

The Monte Carlo simulation of the Borexino detector

M. Agostini,¹ K. Altenmüller,² S. Appel,² V. Atroshchenko,³ Z. Bagdasarian,⁴ D. Basilico,⁵ G. Bellini,⁵ J. Benziger,⁶ D. Bick,⁷ G. Bonfini,⁸ L. Borodikhina,³ D. Bravo,^{9,5} B. Caccianiga,⁵ F. Calaprice,¹⁰ A. Caminata,¹¹ S. Caprioli,⁵ M. Carlini,⁸ P. Cavalcante,^{8,9} A. Chepurinov,¹² K. Choi,¹³ D. D'Angelo,⁵ S. Davini,¹¹ A. Derbin,¹⁴ X.F. Ding,¹ L. Di Noto,¹¹ I. Drachnev,^{1,14} K. Fomenko,¹⁵ A. Formozov,⁵ D. Franco,¹⁶ F. Froberg,¹⁰ F. Gabriele,⁸ C. Galbiati,¹⁰ C. Ghiano,¹¹ M. Giammarchi,⁵ M. Goeger-Neff,² A. Goretti,¹⁰ M. Gromov,¹² C. Hagner,⁷ T. Houdy,¹⁶ E. Hungerford,¹⁷ Aldo Ianni,⁸ Andrea Ianni,¹⁰ A. Jany,¹⁸ D. Jeschke,² V. Kobychyev,¹⁹ D. Korabiev,¹⁵ G. Korga,¹⁷ D. Kryn,¹⁶ M. Laubenstein,⁸ E. Litvinovich,^{3,20} F. Lombardi[†],⁸ P. Lombardi,⁵ L. Ludhova,^{4,21} G. Lukyanchenko,³ I. Machulin,^{3,20} G. Manuzio,¹¹ S. Marcocci,^{1,11} J. Martyn,²² E. Meroni,⁵ M. Meyer,⁷ L. Miramonti,⁵ M. Misiaszek,¹⁸ V. Muratova,¹⁴ B. Neumair,² L. Oberauer,² B. Opitz,⁷ F. Ortica,²³ M. Pallavicini,¹¹ L. Papp,² A. Pocar,²⁴ G. Ranucci,⁵ A. Re,⁵ A. Romani,²³ R. Roncin,^{8,16} N. Rossi,⁸ S. Schönert,² D. Semenov,¹⁴ P. Shakina,¹⁴ M. Skorokhvatov,^{3,20} O. Smirnov,¹⁵ A. Sotnikov,¹⁵ L.F.F. Stokes,⁸ Y. Suvorov,^{25,3} R. Tartaglia,⁸ G. Testera,¹¹ J. Thurn,²⁶ M. Toropova,³ E. Unzhakov,¹⁴ A. Vishneva,¹⁵ R.B. Vogelaar,⁹ F. von Feilitzsch,² H. Wang,²⁵ S. Weinz,²² M. Wojcik,¹⁸ M. Wurm,²² Z. Yokley,⁹ O. Zaimidoroga,¹⁵ S. Zavatarelli,¹¹ K. Zuber,²⁶ and G. Zuzel¹⁸

(Borexino Collaboration*)

¹Gran Sasso Science Institute (INFN), 67100 L'Aquila, Italy

²Physik-Department and Excellence Cluster Universe, Technische Universität München, 85748 Garching, Germany

³National Research Centre Kurchatov Institute, 123182 Moscow, Russia

⁴IKP-2 Forschungszentrum Jülich, 52428 Jülich, Germany

⁵Dipartimento di Fisica, Università degli Studi e INFN, 20133 Milan, Italy

⁶Chemical Engineering Department, Princeton University, Princeton, NJ 08544, USA

⁷Institut für Experimentalphysik, Universität Hamburg, 22761 Hamburg, Germany

⁸INFN Laboratori Nazionali del Gran Sasso, 67010 Assergi (AQ), Italy

⁹Physics Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

¹⁰Physics Department, Princeton University, Princeton, NJ 08544, USA

¹¹Dipartimento di Fisica, Università degli Studi e INFN, 16146 Genova, Italy

¹²Lomonosov Moscow State University Skobeltsyn Institute of Nuclear Physics, 119234 Moscow, Russia

¹³Department of Physics and Astronomy, University of Hawaii, Honolulu, HI 96822, USA

¹⁴St. Petersburg Nuclear Physics Institute NRC Kurchatov Institute, 188350 Gatchina, Russia

¹⁵Joint Institute for Nuclear Research, 141980 Dubna, Russia

¹⁶AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/IRFU, Observatoire de Paris, Sorbonne Paris Cité, 75205 Paris Cedex 13, France

¹⁷Department of Physics, University of Houston, Houston, TX 77204, USA

¹⁸M. Smoluchowski Institute of Physics, Jagiellonian University, 30059 Krakow, Poland

¹⁹Kiev Institute for Nuclear Research, 03680 Kiev, Ukraine

²⁰National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia

²¹RWTH Aachen University, 52062 Aachen, Germany

²²Institute of Physics and Excellence Cluster PRISMA, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

²³Dipartimento di Chimica, Biologia e Biotechnologie, Università degli Studi e INFN, 06123 Perugia, Italy

²⁴Amherst Center for Fundamental Interactions and Physics

Department, University of Massachusetts, Amherst, MA 01003, USA

²⁵Physics and Astronomy Department, University of California Los Angeles (UCLA), Los Angeles, California 90095, USA

²⁶Department of Physics, Technische Universität Dresden, 01062 Dresden, Germany

We describe the Monte Carlo (MC) simulation package of the Borexino detector and discuss the agreement of its output with data. The Borexino MC “ab initio” simulates the energy loss of particles in all detector components and generates the resulting scintillation photons and their propagation within the liquid scintillator volume. The simulation accounts for absorption, reemission, and scattering of the optical photons and tracks them until they either are absorbed or reach the photocathode of one of the photomultiplier tubes. Photon detection is followed by a comprehensive simulation of the readout electronics response. The algorithm proceeds with a detailed simulation

[†] Also at: Laboratorio Subterráneo de Canfranc, Paseo de los Ayerbe S/N, 22880 Canfranc Estacion Huesca, Spain

[‡] Present address: Physics Department, University of California, San Diego, CA 92093, USA

* Corresponding author: spokeperson-borex@lngs.infn.it

of the electronics chain. The MC is tuned using data collected with radioactive calibration sources deployed inside and around the scintillator volume. The simulation reproduces the energy response of the detector, its uniformity within the fiducial scintillator volume relevant to neutrino physics, and the time distribution of detected photons to better than 1% between 100 keV and several MeV. The techniques developed to simulate the Borexino detector and their level of refinement are of possible interest to the neutrino community, especially for current and future large-volume liquid scintillator experiments such as Kamland-Zen, SNO+, and Juno.

PACS numbers: 02.70.Uu, 14.60.Lm, 26.65.+t, 29.40.Mc, 78.70.Ps, 96.60.Jw

I. INTRODUCTION

Large volume, liquid scintillator detectors have contributed greatly to neutrino physics in recent years. Starting twenty-five years ago, Borexino pioneered the development of ultra-low radioactive background detectors that could push the energy threshold down to 100 keV allowing the measurement of most of the solar neutrino spectrum [1]. A wide variety of liquid scintillator detectors, listed in Table I, are either running or being designed for measurements of neutrino oscillations, for neutrino-less double beta decay searches, or for use as active γ -ray and neutron veto devices for direct WIMP dark matter experiments [2].

The accurate modeling by means of Monte Carlo (MC) methods of the entire detection chain is essential to extract quality data from running detectors and to inform the design of future experiments. The simulation needs to include the energy lost by primary particles, the generation and transport of scintillation light in the large volume of liquid scintillator and its detection by photo-multipliers, and the detector response.

In this paper we describe the MC modeling of the Borexino experiment and we compare its output to data. The methods and the results presented here were developed and specifically optimized for the Borexino detector and scintillator mixture, but they can be easily adapted to other liquid scintillators and detector geometries.

A collection of experiments employing large liquid scintillator detectors is shown in Table I. A feature common to all of them is the focus on events of relatively low energy, up to a few tens of MeV. In this energy range it is possible to perform detailed simulations that generate and track individual optical photons emitted produced in scintillation events, with typical photon yield of $\simeq 10^4$ photons/MeV. The MC model can in turn very accurately reproduce physical events with minimal use of effective parameters, provided that all the physics processes are included and realistically described (e.g. light emission, absorption, reemission, scattering, reflection, ...).

Borexino is an un-segmented calorimeter consisting of ~ 278 t of organic liquid scintillator [2]. The detector has been in continuous operation since May 2007 at the underground Laboratori Nazionali del Gran Sasso (LNGS) in Italy. It is designed to measure the low energy portion of the solar neutrino spectrum (see the top left panel of Fig. 1) and, in particular, the 0.862 MeV ${}^7\text{Be}$ solar

neutrinos, whose interaction rate was measured with 5% precision [3] with no observed day-night rate asymmetry [4].

The extremely low levels of radioactivity achieved within the scintillator allowed Borexino to broaden its science reach beyond the original design goal. Borexino reported first evidence for the observation of the 1.44 MeV *pep* neutrinos and holds the current best upper limit on the solar CNO flux [5]. At higher end of the solar neutrino spectrum, the electron recoil induced by ${}^8\text{B}$ neutrinos was measured with a record-low threshold of 3 MeV [6]. A scintillator purification campaign performed ~ 3 years into data taking, which further reduced radioactive contamination, allowed Borexino to measure the spectrum of *pp* solar neutrinos (with end-point energy of 0.42 MeV) [7]. Further physics results of Borexino include the detection of geo-neutrinos in the range between 1.8 and 3 MeV [8–10], searches for anti-neutrinos from astrophysical sources up to 15 MeV [11, 12], searches for solar axions at 5.5 MeV [13], as well as other exotic searches [14–16]. The investigation of short baseline anti-neutrino oscillations into light sterile neutrinos using an artificial ${}^{144}\text{Ce}$ - ${}^{144}\text{Pr}$ source is planned for the near future within the SOX project [17].

Figure 1 provides an overview of the signals of interest in Borexino, their energy ranges, and most relevant backgrounds. The Borexino data analysis calls for an accurate and complete understanding of the detector response over these energy ranges. This is achieved through meticulous energy calibration and the understanding of the degree of uniformity of the energy response of the detector over its volume, of the trigger efficiency, of the detector stability over time, and of the spatial reconstruction and time response for events of different type occurring at different positions, including the scintillation pulse shape characteristics. The Borexino MC simulation reproduces physics and calibrations data with 1% precision or better for all quantities relevant for experiment's physics program.

Short descriptions of the Borexino detector and of the calibration hardware are presented in Sec. II and Sec. III, respectively. We describe the MC package and the physical models it implements in Sec. IV, Sec. V, Sec. VI, and Sec. VII). In Sec. VIII we show how we tuned the simulation parameters with the aid of laboratory measurements, source calibration data, and physics-mode data. Finally, we compare MC results and data in Sec. IX.

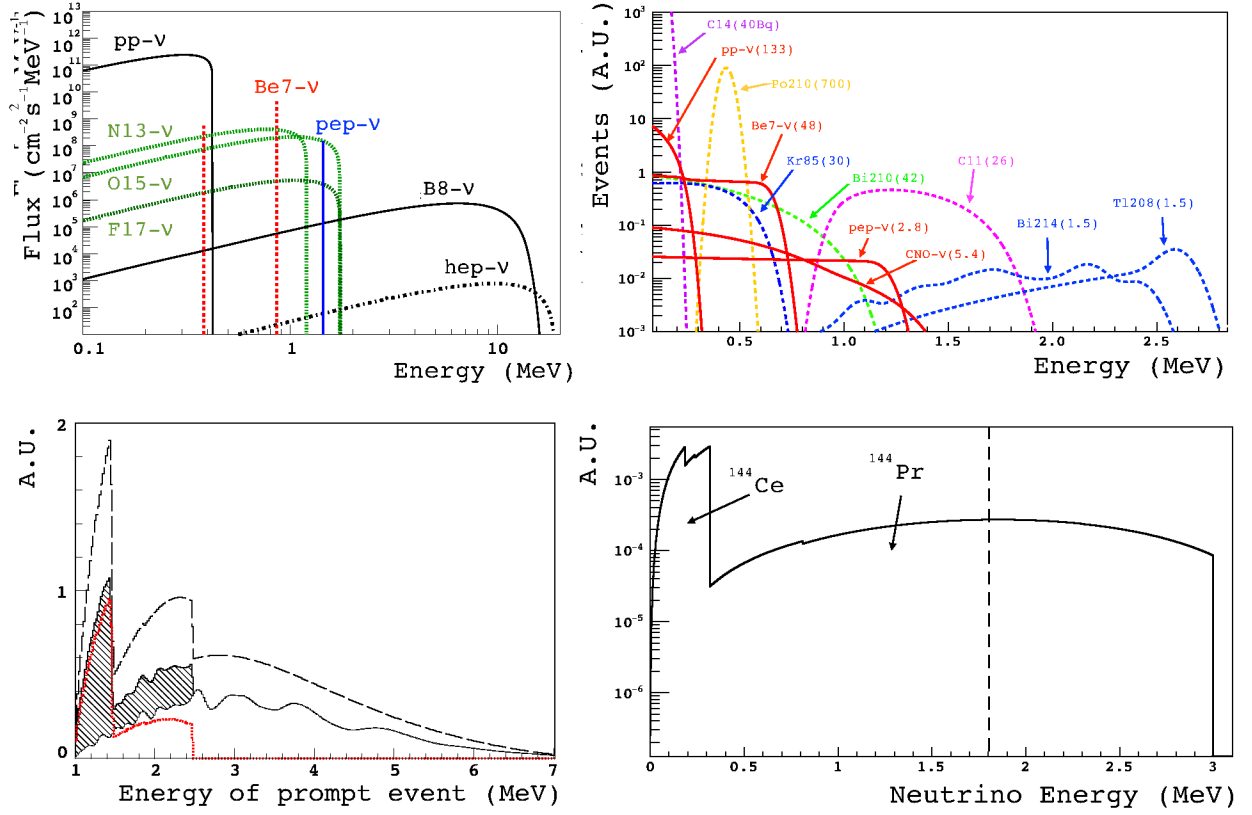


FIG. 1. Sources of signals in Borexino. *Top left*: expected spectrum of solar neutrinos [18]. The flux (vertical scale) is given in $\text{cm}^{-2} \text{s}^{-1}$ for monochromatic lines. *Top right*: energy spectrum of electron recoils induced by the interaction of solar neutrinos in Borexino together with the spectral shapes of some known radioactive contaminants in the scintillator. The numbers shown in parenthesis are in counts per day per 100 ton for all species apart from ^{14}C , for which Bq per 100 ton is used. *Bottom left*: expected spectrum of electron anti-neutrinos in Borexino. The horizontal axis shows the kinetic plus annihilation (1.022 MeV) energy of the prompt positron event. Dashed line: total geo- $\bar{\nu}_e$ plus reactor- $\bar{\nu}_e$ spectrum without oscillations. Solid thick lines: geo- $\bar{\nu}_e$ and reactor- $\bar{\nu}_e$ with oscillations. Dotted line (red): geo- $\bar{\nu}_e$ with the high (low) energy peak due to decays in the ^{238}U chain (^{238}U and ^{232}Th chains). Solid thin line: reactor- $\bar{\nu}_e$. *Bottom right*: energy spectrum of the emitted $\bar{\nu}_e$ in the decay of the ^{144}Ce - ^{144}Pr for the SOX experiment [17]. Only the portion of the $\bar{\nu}_e$ spectrum above 1.8 MeV (dashed vertical line) can be detected via inverse β decay on protons.

II. SHORT DESCRIPTION OF THE BOREXINO DETECTOR

The Borexino detector is schematically depicted in Fig. 2 and widely described in Ref. [24]. The inner detector is enclosed by a stainless steel sphere (SSS) that serves both as the container of the scintillator and buffer liquids and as the mechanical support of the photomultipliers (PMTs). Within this sphere, two nylon vessels (0.125 mm thick) separate the volume in three shells of radii 4.25 m, 5.50 m and 6.85 m, the latter being the radius of the SSS itself. The inner nylon vessel contains the liquid scintillator solution ($\sim 278 \text{ t}$ or $\sim 300 \text{ m}^3$), namely PC (pseudocumene, 1,2,4-trimethylbenzene $\text{C}_6\text{H}_3(\text{CH}_3)_3$) as a solvent and the fluor PPO (2,5-diphenyloxazole, $\text{C}_{15}\text{H}_{11}\text{NO}$) as a solute at a concentration of 1.5 g/l (0.17 % by weight). The second and the third shells (buffer regions) contain PC with

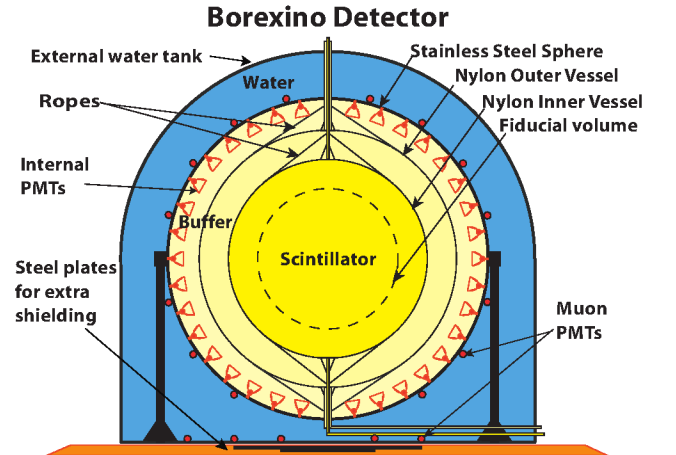


FIG. 2. Schematic drawing of the Borexino detector.

Experiment	Mass	Physics investigation	Status	Reference
Chooz	5 t LS+Gd (0.1%) + 107 t LS	ν oscillations	past	[19]
KamLAND	1 kt LS	ν oscillations	past	[20]
Karmen	56 t +Gd foils	ν oscillations	past	[21]
LSND	167 t LS	ν oscillations	past	[22]
Palo Verde	11 t LS+Gd (0.1%)	ν oscillations	past	[23]
Borexino	278 t LS	ν oscillations	ongoing	[24]
Daya Bay	20 t LS+Gd (0.1%) + 20 t LS	ν oscillations	ongoing	[25]
Double Chooz	8 t LS+Gd (0.1%) + 18 t+80 t LS	ν oscillations	ongoing	[26]
Reno	16 t LS+Gd (0.1%) + 30 t LS	ν oscillations	ongoing	[27]
LENS	125 t LS+In (8%)	ν oscillations	future	[28]
Juno	20 kt LS	ν oscillations	future	[29]
Reno-50	18 kt LS	ν oscillations	future	[30]
KamLAND-Zen	13 t LS+Xe (2.9%) + 1 kt LS	$0\nu\beta\beta$ decay	ongoing	[31]
SNO+	780 t LS+Te (0.5%)	$0\nu\beta\beta$ decay	commissioning	[32]
LVD	1.8 kt LS	SN ν	ongoing	[33]
Nucifer	0.75 t LS+Gd (0.2%)	reactor monitoring	ongoing	[34]
Neos	1 t LS+Gd (0.5%)	sterile ν	ongoing	[35]
Neutrino-4	0.35 t LS+Gd (0.1%)	sterile ν	ongoing	[36]
Prospect	3–13 t LS+ ^6Li	sterile ν	ongoing	[37]
Stereo	1.8 t LS+Gd (0.2%)	sterile ν	commissioning	[38]
SOX	278 t LS	sterile ν	future	[17]
Dark Side-50	30 t LS+TMB (5%)	DM veto	ongoing	[39]
Dark Side-20k	~ 250 t LS+TMB (20%)	DM veto	future	[40]
LZ	20.8 t LS+Gd (0.1%)	DM veto	future	[41]
SABRE	2 t LS	DM veto	future	[42]

TABLE I. Compilation of past, present and future liquid scintillator based experiments excluding Borexino. In the table, “LS” stands for liquid scintillator, “Gd” (gadolinium) and “TMB” (trimethylborate) are used for enhancing neutron captures, “In” is the chemical element indium, “DM” signifies dark matter, “SN” stands for supernova and $0\nu\beta\beta$ for neutrino-less double beta decay.

a small amount (5 g/l)¹ of DMP (dimethylphthalate, $\text{C}_6\text{H}_4(\text{COOCH}_3)_2$) as a light quencher added to further reduce the scintillation yield of pure PC [43].

