

# **Looking inside the Sun with the Borexino experiment: detection of solar neutrinos from the proton-proton chain and the CNO cycle**

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH  
Aachen University zur Erlangung des akademischen Grades einer Doktorin der  
Naturwissenschaften genehmigte Dissertation

vorgelegt von

**Mariia Redchuk, M.Sc.**

aus Kyiv, Ukraine

Berichter: Univ.-Prof. Dr. rer. nat. Livia Ludhova

Univ.-Prof. Dr. rer. nat. Achim Stahl

Tag der mündlichen Prüfung: 31.08.2020

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.



**Type I error**

False positive

**Type II error**

False negative

**Type III error**

True positive for incorrect reasons

**Type IV error**

True negative for incorrect reasons

**Type V error**

Incorrect result which leads you to a correct conclusion due to unrelated errors

**Type VI error**

Correct result which you interpret wrong

**Type VII error**

Incorrect result which produces a cool graph

**Type VIII error**

Incorrect result which sparks further research and the development of new tools which reveal the flaw in the original result while producing novel correct results





# *Abstract*

The **Sun** is fueled by fusion processes occurring in its core that convert hydrogen into helium. Photons produced in these reactions take an order of billion years to reach the surface. However, there is another byproduct of nuclear fusion: neutrinos. They are light and electrically neutral, and, unlike photons, escape the Sun in a matter of seconds. These so-called **solar neutrinos** are the only carriers of real-time information about the core of our Star. We know that at least 99% of solar energy is generated through the **proton-proton ( $pp$ ) fusion chain**. One more process through which hydrogen-to-helium fusion may occur is the catalytic **carbon-nitrogen-oxygen (CNO) cycle**. As it is hypothesized to be the main source of energy in heavier stars, its discovery would carry implications in astrophysics, and provide insights about the chemical composition of the core of the Sun, which is not yet fully understood. Moreover, we can exploit this intense natural beam of neutrinos radiated by the Sun to study the phenomenon of neutrino oscillation, the discovery of which was achieved thanks to solar neutrino data.

The **Borexino detector** was designed with the primary goal of detecting the so-called  ${}^7\text{Be}$  neutrinos, originating from the  $pp$  chain. It is particularly suitable for solar neutrino measurement due to its unprecedented radiopurity and resolution at low energies. After ten years of data taking, the Borexino experiment has comprehensively studied all  $pp$ -chain neutrinos, not only fulfilling but even surpassing its purpose. This thesis presents the results and implications of this measurement, as well as the analysis behind it. The next milestone of Borexino was to probe the existence of the CNO cycle in the Sun through the detection of neutrinos produced in it. I will describe my work on the methods of monitoring the evolution of the detector and improving the quality of its data, which was deemed crucial for the CNO neutrino analysis. Concluding my thesis, I will present the analysis methods and preliminary results, which show evidence of the existence of CNO neutrinos.

The work described in this thesis and my accomplishments are achieved thanks to the collective effort of the Borexino collaboration.



# *Zusammenfassung*

Die **Sonne** wird von in ihrem Kern stattfindenden Fusionsprozessen angetrieben, die Wasserstoff in Helium umwandeln. Die in diesen Reaktionen erzeugten Photonen benötigen etwa eine Milliarde Jahre, um die Oberfläche zu erreichen. Es gibt aber ein weiteres Nebenprodukt der Kernfusion: Neutrinos. Sie sind leicht, elektrisch neutral, und im Gegensatz zu Photonen entkommen sie der Sonne in Sekundenschnelle. Diese sogenannten **solaren Neutrinos** sind die einzigen Träger von Echtzeitinformationen über den Kern unseres Sterns. Wir wissen, dass mindestens 99% der Sonnenenergie durch die **Proton-Proton-Fusionskette** ( $pp$ -Kette) erzeugt wird. Ein weiterer Prozess, durch den eine Wasserstoff-Helium-Fusion stattfinden kann, ist der katalytische **Kohlenstoff-Stickstoff-Sauerstoff-Zyklus** (CNO-Zyklus). Da angenommen wird, dass dieser Zyklus die Hauptenergiequelle in schweren Sternen ist, würde seine Entdeckung Auswirkungen auf unser Verständnis der Astrophysik haben. Des Weiteren würde er Erkenntnisse über die, noch nicht vollständig verstandene, chemische Zusammensetzung des Sonnenkerns liefern. Außerdem können wir diesen intensiven natürlichen von der Sonne ausgestrahlten Neutrinostrahl zur Untersuchung der Neutrinooszillation nützen, dessen Entdeckung dank der Daten der solaren Neutrinos erreicht wurde.

Der **Borexino-Detektor** wurde mit dem primären Ziel entwickelt, die sogenannten  ${}^7\text{Be}$ -Neutrinos, die aus der  $pp$ -Kette stammen, zu detektieren. Aufgrund seiner beispiellosen radioaktiven Reinheit und Auflösung bei niedrigen Energien ist er besonders geeignet für die Messung der solaren Neutrinos. Nach zehn Jahren der Datenerfassung hat das Borexino-Experiment alle Neutrinokomponenten der  $pp$ -Kette umfassend untersucht und damit seinen Zweck nicht nur erfüllt sondern sogar übertroffen. Diese Dissertation befasst sich mit den Ergebnissen und Implikationen dieser Messung, sowie der zugrundeliegenden Analyse. Der nächste Meilenstein von Borexino war die Erforschung der Existenz des CNO-Zyklus in der Sonne durch die Detektion der in ihm erzeugten Neutrinos. Ich werde meine Arbeit über die Methoden zur Überwachung der Entwicklung des Detektors und zur Verbesserung der Qualität seiner Daten beschreiben, die als entscheidend für die CNO-Neutrino-Analyse erachtet wurde. Abschließend präsentiere ich die Analysemethoden und die vorläufigen Ergebnisse, die die Existenz von CNO-Neutrinos deuten.

Die in dieser Dissertation beschriebene Arbeit und meine Leistungen wurden dank der kollektiven Anstrengung der Borexino Kollaboration erreicht.



# Acknowledgements

Looking back at the last three years, I realize just how many people contributed to the research behind this thesis, the process of its writing, and most importantly, to my scientific inspiration. This work would not have existed without them.

First of all, I would like to thank my parents, Oleksandr and Olena Redchuk, for giving me the gift of scientific curiosity. They raised me with the habit of asking questions about everything, but also trying to answer them myself first. They gave me the best education I could have had, and made learning fun and engaging.

I would like to acknowledge my supervisor, Liva Ludhova, for the research projects she has given me these years. They helped me develop the experimental physics expertise and analysis skills that will without doubt prove useful in my scientific career. Thanks to her support, I also got to attend analysis meetings, lectures, underground shifts, conferences, and a summer school, all of which packed my brain with knowledge and experience.

And what would I do without my lovely groupmates! Together we have formed a team forged under the challenges of PhD and postdoc life, strengthened by mutual support, and sprinkled with humor and memes. I will never forget the giggles I've had with Zara, Ömer, Sindhu, Sale, Alex, Apeksha and Giulio; and with Philipp, Fabian and Andrea, who don't belong to our Borexino group, but they were basically adopted. I'm very grateful to all the Borexino and JUNO people in Livia's group for being part of my life. I want to say special thanks to Michaela and Christoph, who gave me priceless advice on thesis writing, and prepared me to what was coming.

My research group is a part of a bigger group, the BxAnalysis group, which is, in turn, a part of the Borexino collaboration. All the work I have done, and everything that's in my thesis, is the result of the collective effort of the BxAnalysis team. I will never forget one of the first times I received unsandwiched praise from someone more experienced than me. One older PhD student just told me, "Really nice talk." That was the first time I thought, "Oh my god, I can do things". Many thanks to Barbara, Chiara, Alina, Oleg, Alessandra, Xuefeng, Davide, Alessio, and many others. I will cherish the times when we were indulging in aperitivo, or going hiking in Gran Sasso, as much as working from morning till night together, or having heated discussions during phone calls.



I would like to express my appreciation to Alessandro, for being an inspiration, and my wise mage in this quest. I could wake him up at 3 a.m., and ask him to list all reactor neutrino experiments, and he would. Then he would list all the solar neutrino experiments, and then go on a rant about how awesome Super-K is. And *then* he would conclude by saying that he's a really bad physicist. You are impossible, Ale. We have brainstormed on so many motivation letters and research statements together, that it actually improved my thesis writing skills. Oh, and did I say thanks for reviewing my text, sharing academia experience, dealing with my work-related stress? Thanks! The last weeks were the worst, ey? Thanks also for the mental support, for being my rock.

As for mental support... I would like to thank Sindhu, and the members of the “Taureau” support group, for being patient with my rants, and contributing with their own rants. One for all, all for one. I would also like to mention my therapist, Vicki I. Lännerholm PsyD., LPCC. She has given me amazing, wonderful guidance. When I was anxious or stressed because of work, she managed to tell me that it’s alright to be anxious, and at the same time help me stop feeling that way. Vicki taught me how to be kind to myself, which is a surprisingly hard task, and how to observe and take time to process things, which I was really horrible at. She also gave me good advice on how to approach thesis writing. The last half a year of my PhD contract would have been a much worse experience without her.

A shout-out goes to Silvia, Simone, Davide, Ale, Xuefeng and Pablo Mosteiro (whom I’ve never even met), for writing their nice theses, which I have looked at more than once to get an understanding of how a thesis should look like. I would like to take a moment to give special acknowledgement to Simone. He has been a true Borexino warrior, and we will all remember him and his work.

Lastly, I would like to credit dozens of people who provided me with the resources, without which my work wouldn’t have been possible, that is, the members of Forschungszentrum Jülich, Institut für Kernphysik, Jureca, INFN, CNAF, III. Physikalisches Institut B, and RWTH Aachen.



The Borexino collaboration, June 2017, Gran Sasso.

# Introduction

In this thesis, I am going to present my work on the measurement of solar neutrinos with the Borexino experiment.

**The first four chapters will give you the information needed to understand the background of my research and the motivation behind it.** In Chapter [1 Neutrino physics](#), I will introduce you to the elementary particles called **neutrinos**. I will start by talking about the hundred years of history of neutrino physics, from the time these particles were postulated, to the discovery of their properties, the study of which is still ongoing. In this chapter, you will also find out about different natural and artificial sources of neutrinos, and their role in the Standard Model of particle physics. I will then proceed to describe the mathematical framework behind the phenomenon of *neutrino oscillation*, which plays a key role in theories focused on these particles, as well as experiments aimed at their measurement. In the context of the latter, I will explain the principles behind neutrino detection.

Among various neutrino sources in the universe is the **Sun**. The study of neutrinos originating in it, called **solar neutrinos**, is the main topic of my research, and the subject of Chapter [2 The Sun and solar neutrinos](#). To begin, I will describe the two nuclear fusion processes that can convert hydrogen to helium, the **proton-proton chain**, and the **carbon-nitrogen-oxygen cycle**. Both processes are incorporated into the *Standard Solar Models*, which describe the state and evolution of the Sun, among other assumptions and experimental inputs; and in both, solar neutrinos are emitted as byproducts. By detecting these neutrinos and measuring their fluxes, we can acquire unique information about the energy production in the solar core, which allows us to make conclusions about the thermodynamic equilibrium of the Sun, as well as its chemical composition. Conversely, we can use the intense solar neutrino beam to probe the mechanism behind flavor transformation in the dense medium of the Sun, and its implications for neutrino physics. The information contained in this chapter will be useful to understand the motivation behind the core of my research, and in particular, the analysis shown in Chapter [7](#) and Chapter [8](#).

One of the experiments designed with the goal of detecting and studying solar neutrinos, is the **Borexino experiment**, which will be described in Chapter [3 The Borexino experiment](#). This part of my thesis will get you acquainted with the scientific agenda of this project, and the structure of the Borexino detector. It will provide you with technical information about Borexino hardware and electronics, the data taking procedure, and the methods of data processing, which is needed for the context of the work presented in Chapter [5](#). Last but not least, I will introduce the Monte Carlo simulation of the detector.

In Chapter [4 Backgrounds in Borexino](#), you will find useful information about various isotopes that constitute **background** relevant for solar neutrino analysis, and methods of their evaluation and rejection. In addition, some sources of background events can be exploited for our benefit in different studies and analyses. In Chapter [5](#), I will demonstrate how we use backgrounds to monitor the stability of the detector, while the work done in Chapter [6](#) utilizes the  $^{14}\text{C}$  isotope.

**In the last four chapters, I will dive into the details of my contribution to the Borexino experiment, and the solar neutrino measurement.**

In Chapter 5 [Detector stability and data quality](#), I will present my work on what stands behind the curtains of the solar neutrino analysis stage: ensuring **stable performance of the detector**, and good **reliable data**. To achieve this, the Borexino collaboration has implemented procedures aimed at monitoring data acquisition, and validating the quality of the data. I will begin this chapter by presenting my work focused on the improvement of the data validation procedure. In Borexino, neutrino signals are detected by specially designed photosensors, called *photomultiplier tubes* (PMTs), the time response of which is calibrated on a weekly basis. In the second part of this chapter, I will present the framework I have developed for the purposes of verifying the success of such calibration procedures; while the third and final part will describe my studies on the quality of the PMTs.

