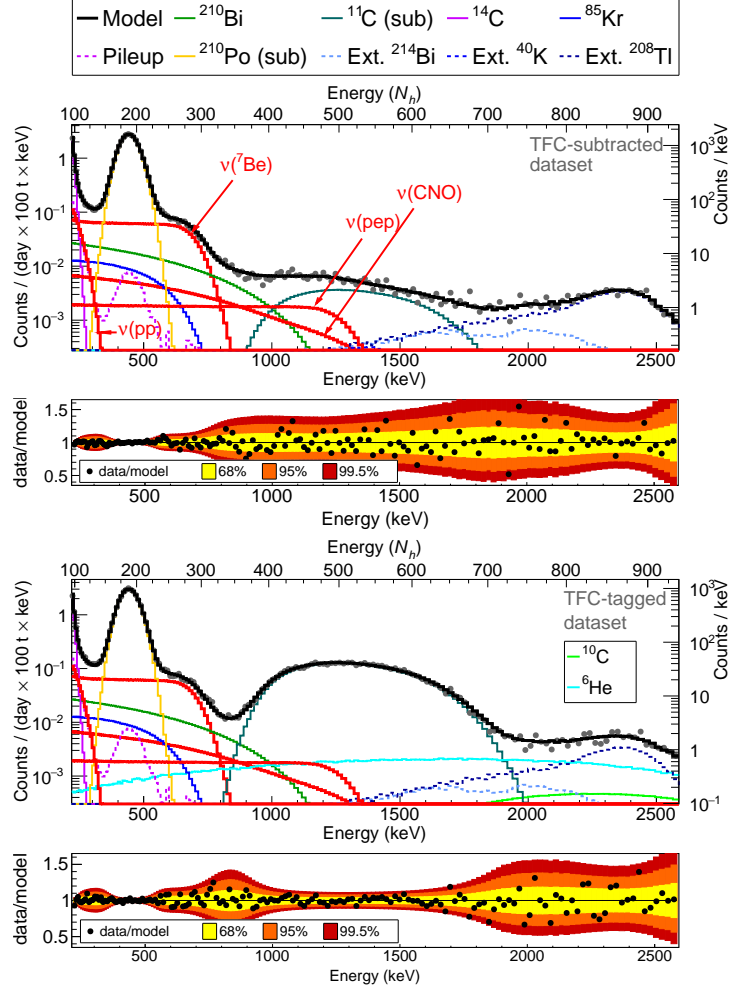
Figure 5.2: Result of the multivariate fit shown in the energy projection. The top and bottom panel represent the TFC-subtracted and -tagged datasets respectively.

constraint and while the reconstructed rates of *pp* and ^{7}Be neutrinos are not affected by the different CNO rate predicted by the HZ/LZ SSM, the correlation with the CNO rate makes the best fit value of the *pep* neutrino rate move. All the likelihood profiles are parabolic, as it is expected from the fact that the distributions of the maximum likelihood estimators shown in Fig. 4.8 are described by Gauss distributions.

pp neutrinos

Neutrinos from proton–proton fusion mark the start of the chain of reactions responsible for most of the Sun power and are the dominant contribution to the solar neutrino spectrum, but their very low energy makes the measurement very delicate for the presence of the high ^{14}C background. The measured rate is $133.2 \pm 12.8(\mathrm{stat})\,\mathrm{cpd}/100\,\mathrm{t}$, in good agreement with the previous Borexino
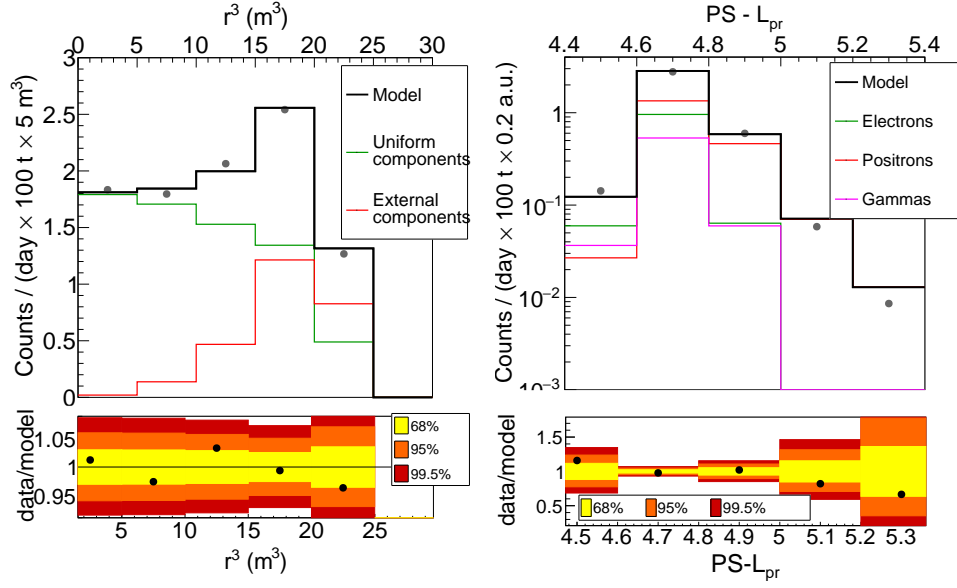
Figure 5.3: Result of the multivariate fit of the TFC-subtracted dataset shown in the radial (*left*) and PS (*right*) dimension.

result [86]. In this analysis the most relevant systematic comes from uncertainties in the detector response, specifically in the modelling of the detector non-linearity which causes the "migration" of a large number of ^{14}C events inside (or outside) the energy rage with the highest sensitivity to the *pp* neutrino signal.

^{7}Be *neutrinos*

The rate of neutrinos produced in the electron capture on ^{7}Be has been constrained to $48.1 \pm 1.1\,(\text{stat})\,\text{cpd}/100\,\text{t}$. The impact of systematics is found to
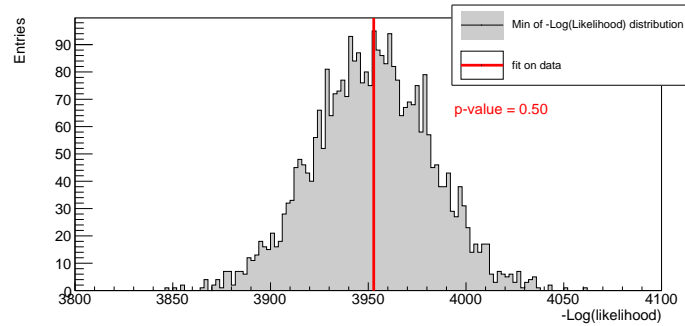
Figure 5.4: Distribution of the minimized negative log-likelihood obtained from thousands of fits performed on pseudo-data compared to the value obtained from the data realization. A *p*-value of 0.5 has been obtained, showing no signal of incompatibility between the data and the model.
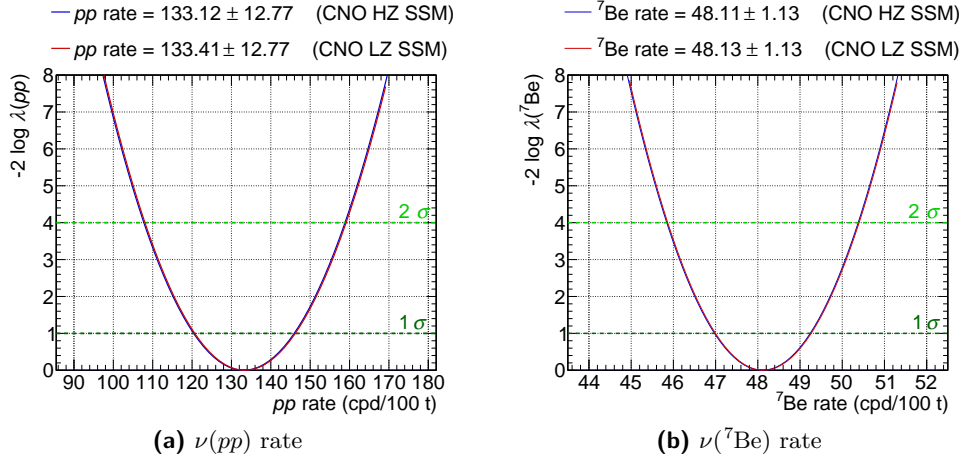
Figure 5.5: Profile likelihood of the *pp* and ^{7}Be neutrino rates obtained constraining the CNO neutrino rate to the values expected from the HZ (red) and LZ (blue) Standard Solar Model. For these components the impact of a different constraints on the CNO rate is negligible.

be at the level of 1.5% and is dominated by the uncertainties in the detector response modelling and in the definition of the fiducial volume. The overall accuracy of measured interaction rate is as low as 2.7%, the best result achieved to date. This marks a substantial improvement with respect to the past measurement of the ^{7}Be neutrino fluxes performed by Borexino [80–82] and KamLAND [87] and opens the way to precision neutrino physics at energies lower than 1 MeV.

pep and CNO *neutrinos*

As discussed in Sec. 4.3, the strong correlation among the rate of ^{210}Bi and *pep* and CNO neutrinos reduces significantly the sensitivity of the analysis. This situation can be solved in principle if an independent measurement of the ^{210}Bi is available and this possibility is indeed extensively discussed in Chap. 6 and Chap. 7, but in the absence of such information it has been decided to exploit some inputs from the SSM to improve the sensitivity to these components.

To measure the *pep* neutrino rate, the HZ and LZ SSM has been used to constrain the CNO neutrino rate. The large difference in the prediction of the SSM affects significantly the background level for *pep* neutrinos, whose rate results 2.38 ± 0.36 cpd/100 t for the CNO neutrino rate predicted by the HZ SSM or 2.61 ± 0.35 cpd/100 t for the LZ SSM (Fig. 5.6). The systematic budget for *pep* neutrinos is dominated by the uncertainty in the modelling of the PS-$\mathcal{L}_{PR}$ variable, but the overall uncertainty still makes it possible to rule out the hypothesis of absence of *pep* neutrino at more than 5σ C.L. with the characteristic shoulder clearly visible in Fig. 5.7a.

(a) ν(CNO) rate (constrained)   **(b)** ν(pep) rate
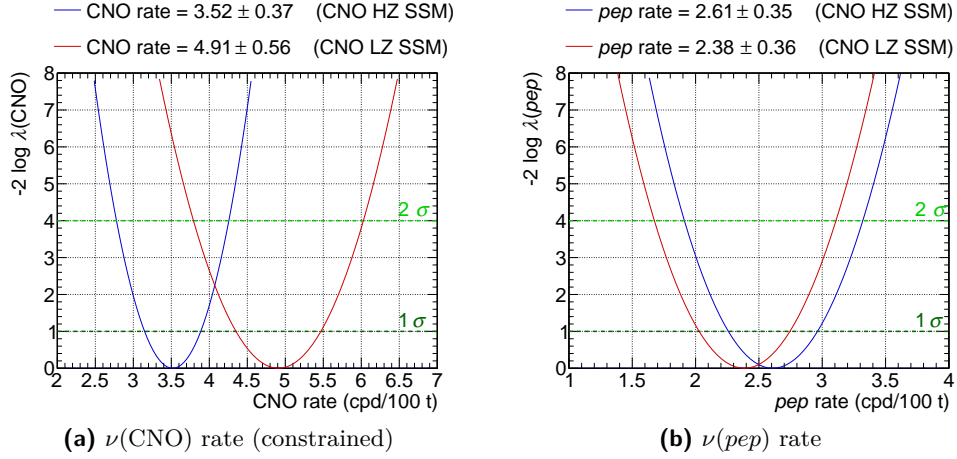
Figure 5.6: Profile likelihood of the *pep* and CNO neutrino rates obtained constraining the CNO neutrino rate to the values expected from the HZ (red) and LZ (blue) Standard Solar Model.

A dedicated fit has been performed as a second step of the analysis to constrain the rate of CNO neutrinos. In this fit the CNO neutrino rate is left free and the *pep* neutrino rate is constrained by fixing the ratio with the *pp* neutrino rate to the value of 47.7 ± 0.8 that is predicted by the SSM under solid nuclear physics arguments as discussed in Sec. 6.3.1. The residual correlation with the ^{210}Bi prevents a measurement of the CNO neutrino rate and an upper limit at $8.1 \, \text{cpd}/100 \, \text{t}$ (95% C.L., including both statistical and systematic uncertainties) has been obtained. The profile likelihood of the CNO neutrino rate is shown in Fig. 5.7b. This result is consistent with the limit reported by Borexino during Phase I [84] using a stronger assumption on the *pep* neutrino rate that was fixed to the value predicted by the SSM.

5.1.4 *Additional consistency checks*

To check the solidity of the results, the analysis has been performed using different datasets. The fit has been repeated on datasets built using a different implementation of the TFC veto and using a different binning of the energy observable, always obtaining fully consistent results.

The stability of the fit over the choice of the lower edge of the fit range has been tested by fitting the data varying the fit starting point around the value of $107 \, N_h$ that was chosen to minimize the overall uncertainty on the measured *pp* rate (see Sec. 4.5. Indeed, while all the neutrino components of the Borexino spectrum are not affected by the energy range considered in the fit, the very low energy *pp* neutrino component can be influenced by a change in the lower edge of the fit range in the analysis. With starting values in a neighbourhood of the chose point the analysis is not very sensitive to the spectral shape of ^{14}C and of

(a)   **(b)**
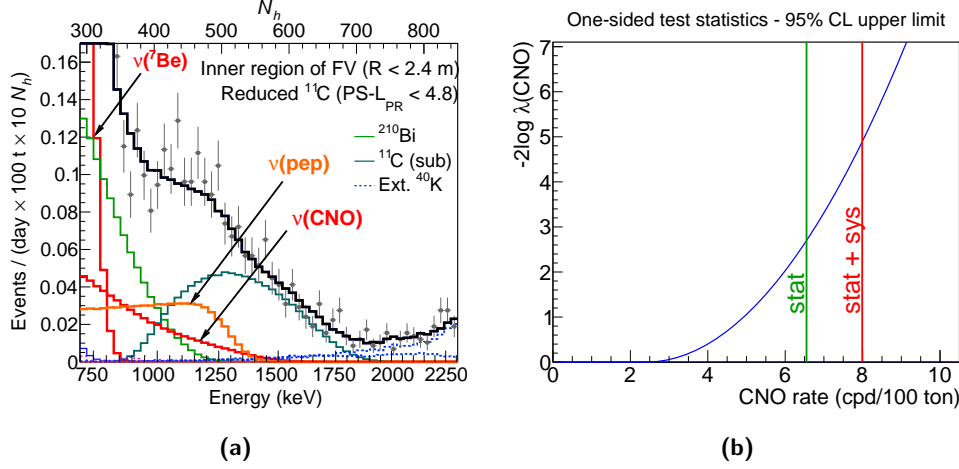
Figure 5.7: (a) Visible energy spectrum of events of the TFC-subtracted dataset selected in the innermost region of the FV ($r < 2.4$ m) and with PS-$\mathcal{L}_{PR} < 4.8$ to further reduce the ^{11}C background. The fit is performed constraining the CNO neutrino rate according to the HZ SSM prediction, and the characteristic shoulder of the *pep* neutrino signal is clearly visible. **(b)** Profile likelihood of the CNO neutrino rate obtained constraining the ratio between the *pp* and *pep* neutrino rate in the fit.

the pile-up. Their rate is hence determined by the independent measurements used to constrain their values without being influenced by the actual data. As Fig. 5.8 shows, the *pp* neutrino rate returned by the fit is fairly stable. However at energies lower than 98 N_h, the information of the spectral shape exploited by the fit starts being in tension with the independent constraints. The relatively high energy starting point however keeps the measurement of the *pp* neutrino rate safe from possible systematics in the modelling of the detector response at very low energy.

5.1.5 *Combination of Monte Carlo and Analytical fit results*

The same analysis has been performed within the collaboration using the previous fit methods (both Monte Carlo and Analytical) and different energy estimators to cross-check the results.

Results from different analysis methods have been combined following the procedure described in [119] to obtain the final results presented in [119, 135] and reported in Tab. 5.3, marking a significant improvement with respect to previous Borexino findings. Table 5.4 shows the measured rate of the background components.

The results obtained with `bx-stats` in this analysis and those provided by the previous fit all agree with each other, but different methods have slightly different performance. The `bx-stats` analysis provides a more accurate measurement of the ^{7}Be and *pep* neutrinos, while the Analytical fit is better at
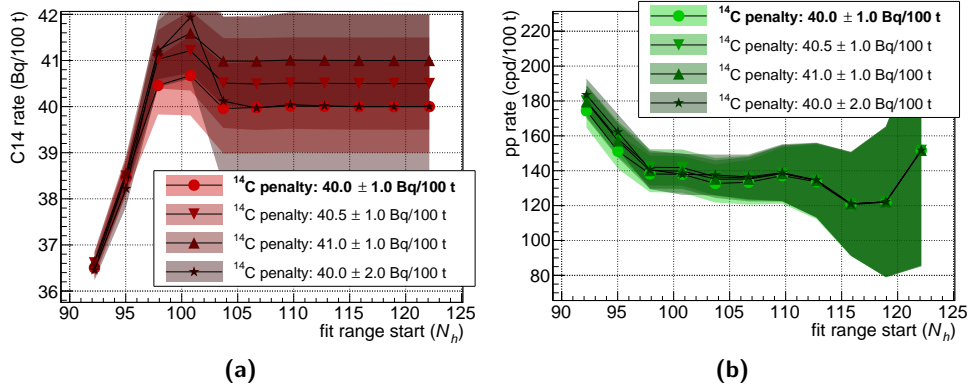
Figure 5.8: Reconstructed rate of ^{14}C **(a)** and pp neutrinos **(b)** as a function of the lower edge of the fit range for different configurations of the constraint on the ^{14}C rate (the one used on data is reported in bold). The shaded area represents the 1σ C.I. returned by the fit.

describing the low energy part of the energy spectrum and measures the rate of pp neutrino more precisely.

The small differences between the Monte Carlo and Analytical fit results, along with the difference between the results obtained using different energy estimators, are accounted in the total systematic uncertainty as shown in Tab. 5.5. The evaluation of systematic uncertainties of the previous fit methods introduces other voices in the overall systematic budget. The pile-up modelling item in Tab. 5.5 refers to different treatment of random coincidences: differently from the Monte Carlo analyses where the contribution of event pileup is treated as a independent component in the fit model, in the Analytical fit events random coincidences are taken into account either by convolving the spectral shape of each component with the random trigger spectrum or by including the "synthetic pile-up" in the fit model as an additional component (Sec. 3.3.3). These three approaches lead to slightly different results on the pp neutrino rate and the difference is quoted as systematic error.

The results reported in Tab. 5.3 are to be considered as the final results presented by the Collaboration, therefore from here on they will be taken as a reference to test the predictions of the SSM and the MSW-LMA solution of neutrino oscillation.

5.2 IMPACT ON NEUTRINO AND SOLAR PHYSICS

The results presented in the previous section are the most accurate measurements of the low energy solar neutrino interaction rate up to date, and combined with an independent measurement of the ^{8}B neutrinos [134] make Borexino the only experiment capable to perform a comprehensive measurement of

Solar ν	**Borexino results** Rate (cpd/100 t)	**HZ SSM** Rate (cpd/100 t)	**LZ SSM** Rate (cpd/100 t)
pp	$134 \pm 10^{+6}_{-10}$	131.0 ± 2.4	132.1 ± 2.3
^{7}Be	$48.3 \pm 1.1^{+0.4}_{-0.7}$	47.8 ± 2.9	43.7 ± 2.6
pep (HZ)	$2.43 \pm 0.36^{+0.15}_{-0.22}$	2.74 ± 0.05	2.78 ± 0.05
pep (LZ)	$2.65 \pm 0.36^{+0.15}_{-0.24}$		
CNO	< 8.1 (95% C.L)	4.91 ± 0.56	3.52 ± 0.37
	Flux cm^{-2}s^{-1}	Flux cm^{-2}s^{-1}	Flux cm^{-2}s^{-1}
pp	$(6.1 \pm 0.5^{+0.3}_{-0.5}) \times 10^{10}$	$5.98(1 \pm 0.006) \times 10^{10}$	$6.03(1 \pm 0.005) \times 10^{10}$
^{7}Be	$(4.99 \pm 0.13^{+0.07}_{-0.10}) \times 10^{9}$	$4.93(1 \pm 0.06) \times 10^{9}$	$4.50(1 \pm 0.06) \times 10^{9}$
pep (HZ)	$(1.27 \pm 0.19^{+0.08}_{-0.12}) \times 10^{8}$	$1.44(1 \pm 0.009) \times 10^{8}$	$1.46(1 \pm 0.009) \times 10^{8}$
pep (LZ)	$(1.39 \pm 0.19^{+0.08}_{-0.13}) \times 10^{8}$		
CNO	$< 7.9 \times 10^{8}$ (95% C.L.)	$4.88(1 \pm 0.11) \times 10^{8}$	$3.51(1 \pm 0.10) \times 10^{8}$

Table 5.3: Results of Borexino-Phase II on the *pp*, ^{7}Be, *pep* and CNO solar neutrinos interaction rate. The first term of the associated uncertainty represents the statistical error obtained with a likelihood profile under the Wilk's approximation, while the second indicates the systematic uncertainty that is detailed in Tab. 5.5. The *pep* neutrino rate depends on the different SSM used to constrain the CNO neutrino flux, and therefore two results are quoted. The fluxes reported in the lower part of the table are computed assuming the MSW-LMA oscillation parameters [137].

Table 5.4: Measured background rate in Borexino-Phase II. The quoted uncertainty includes both the statistical and systematic errors.

Component	Rate (cpd/100 t)
^{14}C	40.0 ± 2.0
^{85}Kr	6.8 ± 1.8
^{210}Bi	17.5 ± 1.9
^{11}C	26.8 ± 0.2
^{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

Source of uncertainty	$\nu(pp)$		$\nu(^7\mathrm{Be})$		$\nu(pep)$	
	$-\%$	$+\%$	$-\%$	$+\%$	$-\%$	$+\%$
Fit method (analytical/MC)	-1.2	$+1.2$	-0.2	$+0.2$	-4.0	$+4.0$
Choice of energy estimator	-2.5	$+2.5$	-0.1	$+0.1$	-2.4	$+2.4$
Pile-up modelling	-2.5	$+0.5$	-0.0	$+0.0$	-0.0	$+0.0$
Fit range and binning	-3.0	$+3.0$	-0.1	$+0.1$	-0.1	$+0.1$
Fit models (see text)	-4.5	$+0.5$	-1.0	$+0.2$	-6.8	$+2.8$
Inclusion of ^{85}Kr constraint	-2.2	$+2.2$	-0.0	$+0.4$	-3.2	$+0.0$
Fiducial Volume	-1.1	$+0.6$	-1.1	$+0.6$	-1.1	$+0.6$
Live time	-0.05	$+0.05$	-0.05	$+0.05$	-0.05	$+0.05$
Scintillator density	-0.05	$+0.05$	-0.05	$+0.05$	-0.05	$+0.05$
Total Systematics (%)	-7.1	$+4.7$	-1.5	$+0.8$	-9.0	$+5.6$

Table 5.5: Systematics budget for the solar neutrino components measured in Borexino-Phase II. The terms due to the different pile-up modelling, fit range and binning contributes to the systematics budget of the previous analysis methods that contributed to the results presented in Tab. 5.3 [119, 135].

the solar neutrinos emitted throughout the entire *pp* chain, with the only exception of hep neutrinos.

Borexino measurements of the solar neutrino interaction rate can be used to test either neutrino or solar phenomenology. Assuming the neutrino fluxes predicted by the SSM one can infer from the measured rate (mostly due to electron-neutrinos given the larger cross section, see Sec. 2.1) the ν_e survival probability, and test the agreement with the MSW-LMA solution of neutrino oscillation. Assuming the neutrino oscillation framework the measured interaction rates can be converted in the integral flux of solar neutrinos and be used to test the predictions of the SSM. In particular, a measurement of the solar neutrino fluxes can be crucial to solve the controversy about the Sun metallicity that led to the introduction of the competing HZ and LZ SSM.

5.2.1 *Study of ν_e survival probability*

As discussed in Sec. 1.1.2 solar neutrinos were crucial for the discovery of flavour oscillations in the neutrino sector. After the measurement of ^{8}B neutrinos performed by SNO [76, 77] exploiting both charged and neutral-current processes, the neutrino oscillation parameters have been measured with great accuracy by a number of experiments [8].

The case of solar neutrinos is however complicated by the presence of a dense medium in which neutrinos propagate and that alters the Hamiltonian with the addition of an interaction potential in a energy-dependent way (Sec. 1.1.3).

Borexino is the only experiment able to measure solar neutrinos in the $200\,\mathrm{keV}$ ($\nu(pp)$) – $16\,\mathrm{MeV}$ ($\nu(^8\mathrm{B})$) energy range, and its results can be used to reconstruct the ν_e survival probability curve for solar neutrinos in the largest energy range ever covered by a single experiment. Since only the ^{7}Be and

Solar ν	P_{ee} BX HZ SSM	P_{ee} BX LZ SSM	P_{ee} MSW-LMA	P_{ee} Vacuum-LMA
pp	0.57 ± 0.09	0.57 ± 0.09	0.5457 ± 0.0087	
^{7}Be	0.53 ± 0.05	0.61 ± 0.05	0.5269 ± 0.0079	0.5378 ± 0.0268
pep	0.43 ± 0.11	0.42 ± 0.11	0.5165 ± 0.0073	
^{8}B	0.37 ± 0.08	0.48 ± 0.08	0.3390 ± 0.0082	

Table 5.6: Values of the electron neutrino survival probabilities. Columns 2 and 3 reports the value of P_{ee} inferred by the latest Borexino results [134, 135] assuming the solar fluxes predicted by the HZ and LZ SSM. The values expected from the MSW-LMA paradigm are reported in column 4 while the expectation from a Vacuum–LMA scenario (energy independent) is shown in column 5.

pep neutrinos are mono-energetic, for pp and ^{8}B neutrinos the results have been computed according to the average energy of neutrinos that produce a deposited energy in the range used for the analysis. The electron neutrino survival probability (P_{ee}) is computed from the measured interaction rate in Borexino (R^{BX}) assuming the neutrino fluxes as predicted by the HZ and LZ SSM (Φ^{SSM}) through the formula [86]

$$P_{ee} = \frac{R^{\mathrm{BX}} - \Phi^{\mathrm{SSM}} n_e \sigma_\mu}{\Phi^{\mathrm{SSM}} n_e (\sigma_e - \sigma_\mu)} \tag{5.1}$$

where $n_e = (3.307 \pm 0.003) \times 10^{31}$ is the number density of target electrons in 100 t of Borexino liquid scintillator while σ_e and σ_μ indicate the ν_e and ν_μ, τ elastic scattering cross sections (Eq. 2.2) integrated over the neutrino energy spectrum.

The values of P_{ee} obtained by Borexino are reported in Tab. 5.6 including in the uncertainties both the accuracy of the measured rate and the precision of the SSM predictions. The results of Borexino are show in Fig. 5.9 along with the prediction of the MSW-LMA model (Eq. 1.38) and the expectation from LMA oscillation in vacuum. The P_{ee} expected from the LMA-MSW oscillation model is computed using the interaction potentials reported in [138] and the oscillation parameters presented in [132], while the one resulting from neutrino oscillation in vacuum is obtained assuming the oscillation parameters measured with reactor antineutrinos [139, 140].

The expected survival probability for low-energy solar neutrinos (pp, ^{7}Be, pep) is not affected by the matter effect in the Sun and can be described using assuming the production and propagation of solar neutrino as in vacuum. In the ^{8}B energy range the interaction potential of electrons in the Sun core suppresses the survival probability, therefore the estimate of the ^{8}B survival probability is crucial to establish the effect of matter in neutrino oscillations. To do so, a frequentist hypothesis test has been performed considering as two

(a) HZ SSM fluxes   **(b)** LZ SSM fluxes
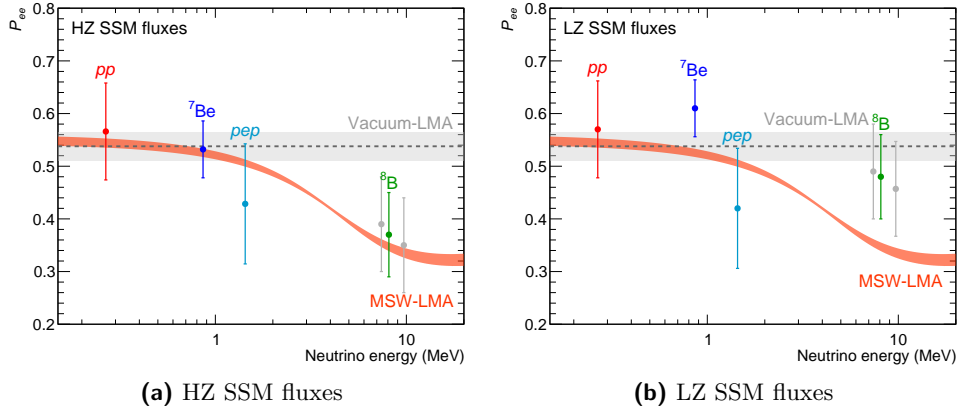
Figure 5.9: Electron neutrino survival probability (P_{ee}) as a function of the neutrino energy. The survival probabilities for the solar-ν are obtained combining the interaction rates measured by Borexino with the fluxes predicted from the **(a)** HZ and **(b)** LZ SSM. The expected P_{ee} computed including the MSW effect is shown as an orange band representing the $\pm 1\sigma$ C.I., while the grey band represents the expected behaviour of neutrinos oscillating in vacuum. The two measurement of ^{8}B neutrinos in grey refers to two different energy ranges used in the analysis, and the final number (in green) shows the combination of these two results [134].

alternative hypothesis the MSW-LMA model vs. the Vacuum–LMA model. The test is based on chi-square difference test statistics [141]

$$t = \chi^2(\text{MSW}) - \chi^2(\text{vacuum}) \tag{5.2}$$

where the χ^2 is computed as

$$\chi^2 = \sum_{i=pp,\,^7\text{Be},\dots} \frac{(x_i - \mu_i)^2}{(\sigma_{x_i}^2 + \sigma_{\mu_i}^2)^2} \tag{5.3}$$

where x_i indicates the value of the ν_e survival probability obtained by Borexino and μ_i the expectation by the model under consideration for the neutrino component i. The uncertainty of Borexino results and of the model predictions are represented by σ_{x_i} and σ_{μ_i} respectively. The distribution of the test statistics t has been built with a MC method: thousands of sets of P_{ee} measurements were generated according to one of the two possible scenarios taking into account both experimental and theoretical uncertainties, computing the value of t for each set. The probability distribution function of t obtained for the hypotheses of MSW-LMA $f(t|\text{MSW})$ and Vacuum–LMA $f(t|\text{vacuum})$ are shown in Fig. 5.10.

The values of the survival probability x_i obtained by Borexino depends on the flux of solar neutrinos assumed from the SSMs. In particular the flux of ^{8}B neutrinos predicted by the HZ and LZ SSM differs by 18% and causes a large difference in the reconstructed values of the ^{8}B survival probability.