The 2212 internal PMTs (8” ETL 9351, formerly Thorn EMI) mounted on the inner side of the SSS detect the scintillation light. In order to enhance the photon detection efficiency, 1828 PMTs out of the total 2212 were equipped with aluminum optical concentrators [44]. Their field of view is designed to collect all scintillation photons incident under an angle of less than $\delta_{\text{max}} = 44^\circ$, corresponding to scintillator signals from within the inner vessel. Photons hitting the concentrator surfaces with angles greater than δ_{max} are reflected back into the detector.

A 18 m in diameter, 16.9 m height domed cylinder filled

by ultra-pure water contains the SSS and acts both as radiation shielding and as Čerenkov detector (outer detector) for identifying and vetoing cosmic muons. For this purpose, 208 additional PMTs are mounted on the outer side of the SSS and on the water tank floor. A detailed description of the detector, of the electronics, and of the purification plants used to prepare the scintillator and fill the detector can be found in Ref. [24] and Ref. [45]. The muon detector design and performances are detailed in Ref. [46].

Solar neutrinos are detected through their elastic scattering on electrons. The measurement of different solar neutrino components is possible through a fit of the electron recoil energy spectrum (see the top right panel of Fig. 1), aiming at disentangling the contribution of solar neutrinos and that of background signals [47].

Anti-neutrinos ($\bar{\nu}_e$) are detected via inverse β decay:

$$\bar{\nu}_e + p \rightarrow e^+ + n, \quad (1)$$

with a threshold of 1.806 MeV. The positron promptly comes to rest in the liquid scintillator and annihilates

¹ The DMP concentration was 5 g/l at the beginning of data taking. After the discovery of a tiny leak in the inner vessel, the concentration was reduced to 2.8 g/l in order to decrease the pressure difference between the inner vessel and the buffer.

Isotope	Type	Energy (keV)
^{57}Co	γ	122
^{139}Ce	γ	165
^{203}Hg	γ	279
^{85}Sr	γ	514
^{54}Mn	γ	834
^{65}Zn	γ	1115
^{60}Co	γ	1173 - 1332
^{40}K	γ	1460
^{222}Rn	α/β	$0 \div 3200$
^{14}C	β	$0 \div 156$
$^{241}\text{Am}-^9\text{Be}$	n	$\sim 0 \div 10000$
Ext. ^{228}Th	γ	2615

TABLE II. Radioactive isotopes used for Borexino calibration. The $^{241}\text{Am}-^9\text{Be}$ source allows to study the thermalization of neutrons in the scintillator and the neutron captures on H and C.

emitting two 511 keV γ -rays, yielding a prompt event with a visible energy of $E_{\text{prompt}} = E_{\bar{\nu}_e} - 0.782 \text{ MeV}$. The free emitted neutron is typically captured on protons within a mean time $\tau \sim 256 \mu\text{s}$ [46], resulting then in the emission of a 2.22 MeV de-excitation γ -ray, which provides a coincident delayed event. The characteristic time and spatial coincidences of prompt and delayed events offer a clean signature of the $\bar{\nu}_e$ detection.

III. THE SOURCE CALIBRATION CAMPAIGNS

A series of calibration campaigns based on various types of radioactive sources inserted directly into the detector volume were performed in November 2008, January 2009, and June-July 2009. In July 2010, an external ^{228}Th γ -ray source was positioned in dedicated pipes close to the SSS in order to study the response to the external γ -rays. The hardware used for a safe, airtight, clean, and accurate deployment of small radioactive sources in several locations within the scintillator target is described in Ref. [48], while the design, construction and performance of the ^{228}Th source can be found in Ref. [49].

Calibration data have been essential in validating the physics model adopted for the description of the scintillation light emission, propagation, and detection by PMTs. The results are in full agreement with observations in the Counting Test Facility, the small 4 t Borexino prototype [50].

The radioactive sources deployed inside the scintillator volume (see Table II) were selected to study the detector response to α and β particles, γ -rays, and neutrons in a wide energy region from 122 keV and $\sim 10 \text{ MeV}$.

More specifically, the goals of the Borexino calibration campaigns included:

- the measurement of the position reconstruction accuracy and resolution for events distributed in the whole inner vessel and over the energy range.
- the calibration of the absolute energy scale and resolution especially in the energy region of interest for studying ^7Be , pep , and pp solar neutrinos interactions (i. e. $\sim 100 \text{ keV} \div \sim 1.5 \text{ MeV}$).
- the measurement of the non-uniformity of the energy response as a function of the event position and energy.
- the production of signals mimicking the external background.

Characterizing the position reconstruction is of high importance for all Borexino analyses. It permits to define a fiducial volume in which the background from the radioactive contaminants on the vessel surfaces or γ -rays originating from the outer parts of the detector are minimized. The optimal choice of the fiducial volume depends on the type of analysis performed, as explained in Ref. [47]. In addition, the accuracy of the reconstruction of the scintillation vertex determines the precision by which the target volume is defined, thus directly affecting the uncertainty of the measurement of absolute neutrino or anti-neutrino fluxes. In SOX, the search for the short-baseline oscillation pattern possibly induced by sterile neutrinos relies heavily on the accurate spatial reconstruction of the inverse β decay events.

Spectral distortions in the energy response, due to effects like the ionization quenching or the dependence of the light collection efficiency on the vertex position, are relevant for the analysis. A correct understanding of such effects is of primary importance for spectroscopic measurements, in particular the precision measurement of the ^7Be solar neutrino interaction rate, which relies on an absolute energy calibration on a level better than 1%. Finally, an accurate understanding of the external background is fundamental for the detection of pep , CNO, and ^8B neutrinos in Borexino.

The source vials were carefully designed to mitigate the risk of introducing unwanted contaminations into the Borexino scintillator target. Spherical quartz vessels (1" diameter) were either filled by ^{222}Rn loaded liquid scintillator (identical to the Borexino one) or γ -ray emitters in aqueous solution. The sources were deployed along the polar axis of the inner vessel and in several positions off axis. An optical system consisting of a LED mounted on the source support and of a series of cameras allowed to measure the position of the source within the inner vessel with $\sim 2 \text{ cm}$ accuracy [48].

A dedicated procedure was developed for loading scintillator inside the vials with ^{222}Rn , minimizing the quenching due to oxygen contamination. The comparison of the ^{214}Po energy peak from within the source with

the ^{214}Bi - ^{214}Po tagged coincidences from contaminants dissolved in the detector gave evidence of small but non-negligible light quenching. Fast coincidence decays from the ^{222}Rn source were used to study the accuracy of the position reconstruction algorithm and to characterize the energy response uniformity in the detector volume.

Pure beta sources could not be introduced directly in the detector because of the high risk of contaminating the scintillator. Instead, beta emitters were dissolved in a non-scintillating solution, for suppressing the pure β component and allowing, at the same time, γ -rays to escape the vial and to induce electrons in the scintillator by Compton scattering or by photoelectric effect. The γ -rays provided the absolute energy scale calibration over the whole region of interest.

Moreover, we used a commercial ^{241}Am - ^9Be neutron source inserted in a properly designed shielding. The neutron scattering on protons during thermalization provided a calibration for proton ionization quenching. The γ -rays emitted by the nuclei capturing neutrons (mostly hydrogen and ^{12}C in the scintillator or in the materials of the source shield) allowed calibration signals in the highest energy range.

IV. THE MONTE CARLO SIMULATION CODE OF THE BOREXINO DETECTOR

Particles depositing energy in the inner vessel or in the buffer regions produce scintillation and Čerenkov photons, which propagate inside the detector and possibly reach the PMTs, yielding a detectable signal. The agreement between measured observables (energy estimators and PMT pulse times [47]) and the physical quantities (deposited energy, position, type of particles generating the signal) depends on the knowledge and understanding of all the physical processes governing the particle energy loss in the various materials, the scintillator light production, propagation, and detection. Besides, it depends on the characteristics of the electronics and of the triggering system. The Borexino MC simulation was designed and optimized to fully model and reproduce all these processes up to the signal detection.

The MC simulation chain consists of a set of numerical codes that:

1. provide a wide range of event generators, from solar neutrino interactions, to radioactive decays, geo-neutrinos, and calibration source events.
2. simulate the energy loss of each specific particle in every material present in the detector, either active (the scintillator, buffer liquid, and water in the muon detector) or passive.
3. generate a number of scintillation or Čerenkov photons considering the particle energy loss in the media and the properties of the scintillator and/or the buffer.

4. track each single optical photon including its interactions with the scintillator and with the materials, until a PMT is reached or the photon is absorbed.
5. generate the PMT response for photons absorbed at the PMT cathode, considering the quantum efficiency of each individual PMT.
6. generate the PMT pulse signal taking into account the specific design of the front end and of the digital electronics chain of Borexino.
7. simulate the trigger generation and save the final output for triggering events.
8. produce a raw data file with the same structure as the one produced by the Borexino data acquisition system.

Technically, the code is structured in three separate programs, which have to be chained:

- the event generation and light tracking (described in Sec. V), accomplishing the tasks 1 to 5.
- the simulation of the electronics response (discussed in Sec. VI), realizing the items 6 to 8.
- the event reconstruction, that converts the binary data generated by the DAQ into physical observables, such as number of photoelectrons, event position, pulse shape variables. This code is the same for both real and MC data. Its short description can be found in Sec. VII.

The code developments were driven by the constant comparison of simulations with calibration and real data.

V. EVENT GENERATION AND LIGHT TRACKING

The event generation is implemented within the Geant4 package and uses the standard libraries therein [51]. Geant4 is an object oriented C++ toolkit for the simulation of the passage of particles through matter.

A high precision simulation depends on a careful description of the system geometry and construction materials. The rich geometry tools of the package allow to obtain the necessary accurate description of the elements of the detector and of the physical properties of solid components, liquids, and light sensors.

A. Outer detector geometry

The Geant4 implementation of the outer detector geometry is shown in Fig. 3. Important elements are the legs supporting the stainless steel sphere (SSS) and the steel platforms at the bottom of the detector, which were inserted as shielding against the rock radioactivity [24].

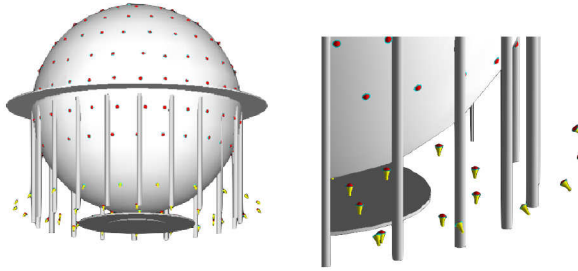


FIG. 3. Simulated geometry for the Borexino outer detector in Geant4.



FIG. 4. Comparison between the schematization of the outer detector PMT in the simulation (left) and the real object (right).

The simulation precisely follows the detector geometry. PMTs are placed on the floor of the water tank and on the outer surface of the SSS. The schematization of the outer detector PMT geometry directly follows the real design, as it is shown in Fig. 4. PMTs are enclosed in an outer shielding against water and pressure, and this is reproduced in the simulation. For PMTs attached to the SSS, reflective tyvek foils cover their surface leaving free only the photocathodes facing the water. The tyvek foils placed on the inner surface of the water tank are also simulated.

Besides the geometrical details implemented in the simulation, the optical properties of all the materials involved are taken into account according to specific measurements or data available in the literature. In the specific case of the outer detector simulation, the tyvek reflectivity dependence upon the wavelength, the PMT quantum efficiencies, and the water properties are of particular importance.

B. Inner detector geometry

Figure 5 shows a cross section of the Borexino inner detector as implemented in the simulation. All the most relevant geometrical features are included, such as the PMTs with their real shapes and positions, the nylon vessels, and the holding endcaps. Special tools were added to consider the real shape of the inner vessel and its time dependent shape that deviates from a perfect sphere. As

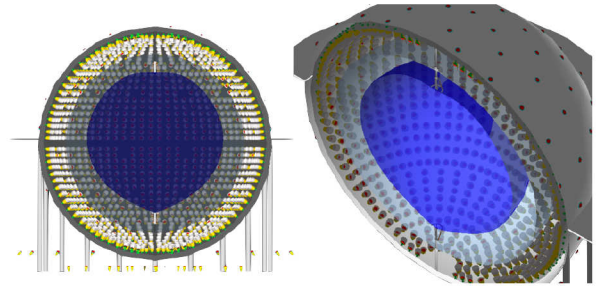


FIG. 5. Simulated geometry for the Borexino inner detector in Geant4. The realistic inner vessel shape (dark blue) is updated on a weekly basis.



FIG. 6. *Left Panel:* Real Borexino inner detector PMTs mounted on the stainless steel sphere. *Right Panel:* MC schematization of the PMTs with (top) and without (bottom) light concentrators.

described in Ref. [47], the inner vessel profile is determined on a weekly basis by using background events from vessel contamination and external γ -rays. The $r - \theta$ distribution of these events is fitted and used as input of the MC simulation for the proper simulation of the vessel shape. Therefore, when simulating data throughout a long period of time, the vessel evolution is taken into account. This is important, since the vessel shape can affect both the amount of external contaminations reaching the standard fiducial volume (FV) and the energy response of the detector (see e.g. Sec. VIII A and Sec. IX C). The granularity with which the real vessel shape is simulated can be adjusted by the user with optimal values usually around ~ 5 cm (i.e. the shape is approximated with a polygonal shape in the $r - \theta$ plane with sides of ~ 5 cm). This is a good tradeoff between the intrinsic uncertainty of the vessel shape determination (a few cm [47]) and an optimized code performance.

In Fig. 5 the nylon vessel endcaps are also visible. They support the vessels and provide the connection between the innermost volumes and the Borexino purification plants. A careful simulation of the endcaps is important, since they induce light shadowing and reflections, which affect the reconstruction for events close to the north and the south poles of the inner vessel.

The key components of the Borexino geometry simulation are the PMTs of the inner detector. Their geometry, materials, and optical properties highly impact the detector response. In Fig. 6, a comparison between the real PMTs mounted on the inner surface of the SSS and the MC modeling is shown. As anticipated in Sec. II, 1828 PMTs are equipped with light concentrators, while 384 are not. The two different configurations are shown both in the left and right panels of Fig. 6. PMTs with concentrators are simulated as shown at the top of the right panel of Fig. 6, while PMTs without concentrators are simulated as shown at the bottom of the same panel. The schematization of the inner detector PMTs is made of a common PMT base (almost invisible, but in green in the right panel of Fig. 6) on top of which the photocathode is placed (red in the picture). The photocathode is made of bialkali and it is a spherical cap. The whole PMT body is surrounded by a conical μ -metal shielding (yellow in the picture), which reduces the effect of the Earth magnetic field on the PMT response.

The light concentrator shape is reproduced in the simulation with sub-cm precision, since it significantly affects the amount of light collected as a function of the event position. PMTs without concentrators are equipped with a small steel ring which surrounds the photocathode and supports it. Figure 7 shows the wavelength dependence of the aluminum reflectivity of light concentrators as implemented in the simulation.

Particular care was devoted to the implementation of the optical properties of the simulated materials in the inner detector, since they are the most relevant for solar neutrino analyses. Apart from the scintillator and buffer properties, which are discussed in more detail in Sec. VE, the nylon vessel absorption length as a function of wavelength and its index of refraction are also included in the simulation. The vessel transparency was measured with a dedicated campaign carried out at the time of the vessel construction. The PMT quantum efficiency spectral dependence was implemented according to the specifics provided by the manufacturer [24] and it

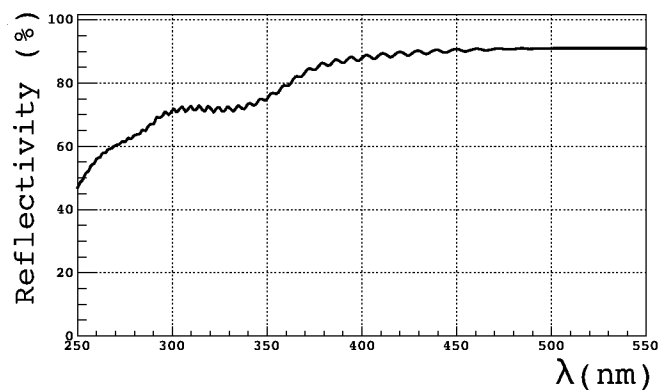


FIG. 7. Reflectivity of polished aluminum light guides as a function of photon wavelength [44].

is shown in Fig. 8. Furthermore, the light concentrators and the SSS inner surface are modeled with a finite probability that the incoming light is reflected or absorbed. In case of reflection, the photon can be reflected specularly or diffusively (Lambertian diffusion). The same holds for the steel rings of the PMTs not equipped with cones, since they act similarly to light concentrators albeit at a reduced efficiency. The precise determination of all these coefficients is performed through the tuning procedure outlined in Sec. VIII.

C. The particle energy loss

The Geant4 libraries allow to simulate radioactive decay chains and energy losses of every type of particle (electrons, γ -rays positrons, neutrons, α -particles...) down to 100 eV in each material of the detector. For all charged particles, the physics of interactions and propagation is completely handled through Geant4 libraries. Special modules developed within the collaboration generate scintillation and Čerenkov photons according to the particle type and to the models of light generation and propagation described in Sec. VE.