One of the parameters used to evaluate their quality will be discussed extensively in Chapter 6 [Effective quantum efficiency of the Borexino photomultiplier tubes](#). After I explain the meaning of **effective quantum efficiency** (EQE), its role in the detector Monte Carlo simulation, and the motivation behind my studies, I will describe different stages of obtaining EQE. My work included improving some of these stages, such as the selection of the data sample that is used for EQE estimation, and the calculation procedure itself; as well as devising methods of ensuring sufficient statistics, necessary to guarantee  $\sim 1\%$  precision for the simulation. At the end of this chapter, you will find the final results based on my work, various ways in which these results have been validated, and their pivotal impact on the accuracy of the detector simulation.

The next two chapters will focus on the approach behind the **solar neutrino measurement with Borexino**.

The **proton-proton chain** (or *pp* chain), mentioned earlier, is known to be responsible for at least 99% of the energy production in the Sun. The analysis that is focused on the measurement of solar neutrinos originating in this fusion process will be presented in Chapter 7 [Comprehensive measurement of the proton-proton chain neutrinos](#). It consists of two separate analyses, that of the low energy region, and the high energy region of the *pp*-chain neutrinos. They will be explained in detail in this chapter, together with the demonstration of my contribution to both. I will conclude this chapter with a discussion on the final results of this measurement, and their implications regarding Standard Solar Models and neutrino physics.

Unlike the *pp* chain, the second fusion process, namely, the **carbon-nitrogen-oxygen (CNO) cycle**, has been residing in the realm of hypotheses until now. Chapter 8 [First evidence of solar neutrinos from the CNO cycle](#) will present the efforts of the Borexino collaboration that resulted in evidence of the existence of the CNO cycle in the Sun. This chapter will guide you through all the aspects of the CNO neutrino analysis: the challenges of CNO neutrino detection, the strategy developed to combat them, and the studies of the sensitivity of the Borexino detector to CNO neutrinos. The research I conducted in Chapter 6 related to the detector simulation was deemed crucial for the analysis of the recent Borexino data, and found its natural continuation in the work presented in this chapter. My contribution to the CNO neutrino analysis consists in the application of the approach based on the Monte Carlo simulation to obtain preliminary results, as well as the evaluation of their significance and related systematic uncertainties.

At the end of my thesis, you will find [Conclusions](#) and [Outlook](#). Enjoy!



# Contents

<b>Abstract</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>Introduction</b>	<b>vii</b>
<b>Abbreviations</b>	<b>xiii</b>
<b>1 Neutrino physics</b>	<b>1</b>
1.1 The history of neutrino . . . . .	1
1.1.1 Discovery of neutrino . . . . .	2
1.1.2 The solar neutrino problem . . . . .	2
1.2 Neutrino sources . . . . .	5
1.3 Neutrinos in the Standard Model of particle physics . . . . .	7
1.4 Neutrino oscillation . . . . .	8
1.4.1 Neutrino oscillation in vacuum . . . . .	8
1.4.2 Neutrino oscillation in matter . . . . .	11
1.5 Neutrino detection . . . . .	12
<b>2 The Sun and solar neutrinos</b>	<b>15</b>
2.1 The proton-proton chain . . . . .	15
2.2 The carbon-nitrogen-oxygen cycle . . . . .	16
2.3 Standard Solar Models . . . . .	17
2.4 Solar neutrinos . . . . .	20
2.4.1 Thermodynamic equilibrium of the Sun . . . . .	22
2.4.2 Ratio of $pp$ and $pep$ neutrino interaction rates . . . . .	22
2.4.3 Probing solar metallicity . . . . .	23
2.4.4 Studying the $pp$ chain terminations . . . . .	24
2.4.5 Probing flavor transformation in matter . . . . .	24
<b>3 The Borexino experiment</b>	<b>27</b>
3.1 Introduction . . . . .	27
3.2 The Borexino detector . . . . .	29
3.2.1 Detector structure . . . . .	29
3.2.2 Particle detection principles . . . . .	30
3.2.3 The Borexino photomultiplier tubes . . . . .	31
3.2.4 The laser calibration system . . . . .	32
3.2.5 Source calibrations . . . . .	33
3.2.6 Purification campaign . . . . .	33

3.2.7	Thermal insulation . . . . .	34
3.3	Data acquisition . . . . .	35
3.3.1	The electronics layout . . . . .	35
3.3.2	The trigger system . . . . .	36
3.3.3	Main DAQ procedure . . . . .	37
3.3.4	Complementary FADC DAQ procedure . . . . .	38
3.4	Data processing and data structure . . . . .	39
3.4.1	Laser timing calibration . . . . .	39
3.4.2	Physical interpretation of raw data . . . . .	40
3.4.3	Energy estimators . . . . .	41
3.4.4	Event types . . . . .	42
3.4.5	Particle identification . . . . .	43
3.5	The Borexino detector simulation . . . . .	46
3.5.1	Geant4 based Monte Carlo simulation . . . . .	46
3.5.2	Electronics simulation . . . . .	47
<b>4</b>	<b>Backgrounds in Borexino</b>	<b>49</b>
4.1	Fiducial volume . . . . .	50
4.2	Isotopes from the $^{238}\text{U}$ chain . . . . .	52
4.2.1	Radon isotope $^{222}\text{Rn}$ . . . . .	53
4.2.2	Pairs of $^{214}\text{Bi}$ - $^{214}\text{Po}$ . . . . .	53
4.2.3	Bismuth isotope $^{210}\text{Bi}$ . . . . .	54
4.2.4	Polonium isotope $^{210}\text{Po}$ . . . . .	54
4.3	Surface and external backgrounds . . . . .	55
4.4	Internal background . . . . .	56
4.4.1	Carbon isotope $^{14}\text{C}$ . . . . .	57
4.4.2	Krypton isotope $^{85}\text{Kr}$ . . . . .	58
4.4.3	Bismuth isotope $^{210}\text{Bi}$ . . . . .	58
4.4.4	Polonium isotope $^{210}\text{Po}$ . . . . .	59
4.5	Cosmogenic background . . . . .	59
4.5.1	Muon detection . . . . .	60
4.5.2	Cosmogenic neutrons . . . . .	61
4.5.3	Carbon isotope $^{11}\text{C}$ . . . . .	61
4.6	Event pileup . . . . .	63
4.6.1	Data driven methods . . . . .	64
4.6.2	Monte Carlo simulation of pileup . . . . .	66
<b>5</b>	<b>Detector stability and data quality</b>	<b>67</b>
5.1	DAQ monitoring and data validation . . . . .	67
5.1.1	Debugging . . . . .	69
5.1.2	New quality checks . . . . .	70
5.2	Laser PMT timing calibration . . . . .	71
5.2.1	Channel time offset . . . . .	71
5.2.2	Goodness of alignment . . . . .	73
5.2.3	Single channel peak deformation . . . . .	74
5.2.4	Laser intensity . . . . .	76

5.3	PMT quality . . . . .	77
5.3.1	Correlation between PMT quality and livetime . . . . .	77
5.3.2	Stable PMT subsets . . . . .	78
<b>6</b>	<b>Effective quantum efficiency of the Borexino photomultiplier tubes</b>	<b>81</b>
6.1	Motivation and the general approach . . . . .	81
6.2	Selection of candle events . . . . .	83
6.2.1	Source of candle events . . . . .	83
6.2.2	Untriggered $^{14}\text{C}$ events . . . . .	83
6.2.3	Standard approach . . . . .	84
6.2.4	Improved approach . . . . .	85
6.3	EQE calculation . . . . .	86
6.3.1	Dark noise correction . . . . .	86
6.3.2	Correction for conic concentrators . . . . .	87
6.3.3	Chosen PMT correction . . . . .	88
6.4	Collecting sufficient statistics . . . . .	91
6.4.1	Time window and radius trade-off . . . . .	91
6.4.2	Dealing with low statistics . . . . .	93
6.5	Conversion of EQE to REQE . . . . .	95
6.6	Results . . . . .	96
6.7	Validation of the new REQE . . . . .	99
6.7.1	Monte Carlo simulation of source calibrations . . . . .	99
6.7.2	Effective light yield . . . . .	100
6.7.3	Phase-II multivariate fit with MCA based on new REQE . . . . .	101
6.7.4	$^{210}\text{Po}$ shift with new REQE . . . . .	102
6.8	Summary . . . . .	103
<b>7</b>	<b>Comprehensive measurement of the proton-proton chain neutrinos</b>	<b>105</b>
7.1	Low energy region analysis . . . . .	105
7.1.1	Data selection . . . . .	106
7.1.2	Fitting method . . . . .	107
7.1.3	Analytical and Monte Carlo approaches . . . . .	111
7.1.4	Systematic uncertainties . . . . .	113
7.2	High energy region analysis . . . . .	115
7.2.1	Data selection . . . . .	115
7.2.2	Fit technique of the $^8\text{B}$ neutrino analysis . . . . .	116
7.2.3	The <i>hep</i> neutrino counting analysis . . . . .	116
7.3	Results and Implications . . . . .	118
7.3.1	Thermodynamic equilibrium of the Sun . . . . .	119
7.3.2	Probing solar metallicity . . . . .	120
7.3.3	Studying the <i>pp</i> chain terminations . . . . .	120
7.3.4	Probing flavor transformation in matter . . . . .	121
<b>8</b>	<b>First evidence of solar neutrinos from the CNO cycle</b>	<b>123</b>
8.1	Challenges of the CNO neutrino detection . . . . .	123
8.2	Phase-II CNO analysis . . . . .	124

8.3	CNO neutrino measurement strategy . . . . .	125
8.3.1	Upper limit on the rate of $^{210}\text{Bi}$ . . . . .	125
8.3.2	Constraint on the rate of <i>pep</i> neutrinos . . . . .	128
8.3.3	Counting analysis . . . . .	128
8.4	Sensitivity studies . . . . .	130
8.5	Phase-III “high setting” multivariate fit . . . . .	132
8.5.1	Multivariate fit with the Analytical Approach . . . . .	133
8.5.2	Multivariate fit with the Monte Carlo Approach . . . . .	134
8.6	Studies of the systematic effects . . . . .	136
8.6.1	Systematic effects in data originating from the MV fit parameters . . . . .	136
8.6.2	Systematic effects based on Monte Carlo datasets . . . . .	138
8.7	Results and conclusions . . . . .	139
8.8	Towards the absolute measurement of the CNO neutrino flux . . . . .	140
	<b>Conclusions</b>	<b>143</b>
	<b>Outlook</b>	<b>145</b>
	<b>A Comprehensive manual of the RunValidation procedure</b>	<b>149</b>
	<b>B Additional information on the EQE of the Borexino PMTs</b>	<b>175</b>
B.1	Time evolution of REQE of single PMTs . . . . .	175
B.2	Evolution of live PMTs . . . . .	179
	<b>C Analytical approach to reference shape construction</b>	<b>181</b>
	<b>Bibliography</b>	<b>193</b>
	<b>Eidesstattliche Erklärung</b>	<b>195</b>

# Abbreviations

<b>AA</b>	Analytical Approach	<b>MCA</b>	Monte Carlo Approach
<b>BTB</b>	Borexino Trigger Board	<b>MLP</b>	Multilayer perceptron
<b>CC</b>	Charged current	<b>MTF</b>	Muon Trigger Flag
<b>CNO</b>	Carbon-nitrogen-oxygen (cycle)	<b>MCF</b>	Muon Cluster Flag
<b>DAQ</b>	Data acquisition	<b>MSW</b>	Mikheev-Smirnov-Wolfenstein (effect)
<b>DMP</b>	Dimethylphthalate	<b>MV</b>	Multivariate (fit or likelihood)
<b>DN</b>	Dark noise	<b>NC</b>	Neutral current
<b>ES</b>	Elastic scattering	<b>NV</b>	Nylon Vessel
<b>EQE</b>	Effective quantum efficiency	<b>OD</b>	Outer Detector
<b>FV</b>	Fiducial volume	<b>OV</b>	Outer Vessel
<b>HER</b>	High Energy Region	<b>p.e.</b>	Photoelectron
<b>HM</b>	High metallicity	<b>PMT</b>	Photomultiplier tube
<b>ID</b>	Inner Detector	<b><i>pp</i></b>	proton-proton (chain)
<b>IDF</b>	Inner Detector Flag	<b>PS</b>	Pulse shape
<b>IV</b>	Inner Vessel	<b>RD</b>	Radial distribution
<b>LER</b>	Low Energy Region	<b>REQE</b>	Relative effective quantum efficiency
<b>LMA</b>	Large mixing angle	<b>SM</b>	Stainless Model
<b>LC</b>	Laser (timing) calibration	<b>SSM</b>	Stainless Solar Model
<b>LM</b>	Low metallicity	<b>SSS</b>	Stainless Steel Sphere
<b>LS</b>	Liquid scintillator	<b>TFC</b>	Three-fold coincidence (method)
<b>LT</b>	Livetime	<b>TT</b>	Trigger type
<b>LY</b>	Light yield	<b>QE</b>	Quantum efficiency
<b>MC</b>	Monte Carlo	<b>WE</b>	Water extraction



# Chapter 1

## Neutrino physics

In this chapter, I will introduce you to the particles called *neutrinos*, the detection of which is the main focus of the Borexino experiment and the analysis conducted using its data, presented in Chapter 7 and Chapter 8. I will start with a historical introduction in Sec. 1.1, summarizing the landmark experiments and theories leading to the discovery of neutrinos and their properties; and the summary of various natural and artificial neutrino sources in Sec. 1.2. Next, I will discuss the role of these particles in the Standard Model (Sec. 1.3), and the most important phenomenon related to neutrinos, namely, *neutrino oscillation* (Sec. 1.4). Sec. 1.5 will conclude this chapter with the description of neutrino detection principles in the context of the Borexino experiment.