Considering the HZ SSM, which is favoured by Helioseismology, the value of t obtained from the Borexino results is $t_{\mathrm{BX}} = -4.16$ and is indicated in Fig. 5.10a with a dashed black line. The value of t_{BX} is perfectly compatible with the distribution of the test statistics derived assuming the MSW-LMA as true, and lies on the tail of the distribution of t obtained according to the Vacuum–LMA hypothesis: the probability of observing a result more extreme of the data is given by the p-value

$$p = \int_{-\infty}^{t_{\mathrm{BX}}} f(t|\mathrm{vacuum})\mathrm{d}t = 0.018 \qquad (5.4)$$

Such a small p-value suggests that the Vacuum–LMA model is not very plausible. It is custom to convert the p-value into a number of Gaussian standard deviations to express the significance of the result. In this thesis the convention of [142] is adopted, where one-sided Gaussian limits are used. In this context, a p-value of 0.018 corresponds to rejecting the Vacuum–LMA hypothesis with a significance of 2.1σ.

(a) HZ SSM fluxes   **(b)** LZ SSM fluxes
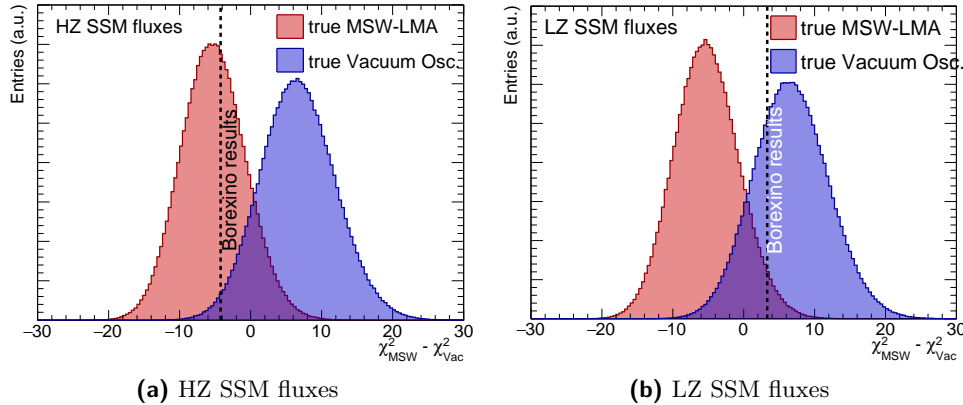
Figure 5.10: Distribution of the test statistics t used to perform a frequentist hypothesis test between the MSW-LMA and the Vacuum–LMA oscillation scenarios. The value of t obtained from Borexino results assuming the solar neutrino fluxes predicted by the **(a)** HZ SSM and **(b)** LZ SSM is shown as a dotted line.

Considering the neutrino fluxes predicted by the LZ SSM, the reduced ^{8}B neutrino flux makes the corresponding P_{ee} closer to the value expected from the Vacuum–LMA scenario than from the MSW-LMA model. As a consequence, repeating the hypothesis test with the value of the survival probability obtained assuming the LZ SSM fluxes, a value $t_{\mathrm{BX}} = 3.32$ is obtained (Fig. 5.10b), favouring the Vacuum–LMA scenario and excluding the MSW-LMA with a 1.8σ significance (0.034 p-value).

5.2.2   *Test of the Standard Solar Model*

The results provided by Borexino have a significant impact in the field of solar physics as they can provide an important test of the current SSM predictions. Borexino is not equally sensitive to all the neutrino flavours since the elastic scattering cross-section of ν_e with electrons is $5-6$ times larger than the one of $\nu_{\mu,\tau}$, hence it is necessary to assume the neutrino oscillation framework to extract the total flux of neutrinos to be compared with the expectations of the SSM.

These results are used to probe three different aspect of the SSM: first, the measurement of the *pp*-chain solar neutrinos can be converted into a measurement of the Sun luminosity in the neutrino channel, which can be compared with the value measured using photons; second, the measurement of the ^{7}Be interaction rate provides information on the ratio between the two main termination branches of the *pp* chain; and last, the neutrino fluxes inferred from the Borexino measurements can be compared to the different predictions of the HZ and LZ SSM and potentially disfavour one of the two models.

5.2.2.1   *Neutrino Luminosity of the Sun*

Neutrinos produced in the core of the Sun are direct messengers of our star. Indeed, every neutrino emitted by the Sun marks a specific reaction that took place in its core. Since the amount of energy released in every single nuclear reaction is well known from nuclear physics, it is possible to compute the total power generated by nuclear processes in the Sun using the neutrino fluxes inferred from the Borexino measurements.

More explicitly, the solar luminosity measured in the neutrino channel $L_\odot^{(\nu)}$ can be expressed as

$$L_\odot^{(\nu)} = 4\pi(1\,\mathrm{A\,U})^2 \times \sum_i \alpha_i \Phi_i \tag{5.5}$$

where α_i indicates the thermal energy released by the reaction associated to the neutrino flux labelled with i. Using the values of α_i computed in [143] under mild assumptions[2] and the results presented in the previous section, the luminosity is equal to $(3.89^{+0.35}_{-0.42}) \times 10^{33}$erg/s, where the dominant contribution to the Sun power comes from the *pp* neutrinos[3]. This result is in good agreement with the value measured through photons that is $(3.846 \pm 0.015) \times 10^{33}$erg/s [145]. This confirms the nuclear origin of the Sun power with the best precision obtained so far by a single experiment [135]. Unlike neutrinos that can exit the Sun in just a few minutes after being produced, radiation takes 10^4-

---

2 In particular, to compute the thermal energy associated to each neutrino flux it is crucial to assume that ^{2}H and ^{3}He are in local kinetic equilibrium in the Sun, meaning that their creation rate is equal to their destruction rate. As it will be detailed in Sec. 6.3.1, this assumption is well motivated in the framework of the current knowledge of the Sun.

3 Using a recent re-evaluation of the α_i coefficients presented in [144] does not affect the estimate of $L_\odot^{(\nu)}$ since the computation of the energy released in the dominant *pp* fusion process is unaltered.

10^5 years to reach the photosphere from the energy production zone, thus the correspondence between the luminosity measured in the neutrino and photon channel also proves stability of the Sun energy production over such a period of time.

5.2.2.2 *Termination of the pp chain*

The measurement of both the *pp* and ^{7}Be neutrino rate constrain the relative intensity between the first and the second termination of the *pp* chain. After the formation of deuteron via *p-p* fusion, ^{3}He is produced in the Sun core and can react with another ^{3}He nucleus terminating the chain without emitting any more neutrinos or, less likely, can interact with a ^{4}He nucleus producing a ^{7}Be neutrino.

Under the same assumptions needed to compute the thermal energy α_i associated to each neutrino flux that was introduced in the previous section, it is possible to define a fictitious neutrino flux $\Phi_{^3\text{He}-^3\text{He}}$ related to the ^{3}He-^{3}He reaction of the *pp* chain that is expressed as [143]

$$\Phi_{^3\text{He}-^3\text{He}} = \tfrac{1}{2}\left(\Phi_{pp} + \Phi_{pep} - \Phi_{^7\text{Be}} - \Phi_{^8\text{B}} - \Phi_{\text{hep}}\right) \tag{5.6}$$

Neglecting the very low *pep*, ^{8}B and hep fluxes, the ratio between the ^{3}He-^{3}He and the ^{3}He-^{4}He reaction rates can be written as [146]

$$R_{\text{I/II}} = \frac{2\Phi_{^7\text{Be}}}{\Phi_{pp} - \Phi_{^7\text{Be}}} \tag{5.7}$$

The measurement of the *pp* and ^{7}Be interaction rate presented in the previous section results in $R_{\text{I/II}}^{(\text{BX})} = 0.178^{+0.027}_{-0.023}$, in good agreement with the values predicted by the HZ and LZ SSM that are $R_{\text{I/II}}^{(\text{HZ})} = 0.180 \pm 0.011$ and $R_{\text{I/II}}^{(\text{LZ})} = 0.161 \pm 0.010$.

5.2.2.3 *High- or Low-Metallicity?*

The measurement of the solar neutrino fluxes can be used to test the compatibility of the predictions of the HZ/LZ SSM, and possibly solve the solar metallicity puzzle.

Indeed, as described in Sec. 1.2.3, the different abundance of the heavy elements affects the opacity of the plasma which determines the temperature profile and, as a consequence, the reaction rate that generates solar neutrinos. The two reactions most sensitive to temperature variation in the *pp* chain are those leading to ^{7}Be and ^{8}B neutrinos. As shown in Tab. 5.3 the flux of ^{7}Be and ^{8}B can vary of about 9% and 18% between the HZ and LZ model. Figure 5.11a shows the parameters space allowed by the SSM obtained including the correlation factor provided by the authors of [34][4]. The theoretical uncer-

---

4 Available at `http://www.ice.csic.es/personal/aldos/Solar_Data.html`.

tainties on the ^{8}B and ^{7}Be flux predictions are dominated by the uncertainty on the opacity of the Sun and by the astrophysical S-factors of the ^{3}He + ^{4}He and p + ^{7}Be reactions. The results of Borexino on the ^{7}Be and ^{8}B are also shown. It is worth to stress that the SSM uncertainties are twice as large as the accuracy of the measured ^{7}Be flux and compatible with the precision of the ^{8}B flux respectively.

It is possible to estimate the compatibility of the data with the two models by performing a χ^2 test, with

$$\chi^2 = \mathbf{\Delta}^T \left(\mathsf{V}_{\mathrm{SSM}} + \mathsf{V}_{\mathrm{BX}}\right)^{-1} \mathbf{\Delta} \tag{5.8}$$

where $\mathbf{\Delta}$ represents the difference between the measured neutrino fluxes and the SSM predictions

$$\mathbf{\Delta} = \left(\Phi_{^8\mathrm{B}}^{\mathrm{SSM}} - \Phi_{^8\mathrm{B}}^{\mathrm{BX}}, \Phi_{^7\mathrm{Be}}^{\mathrm{SSM}} - \Phi_{^7\mathrm{Be}}^{\mathrm{BX}}\right) \tag{5.9}$$

while V_{BX} and $\mathsf{V}_{\mathrm{SSM}}$ are the covariance matrices of the Borexino results and of the SSM in question that accounts respectively for the experimental and theoretical uncertainties. In their explicit form they reads

$$\mathsf{V}_{\mathrm{BX}} = \begin{pmatrix} (\sigma_{^8\mathrm{B}}^{\mathrm{BX}})^2 & 0 \\ & \\ 0 & (\sigma_{^7\mathrm{Be}}^{\mathrm{BX}})^2 \end{pmatrix} \quad \text{and}$$

$$\mathsf{V}_{\mathrm{SSM}} = \begin{pmatrix} \sigma_{^8\mathrm{B}}^{\mathrm{SSM}}\sigma_{^8\mathrm{B}}^{\mathrm{SSM}} & \rho^{\mathrm{SSM}}\sigma_{^8\mathrm{B}}^{\mathrm{SSM}}\sigma_{^7\mathrm{Be}}^{\mathrm{SSM}} \\ \rho^{\mathrm{SSM}}\sigma_{^8\mathrm{B}}^{\mathrm{SSM}}\sigma_{^7\mathrm{Be}}^{\mathrm{SSM}} & \sigma_{^7\mathrm{Be}}^{\mathrm{SSM}}\sigma_{^7\mathrm{Be}}^{\mathrm{SSM}} \end{pmatrix} \tag{5.10}$$

where σ_{BX} is the Borexino experimental uncertainty which include both statistical and the symmetrized systematic error, while σ_{SSM} and ρ_{SSM} indicates the uncertainty and the correlation factor of the SSM under consideration. When the data come from the model considered in this test, the value of the test statistics in Eq. (5.8) is distributed as a χ^2 with 2 degrees of freedom, and it is therefore trivial to compute the corresponding p-value. When testing the prediction of the HZ SSM a p-value of 0.96 is found while for the LZ SSM expectations it is 0.17, which indicates a good compatibility of the data with both the SSMs.

The fact that the data are compatible with both models does not prevent to establish which one is more plausible. To do so, a frequentist hypothesis test similar to the one discussed in Sec. 5.2.1 has been performed. In this case the two alternative hypotheses are the HZ and LZ SSM models, and the test statistic t is

$$t = \chi^2(\mathrm{HZ}) - \chi^2(\mathrm{LZ}) \tag{5.11}$$

where $\chi^2(\mathrm{HZ})$ and $\chi^2(\mathrm{LZ})$ are defined in Eq. 5.8 and the HZ/LZ argument refers to the values of $\Phi_{^8\mathrm{B}}^{\mathrm{SSM}}$, $\Phi_{^7\mathrm{Be}}^{\mathrm{SSM}}$ and ρ^{SSM} that are used in the computation. The probability distribution function of the test statistics for both the
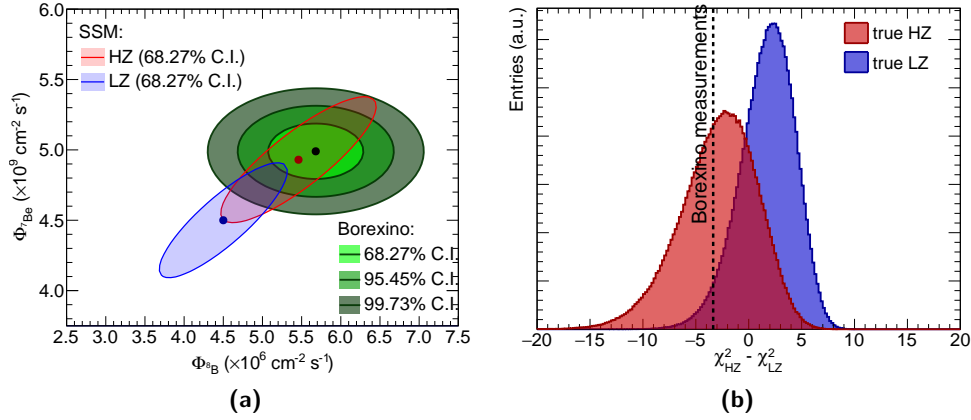
Figure 5.11: (a) Borexino results for $\Phi_{^8\mathrm{B}}$ and $\Phi_{^7\mathrm{Be}}$ (green point and shaded area) compared to the theoretical predictions of the HZ and LZ SSM. **(b)** Distribution of the test statistics t used to perform a frequentist hypothesis test between the HZ and LZ SSM. The value of t obtained from the Borexino measurements is shown as a dashed line.

hypotheses has been derived with a toy-MC technique. Thousands of pairs of $^7\mathrm{Be}$ and $^8\mathrm{B}$ fluxes were generated assuming the HZ/LZ model and the respective uncertainties that have been combined with the experimental uncertainties of Borexino measurements[5]. For each pair of values of $\Phi_{^7\mathrm{Be}}$ and $\Phi_{^8\mathrm{B}}$ the value of t is computed and used to build the distributions that are shown in Fig. 5.11b. The value of the test statistics obtained using Borexino results is $t_{\mathrm{BX}} = -3.49$, which is fully consistent with the HZ model but lies on the tail of the distribution expected from the LZ SSM. The probability of obtaining a result more extreme than the observed one is 0.034, which means that the LZ SSM hypothesis is rejected with 1.8σ significance.

To cross check this result, a Bayesian hypothesis test has also been performed. In the Bayesian framework the fluxes of $^7\mathrm{Be}$ and $^8\mathrm{B}$ are fitted against Borexino measurements assuming the SSM HZ and LZ predictions (with the corresponding uncertainties) as prior probability distributions. In the analysis, performed using the `BAT` package [147], the likelihood of the data is constructed as the sum of two uncorrelated Gaussian measurements, one for $\Phi_{^7\mathrm{Be}}$ and one for $\Phi_{^8\mathrm{B}}$. The resulting posterior probability distributions are reported in Fig. 5.12 along with both the HZ and LZ prior probability distribution to show the update of knowledge achieved after the Borexino measurements.

---

5 The precision of the Borexino measurements has been assumed to be independent from the simulated flux in the range allowed by the SSMs.

To test which one of the HZ/LZ SSM is more likely given the Borexino measurements, the Bayes factor is computed. The Bayes factor K is defined as the ratio of the marginal likelihood $P(\text{data}|\text{HZ/LZ})$, namely

$$
\begin{aligned}
K &= \frac{P(\text{data}|\text{HZ})}{P(\text{data}|\text{LZ})} \\
&= \frac{\int P(\text{data}|\boldsymbol{\Phi}, \text{HZ}) \cdot P(\boldsymbol{\Phi}|\text{HZ}) \mathrm{d}\boldsymbol{\Phi}}{\int P(\text{data}|\boldsymbol{\Phi}, \text{LZ}) \cdot P(\boldsymbol{\Phi}|\text{LZ}) \mathrm{d}\boldsymbol{\Phi}} \cdot \frac{P(\text{HZ})}{P(\text{LZ})}
\end{aligned}
\tag{5.12}
$$

where $\boldsymbol{\Phi}$ indicates the pair $(\Phi_{8_\text{B}}, \Phi_{7_\text{Be}})$. Assuming the HZ and LZ SSM to be equally probable, *i.e.* imposing $P(\text{HZ})/P(\text{LZ}) = 1$ in Eq. 5.12, the analysis yields a Bayes factor of 4.9 or, equivalently, states that the odds of HZ against LZ are $5 : 1$, thus confirming a mild preference for the HZ SSM.

For sake of completeness, a global frequentist analysis was performed by the Collaboration following the procedure already adopted in [82] which takes into account all the solar neutrino data available and KamLAND data, leaving as free parameters both the ^{8}B and ^{7}Be neutrino fluxes and the oscillation parameters. The best fit value in the Φ_{8_B}–Φ_{7_Be} space as well as the confidence intervals are shown in Fig. 5.13. The value of Φ_{7_Be} is basically fixed by the Borexino measurement presented in the previous section, while the results of SNO and Super-Kamiokande lead to a slightly smaller flux of ^{8}B neutrinos, visibly reducing the significance of the hint provided by Borexino in favour of the HZ SSM.

At this point it is important to stress that for the first time the analysis is limited by the uncertainties of the SSM predictions. A very significant contribution to the determination of the solar metal abundances might come from the measurement of the CNO neutrino flux. Indeed, while a change in the solar composition produces the same effects on the *pp* chain neutrino fluxes of a modification of the opacity profile (which does not only depend on the metal abundances), CNO neutrinos can be used to provide a direct determination of the metal abundances in the solar core [148, 149]. As discussed in Sec. 5.1.3 the present analysis of Borexino data is not sensitive enough to provide a measurement of the CNO neutrino flux, but a strategy to detect CNO neutrinos is presented in the following chapters.

5.3 CONCLUSIONS

In this Chapter the first simultaneous measurement of *pp*, ^{7}Be and *pep* neutrinos along with the best upper limit of the CNO neutrino rate have been presented. These measurements represents the most accurate determination of the low energy solar neutrino interactions rate to date, and combined with an independent measurements of ^{8}B neutrinos make Borexino the only experiment with the potential to study the entire spectrum of solar neutrinos.

These results have been used to test both neutrino and solar phenomenology. Assuming the solar neutrino fluxes predicted by the HZ SSM, that is favoured by helioseismological observation respect to the LZ SSM, the presence of a matter effect in neutrino oscillation is established at with a significance of 2.1σ.

Postulating the neutrino oscillation framework, the inferred fluxes of solar neutrinos have been used to probe the SSM predictions. The resulting neutrino luminosity, which is dominated by the contribution of *pp* neutrinos, is in excellent agreement with the solar luminosity measured with photons. A hypothesis test based on the measurements of ^{8}B and ^{7}Be neutrinos has been performed between the HZ and LZ SSM. The different chemical composition of the Sun in the HZ and LZ SSM modifies the plasma opacity that in turn affects the temperature of the core. The fluxes of ^{8}B and ^{7}Be neutrino exhibit the stronger temperature dependence in the *pp* chain, and values measured by Borexino are consistent with the core temperature predicted by the HZ SSM, rejecting the LZ SSM prediction with a 1.8σ significance. Similarly, a Bayesian hypothesis test confirms a mild preference for the HZ SSM.

While these results do not give a definitive answer to the solar metallicity puzzle, a measurement of the CNO neutrino flux would probe directly the abundance of C and N in the Sun. A strategy to achieve the first detection of CNO neutrinos in Borexino is outlined in the following chapters.
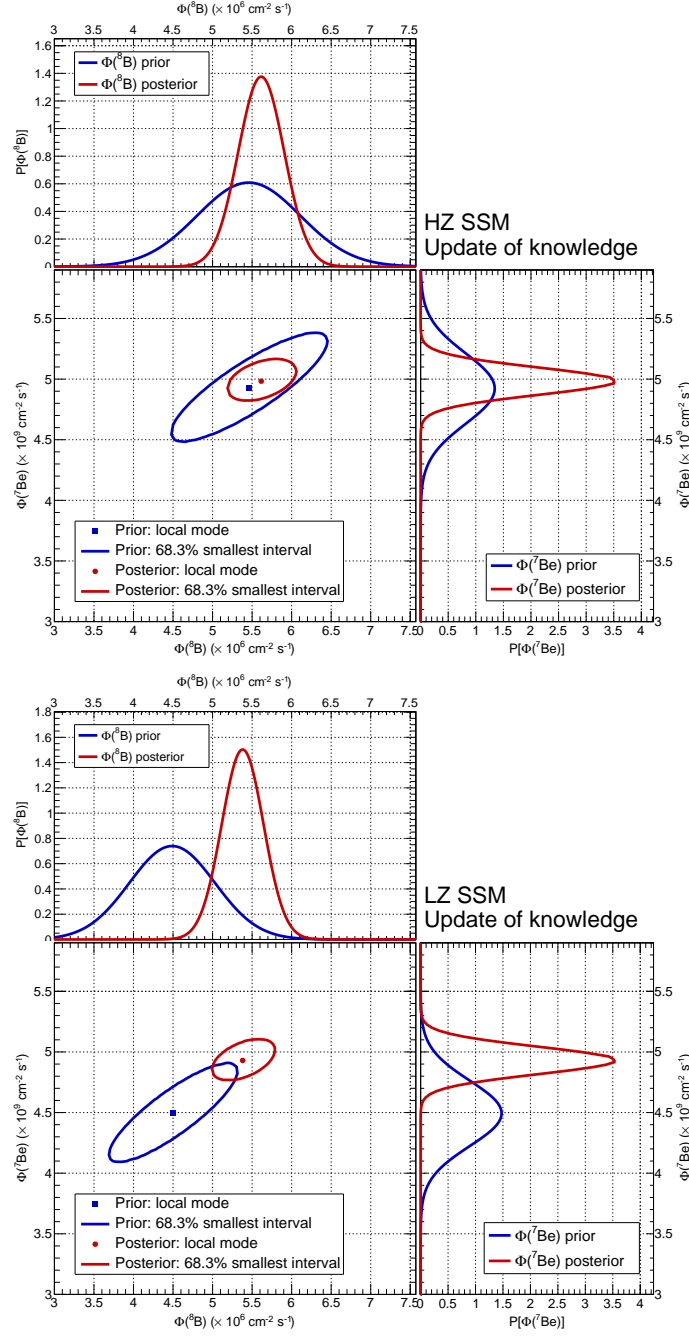
Figure 5.12: Posterior probability distribution for $\Phi_{^8B}$ and $\Phi_{^7Be}$ from a Bayesian fit of Borexino data using as priors the expectations from the HZ (top panel) and LZ (bottom panel) SSM.

Figure 5.13: ^{8}B and ^{7}Be fluxes determined from a global analysis of all the available solar neutrino and KamLAND data. The expected values from the HZ and LZ SSM, along with the latest Borexino results are also shown for comparison.
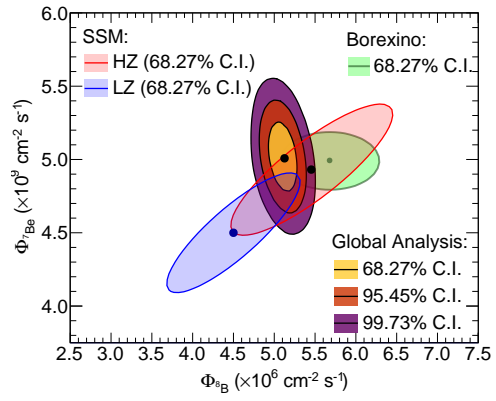
CHAPTER 6

THE SEARCH FOR CNO NEUTRINOS IN BOREXINO

Neutrinos from the sub-dominant CNO cycle are the only low-energy component of the solar neutrino spectrum that has not been detected yet. The analysis presented in the last Chapter provides the most stringent upper limit on the CNO neutrino flux to date and Borexino is the only running experiment with the potential to measure for the first time this elusive component. The discovery of CNO neutrinos is currently one of the main physics goal of the experiment and in this Chapter a strategy to achieve this objective is proposed.

After recalling the relevance of CNO neutrinos for solar physics in Sec. 6.1, Sec. 6.2 presents a comprehensive sensitivity study which shows the crucial importance of external constraints on the ^{210}Bi and *pep* background on the sensitivity to the CNO signal. The strategy to achieve an independent assessment of the background rate is outlined in Sec. 6.3. The rate of *pep* neutrinos can be constrained using argument from nuclear and solar physics, while the supposed equilibrium of the ^{210}Pb decay chain can be exploited to obtain a measurement of the ^{210}Bi rate. A preliminary evaluation of possible systematic uncertainties is discussed in Sec. 6.5. Section 6.6 finally presents an alternative method to break the correlation between the ^{210}Bi and CNO rate exploiting the different time evolution of the two components. While this method does not improve the sensitivity with the current dataset, it can be used in the future if the Borexino data taking will continue beyond 2022.

6.1 THE IMPORTANCE OF CNO NEUTRINOS FOR SOLAR PHYSICS

The CNO cycle is a subdominant process in the Sun accounting for only $\approx 1\%$ of the total energy production, but the measurement of neutrinos produced in these reactions is one of the main goal of current and future solar neutrino experiments.

The detection of CNO neutrinos would be the first direct evidence of the occurrence of the CNO cycle, that is supposed to be the driving energy production mechanism for stars that are ≈ 1.3 times heavier than the Sun [150]. As anticipated in Sec. 5.2.2, a measurement of CNO solar neutrinos would also be extremely meaningful for solar physics as well. The flux of neutrinos emitted in the CNO cycle retains a linear dependence on the abundance of C

and N in the Sun core, and therefore a measurement of CNO neutrinos can be crucial to solve the solar metallicity puzzle introduced in Sec. 1.2.3.

In the framework of the SSM, the fluxes of solar neutrinos –including the CNO cycle ones– depend on the input parameters of the model, which include the solar luminosity, opacity, diffusion as well as the initial abundances of He and other heavier elements. The absolute flux of each component of the solar neutrino spectrum is however correlated to the Sun core temperature $T_\odot$, which describes the velocity of the nuclei in the solar plasma and thus regulates the nuclear reaction rates.

Serenelli et al. [149] show that the flux of ^{8}B, thanks to its strong temperature dependence, can be used as a "thermometer" to remove the dependence of "environmental" parameters of the SSM (*i.e.* those directly affecting the temperature profile) and highlight the parameters that act differently on the individual components of the solar neutrino flux [35]. A weighted ratio of ^{15}O and ^{8}B neutrino fluxes that is almost independent on temperature can be used to isolate the linear dependence on the core abundance of C and N of CNO neutrinos [35]

$$
\left(\frac{\Phi(^{15}\text{O})}{\Phi(^{15}\text{O})^{\text{SSM}}} \right) \Big/ \left(\frac{\Phi(^{8}\text{B})}{\Phi(^{8}\text{B})^{\text{SSM}}} \right)^{0.785} =
$$
$$
\left[\frac{N_\text{C} + N_\text{N}}{N_\text{C}^{\text{SSM}} + N_\text{N}^{\text{SSM}}} \right] \times (1 \pm 2.6\%(\text{Diff}) \pm 10.6\%(\text{Nucl})) \quad (6.1)
$$

where the "SSM" superscript indicates the SSM prediction. Therefore, a measurement of the CNO neutrino flux can be converted into a measurement of the abundance of C and N in the Sun core (N_C and N_N), with the uncertainty given by a term due to the settling of heavy elements in the Sun core (Diff) and the other (Nucl) due to uncertainties in the cross sections of the $^7\text{Be}(p,\gamma)^8\text{B}$ and $^{14}\text{N}(p,\gamma)^{15}\text{O}$ reactions.

6.2   BOREXINO SENSITIVITY TO CNO NEUTRINO

The detection of CNO neutrinos in Borexino is particularly challenging for three main reasons. First, the CNO cycle is responsible for only $\approx 1\%$ of the total energy production of the Sun and as a consequence the flux of CNO neutrinos is much smaller than for the neutrinos emitted in the *pp* chain. Second, the electron recoil spectrum due to CNO neutrino interaction does not show any peculiar spectral feature that would stand out in a spectral analysis. Third, the signal of CNO neutrinos is always smaller than background fluctuations event in the region where the CNO signal is maximum.

Figure 4.1 shows the expected contribution from both neutrino and background components to the Borexino spectrum. The spectral shape of the β decay of ^{210}Bi contaminations in the scintillator is very similar to the low energy part of the CNO neutrino spectrum, while for the higher energy tail of

the $\nu(\text{CNO})$ spectrum the main background are $\nu(pep)$ events and the cosmogenic ^{11}C background. The CNO neutrino signal can thus be mimicked by the interplay between the ^{210}Bi and $\nu(pep)$ events distributions, and their spectral degeneracy induces correlations in the analysis such that a spectral fit of Borexino data has practically no sensitivity to a CNO neutrino signal.

In order to measure the rate of CNO neutrinos, the background rate due to ^{210}Bi and *pep* events must be constrained. In this Section the sensitivity of Borexino under different assumptions on the external constraints is discussed. A simplified model of the sensitivity based on a simple rate analysis is outlined in Sec. 6.2.1, while Sec. 6.2.2 presents the expected statistical uncertainty on the CNO neutrinos rate along with the discovery potential of the experiment for different configurations of the external constraints. The impact of an upper limit on the ^{210}Bi rate on the sensitivity to the CNO neutrinos rate is discussed in Sec. 6.2.3.

6.2.1 *Borexino as a counting experiment*

The similarities in the event distribution of CNO neutrinos and ^{210}Bi and *pep* background make impossible to a spectral to resolve the tree component individually. Given such limited sensitivity, it is possible to study the impact of external constraints on the ^{210}Bi and *pep* background on the sensitivity to CNO neutrinos neglecting the information from the energy spectrum and treating Borexino as a counting experiment.
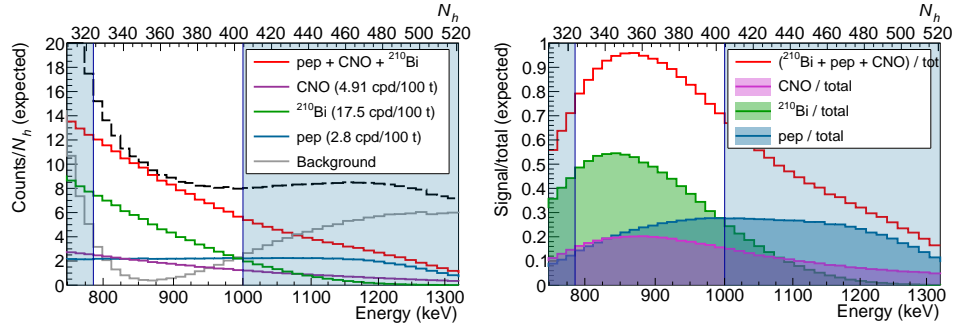
Figure 6.1: *Left Panel*: Expected Borexino spectrum assuming the rates predicted by the HZ SSM for $\nu(\text{CNO})$ and $\nu(pep)$ and a rate of 17.5 cpd/100 t for ^{210}Bi. *Right Panel*: Corresponding fraction of counts as a function of energy for selected spectral components.

