Data selection strategy for solar neutrino analysis with Borexino

Journal of Physics: Conference Series **1342** (2020) 012110 doi:10.1088/1742-6596/1342/1/012110

Data selection strategy for solar neutrino analysis with Borexino

S Caprioli[1], C Ghiano[2], A Re[1] and M Redchuck[3]
(on behalf of the Borexino Collaboration*)

[1] University and INFN Milano, [2]INFN Laboratori Nazionali del Gran Sasso, [3]IKP-2 Forschungzentrum Jülich

E-mail: silvia.caprioli@mi.infn.it

***The Borexino Collaboration**:
M Agostini, K Altenmüller, S Appel, V Atroshchenko, Z Bagdasarian, D Basilico, G Bellini, J Benziger, D Bick, G Bonfini, D Bravo, B Caccianiga, F Calaprice, A Caminata, S Caprioli, M Carlini, P Cavalcante, A Chepurnov, K Choi, O Cloué, L Collica, M Cribier, D D'Angelo, S Davini, A Derbin, X F Ding, A Di Ludovico, L Di Noto, I Drachnev, M. Durero, S Farinon, V Fischer, K Fomenko, A Formozov, D Franco, F Gabriele, J Gaffiot, C Galbiati, M Gschwender, C Ghiano, M Giammarchi, A Goretti, M Gromov, D Guffanti, C Hagner, T Houdy, E Hungerford, Aldo Ianni, Andrea Ianni, N Jonquères, A Jany, D Jeschke, V Kobychev, D Korablev, G Korga, V Kornoukhov, D Kryn, T Lachenmaier, T Lasserre, M Laubenstein, E Litvinovich, F Lombardi, P Lombardi, L Ludhova, G Lukyanchenko, L Lukyanchenko, I Machulin, G Manuzio, S Marcocci, J Maricic, G Mention, J Martyn, E Meroni, M Meyer, L Miramonti, M Misiaszek, V Muratova, R Musenich, B Neumair, L Oberauer, B Opitz, V Orekhov, F Ortica, M Pallavicini, L Papp, Ö Penek, N Pilipenko, A Pocar, A Porcelli, G Ranucci, A Razeto, A Re, M Redchuk, A Romani, R Roncin, N Rossi, S Rottenanger, S Schönert, L Scola, D Semenov, M Skorokhvatov, O Smirnov, A Sotnikov, L F F Stokes, Y Suvorov, R Tartaglia, G Testera, J Thurn, M Toropova, E Unzhakov, C Veyssiére, A Vishneva, M Vivier, R B Vogelaar, F von Feilitzsch, H Wang, S Weinz, M Wojcik, M Wurm, Z Yokley, O Zaimidoroga, S Zavatarelli, K Zuber and G Zuzel

Abstract. The Borexino collaboration has recently released the first simultaneous measurement of the interaction rates of *pp*, ^{7}Be and *pep* solar neutrinos. This result was made possible by the unprecedented low background of the scintillator during Phase-II, together with new data analysis techniques. We present the data selection strategy of the Borexino solar neutrino analysis: we describe how we select the neutrino-like scintillation events according to event-based cuts which eliminate most of the external and cosmogenic backgrounds. We describe also how the spatial distribution of events and a β^+/β^- pulse shape discrimination variable are used in a multivariate fit approach to additionally constrain the residual cosmogenic ^{11}C and external backgrounds.

1. The Borexino Detector

Borexino is a large volume liquid scintillator detector, located deep underground at the Gran Sasso National Laboratories (LNGS, Italy), designed to perform sub-MeV real-time solar neutrino spectroscopy [1]. The active mass consists of 278 t of pseudocumene (PC), doped with 1.5 g/l of PPO. The scintillator is contained in a thin (125 μm) nylon Inner Vessel (IV, 8.5 m in diameter) which is surrounded by a PC buffer doped with a light quencher. The scintillator and buffer are contained in a stainless steel sphere (SSS) with