1. Positron annihilation

Positrons slow down in the scintillator and a fraction of them are captured by electrons before annihilation, forming para- or ortho-positronium. In the former case, the emission of two 511 keV γ -rays is almost immediate, while in the latter it is delayed by a few ns. The ortho-positronium lifetime in the PC+PPO mixture was measured in Ref. [52] and is ~ 3 ns, while the ortho-positronium formation probability was found to be close to 50%. The ortho-positronium lifetime is comparable to the fastest scintillation time constant ($\simeq 4$ ns, see Sec. VIII B) and therefore its formation has a significant impact on the light pulse shape. This phenomenon can be used to identify and subtract β^+ radioactive backgrounds [47]. We developed a custom physics module that handles this effect. According to the measured formation probability, we model the positron energy loss and either the para- or ortho-positronium formation and decay. The two processes are simulated as three-body vertexes, composed by the positron track and two delayed annihilation γ decay. The difference between the former and the latter is that the annihilation happens immediately after the positron is stopped. Ortho-positronium in vacuum decays emitting three γ -rays while in the liquid scintillator it mainly annihilates emitting only two γ -rays. This is due to the *pick-off* effect, for which the positron of an ortho-positronium state annihilates with an electron from the surrounding scintillator. The delay of the γ -ray generation is described by a simple exponential law depending on the ortho-positronium mean life in the scintillator.

D. Primary event generators

Several generators were developed to properly simulate radioactive decays inside the scintillator, solar neutrinos, radioactive sources encapsulated as in the calibration campaign, the ^{241}Am - ^9Be source, and anti-neutrino interactions.

1. Radioactive decays

Standard Geant4 classes manage radioactive decays providing the correct daughter spectra and branching ratios for most of the radionuclides. The radioactive decays which we treat differently from the Geant4 standard approach are those of ^{14}C and ^{210}Bi .

The beta decay of ^{14}C into ^{14}N is an allowed ground-state-to-ground-state transition. This decay was investigated both theoretically and experimentally by many groups, but some unsatisfactory features remain. For instance, its anomalously long half-life (~ 5730 years) with respect to “standard” beta decays has been subject of considerable interest. In addition, there have been different experimental investigations aiming to assess the deviations from the expected allowed decay spectrum [53]. We developed a generator which allows to simulate the ^{14}C beta decay spectrum with a shape factor, i. e. a quantification of the deviation from the allowed shape, either from Ref. [53] or from Ref. [54].

The ^{210}Bi decay is a first-forbidden beta decay and thus its spectral shape cannot be predicted in a straightforward way. In our simulation, the ^{210}Bi event generation is handled in such a way that the shape factor can be modified, in order to use the differences between the various models to evaluate systematic uncertainties. The standard spectral shape which we use is that reported in Ref. [55]. An additional spectral shape is the one proposed in Ref. [56], which differs by a few % from the standard one.

^{11}C is simulated considering the probability of ortho-positronium formation as discussed in Sec. VC.

2. Solar neutrinos

We developed a custom generator for simulating solar neutrino elastic scattering off electrons. The primary solar neutrino energy spectra are those computed in Ref. [18, 57]. The electron solar neutrino survival probability is computed according to Ref. [58] and the mixing parameters are those reported in Ref. [59]. The $\nu_e - e$ and $\nu_{\mu/\tau} - e$ elastic scattering cross sections as a function of energy are computed following Ref. [60]. Finally, the electron recoil energy is sampled including radiative corrections at next-to-leading order as calculated in Refs. [60–62].

3. ^{241}Am - ^9Be neutron source

The ^{241}Am - ^9Be neutron source provided calibration events up to about 10 MeV, due to neutron captures on different nuclei. Neutrons are mainly produced by the following reactions: $^9\text{Be}(\alpha, n)^{12}\text{C}_{\text{gs}}$ and $^9\text{Be}(\alpha, n)^{12}\text{C}^*$. The output neutron energy spectra are described in Refs. [63, 64]. They are used for simulating the primary neutrons emitted by the source. The $^{12}\text{C}^*$ de-excitation through the emission of 4.44 MeV γ -ray is simulated as well. The neutron scattering during the thermalization and its subsequent capture by protons or nuclei are directly managed by the Geant4 libraries as well as the de-excitation of the daughter nucleus.

4. Geo and reactor anti-neutrinos

Geo-neutrino energy spectra are computed from the β^- spectra of the ^{238}U and ^{232}Th decay chains. Reactor anti-neutrino spectra are obtained from Ref. [65]. Since anti-neutrinos are detected via inverse beta decay, the primary vertex is simulated as a positron-neutron pair emitted simultaneously from the same vertex.

5. SOX anti- ν source

A custom generator was developed to simulate the SOX anti-neutrino source [66]. Anti-neutrinos emitted by ^{144}Ce - ^{144}Pr are generated according to Fermi’s theory of β decay, including corrections for the finite size and mass of the nucleus, weak interaction finite size corrections, radiative corrections and screening corrections. Further studies are currently under way to measure the shape factor of the forbidden beta decay.

6. External background

As described in Ref. [67], the SSS, the PMTs, and the light concentrators are contaminated by non-negligible amounts of ^{232}Th , ^{238}U , and ^{40}K . The long-ranged γ -rays from the daughter nuclides ^{208}Tl (2.61 MeV) and ^{214}Bi (up to 3.27 MeV) are the main contribution to the “external background” reaching the inner part of the detector from further-out components. The knowledge of their spectral shapes is of paramount importance for several solar neutrino analyses, and particularly for the measurement of the *pep* and ^8B neutrino interaction rates [5, 6], as well as for constraining the CNO neutrino interaction rate. An appropriate understanding of the external background was achieved by combining the information obtained from the external calibration of the detector [48] with the development of a special simulation procedure for the γ -rays originating in the PMTs and in the SSS.

The rate of the external γ -rays reaching a fiducial sphere of 3m radius is reduced by a factor $\sim 10^7$ with respect to that directly generated by the PMTs and the SSS. This is estimated by considering the attenuation length of about 23cm measured with the calibration data. The multiple interactions of γ -rays in the liquids produce an energy spectrum that cannot be modeled analytically. However, the attenuation factor is so large that it is impossible in practice to perform a brute force simulation, i.e. it is not feasible to generate events in the periphery of the detector and then to trace the minute fraction of γ -rays reaching the fiducial volume. Therefore, aiming at a considerable reduction of the computing time, we applied the *importance sampling algorithm* [68] (implemented inside the Geant4 libraries) as variance reduction method to our MC simulation. These methods are typically used to simulate radiation transport through thick shields of radioactive sources.

The algorithm considers the volume surrounded by the SSS as divided into spherical concentric shells. A number called “importance” is associated to each shell. Every time a photon crosses a boundary between two shells, the particle is either split or killed with a given probability depending on whether the importance of the entering volume is higher or lower than the exiting one.

The simulation is performed in three steps. First, radioactive decay events are simulated on the PMT photocathode surface and on the SSS and the importance sampling algorithm is applied. All γ -rays are propagated until they either are fully absorbed or reach the inner vessel. In case a γ -ray reaches the inner vessel, the kinematics of the first interaction inside this volume is saved. In the second step, the importance sampling is not applied and neither scintillation nor Čerenkov light emissions are simulated. All γ -rays found to interact at least one time in the inner vessel are generated according to the previously saved vertices (step one) and all the energy deposits are simulated until the γ -ray is fully absorbed. If some of the energy deposits are contained inside a sphere whose radius is typically 0.5m greater than the FV radius, the third and final step is performed: optical photons are generated according to the previously computed energy deposits and tracked until they reach the PMTs, like in standard simulations. This algorithm simulates external background events interacting inside the 3m spherical FV at a rate of about 0.5 events per second. More details can be found in Ref. [69].

7. GENIE and other event generators interface

A custom interface was written to simulate events produced by GENIE Neutrino Monte Carlo Generator [70]. Its output data format is a ROOT TTree [71] with information on the particles produced and their interactions. The interface reads the characteristics of the particles, such as type, energy, direction, position, and time, and generates corresponding Geant4 events. In addition to

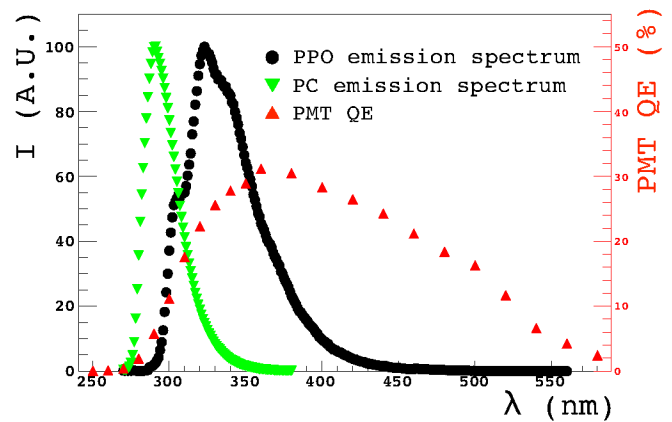


FIG. 8. PPO (black) and PC (green) emission spectra. The two curves are normalized to 100 at the maximum. The red points show the PMT Quantum Efficiency (QE) as a function of the wavelength.

GENIE, any external event generator using TTree data format can be used.

E. Generation of scintillation light

Scintillation photons are sampled according to the emission spectrum of PPO if they are generated in the inner vessel and according to that of PC if they are generated in the buffer region [72]. These spectra are shown in Fig. 8.

The number of emitted photons and their time distribution in the PC+PPO mixture depend on the details of the charged particle energy loss processes and on the molecular interactions between solvent (PC) and fluor (PPO) in the scintillator. We briefly remind that the fluorescence is a property of single molecules in organic scintillators and it depends on the features of the carbon bonds called π bonds. The π electron energy levels are quantized in a series of singlet S_{ij} and triplet T_{ij} states, where $i = 0, 1, 2, \dots$ denotes the electron energy levels and $j = 0, 1, 2, \dots$ vibrational sub-levels. π electrons are those responsible for the fluorescence mechanism. Scintillation light is emitted when π electrons decay from the first excited state S_1 to one of the vibrational energy levels of the fundamental state. More information on the organic scintillation processes can be found in Ref. [73].

Charged particles lose energy in the scintillator mainly exciting or ionizing the solvent. The energy is then transferred to the fluor either by radiative (emission of photons from the solvent and subsequent absorption from the fluor) or non radiative processes (dipole-dipole interactions between excited solvent and solute molecules). For fluor concentrations like the one of Borexino, non radiative transfer dominates. Once PPO is excited, it can emit fluorescence light only through the de-excitation of π electrons. Excited electrons of the fluor promptly reach (through non radiative processes) the first excited state

S_{10} , from where they decay to the fundamental state S_0 , emitting scintillation light, following an exponential time distribution with a time constant of ~ 1.6 ns. If some of the excitation energy is lost by collisions, and the electron is transferred to a triplet excitation state, T_{10} , it is impossible for the electron to de-excite emitting light but also return to the singlet state by non-radiative transition (being $E_{T_{10}} < E_{S_{10}}$). The only process available is the interaction with another solvent molecule in the same triplet state and this is the mechanism generating the delayed fluorescence.

The time distribution and the light yield of the emitted fluorescence photons depend on the various molecular processes taking place as a consequence of the particle energy loss. In addition, they do not only depend on the energy deposit in the scintillator but mainly on how the energy is released, i.e. on the value of dE/dx and on the type of incident particle. Heavy ionizing particles like α 's feature a large dE/dx , and produce large ionization or excitation density, thereby increasing the probability to get the triplet excitation state T_{10} and thus the presence of delayed fluorescence. In addition such large ionization or excitation densities favor molecular processes in which the energy is dissipated in non radiative ways, resulting in the quenching of the scintillation light (ionization quenching) [74].

The light emission cannot be simulated at a molecular level, because too many molecules and processes should be considered, resulting in a huge simulation time. We modeled the emission of fluorescence light by considering the difference in the effective emission time profiles induced by different types of particles and by including the ionization quenching effect.

The light emission time τ in the PC+PPO scintillator is generated in the MC according to the following phenomenological distribution:

$$P(\tau) = \sum_{i=1}^4 \frac{w_i}{\tau_i} \exp -\tau/\tau_i, \quad (2)$$

where the τ_i values and their weights w_i were obtained fitting the data from a dedicated experimental setup and then optimized during the tuning of the MC (see Sec. VIII B).

The light quenching due to ionization is modeled by the Birks formula [74] that links the scintillation light yield dY^{ph} produced when a particle loses energy over a distance dx with a stopping power dE/dx :

$$\frac{dY^{ph}}{dx} = \frac{Y_0^{ph} dE}{1 + k_B \cdot dE/dx}, \quad (3)$$

where the material-dependent Birks factor k_B (of the order of 10^{-2} cm/MeV) and the primary scintillation yield Y_0^{ph} have to be determined for every particular scintillator and incident particle. The total number of emitted scintillation photons is obtained by integrating Eq. (3).

We define the quenching factor $Q_p(E)$ as

$$Q_p(E) = \frac{1}{E} \int_0^E \frac{dE}{1 + k_B dE/dx}. \quad (4)$$

$Q_p(E)$ is always lower than 1. The suffix p recalls that $Q_p(E)$ depends on the particle type p (α , β , or γ) through k_B and dE/dx . Only if $k_B \cdot dE/dx \ll 1$ (for instance for electrons of some MeV) the light yield $Y_p^{ph}(E)$ is approximately proportional to the initial particle energy, while in general the following relation holds:

$$Y_p^{ph}(E) = Y_0 Q_p(E) E. \quad (5)$$

The above equation evidences the intrinsic non linearity between the deposited energy E and the emitted scintillation $Y_p^{ph}(E)$. Deviations of the light yield as a function of the energy deposit from a linear law are increasingly important for protons, α -particles, and nuclear fragments, because of their high ionization yield per unit length. The quenching effect is relevant also for γ -rays. As discussed in Ref. [47], the amount of scintillation light emitted when a γ -ray with energy E is fully absorbed in the scintillator is slightly lower than that emitted by an electron with the same energy E (e.g. the light emitted by a 256 keV γ -ray is similar to the light emitted by a 220 keV electron [14]). The scintillation yield Y_γ and the quenching factor $Q_\gamma(E)$ for γ -rays are given by

$$Y_\gamma = Y_0 \sum_i E_i Q_\beta(E_i) \equiv Y_0 \cdot Q_\gamma(E) \cdot E, \quad (6)$$

where the sum is performed over all the electrons generated during the γ -ray interaction. As $Q_\beta(E)$ decreases as a function of the energy, it results that $Q_\gamma(E)$ is smaller than $Q_\beta(E)$ for the same energy E . As a result, the quenching factor is not negligible for γ -rays with E in the MeV range.

The Birks parametrization is a macroscopic description of quenching and it can not be used directly in a stochastic approach, such as a MC simulation. In the Birks model, the quenching effect is obtained by comparing the initial energy of the primary particle with the energy deposited in the detector. In particular, all energy deposits due to secondary particles (like δ -rays or X-rays) are assumed to belong to the primary. On the other side, in the MC approach each particle is treated independently. Therefore, a correct parametrization of the Birks formalism requires to make the model compatible with the Geant4 framework, by evaluating the quenching factor for the primary ionizing particle, and making each daughter inherit the same factor. The ingredients for this simulation approach are: an *a priori* parametrization of the energy loss (as required by Eq. (4)) and a table of quenching factors as function of the energy (built at the initialisation phase, to speed up the simulation).

The light generation in the detector buffer medium (PC+DMP) is modeled according to the available measured data [43]. The total scintillation yield in the buffer

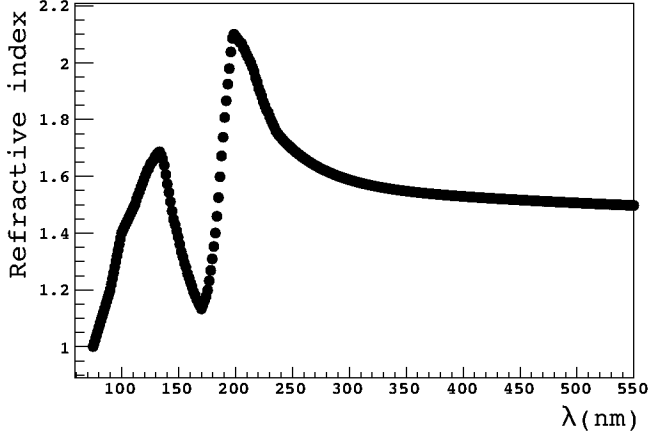


FIG. 9. Refractive index of the Borexino scintillator as a function of the optical photon wavelength.

is about $\sim 4\%$ of that of Borexino scintillator. The scintillation time constant in the buffer is 2.8 ns and the wavelength spectrum of buffer scintillation photons (which coincides with the PC emission spectrum) is presented in Fig. 8.

F. Generation of Čerenkov light

The primary spectrum of Čerenkov light is simulated by generating a number of photons $N_{\check{C}}$ per unit length and wavelength according to the Frank-Tamm equation [61]:

$$\frac{d^2 N_{\check{C}}}{dx d\lambda} = \frac{2\pi\alpha}{\lambda^2} \left(1 - \frac{c^2}{v^2 \cdot n^2(\lambda)} \right), \quad (7)$$

where α is the fine structure constant, $n(\lambda)$ is the wavelength dependent refraction index in the scintillator and v is the particle velocity in the scintillator. Note that the refractive index dependence on the wavelength λ causes a dependence of the Čerenkov spectrum on the particle velocity since the condition

$$\left(1 - \frac{c^2}{v^2 \cdot n^2(\lambda)} \right) > 0 \quad (8)$$

must be satisfied.

The spectrum of Čerenkov photons extends into the ultraviolet region, which is not directly detectable by the PMTs. The mean free path of this ultraviolet light in the scintillator is very short (sub mm), therefore, these photons are almost completely absorbed by the scintillator. However, the scintillator reemits a fraction of these photons at longer wavelengths, allowing for an indirect detection.

As evident from Eq. (7), the knowledge of the dispersion relation $n(\lambda)$ is essential in order to properly simulate the Čerenkov effect. Direct measurements allowed

to obtain $n(\lambda)$ for the Borexino liquid scintillator in the wavelength range $\lambda \in (245, 1688)$ nm. The number of generated Čerenkov photons is proportional to λ^{-2} and therefore an accurate knowledge of $n(\lambda)$ in the UV region is important. Since direct measurements at so short wavelengths are not available for the Borexino scintillator, the measured $n(\lambda)$ was extended in the UV region using the benzene data in literature [75, 76], considering that benzene and PC have similar refractive indices. The curve was furthermore extrapolated to the value $n = 1$ in the deep ultraviolet region. The knowledge of the scintillator properties in the deep ultraviolet region is not accurate and thus both the refractive index and the probability P_{rem} of light reemission by PPO in this photon energy region are experimentally unbounded parameters, which are fixed at the tuning stage of the simulation optimization (see Sec. V G and Sec. VIII). Figure 9 shows the resulting $n(\lambda)$, which is used both for Čerenkov photon generation and optical photon tracking (see Sec. V G).

The standard Geant4 G4Cerenkov library assumes that $n(\lambda)$ is an increasing function of the photon energy. As it is shown in Fig. 9, this is not our case and thus the standard library was extended to allow the usage of an arbitrary $n(\lambda)$.

G. The light tracking

The Borexino MC tracks each optical photon individually, considering its interactions with the single components of the scintillator and the buffer. These processes include elastic Rayleigh scattering, absorption and re-emission of photons by PPO molecules, absorption of photons by DMP, and also photon absorption in the thin nylon vessels. The cross-sections (or equivalently the wavelength dependent attenuation length) for these interactions were obtained with dedicated spectrophotometric measurements as shown in Fig. 10.