In the energy window between 0.79 and 1 MeV where the sum of ^{210}Bi, *pep* and CNO events dominates the spectrum (Fig. 6.1), the total number of events is given by

$$N_{\text{tot}} = E \cdot (\varepsilon_{\text{CNO}} R_{\text{CNO}} + \varepsilon_{pep} R_{pep} + \varepsilon_{\text{Bi}} R_{\text{Bi}}) \quad (6.2)$$

where E is the exposure, R_i is the rate of the component i and the coefficient ε_i is the fraction of events of i falling into the energy range under study. Assuming to have an independent estimation of the rate of *pep* neutrinos

$(\tilde{R}_{pep} \pm \tilde{\sigma}_{pep})$ and of ^{210}Bi $(\tilde{R}_{\mathrm{Bi}} \pm \tilde{\sigma}_{\mathrm{Bi}})$ from Eq. 6.2 it is possible to obtain the CNO neutrino rate and its uncertainty, that are given by

$$R_{\mathrm{CNO}} = \frac{N_{\mathrm{tot}}}{E \cdot \varepsilon_{\mathrm{CNO}}} - \frac{\varepsilon_{\mathrm{Bi}}}{\varepsilon_{\mathrm{CNO}}} \tilde{R}_{\mathrm{Bi}} - \frac{\varepsilon_{pep}}{\varepsilon_{\mathrm{CNO}}} \tilde{R}_{pep} \qquad (6.3)$$

$$\sigma_{\mathrm{CNO}} = \frac{1}{E \cdot \varepsilon_{\mathrm{CNO}}} \sigma_{N_{tot}} \oplus \frac{\varepsilon_{\mathrm{Bi}}}{\varepsilon_{\mathrm{CNO}}} \tilde{\sigma}_{\mathrm{Bi}} \oplus \frac{\varepsilon_{pep}}{\varepsilon_{\mathrm{CNO}}} \tilde{\sigma}_{pep} \qquad (6.4)$$

For a given number of counts N_{tot}, both the CNO neutrino rate and its accuracy are thus determined by the independent estimations of the background rate $(\tilde{R}_{\mathrm{Bi}}, \tilde{R}_{pep})$ and their uncertainties $(\tilde{\sigma}_{\mathrm{Bi}}, \tilde{\sigma}_{pep})$.

It is then possible to estimate the expected statistical error on the ν(CNO) rate as a function of σ_{Bi} and σ_{pep} assuming a plausible rate of $17.5 \,\mathrm{cpd}/100 \,\mathrm{t}$ for ^{210}Bi, the rate of pep neutrinos as predicted from the SSM and an exposure of ≈ 3 years. The statistical uncertainty on the rate of CNO neutrinos is shown as a function of the uncertainty on the ^{210}Bi estimation in Fig. 6.2, where the relative error on the ν(pep) rate has been set at 1 and 10%. The term $\sigma_{N_{\mathrm{tot}}}$ that takes into account statistical fluctuations in the total number of counts has been evaluated with a toy-MC technique in order to consider residual exposure in the ^{11}C-tagged dataset.
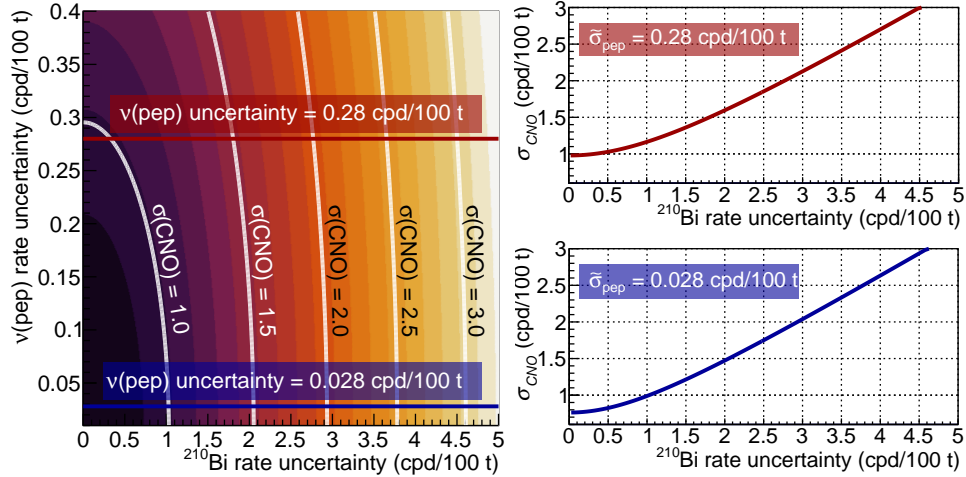
Figure 6.2: *Left Panel* Uncertainty of the CNO neutrino rate as a function of the error in in the determination of the ^{210}Bi and ν(pep) rate computed with a simple rate analysis (Eq. (6.4)). *Right Panel* Expected statistical uncertainty on the CNO neutrino rate as a function of the ^{210}Bi rate accuracy computed using the same simplified rate analysis a assuming a 10% (top) and 1% (bottom) precision on the ν(pep) rate.

The left panel of Fig. 6.2 shows that the CNO uncertainty is limited by the uncertainty on the ^{210}Bi rate. Improving the precision on the ν(pep) rate reduces the uncertainty on the CNO only when the ^{210}Bi rate is known with good precision.

Including additional exposure will not bring significant improvement to the CNO uncertainty: enhanced statistics will only affect the term $\sigma_{N_{\text{tot}}}$ in Eq. (6.4), that represents a subdominant contribution as compared with the uncertainty of the ^{210}Bi rate. Figure 6.3 shows the resulting statistical uncertainty for CNO considering from 6 moth to 5 years of exposure in addition to the 3 years initially considered. The largest impact of the increased statistics is visible where the ^{210}Bi rate is known with good precision, while for $\tilde{\sigma}_{\text{Bi}} \geq 2.0\,\text{cpd}/100\,\text{t}$ the effect does not change significantly the result. This way to consider additional exposure is of course over-simplistic. Systematic uncertainties are not taken into account, as the fact that a spectral analysis might benefit from increased statistics is not considered[1]: nevertheless, this simplified model shows that increasing the exposure will not improve the sensitivity as much as an accurate measurement of the background rate will.
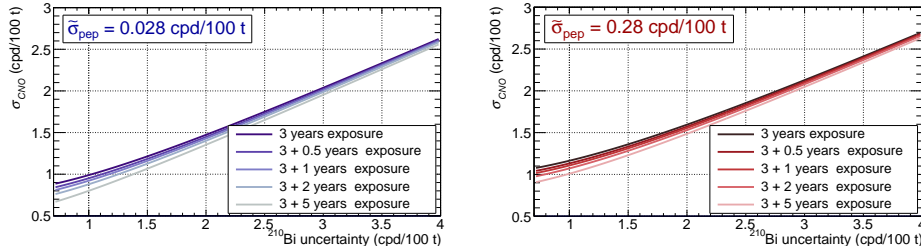
Figure 6.3: Statistical uncertainty of the CNO rate obtained with a simple rate analysis, considering from 6 month to 5 years of additional exposure assuming a 1% (left) and 10% accuracy on the $\nu(pep)$ rate.

A simple rate analysis is at the cornerstone of the sensitivity of many experiments: it is extremely useful to identify the key elements that can make a measurement possible that are often masked by the complications brought by more refined analysis method. Calling Borexino a counting experiment is of course reductive, as it can exploit the specific spectral signature of signal and background component. However this simplified model gives already a feeling of the precision in the background estimation required to achieve a given sensitivity to the CNO rate: the predictions of the HZ and LZ SSMs of the CNO rate in Borexino ranges between 3.1 and 5.5 cpd/100 t, therefore a value of $\sigma_{\text{CNO}} \lesssim 1.5\,\text{cpd}/100\,\text{t}$ is needed to have an adequate sensitivity to the CNO signal, and the contour plot on the left panel of Fig. 6.2 indicates the accuracy in the background determination required to achieve it.

6.2.2 *Sensitivity to CNO neutrinos with background constraints*

The simple rate analysis discussed in the previous section shows that an independent estimation of the background rate is crucial to improve the sensitivity of the analysis to the CNO neutrino signal. In this section the effects of such

---

1 Even though the detector resolution is expected to decrease in time due to the progressive reduction of the number of working PMTs.

additional constraints to the multivariate analysis introduced in the previous Chapters are discussed.

All the studies presented in this section considers the period of time between July 2013 and May 2016[2] and have been performed assuming the rate of *pep* neutrinos predicted by the SSM (2.80 cpd/100 t) and a rate of 17.5 cpd/100 t for ^{210}Bi. Six different configurations of the external constraints has been considered: the rate of *pep* neutrinos has been constrained with an accuracy of 1% and 10% (0.028 and 0.28 cpd/100 t), while the precision of the constraints on ^{210}Bi is 20%, 14% and 10% (3.50, 2.45 and 1.75 cpd/100 t). These values are within reach of the methods to obtain an independent assessment of the background rate that are discussed in Sec. 6.3.

The statistical uncertainty achievable on the CNO rate when both the *pep* and ^{210}Bi background are constrained is presented in Sec. 6.2.2.1. The discovery potential to a CNO neutrino signal is discussed separately in Sec. 6.2.2.2.

6.2.2.1 *Statistical uncertainty of the CNO rate*

The expected statistical uncertainty of the CNO neutrino rate achievable performing the multivariate analysis developed in Chap. 3 has been evaluated in the same way used to establish the sensitivity of the low-energy spectroscopy of solar neutrinos presented in Sec. 4.3, *i.e.* performing thousands of simultaneous fit of pseudo–dataset and considering the width of the best fit estimate distribution as the expected statistical uncertainty. These analyses has been performed fitting of simultaneously the TFC-subtracted and tagged pseudo–data considering the events energy spectrum and radial distribution. The independent estimations of the ^{210}Bi and $\nu(pep)$ background rate has been folded in the analysis by adding two multiplicative Gaussian penalty terms to the fit likelihood according to the procedure discussed in Sec. 3.2.1.

The resulting sensitivity to the CNO neutrinos signal combines the *rate* analysis, recalled by the introduction of the external constraints on the background, with the events distribution *shape* information exploited by the fit. The results on the expected statistical uncertainty are shown Fig. 6.4, where the expectation from the rate analysis presented in the previous section is also reported for comparison. The outcomes of this refined analysis are fully compatible with the results of a rate analysis when the uncertainty on the independent measurement of the ^{210}Bi rate is better than ≈ 2.45 cpd/100 t, while the impact of the shape information becomes non-negligible for looser constraints. The improvement due to the spectral analysis depends on the precision of the estimation of the $\nu(pep)$ rate.

The substantial agreement between the two analysis methods shown in Fig. 6.4 clearly shows that the bulk of the sensitivity to the CNO neutrino signal comes from a simple rate analysis, and that the spectral fit brings a

---

2 The first year and a half of Phase II has not been considered because there are evidence that the concentration of ^{210}Pb (and ^{210}Bi) was not uniform in the detector. This fact would prevent the measurement
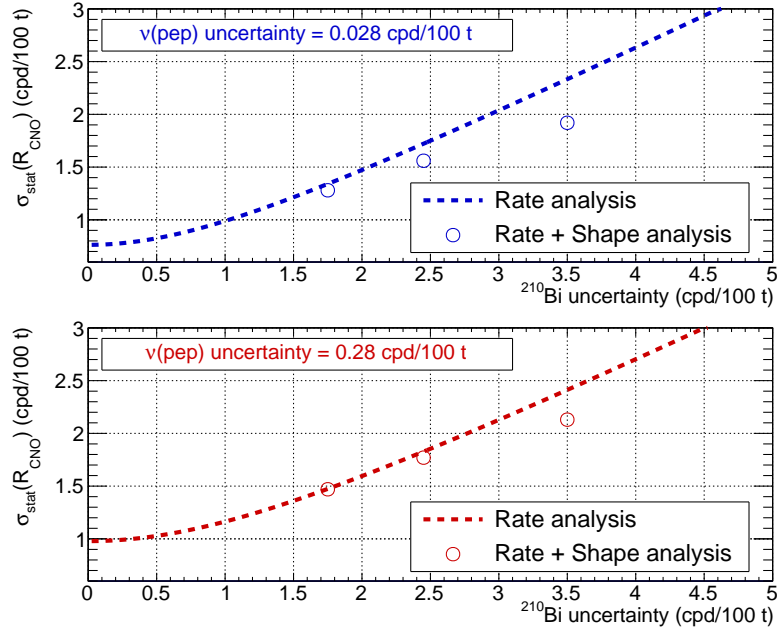
Figure 6.4: Expected statistical uncertainty on the rate of CNO neutrinos obtained from a MV analysis with both the $\nu(pep)$ and the ^{210}Bi rate constrained (Rate + Shape analysis). The expectations from the rate analysis developed in Sec. 6.2.1 are also shown for comparison.

mild improvement only if the constraint on the ^{210}Bi rate is relatively loose. This is a strong limitation of the analysis, as the estimations of the background rate will almost determine the best fit value of the CNO rate and its uncertainty. This also means that any possible bias in the determination of the background will be transferred to the CNO measurement. In the rate analysis scenario, the rate of ^{210}Bi and *pep* are fixed by the mode of the external constraints as the constraints are the only information on these parameters. A bias on the mode of the constraints will hence transfer linearly to the CNO reconstructed rate according to Eq. (6.3). The situation becomes slightly more complicated when a real spectral analysis combined with some constraints on the background is performed. In this case the shape information will be in tension with the constraints if the latter is biased, mitigating the impact on the CNO extracted rate. Figure 6.5 shows the bias on the reconstructed CNO rate induced by a difference between the injected ^{210}Bi rate and the mode of the constraint. As expected, the shape information mitigates the impact of bias but only marginally. Independently from the analysis, a bias of $1\,\text{cpd}/100\,\text{t}$ on the ^{210}Bi rate results in a $0.5\,\text{cpd}/100\,\text{t}$ bias on the extracted CNO rate.

6.2.2.2 *Discovery potential to a CNO neutrino signal*

In the previous section it has been shown how the uncertainty on the CNO rate can be reduced to the level of 1–$3\,\text{cpd}/100\,\text{t}$ depending on the strength
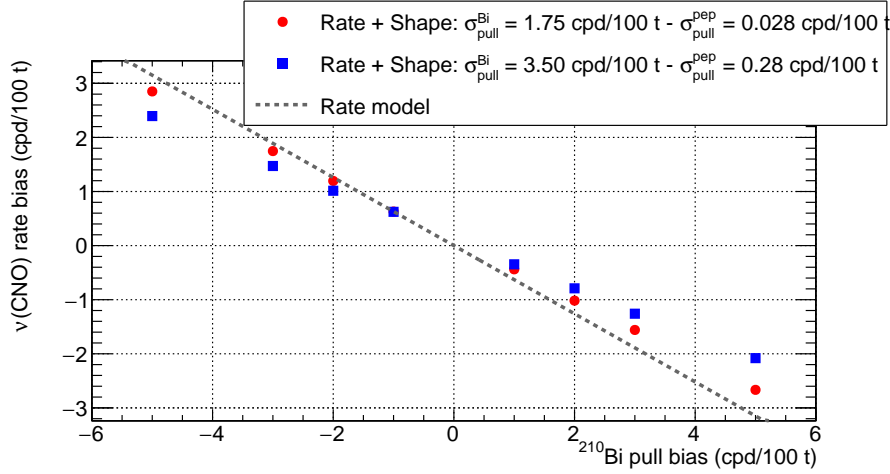
Figure 6.5: Reconstructed CNO rate bias as a function of the difference between the ^{210}Bi constraint mean and the injected value. The results are shown for different configuration of the constraints accuracy along with the expected bias from the rate analysis presented in Sec. 6.2.1.

of the constraints on the background. This implies that the strength of the CNO signal could differ from zero up to 3 (5) σs assuming the value predicted by the LZ (HZ) Standard Solar Model. This information is however not what should be used to claim the evidence for a signal as it indicates only the signal strength and not the probability that the background can mimic a signal.

To properly assess the discovery power of the analysis a frequentist hypothesis test has been performed using the profile likelihood ratio q_0 as a test statistics [142] (See App. A). In the search for the CNO signal two hypothesis are considered: the *null* hypothesis H_0 where only the background is present, and the *alternative* hypothesis H_1 that includes the presence of a signal in addition to the background.

The significance of the signal is quantified by evaluating the compatibility of the observed data with the background hypothesis H_0 that is expressed through a *p*-value. It is then useful to define the *median discovery power* of a measurement as the median *p*-value obtained testing H_0 under the assumption of signal H_1 as sketched in Fig. 6.6. This requires to know the PDF of q_0 for the background only hypothesis $f(q_0|\text{no CNO})$ and in case of the presence of a given CNO signal $f(q_0|\text{HZ/LZ CNO})$.

The PDFs of the test statistics q_0 have been obtained analysing pseudo–dataset as discussed in App. A for both the simplified rate analysis presented in Sec. 6.2.1 and the multivariate fit.

The discovery power has been estimated assuming the CNO neutrino rate predicted by the HZ and LZ SSM under the same configuration of the external constraints used to study the expected statistical uncertainty on the CNO rate in the previous section. The results are shown in Fig. 6.7, where the discovery
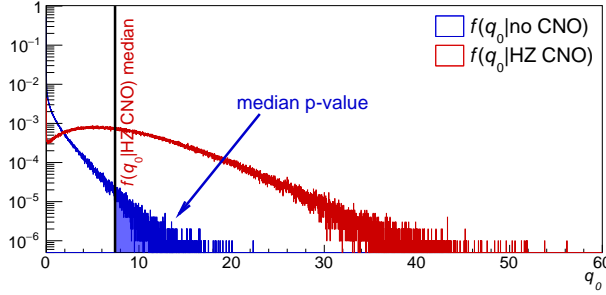
Figure 6.6: Illustration of the definition of the median discovery power. The median p-value is computed testing the background hypothesis $f(q_0|\text{no CNO})$ for the median of the test statistics distribution obtained assuming the CNO rate predicted by the HZ SSM.

power for simple rate analysis and for the multivariate fit are indicated by empty and filled marker respectively.

The discovery power depends on the strength of the signal and on the accuracy of the external constraints: an higher rate of CNO neutrinos, as well as stronger constraints result in higher sensitivity. For an uncertainty on the *pep* and ^{210}Bi background rate of 0.0028 cpd/100 t and 2.45 cpd/100 t respectively, in 50% of the case the significance of the CNO signal will be higher than 3σ (2.1σ) for the CNO rate predicted by the HZ (LZ) SSM.

The difference between the median p-value from the rate analysis and the one obtained including the spectral information gets smaller increasing the precision of the pulls. This effect is coherent with the results on the CNO statistical uncertainty presented in Sec. 6.2.2.1, indicating that the impact of the shape information is larger when constraints are relatively weak while it becomes negligible when external constraints gets more stringent.

6.2.3   *Impact of an upper limit on* ^{210}Bi *rate*

The method for the independent measurement of the ^{210}Bi rate presented in Sec. 6.3.2 under less stringent assumptions can be used to provide an upper limit to the ^{210}Bi background, measuring the sum of the ^{210}Bi rate and a positive contribution that does not affect the region of the spectrum sensitive to CNO.

This information is included in the analysis applying a constraint on the sum of the ^{210}Bi rate and an "invisible" term that acts as a nuisance parameter [3]. This construction of the constraint is equivalent to applying a one-sided Gaussian penalty which enforces only an upper limit on the ^{210}Bi rate. The upper limit on the ^{210}Bi rate is transmitted as a lower limit to CNO rate. Therefore, if the method to measure the ^{210}Bi rate gives exactly the rate of the ^{210}Bi with no additional contribution, the discovery power will be the same one obtained

---

3 An invisible parameter has been added to the fit model including a fictitious spectral component with no entries in the energy range considered in the analysis and requiring its rate to be positively defined.
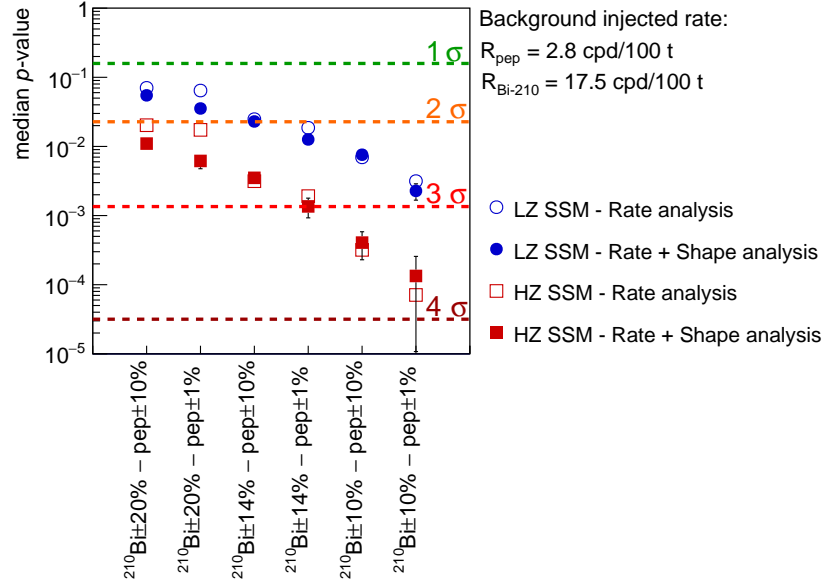
Figure 6.7: Median discovery power of a CNO neutrino signal as predicted by the HZ and LZ SSM for different accuracies in the determination of the background rate. The results from the rate analysis are indicated with empty markers, while those obtained including the spectral information are reported using filled markers.

using a symmetric constraint on the ^{210}Bi rate. This result has been confirmed performing an hypothesis test analogous to the one described in Sec. 6.2.2.2. The sum of ^{210}Bi and the nuisance parameter has been constrained applying a Gaussian penalty with mean given by the ^{210}Bi injected rate and varying accuracy. The results are shown in Fig. 6.8, where the discovery power obtained from the rate analysis and applying a symmetric constraint are also reported for comparison. In case the measurement of ^{210}Bi background includes an additional contribution the discovery power will decrease.

An upper limit on the ^{210}Bi rate can be enough to claim for an evidence of CNO neutrinos, but it will not allow to obtain a precise measurement of the CNO rate that is needed in order to contribute to the solution of the solar metallicity puzzle. Since a lower limit on the ^{210}Bi background is not provided, the upper limit of the CNO neutrinos rate will not be stringent. Figure 6.9 shows the distribution of the best fit estimates obtained from the fit of pseudo–datasets applying a constraint of the *pep* neutrinos rate and leaving the ^{210}Bi rate free (Fig. 6.9a), constraining the ^{210}Bi rate with a symmetric penalty (Fig. 6.9b) and applying the upper limit described in this section (Fig. 6.9c).

The expected confidence intervals in Fig. 6.9c are asymmetric: for large values of the CNO rate they are like those obtained leaving the ^{210}Bi free, while for small values the intervals reflect the ones returned by constraining the ^{210}Bi rate.
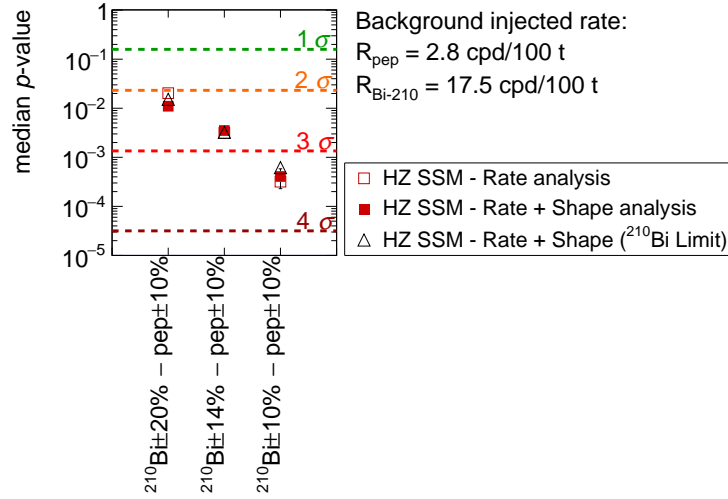
Figure 6.8: Median discovery power obtained assuming the CNO rate predicted by the HZ SSM and constraining the ^{210}Bi with an upper limit and a symmetric penalty.

6.3 STRATEGY FOR AN INDEPENDENT BACKGROUND ASSESSMENT

The sensitivity studies presented in the previous Section shows that Borexino has the potential to detect CNO neutrinos even in the case where the spectral shape does not provide any information to the analysis, as long as the *pep* and ^{210}Bi background rate can be independently assessed.

The rate of *pep* neutrinos can be constrained exploiting its connection with the *pp* neutrino flux as already done in the analysis presented in Chap. 5, and its accuracy can be improved by implementing a constraint based on solar luminosity as explained in Sec. 6.3.1.

As regards the ^{210}Bi background rate, a measurement can be achieved exploiting the equilibrium with its parent and daughter nuclei in the decay chain as discussed in Sec. 6.3.2.

6.3.1 *pep neutrino background estimation*

Neutrinos from the *pep* reaction are produced at the beginning of the *pp* chain but contributes for only $\approx 1\%$ of the total production rate of ^{2}H in the Sun, while most of it is due to the reaction generating *pp* neutrinos.

The ratio Φ_{pp}/Φ_{pep} is well known and basically fixed by nuclear physics. For the reactions generating *pp* (Eq. 1.30) and *pep* (Eq. 1.31) neutrinos the matrix element entering in the cross section calculation is the same and the dominant difference in the thermal averaged cross section $\langle \sigma v \rangle$ (Eq. 1.29) is given by the phase space available. Since *pp* and *pep* neutrinos are produced almost in the same region of the Sun the proton density for the two reactions is very close. Thus, recalling Eq. (1.29) describing the reaction rate between two nuclei in the Sun, the ratio between the rate of the *pp* and *pep* reaction results basically

(a) ^{210}Bi rate free   **(b)** ^{210}Bi rate constrained   **(c)** ^{210}Bi upper limit
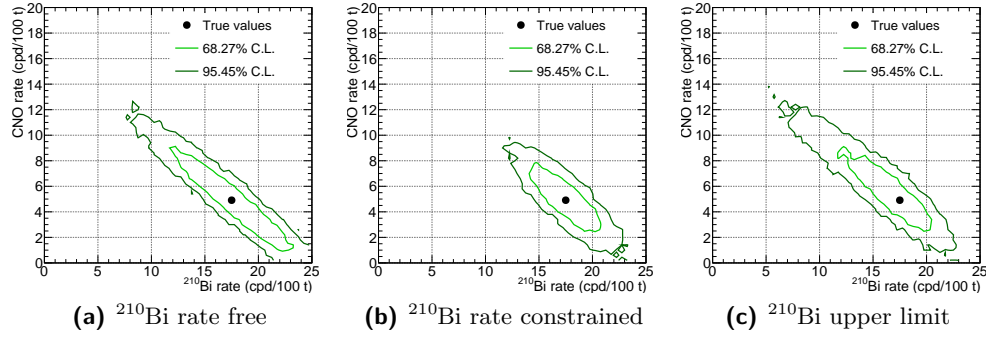
Figure 6.9: Distribution of the ^{210}Bi and CNO best fit estimates obtained fitting thousands of pseudo-datasets applying a 10% constraint on the *pep* rate and leaving the ^{210}Bi free **(a)**, constraining it with a 14% precision **(b)**, and imposing an upper limit **(c)**.

determined by the ratio of the reactions phase space [151]. This fact motivates the approach adopted in [132] that has been used also in Chap. 5 to constraint Φ_{pep} and improve the sensitivity of Borexino to the CNO signal.

The accuracy on the *pep* neutrino flux can be further improved if an additional constraint based on the solar luminosity is also implemented. The idea, formalized by Bahcall in [146] and recently reviewed in [144], consists in enforcing the equivalence between the solar luminosity observed in the neutrino channel and in the well measured photon channel. In this way it is possible to invert Eq. 5.5 and obtain a constraint on the sum of the neutrino fluxes. Since the precision on the solar photon luminosity is of the order of $\approx 4‰$ and the one of the α coefficient is of the order of $\approx 1‰$, given the current accuracy on the other solar neutrino fluxes, a $\approx 1\%$ uncertainty on Φ_{pp} can be achieved [132, 152]. Constraining the ratio between the Φ_{pp} and Φ_{pep} leads to an accuracy of $\approx 1\%$ on the *pep* rate as well.

The correspondence of the neutrino and photon luminosity relies on the following assumptions[143]:

- The energy in the Sun is produced only by hydrogen burning through the *pp* chain and the CNO cycle, therefore excluding exotic phenomena that might contribute to the total solar power.

- The Sun in in equilibrium, thus the observed luminosity is constant over a $\sim 10^5$ years time scale

- ^{2}H and ^{3}He are in local kinetic equilibrium in the Sun, meaning that their creation and destruction rate are equal. Which is a good approximation since the mean lifetime of ^{2}H and ^{3}He in the solar core is of $\sim 10^{-8}$ and $\sim 10^5$ years respectively, to compared to the $\sim 10^{10}$ years long lifetime of a proton that set the Sun evolution time scale.

These assumptions are considered solid in the current framework of our knowledge of the Sun and will be used in the following.

It is worth to notice that the luminosity constraint is not significantly affected by the Sun metallicity. Indeed the SSM predictions of the *pp* and *pep* neutrino fluxes, which include the luminosity constraint in its construction, differ for $\lesssim 1\%$ between the HZ and LZ SSM.

6.3.2 *Measurement of* ^{210}Bi *background rate*

The ^{210}Bi background comes from contaminations in the liquid scintillator left after the purification campaign in fall 2011. The background due to the short-living ^{210}Bi ($\tau = 7.2\,\mathrm{days}$) in Borexino is supported by its long-living parent nucleus ^{210}Pb ($\tau = 32.2\,\mathrm{years}$). The presence of other isotopes from the ^{238}U chain is not substantiated by any evidence: especially, the long-living ^{226}Ra nor ^{222}Rn, easily detectable trough the ^{214}Bi–^{214}Po fast coincidence following their decays, were never observed in the fiducial volume during the entire Phase II.

Bismuth-210 decays into ^{210}Po via β-decay generating events in Borexino with a spectrum very close to the one expected from a CNO signal. The daughter nucleus ^{210}Po terminates the chain by decaying on the stable ^{206}Pb by emitting a 5305 keV α particle with a lifetime of 199.6 days, thus it is possible to summarize the expected chain of decays in the liquid scintillator as

$$^{210}\mathrm{Pb} \xrightarrow[32\,\mathrm{y}]{\beta} {}^{210}\mathrm{Bi} \xrightarrow[7.23\,\mathrm{d}]{\beta} {}^{210}\mathrm{Po} \xrightarrow[199.6\,\mathrm{d}]{\alpha} {}^{206}\mathrm{Pb}\ (\text{stable}) \tag{6.5}$$
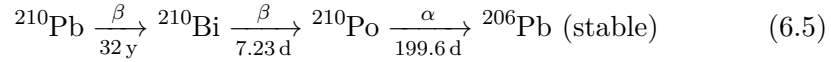
Considering a situation in which only ^{210}Pb is present and no source for either one of the isotopes considered in Eq. (6.5), the evolution of the activity can be computed by solving a system of Bateman equation, whose solution is shown in Fig. 6.10b. Given its very short lifetime ^{210}Bi immediately reach the secular equilibrium with the parent ^{210}Pb, followed by ^{210}Po for its longer lifetime.