13.7 m diameter. The SSS is finally enclosed in a Water Tank (WT), containing 2100 tons of ultra-pure water as an additional shield against backgrounds from the laboratory environment. Low energy neutrinos of all flavors are detected in Borexino by means of their elastic scattering on electrons. The electron recoil energy is converted into scintillation light, which is then collected by a set of 2212 photomultiplier tubes (PMTs) and used to reconstruct energy and position of each event. Liquid scintillator detectors have the great advantage of low threshold ($\sim$150 keV), but, since the emission of scintillation light is isotropic, any information about the initial direction of solar neutrinos is lost. This means that neutrino-induced events in Borexino are intrinsically indistinguishable on an event-by-event basis from the background due to β and γ radioactivity. Together with the very low solar ν interaction rate expected in Borexino (from $\sim$few counts to $\sim$100 counts per day (cpd) in 100 t depending on the ν component), this demands extremely low background levels. This has been achieved thanks to Borexino's deep underground location at LNGS, its graded shielding design and restrictive requirements in term of radiopurity of the scintillator and of the detector construction materials. Moreover, in 2010-2011 the scintillator has been further purified through several cycles of water extraction, starting the Phase-II of Borexino. The increased radiopurity, together with a deeper understanding of the backgrounds and an accurate modeling of the detector response in a wide energy range, have recently led to the first simultaneous precision spectroscopy of *pp*, ^{7}Be and *pep* solar neutrinos [2].

2. Data selection for the solar neutrino analysis

The first step of the solar neutrino analysis is to properly select data in order to obtain a spectrum as clean from backgrounds as possible. A series of cuts applied on event-by-event basis has been developed in order to remove taggable backgrounds and non-physical events. The main cuts are listed below and their effect on the spectrum is shown in Fig.2. The fraction of good events removed by these cuts, estimated using Monte Carlo simulations [3] and calibration data [4], is $\sim$0.1%.

2.1. The muon and muon daughter cuts

The 1200 m (3800 m.w.e) thick rock shield reduces the cosmic muon flux by a factor of 10^6 in LNGS, compared to the sea level. However, there are still about $\sim$4300 muons/day crossing Borexino. The left panel of Fig.1 shows the possible scenarios of muons crossing the detector. Muons passing through the scintillator of the Inner Vessel (IV μ's) deposit hundreds of MeV of ionization energy, far above the endpoint of the ^{8}B-ν recoil energy spectrum. For this reason, IV μs per se are not a serious background for solar neutrino observation, but they can produce daughters through their interaction in (or around) the scintillator. Muons crossing the buffer without touching the IV (Bμ's) generate light partly by the Čerenkov effect, and partly by the residual scintillation of the buffer liquid. These situations are dangerous as the Bμ light output is low enough to mimic solar neutrino signal. The amount of produced light depends almost linearly on the length of the muon trajectory in the buffer and cover the whole range of solar neutrino recoil energy. Finally, muons crossing the Outer Detector only (OD μ's) are not worrysome per se but their daughters may be. Muon identification is therefore fundamental in order to both remove the Bμ's and tag possible cosmogenic radioisotopes produced by a muon interacting in the detector. Muon identification relies on two facts [5]: *i)* with a mean energy of 280 GeV, muons crossing Borexino's water tank are well above the Čerenkov threshold of 160 MeV, thus emitting Čerenkov light. This light is detected by an array of 208 PMTs mounted on the outside of the SSS and the Water Tank floor. *ii)* α, β and γ particles' energy deposit within the scintillator is point-like, and their hits arrive almost simultaneously at the PMTs. On the other hand, muons emit light along the track, producing significantly different pulse shape. The muon veto efficiency is calculated to be 99.992% or better.
Hadronic and electromagnetic showers generated by the muons can produce neutrons and a multitude of radioactive isotopes. They can be easily tagged and vetoed via the time coincidence with the parent muon: events occurring in a 300 ms time window after a muon are vetoed. This results in a live time reduction of about 1.5%.