Photons emitted by scintillation or Čerenkov processes can interact with PC, PPO, DMP, or nylon molecules on their paths to the PMTs. In the model implemented in our MC simulation, if photons interact with DMP or nylon molecules, they are always absorbed. In case of interaction with PC, two distinct cases are considered, and their definitions are related to the PC emission spectrum shown in Fig. 8. Optical photons with a wavelength $\lambda > 310$ nm undergo Rayleigh scattering with an angular distribution $P(\theta) = 1 + \cos^2 \theta$, without time delay and without shift in energy. At shorter wavelengths, the interaction with PC is simulated as absorption by PC molecules, followed by the energy transfer to PPO (if the interaction happens in the inner vessel). With a probability of 82%, the PPO subsequently reemits the photon with an exponentially distributed emission time with a constant $\tau_{PC \rightarrow PPO} = 3.6$ ns. In the buffer medium, where PPO is not present, the interaction with PC for $\lambda < 310$ nm is followed by subsequent PC scintillation according to the PC spectrum of Fig. 8 with a probability

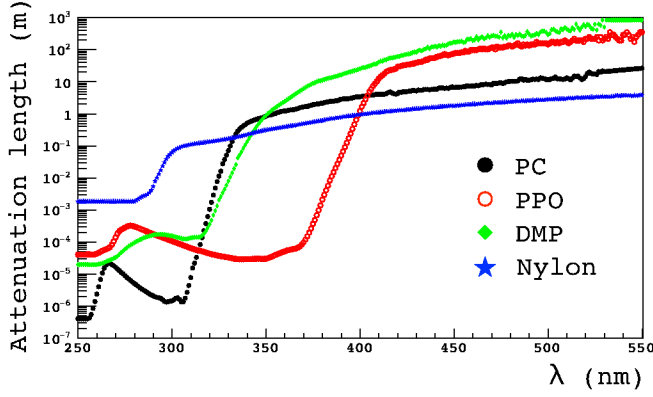


FIG. 10. Attenuation length of PC (black disks), PPO (red circles), DMP (green diamonds), and nylon (blue stars) as a function of the wavelength. As a reference, the PPO emission spectrum peaks between 300 and 400 nm (see Fig. 8).

Optical photon wavelength λ (nm)	Reemission probability P_{rem}
$\lambda < 320$ nm	0.53*
$320 \text{ nm} < \lambda < 375$ nm	0.839
$\lambda > 375$ nm	0.15

TABLE III. Model for the three-plateau reemission probability function after absorption by PPO molecules as a function of the wavelength, as it is implemented in the MC. The value highlighted with ‘*’ is obtained through the tuning procedure (see Sec. VIII C).

of 0.04 and with an exponential time distribution with $\tau_{PC} = 2.8$ ns.

Even if PC molecules are the most abundant, most of the optical photons interact with PPO. This is caused by the substantially shorter attenuation length of PPO in the wavelength range (300-400 nm) corresponding to the PPO emission spectrum (compare Fig. 10 with Fig. 8). Therefore, the absorption and eventual reemission of optical photons by PPO is the most relevant effect on the light propagation in Borexino. While the wavelength-dependent attenuation length of PPO was measured with great accuracy (Fig. 10), an experimental determination of the reemission probability as a function of the optical photon energy $P_{rem}(\lambda)$ is challenging. The optical model describing P_{rem} is summarized in Table III. In practice, it is a three-plateau function in the wavelength range of interest. No experimental data are available for $\lambda < 320$ nm, and thus the value of P_{rem} for this wavelength range was determined through the tuning with calibration data described in Sec. VIII C. Photons reemitted by PPO are produced with an isotropic distribution, according to the standard PPO energy spectrum and with an exponential time distribution with a time constant of $\tau_{PPO} = 1.6$ ns.

The MC simulation considers also the interactions of optical photons with interfaces (e.g. the refraction in-

duced by the presence of the nylon vessels) and surfaces (reflections on the light concentrator surfaces, on the SSS, on the PMT photocathodes).

This light propagation model was already validated in the CTF detector (Borexino’s prototype) [50] except for the treatment of UV photons due to Čerenkov production, which is a novel development introduced in the Borexino MC.

Photons reaching the photocathodes are detected with a probability corresponding to the wavelength-dependent PMT quantum efficiency depicted in Fig. 8. The same spectral dependence of the quantum efficiency (QE), but different peak values, are assumed for all PMTs.

H. Effective quantum efficiencies

All Borexino PMTs have slightly different light detection efficiencies, caused by different intrinsic properties of the photocathodes. These relative quantum efficiencies were measured prior to the PMT installation in Borexino. The peak values of the PMT quantum efficiency versus the wavelength (shown in Fig. 8), normalized to the best PMT in the set, are shown in Fig. 11. The intrinsically different quantum efficiency values for the PMTs are not sufficient to properly characterize the time behavior of the detector, since the global detection efficiency of a single channel might vary with time. For this reason, we introduced an “effective quantum efficiency” that globally describes the efficiency of the PMT, of the front end, and of the digital electronics.

The “effective quantum efficiency” is the frequency of a signal detection in case of a photoelectron, normalized to the solid angle of observation and to the total number of photons in that event. This definition, besides considering the intrinsic quantum efficiency differences among the PMTs, also describes light propagation non uniformities, electronic channel properties, and similar effects.

The effective quantum efficiencies are computed based on a selection of real data reconstructed close to the center (within a spherical radius of 2 m). In this way, solid angle differences among the PMTs are negligible. PMTs with and without concentrators are treated separately, as the detection efficiency is much higher for PMTs equipped with a cone. More specifically, we use ^{14}C events because of their relatively high statistics, uniform volume distribution, steady rate and low probability of multiple hits on the same PMT. The effective quantum efficiency is assumed proportional to the physical count rate on the respective channel, caused by ^{14}C selected in the innermost 2 m sphere. The dark noise of each PMT, which is evaluated weekly during the electronics calibration campaign, is statistically subtracted from the count rate. While some time dependent non-uniformities are observed for the resulting effective quantum efficiencies, they are negligible compared to the intrinsic differences among the PMTs, shown in Fig. 11.

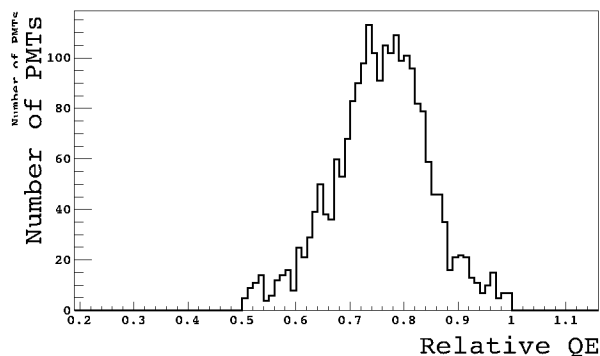


FIG. 11. Relative QE among the whole set of Borexino PMTs. The values are normalized so that 1 corresponds to the highest measured QE.

VI. SIMULATION OF THE READOUT ELECTRONICS

The Geant4 based tracking code simulates optical photons until they are detected by the PMTs. Our electronics simulation code (developed in C++) reproduces the electronics chain and the trigger system response, based on the information of the PMT pulse times. One of its most important features is the ability to simulate the detector electronics following its operating status over the entire data taking period. The detector is continuously monitored over time to assure the data quality and its stability. Because of the intrinsic loss of PMTs and since some acquisition boards might be disabled during part of the data taking, it is crucial that the electronics simulation follows the real time evolution of the operating PMTs. This is evaluated on an event-by-event basis, thanks to the information stored in a database during the standard data taking. The dark rate of individual PMTs, the effective quantum efficiency for each channel, and the PMT gains are also saved in the database and are injected in the simulation on a weekly basis. In addition, the code includes the effect of bad channels and detector inefficiencies by default in the same time dependent way.

Dark counts are generated according to the measured dark rate of each PMT and are superimposed to the signals due to scintillation. The typical dark noise rate is around a few hundreds of Hz for most of the PMTs. During data taking, time alignment is achieved by a common laser pulse, distributed to all PMTs via fiber optics with a timing precision of better than 0.5 ns. Intensities are chosen to detect an average of less than one photon per PMT per pulse, allowing the single photoelectron calibration in the sub-Poissonian regime. The accuracy of the relative time alignment is better than the PMT time jitter and so no other effects must be included in the simulation. In the electronics simulation the effective quantum efficiency is applied to each optical photon reaching a PMT. In case of detection, the code fully simulates the response of the analog and of the digital electronic chains

connected to every single PMT.

The front end analogue board generates two signals: a linear amplification of the PMT pulse, and a gate-less integrated signal (as described in Ref. [77]). The linear signal is discriminated by the digital board by means of a double threshold discriminator. The higher threshold allows to discriminate if the PMT was fired, while the lower one is used for timing purposes (with 0.5 ns resolution), thus reducing the time jitter dependence on the pulse height. We commonly define as “hit” a PMT pulse which crosses the higher discriminator threshold. The integrated signal is sampled by an Analog to Digital Converter (ADC) thus measuring the charge of each pulse. The integrator circuit is designed to keep the signal almost constant for 80 ns after the pulse, followed by an exponential decay with a time constant of 500 ns due to the capacitive coupling between PMTs and front-ends. In order to minimize the noise in the charge measurement, for each hit the integrated signal is sampled twice with an 8-bit ADC, once before the signal rise and a second time after 80 ns. The charge is proportional to the difference of these two ADC values.

After a channel detects a hit, the digital board has a dead time of 140 ns. This means that if two hits are closer than 140 ns in time for the same PMT, the time information of the second hit is lost. However, the system is able to measure its charge if it occurs less than 80 ns after the first hit. The occurrence of multiple hits on the same PMT is small but not null and its value must be accurately simulated to reproduce the energy response of the detector. The probability of double hits depends on the event position (if the scintillation event is not in the center of Borexino, the PMTs closer to the event detect more light) and on the energy, introducing a small non linearity and non uniformity which is included in the simulation. The response of the integrator to multiple hits is shown in Fig. 12 for a few example cases and it is implemented in the simulation.

In the electronics simulation code, the charge associated to each photoelectron is sampled from the single photoelectron response of each PMT modeled as an exponential plus Gaussian distribution. The specific parameters describing this curve for each PMT (different in the case of dark noise events and real scintillation events) were measured channel by channel. One example is shown in the left panel of Fig. 14. After the value of the charge is sampled, a corresponding PMT pulse is generated according to a reference shape for the signal which was acquired directly from the output of the front end. An averaged shape is applied to all the PMTs and it is shown in the right panel of Fig. 14. The generation of the analog output of the inverted and amplified PMT pulse, as it is produced by the front end, allows the simulation of the double threshold logic of the discriminators.

For each hit, the PMT transit time spread is simulated. The generation of an after-pulse is also considered with a fixed probability of 0.028. The probability density function used for extracting the after-pulse hit delay is shown

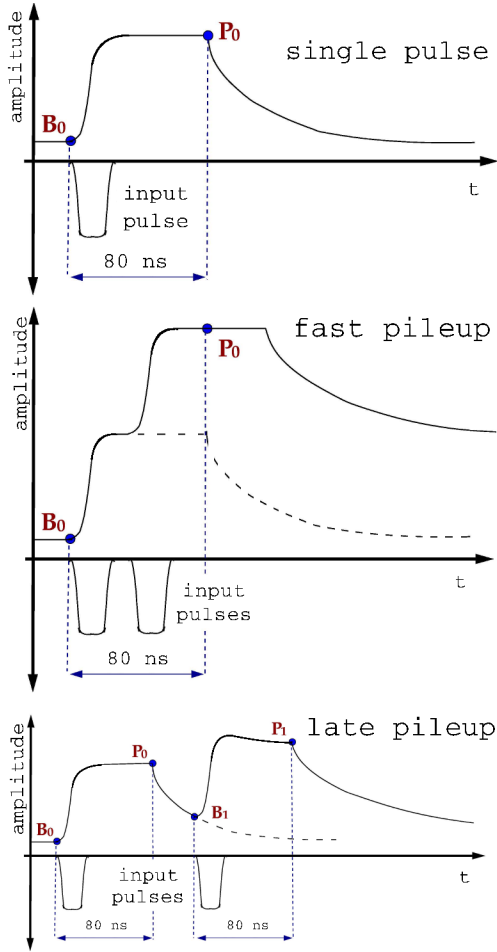


FIG. 12. Principle of operation of the gate-less charge integrator used in Borexino. The integrated output is shown as a function of time for various photoelectron arrival time combinations. Note that in the last case, after the second photoelectron arrival, the integrated output is not exactly constant. This effect is both taken into account in data analysis and in MC simulations.

in Fig. 13. The program also reproduces the gate-less integrator circuit previously described and the 140 ns dead time of the digital board is included in the simulation. The correct reproduction of the response of the front end and of the digital modules is crucial to obtain an accurate simulation of the position-dependent energy response of the detector.

A physical trigger in Borexino and in the MC simulation happens when N hits are detected within a time window Δt . Typically, trigger threshold and time window are set to $N = 20 \div 30$ and $\Delta t = 100$ ns. The threshold N changed within Borexino's history (mainly because of the loss of live PMTs) and is correctly taken into account in the MC.

The electronics simulation also reproduces the time structure of the acquired data which is due to the design details of the digital boards, providing an accurate repro-

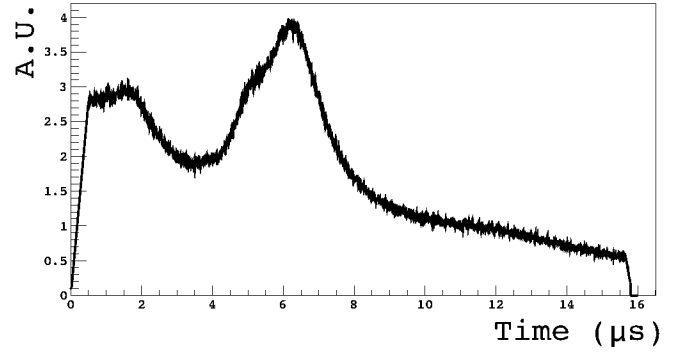


FIG. 13. Probability density function of after-pulse hit delays. This curve was measured for Borexino's PMTs in a time window of $16 \mu\text{s}$, which corresponds to the trigger acquisition gate duration.

duction of the output of the real system. This allows the trigger evaluation conditions to be the same both in the detector and in the simulation. There are two different strategies for the search of the trigger condition in the simulation: a time flow based and a “boosted” one. The former consists in letting the time flow with the steps of the time resolution of the trigger board. The condition for the occurrence of the trigger is evaluated as the time flows. This has the advantage of being able to generate a trigger event even if the “real” event generated with the tracking code is not present and the hits are due to the dark rate or the noise. The disadvantage consists in a very slow simulation. The boosted technique, instead, looks for the occurrence of the trigger condition in a time window around the physical Geant4 event. This is much faster, but the generation of the triggered event is related to the presence of the simulated vertex. In normal simulations, the boosted technique is the working condition, while the first could be used for trigger efficiency studies.

When the trigger is generated, all the hits within the acquisition gate of $\sim 16.5 \mu\text{s}$ ($\sim 6.9 \mu\text{s}$ before Dec. 2007) are saved. The information of each hit is stored using the real data format and is processed by the same reconstruction code. For each event, in addition to the electronics response of each channel, the information of the “MC truth” (information regarding some true input parameters of the simulated event) is stored in a dedicated structure, that can be accessed after the processing of the simulated events through the reconstruction code.

The primary events generated by the Geant4 code are not correlated in time. The only exceptions are radioactive decay chains and neutron captures. The time information is added by the electronics simulation, which assigns a time-stamp to each event, by sampling from an exponential distribution with a custom time constant, usually chosen to reproduce the observed event rate in real data. Coincident scintillation events caused by fast radioactive decays are sometimes detected within the same acquisition gate. The MC simulation implements this effect by copying the timing information from the Geant4

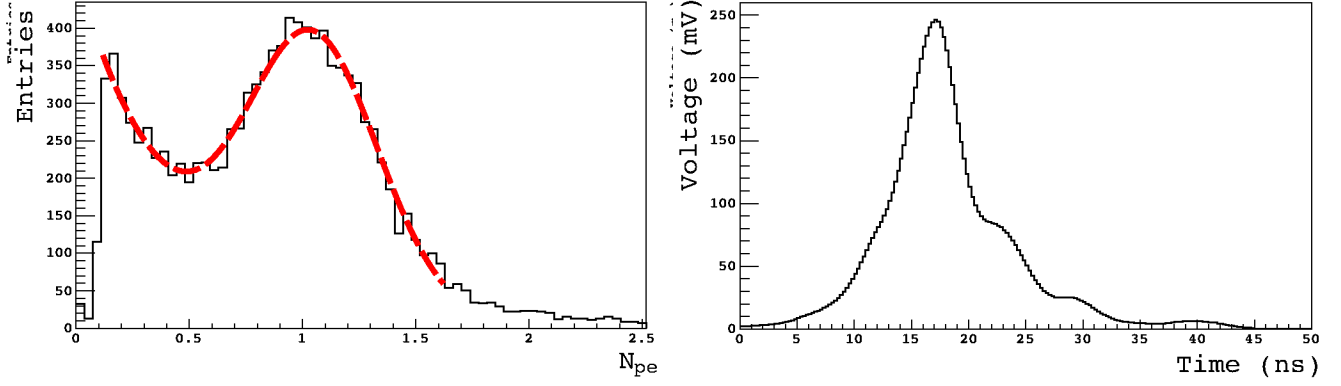


FIG. 14. *Left Panel:* Example of single photoelectron charge response of a PMT measured with a special laser calibration run with a particularly low discriminator threshold. *Right Panel:* Mean single photoelectron waveform, output of the front end, rescaled according to the specific hit charge and used for the discriminator threshold condition evaluation.

tracking code to the electronics simulation, which correlates the events in time.

The effects and asymmetries induced by dead PMTs and effective quantum efficiencies can be studied best by comparing different configurations with an ideal detector featuring a uniform distribution of equal PMTs, an option implemented in the MC.

VII. THE RECONSTRUCTION CODE

The reconstruction in Borexino is based on the amount of light collected by each PMT (related to the event energy) and on the relative detection times of the hits on the PMTs (used for event position and particle identification). A dedicated algorithm (called “clustering”) identifies the number of hits within the gate that correspond to a single scintillation event. Occasionally more than one cluster may be present during the gate duration. We define (both in data and in MC) different energy estimators called N_p , N_h , and N_{pe} [47]. They correspond to measured quantities such as the number of hit PMTs (N_p), the total number of hits (N_h), or the number of photoelectrons (N_{pe}) detected during the duration of a cluster or within proper fixed time interval. N_p is basically the number of PMTs that detected at least one hit, and in general it differs from N_h . Both N_p and N_h are computed starting from the measured values N_p^m and N_h^m [47]:

$$N_p^m = \sum_{j=1}^{N'} p^j \quad (9)$$

$$N_h^m = \sum_{j=1}^{N'} h^j, \quad (10)$$

where $p^j = 1$ when at least one photon is detected by the PMT j and $p^j = 0$ otherwise, $h^j = 0, 1, 2, \dots, n$ corresponds to the number of detected hits and N' is the number of

operating channels and it is evaluated on a nearly event-by-event basis using calibration events acquired during the run. N_p and N_h are then obtained after normalizing the measured values to $N_{tot} = 2000$ working channels through the relations [47]:

$$N_{p,h} = \frac{N_{tot}}{N'(t)} N_{p,h}^m. \quad (11)$$

N_p and N_h are obtained by the clustering algorithm without choosing a priori the time duration of the cluster (the maximum allowed value is $1.5 \mu s$). For some particular analyses (e.g. those extending to “low energy” such as the pp neutrino flux measurement [7]), additional energy estimators were introduced to perform the sum of Eq. (11) within a fixed time window measured from the first hit belonging to the cluster. We refer to them as $N_{p,h}^{dti}$, where $dti = 230$ or 400 ns and it corresponds to the fixed time interval over which the sum is computed. The time interval begins with the first hit of the first cluster and if needed it includes hits belonging to any additional clusters if present.