### 1.1 The history of neutrino

Table 1.1 presents a brief summary of the important events in the history of neutrino, on which I will elaborate in the sections below.

1914	J. Chadwick shows that electrons emitted in beta decay have a continuous energy spectrum.	[1]
1930	W. Pauli proposes a new particle to explain the phenomenon.	[2]
1933	E. Fermi postulates the first theory of nuclear beta decay.	[3]
1946	B. Pontecorvo proposes radiochemical method of neutrino detection.	[4]
1956	C. Cowan and F. Reines detect reactor antineutrinos.	[5, 6]
1957	B. Pontecorvo discusses the possibility of oscillation $\nu \leftrightarrow \bar{\nu}$ .	[7, 8]
1962	Discovery of muon neutrino by L. Lederman, M. Schwartz, and J. Steinberger.	[9, 10]
1962	Z. Maki, M. Nakagawa, and S. Sakata introduce <i>neutrino oscillation</i> hypothesis and formalism.	[11, 12]
1968	R. Davis and M. Koshiba pioneer the solar neutrino problem with the Homestake experiment.	[13, 14]
1975	Discovery of tau lepton and hypothesis of existence of tau neutrino.	[15]
1985	B. Pontecorvo formulates three-flavor neutrino oscillation.	[16]
1998	The Super-Kamiokande experiment provides strong evidence for flavor oscillation of atmospheric neutrinos.	[17, 18]
2000	Discovery of the tau neutrino by the DONUT collaboration.	[19]
2002	The SNO experiment presents direct evidence of neutrino flavor transformation based on solar $^8\text{B}$ neutrino data.	[20, 21]

TABLE 1.1: Brief history of neutrino.

### 1.1.1 Discovery of neutrino

In the early 1900s, electrons emitted in beta decays used to be considered to be monoenergetic, same as  $\gamma$  and  $\alpha$  particles coming from other known radioactive processes, carrying the full energy resulting from the difference of the final and initial state. The discovery of James Chadwick in 1914 that this is not true and the  $\beta$  spectrum is continuous [1] incited speculations about energy conservation laws on subatomic level. This even prompted Niels Bohr to suggest that conservation laws might hold only on statistical level, and be violated in single decays. Only fifteen years later, Wolfgang Pauli proposed a more reasonable solution in his "Letter to the Radioactive Group" [2]. He postulated a light neutral particle that would be emitted simultaneously together with the electron, carrying the remaining energy and explaining the continuous spectrum.

His idea was further developed by Enrico Fermi, who called the particle *neutrino* and incorporated it into his landmark theory of nuclear beta decay [3]. The prospect of actually detecting this elusive particle seemed dim, due to it interacting with matter only through weak force. However, two decades later, Clyde Cowan and Fred Reines succeeded to detect antineutrinos coming from a nuclear reactor in the historical *Cowan–Reines neutrino experiment* [5], exploiting the reaction we now know as *inverse beta decay* (IBD):

$$\bar{\nu} + p \rightarrow n + e^+. \quad (1.1)$$

At that time, it was not yet known that neutrinos come in three flavors,  $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$ . In 1995, F. Reines shared the Nobel Prize award “for the detection of the neutrino” [6].

### 1.1.2 The solar neutrino problem

In 1946, Bruno Pontecorvo proposes the *radiochemical method* of neutrino detection [4] based on the reaction:

$$\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}. \quad (1.2)$$

This reaction has a threshold of 0.814 MeV, and was used in the *Homestake experiment*, conducted in 1968. Apart from pioneering the detection of neutrinos coming from the Sun (called *solar neutrinos*), which was awarded with a Nobel Prize in 2002 [14], the important implication of this experiment was the so-called *solar neutrino problem*, as it reported a 70% deficit of solar neutrinos compared to solar model predictions [13].

Much earlier, in 1957, Bruno Pontecorvo formulated a hypothesis of the possibility of neutrinos oscillating to antineutrinos [7, 8], based on the similar phenomenon of *neutral kaon oscillation*  $K^0 \leftrightarrow \bar{K}^0$  [22]. The phenomenon of oscillation implied *neutrino mixing*, i.e. the necessity that neutrino be a *superposition* of several eigenstates, similar to  $K^0$  being a “mixed” state of  $K_S$  and  $K_L$ , short- and long-lived weak eigenstates. However, in his scientific works, Pontecorvo considered only one type of neutrino,  $\nu$ , and discussed oscillation  $\nu \leftrightarrow \bar{\nu}$ , as different flavors of neutrino has not been observed yet and resided in the realm of hypotheses.

His idea became the precursor of the theory of *neutrino flavor oscillation*, which was first formulated after the discovery of the *muon neutrino* ( $\nu_\mu$ ) in 1962 [9, 10], when Ziro Maki, Masami



Nakagawa, and Shoichi Sakata have developed a formalism to describe oscillation  $\nu_e \leftrightarrow \nu_\mu$  [11, 12] in its initial simple form:

$$\begin{cases} \nu_e = \nu_1 \cos \delta - \nu_2 \sin \delta, \\ \nu_\mu = \nu_1 \sin \delta + \nu_2 \cos \delta, \end{cases} \quad (1.3)$$

based on the concept of mixing, defining  $\nu_1$  and  $\nu_2$  as *mass eigenstates*, and introducing a single “mixing angle”  $\delta$ . Since then, the theory has been developed and updated to accommodate tau neutrinos, and all the mixing angles measured, more on which will be presented in Sec. 1.4.1.

The significant deficit of solar neutrinos observed by the Homestake experiment mentioned above was actually the first hint supporting this prediction. However, at the time, this result was either put under doubt, as the deficit was so huge, or perceived as an indication that the existing solar models are flawed. It is only much later, in 1985, that Bruno Pontecorvo proposed neutrino oscillation again, incorporating also the third neutrino flavor,  $\nu_\tau$ , hypothesized after the discovery of the tau lepton [15, 16]. The rest of the physics community believed neutrinos to be massless and their flavors to be unchanging. Nevertheless, many experiments followed in the next decades with the aim of testing the Homestake results and resolving the solar neutrino problem. A group of projects called *gallium experiments* were designed with the specific based on the capture of electron neutrinos on gallium:



This reaction has a threshold of 0.233 MeV, lower than that in Eq. 1.2. The first one of them, *Soviet–American Gallium Experiment* (SAGE), began in 1989, and reported its first results after five years of data taking, observing a 40–44% deficit of solar neutrinos [23]. The aptly named *Gallium Experiment* (GALLEX) operated between 1991 and 1997, and was succeeded by the *Gallium Neutrino Observatory* (GNO) experiment running in 1998–2003, both reporting deficit as well.

At that point, the community started to consider the hypothesis of neutrino oscillation seriously. Another experiment called *Super-Kamiokande*, started its operation in 1996, with the primary goal of detecting proton decay, following its predecessor, *KamiokaNDE*. The experiment joined the solar neutrino problem efforts, and already in 1998 provided strong evidence proving the existence of neutrino flavor oscillation [17, 18]. Since Super-Kamiokande is a Cherenkov detector, it has the possibility to measure the direction of the observed neutrinos. This allowed the experiment to show the disappearance of atmospheric muon neutrinos traveling longer distance traversing the Earth, compared to the ones coming from the atmosphere directly above. The final results from the radiochemical experiments and Super-Kamiokande are shown in Table 1.2.

The final resolution of the solar neutrino problem came from the *Sudbury Neutrino Observatory* (SNO), a project designed to detect the so-called solar  ${}^8\text{B}$  neutrinos through their interactions with heavy water (deuterium oxide,  ${}^2\text{H}_2\text{O}$ ). The solar neutrino flux was estimated using three different interactions of neutrinos in heavy water: charged-current (CC) and neutral-current (NC) interactions, as well as elastic scattering (ES) on electrons. The CC interaction provides the measurement of the  $\nu_e$  flux, since only  $\nu_e$  take part in this reaction; while NC provides the total flux of all neutrino flavors. ES interactions involve all flavors as well, however, the scattering cross section on  $e^-$  is six times larger for  $\nu_e$ , making them the dominant contribution

Experiment	Detection method	Measured flux	$\Phi_{obs}/\Phi_{SSM}$
Homestake	$\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}$	$2.56 \pm 0.16 \pm 0.15$	$0.30 \pm 0.03$
Super-Kamiokande	$\nu_{e,\mu,\tau} + e^- \rightarrow \nu_{e,\mu,\tau} + e^-$	$2.32 \pm 0.04 \pm 0.05$	$0.41 \pm 0.01$
Gallium experiments	$\nu_e + {}^{71}\text{Ga} \rightarrow {}^{71}\text{Ge} + e^-$	$69.3 \pm 4.1 \pm 3.6$	$0.52 \pm 0.03$
SNO	(NC) $\nu_{e,\mu,\tau} + d \rightarrow \nu_{e,\mu,\tau} + n + p$	$5.54^{+0.33+0.36}_{-0.31-0.34}$	$1.00 \pm 0.08$
	(CC) $\nu_e + d \rightarrow e^- + p + p$	$1.67^{+0.05+0.07}_{-0.04-0.08}$	$0.29 \pm 0.02$
	(ES) $\nu_{e,\mu,\tau} + e^- \rightarrow \nu_{e,\mu,\tau} + e^-$	$1.77^{+0.24+0.09}_{-0.21-0.10}$	$0.31 \pm 0.05$

TABLE 1.2: Results on the solar neutrino flux from various historical experiments. The third column reports measured fluxes in the units of SNU for radiochemical experiments, and  $10^6 \text{ cm}^{-2} \text{ s}^{-1}$  for the rest [24–28].

to the interaction rate. The resulting prediction combining all the three measurements is shown in Fig. 1.1, while the corresponding reactions and most recently updated flux values are reported in Table 1.2.

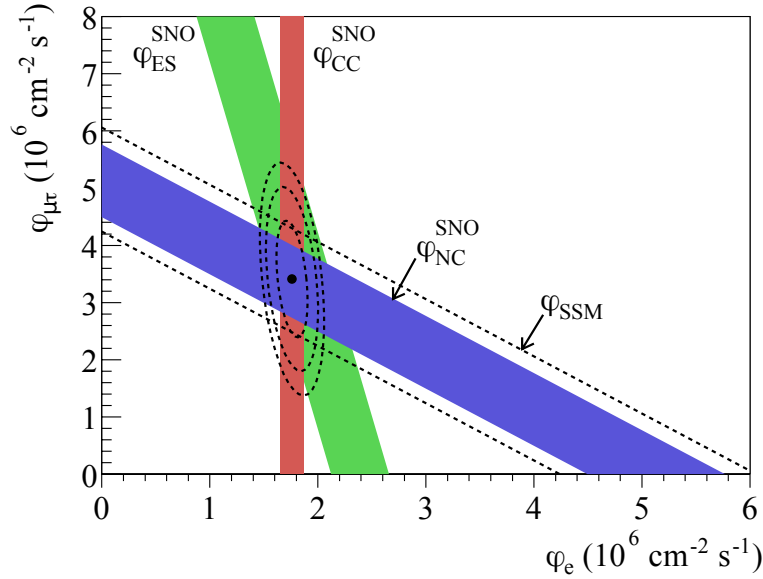


FIGURE 1.1: SNO measurement of the solar  ${}^8\text{B}$  neutrino flux as calculated based on CC, NC and ES detection channels [20, 29].

The neutrino flux measured by SNO through the NC interaction was in accordance with the solar model prediction; while the one measured through the CC channel showed a significant deficit. This effect happens due to the aforementioned fact that all neutrino flavors interact through the NC channel, while only  $\nu_e$  interacts via CC, the deficit representing  $\nu_e$  disappearance due to oscillation to  $\nu_\mu$  and  $\nu_\tau$ . This measurement and the conservation of the total flux of all flavors provided unambiguous proof that part of the solar  $\nu_e$  transform into other flavors during propagation to the Earth. Further measurements and analysis resulted in the confirmation that flavor transformation occurs via the mechanism of neutrino oscillation, which will be described in detail in Sec. 1.4.

## 1.2 Neutrino sources

Neutrinos and antineutrinos are created in processes that involve weak interaction. Figure 1.2 depicts theoretical and experimentally measured fluxes of the neutrinos originating from known natural and artificial sources. Below, I will discuss the major neutrino sources, and the properties of neutrinos they produce.

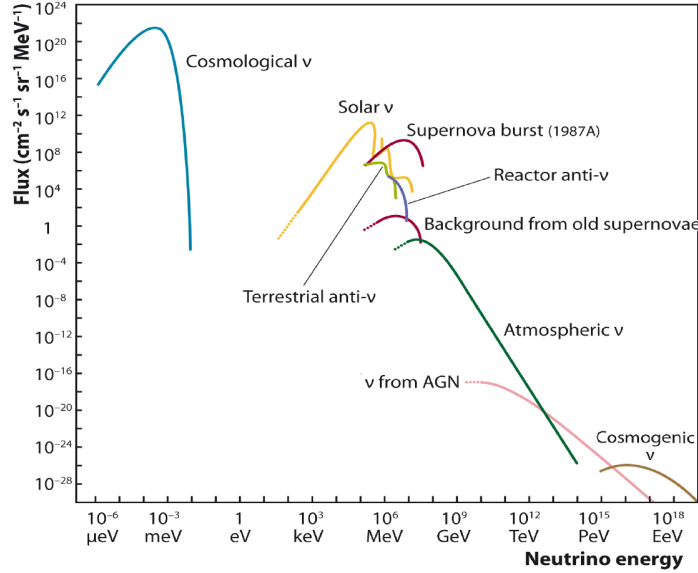


FIGURE 1.2: Measured and expected fluxes of neutrinos and antineutrinos originating from natural and artificial sources [30].