2.2. The Fiducial Volume cut

The photomultiplier tubes, light cones and other materials connected to the SSS, together with the end-cap regions of the inner and outer nylon vessels are sources of external background γ-rays (mostly from ^{208}Tl, ^{214}Bi and ^{40}K decays). These γ-rays (E $\sim$ few MeV), could in principle travel the buffer distance

and reach the Inner Detector, depositing at least part of their energy in the active part of the detector. Since the event rate induced by external γ-rays in the IV reduces esponentially going towards the center of the detector (see right panel of Fig.1), the most effective method to reduce them is through the selection of a fiducial volume (FV). The FV is defined by searching for a volume where the external background is negligible or, at least, strongly suppressed. Of course, the optimal choice of FV should be a compromise between a good signal to background ratio and a non dramatic loss of statistics. The chosen fiducial volume selects the innermost region of the scintillator (71.3 t) contained within the radius $R < 2.8$ m and the vertical coordinate $-1.8 < z < 2.2$ m.

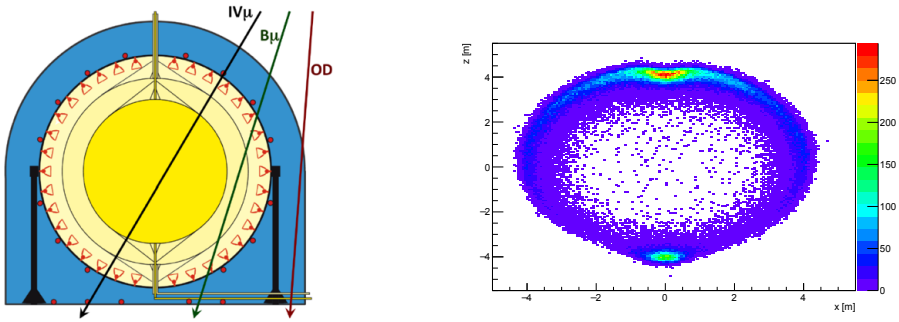

Figure 1: *Left*: Possible configurations of muons crossing the Borexino detector. Muons crossing only the Outer Detector are abbreviated as ODμ's. Muons traversing the Inner Detector can cross the Inner Vessel (IV μ's) or the Buffer only (Bμ's). *Right*: spatial distribution of events in the Inner Vessel (R=4.25 m) in the 900-1500 N_{pe} (number of photoelectrons) energy region, which is dominated by external γ-rays from ^{208}Tl and ^{214}Bi.

2.3. The Three-Fold Coincidence veto

The ^{11}C isotope is continuously produced by muons through spallation on ^{12}C, together with a neutron:

$$\mu + {}^{12}C \rightarrow \mu + {}^{11}C + n. \tag{1}$$

The neutrons are captured with a mean life-time of (254.5±1.8) μs on hydrogen, emitting characteristic 2230 keV γ-rays. The ^{12}C decays with a mean-life τ=29.4 min via positron emission:

$${}^{11}C \rightarrow {}^{11}B + e^{+} + \nu_e. \tag{2}$$

The mean-life of ^{11}C is too long to apply a veto on the whole detector after every muon. The ^{11}C contamination therefore has to be removed with a more sophisticated approach. The three-fold coincidence (TFC) algorithm vetoes space-time regions of the detector after coincidences between the neutron capture and the parent muon, in order to exclude the subsequent ^{11}C (β^+) decay. The reconstruction of the interaction positions of the gamma ray from neutron capture and the tracks of parent muons are crucial for the success of the TFC technique. The guiding principle for the determination of the most appropriate parameters is the search for the optimal compromise between ^{11}C rejection and preservation of the residual exposure after the veto cuts. The details about the algorithm can be found in [6] and [7]. The TFC technique has been recently improved by implementing a new algorithm, which evaluates the likelihood that an event is a ^{11}C candidate, considering relevant observables such as distance in space and time from the parent muon, distance from the neutron, neutron multiplicity, muon dE/dx and the number of muon clusters in an event. Based on this probability, the data-set is divided in two samples: one depleted (*TFC-subtracted*) and one enriched (*TFC-tagged*) in ^{11}C. These two sets are fitted separately in a multivariate fit. The new TFC algorithm has (92±4)% ^{11}C-tagging efficiency, while preserving (64.28±0.01)% of the total exposure in the TFC-subtracted spectrum.