The third energy variable, N_{pe} , is the total number of collected photoelectrons normalized to N_{tot} channels. Following the same procedure explained above for $N_{p,h}$, the measured charge N_{pe}^m is calculated by summing the hit charges q_i^j expressed in photoelectrons:

$$N_{pe}^m = \sum_{i=1}^{N_h^m} q_i^j. \quad (12)$$

The number of channels with working ADC’s and charge readout, N'' , is used for normalizing N_{pe}^m to N_{tot} working channels:

$$N_{pe} = \frac{N_{tot}}{N''(t)} N_{pe}^m. \quad (13)$$

N'' is normally lower than N' by a few tens of channels.

The typical light yield in Borexino is of the order of 500 photoelectrons per MeV for events uniformly distributed

in the fiducial volume. Therefore, the number of hits recorded by each PMT approximately follows a Poisson distribution whose mean value is much lower than 1 up to an energy of a few MeV. Only for relatively high energy events ($E \gtrsim 3$ MeV) the mean value is close to 1 or greater.

The position reconstruction algorithm determines the most likely vertex position \vec{r}_0 for the scintillation event based on the knowledge of the arrival times t_i^j of the detected hits on the PMTs. For each measured t_i^j , the algorithm subtracts a position dependent time-of-flight T_{flight}^j from the interaction point to the PMT j placed at \vec{r}^j , i.e.:

$$T_{flight}^j(\vec{r}_0, \vec{r}^j) = |\vec{r}_0 - \vec{r}^j| \frac{n_{\text{eff}}}{c}. \quad (14)$$

Then, the likelihood $L_E(\vec{r}_0, t_0 | (\vec{r}^j, t_i^j))$ that the event occurs at the time t_0 in the position \vec{r}_0 given the measured hit space-time pattern (\vec{r}^j, t_i^j) is maximized. The probability density function for each hit used in the likelihood computation depends on the total number of collected photoelectrons in that same hit. More details can be found in Ref. [48].

The quantity n_{eff} appearing in Eq. (14) is called “effective refractive index” and it defines an effective velocity for the optical photons. This single parameter globally considers that photons with different wavelengths travel with different group velocities and that photons can be scattered or reflected from the emission to the detection points.

The position reconstruction code is used in the same way both for real data and MC ones. However, a small difference in the optimal value of n_{eff} for data and MC is observed. Using calibration data (see Sec. III) we determined $n_{\text{eff}}^{\text{data}} = 1.68$ for real data, while for MC events the best value is $n_{\text{eff}}^{\text{MC}} = 1.66$. This difference (about 1.2%) is consistent with the instrumental error on the measured values of $n(\lambda)$, obtained with ellipsometric techniques.

VIII. TUNING OF THE MONTE CARLO

The MC simulation forms the basis of all data analyses in Borexino. The best agreement between MC and real data is required in case of the solar neutrino analyses, where all the rates different from that of ^8B neutrinos, are obtained through a fit of the energy spectrum of the detected events passing criteria of data selection and quality cuts [47]. The MC simulation is used to produce the probability density functions (PDFs) to be used in the fit. The background spectra or the neutrino induced electron recoil spectra are converted into PDFs dependent on one of the energy estimators discussed in Sec. VII, using the MC to simulate the detector response. The accuracy of the conversion directly relates to the systematic uncertainty on the solar neutrino rate measurements.

The MC was tuned and validated using data sets independent from the one on which the actual fit was per-

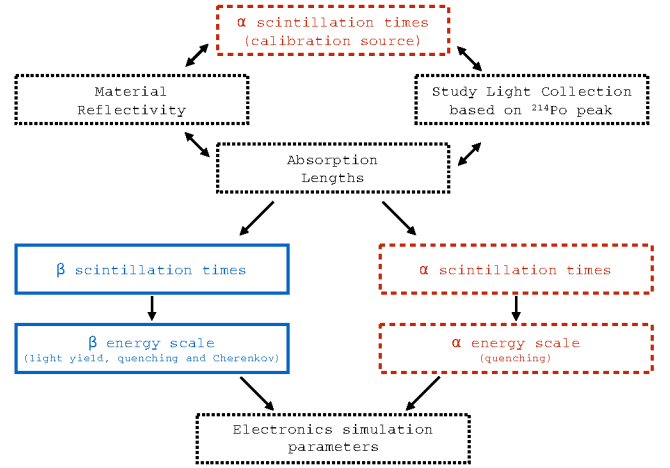


FIG. 15. Sketch of the “MC tuning strategy”. Continuous contours (blue) are relative to β particles, while dashed ones (red) refer to α ’s. Finer dashed contours (black) indicate sets of response parameters common to both α ’s and β ’s.

formed, such as calibration data and service triggers. The most important features of the events, i.e. the time-related distributions and the position-dependent number of hits and photoelectrons detected by each PMT, are reproduced with an accuracy of better than 1% within a 3 m radius fiducial volume in the energy range of solar neutrino analysis.

The full simulation of the scintillation and light propagation requires the input of a large number of parameters describing various features of the scintillator and of the materials. Examples are the attenuation lengths, the reflectivity, the reemission probabilities and the refractive indexes. The calibration campaigns [48] provided clean data samples, which were used to check the accuracy of the MC simulation. A portion of the calibration data was used as a tuning sample (to optimize some input parameters), while another independent set of calibration data (or data coming from the standard data taking) was used for testing the performance of the simulation. Great care was devoted to simulating the source runs and the vial geometry, accounting for all the known non idealities of the detector and its operating status as a function of time, as discussed in the previous sections.

Most of the input parameters of the simulation were measured in dedicated laboratory setups. However, laboratory conditions might sometimes have deviated from the operating conditions of the Borexino detector. For instance, all the laboratory characterization of the Borexino liquid scintillator might result inaccurate, simply because of the optical purity of the sample could be not the same as the one achieved inside Borexino.

The main difficulty in the tuning procedure is posed by the correlations among the parameters. Most of the physical effects (which are connected to the material properties) are mutually dependent. As an example, the energy

response is strictly related to the time response², which in turn depends on light propagation effects and transparencies. For these reasons, the tuning was performed iteratively and following the strategy outlined in Fig. 15.

At the beginning, all the parameters are initialized to measured values and a global check of the simulation is performed by direct comparison with results from calibrations. In this condition, the simulated detector response differs from the measured one by at most a few tens of percent. As shown in Fig. 15, the first step consists in varying attenuation lengths and reflectivity, in order to roughly reproduce the amount of light collected by the PMTs when the scintillation happens at various points inside the detector volume (see Sec. VIII A). Then, the time response for α -particles (see Eq. (2)) contained in the vial used for calibrations (see Sec. III) is optimized, so that these calibration points can be used for a precise tuning of the reflectivity and attenuation lengths. This part of the tuning of the parameters is referred to as “ α scintillation times (calibration source)” and “study light collection based on ^{214}Po peak” in Fig. 15.

Then, the procedure restarts from the beginning, using β -like events to cross check the time response of the simulation with the tuned attenuation lengths (because of absorption and reemission, attenuation lengths influence the time response) and reflectivity. Then, the specific parameters of the scintillation times for β 's and α 's are tuned, as well as the energy scale. The energy scale for β events is tuned together with some parameters of the electronics simulation, as the amount of reconstructed photons depends both on the intrinsic scintillator response and on the electronics. In the next sections, more details on all the steps of the tuning are described.

A. Non uniformity of the energy response within the inner vessel

The first step consists in the tuning of the optical attenuation lengths and of material reflectivity. These parameters impact on the uniformity of the energy response of the detector. The mixed ^{14}C - ^{222}Rn source, deployed in ~ 200 different positions, was used to perform such a study. Particularly, the chosen reference is the position of the ^{214}Po α energy peak. Because of the quenching of the scintillator contained in the vial, the precision tuning and study of the light collection was performed after a dedicated tuning of the time response of the scintillator contained in the source vial. ^{214}Po events are point like with respect to the Borexino reconstruction resolution and produce about $\sim 300 N_p$. Since the mean channel occupancy is ~ 0.15 , the ^{214}Po allows to operate in the single photoelectron regime, where the N_p and N_h energy estimators have a negligible dependence on electronic ef-

fects, allowing to understand the light propagation almost independently of the electronics simulation.

The attenuation lengths for PC, PPO, DMP, and nylon were measured as a function of the photon wavelength with spectrophotometric techniques using samples with small volumes of liquid. During the tuning procedure, the absolute values of the attenuation lengths were slightly adjusted by scaling them with multiplicative factors (Λ). Other important parameters responsible of the asymmetry in the light collection inside the detector are the values of reflectivity of the light concentrators and of the stainless steel sphere. Besides a good reproduction of the shape of the concentrators, the aluminum cone reflectivity and the ratio between specular and diffusive reflection³ components play a crucial role and have to be determined with high accuracy. The key observables allowing the tuning of these parameters are basically two:

- the ratio of the fired PMTs with and without light concentrators (“cone-no cone ratio”).
- the ratio of the fired PMTs far (> 4 m) and close (< 4 m) from the source (“far-near ratio”).

These observables have the advantage to be energy independent, allowing us to avoid degeneracies due to the non linear behavior of the energy scale and also tolerate differences on the absolute energy scale at this stage of the tuning. The cone-no cone ratio allows to determine the reflectivity of the inner and outer surfaces of the light concentrators and the specular/diffusive ratio. The same holds for the reflectivity of the steel ring mounted on the PMTs without concentrator. The far-near ratio allows to study the attenuation lengths. Actually, both the attenuation lengths and the reflectivity affect the two ratios at the same time, since the concentrator efficiency depends on the distance of the event from the PMT. The two mentioned distributions and their correlation were compared with simulation data for ~ 20 different source position inside the inner vessel. The best parameters have been selected using the χ^2 and the Kolmogorov-Smirnov test statistics. Figure 16 shows examples of the cone-no cone ratio and far-near ratio distributions for selected points inside the inner vessel. In the plot, the results of the Kolmogorov-Smirnov and χ^2 tests are also reported. Figure 17 shows the combination of the two ratios, and particularly the different cone-no cone ratio distributions for near and far PMTs. The two distributions show different features, ideal candles to fully characterize the detector response as function of the vertex position.

The results of the attenuation length and reflectivity tuning are shown in Table V, where the nominal and tuned values are reported for various parameters. As already mentioned, the values of the reflectivity are also influenced by the time response of the detector (see Sec.

² This is mostly related to the electronics response, see Sec. VI.

³ Diffusive reflection is modeled as Lambertian reflection.

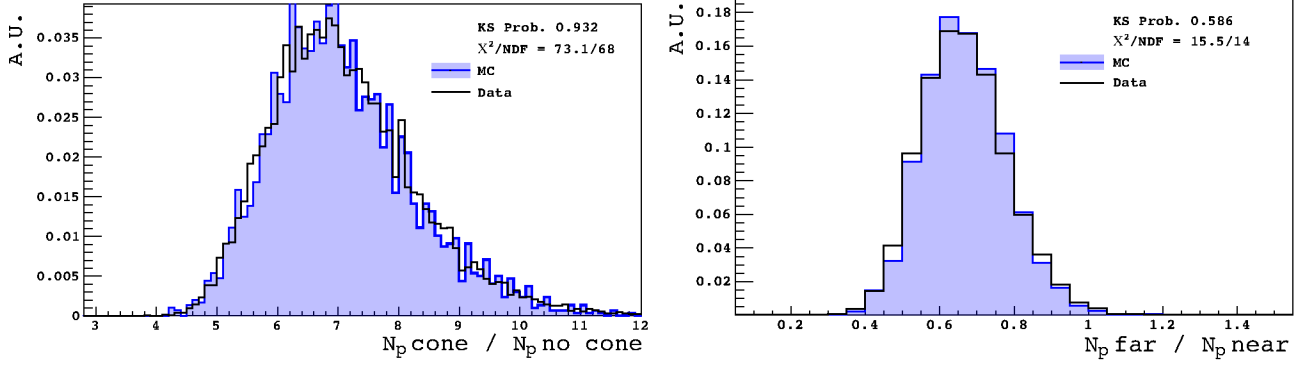


FIG. 16. *Left Panel:* MC-data comparison of the cone-no cone ratio in a selected point in the inner vessel. The first entry in the legend indicates the outcome of a Kolmogorov-Smirnov test. *Right Panel:* MC-data comparison of the far-near ratio in a selected point in the inner vessel.

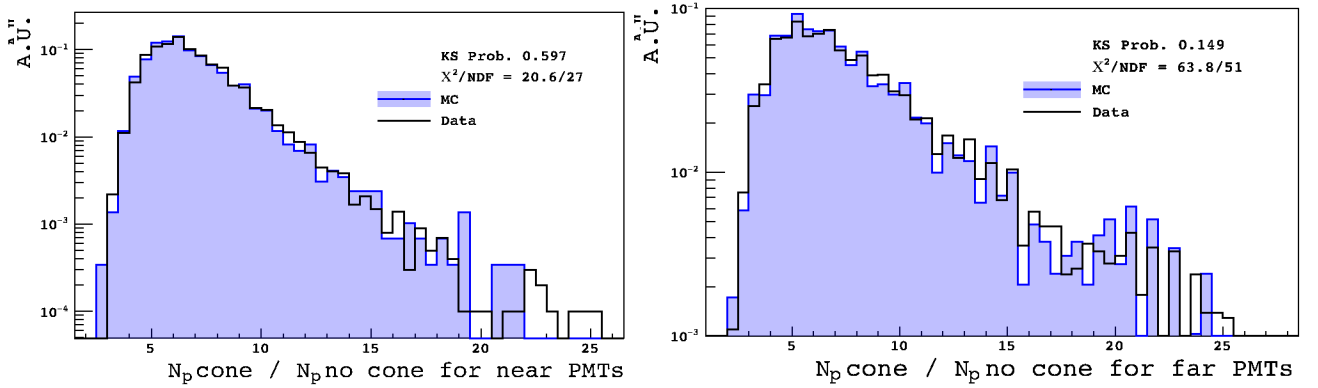


FIG. 17. Cone-no cone ratios for near (left) and far (right) PMTs. The first entry in the legend indicates the outcome of a Kolmogorov-Smirnov test.

VIII B) and thus the final result was reached through the iterative approach described in this section. The goal set on the accuracy of the light collection is $\lesssim 1\%$ and the comparison between data and simulation is shown in the bottom panel of Fig. 18, where the relative discrepancy in the case of N_h is plotted as a function of the source radial position. Inside the 3.5 m sphere, all the points are contained in the $\pm 0.5\%$ band, thus showing that the precision goal was reached. The top panel of Fig. 18 shows the comparison between data and MC of the ^{214}Po peak position in the N_h energy estimator. Similar results were achieved both with N_p and N_{pe} . These results demonstrate a strong improvement in the comprehension of the detector response with respect to the Borexino calibration results already published in Ref. [47].

B. Time distribution of the collected light

The main parameters describing the time response of the detector are the scintillation time constants, the absorption-reemission delays, and the electronics response of the detector. In order to decouple scintillation

and light propagation effects, “low energy” sources in the center of the detector were chosen both for α and β -particles. This choice allows to study the response of the system with all the PMTs in the single photoelectron condition. Particularly, ^{210}Po α decays from data were selected using a pulse shape discriminator parameter [47] in the innermost 1 m sphere. Data acquired with the ^{85}Sr source in the center of the detector were used as benchmark for the simulations of β -particles. Both ^{85}Sr and ^{210}Po have a visible energy of around $\sim 200 N_h$, which is relatively low and thus the electronics effects are minimized. Unfortunately, it is not possible to precisely tune the response to α particles using any of the data from the calibration campaign, since in that case the scintillation happened in the vial and not in the actual Borexino scintillator.

The tuning is performed for these classes of events in the center, and then the agreement is checked for different source positions. This also allows to test whether the attenuation lengths were properly selected at the previous stage, since the time response is affected by the amount of absorptions and reemissions. Figure 19 compares the MC and data time distributions of the hits

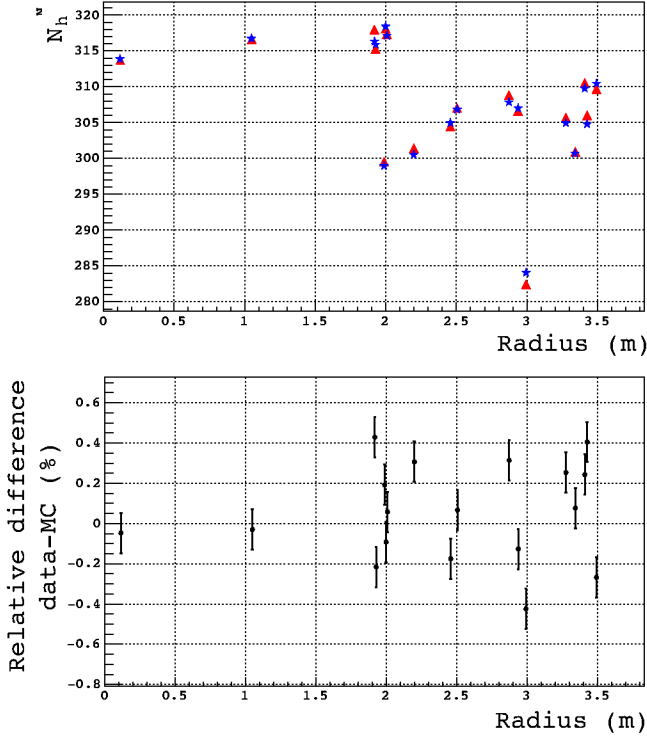


FIG. 18. *Top Panel:* ^{214}Po energy peak in MC (blue stars) and data (red triangles) as a function of the radial distance of the source from the center. The statistical uncertainties on the points are smaller than the markers. *Bottom Panel:* Relative difference of the ^{214}Po energy peak in MC and data (N_h energy estimator) as a function of the radial distance of the source from the center.