**Cosmological neutrinos**, also called *relic neutrinos*, are the primordial neutrino background, remnant from the earliest stages of the Universe, analogous to the *cosmic microwave background* (CMB). Cosmological neutrinos have not been detected yet, as their energies, spanning from  $\mu\text{eV}$  to  $\text{meV}$ , make their interaction cross sections incredibly small, even though their fluxes are expected to be orders of magnitude above all other known neutrinos, as demonstrated in Fig. 1.2 by the blue curve [31].

Not to be confused with cosmological neutrinos, **cosmogenic neutrinos** are produced in interactions of ultra-high-energy cosmic rays (dominated by protons) with the CMB. They are, in fact, a complete opposite of cosmological neutrinos, reaching extremely high energies of order of  $\text{PeV}$  and higher. The highest energy neutrinos ever observed were registered by the IceCube detector [32], depositing energy of  $2.6 \text{ PeV}$ . Nevertheless, even nine years of IceCube data taking are not yet enough to claim discovery of this type of neutrinos, and can only provide the most stringent limits so far.

**Atmospheric neutrinos** cover the widest energy spectrum, shown in green in Fig. 1.2. They are created in interactions of cosmic rays with the atmosphere, which produce cascades of particles. Among them, charged pions  $\pi^\pm$  decay into muons, producing muon neutrinos:

$$\begin{aligned}\pi^- &\rightarrow \mu^- + \bar{\nu}_\mu, \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu,\end{aligned}\tag{1.5}$$

with a branching ratio of 99.99%. Muons themselves decay as well, producing neutrinos:

$$\begin{aligned}\mu^- &\rightarrow e^- + \bar{\nu}_e + \nu_\mu, \\ \mu^+ &\rightarrow e^+ + \nu_e + \bar{\nu}_\mu.\end{aligned}\tag{1.6}$$

Atmospheric neutrinos played an important role in the observation of neutrino oscillation, as will be mentioned in Sec. 1.4.1.

Another type of neutrinos often used to study neutrino physics are **reactor neutrinos**, produced in  $\beta$  decays of radioactive materials in nuclear reactors, which can be described in general by the formula:

$${}^A_ZX \rightarrow {}^A_{Z+1}X' + e^- + \bar{\nu}_e,\tag{1.7}$$

where  $A$  and  $Z$  are the mass number and the atomic number of an isotope  $X$ , respectively. Reactor neutrinos have been used in a multitude of studies aimed at the measurement of the parameters that govern neutrino oscillation (more in Sec. 1.4.1).

Aside from human-made nuclear reactors,  $\beta$  decays occur very commonly in nature. **Geoneutrinos** are electron antineutrinos produced in decays of long-lived isotopes present in the crust and mantle of the Earth, for instance,  ${}^{238}\text{U}$  and  ${}^{232}\text{Th}$ :

$$\begin{aligned}{}^{238}\text{U} &\rightarrow {}^{206}\text{Pb} + 8\alpha + 8e^- + 6\bar{\nu}_e, \\ {}^{232}\text{Th} &\rightarrow {}^{208}\text{Pb} + 6\alpha + 4e^- + 4\bar{\nu}_e.\end{aligned}\tag{1.8}$$

The measurement of their fluxes provides information about the distribution of these isotopes, and estimate their contribution to the total heat output of the Earth. Geoneutrinos have been detected by the KamLAND experiment [33], and comprehensively studied by Borexino [34–37]

**Supernova neutrinos** are products of *supernovae*, powerful explosions that accompany the collapse of stars, for which the nuclear fusion cannot balance out the gravitational force anymore. Such events are very rare and short, but release tremendous amounts of energy, the vast majority of which is emitted in a burst of neutrinos. The only times supernova neutrinos were detected was during the so-called supernova 1987A, depicted in Fig. 1.2. Neutrinos emerge from the core of the exploding star sooner than photons, and provide crucial information about the event. In order to detect such possible neutrino bursts, and provide an early warning of a galactic supernova, in 2004 an international network of experiments founded a *Supernova Early Warning System* (SNEWS) [38]. It involves seven neutrino experiments, including Borexino.

Last but not least, **solar neutrinos**, are produced in the fusion reactions occurring in the core of the Sun. These neutrinos and their importance in neutrino physics have already been mentioned in Sec. 1.1.2. The fusion processes, the Sun, solar neutrinos and their properties will be discussed extensively in Chapter 2.

### 1.3 Neutrinos in the Standard Model of particle physics

The **Standard Model** (SM) of particle physics describes three fundamental forces, namely, electromagnetic, weak, and strong (excluding only gravitational force). These forces are mediated by *force carriers*, spin-0 particles called *bosons*: photon,  $W^\pm$  and  $Z$  bosons, and gluon, respectively; with the recent addition of the *Higgs boson*, responsible for the mechanism behind mass generation. The SM also classifies twelve spin- $\frac{1}{2}$  elementary particles called *fermions*, six *quarks* and six *leptons*, which are in turn divided into *generations* based on their interactions and behavior. This classification is depicted schematically in Fig. 1.3.

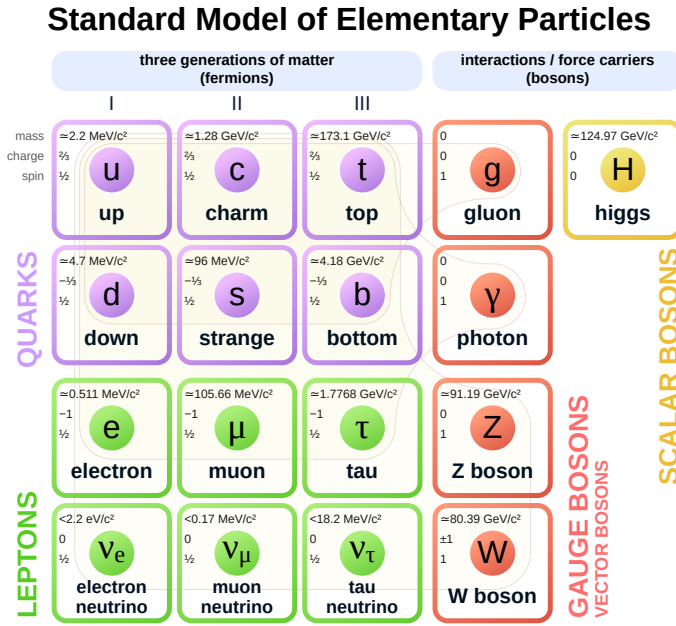


FIGURE 1.3: Elementary particles classified by the Standard Model (*credit: Fermilab, Wikimedia Commons*).

**Neutrinos** belong to the lepton sector of the SM, and appear in three *flavors*: **electron neutrino**  $\nu_e$ , **muon neutrino**  $\nu_\mu$ , and **tau neutrino**  $\nu_\tau$ , corresponding to the three charged leptons, electron  $e^-$ , muon  $\mu^-$ , and tau  $\tau^-$ , respectively. Unlike the latter, neutrinos are electrically neutral and interact only via the weak force and gravity. One can define a quantum number for each generation of leptons: *electronic* ( $L_e = +1$  for  $\nu_e$  and  $e^-$ ), *muonic* ( $L_\mu = +1$  for  $\nu_\mu$  and  $\mu^-$ ), and *tauonic* ( $L_\tau = +1$  for  $\nu_\tau$  and  $\tau^-$ ) *lepton number*. The corresponding antileptons,  $\bar{\nu}_\alpha$  and  $\alpha^+$  for  $\alpha = e, \mu$ , and  $\tau$ , are assigned  $L = -1$ . In each interaction involving leptons, the leptonic number is conserved. Equations 1.9, 1.10, and 1.11 show examples of conservation of  $L_e$ ,  $L_\mu$ , and  $L_\tau$ , respectively:

$$\bar{\nu}_e + p \rightarrow n + e^+, \quad -1 + 0 = 0 + -1 \quad (1.9)$$

$$\pi^+ \rightarrow \mu^+ + \nu_\mu, \quad 0 = -1 + 1 \quad (1.10)$$

$$\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau. \quad 1^\tau = 1^e + (-1)^e + 1^\tau \quad (1.11)$$

Throughout the years, the SM has proven to be a reliable model of particles. Not once have new elementary particles been discovered after being postulated based solely on its coherent symmetrical structure. For example, the tau neutrino was proposed almost immediately after the discovery of the tau lepton [15], its existence logically following from the  $e-\nu_e$  and  $\mu-\nu_\mu$  doublets. Its discovery 25 years later [19] added to the confirmation of the robustness of predictions based on the SM.

Nevertheless, there are still phenomena not described by the SM, many of them related to neutrinos. Even though Eqs. 1.9 - 1.11 above demonstrate the conservation of lepton numbers, this law is put into question by *neutrino flavor transformation*. This phenomenon consists in that neutrinos may change their flavor during propagation, which does not conserve lepton number, making lepton flavor only approximately conserved. Moreover, the discovery of the mechanism behind flavor transformation, called *neutrino oscillation*, implied the existence of three nonzero neutrino mass eigenstates. However, in the SM neutrinos are massless particles.

This shows that the SM does not yet fully incorporate neutrinos, having no framework to accommodate neutrino masses and oscillating flavors, about which I will talk in detail in Sec. 1.4.

## 1.4 Neutrino oscillation

As mentioned in Sec. 1.1.2, the phenomenon of **neutrino oscillation** was conceived and discovered as a solution to the *solar neutrino problem*, explaining the deficit in solar electron neutrinos. In Sec. 1.4.1, I will describe the mathematical framework behind oscillation that occurs when neutrinos propagate in *vacuum*. When they travel through *matter*, the oscillation is subject to the so-called *Mikheev-Smirnov-Wolfenstein effect*, which results in **matter-enhanced oscillation**, which will be discussed in Sec. 1.4.2.

Neutrino oscillation can occur only with the existence of **neutrino mixing**, that is, the neutrino *flavor* states,  $|\nu_e\rangle$ ,  $|\nu_\mu\rangle$ , and  $|\nu_\tau\rangle$ , being a superposition of *mass* states,  $|\nu_1\rangle$ ,  $|\nu_2\rangle$ , and  $|\nu_3\rangle$  (and vice versa). In addition, it implies that the latter have nonzero masses. In this section, I will use the notation  $|\nu_\alpha\rangle$  for flavor states  $\alpha = e, \mu, \tau$ ; and  $|\nu_i\rangle$  for mass states  $i = 1, 2, 3$ . Moreover, when talking about experimental measurements, **disappearance** of a flavor  $\alpha$  denotes a measurement indicating a reduced  $\nu_\alpha$  flux compared to the initial one due to transformation into other flavors; while **appearance** of a flavor  $\beta$  stands for presence of  $\nu_\beta$  in an originally pure  $\nu_\alpha$  flux, due to flavor transformation  $\nu_\alpha \rightarrow \nu_\beta$ .

### 1.4.1 Neutrino oscillation in vacuum

After a neutrino of a flavor  $\alpha$  is created, it propagates as a mixture of mass eigenstates:

$$|\nu_\alpha\rangle = \sum_{i=1,2,3} U_{\alpha i}^* |\nu_i\rangle, \quad (1.12)$$

$$|\nu_i\rangle = \sum_{\alpha=e,\mu,\tau} U_{\alpha i} |\nu_\alpha\rangle, \quad (1.13)$$

where  $U$  is the *neutrino mixing matrix* also known as the *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) matrix, named after Z.Maki, M. Nakagawa, and S. Sakata, who first introduced the

mathematical mixing-oscillation formalism, mentioned in Sec. 1.1.2 [11, 12]; and B. Pontecorvo, who further developed it, incorporating also the third neutrino flavor [16]. For antineutrinos this relation differs only by the complex conjugate being present in the second equation rather than the first. The massive states  $|\nu_i\rangle$  are eigenstates of the vacuum Hamiltonian  $H_0$ :

$$H_0|\nu_i\rangle = E_i|\nu_i\rangle. \quad (1.14)$$

Note that I am going to use *natural units* ( $c \equiv \hbar \equiv 1$ ) in this and the following chapters. The Schrödinger equation for a the neutrino mass state:

$$i\frac{d}{dt}|\nu_i(t)\rangle = H_0|\nu_i(t)\rangle, \quad (1.15)$$

implies that the evolution of  $|\nu_i\rangle$  in time can be approximated as a plane wave:

$$|\nu_i(t)\rangle = e^{-iE_it}|\nu_i\rangle, \quad (1.16)$$

where  $|\nu_i\rangle$  denotes the initial state  $|\nu_i(t=0)\rangle$ . Considering  $t \approx L$ , and substituting Eq. 1.16 into Eq. 1.12, one obtains:

$$|\nu_\alpha(L)\rangle = \sum_i U_{\alpha i}^* e^{-iE_i L} |\nu_i\rangle. \quad (1.17)$$

Finally, using Eq. 1.13 and Eq. 1.17:

$$|\nu_\alpha(L)\rangle = \sum_\beta \sum_i U_{\alpha i}^* e^{-iE_i L} U_{\beta i} |\nu_\beta\rangle. \quad (1.18)$$

Eq. 1.18 mathematically demonstrates the phenomenon that the neutrino created with flavor  $\alpha$  becomes a superposition of flavor states after it travels a certain distance  $L$ . The probability of the flavor  $\alpha$  being detected as flavor  $\beta$  after propagation is:

$$P_{\alpha\beta}(L) = |\langle\nu_\beta|\nu_\alpha(L)\rangle|^2 = \sum_{i,j} U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^* e^{-i\Delta E_{ij}L}, \quad (1.19)$$

where  $\Delta E_{ij} = E_i - E_j$ . Using the relation  $E_i = \sqrt{|\vec{p}_i|^2 + m_i^2}$ , and considering neutrinos to be ultrarelativistic, i.e.  $|\vec{p}_i| \gg m_i$ , the expression for  $E_i$  can be approximated through Taylor expansion as:

$$E_i \approx |\vec{p}_i| + \frac{m_i^2}{2|\vec{p}_i|}. \quad (1.20)$$

Using the result in Eq. 1.20, and combining it with the approximation  $|\vec{p}_i| \approx E$ , where  $E$  is the total energy of the neutrino,  $\Delta E_{ij}$  can be expressed as:

$$\Delta E_{ij} \approx \frac{\Delta m_{ij}^2}{2E}, \quad (1.21)$$

where  $\Delta m_{ij} = m_i^2 - m_j^2$ .