The ^{210}Bi and ^{210}Po are thus expected to follow the same time evolution of the parent ^{210}Pb contamination, which decays exponentially with a lifetime of 32.2 years. In principle it is possible to exploit this feature to separate the ^{210}Bi from the ν(CNO) signal, but, as it will be shown in the Sec. 6.6, this would require a long data taking period under very stable conditions. A measurement of the ^{210}Bi background rate can also be achieved by measuring the ^{210}Po daughter rate. This method was first suggested by Villante et al. in [153] and it is particularly interesting because ^{210}Po α decays can be identified on a event–by–event basis exploiting the pulse shape discrimination classifier presented in Sec. 2.6.4.1.

At the end of the purification campaign a large, out of equilibrium contamination of ^{210}Po was present in the scintillator. When the initial rate of ^{210}Po

(a) ^{210}Pb decays scheme

(b) ^{210}Pb chain evolution
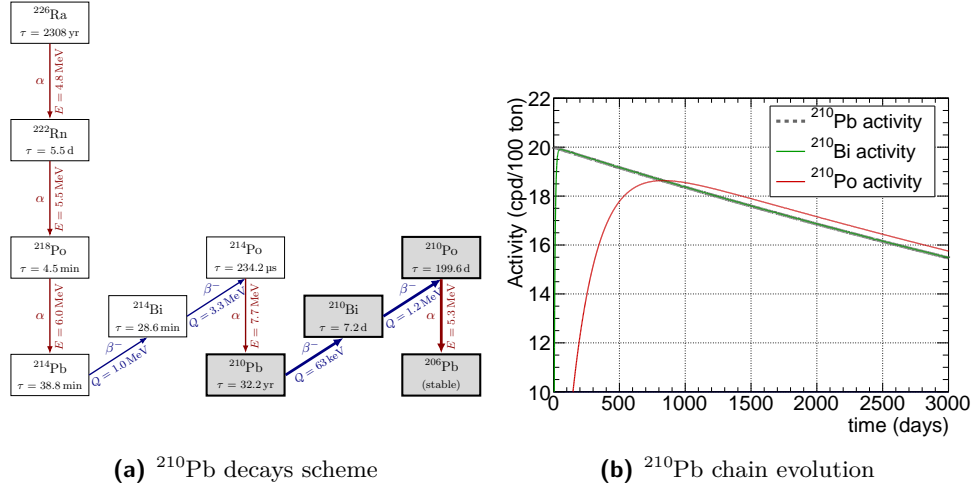
Figure 6.10: (a) Decay chain of ^{238}U following ^{226}Ra. For each isotope the main decay mode is shown, along with the energy of the α particle or the β spectrum end-point. The isotopes for which there are evidence of presence inside Borexino are highlighted using a shaded grey area. **(b)** Expected evolution of the rate of ^{210}Pb, ^{210}Bi and ^{210}Po in case of an initial contamination of ^{210}Pb. The trend of ^{210}Pb is masked by the one of ^{210}Bi that immediately reach the equilibrium with the parent nucleus.

is different from zero and assuming no sources of ^{210}Bi and ^{210}Pb, the time evolution of the ^{210}Po rate can be expressed as

$$R_{\mathrm{Po}}(t) = \left(R_{\mathrm{Po}}^{(u)} - R_{\mathrm{Po}}^{(s)} \right) e^{-t/\tau_{\mathrm{Po}}} + R_{\mathrm{Po}}^{(s)} \tag{6.6}$$

where the $R_{\mathrm{Po}}^{(s)}$ indicates the ^{210}Po rate *supported* by the ^{210}Pb decay chain and in equilibrium with ^{210}Bi, while $R_{\mathrm{Po}}^{(u)}$ stands for the *unsupported* initial contamination. The supported ^{210}Po rate can thus be obtained from a fit of the time evolution of the ^{210}Po rate like the one sketched in Fig. 6.3.2, providing an indirect measurement of the ^{210}Bi event rate.

The accuracy in the determination of the ^{210}Po supported rate depends on the magnitude of the initial out-of-equilibrium contamination, on the mass of the considered fiducial volume and on the observation time. Figure 6.3.2 shows the projected sensitivity on the ^{210}Po supported rate as a function of the fiducial volume mass and of the observation time. This result has been obtained fitting simulated pseudo–dataset that were generated assuming the model in Eq. (6.6), with an unsupported rate of 50 cpd/100 t and a supported rate of 17.5 cpd/100 t. According to this study, an accuracy of the order of 10 % (corresponding to 1.75 cpd/100 t) on the supported rate can be achieved in approximatively 8 months of data taking if the 71.3 t *pep* fiducial volume is used.

This method is considered the most promising strategy to improve the sensitivity to the CNO neutrino signal, but despite appearing quite straightforward, relies on a set of assumptions that will be discussed in detail in Sec. 7.1 and
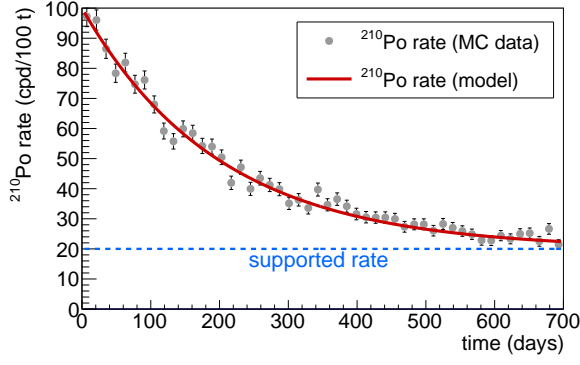
Figure 6.11: Expected trend of the ^{210}Po rate when both ^{210}Po and ^{210}Pb are present at $t = 0$ (Eq. 6.6). The contribution of the ^{210}Pb decay chain to the overall ^{210}Po rate (*supported* rate) is highlighted.
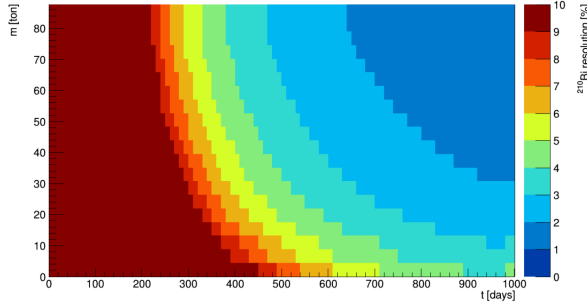
Figure 6.12: Projected sensitivity to the ^{210}Po supported rate as a function of the exposure obtained assuming the exponential model in Eq. (6.6). Courtesy of N. Rossi.

must be verified. As an example, Eq. 6.6 was derived assuming the spatial uniformity of ^{210}Po that in reality depends on the motion of the fluid inside the detector caused by thermal instabilities. The convective motion of the fluid may indeed carry the ^{210}Po emanated by the nylon of the inner vessel towards the detector centre where the fiducial volume is defined, providing continuous contribution of unsupported ^{210}Po to the overall ^{210}Po rate.

The collaboration made a huge effort to stop the fluid motions through a complete thermal stabilization of the detector that is described in the next Chapter. The possibility of a continuous injection of unsupported ^{210}Po into the fiducial volume motivated the development of a method to identify the region of the detector where this effect has the minimum impact. The method is discussed in Sec. 7.5, but it is important to notice that finding regions of the detector where the migration of ^{210}Po has the minimum impact does not mean this effect is not present.

6.4   IMPACT OF PURIFICATION

In this section the impact of a new purification campaign on the sensitivity to a CNO neutrinos is discussed.

A further purification of the liquid scintillator can reduced the content of ^{210}Bi, but the shape of its energy spectrum will still be degenerate with the one due to CNO neutrinos and thus a lower content of ^{210}Bi does not bring an improvement to the sensitivity. A prove of this is visible in Fig. 6.13, that shows the distribution of the ^{210}Bi and CNO rate reconstructed performing a

multivariate analysis on pseudo–datasets when the *pep* neutrino rate has been constrained with a 10% accuracy and the rate of ^{210}Bi is assumed to be as low as 10, 5, 2 and 0 cpd/100 t.

Thus, in order to acquire some sensitivity to the CNO signal, an independent evaluation of the ^{210}Bi background rate is still needed. From an study analogue to the one presented in Sec. 6.2.2.1 it has been found that even with a smaller background the statistical uncertainty of the CNO rate can still be described quite accurately with the rate analysis sketched in Sec. 6.2.1 when the background rate of both ^{210}Bi and *pep* neutrinos is constrained. It is worth to notice that in this framework the accuracy on the CNO depends on the absolute values of the background uncertainties (Eq. 6.4), thus the relative accuracy in the determination of the ^{210}Bi rate can be lower.
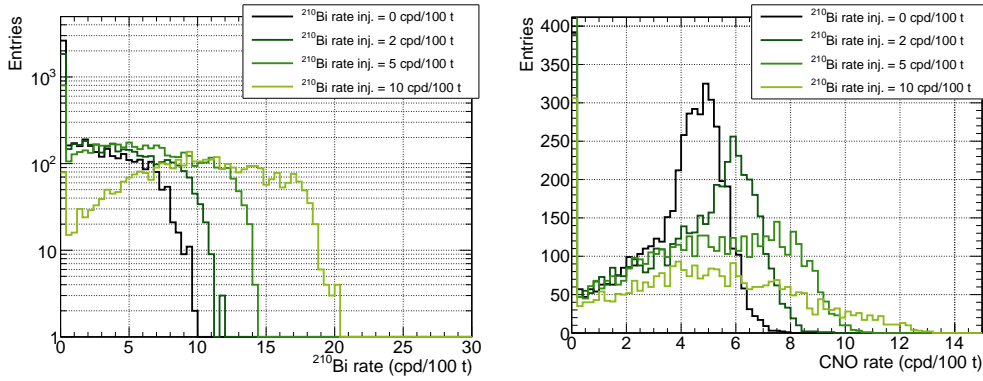
Figure 6.13: Distribution of the ^{210}Bi and CNO best estimates obtained performing a multivariate fit on ensembles of pesudo-dataset. The pesudo-data have been simulated injecting a ^{210}Bi rate of 10, 5, 2 and 0 cpd/100 t.

A smaller contamination of ^{210}Bi without external constraints will not impact the discovery power of the experiment, but can in principle improve the upper limit on the CNO rate. A more stringent limit on the CNO neutrino flux may be important to exclude the HZ SSM prediction and to confirm the most recent observation of photospheric chemical abundances. Figure 6.14 shows the distribution of the 95% C.L. upper limit obtained from the analysis of pseudo-datasets generated assuming the CNO rate as predicted by the LZ SSM and a ^{210}Bi background of 10, 5, and 2 cpd/100 t. Reducing the ^{210}Bi background, the median of the upper limit distribution tends to lower values, passing from 8.4 cpd/100 t for a 10 cpd/100 t ^{210}Bi rate to 5.6 cpd/100 t when ^{210}Bi is reduced to 2 cpd/100 t. The comparison with the rate of CNO neutrinos expected from the HZ SSM shows that the probability of excluding its central value is not very significant.

6.5   IMPACT OF SYSTEMATIC UNCERTAINTIES

The sensitivity studies presented in Sec. 6.2 showed that the bulk of the sensitivity to a CNO signal comes from the simple rate analysis introduced in
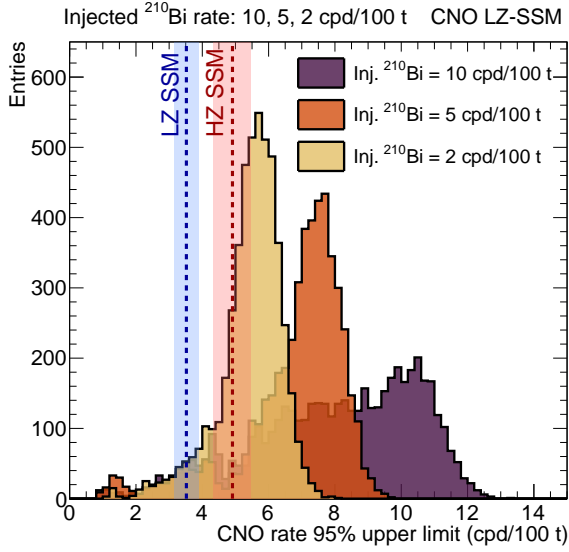
Figure 6.14: Distribution of the CNO 95% C.L. upper limit for different ^{210}Bi scenario obtained from the fit of an ensemble of simulated pseudo-datasets. The rate of *pep* neutrinos has been constrained with a 1% accuracy in the analysis.

Sec. 6.2.1 and the spectral information only plays a mild role when the background rate is known with poor accuracy. This is a severe limitation of the analysis since it means that the data provide only a marginal amount of information and that the measurement is practically determined by the estimation of the background rate.

In this framework, the impact of systematic uncertainties connected with the modelling of the detector response discussed in Sec. 4.4 is expected to be small, but the importance of the shape of the ^{210}Bi β spectrum becomes more relevant. The coefficient $\varepsilon_{^{210}\text{Bi}}$ will be altered by any distortion of the ^{210}Bi spectrum, including those which do not affect the small energy window where most of the sensitivity to the CNO comes from (Fig. 6.1). Indeed, a measurement of the ^{210}Bi background through the ^{210}Po supported rate provides the integral rate of ^{210}Bi, regardless of the energy distribution of the produced electrons. Therefore a distortion of the energy spectrum, even below the detection threshold of the original ^{210}Bi spectrum measurement in [133], will affect the reconstructed number of ^{210}Bi events in the CNO region of interest producing a bias in the CNO reconstructed rate. This effect motivates even more the current efforts for a new, low threshold measurement of the ^{210}Bi spectrum that are pursued by the Dresden and St. Petersburg-PNPI groups of the Borexino collaboration.

Possible systematics affecting the determination of the background rate contribute to the overall error budget in the same way the statistical errors of the estimation of the *pep* and ^{210}Bi rate do. It is then possible to include the systematic uncertainties on the background rate in the total uncertainties $\tilde{\sigma}_{pep}$ and $\tilde{\sigma}_{^{210}\text{Bi}}$ that are used to constrain it.

6.6 TIME ANALYSIS

In this section an alternative method to improve the sensitivity to CNO neutrinos is presented. The method exploits the different time evolution of the ^{210}Bi background and of the CNO neutrino rate to constrain the two components and break their correlation.

From the considerations presented in Sec. 6.3.2, if the ^{210}Pb is uniformly distributed in the detector, the rate of ^{210}Bi is decays exponentially with the lifetime of the parent ^{210}Pb nucleus. On the other hand, also the rate of neutrino components is not constant in time since the eccentricity of the Earth orbit around the Sun gives a sensible seasonal modulation to the measured neutrino flux [116], while the rate of ^{11}C is affected by the temperature variations of the atmosphere [154].

To understand the contribution to the sensitivity of the time information a simplified analysis model has been developed. All the components of the spectrum have been considered constant with the exception of ^{210}Bi that decays with $\tau = \tau_{^{210}\text{Pb}} = 32.2$ years. The evolution of the event rate within a given energy range w is shown in Fig. 6.15a and expressed as

$$
\begin{aligned}
R(t) &= f_{\text{DC}} M \times \sum_i \varepsilon_i(w) R_i(t) \\
&= f_{\text{DC}} M \times \left(\varepsilon_{\text{Bi}}(w) \cdot R_{\text{Bi}}^{(0)} \cdot e^{-t/\tau_{\text{Pb}}} + \sum_{i \neq \text{Bi}} \varepsilon_i(w) R_i \right)
\end{aligned} \tag{6.7}
$$

where the index i runs over the components of the Borexino signal, R_i and $\varepsilon_i(w)$ indicates the corresponding rate and fraction of events in the energy range w, M is the fiducial volume mass and f_{DC} represents the duty cycle of the measurement.

If the considered time interval is divided into two halves each long Δt, it is possible to show that the difference between the number of events recorded in the first (N_1) and in the second half (N_2) is

$$
\Delta N(w) = N_1(w) - N_2(w) = f_{\text{DC}} M \varepsilon_{\text{Bi}}(w) R_{\text{Bi}}^0 \Delta t \left[1 - \exp\left(-\Delta t/\tau_{\text{Pb}} \right) \right] \tag{6.8}
$$

from which one can obtain the rate of Bi as

$$
R_{\text{Bi}} \approx \frac{\Delta N(w)}{M \varepsilon_{\text{Bi}}(w)} \cdot \frac{\tau_{\text{Pb}}}{(\Delta t)^2} \tag{6.9}
$$

where the approximation $\Delta t \ll \tau_{\text{Pb}}$ is used. The corresponding uncertainty is

$$
\sigma_{\text{Bi}} \approx \frac{\sqrt{N_1(w) + N_2(w)}}{f_{\text{DC}} M \cdot \varepsilon_{\text{Bi}}(w)} \frac{\tau_{\text{Pb}}}{(\Delta t)^2} \tag{6.10}
$$

from which it is possible to see that the sensitivity to the ^{210}Bi decay is inversely proportional to the squared of the observation time, to the exposure

per unit time $(f_{DC}M)$ and to the fraction of ^{210}Bi events considered in the energy range under study (ε_{Bi}). Given its peculiar spatial evolution (Sec. 7.2.2), ^{210}Po has not been considered in this study as it is removed supposing a 100% efficient α/β discrimination based on the MLP parameter presented in Sec. 2.6.4.1. Assuming the rates measured in [119, 135] for the neutrino and background components and a plausible rate of 17.5 cpd/100 t for ^{210}Bi it is possible to have a rough estimation of the sensitivity of the time analysis: considering the small energy range between the $\nu(^{7}$Be$)$ shoulder and the ^{11}C peak used in the rate analysis for the sensitivity to CNO neutrinos in Sec. 6.2.1 and shown in Fig. 6.15b, the error on the ^{218}Bi rate obtained from the simplified time analysis is 21.7, 10.6 and 7.5 for 5, 8 and 10 years of data taking respectively. The sensitivity to the time decay strongly depends on the fraction of ^{210}Bi events $\varepsilon_{Bi}(w)$ considered in the analysis: indeed, if one considers an energy interval including the low energy part and extended up to the ^{210}Bi endpoint (340–1340 keV, labelled LE+ in Fig. 6.15b), the uncertainties for 5, 8 and 10 years become respectively 11.1, 5.4 and 3.9 cpd/100 t.

(a) Expecetd count rate evolution

(b) Energy windows for the time analysis
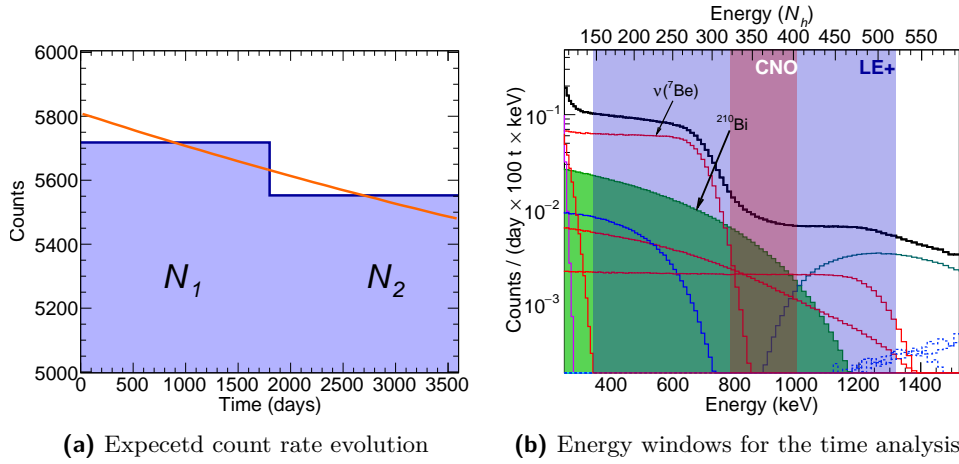
Figure 6.15: (a) Expected counts in the two halves of a 10 years long data taking. The count evolution is also shown. **(b)** Regions of the energy spectrum considered for the sensitivity study of a time analysis. The ^{210}Bi component is highlighted with a green filled area and the ^{210}Po has been removed supposing a 100% efficient α/β pulse shape discrimination. The red shaded area (CNO window) indicates the energy range used in Sec. 6.2.1 for the simple rate analysis, while the blue one (LE+) includes a larger fraction of ^{210}Bi and is therefore more sensitive to the time behaviour.

The time information can thus provide a substantial contribution to the sensitivity to the ^{210}Bi rate if the data taking in stable conditions is longer than 8–10 years. To study in detail the impact of time to the analysis when combined with the spectral information, the time variable has been added to the MV analysis developed in Chap. 3. Figure 6.16 shows the projection of

the PDF[4] in the time-energy space for ^{210}Bi, which decays with $\tau_{210\text{Pb}}$, and for *pep* neutrino, where a seasonal modulation with a 7.1% amplitude [116] has been imposed.

(a) ^{210}Bi   **(b)** *pep* neutrinos
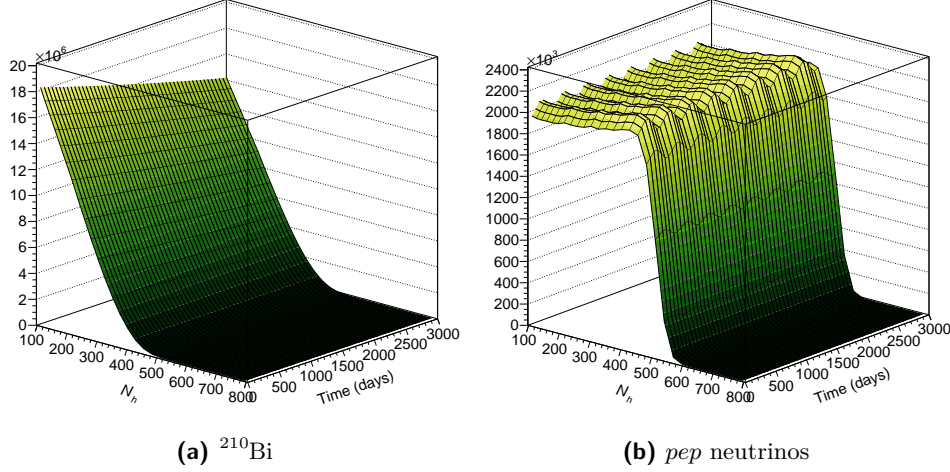
Figure 6.16: Projection of the multivariate PDFs in the energy-time space for ^{210}Bi and *pep* neutrinos (note that the PDFs are shown in linear scale).

The sensitivity to the ^{210}Bi decay has been evaluated in the same way presented in Sec. 4.3, *i.e.* generating thousands of pseudo–datasets injecting the rate measured in Phase II [119, 135] and building the distribution of the best fit estimates. In this study, for each dataset a simultaneous multivariate fit on the TFC-subctracted and tagged datasets has been performed considering the events radial position r^3 in the energy range between 290 and 937 N_h, while the time information is studied in the LE+ energy window (Fig. 6.15b) using 6 months width bins[5]. The improvement brought to the ^{210}Bi sensitivity by the time information is clearly visible in Fig. 6.17: the top panel shows the distribution of the best fit estimates obtained fitting pseudo-datasets generated with an exposure of 5, 8 and 10 years without considering the events time distribution. In the panel the distributions are obtained including the time in the analysis. Including the time behaviour to the analysis lead to an accuracy in the ^{210}Bi as good as 2.7 cpd/100 t for a 10 years long exposure without imposing any constraint on the *pep* rate. Also the sensitivity to CNO

---

4 The PDFs used in this study have been built without considering the change of the detector response in time, and in particular the loss of resolution due to PMTs failures. However, as discussed in Sec. 6.2.2, the spectral information is not expected to affect significantly the sensitivity to CNO neutrinos and therefore these PDFs allows a reasonable approximation of the actual sensitivity.

5 The use of a 6 months binning in the time dimension has the effect to damp the seasonal modulation of ^{11}C and of the neutrino components if not in phase with the calendar. It can be shown that including this information (*i.e.* using shorter time bins) contributes only marginally to the sensitivity to ^{7}Be, while it brings an even smaller improvement to CNO and *pep* neutrinos.

neutrinos is improved, although the residual correlation with the *pep* neutrino rate limits the uncertainty at $3.0\,\mathrm{cpd}/100\,\mathrm{t}$. It is important to stress that this result has been obtained under the optimistic assumption that the detector resolution remains stable over such a long period of time, but on the other hand it is reasonable to expect that the time analysis would not be strongly affected by a loss of resolution.

(a) MV fit: $N_h + r^3$

(b) MV fit: $N_h + r^3 + \text{time}$
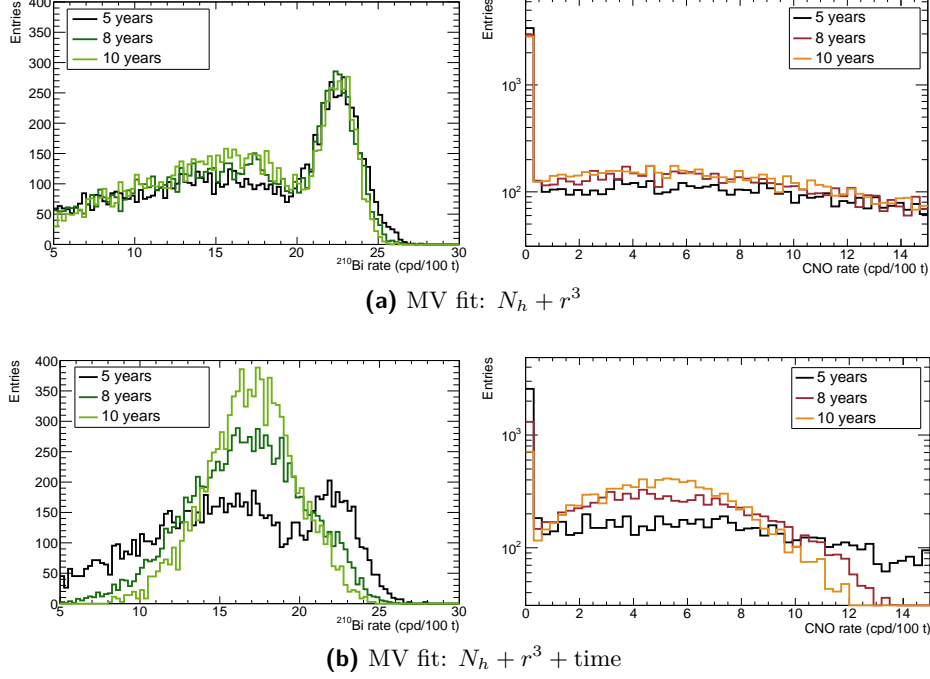
Figure 6.17: Distribution of ^{210}Bi and CNO best fit estimates obtained analysing pesudo-datasets without considering the time information **(a)** and considering it in the fit **(b)**. No external constraints is applied in the analysis.

The accuracy of the ^{210}Bi and CNO rate can be further improved if one constrained the *pep* neutrino rate, which enhance the sensitivity to the spectral shape as shown in Sec. 4.3. In addition, the time analysis can also be optimized. Equation (6.10) indicates that the accuracy of the rate is inversely proportional to the mass of the liquid scintillator considered, thus a larger fiducial volume would improve the sensitivity to the ^{210}Bi decay in time. A similar situation was faced in the search for the seasonal modulation of the ^{7}Be neutrino signal [116], for which a $98.6\,\mathrm{t}$ fiducial volume was defined (see Sec. 2.6.2). A preliminary analysis performed using this extended fiducial volume shows that for a 10 year long data taking, under the assumptions of stable energy resolution, the rate of ^{210}Bi and CNO neutrinos can be measured with a statistical accuracy of 2 and $1.6\,\mathrm{cpd}/100\,\mathrm{t}$ respectively if the *pep* rate is constrained with a 10% precision.

These projections, although relying on some optimistic assumptions, indicates that the events time information can improve the measurement of ^{210}Bi,

but only if the data taking will extend beyond 2022, motivating possible more refined future analyses. With the current dataset the inclusion of time in the multivariate analysis does not bring significant improvement to the ^{210}Bi sensitivity, therefore most of the collaboration efforts are devoted to obtain a measurement of the ^{210}Po rate as an indirect estimation of the ^{210}Bi rate.

6.7   CONCLUSIONS

The search for CNO neutrinos is currently one of the most prominent topic in solar neutrino physics. Borexino, with its unique radiopurity, is the only running experiment with the potential to study CNO neutrinos, but such an analysis is severely affected by the strong correlation induced by the degeneracy between the energy distribution of CNO events and the sum of ^{210}Bi and *pep* contributions.

In order to break the correlations, the *pep* and ^{210}Bi background rate must be independently constrained. A detailed sensitivity study has been performed, showing that if the background is constrained the bulk of the sensitivity to the CNO neutrino signal comes from a simple rate analysis. The multivariate analysis used to obtain the results presented in the previous chapter only brings mild improvement on the sensitivity when the accuracy of the constraints on the background rate is relatively loose. The statistical uncertainty on the CNO rate depends on the uncertainty in the determination of the ^{210}Bi and *pep* background. If the ^{210}Bi and *pep* neutrinos rate are constrained with a 2.45 cpd/100 t and 0.028 cpd/100 t accuracy, in 50% of the cases CNO neutrinos can be detected with a significance higher than 3σ if the CNO neutrino flux is the one predicted by the HZ SSM. The median significance is reduced to 2.1σ for the CNO flux expected by the LZ SSM. The sensitivity does not improve significantly if a new purification campaign further reduce the ^{210}Bi content, since the it will not remove the degeneracy of the events distribution shapes.

It is then crucial to obtain an independent assessment of the background. The rate of *pep* neutrinos can be constrained exploiting its relation with the *pp* neutrino rate using solid argument from nuclear physics, as already done in the analysis presented in the previous Chapter. The uncertainty on the *pep* rate can be reduced to a few per-cent level by imposing an additional constrained based on solar luminosity.

The ^{210}Bi background is supported by the its long-living parent ^{210}Pb ($\tau_{\mathrm{Pb}} = 32.2$ yr). If ^{210}Bi is in equilibrium with ^{210}Pb it decays with the same lifetime of the parent nucleus. It has been shown that it is possible to break the correlation with the CNO rate exploiting the different time evolution of the two components, but it would require a 10 year long under stable conditions and it does not improve the sensitivity with the current dataset. The most promising method to constrain the ^{210}Bi background relies on a measurement of the ^{210}Po daughter, which is expected to be in equilibrium with the ^{210}Pb chain. Polonium-210 events are efficiently tagged using pulse shape discrimination

techniques, but this measurement is complicated by fluid motions inside the detector that affects the its spatial distribution as it will be discussed in the next Chapter.

CHAPTER 7

TOWARDS A DETERMINATION OF THE ^{210}Bi BACKGROUND

The measurement of the ^{210}Bi background rate has a crucial impact on the sensitivity to the CNO neutrino signal. The most promising strategy that is currently pursued by the Collaboration relies on the equilibrium between ^{210}Bi, its parent nucleus ^{210}Pb and the ^{210}Po daughter that allows to extract the ^{210}Bi rate from a measurement of the ^{210}Po which can be tagged with high efficiency thanks to pulse shape discrimination techniques.