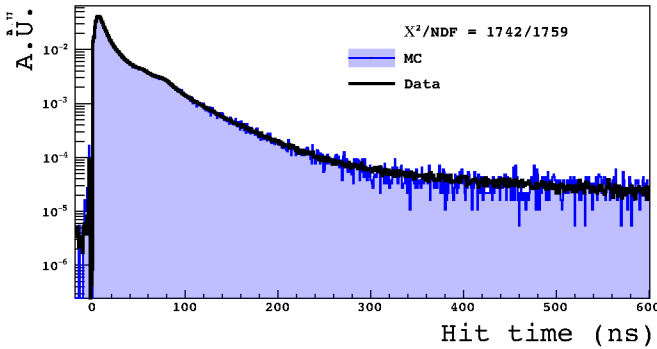


FIG. 19. Light curve (i. e. hit times with respect to the first hit in the cluster) for ^{85}Sr events in the center of the detector.

detected by the PMTs for events from the ^{85}Sr calibration source in the center. This curve is the convolution of the four exponential response function used to describe the scintillation process (see Ref. [47]) and the simulated absorptions and reemissions. It is possible to distinguish the shape due to the sum of the four exponentials, but some additional features are evident, such as the structure at approximately ~ 100 ns, which is due to the re-

flection of photons on the stainless steel sphere, and the tail at larger times, which is due to the contribution of after-pulse hits and dark noise. The correct reproduction of the hit time distribution ensures that other more complex variables reconstructed from the hit times (such as the reconstructed event position and the Gatti α/β pulse shape discrimination variable [47]) are also correctly reproduced by the MC, as shown for example in Fig. 20.

The statistical tests adopted to select the best parameters were again the Kolmogorov-Smirnov and χ^2 tests. The results of the tuning of the parameters are reported in Table VI. The exponential parameters were only slightly adjusted with respect to the experimentally determined ones (coming from Ref. [78]). This is acceptable, because of the optical purity of the laboratory scintillator sample (see beginning of Sec. VIII). Furthermore, while the full Borexino MC is essentially an ab initio simulation, the model adopted in Ref. [78] for estimating the scintillation time constants is based on phenomenology. Therefore, a little discrepancy might also be due to some small propagation effects which are not accounted for in the laboratory setup measurement. However, it was not necessary to modify the characteristic times of absorption and reemission. In this respect, the most important parameter is τ_{PPO} , which affects significantly the time distribution of photons (most of the primary light is absorbed and reemitted within a few cm from the interaction point). τ_{PPO} is quite well established in the literature, and thus the absence of the need of a modification is an argument for the solidity of the optical model of the detector in the tracking code.

C. The energy scale

The last step of the tuning of the simulation is the reproduction of the absolute energy scale. All the input parameters, apart from the ones describing the quenching and the scintillation times, are common to the β and α energy response. Crucial parameters affecting the energy scale are the absolute scintillation yield Y_0^{ph} , the Birks quenching parameter k_B , and the yield of the Čerenkov light. As discussed in Sec. V F, there are no experimental data available for P_{rem} for $\lambda < 320$ nm. In addition, the index of refraction in the near UV is not accurately constrained. As a consequence, P_{rem} is considered as an effective parameter. Furthermore, some electronics fine tuning was required for a high precision simulation of N_{pe} .

The reproduction of the energy response of β particles was optimized scanning the light yield Y_0^{ph} , the Birks quenching factor k_B , the reemission probability at low wavelengths, and an overall multiplicative parameter (G_s^{PMT}) rescaling the simulated single photoelectron gains of the PMTs. G_s^{PMT} is needed since the lack of some details in the electronics simulation requires a small adjustment. In particular, it is necessary to overcome the approximation used in the electronics simulation where

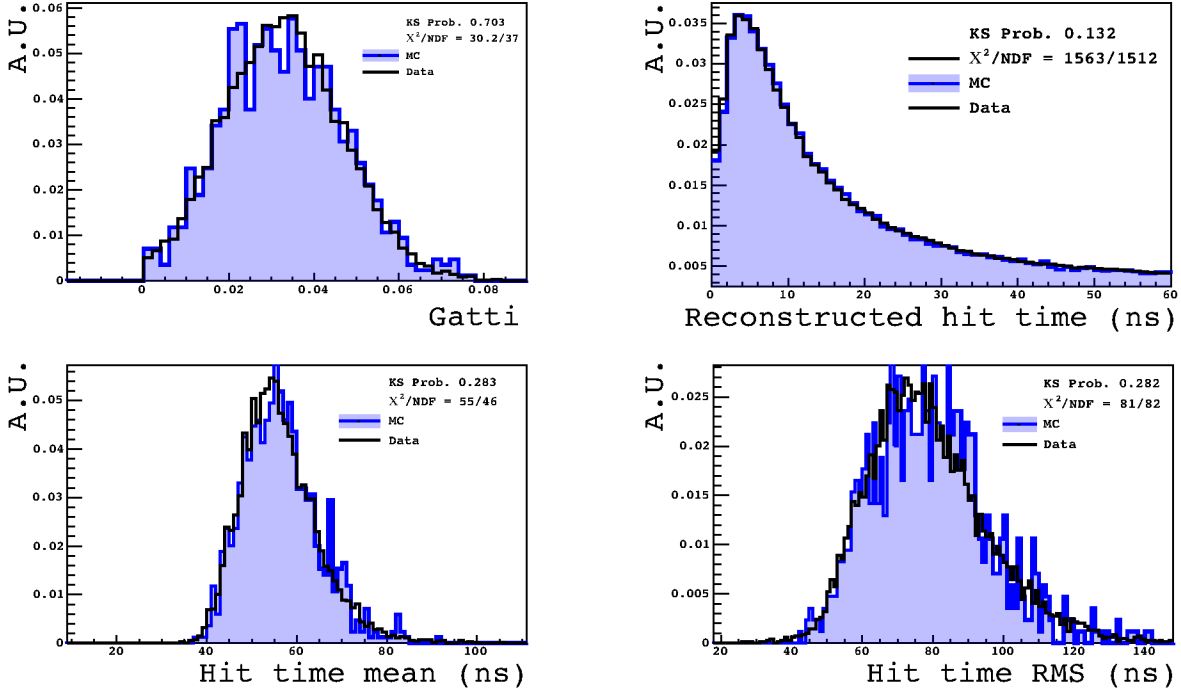


FIG. 20. Data-MC agreement for some timing variables for ^{210}Po events in the center. In the top left panel the α/β Gatti [47] distribution is shown, the top right panel contains the reconstructed hit time distribution (hit time distribution corrected for the time of flight), the bottom left (right) panel shows the mean (RMS) time of the hits in the cluster. The first entry in the legend indicates the outcome of a Kolmogorov-Smirnov test.

the reference shape for the pulse of the PMTs is unique for all the channels. All the other parameters in the simulation were kept fixed and the measured energy estimators were compared with the simulated ones for events in the center of the detector. A comparison of the γ -ray peak positions for the N_h , N_p , and N_{pe} variables is shown in Fig. 21.

The energy spectra of ^{57}Co , ^{139}Ce , and ^{203}Hg overlap with the tail of ^{14}C spectrum. A volume cut is not sufficient to obtain clean peaks and thus a proper subtraction of background events (acquired before and after the source insertion) was performed. The agreement between data and simulation is evaluated comparing the reconstructed energy peaks of each γ -ray source. They are fitted with a Gaussian function plus a line (the irreducible background due to radioactive decays or neutrinos) to obtain the mean value and the RMS of the energy estimator. The goodness of the energy simulation was defined with a custom estimator:

$$\epsilon = \sum_{\text{sources}} \frac{(\mu^{\text{data}} - \mu^{\text{MC}})^2}{e_{\mu^{\text{data}}}^2 + e_{\mu^{\text{MC}}}^2 + 1}, \quad (15)$$

where μ is the mean value of the Gaussian function from the fit (i.e. the peak position in the energy spectrum) and e_{μ} its error. Such definition of the denominator avoids biases due to very small errors on μ^{data} and μ^{MC} . The optimized set of Y_0 , k_B , reemission probability, and G_s^{PMT} was found looking for the minimum value of ϵ .

Parameter	Expected value	Tuned value
Y_0^{ph} (photons/MeV)	$O(10^4)$	13600
$k_B(\beta, \text{cm MeV}^{-1})$	$O(10^{-2})$	0.01098
$k_B(\alpha, \text{cm MeV}^{-1})$	$O(10^{-2})$	0.01055
G_s^{PMT}	1	1.015
$P_{\text{rem}}(\lambda < 320 \text{ nm})$	-	0.53

TABLE IV. Results from the tuning of the parameters describing the energy response. G_s^{PMT} allows to slightly vary the gains of the PMTs in the simulation. $P_{\text{rem}}(\lambda < 320 \text{ nm})$ is the parameter discussed in Sec. V E. The orders of magnitude of the starting points of the light yield Y_0^{ph} and of the quenching parameters k_B were inferred from standard properties of organic liquid scintillators in the literature when available.

This procedure was followed for each energy estimator, showing that the results are independent on it.

The energy response (peak and resolution) of all the sources located in the center of the detector is reproduced with a precision better than 0.8%. The resulting tuned parameters are given in Table IV. The procedure is then repeated once again for α particles. In this case, only the Birks quenching parameter (k_B for α) is free, as all the others are fixed by the previous iteration.

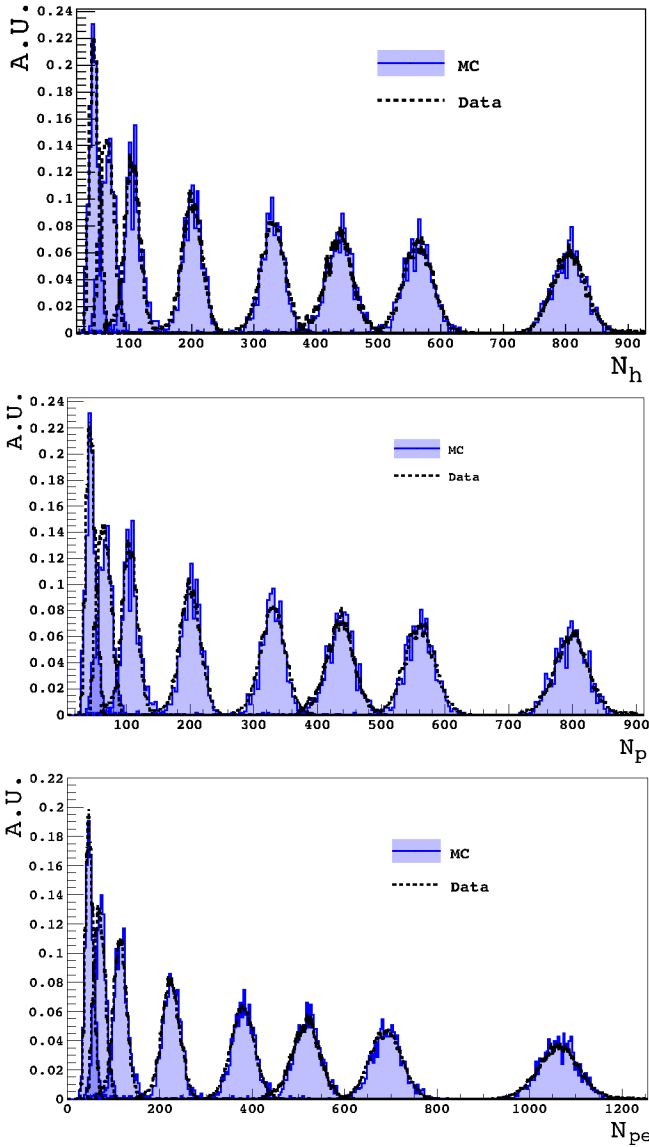


FIG. 21. *Top panel:* Data-MC agreement for the N_h energy estimator for the γ -ray sources listed in Sec. III with the addition of the 2.22 MeV γ -ray events from the neutron capture of the ^{241}Am - ^9Be source in the center of the detector. *Middle Panel:* Data-MC agreement for the N_p energy estimator for the γ -ray sources listed in Sec. III with the addition of the 2.22 MeV γ -ray events from the neutron capture of the ^{241}Am - ^9Be source in the center of the detector. *Bottom Panel:* Data-MC agreement for the N_{pe} energy estimator for the γ -ray sources listed in Sec. III with the addition of the 2.22 MeV γ -ray events from the neutron capture of the ^{241}Am - ^9Be source in the center of the detector.

Parameter	Expected value	Tuned value
Λ_{PC}	1	1.35
Λ_{nylon}	1	0.75
Λ_{DMP}	1	1.1
Λ_{PPO}	1	1.15
R_{SSS}	0.49*	0.55
$R_{\text{SSS}}^{\text{spike}}$	0.4*	0.12
R_{cathode}	0.1*	0
$R_{\mu\text{-metal}}$	0.4*	0.4
$R_{\text{PMT ring}}$	0.6*	0.75
$R_{\text{PMT ring}}^{\text{spike}}$	0.8*	0.88
$R_{\text{conc. int.}}$	0.88 [44]	0.95
$R_{\text{conc. int.}}^{\text{spike}}$	0.8*	0.985
$R_{\text{conc. ext.}}$	0.88 [44]	0.95
$R_{\text{conc. ext.}}^{\text{spike}}$	0.8*	0.975
$R_{\text{nylon ring}}$	0.4*	0.3

TABLE V. Results on the tuning of the parameters related to attenuation lengths and reflectivity. Λ is the multiplicative coefficient which multiplies the measured curves of the attenuation lengths, R is the reflectivity and R^{spike} is the fraction of specular reflection with respect to diffusive reflection. “PMT ring” denotes the stainless steel small ring mounted on the PMTs without concentrators. $R_{\text{nylon ring}}$ is the value of the reflectivity of the nylon endcaps on top and bottom of the inner vessel. The expected values for the parameters highlighted with “*” are inferred from typical values for those materials as found in the literature.

IX. VALIDATION OF THE TUNED MONTE CARLO

A. ^{210}Po events and energy resolution

α decays of ^{210}Po in the scintillator represent a good benchmark for the validation of the tuned MC. ^{210}Po is not uniformly distributed in the detector and its activity and space distribution change as a function of time because of its decay and of convective movements of the scintillator. Furthermore, Borexino experienced a quite significant loss of live PMTs along the almost 10 years of data taking, which resulted in a time varying energy response and in a time dependence of the non uniformity of the energy response [47]. The distribution of dead PMTs changed over time (at the beginning, most of the dead PMTs were at the bottom of the detector, while in the last years the situation evolved towards a more uniform distribution). The accurate MC tuning allows to obtain a good agreement in the reproduction of the detector en-

Parameter	Measured value	Tuned value
$\tau_{1,\beta}$ (ns)	3.95	3.7
$\tau_{2,\beta}$ (ns)	23.56	24
$\tau_{3,\beta}$ (ns)	78.86	60
$\tau_{4,\beta}$ (ns)	546.39	600
$q_{1,\beta}$	0.933	0.889
$q_{2,\beta}$	0.024	0.055
$q_{3,\beta}$	0.022	0.027
$q_{4,\beta}$	0.021	0.029
$\tau_{1,\alpha}$ (ns)	4.15	3.9
$\tau_{2,\alpha}$ (ns)	19.90	26
$\tau_{3,\alpha}$ (ns)	99.91	110
$\tau_{4,\alpha}$ (ns)	617.96	630
$q_{1,\alpha}$	0.679	0.674
$q_{2,\alpha}$	0.144	0.146
$q_{3,\alpha}$	0.102	0.103
$q_{4,\alpha}$	0.075	0.077
τ_{PC} (ns)	28	28
τ_{PPO} (ns)	1.6	1.6
$\tau_{PCtoPPO}$ (ns)	3.6	3.6

TABLE VI. Results on the tuning of the scintillation time parameters for α and β particles. It has to be noted that all these numbers are correlated to some of the reflectivity values (especially stainless steel sphere's and concentrator's). The τ_i and q_i values are the time constants and the weights describing the four exponentials modeling the primary scintillation light generation (see Ref. [47]). The starting values for the exponential parameters come from Ref. [78] (see the discussion in text for the explanation of the differences with the tuned values). The three last rows describe absorption and reemission time constants, as discussed in Sec. V E.

ergy response to ^{210}Po α decays without any change of the input parameters found at the tuning stage.

We investigated the agreement between data and MC for the period 2012–13. On a monthly basis, ^{210}Po events were selected in the data in cubic voxels of ~ 1.5 m edge using a suitable pulse shape parameter and appropriate energy cuts [47]. We separately regarded 59 cubic voxels within the fiducial volume. As a second step, we simulated ^{210}Po data sets reproducing the time and space distributions observed, including the number of operating channels in this data taking period. We then compared the spectral position of the ^{210}Po peak in data and MC and found an agreement (see Fig. 22) of better than 1%. This result demonstrates that the simulation that is tuned on discrete source positions is able to reproduce

the non uniformity of the energy response observed in data. In addition, it shows that the scintillator response is stable over time as calibration data taken in the year 2009–2011 furnish MC parameters that accurately reproduce data measured several years later⁴.

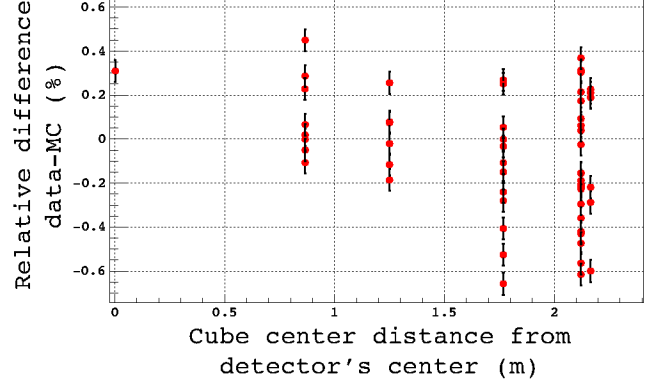


FIG. 22. Relative difference between the ^{210}Po energy peak (evaluated with the N_h energy estimator) in MC and in data as a function of the cube center position (radius from the center of the detector). The cubes cover completely the innermost 3 m of Borexino's inner vessel.

An important crosscheck of the MC modeling is based on the comparison of the predicted energy resolution with the one measured in data, as no additional parameter in the simulation is required for tuning the energy resolution. The reproduction of the correct value is a by-product of the simulation of all relevant scintillation processes (charged particle energy loss, scintillation and Čerenkov light generation, absorption, reemission, and light propagation) and of the proper schematization of the electronics effects (PMT jitter, electronics features and so on). Testing the ability of the MC simulation to reproduce the ^{210}Po peak position and width allows to test the combination of all the above mentioned effects. Both the peak position and its width are strictly correlated with the non uniformity simulation both in time and in space of the detector response. In addition, the peak width accounts for the energy resolution effects providing a benchmark for testing the whole MC simulation.

B. High energy light collection

Light collection tests were also performed studying the 2.2 MeV γ -ray emission after neutron capture events on hydrogen from an ^{241}Am - ^9Be neutron source. These γ -ray events have higher energy with respect to the typical energies of analysis for solar neutrino measurements⁵.

⁴ Note that the ^{210}Po events selected for the α particle tuning are different from those used for this test and shown in Fig. 22.

⁵ With the exclusion of the ^8B neutrino analysis.