Finally, the phase  $-\Delta E_{ij}t$  in Eq. 1.19 can be written as  $-i\frac{\Delta m_{ij}L}{2E}$ . For convenience, one can express Eq. 1.19 as:

$$P_{\alpha\beta}(L) = \delta_{\alpha\beta} - 4 \sum_{i>j} \Re(A) \sin^2 \frac{\Delta m_{ij}L}{4E} + 2 \sum_{i>j} \Im(A) \sin \frac{\Delta m_{ij}L}{2E}, \quad (1.22)$$

where  $A = U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*$ . The factor inside the sine functions in Eq. 1.22 already tells us several things:

- the observation of oscillation implies that neutrinos are massive, otherwise  $\Delta m_{ij} = 0$  for all  $i, j$ , which would imply  $P(\nu_\alpha \rightarrow \nu_\beta) = \delta_{\alpha\beta}$ , giving  $P(\nu_\alpha \rightarrow \nu_\beta) = 0$  for  $\alpha \neq \beta$ ;
- for the same reason, mass states  $|\nu_i\rangle$  have different masses  $m_i \neq m_j$  for  $j \neq i$ ;
- $L/E$  is the deciding factor for the design of the experiments aimed at observing oscillation.

The very first estimations of  $\Delta m_{ij}$  and  $U_{\alpha i}$  used a simplified two-flavor model. With this assumption, the PMNS matrix  $U$  is parametrized using a single **mixing angle**  $\theta$ , as shown previously in Eq. 1.3:

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (1.23)$$

Using Eq. 1.22, the resulting probability of flavor transformation  $\alpha \rightarrow \beta$  is:

$$P_{\alpha\beta}(L) = \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E}, \quad (1.24)$$

and the corresponding *survival probability* can be obtained simply as  $P_{\alpha\alpha} = 1 - P_{\alpha\beta}$ .

This approximation was used by the KamLAND experiment, assuming a two-neutrino model with only  $\nu_e$  and  $\nu_\mu$ , and observing reactor antineutrino disappearance [39, 40]. The measurement was also performed on solar neutrino data by SNO and Super-Kamiokande, mentioned in Sec. 1.1.2, based on the same two-flavor simplification [41]. The first precision measurement of this mixing angle  $\theta$ , and the **mass splitting**  $\Delta m$  resulted from the combination of KamLAND and solar neutrino experimental results [42, 43]. Note that for solar neutrinos, one has to account for the contribution from *matter-enhanced oscillation* in the Sun (more in Sec. 1.4.2).

The full parametrization of  $U$ , including all neutrino flavors, involves three mixing angles and two mass splittings, and is expressed as follows:

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{\text{atmospheric}} \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix}}_{\text{reactor}} \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{solar}}, \quad (1.25)$$



where  $c_{ij}$  and  $s_{ij}$  stand for  $\cos \theta_{ij}$  and  $\sin \theta_{ij}$ , respectively, and the parameter  $\delta_{\text{CP}}$  is the so-called CP-violating phase. The mixing angles  $\theta_{ij}$  govern the amplitude of the oscillation, while mass splittings  $\Delta m_{ij}$  govern the oscillation frequency. As one can see, the last submatrix in Eq. 1.25, marked “solar”, is indeed the one shown in Eq. 1.23, since it describes the case  $\theta_{13} = \theta_{23} = 0$ , where  $\nu_\tau$  does not participate in any flavor transformation. For this reason,  $\theta_{12}$  and  $\Delta m_{12}$  are often called the *solar mixing angle*  $\theta_{\text{sol}}$  and *solar mass splitting*  $\Delta m_{\text{sol}}$ , respectively, as their first estimations were obtained based on the famous solar neutrino deficit.

In the same fashion, the first measurements of  $\theta_{23}$  and  $\Delta m_{23}$  were performed approximating the PMNS matrix  $U$  as the leftmost submatrix in Eq. 1.25 using a simplified two-flavor model for  $\nu_\mu$  and  $\nu_\tau$ . They were done based on the observation of atmospheric  $\nu_\mu$  and  $\bar{\nu}_\mu$  disappearance by Super-Kamiokande [17], as well as  $\nu_\mu$  disappearance in accelerator-produced beams in the K2K [44, 45] and MINOS [46] experiments. Consequently, these oscillation parameters are often called *atmospheric*, and are denoted as  $\theta_{\text{atm}}$  and  $\Delta m_{\text{atm}}$ .

These legendary works resulted in yet another neutrino Nobel Prize in 2015 awarded to Takaaki Kajita and Arthur B. McDonald, leading the Super-Kamiokande and the SNO experiments, “for the discovery of neutrino oscillation, which shows that neutrinos have mass” [21].

The last mixing angle to have been measured was  $\theta_{13}$  by Daya Bay [47, 48], Double Chooz [49, 50], and RENO [51, 52], which observed disappearance of electron antineutrinos from nuclear reactors. Experiments with a baseline of  $\sim 1$  km can neglect the disappearance of  $\bar{\nu}_e$  governed by  $\theta_{12}$  and  $\Delta m_{12}$ , using the approximation for  $\theta_{13}$  similar Eq. 1.24. Nowadays, the mixing parameters are determined simultaneously through a global fit of accelerator, solar, reactor, and atmospheric neutrino data [53, 54].

### 1.4.2 Neutrino oscillation in matter

When neutrinos travel through matter, rather than vacuum, they are subject to the so-called **Mikheyev-Smirnov-Wolfenstein effect**, which enhances and modifies flavor transformation. The theory was first introduced by Lincoln Wolfenstein in 1977 [55], who suggested that the effect of coherent forward scattering must be taken into account, when considering flavor transformation of neutrinos traveling through matter. Due to this, mass eigenstates  $|\nu_i\rangle$  that propagate in a certain medium acquire different *effective masses* than in vacuum. The theory was later developed by Stanislav Mikheev and Alexei Smirnov [56]. Mathematically speaking, in addition to the vacuum Hamiltonian  $H_0$ , introduced in Eq. 1.14, while propagating through the medium, the neutrino of flavor  $\beta$  is going to feel an effective potential  $V_\beta$ , which would modify Eq. 1.18:

$$|\nu_\alpha(L)\rangle = \sum_\beta \sum_i \left( U_{\alpha i}^* e^{-iE_i L} U_{\beta i} + V_\beta \right) |\nu_\beta\rangle. \quad (1.26)$$

The effect of the potential  $V_\alpha$  can be “absorbed” in the following way:

$$|\nu_\alpha(L)\rangle = \sum_\beta \sum_i U_{\alpha i}^{m*} e^{-iE_i L} U_{\beta i}^m |\nu_\beta\rangle. \quad (1.27)$$

The appearance of the matter potential effectively modifies the initial matrix  $U$ . In this representation, the matrix  $U^m$  has the same form as in the vacuum case, but contains modified mixing angles  $\theta_{ij}^m$ , which depend on the vacuum angles  $\theta_{ij}$ , and parameters defining  $V_\beta$ . In the simplified case of two-flavor neutrino mixing, the relationship between  $\theta^m$  and  $\theta$  is derived using Eq. 1.26 and Eq. 1.27 as follows [57]:

$$\tan 2\theta^m = \frac{\tan 2\theta}{1 - \frac{V_e}{\Delta m^2 \cos 2\theta}}. \quad (1.28)$$

In this equation,  $V_e$  is the potential of the medium felt by  $\nu_e$ :

$$V_e = \sqrt{2}G_F n_e, \quad (1.29)$$

where  $G_F$  is the *Fermi coupling constant*, and  $n_e$  is the electron density of the medium.

This finding was crucial in the context of the solar neutrino problem, discussed in Sec. 1.1.2 [56, 58]. A global analysis of all the available solar neutrino data in terms of matter-enhanced oscillation in a simplified two-flavor model strongly favored the so called the Large Mixing Angle (LMA) solution of the MSW effect. In this context, LMA stands for the value of the vacuum mixing angle  $\theta_{sol}$ , which has been found to be relatively large, as the global fit excluded the region of  $\tan\theta_{sol} > 1$  [59]. In the solar neutrino analysis based on the Borexino experiment, which I will present in Chapter 7 and Chapter 8, one considers the MSW effect on solar neutrinos propagating through the dense medium of the Sun. More details on the specifics of the MSW effect in the Sun will be shown in Sec. 2.4.5.

## 1.5 Neutrino detection

There are different methods of neutrino detection, chosen by one or another experiment depending on their goal and the type of neutrino. Several methods of neutrino and antineutrino detection, like inverse beta decay (Eq. 1.1) or radiochemical detection (Eq. 1.2), have already been mentioned in Sec. 1.1.1 and Sec. 1.1.2, and summarized in 1.2. Liquid scintillator experiments, such as Borexino, detect neutrinos via their elastic scattering on electrons, present in the scintillator:

$$\nu_{e,\mu,\tau} + e^- \rightarrow \nu_{e,\mu,\tau} + e^-. \quad (1.30)$$

The initial neutrino energy  $E_\nu$  is estimated through the measured kinetic energy of the recoiled electrons  $T_e$ :

$$E_\nu = T_e + E'_\nu, \quad (1.31)$$

where  $E'_\nu$  is the energy that the neutrino carries away after scattering. The relationship between  $E'_\nu$  and  $E_\nu$  is expressed as follows:

$$E'_\nu(\theta) = \frac{E_\nu}{1 + \frac{E_\nu}{m_e c^2}(1 - \cos \theta)}, \quad (1.32)$$

where  $m_e$  is the electron mass, and  $\theta$  is the scattering angle.

While all neutrino flavors can participate in this reaction via *neutral current* (NC) interaction, which happens through the exchange of the neutral boson  $Z^0$ , electron flavor neutrinos also take part in the *charged current* (CC) interaction, mediated by  $W^\pm$ . Both types of scattering are illustrated by Feynman diagrams in Fig. 1.4.

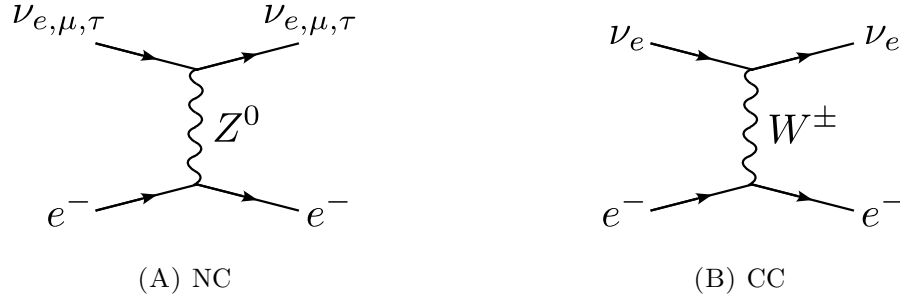


FIGURE 1.4: Feynman diagrams of NC and CC neutrino-electron scattering.