In this Chapter the strategy for the measurement of the ^{210}Po rate supported by the ^{210}Pb decay chain is described along with some preliminary results. Section 7.1 recalls the link between the ^{210}Bi and ^{210}Po contaminations, making explicit the underlying assumptions that must be verified. The ^{210}Po distribution inside the detector is discussed in Sec. 7.2, where the role of fluid motion which transport ^{210}Po along the detector is highlighted. The complex dynamic of the ^{210}Po concentration makes it very difficult to obtain a stationary model of the ^{210}Po spatial distribution, but in Sec. 7.3 it is shown that in a volume where both the ^{210}Po concentration gradient and Laplacian are zero the rate of ^{210}Po should be only due to the contribution of the term supported by the ^{210}Pb chain. After a short presentation of a method developed within the collaboration to reconstruct the ^{210}Po density minimum in Sec. 7.4, Sec. 7.5 describes the development of a model-independent method to identify "plateau" region that satisfies the criteria discussed above. Preliminary results of the application of this method on the Borexino data are reported in Sec. 7.6.

7.1 THE METHOD CONCEPTS AND FUNDAMENTAL ASSUMPTIONS

The most promising method to achieve an independent measurement of the ^{210}Bi in Borexino was suggested in [153] and has been briefly introduced in Sec. 6.3.2. It relies on the equilibrium of ^{210}Bi with its parent nucleus ^{210}Pb and its ^{210}Po daughter, that can be reached given the very long lifetime of ^{210}Pb (32.2 years) compared to the one of ^{210}Bi and ^{210}Po (7.2 and 199.6 days respectively). If the equilibrium is reached, the rate of ^{210}Bi is equal to the rate of ^{210}Po, that can be measured with good accuracy exploiting the MLP–based pulse shape discrimination tagging of its α decay into the stable ^{206}Pb.

The rate of the isotopes belonging to the ^{210}Pb chain can be computed by solving a system of coupled differential equations, each one describing the evolution of the activity A_i of the isotope i as

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = -\frac{A_i}{\tau_i} + \frac{A_{i-1}}{\tau_{i-i}} + S_i \tag{7.1}$$

where A_{i-1} indicates the activity of the parent nuclei and S is a generic *source* term which accounts for the possible migration of i nuclei inside the considered fiducial volume.

In order to exploit the link between the ^{210}Pb, ^{210}Bi and ^{210}Po emerging from the system of differential equations in (7.1) and extracting information on the ^{210}Bi rate the following working assumptions are made:

i ^{210}Pb *is out of equilibrium*
The radioactive contaminations in the liquid scintillator does not include isotopes above ^{210}Pb in the ^{238}U decay chain. As already discussed in Sec. 6.3.2, this hypothesis is supported by the fact that fast coincidences due to ^{214}Bi–^{214}Po decays indicating the presence of ^{222}Rn were not observed after the last purification campaign.

ii *No source of* ^{210}Pb
The amount of ^{210}Pb in the liquid scintillator is fixed except from its natural decay and no further release of ^{210}Pb happened. A possible source can be the nylon of the inner vessel, which has a large contamination of ^{210}Pb that could in principle diffuse into the scintillator. This possibility is very unlikely, since the detachment of ^{210}Pb from nylon is not reported in literature. In addition, a constant release of ^{210}Po in the scintillator would cause an increase of the ^{210}Bi concentration that is not observed in the data up to May 2016.

iii *No source of* ^{210}Bi *inside the Fiducial Volume*
Possible mechanisms injecting ^{210}Bi into the fiducial volume must be excluded, leaving only the ^{210}Bi generated by the ^{210}Pb originally present in the liquid scintillator. As for ^{210}Pb, ^{210}Bi could in principle diffuse from the inner vessel nylon into the fiducial volume, but even if this effect, never reported in literature, actually happens, its short lifetime would prevent the detached ^{210}Bi to reach the fiducial volume.

iv *The inner vessel is the only possible source of* ^{210}Po
Contrarily to ^{210}Pb and ^{210}Bi, ^{210}Po can detach from nylon and diffuse through materials, thus the inner vessel provides a constant injection of ^{210}Po into the liquid scintillator.

v *No dust contaminations are present inside the liquid scintillator*
Dirty materials, possibly stuck on the inner vessel nylon, must not diffuse inside the liquid scintillator. This assumption is justified by the fact that

no ^{222}Rn events, indicating some traces of ^{238}U, were never detected after the last purification campaign.

Under these assumptions, it is possible to extract the amount of ^{210}Bi present in the scintillator from a measurement of the evolution of the ^{210}Po activity described by

$$\frac{\mathrm{d}A_{\mathrm{Po}}}{\mathrm{d}t} = -\frac{A_{\mathrm{Po}}}{\tau_{\mathrm{Po}}} + \frac{A_{\mathrm{Bi}}}{\tau_{\mathrm{Bi}}} + S_{\mathrm{Po}} \tag{7.2}$$

The ^{210}Po source term makes the problem non-trivial, but in case this term is zero inside the fiducial volume and only ^{210}Pb pollutes the liquid scintillator at $t = 0$, then the three species of interest reach the equilibrium in ≈ 2 years as shown in Fig. 6.10b. As already mentioned in Sec. 6.3.2, a large contamination of ^{210}Po was present in the scintillator at the beginning of Phase II: if one considers an initial concentration different from zero, Eq. (7.2) can still be solved (assuming a null source term) yielding the evolution of the rate described by Eq. 6.6

$$R_{\mathrm{Po}}(t) = \left(R_{\mathrm{Po}}^{(u)} - R_{\mathrm{Po}}^{(s)} \right) \cdot \mathrm{e}^{-t/\tau_{\mathrm{Po}}} + R_{\mathrm{Po}}^{(s)} \tag{7.3}$$

where the rate evolution is described by two terms: a *supported* term $R_{\mathrm{Po}}^{(s)}$, which is due to the decay of ^{210}Bi in equilibrium with ^{210}Pb, and an *unsupported* contribution $R_{\mathrm{Po}}^{(u)}$ that indicates the initial, out–of–equilibrium contamination.

Therefore, in absence of a source term, the ^{210}Po rate inside a given fiducial volume is expected to decrease exponentially and reach a plateau given by the contribution supported by the ^{210}Pb chain. Figure 7.1 reports the evolution of the ^{210}Po inside the ^{7}Be fiducial volume starting from the beginning of Phase II: after an initial exponential decay the ^{210}Po rate does not stabilize but shows large fluctuations. This behaviour is explained by the emergence of the source term contributions that are discussed in the next section.
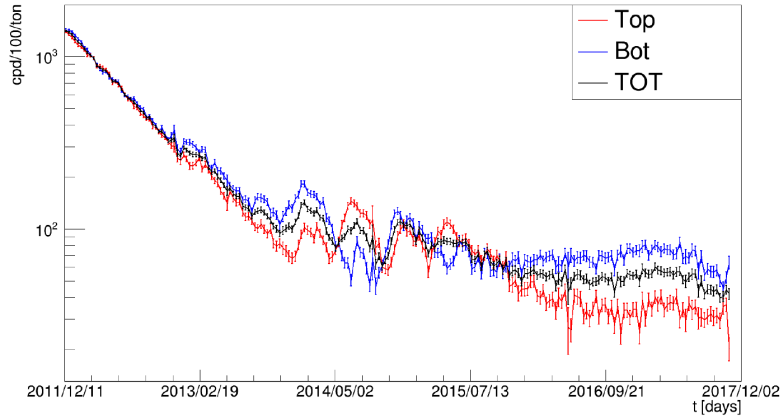
Figure 7.1: Evolution of the ^{210}Po rate inside the ^{7}Be FV ($r < 3.02\,\mathrm{m}$, $|z| < 1.67\,\mathrm{m}$) since the beginning of Phase II. Courtesy of N. Rossi.

7.2   SPATIAL EVOLUTION OF ^{210}Po

The peculiar spatial distribution and evolution of ^{210}Po play a central role in the efforts to obtain some indications on the ^{210}Bi rate. In this section an overview of the ^{210}Po rate spatial dynamic is outlined, starting from the simplest case in which only diffusive motions are considered (Sec. 7.2.1). Section 7.2.2 presents the ^{210}Po data, from which the relevance of the role of the detector fluid dynamics emerges clearly. The collaboration undertook unprecedented efforts in order to understand the fluid behaviour, that is driven by local changes in the detector temperature profile. Section 7.2.3 describes the operations made to damp the temperature effects in the detector, while Sec. 7.2.4 briefly presents the status of the fluid dynamics simulations performed by the Milano-Politecnico and LNGS groups in order to understand in detail the detector fluid dynamics.

7.2.1   *Diffusive case solution*

In this section the ideal case of ^{210}Po detaching from the inner vessel nylon and diffusing inside the liquid scintillator is treated. In this situation the vessel material represents a constant, static source term in Eq. 7.3, but it is not expected to prevent a measurement of the ^{210}Po supported rate. Indeed, if one considers a situation where only ^{210}Po diffusing from the surface of a sphere is present, the Flick's diffusion equation reads

$$\frac{\partial n_{\text{Po}}(\boldsymbol{x}, t)}{\partial t} = D\nabla^2 n_{\text{Po}}(\boldsymbol{x}, t) - \frac{n_{\text{Po}}(\boldsymbol{x}, t)}{\tau_{\text{Po}}} \tag{7.4}$$

with n_{Po} indicating the ^{210}Po spatial density. It can be shown that Eq. 7.4 has solution

$$n_{\text{Po}}(\boldsymbol{x}, t) = n_{\text{Po}}^{(0)}\frac{\sinh(r/\lambda)}{r/\lambda}e^{-t/\tau_{\text{Po}}} \tag{7.5}$$

where $n_{\text{Po}}^{(0)}$ is the density of ^{210}Po on the inner vessel surface and λ is the diffusion length, defined as $\sqrt{D\tau_{\text{Po}}}$ with D indicating the diffusion coefficient of ^{210}Po. The diffusion length represents the characteristic scale interested by the ^{210}Po diffusion: since the diffusion coefficient for ^{210}Po in hydrocarbon –which has a structure similar to the Borexino scintillator– is $\approx 2 \times 10^{-5}\,\text{cm}^{-2}\text{s}^{-1}$, the diffusion length results to be $\approx 20\,\text{cm}$, thus a fiducial volume cut selecting the innermost part of the detector can be enough to remove the ^{210}Po diffusing from the inner vessel nylon.

7.2.2   *Temperature instabilities effect*

In the aftermath of the purification campaign that took place in 2010-2011, the ^{210}Po concentration was ≈ 800 times larger than the one of ^{210}Bi. As discussed in the previous sections, the ^{210}Po was expected to decrease exponentially and

eventually reach a plateau throughout the detector, with the exception of the region closer to the vessel that is interested by the diffusion from the nylon.

Events due to ^{210}Po decay can be effectively tagged on a event by event basis exploiting the MLP pulse shape parameter introduced in Sec. 2.6.4.1. ^{210}Po events has been selected from the data by considering the candidates having MLP < 0.05 and with an energy compatible with the ^{210}Po peak.

Figure 7.2 shows the evolution of the ^{210}Po rate within a 2.75 m radius sphere divided in 31 iso-volumetric layers for the first three years and a half of Borexino Phase II: after an initial decrease seemingly uniform in the detector, a striking "seasonal" migration of ^{210}Po appears.
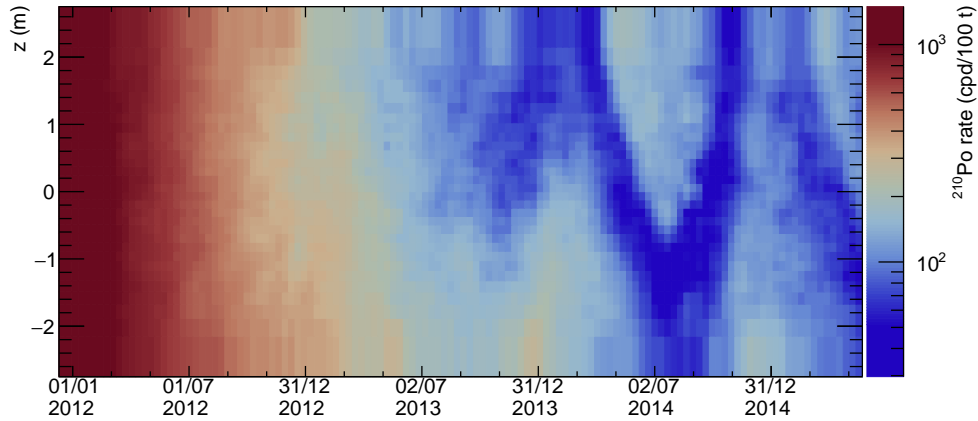
Figure 7.2: Evolution of the ^{210}Po rate within 31 iso-volumetric layers of a 2.75 m radius sphere for the first 1250 days of Borexino Phase II. The histogram has been smoothed to make the seasonal migration effect more clear.

These effects were explained as due to fluid motion caused by temperature instabilities inside the detector. In fact, the motion of the fluid can bring the ^{210}Po from the regions close to the inner vessel where it is emanated towards the detector centre. In order to stop the ^{210}Po migration inside the innermost part of the detector, it was crucial to understand the thermal behaviour of Borexino to stabilize the detector.

7.2.3 *Thermal stabilization of the detector and impact on the ^{210}Po migration*

Before proceeding with any action aiming to stabilize the detector temperature profile, the temperature gradient needed to be studied. The Borexino temperature profile develops in the vertical direction, having its minimum at the bottom –where it is in contact with the basement– and growing monotonically with the height. Temperature variations in the experimental hall due to seasonal effects or to other experimental activities can significantly alter the temperature gradient and activate fluid motion responsible for the ^{210}Po migration.

To carefully monitor possible instabilities in the temperature profile the Latitudinal Temperature Probe System (LTPS) was installed [155, 156]: consisting of 54 probes placed both inside and outside the stainless steel sphere as sketched in Fig. 7.3, it provides a vertical profile monitoring with an azimuthal resolution of 180° that is fundamental to understand the thermal behaviour of the detector.
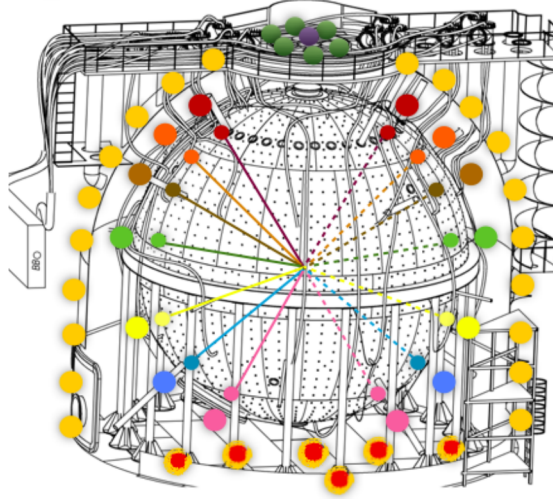
Figure 7.3: Sketch of the Latitudinal Temperature Probe System positioning. Picture from [155].

The effect of the environmental temperature variations emerges clearly from Fig. 7.4 where the trends of the temperature recorder by 14 probes located in re-entrant tubes extended into the detector buffer are reported. In the first period of operation of the temperature monitor large fluctuations affect the top part of the detector, while in the bottom the effect is attenuated.

The temperature profile changes significantly when the activities for the detector stabilization began. The first step, occurred in May 2015, consisted in insulating the detector by wrapping the entire surface of the Water Tank with a 20 cm double layer of mineral wool. Shortly after, the loop continuously changing fluxing water in the external tank was stopped, allowing the bottom of the detector to cool down through heat exchange with the basement. As a further contribution to the temperature stability, the Active Gradient Stabilization System or Temperature Active Control System[1] (TACS) was installed. This system consists of twelve independent water loop circuits operating at controlled temperature, was mounted on the uppermost part of the Water Tank that is the most affected by environmental temperature changes. Between November 2018 and January 2019, the TACS was upgraded with the installation of two additional pipe rings covering part of the detector at smaller latitude.

---

1 In this thesis the active control system will be referred to as TACS and not as Active Gradient Stabilization System since its acronym AGSS is also used to indicate the set of spectroscopic observations of the Sun photosphere that is used as an input in the LZ SSM.
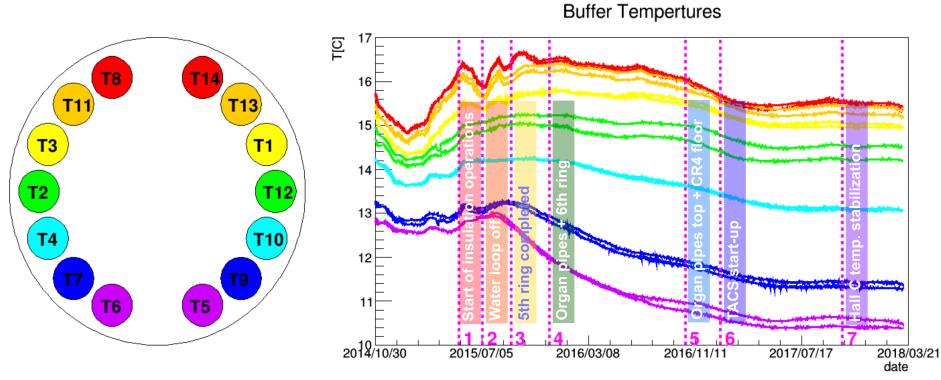
Figure 7.4: Evolution of the temperature inside the detector outer buffer recorded by the probe system. The quota of the different probes is sketched on the left panel.

The temperature stabilization visible in Fig. 7.4 has a clear impact on the ^{210}Po migration. Figure 7.5 shows the evolution of the ^{210}Po rate inside in the same volume of Fig. 7.2 but using the data acquired after the insulation operations. Soon after the beginning of the insulation operations and the water loop stop, the macroscopic seasonal effect disappeared and ^{210}Po became more uniform throughout the detector.
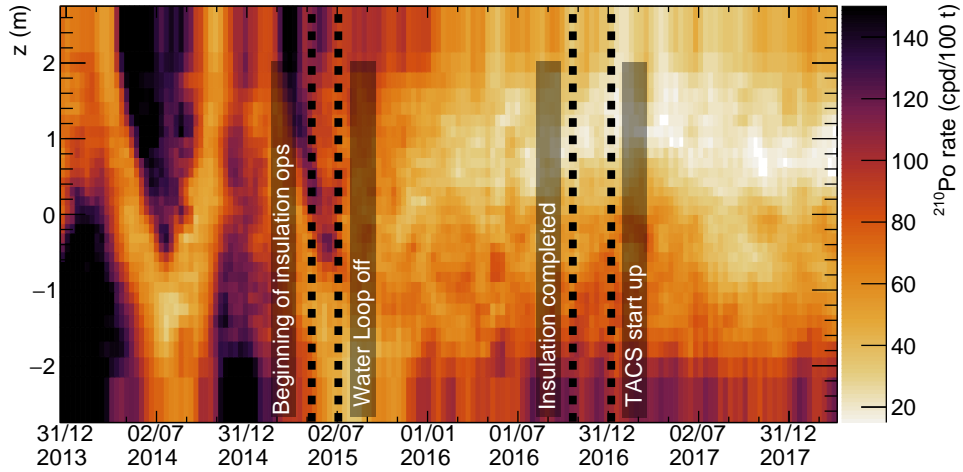
Figure 7.5: Evolution of the ^{210}Po spatial distribution during the thermal stabilization of the detector. Polonium-210 events are selected in 31 iso-volumetric layers of a 2.75 m sphere as done in Fig. 7.2. Immediately after the beginning of the insulation and the stop of the water loop the seasonal motions still evident in 2014 are damped.

The stabilization of the detector accomplished its primary goal to stop the macroscopic migration of ^{210}Po caused by the seasonal variation of the experimental hall temperature. However it is premature to say that it is enough to stop the fluid motion in the detector. As one can see from Fig. 7.5, the bottom part of the detector shows a larger ^{210}Po higher than the top part and

the "cleanest" region remains confined in the top hemisphere. These effects may be due to residual fluid motion that require a numerical simulation of the complex detector fluid dynamics to be understood.

7.2.4 *Fluid dynamics simulations*

The LTPS temperature monitor system was crucial to have an on-line measurement of the detector temperature profile, but in order to have an in-depth understanding of the detector behaviour its fluid dynamic must be studied.

Borexino fluid dynamics studies are performed thanks to the joint efforts of the Milano-Politecnico, Virginia Tech and LNGS groups and provide very useful information about the detector. The Computational Fluid Dynamic (CDF) simulation is indeed crucial to understand the temperature stratification in the detector and even more to ensure the optimal fluid stability through the TACS operations.

A 2 and 3D purely conductive models of the detector were implemented into CDF simulations that were use to study the heat exchange processes among the different parts of the detector, the environment and the basement. The detector cooling that followed the termination of the water loop and the consequent temperature stratification were predicted [155]. This result has been important to show that some initial concerns about an overall cooling of the detector due to the coupling to the cold heat sink of the basement that would have cancelled the temperature gradient between the top and bottom part of the detector were not motivated.

The CDF simulations were also used to study the effect of convective motions due to temperature instabilities. The complexity of the convectional simulation required to used a 2D model of the detector in place of a full 3D description, but still gives very interesting insights. In particular, from CDF simulations it emerges that when imposing the temperature profile measured by the LTPS, fluid currents in the vertical direction are not expected but horizontal fluid motion caused by small asymmetries in the temperature boundary conditions can occur. These motions are predicted to be very slow ($\sim 1 \times 10^{-7}$ m/s) but still they can play a significant role in carrying ^{210}Po from the active volume periphery towards the innermost part of the detector. In fact, recent preliminary results show that when considering the temperature profile recorded in February 2017, the horizontal velocity has a minimum ≈ 1 m above the equatorial plane of the detector, where the temperature gradient has its minimum. The fluid motion were used to test a simplified ^{210}Po transport model in which ^{210}Po and ^{210}Bi are emanated from the inner vessel. The predicted ^{210}Po rate is shown in Fig. 7.6 for an ideally spherical vessel and the actual vessel shape. In both cases the ^{210}Po concentration has a minimum in correspondence of the minimum of the horizontal velocity.

Despite the good qualitative agreement of the results in Fig. 7.6 with the data shown in Fig. 7.5, CDF simulations are not accurate enough to describe the ^{210}Po spatial distribution yet.
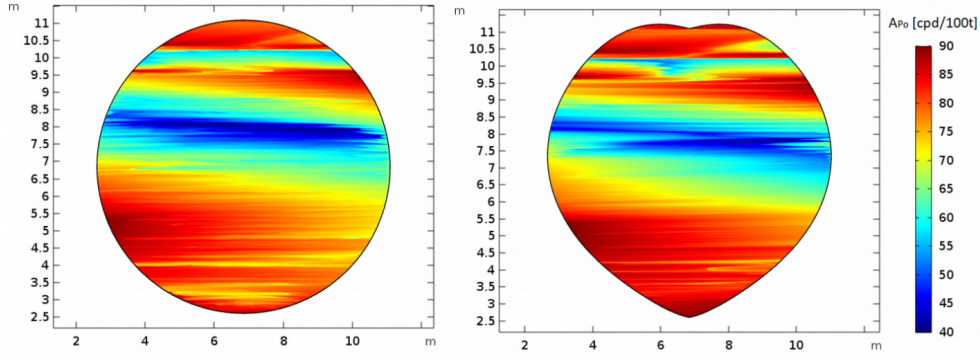
Figure 7.6: Spatial distribution of ^{210}Po obtained implementing a simple transport model in the CDF simulations of the detector assuming an ideal spherical vessel (*left panel*) and the actual vessel shape (*right panel*). The ^{210}Po rate has its minimum $\approx 1\,\text{m}$ above the equator, where the horizontal velocity of the fluid is close to zero. Courtesy of Valentino Di Marcello.

7.3   DETERMINATION OF THE ^{210}PO SUPPORTED RATE FROM A CLEAN VOLUME SELECTION

The complex ^{210}Po dynamic described in the previous section complicates significantly the possibility to achieve a measurement of the ^{210}Bi contamination through the supported ^{210}Po rate. In fact, the effects of fluid motion in the ^{210}Po spatial distribution is hard to model analytically and the available numerical simulations, although in qualitative agreement with the observed data, are not accurate enough to build a reliable model of the ^{210}Po density throughout the detector.

A possible solution to this problem which does not require any assumption on the actual ^{210}Po distribution has been suggested by Ding XF., F. Villante and N. Rossi [93]. Considering only the innermost part of the detector, it is possible to write a continuity equation for the concentration of ^{210}Po $n_{\text{Po}}(\boldsymbol{x}, t)$ and ^{210}Bi $n_{\text{Bi}}(\boldsymbol{x}, t)$ as

$$\frac{\partial n_{\text{Bi}}}{\partial t} = \frac{n_{\text{Pb}}}{\tau_{\text{Pb}}} - \frac{n_{\text{Bi}}}{\tau_{\text{Bi}}} + \boldsymbol{\nabla} \cdot [D_{\text{Bi}} \boldsymbol{\nabla} n_{\text{Bi}} - \boldsymbol{v} n_{\text{Bi}}] \tag{7.6}$$

$$\frac{\partial n_{\text{Po}}}{\partial t} = \frac{n_{\text{Bi}}}{\tau_{\text{Bi}}} - \frac{n_{\text{Po}}}{\tau_{\text{Po}}} + \boldsymbol{\nabla} \cdot [D_{\text{Po}} \boldsymbol{\nabla} n_{\text{Po}} - \boldsymbol{v} n_{\text{Po}}] \tag{7.7}$$

where D_i indicates the diffusion coefficient and $\boldsymbol{v}(\boldsymbol{x}, t)$ is the fluid velocity field. The concentration of ^{210}Po is not known a priori, but it is expected to exhibit a local minimum at time t in $\boldsymbol{x}_0(t)$. If one tracks down the evolution of n_{Po} in its minimum $\boldsymbol{x}_0$, one finds that

$$\frac{\mathrm{d} n_{\text{Po}}}{\mathrm{d} t}(\boldsymbol{x}_0, t) = \left[(\boldsymbol{\nabla} n_{\text{Po}}) \cdot \boldsymbol{v} + \frac{\partial n_{\text{Po}}}{\partial t} \right]_{\boldsymbol{x} = \boldsymbol{x}_0} = \frac{\partial n_{\text{Po}}}{\partial t}(\boldsymbol{x}_0, t) \tag{7.8}$$

thus

$$\frac{\partial n_{\text{Po}}}{\partial t}(\boldsymbol{x}_0, t) = \left\{ \frac{n_{\text{Bi}}}{\tau_{\text{Bi}}} - \frac{n_{\text{Po}}}{\tau_{\text{Po}}} + \boldsymbol{\nabla} \cdot [D_{\text{Po}}(\boldsymbol{\nabla} n_{\text{Po}}) - \boldsymbol{v} n_{\text{Po}}] \right\}_{\boldsymbol{x}=\boldsymbol{x}_0} \tag{7.9}$$

$$= \left\{ \frac{n_{\text{Bi}}}{\tau_{\text{Bi}}} - \frac{n_{\text{Po}}}{\tau_{\text{Po}}} + D_{\text{Po}} \nabla^2 n_{\text{Po}} - (\boldsymbol{\nabla} \cdot \boldsymbol{v}) n_{\text{Po}} - \boldsymbol{v} \cdot (\boldsymbol{\nabla} n_{\text{Po}}) \right\}_{\boldsymbol{x}=\boldsymbol{x}_0}$$
$$\tag{7.10}$$

Since the n_{Po} gradient is zero in $\boldsymbol{x}_0$ by definition and $\boldsymbol{\nabla} \cdot \boldsymbol{v}$ is also null for fluid incompressibility, the last two terms of Eq. (7.10) vanish, leaving with

$$\frac{\partial n_{\text{Po}}}{\partial t}(\boldsymbol{x}_0, t) = \left[\frac{n_{\text{Bi}}}{\tau_{\text{Bi}}} - \frac{n_{\text{Po}}}{\tau_{\text{Po}}} + D_{\text{Po}} \nabla^2 n_{\text{Po}} \right]_{\boldsymbol{x}=\boldsymbol{x}_0} \tag{7.11}$$

Equation (7.11) indicates that at the minimum of the ^{210}Po density fluid motion does not contribute to the ^{210}Po concentration and only diffusion brings additional ^{210}Po to the ^{210}Pb supported term. The effect of diffusion is cancelled when the Laplacian of the ^{210}Po density is zero.

These results motivated the development of a series of methods to trace the evolution of the ^{210}Po density around its minimum. The *unbiased minimum finder* (UMF), briefly introduced in the next section, is a method developed within the Collaboration which finds the minimum of a generic distribution. As mentioned above, the condition of minimum density does not exclude the contribution of ^{210}Po diffusion. However the contribution of diffusion is always positive, thus it is possible to use the results of the UMF as an upper limit on the ^{210}Po supported rate even without quantifying the impact of diffusion. In Sec. 7.5 a new method is presented. This method, called Polonium Plateau Finder, looks for regions of the detector where the impact of both fluid motion and diffusion on the ^{210}Po concentration is minimum. Assuming the ^{210}Bi to be uniformly distributed inside the detector, this situation corresponds to the simple relationship derived in Eq. 7.2 where the source term can be neglected.

7.4   THE UNBIASED MINIMUM FINDER

The *unbiased minimum finder* is a class of methods developed within the Collaboration to search for the minimum of the ^{210}Po distribution. This method is based on a binned weighted likelihood approach. The evolution of the minimum obtained applying the method on iso-volumetric layers of a 2.75 m radius sphere is shown in Fig. 7.7. The continuity of the minimum path suggest that it is following a physical motion of a clean "bubble" which seems to move towards the detector centre in the last year.

The trend of the ^{210}Po rate associated to this volume is shown in Fig. 7.8 and appears to be consistent with the expected trend of the ^{210}Po rate (Eq. 7.3).
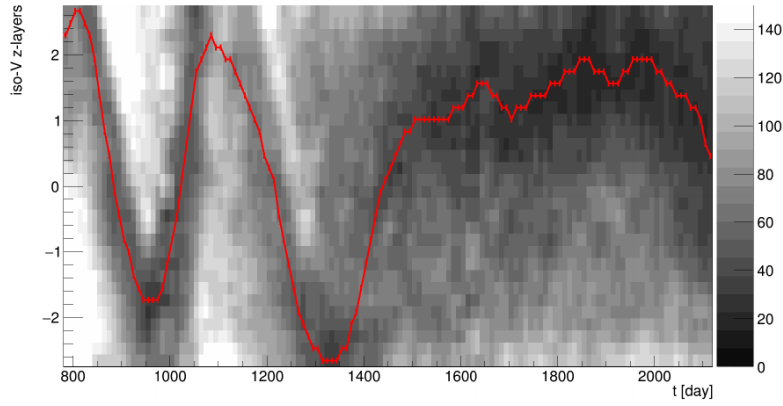
Figure 7.7: Position of the minima returned by the Unbiased Minimum Finder during Phase II insude a 2.5 m radius sphere divided along the vertical axis into 30 iso-volumetric layers. The spatial evolution of the rate is smeared for a better graphical rendering. Courtesy of N. Rossi.
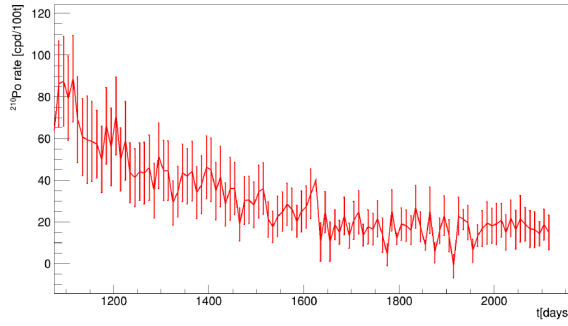
Figure 7.8: Evolution of the ^{210}Po rate at the minima returned by the Unbiased Minimum Finder.