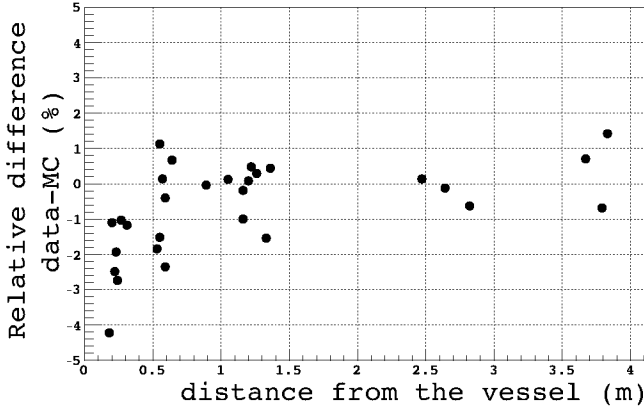


FIG. 23. Relative difference of the 2.2 MeV neutron capture on hydrogen γ -ray energy peak (N_{pe} variable) in MC and data as a function of the distance from the inner vessel. The agreement worsens at short distance from the vessel (distance similar to the γ -ray attenuation length on the scintillator), probably due to uncertainties in the reconstruction of the vessel shape [47].

However, reproducing the detector behavior at energies of a few MeV are of interest for geo-neutrinos and SOX analyses. Since for the former analysis the fiducial volume radius is increased [8], several positions close to the inner vessel were mapped with the ^{241}Am - ^9Be source. Figure 23 shows the relative difference between data and MC as a function of the distance from the vessel using the N_{pe} energy estimator. Although the tracking code was tuned at lower energy, the discrepancy is $\lesssim 2\%$ in the volume of interest. The worsening of the agreement at short distance from the vessel is probably due to uncertainties in the reconstruction of the vessel shape [47].

C. External background

The external γ -ray simulation procedure was validated by direct comparison with data from calibrations. No parameters were tuned or modified to improve the goodness of the simulations, thus the complete data set from external calibrations was used for testing the already tuned MC code.

The external calibration campaign was carried out using a ^{228}Th source ($\tau = 2.76$ yr), with ~ 2.9 MBq activity placed in several positions both in the top and the bottom hemispheres, as listed in Table VII. ^{208}Tl is one of ^{228}Th 's daughters and the emission probability of the 2.615 MeV γ -ray line is 35.6%. α and β particles emitted by daughter nuclides of ^{228}Th are absorbed by the source encapsulation, while low energy γ -rays are absorbed in the buffer, very close to the generation point. Data were acquired with the source in positions on two sides of the detector (north and south, geographically) and with different heights with respect to the ground. In Table VII the “N” (“S”) letter indicates the north (south) side while

the numbers refer to the position's height. It is important to study the detector response to external γ -rays in different positions and especially at different heights to test both the effect of the nylon vessel shape and that of the asymmetry of the live PMTs distribution.

Using the variance reduction procedure described in Sec. VD 6, ^{208}Tl events originating from all source positions (see Table VII) were simulated. The reconstructed energy spectra and event positions were compared to the measured ones, recording an excellent agreement between simulations and real data. Examples are given in Fig. 24, where the event positions reconstructed in data and MC are shown, and in Fig. 25, where the γ -ray energy spectrum built with the N_h variable is compared in data and MC for different fiducial volume cuts. Both the event position and the energy spectrum are well reproduced. Particularly, it is possible to note a good agreement even with the 3.5 m fiducial volume cut (left panel of Fig. 25), which is already well beyond the needs of the solar neutrino analysis (i.e. good light collection reconstruction in the 3 m sphere). The different shape of the energy spectrum between the two plots in Fig. 25 is due to the different fiducial volume cut, and thus to the different ratio between the Compton tail (of events degraded in energy by the buffer) and the full absorption peak.

A more systematic study was performed to assess the solidity of the method in the whole inner vessel volume. In particular, a parameter of interest is the ratio between the number of events in the full absorption peak and in the Compton tail, $\rho_{\text{peak/tail}}$, which mostly depends on the vessel shape non-uniformity. This ratio is quantified by fitting the full absorption peak with a Gaussian (as shown in Fig. 26). The integral of the Gaussian is then compared to the total number of events in the spectrum and thus to those in the Compton tail.

Figure 27 shows the values of $\rho_{\text{peak/tail}}$ as a function of the source position both for MC and data for events selected in a 3 m fiducial sphere. The excellent agreement found demonstrates that the simulation is working properly. Moreover, it proves once more that the dynamical vessel shape reconstruction in Borexino is solid, since $\rho_{\text{peak/tail}}$ as a function of the source position depends on it.

D. Position reconstruction

The ability of the MC to reproduce the position reconstruction was evaluated in two steps. As the first step, a large sample of events homogeneously distributed over the 3.5 m fiducial volume was generated to check for the presence of systematic shifts in the reconstruction. As shown in the top left panel of Fig. 28, the difference between the real vertex position (recorded in the MC truth structure for each event) and the reconstructed position is compatible with zero within the error. As second step, the reconstructed position resolution was compared for

	N1	N2	N3	N4	N5	N6	N7	S3	S5	S7
X [m]	2.63	4.16	5.93	6.53	6.07	4.37	2.55	-6.00	-6.06	-2.21
Y [m]	-0.62	-0.81	-1.44	-1.65	-1.08	-0.80	-0.43	0.98	1.02	-0.27
Z [m]	6.29	5.39	3.11	-1.24	-2.99	-5.22	6.34	3.15	-3.03	-6.48

TABLE VII. Positions of the external calibration ^{228}Th source deployments. The coordinates are relative to the center of the detector. “N” (“S”) letter indicates the north (south) side of the detector while the numbers refer to the position’s height.

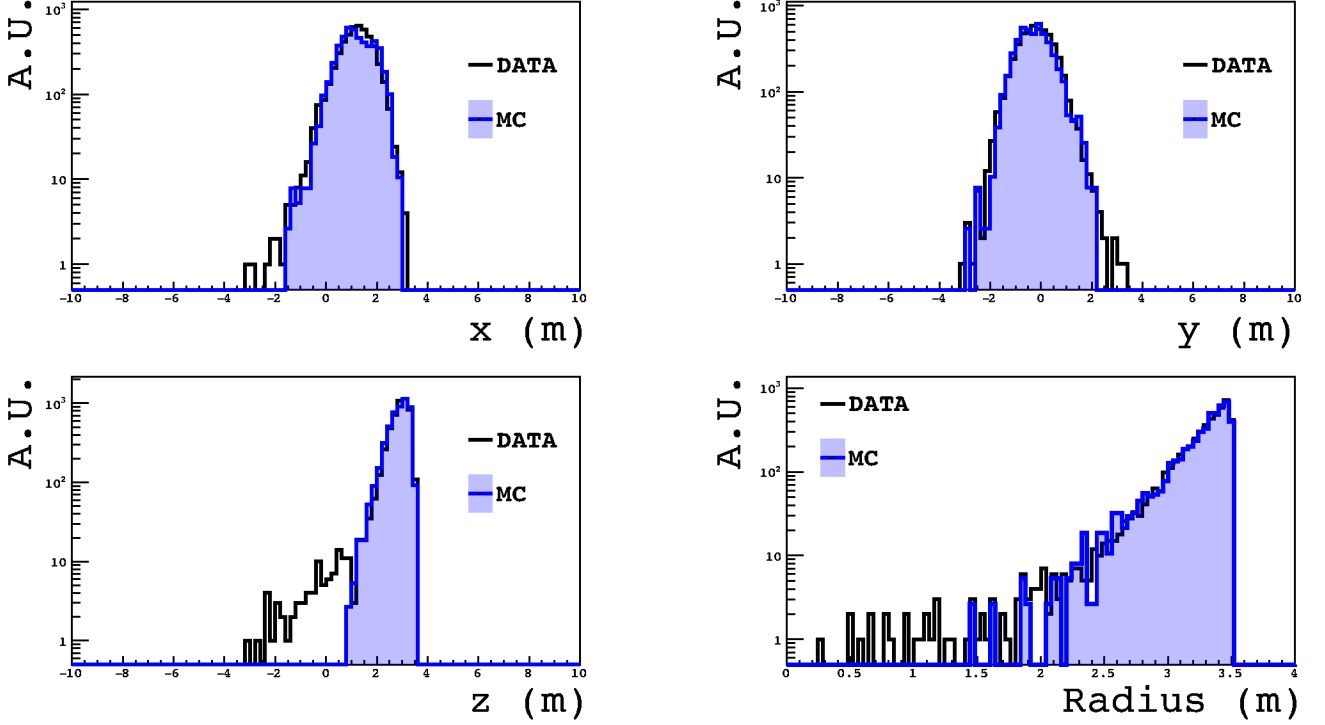


FIG. 24. Comparisons of the reconstructed positions (x , y , and z coordinates and radius) in MC and calibration data for the source deployed in the N3 position from Table VII. The discrepancy between data and MC, mainly visible at negative z values, is due to bulk events (e.g. radioactive decays in the scintillator) which are not external-source related and are not simulated in this specific example.

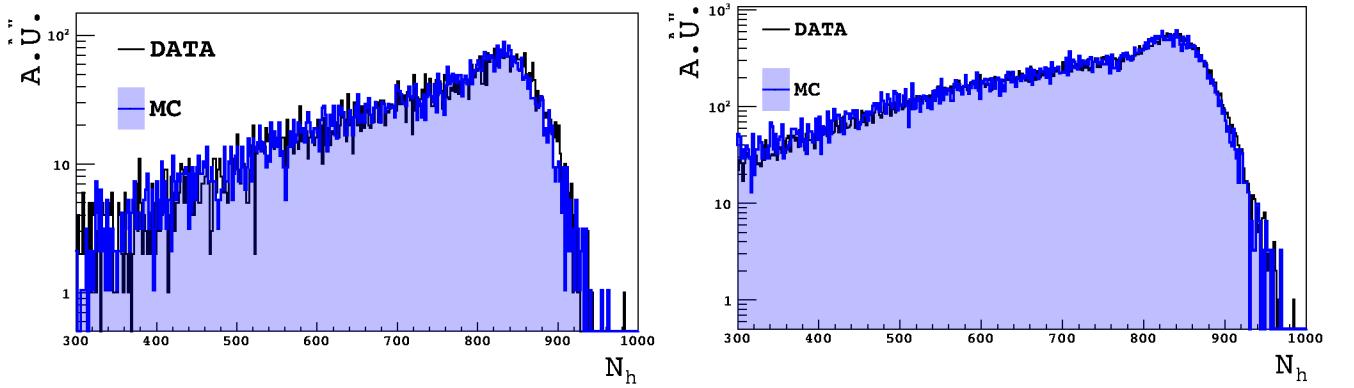


FIG. 25. *Left Panel:* Comparison of the external γ -ray energy spectrum (N_h) for ^{228}Th source-induced events from the N3 position. The fiducial volume is defined as a sphere with 3.5 m radius. *Right Panel:* Comparison of the external γ -ray energy spectrum (N_h) for source-induced events from the N5 position. The fiducial volume cut is a 3 m radius sphere.

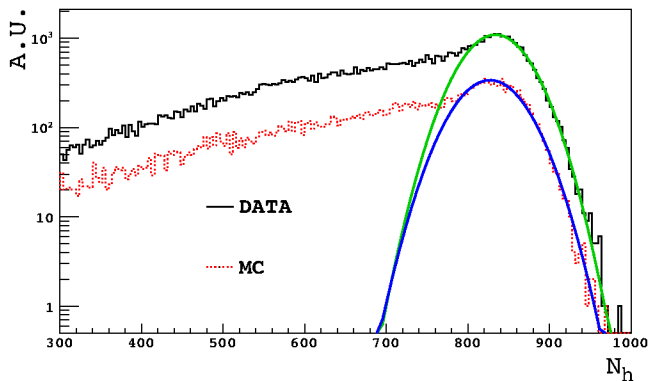


FIG. 26. Gaussian fit examples of the full γ -ray absorption peak for data and MC for the ^{228}Th external calibration source in the N3 position from Table VII. The relative normalization of the two spectra is arbitrary.

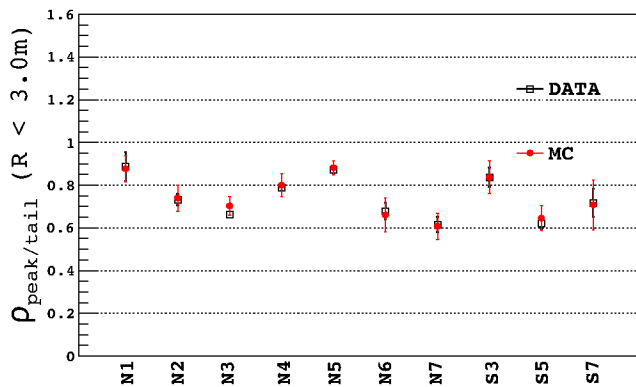


FIG. 27. Value of the full absorption peak/Compton tail ratio as a function of the source position both in MC and in data.

data from calibration sources at defined positions and corresponding MC samples. The agreement between data and MC is shown in Fig. 28. Since no tuning parameters are introduced to reproduce the position reconstruction accuracy, this result is an indication of a reliable modeling of light production, propagation and detection.

In conclusion, the correct reproduction of the position reconstruction resolution validates the accuracy of the MC, while the absence of systematic shifts demonstrates the overall quality of the position reconstruction algorithm.

E. Event pileup

Pileup events are due to two or more scintillation events happening so close in time that the clustering algorithm cannot disentangle them. It is the most critical background for very low energy solar neutrino analyses, particularly for the measurement of the pp neutrino interaction rate [7]. For the analysis presented in Ref. [7], the pileup energy spectrum was built using a data driven

method (we call it “synthetic pileup”) and the energy estimators defined with a fixed cluster duration (i.e. N_p^{dti} , see Sec. VII). We recall that for every triggered event, we register all the hits in a time window of $16.5 \mu\text{s}$, which is much longer than the cluster duration (typically around $1 \mu\text{s}$). The algorithm for the synthetic pileup production considers the hits recorded within a time interval of duration dti in the second half of the gate and superimposes them with the primary cluster, which caused the trigger to occur and is placed at the beginning of the DAQ gate. These synthetic events are then processed with the standard reconstruction code. The pileup spectrum is defined as the set of synthetic events whose energy estimator N_p^{dti} is increased by at least E_{min} , with respect to the corresponding original event, after the overlapping of the additional hits. The major sources of systematics in this method are the choice of E_{min} and the assumption that all pileup events contain at least one primary event which would have triggered on its own. Another critical point is the fact that the PMT dark noise is double counted, since it is automatically contained both in the original clustered event and in the hits which are shifted in time.

The tuned MC simulation allows to reproduce the synthetic pileup and to gain understanding on its different components. In addition, the MC pileup is free of the above mentioned systematic effects.

The MC pileup generation works as follows. Events of given kinds are produced with the tracking code and saved. Two different events are correlated in time by an external application which creates a new file, preserving the same output structure of the tracking code. The time correlation is typically created by sampling a random delay between the events uniformly in a selectable range (typically equal to a few μs). The output so created is then processed through the standard electronics simulation and reconstruction code. More details can be found in Ref. [69]. This procedure has the advantage of replicating what happens in reality: real events overlap at the level of light production in the scintillator, and the already piled-up pattern of hits undergoes the electronics processing and the reconstruction. This is performed in the same way in the simulation. The disadvantage of this method is the need of knowing the most abundant primary classes of events which contribute non negligibly to the pileup. The final pileup spectrum is constructed by summing up the results of the overlap of pairs of events (such as $^{14}\text{C}^{14}\text{C}$, see below for details). As a final remark, this method allows the inclusion in the analysis of those pileup events which might not trigger the detector individually, but which do if they happen close enough in time. This is not possible with the synthetic pileup method.

Among the classes of events with the highest rate, we identified three important ingredients for the pileup: ^{14}C (120 Bq), ^{210}Po (a few hundreds of Bq), and γ -rays of the external background, which interact in the buffer or in the outermost region of the inner vessel. The develop-

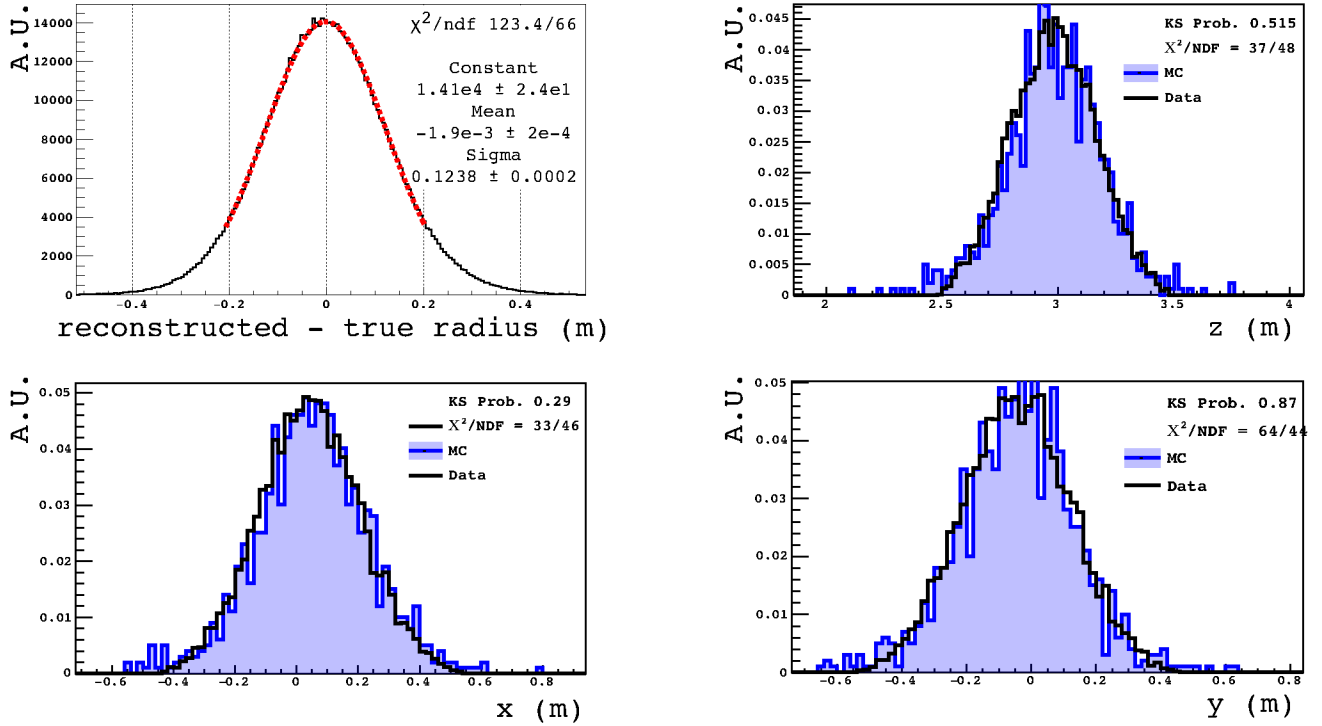


FIG. 28. The top left panel shows the difference between real and reconstructed radius for electron events uniformly simulated in a 3.5 m sphere. The distribution is not Gaussian, thus the χ^2/NDF of the fit is not very good. However, the mean of the distribution is only a couple of mm different from zero. The other panels show the comparison between data and MC of the three cartesian coordinates of the ^{85}Sr source calibration events. The first entry in the legend indicates the outcome of a Kolmogorov-Smirnov test.

ments carried out on the external background simulation (see Sec. VD6) were essential to assess that the order of magnitude of the expected external γ -ray rate is not negligible for the pileup construction. Extrapolating the measurements of the external contaminations inside the inner vessel according to the simulated radial distribution for the outermost shells, we found that the expected count rate of external γ -rays interacting in the whole SSS volume is about 7 kBq. However, it can be easily proven (also by using the MC), that the overlap of two external γ -ray events is negligible in practice. Despite the very high rate of external γ -rays, most of their pileup events are overlaps of a few hits coming from two external γ -ray depositions in the buffer. Given the very low visible energy of these events, almost none of them triggers or is identified as a cluster. Therefore, they can be neglected in the analysis. Thus, the most important components turn out to be:

- ^{14}C overlapped with external γ -rays.
- ^{14}C overlapped with ^{14}C .
- ^{210}Po overlapped with external γ -rays.