Due to this, the cross section of  $\nu_e$  scattering on  $e^-$  ( $\sigma_e$ ) is around six times larger than that of  $\nu_\mu$  ( $\sigma_\mu$ ) and  $\nu_\tau$  ( $\sigma_\tau$ ). Apart from the type of scattering,  $\sigma_\alpha$  depends on the initial neutrino energy  $E_\nu$ , and the scattering angle  $\theta$ . Using Eq. 1.31 and Eq. 1.32,  $\sigma_\alpha$  can be expressed as a function of  $E_\nu$  and  $T_e$ . The minimum and maximum scattering angles  $\theta_{min} = 0$  and  $\theta_{max} = \pi$  correspond to the minimum and maximum kinetic energy  $T_e^{min} = 0$  and  $T_e^{max}$ , which can be obtained from Eq. 1.32:

$$T_e^{max} = \frac{E_\nu}{1 + \frac{m_e c^2}{2E_\nu}}. \quad (1.33)$$

When scattering on electrons is used as the detection method, the expected neutrino interaction rate can be obtained in the following way, accounting for the differences in the NC and CC scattering:

$$R = N_e \int_0^{T_e^{max}} \int \frac{d\Phi}{dE_\nu} \left( \frac{d\sigma_e(E_\nu, T_e)}{dT_e} P_{ee}(E_\nu) + \frac{d\sigma_{\mu,\tau}(E_\nu, T_e)}{dT_e} [1 - P_{ee}(E_\nu)] \right) dE_\nu dT_e, \quad (1.34)$$

where  $N_e$  is the number of target electrons,  $d\Phi/dE$  is the differential neutrino flux, and  $P_{ee}$  is the electron neutrino survival probability, as defined in Eq. 1.19 (accounting for the MSW effect, if applies). Using an experimental measurement on  $R$ , this formula can be applied for various analyses, depending on the knowledge of  $\Phi$  and  $P_{ee}$ :

- by assuming a theoretical prediction on the flux  $\Phi$ , one can obtain  $P_{ee}$ , and use the result to prove the flavor transformation model (vacuum oscillation, MSW, nonstandard interactions, etc.);
- by assuming MSW-LMA values for  $\theta_{ij}$  and  $\Delta m_{ij}$ , one can probe the theoretical prediction on the flux  $\Phi$ .

Indeed, both analyses are performed based on solar neutrino data. More details on this will be described in Sec. 2.4, and the practical application based on Borexino data will be shown in Sec. 7.3.



## Chapter 2

# The Sun and solar neutrinos

In this Chapter, I will talk about the Sun, the fusion processes occurring in its core, and neutrinos that come from these processes, known as *solar neutrinos*.

In Sec. 2.1, I will describe the so-called *proton-proton (pp) chain*, a set of nuclear reactions that convert hydrogen to helium, known to be the mechanism fueling stars with masses similar to that of the Sun. Another possible way to convert hydrogen to helium is through the catalytic *carbon-nitrogen-oxygen (CNO) cycle*, which I will describe in Sec. 2.2. Even though this process has never been observed yet, theoretical models predict that it might be dominant in stars 1.3 times heavier than the Sun. During different stages of the *pp* chain and the CNO cycle, neutrinos are emitted.

In Sec. 2.3, I will talk about the *Standard Solar Models*, which rely on our knowledge of the aforementioned nuclear fusion processes, and other information about the Sun coming from various measurements, to construct a picture of the solar interior. Based on nuclear physics, one can obtain the theoretical spectral shapes of the solar neutrinos emitted in each reaction; while the SSMs predict the total solar neutrino flux, i.e. the normalization of the shape. In Sec. 2.4, I will discuss the spectral shapes and fluxes of solar neutrinos, as well as how these particles can be used to probe Standard Solar Models, and study neutrino physics.

### 2.1 The proton-proton chain

The **proton-proton chain**, or *pp* chain, is a chain of nuclear reactions that results in the fusion of four protons into helium:



This conversion can happen through different alternative chains, 99.7% of which begin with the proton-proton fusion reaction:



or less frequently, through the three-body reaction:



both producing  ${}^2\text{H}$ , that will be converted to  ${}^3\text{He}$ , and finally,  ${}^4\text{He}$  in the next steps of the chain. The full chain with its three alternative endings, *pp*-I, *pp*-II, and *pp*-III, is represented

schematically in Fig. 2.1A, showing also the branching ratios, and highlighting the neutrino-producing reactions. More about these neutrinos and their spectra and fluxes will be discussed in Sec. 2.4. Figure 2.1B depicts a graphic representation of the *pp*-I ending.

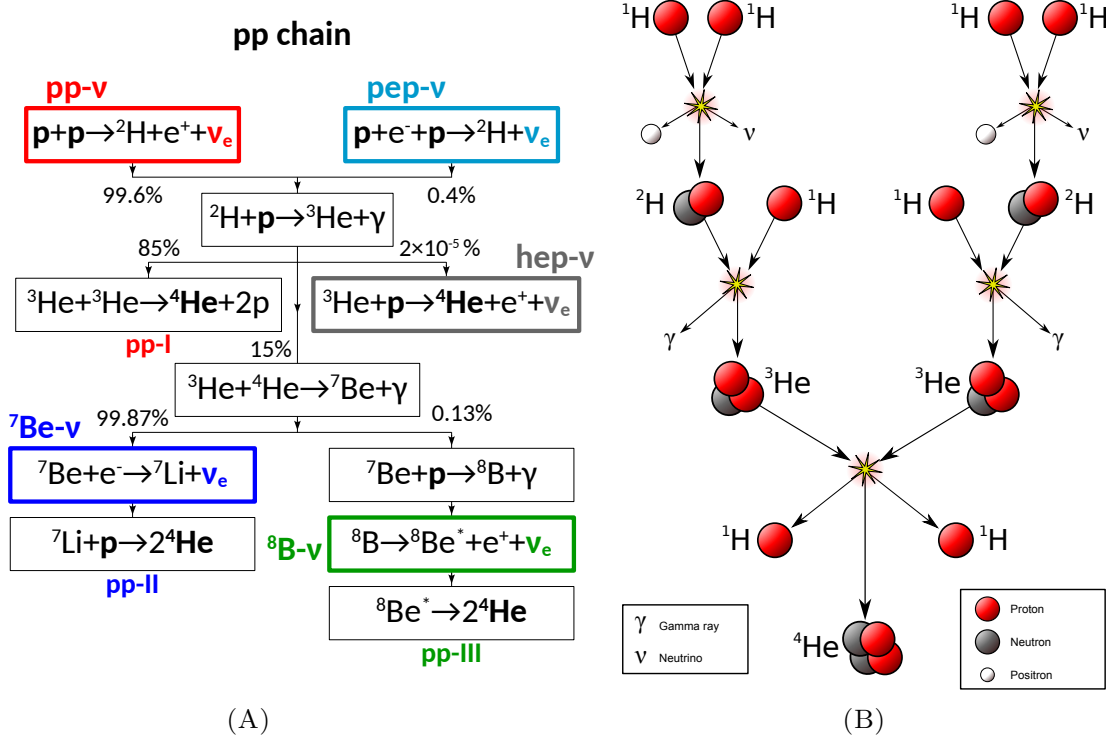


FIGURE 2.1: (A) A schematic representation of the *pp* chain. (B) A graphic representation of the *pp*-I branch (credit: Wikimedia Commons).

The *pp* chain is the dominant process occurring in stars with mass similar or smaller than that of the Sun. The hypothesis that the *pp* chain is the underlying process behind the burning of the Sun and other stars was first formulated by Arthur Stanley Eddington in the 1920s [60]. At that time, this was under doubt, as a simple fusion of two protons would result in an extremely unstable isotope  ${}^2\text{He}$ . This problem was resolved in 1938 by Hans Bethe [61], who proposed that during proton-proton fusion, one proton could decay into a neutron via *inverse beta decay* (Eq. 1.1), resulting in a more stable compound, *deuterium* ( ${}^2\text{H}$ ), as shown in Eq. 2.2.

## 2.2 The carbon-nitrogen-oxygen cycle

Another possible mechanism behind the conversion of hydrogen into helium, alternative to the *pp*-chain, is the closed-loop **CNO cycle**, catalyzed by carbon (C), nitrogen (N), and oxygen (O). It consists of two alternative loops, shown schematically in Fig. 2.2A: CNO-I, having a branching ratio of 99.96%; and CNO-II, with the remaining 0.04%. Since the CNO-I cycle, depicted graphically in Fig. 2.2B, is overwhelmingly dominant, it is usually referred to simply as the CNO cycle. Neutrinos produced in this cycle originate from  $\beta$  decays of  ${}^{15}\text{O}$  and  ${}^{13}\text{N}$ , and have continuous energy spectra with endpoints at around 1.5 MeV and 1.7 MeV, respectively (more in Sec. 2.4).

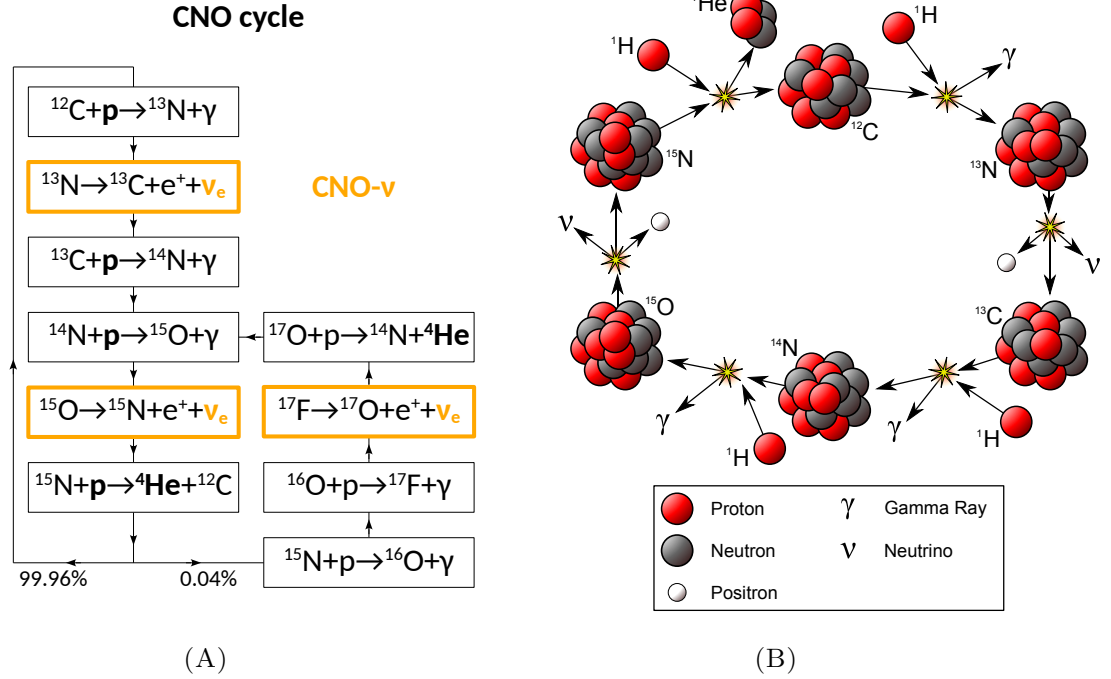


FIGURE 2.2: (A) A schematic representation of the reactions belonging to the CNO cycle, with alternative CNO-I (99.96%) and CNO-II (0.04%) loops. (B) A graphic representation of the CNO-I cycle (*credit: Wikimedia Commons*).

The overall reaction is:

$$4p + 2e^- \rightarrow {}^4\text{He} + 2\nu_e + 7\gamma. \quad (2.4)$$

The CNO cycle is hypothesized to be the dominant energy production mechanism in stars more than 1.3 times heavier than the Sun. It was proposed independently by Carl von Weizsäcker in 1938 [62], and Hans Bethe in 1939 [63]. However, it has never been experimentally observed. If it happens in the Sun, it would account for less than 1% of the solar energy. Proof of the existence of the CNO cycle would carry implications for astrophysics in general, as well as Standard Solar Models, about which I will talk in Sec. 2.3 below.

## 2.3 Standard Solar Models

The structure of the Sun consists of the following concentric layers.

1. **Core**, the innermost region where nuclear fusion occurs.
2. **Radiative zone**, in which energy transfer occurs through radiation of photons.
3. **Convective zone**, in which energy transfer occurs through convection.
4. **Photosphere**, the visible surface of the Sun, from which photons escape.
5. **Atmosphere**, a gaseous halo surrounding the Sun.

The **Standard Solar Model** (SSM) is a mathematical description of the Sun, in which its state and evolution are described by a set of equations based on the following assumptions:

- the **solar energy** is generated through nuclear fusion of hydrogen into helium in the core of the Sun, occurring through the *pp* chain (Sec. 2.1), contributing to  $> 99\%$  of the energy output; and the CNO cycle (Sec. 2.2), the contribution of which is yet unknown;
- the **energy transport** from the core to the surface of the Sun occurs by radiation and convection;
- the **solar composition** is defined by the initial abundance of elements in the Sun, the only changes to it coming from the nuclear fusion reactions;
- the Sun is in a state of **thermal equilibrium**, with the energy produced in the nuclear reactions being carried away by photons and neutrinos;
- the Sun is in a state of **hydrostatic equilibrium**, in which gravitational force is balanced out by the pressure gradient resulting from thermonuclear reactions.

In addition to these assumptions, the models also rely on physical parameters of the Sun that constitute boundary conditions, namely :

- **luminosity**, measured directly through solar irradiance to be  $3.828 \cdot 10^{26} \text{ W}$ ;
- **age**, estimated to be  $4.57 \pm 0.01 \text{ Gyr}$ , based on radioactive dating of meteorites, and rocks on the Earth and the Moon;
- **mass**, computed to be  $1.988 \cdot 10^{33} \text{ kg}$ , based on the orbit equation of the Earth;
- **radius**, calculated to be  $6.955 \cdot 10^8 \text{ m}$ , based on solar luminosity and surface temperature;
- **initial element abundances**, some of which can be inferred through spectroscopy and helioseismology.