7.5 DEVELOPMENT OF A MODEL INDEPENDENT PLATEAU FINDER

The complex ^{210}Po dynamic driven by both diffusion and convective motions in the fluid makes the measurement of the ^{210}Po supported rate not trivial. In Sec. 7.3 it has been shown that where both the gradient and the Laplacian of the ^{210}Po concentration are zero, that region is free from the effects of diffusion and convection and only the ^{210}Po supported by the ^{210}Pb decay chain is present there. These conditions are met in regions with a constant minimum of the ^{210}Po density[2].

Identifying an evolving, "flat" minimum in the data is a non trivial problem. The lack of a model for the ^{210}Po density does not allow to make any assumption on the underlying distribution of the data, and the very low count rate expected (a rate of 17.5 cpd/100 t corresponds to only ≈ 6 events per m^3 per month) makes the impact of statistical fluctuations too strong to be treated with a standard "binned" analysis.

---

2 In such regions all the terms of the Hessian matrix of the ^{210}Po density are zero, that is a sufficient condition for the Laplacian of the density to be zero too.

This section presents the development of a method, also known as *Plateau Finder*, based on Kernel Density Estimator (KDE) to identify regions of the detector compatible with a flat minimum. Section 7.5.1 summarizes the main concepts of kernel density estimators while Sec. 7.5.2 describes the strategy for the definition of flat-compatible regions at the minimum of the distribution. The results obtained testing the method on simulated datasets are discussed in Sec. 7.5.3.

7.5.1 *Introduction to Kernel Density Estimators*

When studying the distribution of the ^{210}Po events to reconstruct its minimum it is not possible to assume its shape from some model, therefore only the data can be used to extract the real ^{210}Po density field. This problem can be approached without making any assumption on the underlying distribution with non-parametric Density Estimators (DE) that are briefly introduced in this section following [157, 158].

Let's treat a simple, uni-dimensional case first. Given a set of data $\{x_n\}; n = 0, 1, \ldots N$ extracted from the unknown distribution $f(x)$, it is possible to estimate the density profile simply filling a histogram. Histogram is indeed the most common non-parametric density estimators: it can be defined evaluating for each event x_i an indicator function $I(x; w)$ that is 1 when x lies in the bin interval $(-w/2, +w/2)$ and 0 outside. The histogram is thus defined as the sum of all the events contributions, that is

$$h(x) = \sum_{i=0}^{N} I(x - \tilde{x}_i; w) \tag{7.12}$$

where $\tilde{x}_i$ is the bin centre in which the event x_i falls.

Histograms have some known drawbacks that can be particularly important when the statistics in the data is limited, namely: (*i*) histograms are discontinuous even if the PDF is continuous, (*ii*) they depend on the bin size and origin, (*iii*) the information from the location of a datum within the bin is lost.

It is possible to generalize the concept of histograms by replacing the indicator $I(x; w)$ with a generic, normalized, smooth *kernel function* $k(x; w)$ of the form

$$k(x - x_n; w) = \frac{1}{w} K\left(\frac{x - x_n}{w}\right) \tag{7.13}$$

for which the corresponding density estimator $\hat{f}_0(x)$ is

$$\hat{f}_0(x) = \frac{1}{Nw} \sum_{n=1}^{N} K\left(\frac{x - x_n}{w}\right) \tag{7.14}$$

which is smooth by construction.

Conceptually this approach makes each event spread over a region of order w. Note that the particular form of the kernel usually does not impact significantly on the estimator, but the most popular choice is a Gaussian kernel. The determination of (Gaussian) kernel "width" w is somehow arbitrary. If the bandwidth w is constant for all i, this kind of estimators are often referred to as *fixed kernel estimate*. It can be shown that in the limit of a large amount N of normally distributed data, an optimal choice of the bandwidth w^* can be obtained minimizing the Mean Integrated Squared Error (MISE) of the estimator, resulting in

$$w^* = \left(\frac{4}{3}\right)^{\frac{1}{5}} \sigma N^{-\frac{1}{5}} \tag{7.15}$$

where σ is the data standard deviation. One can see that the optimal smoothing parameter decreases increasing the sample size N.

However normally distributed data are just a particular case and in general w^* is not known. In order for these methods to be more flexible and treat a wider class of distributions *adaptive kernel estimates* were introduced. The only difference with respect to fixed kernel estimates is that the bandwidth w is not fixed but depends locally on data (w is narrower where the density is higher, wider when data are sparse). Following the relative uncertainty of Poisson statistics, one might assume that in the region containing the event x_n the smoothing parameter should be

$$w(x_n) = \frac{w^*}{\sqrt{f(x_n)}} \tag{7.16}$$

that leave the problem to set w^* without knowing the actual unknown distribution $f(x)$.

It can be shown with a dimensional analysis that $w^* \propto \sqrt{\sigma}$, where in place of $f(x)$ one can use the fixed kernel estimator $\hat{f}_0(x)$ as a good approximation. Then, using the MISE-optimized choice of w^* obtained for the fixed kernel case (Eq. 7.15), the adaptive kernel estimator $\hat{f}_1(x)$ can be written as

$$\hat{f}_1(x) = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{w_n} K\left(\frac{x - x_n}{w_n}\right) \tag{7.17}$$

with

$$w_n = w(x_n) = \left(\frac{4}{3}\right)^{\frac{1}{5}} \rho \sqrt{\frac{\sigma}{\hat{f}_0(x_n)}} N^{-\frac{1}{5}} \tag{7.18}$$

The ρ parameter is typically set to unity and should be adjusted only in extreme situation when the local structure of the data σ_{loc} is two orders of magnitude smaller than the standard deviation of the data.

The distribution of ^{210}Po events is three-dimensional, and in principle adaptive KDE can be implemented in more than one dimension applying a non linear transformation to the data in order to have a diagonal covariance ma-

trix [157]. Once data are properly transformed, the KDE can be expressed as the product of univariate kernels with independent bandwidth. Considering more dimensions comes with the price of enhancing the sparseness of the data which is critical for a low rate analysis like the one of the ^{210}Po distribution. For this reason in this thesis the ^{210}Po concentration is studied in one dimension only as a proof of concept that can serve for future studies which may include more dimensions.

7.5.2 *Method for the definition of a flat region*

In order to identify a flat region in a smooth density profile it is possible to exploit the distribution derivative which is zero where the distribution is constant.

In a realistic scenario, even if events are sampled from a constant distribution, statistical fluctuations in the data induce local changes in the derivative, as it is shown in Fig. 7.9a where a toy distribution was generated sampling uniformly events injecting an average density of 48 counts/m in a 4 m interval. The red line indicates the density estimator obtained with an adaptive kernel, which smooths the statistical fluctuations of the data. The lower panel shows the density estimator derivative. The magnitude of the fluctuations in the derivative is evaluated by randomly sampling its value in 10^3 points for 100 different datasets where the same count density was injected. The distribution of the derivative values obtained for 48 counts/m injected density is shown in Fig. 7.9b and can be used as a reference to estimate how likely is to observe a given value of the derivative when events are extracted from a flat distribution with this count density.

The amplitude of the derivative fluctuations clearly depends on the average injected density: scanning different values of the average events density the standard deviation of the derivative values distribution follows a linear trend that is reported in Fig. 7.10.

Consider now a situation close to the one expected from Borexino data. Looking at the latest data shown in Fig. 7.5 it is possible to expect a plateau in the central part of the detector enclosed between a strong increase of the ^{210}Po rate towards the detector top and bottom extremes like the one represented in Fig. 7.11.

The red line in Fig. 7.11 represent the model assumed to build a toy dataset generated considering an exposure of 25 days and a realistic event density in the plateau region. The density estimator is evaluated using the adaptive kernel method and is shown with a blue line.

In Sec. 7.3 it has been demonstrated that in order to find a region of the detector where the amount of out of equilibrium ^{210}Po injected by diffusion or carried by fluid motion is minimum, a flat minimum of the ^{210}Po spatial distribution should be considered. Starting from the position of the density estimator minimum it is possible to define a region compatible with a flat distribution "expanding" the selected volume in the lower and upper direction

(a)   **(b)**
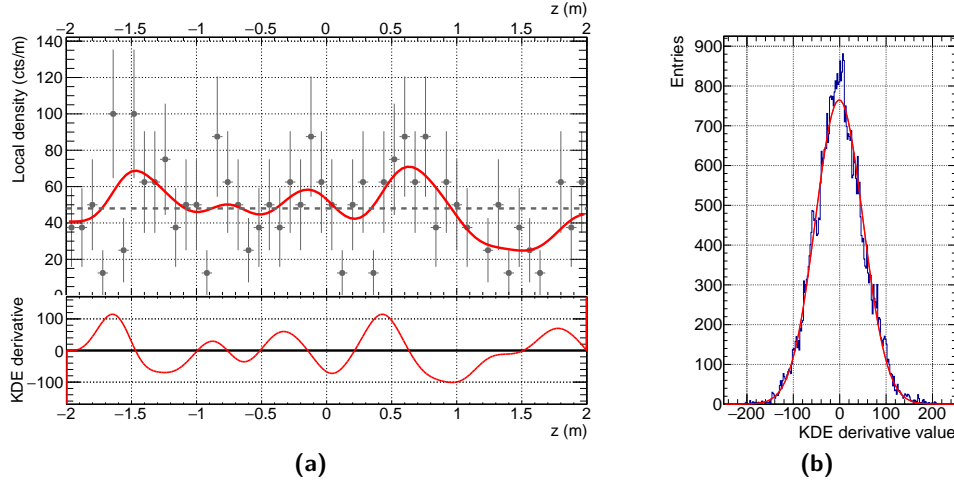
Figure 7.9: (a) Example of flat distribution with average count density of $48\,\mathrm{counts/m}$. The returned density estimator is shown with a red line and its derivative is reported in the bottom panel. **(b)** Distribution of the derivative values obtained from a random sampling of 10^3 points for several different uniform datasets.
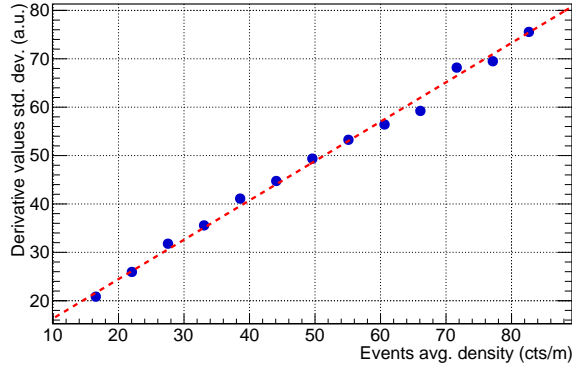
Figure 7.10: Standard deviation of the derivative value distribution as a function of the average event density obtained repeating the procedure shown for Fig. 7.9 for different injected density.

until the density estimator derivative exceeds a certain threshold. In order to set the derivative value threshold and to define the plateau region the following procedure is adopted:

1. The event density $\rho_{\min}$ at the local minimum $z_{\min}$ is evaluated according to the density estimator $\rho_{\min} = \hat{f}(z_{\min})$

2. The expected standard deviation of the derivative values distribution $\sigma(\rho)$ is estimated from a linear interpolation of the trend shown in Fig. 7.10 and a preliminary threshold t_0 is set at $1.7\sigma(\rho_0)$.

3. A temporary plateau is defined expanding expanding a volume from $z_{\min}$ in both directions until the derivative reaches the threshold t_0, as shown in Fig. 7.11 where the selected region is highlighted in orange.

4. The average density in the temporary plateau ρ_{tmp} is computed from the integral of the density estimator in the selected region and it is used to define an updated threshold $t_1 = n\sigma(\rho_{\mathrm{tmp}})$ where n is a parameter that should be optimized.

5. Expand the volume from z_{min} to the point in which the density estimator derivative exceeds the threshold t_1. This region, shown in light blue in the top panel of Fig. 7.11, can be considered as consistent with a flat distribution.
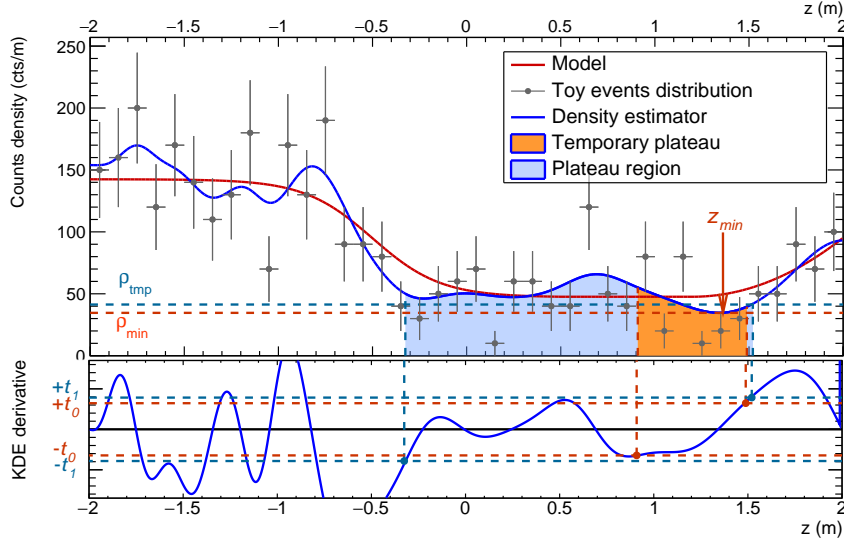
Figure 7.11: Sketch of the procedure for the definition of the plateau region. The top panel shows a distribution of events obtained sampling a given model and the density estimator returned by an adaptive kernel method. The KDE derivative is shown in the bottom panel with the different threshold used to define the plateau region. See text for explanation.

This procedure allows to define a plateau region for a generic distribution without making any assumption on its particular shape nor on the event density at its minimum. The iteration in the definition of the derivative threshold in the procedure just described is introduced to resolve those cases in which the minimum of the KDE falls into a dip with steep walls that keep the reconstructed plateau confined even if it is induced by statistical fluctuations. Evaluating a temporary plateau with an average density ρ_{tmp} (larger than ρ_{min} by definition), allows to set a larger threshold on the derivative value that can help overcoming the barriers due to statistical fluctuations in the data.

7.5.3 *Test and optimization*

The definition of the plateau region introduced in the previous section has been tested and optimized on simulated data. The test model assumed for

this study has been derived from a qualitative comparison with the latest ^{210}Po data and is described with other test assumptions in Sec. 7.5.3.1. The density estimator algorithm used in these tests as well as on the real data is described in Sec. 7.5.3.2, while Sec. 7.5.3.3 presents the main results of this study.

7.5.3.1 ^{210}Po *test density model*

As discussed at the end of Sec. 7.5.1, the search for a constant density region compatible with the minimum of the ^{210}Po concentration in principle should be carried out in three dimensions but the events sparseness issues induced by the increased dimensionality of the problem suggested to make a first attempt using only one dimension. Since the ^{210}Po data shows a peculiar asymmetry between the top and bottom part of the detector the events distribution along the vertical coordinate z has been considered. In order to select a spherical fiducial volume as it is suited for the detector geometry, a transformed vertical coordinate is defined to account for changing volume of the spherical segment enclosed between z and $z + \mathrm{d}z$ along the vertical direction. If one considers a sphere with radius R, the small volume of a spherical layer parallel to the equatorial plane and with distance z from the centre is

$$\mathrm{d}V = \pi(R^2 - z^2)\mathrm{d}z \tag{7.19}$$

and a variable that scales proportionally to the volume like

$$\zeta(z) = \int \pi(R^2 - z^2)\mathrm{d}z = \pi\left(R^2 z - \tfrac{1}{3}z^3\right) \tag{7.20}$$

will be distributed uniformly for events selected inside the sphere. In order to make the range of ζ independent from the particular choice of R, ζ has been normalized imposing $\zeta(\pm R) = \pm 2$, thus

$$\zeta(z) = \left(R^2 - \tfrac{1}{3}z^3\right) \cdot \frac{3}{R^3} \tag{7.21}$$

In absence of a reliable model for the ^{210}Po density profile, the choice of the test model is to some extent arbitrary. On the other hand the data presented in Fig. 7.5 clearly show that a "clean" region is present slightly above the detector equatorial plane, and that out of there the ^{210}Po rate increases (faster at the bottom, slower on the top). For a constant ^{210}Po density between $\zeta^{\mathrm{pl}}_{\mathrm{min}}$ and $\zeta^{\mathrm{pl}}_{\mathrm{max}}$, a plausible model of the increase has been assumed:

$$d(\zeta) = 1 + \begin{cases} \frac{b-1}{1+\exp\{\lambda[\zeta-(\zeta^{\mathrm{pl}}_{\mathrm{min}}-\mu)]\}} - \frac{b-1}{1+\exp\{\lambda\mu\}} & \zeta < \zeta^{\mathrm{pl}}_{\mathrm{min}} \\[2em] \frac{t-1}{\left(2-\zeta^{\mathrm{pl}}_{\mathrm{max}}\right)^2} \cdot \left(\zeta - \zeta^{\mathrm{pl}}_{\mathrm{max}}\right)^2 & \zeta > \zeta^{\mathrm{pl}}_{\mathrm{max}} \end{cases} \tag{7.22}$$

where $b(t)$ are scaling factors representing how large is the difference between the ^{210}Po density at the plateau and at the bottom(top) respectively. The parameter λ is the slope of a sigmoid function with turning point at μ and the second term in the $\zeta < \zeta_{\min}^{\mathrm{pl}}$ case is used to ensure the continuity of $d(\zeta)$ at $\zeta = \zeta_{\min}^{\mathrm{pl}}$.

Given the low event rate expected in the plateau region, a 25 day-long integration time has been chosen in order to have an adequate statistics.

7.5.3.2 *Kernel Density Estimator algorithm settings*

All the studies presented in this Chapter are obtained exploiting the methods of the `TKDE` class of the `ROOT` analysis framework [121], which allow one to build a density estimator using an adaptive kernel starting from an unbinned dataset as described in Sec. 7.5.1.

If the kernel density estimators are thought as a sort of smearing, one can immediately see that some issues arise at the boundaries. Among the techniques reported in literature to deal with this feature of the estimators, the "data reflection" is the one that is most suited for this analysis. As the name suggests, events close to the boundaries are mirrored in order to ensure that the probability of events that cross the boundary is exactly compensated by the same events that are reflected. As a by product of this method, the derivative at the boundaries is always zero. In the cases under study the ρ parameter introduced in Eq. (7.16) was set to 1 according to the prescription in [157].

7.5.3.3 *Test results*

The Plateau Finder has been tested on simulated pseudo-datasets obtained assuming the density model described in Sec. 7.5.3.1 and an exposure of 25 days. In order to account for different shapes of the ^{210}Po spatial distribution, three different configurations of the model have been considered: the first, shown in Fig. 7.12a with a narrow plateau enclosed between two steep rises; a second with a larger plateau and smoother boundaries (Fig. 7.12b); a third, quite unrealistic, case in which the ^{210}Po distribution has a very smooth increase (Fig. 7.12c).

For each of the three models shown in Fig. 7.12 5000 pseudo-datasets have been generated injecting a rate of 17.5 cpd/100 t in the plateau region (highlighted with a grey band). For each dataset the density estimator is build using the methods described in Sec. 7.5.1. In order to reconstruct the plateau region, the adaptive threshold on the derivative value introduced at step 4 of the procedure in Sec. 7.5.2 has been defined setting n at 1.5, 1.75 and 2 so that the volume expansion is stopped when the value of the derivative exceed 1.5, 1.75 and 2σ of its distribution.

The reconstructed plateau regions for a subset of the 100 pseudo-datasets are shown in Fig. 7.13 for the three models. One can see that for the "Narrow" and "Medium" model, the method tends to slightly overestimate the size of the plateau and that the definition of the border is more precise when the

(a) Narrow model   **(b)** Medium model   **(c)** Wide model
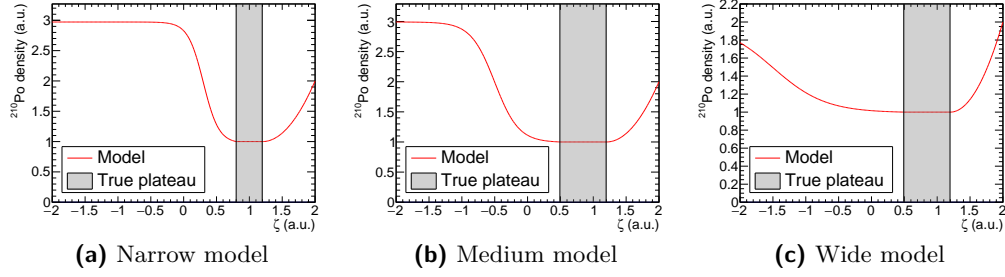
Figure 7.12: The three configurations of the ^{210}Po density model introduced in Eq. 7.22 that have been used to characterize the plateau reconstruction method.

underlying distribution has a steep increase. On the other hand, the "Wide" model case shows that when the slope of the distribution is small, the accuracy in the boundary determination is quite poor and the plateau region can vary significantly between different realization of the data.

(a) Narrow model   **(b)** Medium model   **(c)** Wide model
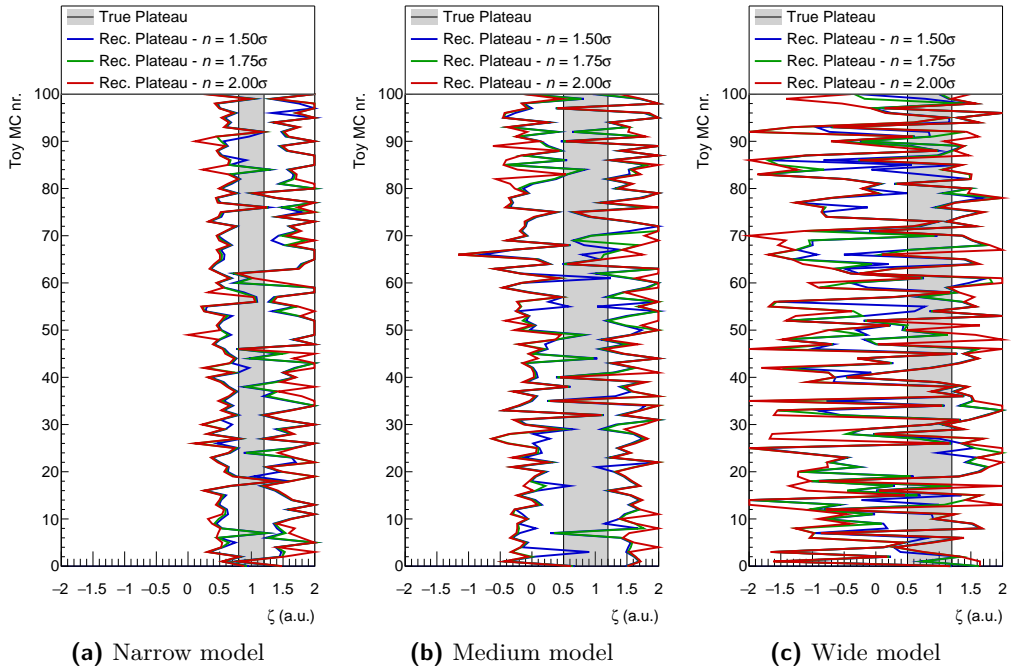
Figure 7.13: Reconstructed plateau region for the three ^{210}Po density models. The real plateau region is highlighted with a grey band and the different line colours refers to different values of the n parameter that defines the value of the derivative threshold.

Once the plateau region is defined, the number of events lying in the area N is defined by the integral of the density estimator between the reconstructed boundaries $\zeta^{\mathrm{pl}}_{\mathrm{min}}$ and $\zeta^{\mathrm{pl}}_{\mathrm{max}}$. The reconstructed plateau rate is computed as

$$R_{\mathrm{Po}} = \frac{N(\zeta^{\mathrm{pl}}_{\mathrm{min}}, \zeta^{\mathrm{pl}}_{\mathrm{max}})}{T \cdot M(\zeta^{\mathrm{pl}}_{\mathrm{min}}, \zeta^{\mathrm{pl}}_{\mathrm{max}})} \tag{7.23}$$

where T is the assumed live time and $M(\zeta^{\mathrm{pl}}_{\mathrm{min}}, \zeta^{\mathrm{pl}}_{\mathrm{max}})$ indicates the mass of the reconstructed plateau.

Figure 7.14 presents the distributions of the plateau rate for the different test models and values of n. The distribution of the rate injected in the actual plateau region is shown as a grey histogram and reflects the statistical fluctuations of the number of events in the real plateau, while the distribution of the results obtained for different values of n are displayed using different line colours. The distribution of the ^{210}Po plateau rate obtained by the method is described by a Gauss distribution that is in good agreement with the expected one. Discrepancies in the median of the reconstructed values with respect to the injected plateau rate are reported in the legend of Fig. 7.14, but are in general small when compared to the statistical uncertainty. When considering the "Narrow" model (Fig. 7.12a), a value of $n = 2$ leads to include significant portion of the rising parts of the density profile that makes the reconstructed rate to be overestimated for $1 \, \mathrm{cpd}/100 \, \mathrm{t}$ while on the other hand having $n = 1.5$ gives almost unbiased results. Conversely, in the situations described by the "Wide" model (Fig. 7.12c) if n is set to 1.5 the probability of being trapped inside a local minimum becomes non negligible and leads to underestimate the reconstructed rate for $2.2 \, \mathrm{cpd}/100 \, \mathrm{t}$, while the bias almost disappears if one considers $n = 2$. As a compromise between these two extreme situations, a value of $n = 1.8$ has been assumed for the study on ^{210}Po data that are presented in the next section.

(a) Narrow model   **(b)** Medium model   **(c)** Wide model
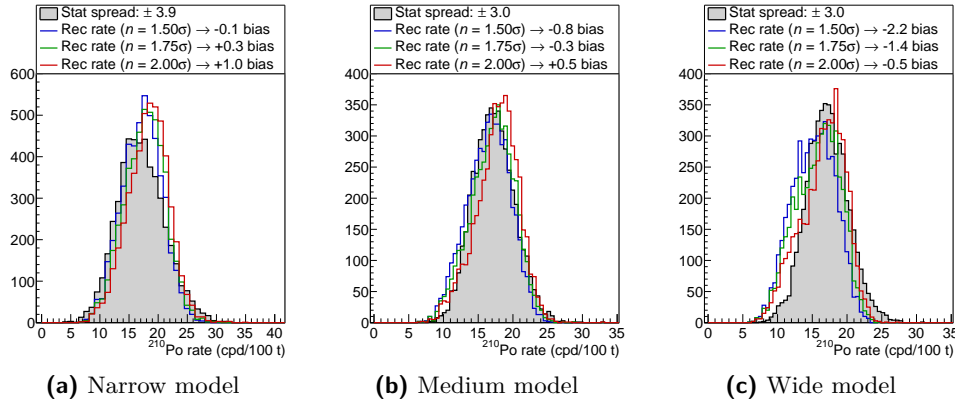
Figure 7.14: Distribution of the injected (filled grey) and reconstructed (coloured lines) ^{210}Po plateau rate for the three density models in Fig. 7.12 and different values of n.

To evaluate the dependence of the Plateau Finder results on the specific test density profile adopted, a study has been performed randomizing the parameter of the model in Eq. (7.22) and building a pseudo-dataset for each model realization with the same event rate injected in the plateau region. The density plot in Fig. 7.15a displays the model considered in this test, while the distribution of the plateau rate returned by the method is reported in Fig. 7.15b. As in the case of the "Narrow" test model, a small plateau size causes the results to be slightly overestimated, but the induced bias is still small compared to the expected statistical uncertainty.
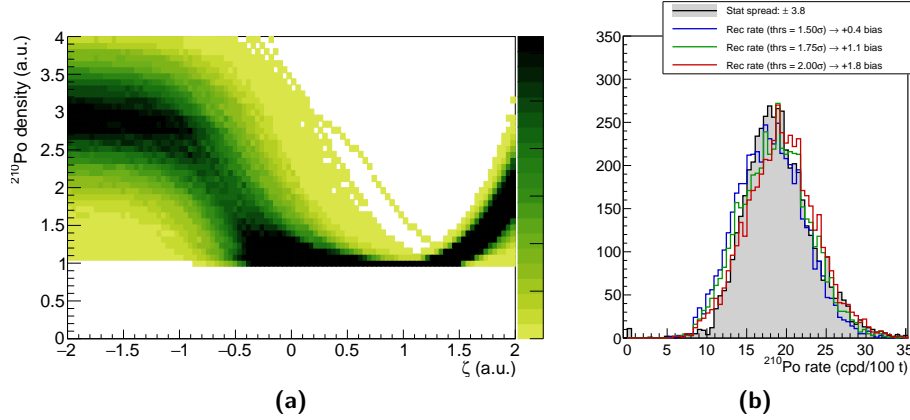
Figure 7.15: (a) 5000 realizations of the test model obtained randomizing the parameters of the "Medium" model template. **(b)** Distribution of the reconstructed plateau rate for different settings of the derivative value threshold.

In order to check the robustness of the method with respect to the injected plateau rate, the study presented above was repeated considering the "Medium" density model (Fig. 7.12b) and injecting a different ^{210}Po rate in the plateau region. This test has been performed assuming a minimum rate of $10\,\mathrm{cpd}/100\,\mathrm{t}$ (small compared to the expectation) and $50\,\mathrm{cpd}/100\,\mathrm{t}$ (≈ 3 times larger than the expected rate). The results are shown in Fig. 7.16 and still show a good agreement with the injected distribution, with deviations that are small compared to the expected statistical spread.

7.6 PRELIMINARY PLATEAU FINDER RESULTS

This section presents some preliminary results obtained by the Plateau Finder on Borexino data. Section 7.6.1 presents the details of the data selection and preparation while the results are discussed in Sec. 7.6.2 along with the limitation of the method and the prospects for future improvements.

Figure 7.16: Distribution of the reconstructed plateau rate for an injected value of 10 and 50 cpd/100 t obtained assuming the "Medium" model in Fig. 7.12b.