2.4. Other cuts

Besides the three main cuts described previously, other minor cuts are applied. Their effect cannot be appreciated visually on the spectrum, but it is important at the level of fit. We remove radon correlated ^{214}Bi- ^{214}Po delayed coincidences from the ^{238}U and unphysical noise events.

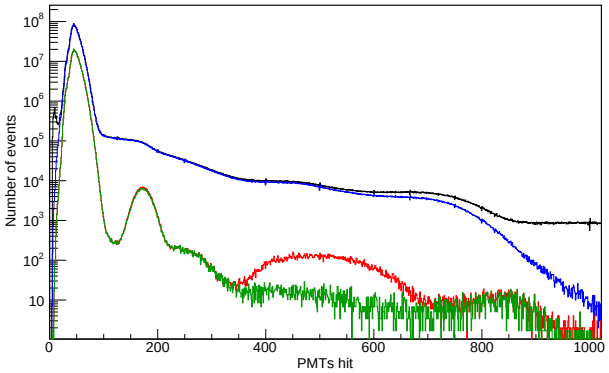

Figure 2: Effect of the cuts on the raw spectrum (shown in black). The blue spectrum shows the effect of the muon and muon daughter cuts. The shape of the final spectrum (in red) is dominated by the effect of the fiducial volume cut. The green spectrum is the TFC-subtracted spectrum, normalized to the same exposure.

3. Residual Backgrounds

The described cuts are very powerful but they are not able to remove all the background events. The residual backgrounds are:

- **^{14}C**: β-emitter with 156 keV end-point and 8270-year mean-life. ^{14}C is chemically identical to ^{12}C and thus it cannot be removed from the organic scintillator through purification. ^{14}C is by far the largest Borexino background and it determines the detector low-energy threshold. The ^{14}C rate is $(3.46 \pm 0.09)\,10^6$ cpd/100 t, about $\simeq 10^5$ times higher than the measured ^7Be–ν rate.
- **^{85}Kr**: β emitter with 687 keV end-point and 15.4-year mean-life. ^{85}Kr rate can be estimated independentely from the spectral fit by detecting the fast β-γ sequence coming from its rare decay into the meta-stable ^{85m}Rb [6].
- **^{210}Bi**: β-emitting daughter of ^{210}Pb with 7.23-day mean-life and 1160 keV end-point. ^{210}Bi background is a very crucial background since its spectrum is very similar to that of CNO neutrinos and the determination of the contribution of both species through the fit is impossible. An indipendent measurement of the ^{210}Bi with a good accuracy is needed to have an handle to single out the CNO contribution. A big effort is currently devoted to developing analysis techiques to determine the content of ^{210}Bi by studying the evolution of its daughter ^{210}Po.
- **^{210}Po**: mono-energetic 5300 keV α-emitter with 200-day mean-life. Due to the strong ionization quenching of the scintillator, the ^{210}Po spectrum appears as a clear peak at about 420 keV, within the ^{7}Be-ν energy region. The identification of ^{210}Po events is quite easy since it decays by alpha emission which can be efficiently separated from β via pulse-shape discrimination techniques.
- **Residual ^{11}C and external γ-rays**: The FV cut and the TFC veto are able to considerably reduce the external and the ^{11}C backgrounds. However, there are still $\sim$2 cpd/100 t of ^{11}C and $\sim$2.5 cpd/100 t of external γ-rays in the final spectrum.

4. Additional constraints on ^{11}C and external backgrounds

The global fit of Borexino data, including the higher energy region for *pep*-ν and CNO-ν detection, needs the additional constraints for the residual ^{11}C and external γ-rays. To do so, the data radial distribution and a β^+/β^- pulse-shape discriminating variable distribution are included in the multivariate fit [2]. The procedures for building the reference PDFs are briefly outlined in the next subsection.