Even though the development of 3D hydrodynamic models of the solar atmosphere has led to a complete revision of solar element abundances [64–66], this parameter remains the least known out of the ones listed above. A crucial parameter related to solar abundance, is the so-called **solar metallicity**, the fraction of elements heavier than helium, called *metals* in the context of solar physics. The newest generation of SSMs, called **B16** [67], employs two central sets of solar abundances, based on photospheric and meteoritic abundances from different inputs, and results in different predictions of solar metallicity:

- **GS98** [68], resulting in a metal-to-hydrogen ratio  $(Z/X)_{\odot} = 0.0229$ ;
- **AGSS09met** [65], resulting in  $(Z/X)_{\odot} = 0.0178$ .



Individual abundances of the metals most relevant for the model are shown in Table 2.1. Based on the different different inputs, the B16 SSMs are divided into two sets, B16-GS98 and B16-AGSS09met.

Element	GS98	AGSS09met
C	$8.52 \pm 0.06$	$8.43 \pm 0.05$
N	$7.92 \pm 0.06$	$7.83 \pm 0.05$
O	$8.83 \pm 0.06$	$8.69 \pm 0.05$
Ne	$8.08 \pm 0.06$	$7.93 \pm 0.10$
Mg	$7.58 \pm 0.01$	$7.53 \pm 0.01$
Si	$7.56 \pm 0.01$	$7.51 \pm 0.01$
S	$7.20 \pm 0.06$	$7.15 \pm 0.02$
Ar	$6.40 \pm 0.06$	$6.40 \pm 0.13$
Fe	$7.50 \pm 0.01$	$7.45 \pm 0.01$
$(Z/X)_\odot$	0.02292	0.01780

TABLE 2.1: Abundances of the solar elements that make the biggest contribution to the uncertainties in SSMs, given as  $\ln(N_i/N_H) + 12$ . The parameter  $Z/X$  denotes the total metal-to-hydrogen ratio [65, 67, 68].

One of the methods of assessing the performance of the SSMs with different abundance inputs, is provided by *helioseismology*, study of the propagation of acoustic waves in the solar surface [69, 70]. Figure 2.3 shows the fractional sound speed difference in the Sun, predicted by different SSMs, depending on the radius.

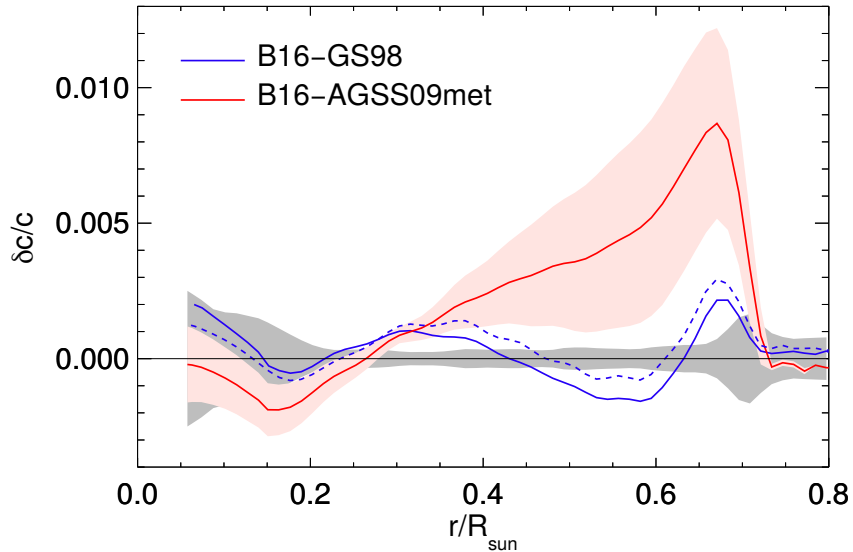


FIGURE 2.3: Fractional sound speed difference  $\delta c/c = (c_\odot - c_{mod})/c_{mod}$ , where  $c_\odot$  is the measured value, and  $c_{mod}$  is the value predicted by the given SSM. The results from B16-AGSS09met (red), B16-GS98 (solid blue), and SFII-GS98 (dashed blue) SSMs are shown. The shadowed red region corresponds to the  $1\sigma$  theoretical uncertainty [67].

The prediction based on B16-GS98 (solid blue) shows the largest discrepancy with the solar sound speed profile in the region  $0.65 < r/R_\odot < 0.7$ , removing which brings the agreement to  $1.4\sigma$ . For B16-AGSS09met (red), however, the disagreement is global, and even removing the region mentioned above leads to a  $2.7\sigma$  discrepancy. This mismatch is caused mainly by the abundance inputs, while the updated SSMs do not cause much difference. For comparison, an older SSM, SFII-GS98 [71], is shown in dashed blue.

In addition, two other helioseismic quantities can be used to judge the quality of SSMs: surface helium abundance ( $Y_s$ ), and the location of the bottom of the convective zone ( $R_{CZ}$ ) relative to the radius of the Sun ( $R_\odot$ ). Both are shown in Table 2.2, together with the average sound speed difference, shown in Fig. 2.3. The AGSS09met inputs are based on the most recent updated revision of solar abundances, and, as one would expect, the B16-AGSS09met predictions for  $Y_s$  and  $R_{CZ}/R_\odot$  show very good agreement with solar values ( $0.5\sigma$ ), compared to B16-G98 ( $2.1\sigma$ ). Nevertheless, the discrepancy with the solar sound speed profile indicates that the revised inputs might not be correct.

Quantity	B16-GS98	B16-AGSS09met	Solar
$Y_s$	$0.2426 \pm 0.0059$	$0.2317 \pm 0.0059$	$0.2485 \pm 0.0035$
$R_{CZ}/R_\odot$	$0.7116 \pm 0.0048$	$0.7223 \pm 0.0053$	$0.713 \pm 0.001$
$\langle \delta c/c \rangle$	$0.0005^{+0.0006}_{-0.0002}$	$0.0021 \pm 0.001$	0

TABLE 2.2: Surface helium abundance ( $Y_s$ ), location of the bottom of the convective zone ( $R_{CZ}$ ) relative to the radius of the Sun, and average fractional sound speed difference, predicted by different SSMs in comparison to the helioseismic measurements [67].

A parameter that can provide us with important information that could solve the solar abundance problem, is the aforementioned *solar metallicity*. As shown in Table 2.1, different abundance predictions result in different values of metal-to-hydrogen ratio  $(Z/X)_\odot$ . Solar metallicity can be probed by looking at the fluxes of neutrinos emitted in the  $pp$  chain and the CNO cycle, highlighted in Fig. 2.1A and Fig. 2.2A, respectively. The flux of CNO neutrinos depends directly on the abundance of carbon, nitrogen, and oxygen; while the flux of neutrinos coming from the  $pp$  chain is determined by its fusion rates, which depend indirectly on solar metallicity, as it influences solar opacity, and, in turn, the core temperature. Based on  $(Z/X)_\odot$ , from now on B16-GS98 will be referred to as **high metallicity** (HM) SSM, and B16-AGSS09met as **low metallicity** (LM) SSM. More on solar neutrinos and how to use their fluxes to probe HM- and LM-SSMs will be described in Sec. 2.4 below.

## 2.4 Solar neutrinos

**Solar neutrinos** are electron flavor neutrinos produced in nuclear fusion processes in the core of the Sun, namely, the  $pp$  chain (Sec. 2.1), and the CNO cycle (Sec. 2.2). The  $pp$ -chain neutrinos were detected for the first time in the 1960s [13], and since then thoroughly studied, and comprehensively measured during the recent years [72]. The first evidence of the existence of solar CNO neutrinos was obtained only in the present day, resulting from the analysis conducted by the Borexino collaboration, which will be presented in this thesis in Chapter 8.

Each neutrino produced in the  $pp$  chain is called by the name of the reaction it was produced in:  **$pp$**  ( $p + p$ ),  **$pep$**  ( $p + e^- + p$ ),  **${}^7\text{Be}$**  ( ${}^7\text{Be} + e^-$ ),  **${}^8\text{B}$**  ( ${}^8\text{B}$  decay), and  **$hep$**  ( ${}^3\text{He} + p$ ) neutrinos. In a similar fashion, neutrinos originating from the CNO cycle are labeled  **${}^{13}\text{N}$** ,  **${}^{15}\text{O}$** , and  **${}^{17}\text{F}$** , corresponding to the isotope in the beta decay of which it was produced. Figure 2.4 shows the theoretical spectral of solar neutrinos, based on nuclear physics [73], with fluxes predicted by B16 SSM [67]. The first calculations of the spectral shapes and predictions on solar neutrino fluxes were made by John Bahcall in the 1960s [74]. Schematic representations of the  $pp$  chain and the CNO cycle, introduced earlier in Fig. 2.1A and Fig. 2.2A, respectively, are repeated for reference.

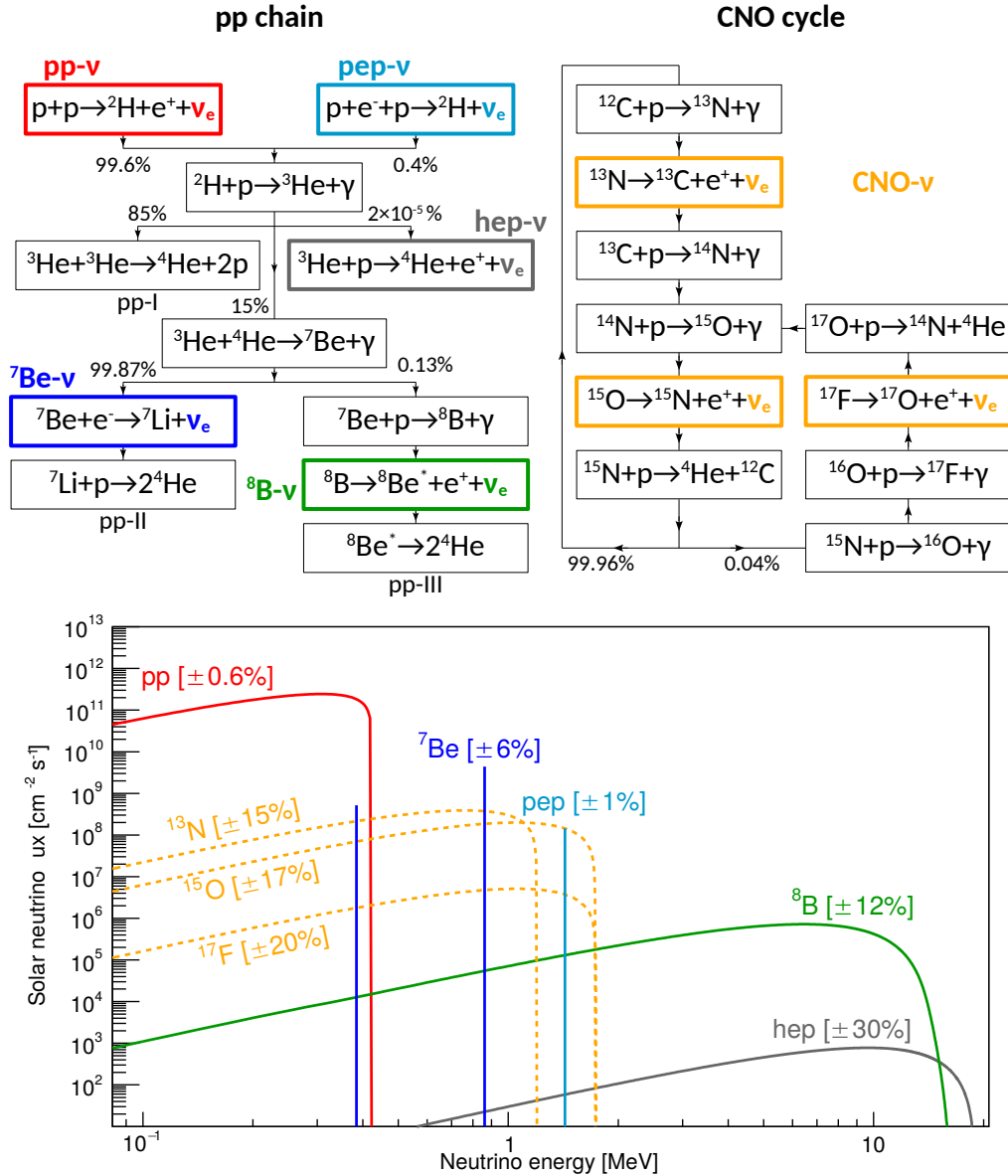


FIGURE 2.4: Theoretical spectral shapes of solar neutrinos produced in the  $pp$  chain and the CNO cycle [73]. The shapes are normalized to fluxes predicted by B16 SSM [67] are shown in units of  $\text{cm}^{-2}\text{s}^{-1}\text{MeV}^{-1}$  for continuous spectra, and  $\text{cm}^{-2}\text{s}^{-1}$  for monoenergetic lines.

In the 1950s, William Fowler and others, who comprehensively studied the  $pp$  chain [75, 76], pointed out that solar neutrinos can be used to test solar models. These particles are the only direct probe of the core of the Sun, since they can escape to the surface in a matter of seconds, as opposed to photons, for which the time scale is of order of  $10^5$  years. In the following subsections, I am going to talk about the ways in which the measurement of solar neutrino fluxes can be used for studying the Sun and probing SSMs, as well as the MSW effect, presented in Sec. 1.4.2.