(a) Inj. rate $= 10$ cpd/100 t   **(b)** Inj. rate $= 50$ cpd/100 t
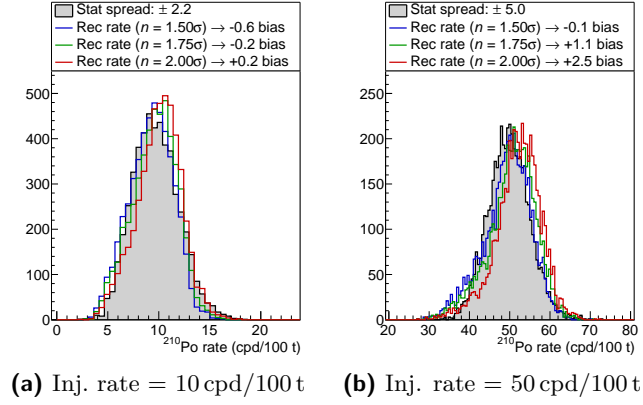
7.6.1 *Data selection and preparation*

This study has been performed on data acquired starting from January 2016, when the stabilization operations already stopped seasonal convective fluid motion. Borexino data are stored in the form of DST files that groups all the runs acquired in a single week and that are filtered following the same procedure outlined in Sec. 2.6.1 to remove crossing muons, short-lived cosmogenic isotopes and noise events. The filtered DST files have been stripped into "light trees" which contain the relevant information for the analysis such as the energy estimators, position reconstruction and pulse shape parameter.

The fiducial volume cut has been chosen aiming to select an adequate number of events with maximum distance from the vessel nylon where ^{210}Po is emanated. As a trade-off between the need of a large exposure and a large distance from the vessel, the fiducial volume has been defined as a 2 m radius sphere with its centre shifted by 40 cm towards the top part of the detector as shown in Fig. 7.17.

Figure 7.17: Fiducial Volume defined for the ^{210}Po analysis. The vessel profiles are drawn in different colours according to trace the evolution of the shape through time. The FV selected for the analysis is shown with a green filled area. A spherical FV centred in the detector coordinates origin is also shown with a dashed line for comparison.
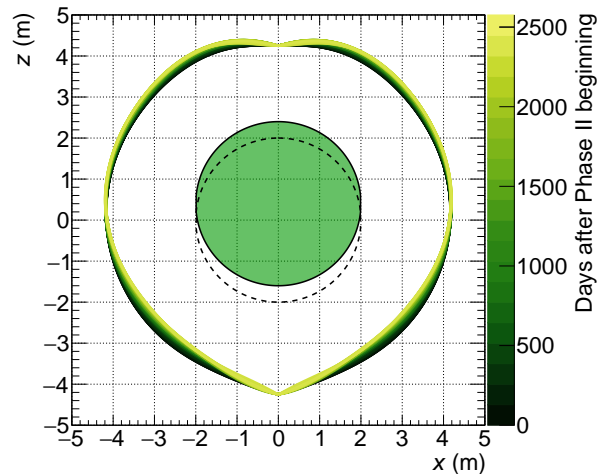
Events selected within the fiducial volume are classified as due to the ^{210}Po decay if the MLP variable is smaller than 0.05 and their energy is compatible with the ^{210}Po peak. Figure 7.18 shows the results of the selection when the fiducial volume cut and the ^{210}Po selection criteria are applied.
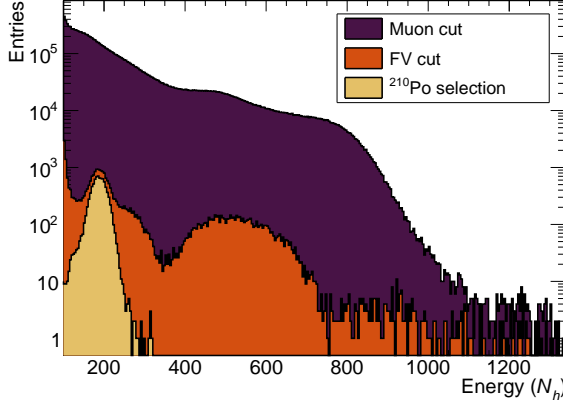
Figure 7.18: Performance of the ^{210}Po selection cuts. The small distortion on the left shoulder of the ^{210}Po peak is due to the inefficiency of the MLP classifier and accounts for $\approx 2\%$ of the ^{210}Po events.

In order to have a sufficient number of ^{210}Po events in the plateau region, groups of 4 subsequent DST files have been jointly analysed. The total live time associated to each group is given by the sum of the individual DST live time, that is computed considering also the impact of the muon and cosmogenic cuts. The typical live time for a group of 4 DSTs is ≈ 25 days.

7.6.2   *Results and prospects*

The full dataset collected between January 2016 and January 2019 has been divided into 29 4-weeks long periods. For each period the ^{210}Po events ζ distribution is built using the definition in Eq. 7.21 and analysed with the Plateau Finder method developed in Sec. 7.5. The ζ distributions for some of the considered time periods are shown as an example in Fig. 7.19.

The bottom panel of Fig. 7.20 shows the spatial evolution of the ^{210}Po rate along with the selected plateau region that is displayed as a shaded area. The ζ variable has been transformed back into the vertical coordinate z with a variable size binning to ensure that the same volume is represented by each bin. In the last months, in agreement with the Unbiased Minimum Finder results, the position of the "cleanest" region gets close to the centre in response to the stabilization of the detector.

Once the plateau is identified, for each period the number of ^{210}Po events in the plateau is computed by taking the integral of the density estimator in the selected region. The ^{210}Po rate in the plateau region is then evaluated as in Eq. (7.23). The statistical uncertainty associated to each measurement of
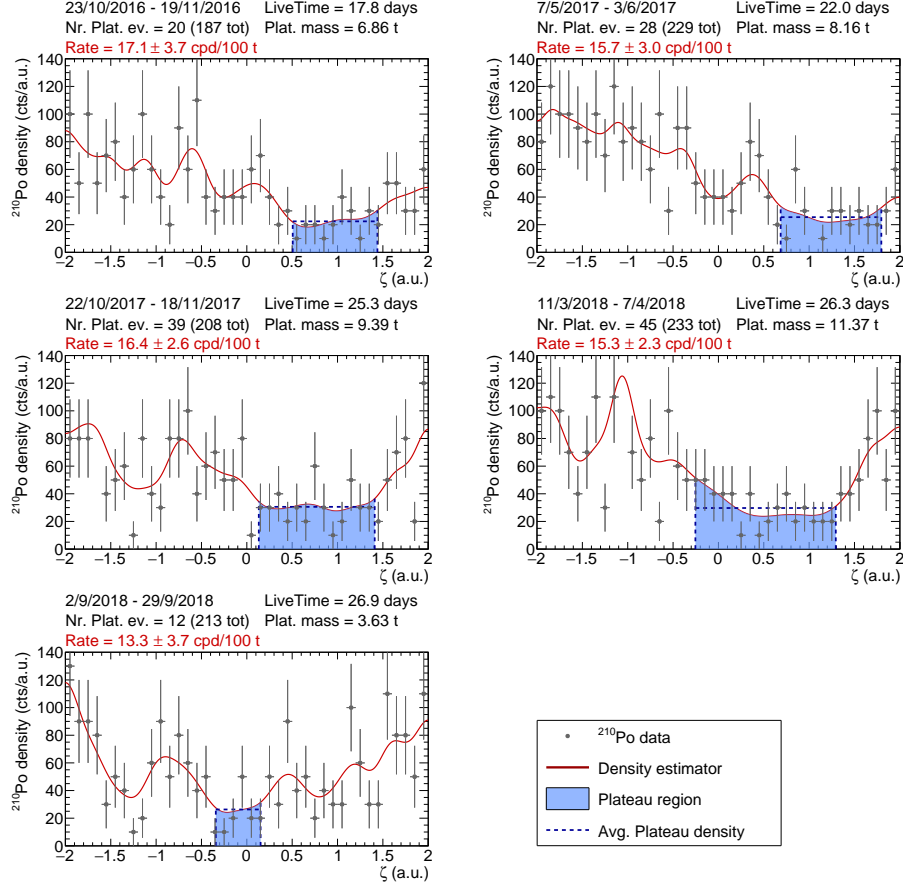
Figure 7.19: Examples of the plateau finder application on the ^{210}Po data. For each time period the density estimator is shown with a red line, and the identified plateau region is highlighted with a light blue area. The reconstructed average ^{210}Po density is reported as a dashed blue line.

the ^{210}Po plateau rate accounts for statistical fluctuations in the number of events in the plateau, which is

$$\sigma_{\text{stat}} = \frac{\sqrt{N(\zeta_{\min}^{\text{pl}}, \zeta_{\max}^{\text{pl}})}}{T \cdot M(\zeta_{\min}^{\text{pl}}, \zeta_{\max}^{\text{pl}})} \tag{7.24}$$

and it is reported along with the measured ^{210}Po rate in Fig. 7.19. The systematic uncertainty intrinsic to the plateau reconstruction method is hard to estimate. In Sec. 7.5.3 it has been shown that in case the ^{210}Po distribution has a wide, almost constant region the reconstructed plateau rate can be underestimated up to $\approx 1 \, \text{cpd}/100 \, \text{t}$ when the threshold is set at 1.8σ of the derivative values distribution. This situation is however not very realistic given the observed distribution of ^{210}Po events. On the contrary, the preliminary results of the fluid dynamics simulations discussed in Sec. 7.2.4 shows a ^{210}Po with a very narrow minimum region where, depending on the slope of the boundaries,

the plateau finder might tend to overestimate the plateau rate. A sound evaluation of the method systematic uncertainties can be obtained only when a more reliable model of the ^{210}Po will be available. At this preliminary stage of the measurement, a conservative systematic uncertainty of $\pm 1\,\mathrm{cpd}/100\,\mathrm{t}$ has been considered to account to possible bias of the method. Since the ^{210}Po spatial distribution changes with time, systematics will not affect all the measured periods coherently and therefore they are added in quadrature to the statistical uncertainty of the reconstructed rate in each period.

The evolution of the ^{210}Po rate in the region identified as a plateau is shown in the top panel of Fig. 7.20, with the error bar representing the statistical uncertainty and a couple of horizontal square brackets to display the total uncertainty which include the systematic uncertainty of the method. The evolution of the ^{210}Po plateau rate is expected to be described by Eq. (7.3), therefore it is possible to extract the ^{210}Po supported rate performing a fit of the rate trend.

The fact that every ^{210}Po rate measurement performed in a single time period, has a different "weight" influenced by its specific exposure demands a special attention in the fit procedure. Both using a binned likelihood and a Least Square method, the value of the supported rate reconstructed by the fit turns out to have a sensitive bias. This issue is overcome by using the Weighted Likelihood approach already implemented in the `ROOT` framework, which was proven to give unbiased results on simulated datasets.

The fit of the rate evolution reconstructs a preliminary value of 16.81 ± 0.87 cpd/100 t for the ^{210}Po supported rate with a $\chi^2/\mathrm{ndf} = 28.31/27$ showing a good agreement between the data and the model. However, the large shifts of the plateau region as well as the variations of its size shows that the method (or the ^{210}Po behaviour itself) is quite unstable. A tentative study has been performed fixing the size of the plateau region to a given value. This approach gives biased results by construction since the method builds the plateau volume starting from the reconstructed minimum of the ^{210}Po density and therefore returns the cleanest volume of that size. This effect is visible in Fig. 7.21a, which shows the fit of the plateau rate evolution obtained with different plateau size. The trend of the supported rate reconstructed by the fit as a function of the plateau size is reported in Fig. 7.21b. Possible solutions to remove this bias (*e.g.* vetoing a fraction of the reconstructed plateau) are currently under study.

In general, the results obtained by the Plateau Finder appear to be consistent with the value obtained in the analysis presented in Chap. 5 constraining the CNO rate and can be taken as a robust upper limit on the ^{210}Po supported rate. Nevertheless they must be considered as preliminary results that can be significantly improved in the next future. In particular, the one-dimensional approach developed in this chapter has a clear limitation whenever the "real" plateau region has a smaller radius than the selected fiducial volume. In this situation the projection on the vertical coordinate brings together the plateau and the dirtier outer region spoiling the measurement. This issues can be over-
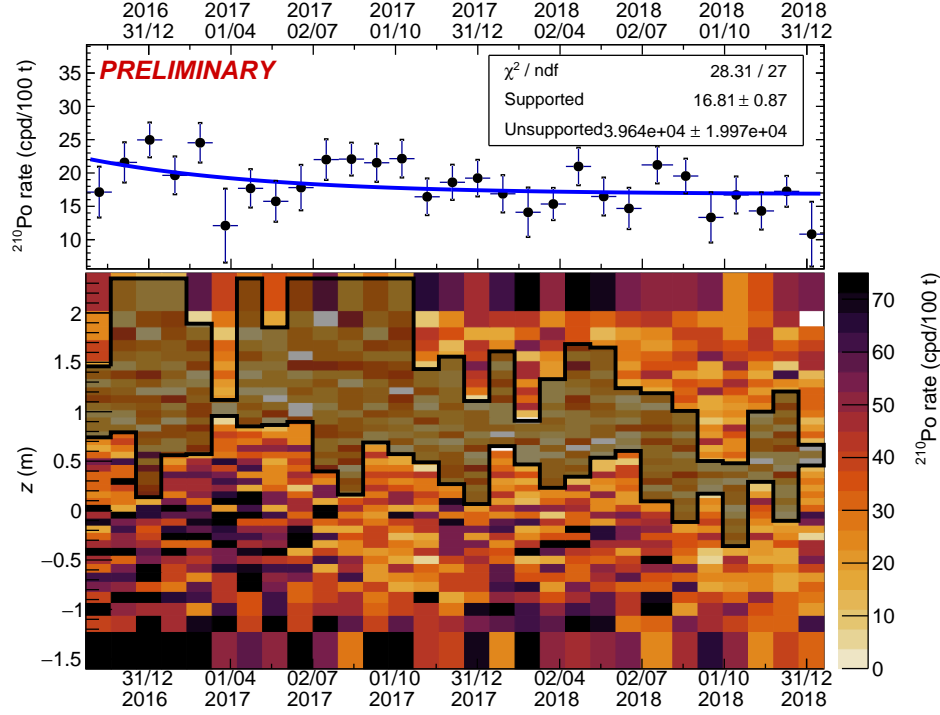
Figure 7.20: *Bottom panel* Spatial evolution of the ^{210}Po rate between January 2016 and January 2019. The plateau region selected by the method developed in Sec. 7.5 is displayed with a shaded area. *Top panel* Evolution of the ^{210}Po plateau rate in the selected plateau region. For each point, the error bar represents the statistical uncertainty, while the impact of a preliminary evaluation of systematic uncertainty is shown by plotting the total error budget with horizontal square brackets. The rate evolution is fitted with the model in Eq. (7.3) in order to extract the ^{210}Po supported rate.

come with the implementation of a multi-dimensional Plateau Finder which can be based on the very same principles of the method presented in this chapter. While a three-dimensional search appears challenging because of the low-event rate and sparseness of the data, the apparent azimuthal symmetry can be exploited to reduce the dimensionality of the problem and study the ^{210}Po distribution in the z–ρ space where $\rho = \sqrt{x^2 + y^2}$.

7.7   CONCLUSIONS AND OUTLOOKS

In this Chapter the strategy for the determination of the ^{210}Bi background rate from a measurement of the daughter ^{210}Po has been presented. The main difficulty of this measurement arises from the complicated spatial dynamics of ^{210}Po that is strongly influenced by fluid motion that carries ^{210}Po emanated from the inner vessel nylon towards the detector centre. In the last four years the collaboration started a series of operations aiming to the thermal
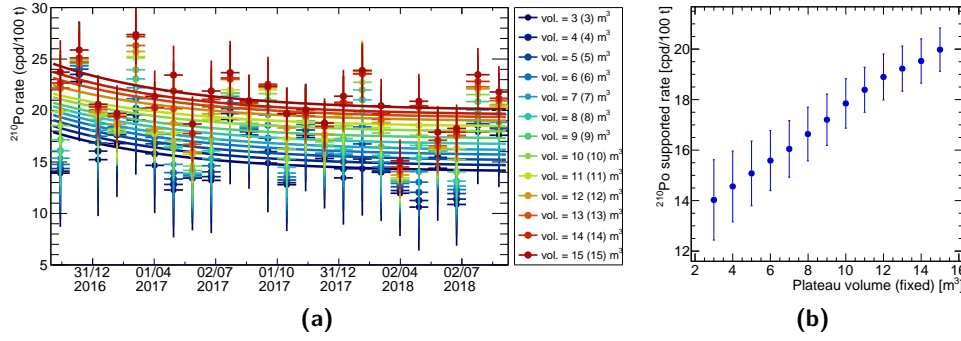
Figure 7.21: (a) Reconstructed rate as a function of time obtained with fixing the size of the plateau region at different values. **(b)** Values of the ^{210}Po supported rate reconstructed by the fit as a function of the chosen plateau size. These results are based on data collected between October 2016 and September 2018.

stabilization of the detector that changed dramatically the behaviour of the fluid, terminating the strong convective currents that were present before 2015.

Even if the fluid dynamics simulations developed by the joint efforts of the Virginia Tech, Milano-Politecnico and LNGS groups are not yet accurate enough to describe the expected ^{210}Po spatial distribution, a measurement of the ^{210}Po rate supported by the ^{210}Pb decay chain can be achieved selecting and tracing in time the volume of the detector in which both the gradient and the Laplacian of the ^{210}Po concentration are null.

In order to identify such *plateau* region without any assumptions on the actual ^{210}Po distribution, a method based on Kernel Density Estimators has been developed and applied on data. The evolution of the ^{210}Po rate inside the plateau is well compatible with the expected trend, but more studies are needed in order to interpret the evolution of the plateau region.

These results must be considered as preliminary and can be improved in the next future. In principle, the volume with the minimum ^{210}Po density should be searched scanning the entire three dimensional space. In this study the ^{210}Po events distribution has been considered only along the vertical axis only as a first step for future studies that will include more dimensions and exploit different density estimation methods. The development of advanced transport models obtained from fluid dynamics simulations may provide an adequate modelling of the ^{210}Po spatial distribution that can be used to further optimize the plateau finder method as well as to constrain the variety of plausible density profiles for the evaluation of systematic uncertainties. A reliable transport model will also allow to estimate the magnitude of the residual ^{210}Po diffusion from the inner vessel nylon into the innermost part of the detector.

It should be underlined that the measurement of the ^{210}Po supported rate is a crucial requirement for an improved result on the CNO neutrino rate, but it is just one piece of the puzzle. In order to use such measurement to constrain the ^{210}Bi background rate, the assumptions listed in Sec. 7.1 must be verified

and intense efforts are ongoing in order to establish if the foundation of the methods are satisfied.

CHAPTER 8

CONCLUSIONS AND OUTLOOKS

Solar neutrinos led to huge discoveries in the last forty years culminated with the discovery of neutrino oscillation. Now, solar neutrino physics is entering a new era of precision measurement that will lead to a deeper understanding of both neutrino properties and the functioning of the Sun by probing the predictions of the Standard Solar Model. Testing the Sun is particularly interesting because of the so-called *solar metallicity puzzle*: in recent years, new, more accurate determinations of the chemical abundances in the photosphere led to a 30%–40% reduction of the estimated content of metals in the Sun, but the resulting Low–metallicity (LZ) SSM turned out to be in strong tension with helioseismological observations, that are instead compatible with the previous generation of High-metallicity (HZ) SSM. Like helioseismological predictions, solar neutrino fluxes are affected by the different chemical composition of the Sun; therefore a precision measurement of solar neurtino fluxes can help determining the actual metal content of the Sun.

Borexino is the first pivotal experiment bringing precision neutrino physics below the MeV threshold. The key of the experimental success of Borexino lies in the outstanding radiopurity of its liquid scintillator. The residual background is however indistinguishable from the signal expected from the interaction of solar neutrinos on a event–by–event basis, and the rate of neutrino and background components is constrained exploiting the events energy, spatial and pulse shape distribution. In this thesis work a new multivariate fit has been developed, tested and benchmarked to perform such analysis. The direct implementation of a multidimensional fit overcomes some limitations of the previous analysis procedure while improving the fit stability and computational performance. Thanks to the sub-percent level accuracy of the Monte Carlo simulation of the experiment, multidimensional PDFs required for the multivariate analysis have been generated from ensembles of simulated data applying an efficient binning to minimize the impact of statistical fluctuations while not losing any physical information in the data.

The similarities between some of the fit components induce correlations that propagate in the reconstructed rates severely affecting the analysis potential. In order to break the correlations, additional information are included in the analysis in the form of external constraints leading to a substantial improvement of the overall sensitivity. In particular, the independent assessment of the ^{14}C and of the random coincidences rates is crucial for the measurement

of the low energy *pp* neutrinos, while at higher energy the interplay between the events distribution of *pep* neutrinos, CNO neutrinos and ^{210}Bi background prevents the analysis to individually resolve the different components and constraints based on the Standard Solar Model predictions had to be implemented. The impact of external constraints and of systematic uncertainties in the fit model has been studied performing the analysis on a large ensemble of pseudo–data generated from a family of models built randomly sampling the detector response parameters within their uncertainties. The resulting sensitivity confirmed Borexino as the most sensitive detector for low energy solar neutrinos.

Differently from previous Borexino analyses, the interaction rates of all the low energy solar neutrinos have been extracted performing a single fit over a wide energy range. The interaction rate of neutrinos produced in the *pp* reaction which drives the energy production in the Sun has been measured to be $134 \pm 10(\text{stat})^{+6}_{-10}(\text{sys})$ counts per day per 100 ton of scintillator (cpd/100 t). The rate of neutrinos produced by the electron capture on ^{7}Be has been constrained to $48.3 \pm 1.1^{+0.4}_{-0.7}$ cpd/100 t that is a factor 2 more accurate than the current prediction of the Standard Solar Model. The presence of neutrinos from the *pep* reaction has been established for the first time with a 5σ significance once the rate of CNO neutrinos has been constrained to the predictions of the Standard Solar Model: the rate of *pep* results $2.43 \pm 0.36^{+0.15}_{-0.22}$ cpd/100 t or $2.65 \pm 0.36^{+0.15}_{-0.24}$ cpd/100 t depending on the central value of the CNO neutrino rate which is computed considering respectively the High and Low metallicity composition model of the Sun. To set an upper limit on the rate of CNO neutrino interactions, the *pep* neutrino rate has been effectively constrained by imposing the ratio between the fluxes of *pp* and *pep* neutrinos that is determined by nuclear physics and is almost independent from the SSM details. The residual correlation with the ^{210}Bi rate prevents a measurement of the CNO neutrino rate and an upper limit at 8.1 cpd/100 t (95% C.L.) has been obtained.

The measurements of the solar neutrino interaction rates performed in this thesis work are the most accurate estimates to date, and combined with an independent measurement of ^{8}B neutrinos makes Borexino the only experiment with the capability to study the entire spectrum of solar neutrinos (with the exception of the very low flux of hep neutrinos). This rich set of measurements has been used to test both the oscillation framework and the predictions of the Standard Solar Models. Indeed, assuming the fluxes predicted by the SSM it is possible to compute from the interaction rates the $\nu_e \rightarrow \nu_e$ oscillation probability in the energy range between 0.2 and 10 MeV probing the transition between the vacuum–dominated low energy region and the matter–enhanced oscillations at high energy. Assuming the predictions of the HZ SSM, Borexino findings rejects the absence of the matter effect with a *p*-value of 0.018. On the other hand, the ^{8}B neutrino flux is reduced by 18% when considering the LZ SSM predictions, leading to a survival probability that is in better agreement with the prediction that does not consider the impact of the interaction potential with the Sun matter in neutrino oscillation.

Assuming the neutrino oscillation models tuned using experiments on Earth, it is possible to compute the solar neutrino fluxes from the measured interaction rates and use them to test the predictions of the SSM. The proton–proton fusion generating *pp* neutrinos is the process driving the energy production of the Sun and the luminosity in the neutrino channel computed from the inferred *pp* neutrino flux results in excellent agreement with the photon luminosity. The accurate measurement of ^{7}Be neutrinos allows to estimate the relative intensity of the two main terminations of the *pp*-chain, finding a result consistent with the predictions of the SSM. Finally, the measurements of the ^{7}Be and ^{8}B neutrino fluxes were combined to test the prediction of the SSM with different metallicity. Indeed, a different content of metals changes the solar plasma opacity and thus the temperature profile of the Sun. All the neutrino fluxes depend coherently on the temperature, therefore the different composition models of the Sun cause a difference in the predicted ^{8}B and ^{7}Be fluxes of 18% and 9% respectively. Performing a frequentist hypothesis test, Borexino results have been found fully compatible with the expectations of the HZ SSM, while the LZ SSM predictions are rejected with a *p*-value of 0.034. Similarly, a Bayesian hypothesis test of the HZ vs. LZ SSM predictions yields a Bayes factor of 4.9 confirming a mild preference for the HZ SSM.

A measurement of the CNO neutrino flux would provide a fundamental piece of information for the solution of the *metallicity puzzle*. Indeed, unlike ^{8}B and ^{7}Be neutrino fluxes that depend on the chemical composition in a indirect way, CNO neutrinos are direct messengers of the abundance of C and N in the Sun. The CNO cycle accounts for only $\approx 1\%$ of the total energy production of the Sun, and the corresponding weak flux of neutrinos has never been observed yet. Borexino is currently the only running experiment with the potential to obtain a first evidence of CNO neutrinos, but this measurement is made extremely challenging by the correlation of the CNO neutrino rate with the *pep* and ^{210}Bi background that can mimic the CNO events distribution. To break the correlations with *pep* and ^{210}Bi an independent assessment of the background rate is needed. An extensive sensitivity study has been performed, showing that if the *pep* and ^{210}Bi can be constrained with some independent estimates, then the Borexino sensitivity to CNO neutrinos is defined from a simple counting analysis and the additional information considered by the multivariate analysis only brings a limited improvement when the accuracy of the constraints on the background is quite loose. A hypothesis test based on a profile likelihood test statistics has been performed to assess Borexino median discovery power for different configurations of the uncertainties in the determination of the background rate. An uncertainty of the order of $2.4\,\mathrm{cpd}/100\,\mathrm{t}$ on the ^{210}Bi rate results in a median significance of 3σ and 2.3σ for the CNO rate predicted by the HZ and LZ SSM respectively.

While the rate of *pep* can be constrained with $\approx 1\%$ accuracy exploiting the aforementioned relationship with *pp* neutrinos and implementing a constraint based on solar luminosity, the background rate of ^{210}Bi must be measured independently. The most promising strategy to obtain a determination of the

^{210}Bi rate relies on the secular equilibrium of ^{210}Bi ($\tau_{\mathrm{Bi}} = 7\,\mathrm{days}$) with the parent ^{210}Pb ($\tau_{\mathrm{Pb}} = 32\,\mathrm{yr}$) and the ^{210}Po daughter ($\tau_{\mathrm{Po}} = 199\,\mathrm{days}$). Differently from ^{210}Bi and ^{210}Pb, ^{210}Po undergoes α decay and can be efficiently tagged exploiting the peculiar time development of the scintillation light. The measurement of the ^{210}Po rate is complicated by fluid motions in the liquid scintillator caused by temperature instabilities that carries ^{210}Po from the periphery of the detector —where it is emanated by the vessel nylon— towards the detector centre.

The spatial distribution of ^{210}Po could be used to separate the contribution of the ^{210}Po in equilibrium in the liquid scintillator from the component migrating from the vessel walls, but a complete description of the ^{210}Po density profile would require the complete description of the detector fluid dynamic that is not available. A model independent method based on kernel density estimators was developed to identify the regions of the detector with a constant ^{210}Po density compatible with the minimum. The so-called Polonium Plateau Finder was tested and optimized on pseudo–data and eventually applied on real data, showing that an uncertainty on the ^{210}Bi rate better than $2\,\mathrm{cpd}/100\,\mathrm{t}$ is in principle achievable. The result presented in this thesis must be however considered as preliminary. Indeed, the plateau finder implementation is currently limited to a single dimension, while the "^{210}Po–clean" scintillator volume has to be defined in a three dimensional space. The development of a multidimensional plateau finder based on the same concepts of the method presented in this thesis is foreseen for the next future and can be used to provide the measurement of the ^{210}Bi rate that is crucial to obtain a first evidence of CNO neutrinos that is currently the most ambitious physics goal of Borexino.

In summary, within the framework of this thesis work the fluxes of low energy solar neutrinos have been measured providing the most accurate determination of the pp and ^{7}Be neutrino fluxes to date. The CNO neutrino flux could only be constrained with an upper limit. A 3σ evidence of CNO neutrinos has been estimated to be within reach introducing a constraint on some background components, that can be achieved using the analysis technique proposed in this thesis. The first detection of CNO neutrinos would be a major scientific breakthrough since it would provide the ultimate evidence of the occurrence of the CNO cycle while giving a direct indication of the chemical composition of the Sun interior.

PROFILE-LIKELIHOOD BASED HYPOTHESIS TEST FOR CNO NEUTRINOS SENSITIVITY STUDIES

This appendix presents the definition of the profile-likelihood test statistics used to estimate the discovery power of Borexino for a CNO neutrino signal. The procedure to build the PDF of the test statistics for the rate analysis is reported in Sec. A.1, while the one followed in the case of the multivariate analysis is described in Sec. A.2.

The profile likelihood ratio $\lambda(\mu)$ is defined as

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{A.1}$$

where μ is the value of the parameter of interest (in this case the rate of CNO neutrino) and $\hat{\hat{\boldsymbol{\theta}}}$ is the set of nuisance parameters (the rates of the other components in the fit model) that maximise the likelihood $\mathcal{L}(\mu, \boldsymbol{\theta})$ for the specific value of μ. The $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ parameters are the maximum likelihood estimators, *i.e.* the values of the CNO rate and of the other parameters corresponding to the global maximum of the likelihood. The presence of a CNO neutrino signal would lead to an increase of the number of the events (the rate of CNO neutrino cannot be negative), therefore a test statistics q_0 defined as

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{A.2}$$

suited for the discovery of a positive signal [142] is chosen. The definition of q_0 in the search for a CNO neutrino signal is equivalent to

$$q_0 = -2\left[\ln\mathcal{L}(0, \hat{\hat{\theta}}) - \ln\mathcal{L}(\hat{\mu}, \hat{\theta})\right] \tag{A.3}$$

where the $\ln\mathcal{L}(0, \hat{\hat{\theta}})$ is the maximised likelihood when the CNO rate is fixed to zero while $\mathcal{L}(\hat{\mu}, \hat{\theta})$ represents the absolute likelihood maximum obtained when the CNO neutrino rate is free in the analysis.

In order to compute the median discovery power to the CNO signal, the PDFs of the test statistics must be known. In this thesis work the PDFs of the test statistics have been built using a Monte Carlo approach, *i.e.* analysing thousands of pseudo-datasets generated injecting a known signal. For each

dataset the value of q_0 is computed and used to build the distribution of the test statistics. However, Cowan et al. in [142] provides an asymptotic expression for probability distribution function of the test statistics q_0:

$$f(q_0|R'_{\text{CNO}}) = \left(1 - \Phi\left(\frac{R'_{\text{CNO}}}{\sigma_{\text{CNO}}}\right)\right)\delta(q_0)+$$
$$+ \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp\left[-\frac{1}{2}\left(\sqrt{q_0} - \frac{R'_{\text{CNO}}}{\sigma_{\text{CNO}}}\right)\right] \quad \text{(A.4)}$$

where R'_{CNO} is the value of the assumed CNO rate, σ_{CNO} is the expected statistical uncertainty and $\Phi(x)$ is the cumulative distribution of a zero mean and unit variance Gaussian. This asymptotic expression has been used to cross-check the distribution obtained with the Monte Carlo method as described in the next sections.