4.1. Radial versus energy distribution of events

The spatial distribution of events in the detector is related to their origin. Neutrino events, cosmogenic and intrinsic backgrounds are uniformly distributed, while external γ's are not, because of the shielding

offered by the scintillator itself. To disentangle external contaminations, the radial versus energy distribution of events is fit using two dimensional PDFs of uniform and external components, obtained with Monte Carlo simulations. Figure 3 shows the radial distributions (integrated over energy) both for uniformly distributed and external background events.

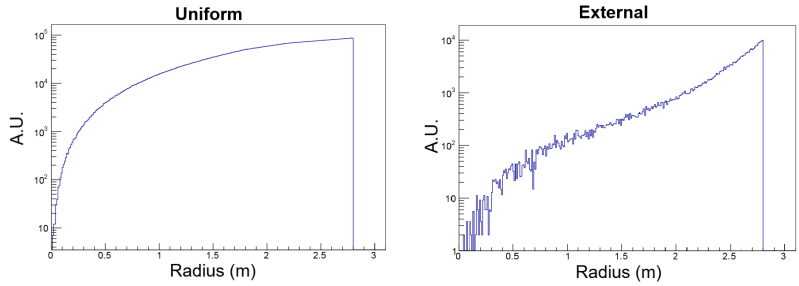

Figure 3: Reference PDFs for the radial distribution (integrated over energy) of uniformly distributed events (*Left*) and of external background events (*Right*) in the fiducial volume.

4.2. β^+/β^- pulse-shape discrimination: the likelihood of the position reconstruction

The scintillation time response produced by e^+ is significantly different compared to that generated by e^-. The e^+ energy deposit, in fact, is not point-like, due to the two back-to-back 511 keV annihilation γ's Moreover, in 50% of the cases, e^+ annihilation is delayed by ortho-positronium formation, the spin-triplet state of positronium. Ortho-positronium has a lifetime of $\tau \sim 3$ ns, which is comparable to the fastest scintillation component of the scintillation pulse, thus introducing a measurable distortion in the time distribution of hit PMTs with respect to pure β^- events [8]. These facts make the β^+ event topology wider in space. To exploit this difference, we use a novel discrimination parameter (PS-$\mathcal{L}_{PR}$), based on the output likelihood of the position reconstruction algorithm [6]. The PS-$\mathcal{L}_{PR}$ reference shape for e^+ is obtained from a pure TFC-selected ^{11}C sample, while for e^- is obtained from simulations (and checked on electron-like events selected from ^{214}Bi-^{214}Po coincidences).

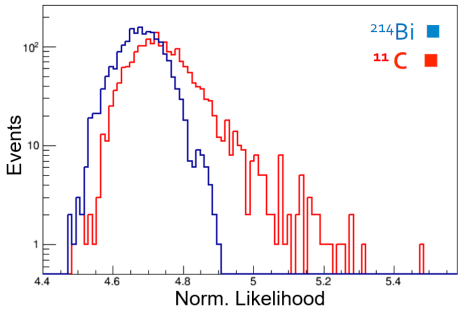

Figure 4: Reference PDFs of the PS-$\mathcal{L}_{PR}$ pulse-shape discriminating variable for β^+ (red) and β^- events. ^{11}C events have in general higher values of the PS-$\mathcal{L}_{PR}$ variable, which indicates a worse reconstructed position.

References

[1] G. Alimonti *et al.* (Borexino Collaboration), *Nuclear Instruments and Methods A*, **600**, 568 (2009).
[2] M. Agostini *et al.* (Borexino Collaboration), arXiv:1707.09279 (2017).
[3] M. Agostini *et al.* (Borexino Collaboration), arXiv: 1704.02291 (2017).
[4] H. Back *et al.* (Borexino Collaboration), *Journal of instrumentation*, **7**, 10018 (2012).
[5] G. Bellini *et al.* (Borexino Collaboration), *Journal of instrumentation*, **6**, 05005 (2011).
[6] G. Bellini *et al.* (Borexino Collaboration), *Physical Review D*, **89**, 112007 (2014).
[7] C. Galbiati *et al.*, *Physical Review C*, **71**, 055805 (2005).
[8] D. Franco, G. Consolati and D. Trezzi, *Physical Review C*, **83**, 015504 (2011).