The pileup with ^{210}Po results in an effective smearing of the ^{210}Po peak width. Furthermore, the contribution of ^{210}Po overlapped with ^{14}C can be neglected, being of the

second order with respect to that of ^{210}Po with external γ -rays.

At this point, the MC production of these three pileup components is straightforward. However, the relative rates must be known in order to sum them up and produce the final pileup spectrum. We obtained them by fitting the synthetic pileup spectrum with a fixed E_{\min} value with the superposition of pileup spectra produced by the MC and obtained with the application of the same E_{\min} cut. The rates of the primary components can be established with $\sim 2\%$ accuracy. Then, MC pileup spectra are produced with no other modifications or assumptions. Note that in the final MC pileup, there is no E_{\min} parameter. The fit to the synthetic pileup spectrum produced for the analysis shown in Ref. [7]⁶ and $E_{\min} = 6 N_p^{230}$ is shown in Fig. 29. The spectral components used for fitting are ^{14}C overlapped with external γ -rays, ^{14}C overlapped with ^{14}C , and ^{210}Po overlapped with external γ -rays. The fit result provides the relative weights of the different components, depurated from threshold and reconstruction effects.

Once the various contributions were established with

⁶ In Ref. [7], the synthetic pileup spectral shape with $E_{\min} = 5 N_p^{230}$ was used to describe the pileup.

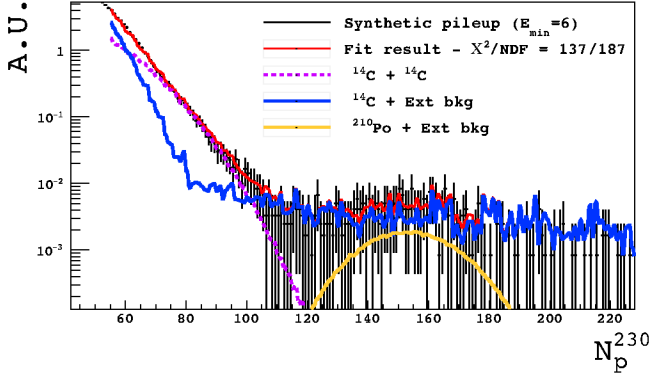


FIG. 29. Fit of the synthetic pileup spectrum with the considered MC pileup spectral components for $E_{min} = 6 N_p^{230}$.

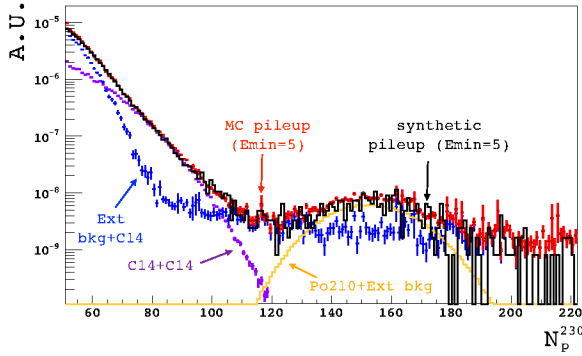


FIG. 30. Synthetic and MC pileup for $E_{min} = 5 N_p^{230}$. The different components are shown in various colors.

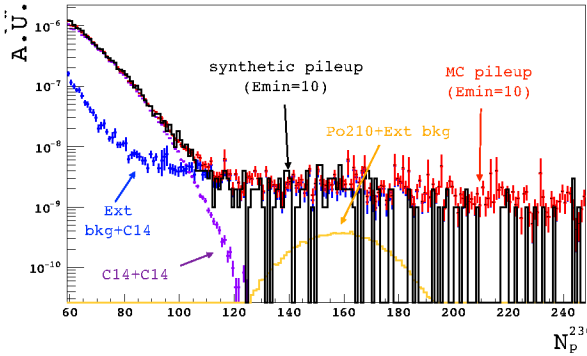


FIG. 31. Synthetic and MC pileup for $E_{min} = 10 N_p^{230}$. The different components are shown in various colors.

$E_{min} = 6 N_p^{230}$, the MC pileup was validated by comparing the predicted pileup shape with that of the synthetic one for other values of E_{min} . The results are shown in Figs. 30 and 31. The pictures show that the synthetic pileup shape changes a lot as a function of E_{min} and the MC pileup is able to reproduce it well. In particular, as E_{min} grows, the contribution of the pileup with external γ -rays is reduced, since the mean visible energy of buffer events is of only a few hits. In the case of $E_{min} = 10 N_p^{230}$, the pileup can mostly be explained by

^{14}C on ^{14}C events, with a higher energy tail of γ -rays interacting in the inner vessel and reconstructed in the FV. The agreement of the MC pileup with the synthetic one tests not only the procedure of the pileup creation and composition, but also the quenching description, the light collection reproduction, the light propagation modeling, and the external background production. Pileup events with a given visible energy are formed by events whose energy, if summed up in real single events, would correspond to a lower quenching, and thus a higher amount of detected light. The energy scale of MC pileup matches well with that of data, giving confidence on the solidity of the quenching description. Furthermore, the hit time distribution is essential for reproducing the behavior of position reconstruction. Pileup events happen “far away” from each other in the detector and thus the behavior of the reconstruction algorithm cannot be established *a priori*, since it is based on a likelihood maximization. The MC pileup spectral composition after the cuts agrees with that of the synthetic pileup, further strengthening the trust in the whole simulation procedure.

X. CONCLUSIONS

In this paper we describe the details of the physical processes and instrumental response implemented in the MC simulation of the Borexino detector. The *ab initio* simulation of the energy loss of particles interacting with the liquid scintillator target and of the generation, propagation and detection of scintillation photons yield better than 1% accurate results for the reconstructed energy and for the time response of the detector for events in the innermost 100 tonnes of Borexino (a 3 m radius sphere). The use of calibration sources has proven essential in providing a large statistical sample of events with well-defined energy and position, complementary to data collected in physics mode.

The Borexino simulations have steadily evolved over time and demonstrate an exquisite level of understanding of the processes governing the light propagation in large volume liquid scintillators, suggesting that the techniques presented in this report could easily be adapted to similar detectors. The implementation of the ionization quenching, the detailed model of the Čerenkov light emission, interaction and timing, and the treatment of the effective quantum efficiency of the PMTs are just a few examples of features that could successfully be exported to other experiments featuring different geometries and scintillator mixtures.

The Borexino MC code is an invaluable tool for all physics analyses and we expect it to play an equivalently crucial role in the forthcoming SOX sterile-neutrino phase of the experiment. The measurement of the solar neutrino interaction rates requires the highest level of accuracy from the MC, which is used to generate the energy spectra of all the components used in the final fit of the recorded energy spectrum. We note that the accuracy of

the energy response reported here exceeds that achieved for previous analyses (see Refs. [3, 47]) while covering a wider energy range (it now extends down to ~ 100 keV). This suggests that a reduction of the systematic uncertainty on the solar neutrino interaction rate over previous measurements is possible.

XI. ACKNOWLEDGEMENTS

The Borexino program is made possible by funding from INFN (Italy), NSF (USA), BMBF, DFG, HGF, and MPG (Germany), RFBR (Grants 16-02-01026 A, 15-02-02117 A, 16-29-13014 ofi-m, 17-02-00305 A) and RSF (Grant 17-02-01009) (Russia), and NCN Poland (Grant No. UMO-2013/10/E/ST2/00180). We acknowledge the generous hospitality and support of the Laboratori Nazionali del Gran Sasso (Italy).

-
- [1] R. S. Raghavan, S. Pakvasa, and B. A. Brown, *Phys. Rev. Lett.* **57**, 1801 (1986).
 - [2] P. Agnes *et al.* (DarkSide), *Phys. Rev.* **D93**, 081101 (2016), [arXiv:1510.00702 \[astro-ph.CO\]](#).
 - [3] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. Lett.* **107**, 141302 (2011).
 - [4] G. Bellini *et al.* (Borexino Collaboration), *Physics Letters B* **707**, 22 (2012).
 - [5] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. Lett.* **108**, 051302 (2012).
 - [6] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. D* **82**, 033006 (2010).
 - [7] G. Bellini *et al.* (Borexino Collaboration), *Nature* **512**, 383 (2014).
 - [8] G. Bellini *et al.* (Borexino Collaboration), *Physics Letters B* **687**, 299 (2010).
 - [9] G. Bellini *et al.*, *Physics Letters B* **722**, 295 (2013).
 - [10] M. Agostini *et al.* (Borexino Collaboration), *Phys. Rev. D* **92**, 031101 (2015).
 - [11] G. Bellini *et al.*, *Physics Letters B* **696**, 191 (2011).
 - [12] M. Agostini *et al.* (Borexino Collaboration), *Astropart. Phys.* **86**, 11 (2017), [arXiv:1607.05649 \[astro-ph.HE\]](#).
 - [13] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. D* **85**, 092003 (2012).
 - [14] M. Agostini *et al.* (Borexino Collaboration), *Phys. Rev. Lett.* **115**, 231802 (2015).
 - [15] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. C* **81**, 034317 (2010).
 - [16] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev. D* **88**, 072010 (2013).
 - [17] G. Bellini *et al.* (Borexino Collaboration), *J. High Energy Phys.* **1308**, 038 (2013).
 - [18] J. Bahcall, "<http://www.sns.ias.edu/jnb>," (2005).
 - [19] M. Apollonio *et al.* (CHOOZ Collaboration), *Eur. Phys. J.* **C27**, 331 (2003), [arXiv:hep-ex/0301017 \[hep-ex\]](#).
 - [20] S. Abe *et al.* (KamLAND Collaboration), *Phys. Rev. Lett.* **100**, 221803 (2008), [arXiv:0801.4589 \[hep-ex\]](#).
 - [21] B. Zeitnitz, *Prog. Part. Nucl. Phys.* **32**, 351 (1994).
 - [22] C. Athanassopoulos *et al.* (LSND Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **388**, 149 (1997), [arXiv:nucl-ex/9605002 \[nucl-ex\]](#).
 - [23] F. Boehm *et al.*, *Phys. Rev.* **D64**, 112001 (2001), [arXiv:hep-ex/0107009 \[hep-ex\]](#).
 - [24] G. Alimonti *et al.* (Borexino Collaboration), *Nucl. Instrum. Meth.* **A600**, 568 (2009), [arXiv:0806.2400 \[physics.ins-det\]](#).
 - [25] F. An *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **685**, 78 (2012).
 - [26] Y. Abe *et al.* (Double CHOOZ Collaboration), *Phys. Rev.* **D86**, 052008 (2012), [arXiv:1207.6632 \[hep-ex\]](#).
 - [27] J. K. Ahn *et al.* (RENO Collaboration), *Phys. Rev. Lett.* **108**, 191802 (2012), [arXiv:1204.0626 \[hep-ex\]](#).
 - [28] C. B. *et al.*, *Nuclear Physics B (Proc. Suppl.)* **118**, 450 (2003).
 - [29] F. An *et al.* (JUNO Collaboration), *J. Phys.* **G43**, 030401 (2016), [arXiv:1507.05613 \[physics.ins-det\]](#).
 - [30] S.-B. Kim, *Proceedings, Neutrino Oscillation Workshop (NOW 2014): Conca Specchiulla, Otranto, Lecce, Italy, September 7-14, 2014*, *Nucl. Part. Phys. Proc.* **265-266**, 93 (2015), [arXiv:1412.2199 \[hep-ex\]](#).
 - [31] A. Gando *et al.* (KamLAND-Zen Collaboration), *Phys. Rev. Lett.* **117**, 082503 (2016).
 - [32] S. Andringa *et al.* (SNO+ Collaboration), *Adv. High Energy Phys.* **2016**, 6194250 (2015).
 - [33] G. Bari *et al.* (LVD Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **A277**, 11 (1989).
 - [34] G. Boireau *et al.* (NUCIFER Collaboration), *Phys. Rev.* **D93**, 112006 (2016), [arXiv:1509.05610 \[physics.ins-det\]](#).
 - [35] Y. J. Ko *et al.*, (2016), [arXiv:1610.05134 \[hep-ex\]](#).
 - [36] A. P. Serebrov *et al.*, *J. Exp. Theor. Phys.* **121**, 578 (2015), [*Zh. Eksp. Teor. Fiz.*148,no.4,665(2015)].
 - [37] J. Ashenfelter *et al.* (PROSPECT Collaboration), (2015), [arXiv:1512.02202 \[physics.ins-det\]](#).
 - [38] J. Hauser (Stereo Collaboration), *Proceedings, 14th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2015): Torino, Italy, September 7-11, 2015*, *J. Phys. Conf. Ser.* **718**, 062023 (2016).
 - [39] P. Agnes and others., *Journal of Instrumentation* **11**, P03016 (2016).
 - [40] DS-20k Collaboration, DarkSide-20k TDR soon released.
 - [41] D. S. Akerib *et al.* (LZ Collaboration), (2015), [arXiv:1509.02910 \[physics.ins-det\]](#).
 - [42] C. Tomei, *14th Vienna Conference on Instrumentation (VCI 2016) Vienna, Austria, February 15-19, 2016*, *Nucl. Instrum. Meth.* **A845**, 418 (2017).
 - [43] M. Chen, F. Elisei, F. Masetti, U. Mazzucato, C. Salvo, and G. Testera, *Nucl. Instrum. Meth.* **A420**, 189 (1999).
 - [44] L. Oberauer, C. Grieb, F. Feilitzsch, and I. Manno, *Nucl. Instrum. Methods Phys. Res., Sect. A* **530**, 453 (2004).
 - [45] G. Alimonti *et al.* (Borexino Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **609**, 58 (2009).
 - [46] G. Bellini *et al.*, *Journal of Instrumentation* **6**, P05005 (2011).
 - [47] G. Bellini *et al.* (Borexino Collaboration), *Phys. Rev.* **D89**, 112007 (2014), [arXiv:1308.0443 \[hep-ex\]](#).

- [48] H. Back *et al.* (Borexino Collaboration), *JINST* **7**, P10018 (2012), [arXiv:1207.4816 \[physics.ins-det\]](#).
- [49] W. Maneschg *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **680**, 161 (2012).
- [50] G. Alimonti *et al.* (Borexino Collaboration), *Nucl. Instrum. Meth.* **A440**, 360 (2000).
- [51] S. Agostinelli *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
- [52] D. Franco, G. Consolati, and D. Trezzi, *Phys. Rev. C* **83**, 015504 (2011).
- [53] V. V. Kuzminov and N. J. Osetrova, *Physics of Atomic Nuclei* **63**, 1292 (2000).
- [54] J. Montara, S. Freedman, B. Fujikawa, and others., *Test of the conserved vector current hypothesis in the beta-decay of ^{14}C* , Tech. Rep. (Argonne Nat. Lab. NSD Report, 1998).
- [55] H. Daniel, *Nuclear Physics* **31**, 293 (1962).
- [56] D. Flothmann, W. Wiesner, R. Löhken, and H. Rebel, *Zeitschrift für Physik* **225**, 164 (1969).
- [57] W. T. Winter, S. J. Freedman, K. E. Rehm, and J. P. Schiffer, *Phys. Rev.* **C73**, 025503 (2006), [arXiv:nucl-ex/0406019 \[nucl-ex\]](#).
- [58] P. C. de Holanda, W. Liao, and A. Yu. Smirnov, *Nucl. Phys.* **B702**, 307 (2004), [arXiv:hep-ph/0404042 \[hep-ph\]](#).
- [59] F. Capozzi, E. Lisi, A. Marrone, D. Montanino, and A. Palazzo, *Nucl. Phys.* **B908**, 218 (2016), [arXiv:1601.07777 \[hep-ph\]](#).
- [60] J. N. Bahcall, M. Kamionkowski, and A. Sirlin, *Phys. Rev.* **D51**, 6146 (1995), [arXiv:astro-ph/9502003 \[astro-ph\]](#).
- [61] K. A. Olive *et al.* (Particle Data Group), *Chin. Phys.* **C38**, 090001 (2014).
- [62] M. Passera, *Phys. Rev.* **D64**, 113002 (2001), [arXiv:hep-ph/0011190 \[hep-ph\]](#).
- [63] A. Vijaya and A. Kumar, *Nucl. Instrum. Methods Phys. Res., Sect. A* **111**, 435 (1973).
- [64] T. Vilaithong *et al.*, *Journal of The Science Society of Thailand* **9**, 129 (1983).
- [65] M. Baldoncini, I. Callegari, G. Fiorentini, F. Mantovani, B. Ricci, V. Strati, and G. Xhixha, *Phys. Rev. D* **91**, 065002 (2015).
- [66] M. Meyer, *SOX - Towards the Detection of Sterile Neutrinos in Borexino : Beta Spectrum Modeling, Monte Carlo Development and Sensitivity Studies for the Sterile Neutrino Search in Borexino*, Ph.D. thesis, Hamburg U., Hamburg (2016).
- [67] C. Arpesella *et al.* (Borexino Collaboration), *Phys. Rev. Lett.* **101**, 091302 (2008).
- [68] G. Collaboration, “*Geant4 user’s guide for application developers*,” .
- [69] S. Marocco, *Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with ^{136}Xe -loaded Liquid Scintillators*, Ph.D. thesis, Gran Sasso Science Institute (2016).
- [70] C. Andreopoulos *et al.*, *Nucl. Instrum. Meth.* **A614**, 87 (2010), [arXiv:0905.2517 \[hep-ph\]](#).
- [71] R. Brun and F. Rademakers, *New computing techniques in physics research V. Proceedings, 5th International Workshop, AIHENP ’96, Lausanne, Switzerland, September 2-6, 1996*, *Nucl. Instrum. Meth.* **A389**, 81 (1997).
- [72] F. Elisei *et al.*, *Nucl. Instrum. Meth.* **A400**, 53 (1997).
- [73] G. F. Knoll, *Radiation detection and measurement; 4th ed.* (Wiley, New York, NY, 2010).
- [74] J. B. Birks, *Proceedings of the Physical Society. Section A* **64**, 874 (1951).
- [75] R. A. MacRae, M. W. Williams, and E. T. Arakawa, *The Journal of Chemical Physics* **61**, 861 (1974).
- [76] R. M. Waxler and C. E. Weir, *Journ. of Res. of the National Bureau of Standards* **67A**, 163 (1963).
- [77] V. Lagomarsino and G. Testera, *Nucl. Instrum. Meth.* **A430**, 435 (1999).
- [78] P. Lombardi, F. Ortica, G. Ranucci, and A. Romani, *Nucl. Instrum. Meth.* **A701**, 133 (2013).