A.1   TEST STATISTICS DISTRIBUTIONS FOR A RATE ANALYSIS

In the simple counting model sketched in Sec. 6.2.1, the region of interest is populated by three species only: two of them (^{210}Bi and *pep*) forming the background are independently constrained, while the signal is free. Given a certain number of counts N, the negative log-likelihood for that specific realization is

$$-\ln L(\mu, \boldsymbol{\theta}) = \frac{[N - E \cdot (\varepsilon_{\text{CNO}}R_{\text{CNO}} + \varepsilon_{\text{Bi}}R_{\text{Bi}} + \varepsilon_{pep}R_{pep})]^2}{2N}$$
$$+ \frac{\left(R_{^{210}\text{Bi}} - \tilde{R}_{\text{Bi}}\right)^2}{2\tilde{\sigma}_{\text{Bi}}^2} + \frac{\left(R_{pep} - \tilde{R}_{pep}\right)^2}{2\tilde{\sigma}_{pep}^2} \quad \text{(A.5)}$$

where the ε coefficients are the ones defined in Sec. 6.2.1 while the last two terms are used to express the constraints on the background rate, that is estimated to be $\tilde{R}_{pep} \pm \tilde{\sigma}_{pep}$ and $\tilde{R}_{\text{Bi}} \pm \tilde{\sigma}_{\text{Bi}}$ for *pep* neutrino and ^{210}Bi respectively.

If one considers the background hypothesis in which CNO neutrinos are not present, the number of counts N is a random variable described by a Poisson PDF with mean given by the sum of the expected number of counts of the background components. The minimum of the log-likelihood function in Eq. (A.5) is always zero when the CNO rate is free, and the returned CNO rate is

$$\hat{R}_{\text{CNO}} = \frac{1}{E\varepsilon_{\text{CNO}}}\left[N - E \cdot (\varepsilon_{\text{Bi}}R_{\text{Bi}} + \varepsilon_{pep}R_{pep})\right] \quad \text{(A.6)}$$

When one assumes the background hypothesis H_0, *i.e.* imposes $R_{\text{CNO}} = 0$ the minimum of the NLL is

$$-\ln L(0, \hat{\hat{\boldsymbol{\theta}}}) = \frac{1}{2} \cdot \frac{[N - E \cdot (\varepsilon_{\text{Bi}}X_{\text{Bi}} + \varepsilon_{pep}X_{pep})]^2}{N + E^2 \cdot (\varepsilon_{\text{Bi}}\sigma_{\text{Bi}} + \varepsilon_{pep}\sigma_{pep})^2} \quad \text{(A.7)}$$

(a) Rate analysis   **(b)** Rate + Shape analysis
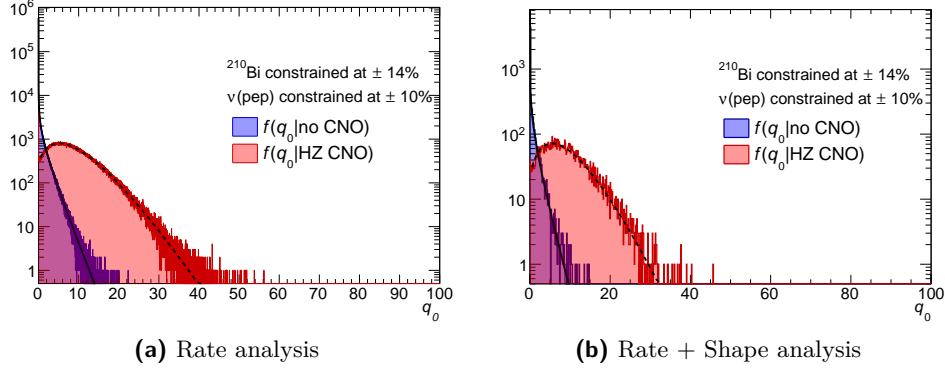
Figure A.1: Distribution of the test statistics q_0 obtained for the background hypothesis (blue) and for the CNO signal predicted by the HZ SSM (red) when the *pep* and ^{210}Bi background are known with a 0.28 and 2.45 cpd/100 t accuracy. In **(a)** the distributions were obtained with the simple rate rate analysis, while the ones in **(b)** were built performing a MV analysis on MC generated datasets. The distributions expected from [142] are shown for the background and HZ CNO as a solid and dashed line.

With Eq. (A.7) in mind, one can thus obtain the PDF of the test statistics q_0 defined in Eq. A.2 under a given assumption on the CNO rate just by taking Eq. (A.7) and sampling N from a Poisson distribution with a mean value depending on the considered hypothesis, and $\tilde{R}_i$ varying according to the corresponding uncertainty $\tilde{\sigma}_i$. An example of the distributions $f(q_0|\text{no CNO})$ and $f(q_0|\text{HZ CNO})$ for the *pep* and ^{210}Bi rate constrained with an accuracy of 0.28 and 2.45 cpd/100 t respectively are shown in Fig. A.1a. The distributions are in excellent agreement with the approximate expression given by Cowan et al.

A.2   TEST STATISTICS DISTRIBUTION FOR THE MULTIVARIATE ANALYSIS

The distributions of the test statistics for the multivariate analysis have been obtained with a Monte Carlo method.

A set of 10^4 pseudo–dataset including the events energy and radial position has been generated according to the background or signal hypothesis considered. To compute the value of q_0 each dataset is fitted twice as indicated in Eq. (A.3), keeping the *pep* and ^{210}Bi background constrained. The first fit is performed with including the CNO neutrinos in the fit model, while they are not considered in the second fit. The value of q_0 is given by the difference of the Negative Log-Likelihood minimized in the two fits.

An example of the test statistics distributions obtained for the background and HZ CNO signal hypothesis when the *pep* and ^{210}Bi background rates are constrained with an accuracy of 0.28 and 2.45 cpd/100 t is shown in Fig. A.1b.

Also in this case the distributions of the test statistics are in good agreement with the expectation from [142] where the σ_{CNO} in this case are the values of the statistical uncertainty shown in Fig. 6.4.

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

[1] H. R. Crane. "The Energy and Momentum Relations in the Beta-Decay, and the Search for the Neutrino". In: *Rev. Mod. Phys.* 20 (1 1948).

[2] J. Chadwick. "Intensitätsverteilung im magnetishen Spektrum der β-Strahlen von Radium $B + C$". In: *Verh. d. deutschen Phys. Ges.* 16 (1914).

[3] J. Chadwick and C.D. Ellis. "A Preliminary Investigation of the Intensity Distribution in the β-Ray Spectra of Radium B and C". In: *Proceedings of the Cambridge Philosophical Society* 21 (1922).

[4] E. Fermi. "Tentativo di una Teoria Dei Raggi β". In: *Nuovo Cim.* 11.1 (1934).

[5] E. Fermi. "Versuch einer Theorie der β-Strahlen. I". In: *Z. Phys.* 88.3 (1934).

[6] F. Reines and C. L. Cowan. "Detection of the Free Neutrino". In: *Phys. Rev.* 92 (3 1953).

[7] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". In: *Science* 124.3212 (1956).

[8] M. Tanabashi and other (Particle Data Group). "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018).

[9] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 1957).

[10] M. Goldhaber, L. Grodzins, and A. W. Sunyar. "Helicity of Neutrinos". In: *Phys. Rev.* 109 (3 1958).

[11] A. Salam and J.C. Ward. "Electromagnetic and weak interactions". In: *Phys. Lett.* 13.2 (1964).

[12] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967).

[13] F.J. Hasert et al. (Gargamelle Collaboration). "Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment". In: *Phys. Lett. B* 138 (46 1973).

[14] B. Pontecorvo. "Mesonium and antimesonium". In: *Sov. Phys. JETP* 6 (1957).

[15] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". In: *Sov. Phys. JETP* 7 (1958).

[16] L. Wolfenstein. "Neutrino oscillations in matter". In: *Phys. Rev. D* 17 (9 1978).

[17] S.P. Mikheyev and A.Y. Smirnov. "Resonance enhancement of oscillations in matter and solar neutrino spectroscopy". In: *Yadernaya Fizika* 42 (1985).

[18] J.N. Bahcall. *Neutrino Astrophysics.* Cambridge Univeristy Press, 1989.

[19] J. N. Bahcall et al. "Solar Neutrino Flux." In: *Astrophys. J.* 137 (1963).

[20] J.N. Bahcall. "Solar models: An historical overview". In: *Nucl. Phys. B Proc. Suppl.* 118 (2003). Proceedings of the XXth International Conference on Neutrino Physics and Astrophysics.

[21] R. Davis, D.S. Harmer, and K.C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968).

[22]   J. N. Bahcall et al. "Solar Neutrinos.IV. Effect of Radiative Opacities on Calculated Neutrino Fluxes". In: *Astrophys. J.* 184 (1973).

[23]   J.N. Bahcall et al. "New Solar-Neutrino Flux Calculations and Implications Regarding Neutrino Oscillations". In: *Phys. Rev. Lett.* 45 (11 1980).

[24]   J.N. Bahcall et al. "Standard solar models and the uncertainties in predicted capture rates of solar neutrinos". In: *Rev. Mod. Phys.* 54 (3 July 1982).

[25]   J.N. Bahcall and R.K. Ulrich. "Solar models, neutrino experiments, and helioseismology". In: *Rev. Mod. Phys.* 60 (2 1988).

[26]   J. N. Bahcall and M. H. Pinsonneault. "Standard solar models, with and without helium diffusion, and the solar neutrino problem". In: *Rev. Mod. Phys.* 64 (4 1992).

[27]   J.N. Bahcall, M.H. Pinsonneault, and G.J. Wasserburg. "Solar models with helium and heavy-element diffusion". In: *Rev. Mod. Phys.* 67 (4 1995).

[28]   J.N. Bahcall, M.H. Pinsonneault, and S. Basu. "Solar Models: Current Epoch and Time Dependences, Neutrinos, and Helioseismological Properties". In: *Astrophys. J.* 555.2 (2001).

[29]   J.N. Bahcall and M.H. Pinsonneault. "What Do We (Not) Know Theoretically about Solar Neutrino Fluxes?" In: *Phys. Rev. Lett.* 92 (12 2004).

[30]   J.N. Bahcall, A.M. Serenelli, and S. Basu. "New Solar Opacities, Abundances, Helioseismology, and Neutrino Fluxes". In: *The Astrophysical Journal* 621.1 (2005).

[31]   J.N. Bahcall, A.M. Serenelli, and S. Basu. "10,000 Standard Solar Models: A Monte Carlo Simulation". In: *Astrophy. J. Supp. S.* 165.1 (2006).

[32]   A.M. Serenelli et al. "New Solar Composition: The problem with solar models revisited". In: *Astrophys. J.* 705.2 (2009).

[33]   A.M. Serenelli, W.C. Haxton, and C. Peña-Garay. "Solar Models with Accretion I. Application to the solar abundance problem". In: *Astrophys. J.* 743.1 (2011).

[34]   N. Vinyoles et al. "A New Generation of Standard Solar Models". In: *Astrophys. J* 835.2 (2017).

[35]   A. Serenelli. "Alive and well: A short review about standard solar models". In: *Eur. Phys. J. A* 55.4 (2016).

[36]   A.S. Eddington. "The Internal Constitution of the Stars". In: *Nature* 106 (1920).

[37]   A.S. Eddington. "The internal constitution of the stars". In: *Observatory* 43 (1920).

[38]   G. Gamow. "Zur Quantentheorie des Atomkernes". In: *Z. Phys.* (1928).

[39]   R. d. Atkinson. "Atomic Synthesis and Stellar Energy I-II". In: *Astrophys. J.* 73 (1931).

[40]   G. Gamow. "Nuclear Energy Sources and Stellar Evolution". In: *Phys. Rev.* 53 (7 1938).

[41]   S. Chandrasekhar. *An introduction to the study of stellar structure.* The University of Chicago press, 1939.

[42] H.A. Bethe and C.L. Critchfield. "The Formation of Deuterons by Proton Combination". In: *Phys. Rev.* 54 (4 1938).

[43] C.F. von Weizsäcker. "Über Elementumwandlungen in Innern der Sterne I". In: *Phys. Z.* (38 1937).

[44] C.F. von Weizsäcker. "Über Elementumwandlungen in Innern der Sterne II". In: *Phys. Z.* (39 1938).

[45] H.A. Bethe. "Energy Production in Stars". In: *Phys. Rev.* 55 (5 1939).

[46] A. Best et al. "Underground nuclear astrophysics: Why and how". In: *Eur. Phys. J. A* 52.4 (2016).

[47] E. G. Adelberger et al. "Solar fusion cross sections. II. The *pp* chain and CNO cycles". In: *Rev. Mod. Phys.* 83 (1 2011).

[48] C. Iliadis. *Nuclear Physics of Stars*. Wiley, 2008.

[49] A. Formicola et al. (LUNA Collaboration). "Astrophysical *S*-factor of ^{14}N$(p,\gamma)^{15}$O". In: *Phys. Lett. B* 591.1 (2004).

[50] M. Marta et al. (LUNA Collaboration). "The ^{14}N$(p,\gamma)^{15}$O reaction studied with a composite germanium detector". In: *Phys. Rev. C* 83 (4 2011).

[51] L. C. Stonehill, J. A. Formaggio, and R. G. H. Robertson. "Solar neutrinos from CNO electron capture". In: *Phys. Rev. C* 69 (1 2004).

[52] F.L. Villante. "ecCNO solar neutrinos: A challenge for gigantic ultra-pure liquid scintillator detectors". In: *Phys. Lett. B* 742 (2015).

[53] R.B. Leighton, R.W. Noyes, and G.W. Simon. "Velocity Fields in the Solar Atmosphere. I. Preliminary Report." In: *Astrophys. J.* 135 (1962).

[54] M. Asplund, N. Grevesse, and A.J. Sauval. "The solar chemical composition". In: *Nucl. Phys. A* 777 (2006). Special Isseu on Nuclear Astrophysics.

[55] M. Asplund, A.J. Grevesse N.and Sauval, and P. Scott. "The Chemical Composition of the Sun". In: *Annu. Rev. Astron. Astrophys.* 47.1 (2009).

[56] Scott, P. et al. "The elemental composition of the Sun - I The intermediate mass elements Na to Ca". In: *Astron. Astrophys.* 573 (2015), A25.

[57] Scott, P. et al. "The elemental composition of the Sun - II The Iron group element from Sc to Ni". In: *Astron. Astrophys.* 573 (2015), A26.

[58] N. Grevesse and A.J. Sauval. "Standard Solar Composition". In: *Space Sci. Rev.* 85.1 (1998).

[59] Christensen-Dalsgaard, J. et al. "On the opacity change required to compensate for the revised solar composition". In: *Astron. Astrophys.* 494.1 (2009).

[60] N.R. Badnell et al. "Updated opacities from the Opacity Project". In: *Mon. Notices Royal Astron. Soc.* 360 (2005).

[61] C.A. Iglesias and F.J. Rogers. "Updated Opal Opacities". In: *Astrophys. J.* 464 (1996).

[62] J.E. Bailey et al. "A higher-than-predicted measurement of iron opacity at solar interior temperatures". In: *Nature* 517 (2015).

[63] N. Song et al. "Helioseismic and neutrino data-driven reconstruction of solar properties". In: *Mon. Notices Royal Astron. Soc.* 477.1 (2018).

[64] F. Vissani and Strumia. *Neutrino masses and mixings and...* 2006. arXiv: `hep-ph/0606054`.

[65]   B.T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". In: *Astrophys. J.* 496.1 (1998).

[66]   W. Hampel et al. (GALLEX Collaboration). "GALLEX solar neutrino observations: results for GALLEX IV". In: *Phys. Lett. B* 447.1 (1999).

[67]   M. Altmann et al. (GNO Collaboration). "Complete results for five years of GNO solar neutrino observations". In: *Phys. Lett. B* 616.3 (2005).

[68]   J. N. Abdurashitov et al. (SAGE Collaboration). "Measurement of the solar neutrino capture rate with gallium metal. III. Results for the 2002–2007 data-taking period". In: *Phys. Rev. C* 80 (1 2009).

[69]   K. S. Hirata et al. "Observation of ^{8}B solar neutrinos in the Kamiokande-II detector". In: *Phys. Rev. Lett.* 63 (1 1989).

[70]   K. S. Hirata et al. "Real-time, directional measurement of ^{8}B solar neutrinos in the Kamiokande II detector". In: *Phys. Rev. D* 44 (8 1991).

[71]   J. Hosaka et al. (Super-Kamiokande Collaboration). "Solar neutrino measurements in Super-Kamiokande-I". In: *Phys. Rev. D* 73 (11 2006).

[72]   J. P. Cravens et al. (The Super-Kamiokande Collaboration). "Solar neutrino measurements in Super-Kamiokande-II". In: *Phys. Rev. D* 78 (3 2008).

[73]   K. Abe et al. (Super-Kamiokande Collaboration). "Solar neutrino results in Super-Kamiokande-III". In: *Phys. Rev. D* 83 (5 2011).

[74]   K. Abe et al. (Super-Kamiokande Collaboration). "Solar neutrino measurements in Super-Kamiokande-IV". In: *Phys. Rev. D* 94 (5 2016).

[75]   J Boger et al. "The Sudbury Neutrino Observatory". In: *Nucl. Instrum. Methods Phys. Res. A* 449.1 (2000). DOI: `https:%%doi.org/10.1016/S0168-9002(99)01469-2`.

[76]   Q.R. Ahmad et al. (SNO Collaboration). "Measurement of the Rate of $\nu_e + d \to p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 2001).

[77]   Q.R. Ahmad et al. (SNO Collaboration). "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002).

[78]   S.N. Ahmed et al. (SNO Collaboration). "Measurement of the Total Active ^{8}B Solar Neutrino Flux at the Sudbury Neutrino Observatory with Enhanced Neutral Current Sensitivity". In: *Phys. Rev. Lett.* 92 (18 2004).

[79]   B. Aharmim et al. (SNO Collaboration). "Independent Measurement of the Total Active ^{8}B Solar Neutrino Flux Using an Array of ^{3}He Proportional Counters at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 101 (11 2008).

[80]   C. Arpesella and others (Borexino Collaboration). "First real time detection of ^{7}Be solar neutrinos by Borexino". In: *Phys. Lett. B* 658.4 (2008).

[81]   G. Bellini et al. (Borexino Collaboration). "Precision Measurement of the ^{7}Be Solar Neutrino Interaction Rate in Borexino". In: *Phys. Rev. Lett.* 107 (14 2011).

[82]   Bellini, G. and others (Borexino Collaboration). "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy". In: *Phys. Rev. D* 89.11 (2014).

[83]   G. Bellini et al. "Absence of a day–night asymmetry in the 7Be solar neutrino rate in Borexino". In: *Phys. Lett. B* 707.1 (2012).

[84]   Bellini, G. and others (Borexino Collaboration). "First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino". In: *Phys. Rev. Lett.* 108 (5 2012).

[85]   G. Bellini and others (Borexino Collaboration). "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).

[86]   G. Bellini et al. (Borexino collaboration). "Neutrinos from the primary proton–proton fusion process in the Sun". In: *Nature* 512 (2014).

[87]   A. Gando et al. (KamLAND Collaboration). "^{7}Be solar neutrino measurement with KamLAND". In: *Phys. Rev. C* 92 (5 2015).

[88]   S. Abe et al. (KamLAND Collaboration). "Measurement of the ^{8}B solar neutrino flux with the KamLAND liquid scintillator detector". In: *Phys. Rev. C* 84 (2011).

[89]   M. Anderson et al. (SNO+ Collaboration). "Measurement of the ^{8}B solar neutrino flux in SNO+ with very low backgrounds". In: *Phys. Rev. D* 99 (1 2019).

[90]   S. Andriga et al. (SNO+ Collaboration). "Current Status and Future Prospects of the SNO+ Experiment". In: *Adv. High Energy Phys.* 2016 (2016).

[91]   Z. Djurcic et al. (JUNO). *JUNO Conceptual Design Report*. 2015. arXiv: `1508.07166 [physics.ins-det]`.

[92]   Fengpeng An et al. (JUNO Collaboration). "Neutrino Physics with JUNO". In: *J. Phys.* G43.3 (2016). DOI: `10.1088/0954-3899/43/3/030401`.

[93]   X.F. Ding. "High Precision Solar Neutrino Spectroscopy with Borexino and Juno". PhD Thesis. Gran Sasso Science Institute, 2019.

[94]   M. Maltoni and Y.A. Smirnov. "Solar neutrinos and neutrino physics". In: *Eur. Phys. J. A* 52.4 (2016).

[95]   K. Abe et al. (Hyper-Kamiokande Collaboration). *Hyper-Kamiokande Design Report*. 2018. arXiv: `1805.04163 [physics.ins-det]`.

[96]   L. Baudis et al. "Neutrino physics with multi-ton scale liquid xenon detectors". In: *J. Cosmol. Astropart. Phys* 2014.01 (2014).

[97]   D. Franco et al. "Solar neutrino detection in a large volume double-phase liquid argon experiment". In: *J. Cosmol. Astropart. Phys* 2016.08 (2016).

[98]   M. Yeh et al. "A new water-based liquid scintillator and potential applications". In: *Nucl. Instrum. Methods Phys. Res. A* 660.1 (2011).

[99]   J.R. Alonso et al. "Advanced Scintillator Detector Concept (ASDC): A Concept Paper on the Physics Potential of Water-Based Liquid Scintillator". In: (2014). physics.ins-det: `1409.5864`.

[100]   G D. Orebi Gann (THEIA Interest Group). *Physics Potential of an Advanced Scintillation Detector: Introducing THEIA*. 2015. arXiv: `1504.08284 [physics.ins-det]`.

[101]   J. Caravaca et al. "Experiment to demonstrate separation of Cherenkov and scintillation signals". In: *Phys. Rev. C* 95 (5 2017).

[102]   R. Bonventre and G.D. Orebi Gann. "Sensitivity of a low threshold directional detector to CNO-cycle solar neutrinos". In: *Eur. Phys. J. C* 78.6 (2018).

[103] T. Kovacs et al. "Borex: Solar neutrino experiment via weak neutral and charged currents in Boron-11". In: *Sol. Phys.* 128.1 (1990).

[104] J. N. Bahcall, M. Kamionkowski, and A. Sirlin. "Solar neutrinos: Radiative corrections in neutrino-electron scattering experiments". In: *Phys. Rev. D* 51 (11 1995).

[105] V. Lagomarsino and G. Testera. "A gateless charge integrator for Borexino energy measurement". In: *Nucl. Instrum. Methods Phys. Res. A* 430.2 (1999).

[106] Alimonti, G. and others (Borexino Collaboration). "The Borexino detector at the Laboratori Nazionali del Gran Sasso". In: *Nucl. Instrum. Methods Phys. Res. A* 600.3 (2009).

[107] J. Benziger et al. "The nylon scintillator containment vessels for the Borexino solar neutrino experiment". In: *Nucl. Instrum. Methods Phys. Res. A* 582.2 (2007).

[108] J. Benziger et al. "A scintillator purification system for the Borexino solar neutrino detector". In: *Nucl. Instrum. Methods Phys. Res. A* 587.2 (2008).

[109] G. Alimonti et al. "The liquid handling systems for the Borexino solar neutrino detector". In: *Nucl. Instrum. Methods Phys. Res. A* 609.1 (2009).

[110] J. Benziger et al. "The fluid-filling system for the Borexino solar neutrino detector". In: *Nucl. Instrum. Methods Phys. Res. A* 608.3 (2009).

[111] G Alimonti et al. "A large-scale low-background liquid scintillation detector: the counting test facility at Gran Sasso". In: *Nucl. Instrum. Methods Phys. Res. A* 406.3 (1998).

[112] M.G. Giammarchi et al. "The scintillator solvent procurement for the Borexino solar neutrino detector". In: *Nucl. Instrum. Methods Phys. Res. A* 648.1 (2011).

[113] J.B. Benziger and F.P. Calaprice. "Large-scale liquid scintillation detectors for solar neutrinos". In: *Eur. Phys. J. A* 52.4 (2016).

[114] Simone Marcocci. "Precision measurement of solar ν fluxes and prospects fot $0\nu\beta\beta$ search with ^{136}Xe-loaded liquid scintillators". PhD Thesis. Gran Sasso Science Institute, 2016.

[115] Stefano Davini. "Measurement of the *pep* and CNO solar neutrino interaction rates in Borexino". PhD Thesis. Università degli Studi di Genova, 2012.

[116] M. Agostini et al. (Borexino collaboration). "Seasonal modulation of the ^{7}Be solar neutrino rate in Borexino". In: *Astroparticle Physics* 92 (2017).

[117] H. Back et al. (Borexino Collaboration). "CNO and *pep* neutrino spectroscopy in Borexino: Measurement of the deep-underground production of cosmogenic ^{11}C in an organic liquid scintillator". In: *Phys. Rev. C* 74 (4 2006).

[118] C. Galbiati et al. "Cosmogenic ^{11}C production and sensitivity of organic scintillator detectors to *pep* and CNO neutrinos". In: *Phys. Rev. C* 71 (5 2005).

[119] M. Agostini et al. (Borexino collaboration). *First Simultaneous Precision Spectroscopy of pp, ^{7}Be, and pep Solar Neutrinos with Borexino Phase-II.* 2017. arXiv: 1707.09279 [physics.hep-ex].

[120] G.F. Knoll. *Radiation detection and measurement.* John Wiley & Sons Inc., 2010.

[121] R. Brun and F. Rademakers. "ROOT — An object oriented data analysis framework". In: *Nucl. Instrum. Methods Phys. Res. A* 1 (1997).

[122] D. Franco, G. Consolati, and D. Trezzi. "Positronium signature in organic liquid scintillators for neutrino experiments". In: *Phys. Rev. C* 83 (1 2011).

[123] G. Bellini et al. (Borexino collaboration). "Observation of geo-neutrinos". In: *Phys. Lett. B* 687.4 (2010).

[124] G. Bellini et al. (Borexino collaboration). "Measurement of geo-neutrinos from 1353 days of Borexino". In: *Phys. Lett. B* 722.4 (2013).

[125] M. Agostini et al. (Borexino collaboration). "Spectroscopy of geoneutrinos from 2056 days of Borexino data". In: *Phys. Rev. D* 92 (3 2015).

[126] M. Agostini et al. (Borexino collaboration). "The Monte Carlo simulation of the Borexino detector". In: *Astroparticle Physics* 97 (2018).

[127] S. Agostinelli et al. (Geant4). "Geant4—a simulation toolkit". In: *Nucl. Instrum. Methods Phys. Res. A* 506.3 (2003).

[128] H. Back et al. (Borexino collaboration). "Borexino calibrations: hardware, methods, and results". In: *J. Instrum.* 7.10 (2012).

[129] R. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". In: *Comp. Phys. Comm.* 77.2 (1993).

[130] G. Cowan. *Statistical tests with weighted Monte Carlo events*. 2012. URL: https://www.pp.rhul.ac.uk/~cowan/stat/notes/weights.pdf.

[131] T. Glusenkamp. "A unified perspective on modified Poisson likelihoods for limited Monte Carlo data". In: (2019). arXiv: 1902.08831 [astro-ph].

[132] J. Bergström et al. "Updated determination of the solar neutrino fluxes from solar neutrino data". In: *JHEP* 2016.3 (2016).

[133] H. Daniel. "Das β-spektrum des RaE". In: *Nucl. Phys.* 31 (1962).

[134] M. Agostini et al. (Borexino collaboration). *Improved measurement of ^8B solar neutrinos with 1.5 kt y of Borexino exposure*. 2017. arXiv: 1709.00756 [physics.hep-ex].

[135] M. Agostini et al. (Borexino collaboration). "Comprehensive measurement of *pp*-chain solar Neutrinos". In: *Nature* 562 (2018).

[136] R. N. Saldanha. "Precision measurement of the ^7Be solar neutrino interaction rate in Borexino". PhD Thesis. Princeton University, 2012.

[137] I. Esteban et al. "Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity". In: *JHEP* 2017.1 (2017).

[138] P.C. de Holanda, Wei Liao, and A.Yu. Smirnov. "Toward precision measurements in solar neutrinos". In: *Nucl. Phys. B* 702.1 (2004).

[139] F.P. An et al. (Daya Bay Collaboration). "Measurement of electron antineutrino oscillation based on 1230 days of operation of the Daya Bay experiment". In: *Phys. Rev. D* 95 (7 2017).

[140] A. Gando et al. (KamLAND Collaboration). "Reactor on-off antineutrino measurement with KamLAND". In: *Phys. Rev. D* 88 (2013).

[141] M. Blennow et al. "Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering". In: *JHEP* 2014.3 (2014).

[142] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71.2 (2011).

[143] J.N. Bahcall. "The luminosity constraint on solar neutrino fluxes". In: *Phys. Rev. C* 65 (2 2002).

[144] F. Vissani. *Luminosity constraint and entangled solar neutrino signals*. 2018. arXiv: 1808.01495 [astro-ph.SR].

[145] C. Fröhlich and J. Lean. "The Sun's total irradiance: Cycles, trends and related climate change uncertainties since 1976". In: *Geophysical Research Letters* 25.23 (1998).

[146] J.N. Bahcall and C. Peña-Garay. "A road map to solar neutrino fluxes, neutrino oscillation parameters, and tests for new physics". In: *JHEP* 2003 (2003).

[147] A. Caldwell, D. Kollár, and K Kröninger. "BAT – The Bayesian analysis toolkit". In: *Comput. Phys. Commun.* 180.11 (2009).

[148] W. C. Haxton and A. M. Serenelli. "CN Cycle Solar Neutrinos and the Sun's Primordial Core Metallicity". In: *Astrophys. J.* 687.1 (2008).

[149] A. Serenelli, C. 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).

[150] M. Salaris and S. Cassisi. *Evolution of Stars and Stellar Populations*. Wiley, 2005.

[151] J.N. Bahcall. "Solar neutrino cross sections and nuclear beta decay". In: *Phys. Rev.* 135.1B (1964).

[152] A. Ianni. "Solar neutrinos". In: *Prog. Part. Nucl. Phys.* 94 (2017).

[153] F.L. Villante et al. "A step toward CNO solar neutrino detection in liquid scintillators". In: *Phys. Lett. B* 701.3 (2011).

[154] M. Agostini et al. (Borexino collaboration). *Modulations of the Cosmic Muon Signal in Ten Years of Borexino Data*. 2018. arXiv: 1808.04207 [physics.hep-ex].

[155] D. Bravo-Berguño et al. "The Borexino Thermal Monitoring & Management System and simulations of the fluid-dynamics of the Borexino detector under asymmetrical, changing boundary conditions". In: *Nucl. Instrum. Methods Phys. Res. A* 885 (2018).

[156] D. Bravo-Berguño. "Precision Background Stability and Response Calibration in Borexino: Prospects for Wideband, Precision Solar Neutrino Spectroscopy and BSM Neutrino Oscillometry Through a Deeper Detector Understanding". PhD Thesis. Virginia Polytechnic Institute, 2016.

[157] K. Cranmer. "Kernel estimation in high-energy physics". In: *Comp. Phys. Comm.* 136.3 (2001).

[158] I. Narsky and F.C. Porter. *Statistical Analysis Techniques in Particle Physics*. Wiley-Vch, 2014.

All of science is nothing more than the refinement of everyday thinking.

— Albert Einstein

ACKNOWLEDGEMENTS