### 2.4.1 Thermodynamic equilibrium of the Sun

As mentioned in Sec. 2.3, one of the fundamental assumptions of SSMs is that the Sun is in a state of thermal equilibrium, meaning that the solar energy measured through photons is equal to the energy produced in the fusion reactions in its core [77, 78]. This assumption can be tested by comparing the total photon output and the measured total solar neutrino flux  $\Phi_\nu$ :

$$\Phi_\nu = \frac{L_\odot}{4\pi d^2 E_\nu}, \quad (2.5)$$

where  $L_\odot = 3.828 \cdot 10^{26}$  W is the previously mentioned solar luminosity,  $d = 1.5 \cdot 10^{11}$  m is the average distance from the Earth to the Sun, and  $E_\nu$  is the average energy of a solar neutrino. Every  $4p \rightarrow {}^4\text{He}$  fusion produces two solar neutrinos, as seen in Eq. 2.1 and Eq. 2.4, with released energy  $Q = 26.7$  MeV, meaning that each neutrino carries away at most  $E_\nu = Q/2 = 13.35$  MeV. Based on this, the total solar neutrino flux can be roughly estimated to be of the order of  $10^{10} \text{ cm}^{-2} \text{ s}^{-1}$ . This prediction is in agreement with the solar neutrino measurements, in particular, ones resulting from the Borexino solar neutrino analysis, more on which will be discussed in Sec. 7.3.1.

Conversely, by taking the thermal equilibrium assumption as valid, one can place a constraint on solar neutrino fluxes. For example, in the so-called Phase-III analysis of the Borexino experiment (which will be introduced in Chapter 3), aimed at CNO neutrino detection, this so-called *solar luminosity constraint* is applied to estimate the flux of  $pep$  neutrinos [79], which will be discussed in Sec. 8.3.2.

### 2.4.2 Ratio of $pp$ and $pep$ neutrino interaction rates

Another estimation that can be used to benefit the CNO neutrino analysis is that of the ratio between the interaction rates of  $pp$  and  $pep$  neutrinos. Since the  $pp$  fusion (Eq. 2.2) and the three-body  $pep$  reaction (Eq. 2.2) depend on the same *nuclear matrix element*, the ratio between the fluxes and, therefore, measured interaction rates is independent from nuclear physics. The ratio is predicted to be  $47.8 \pm 0.8$  and  $47.5 \pm 0.8$  by HM-SSM and LM-SSM, respectively [80, 81]. These values are used in the analysis based on Phase-II data of the Borexino experiment (refer to Sec. 3.1), in order to determine an upper limit on the CNO neutrino rate, which will be discussed in Sec. 8.2.

### 2.4.3 Probing solar metallicity

The difference in the metallicity assumed by HM- and LM-SSMs results in different solar neutrino flux predictions, as shown in Table 2.3. As can be seen, among the  $pp$ -chain neutrinos, the  ${}^7\text{Be}$  and  ${}^8\text{B}$  neutrino fluxes are the most sensitive to metallicity. The measurement of their fluxes can be used to probe this parameter.

Solar $\nu$	B16-GS98 (HM)	B16-AGSS09met (LM)	Difference (%)
$pp$	$5.98 (1 \pm 0.006) \cdot 10^{10}$	$6.03 (1 \pm 0.005) \cdot 10^{10}$	$-0.84$
$pep$	$1.44 (1 \pm 0.01) \cdot 10^8$	$1.46 (1 \pm 0.009) \cdot 10^8$	$-1.39$
$hep$	$7.98 (1 \pm 0.30) \cdot 10^3$	$8.25 (1 \pm 0.30) \cdot 10^3$	$-3.38$
${}^7\text{Be}$	$4.93 (1 \pm 0.06) \cdot 10^9$	$4.50 (1 \pm 0.06) \cdot 10^9$	$+8.72$
${}^8\text{B}$	$5.46 (1 \pm 0.12) \cdot 10^6$	$4.50 (1 \pm 0.12) \cdot 10^6$	$+17.58$
CNO	$4.88 (1 \pm 0.11) \cdot 10^8$	$3.51 (1 \pm 0.10) \cdot 10^8$	$+28.07$

TABLE 2.3: Solar neutrino fluxes, predicted by HM- and LM-SSMs, in units of  $\text{cm}^{-2}\text{s}^{-1}$  [67].

The SSM predictions of  ${}^7\text{Be}$  and  ${}^8\text{B}$  neutrino flux values with allowed  $1\sigma$  regions are shown in Fig. 2.5. The fluxes are normalized to the values based on a global analysis of the data from solar neutrino experiments as of 2016.

As can be seen, experimental measurement of the  ${}^7\text{Be}$  and  ${}^8\text{B}$  neutrino fluxes provides hints towards metallicity. Such analysis is performed in Sec. 7.3.2, including the more recent low-energy solar neutrino results from Borexino. So far, the  ${}^7\text{Be}$  and  ${}^8\text{B}$  neutrino data provides only a weak hint towards HM-SSM. A stronger result that could settle the controversy would come from including the measurement of solar CNO neutrinos, since they are the most sensitive to metallicity, as shown in Table 2.3.

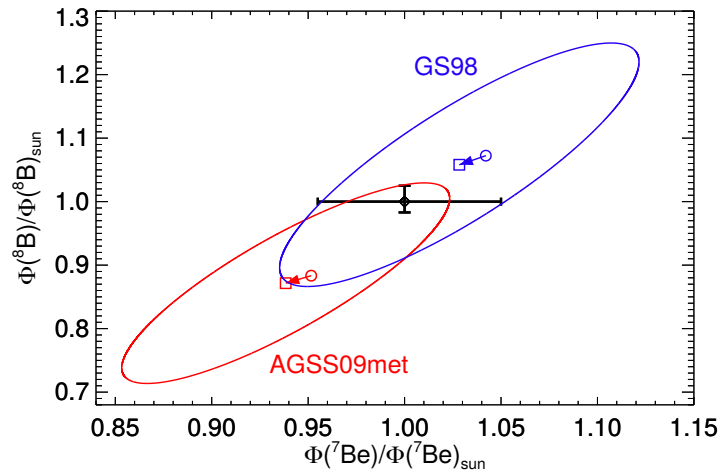


FIGURE 2.5: Fluxes of  ${}^7\text{Be}$  and  ${}^8\text{B}$  neutrinos, predicted by LM- (red) and HM-SSMs (blue), normalized to the measured values based on solar neutrino experiments as of 2016 [79], corresponding to the black circle with error bars. The squares (circles) denote results obtained using the new (old) generation of SSMs, called B16 (SF-II). The ellipses denote theoretical  $1\sigma$  C.L. regions [67].

### 2.4.4 Studying the $pp$ chain terminations

The relative intensity of the  $pp$ -I and  $pp$ -II terminations of the  $pp$  chain (refer to Fig. 2.1A), denoted as  $R_{\text{II/I}}$ , can be calculated on a theoretical basis, and tested by experimental measurements. In addition to nuclear physics predictions, this ratio is influenced by the solar metallicity assumed in the SSMs, as it determines the predicted opacity of the solar plasma, and, in turn, impacts the central temperature of the Sun, regulating the branching ratios of  $pp$ -I and  $pp$ -II. The values predicted by HM- and LM-SSMs are  $R_{\text{II/I}} = 0.180 \pm 0.011$  and  $R_{\text{II/I}} = 0.161 \pm 0.010$ , respectively. Experimentally, this value can be obtained based on the measured fluxes of  $pp$  and  ${}^7\text{Be}$  neutrinos:

$$R_{\text{II/I}} = \frac{2\phi({}^7\text{Be})}{\phi(pp) - \phi({}^7\text{Be})}, \quad (2.6)$$

neglecting the contributions from  $pep$ ,  $hep$  and  ${}^8\text{B}$  neutrinos [82]. Measurement of the fluxes of  $pp$  and  ${}^7\text{Be}$  neutrinos from the Borexino experiment can be used to probe these predictions, as will be shown in Sec. 7.3.3.

### 2.4.5 Probing flavor transformation in matter

As described in Sec. 1.4.2, the MSW effect, occurring due to the influence of the medium on neutrino propagation, effectively modifies the values of the mixing angles  $\theta_{ij}^m$  that determine the elements of the mixing matrix  $U^m$ . Based on the MSW-LMA solution, the survival probability can be approximated as [82]:

$$P_{ee} \approx \frac{1}{2} + \frac{1}{2} \cos 2\theta_{12}^m \cos 2\theta_{12}. \quad (2.7)$$

The mixing angle in matter  $\theta_{12}^m$  is derived as follows:

$$\cos 2\theta_{12}^m = \frac{\cos 2\theta_{12} - \beta}{(\cos 2\theta_{12} - \beta)^2 + \sin^2 2\theta_{12}}. \quad (2.8)$$

The parameter  $\beta$  stands for:

$$\beta = \frac{2\sqrt{2}G_F n_e E_\nu}{\Delta m_{12}^2}, \quad (2.9)$$

where  $G_F$  is the Fermi coupling constant,  $n_e$  is the electron density at the location where the solar neutrino was produced, and  $E_\nu$  is its energy. For small values of  $\beta$ , corresponding to low  $E_\nu$ , namely,  $\beta < \cos 2\theta_{12} \approx 0.4$ ,  $P_{ee}$  is dominated by vacuum oscillation; while for values  $\beta > 1$ , MSW oscillation prevail:

$$P_{ee} = \begin{cases} 1 - \frac{1}{2} \sin^2 2\theta_{12}, & \beta < 0.4 \text{ (vacuum)}, \\ \sin^2 \theta_{12}, & \beta > 1 \text{ (MSW)}. \end{cases} \quad (2.10)$$

The resulting dependence on the solar neutrino energy is shown in Fig. 2.6. The  $E_\nu$  region corresponding to  $\beta \approx 1$  results in the so-called *resonance condition* [58].

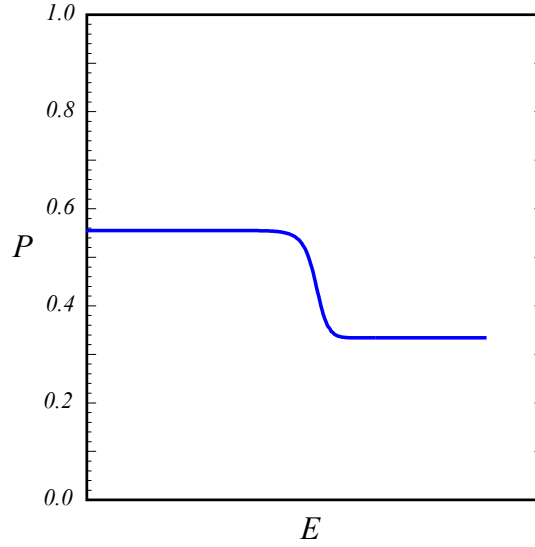


FIGURE 2.6: Schematic trend of the  $\nu_e$  survival probability  $P$ , based on Eq. 2.10, as a function of neutrino energy  $E$  for the MSW-LMA solution [82].

The experimental measurements of solar neutrino fluxes can be used to probe the MSW-LMA solution. One can calculate  $P_{ee}$  based on the measurement of the solar neutrino fluxes, and the initial fluxes predicted by SSMs, using Eq. 1.34. As solar  $pp$ ,  ${}^7\text{Be}$ ,  $pep$ , and  ${}^8\text{B}$  neutrinos cover the energy range from  $\sim 0.1$  to  $10\text{ MeV}$ , this provides values of  $P_{ee}$  both below and above the resonance point, shown in Fig. 2.6, which allows us to probe MSW predictions compared to the ones based on vacuum-only oscillation. In particular, the value of  $P_{ee}$  based on the  ${}^8\text{B}$  neutrino measurement falls in the neutrino energy range where the predictions based on vacuum or MSW oscillation differ the most. The results of this approach based on the Borexino solar neutrino results will be shown Sec. 7.3.4.

In addition, such measurements can be used to probe for deviations from the MSW-LMA theory due to *nonstandard neutrino interactions*, which influence neutrino production, propagation and detection [83].





## Chapter 3

# The Borexino experiment

In this Chapter, I will introduce the Borexino experiment and its scientific program, the Borexino detector, and the methods of processing the data it measures.

I will start with a general introduction to the experiment in Sec. 3.1, and in Sec. 3.2 continue with the description of the structure of the detector and its various parts, and important detector operations. In particular, in Sec. 3.2.3, I will mention the Borexino photomultiplier tubes (PMTs), the information about which will be relevant in Sec. 5.3, where I discuss PMT quality; and in Chapter 6, where I describe my work aimed at obtaining the effective quantum efficiency of the Borexino PMTs. In Sec. 3.2.6, I will describe the Borexino 2010-2011 purification campaign, which characterizes the Borexino Phase-II analysis that I will present in Chapter 7; and the 2015 thermal insulation campaign (Sec. 3.2.7), crucial for the Phase-III CNO neutrino analysis which I will focus on in Chapter 8.

After this, I will continue with an overview of the essential electronics systems that take part in the detection and data acquisition (Sec. 3.3). The knowledge about the main Borexino data acquisition system (Sec. 3.3.1) will be important in Sec. 5.1, where I will talk about my work related to the Borexino data quality. In addition, I will introduce the so-called FADC data acquisition system (Sec. 3.3.4), complementary to the main one, the information from which I have used in my contribution to the *hep* neutrino analysis in Sec. 7.2.3.

I will then proceed to talk about the Borexino data and the way it is treated (Sec. 3.4), as it is important to understand how the raw information is processed to obtain physical parameters describing the events, such as energy, position etc. In particular, I will mention the so-called laser timing calibration in Sec. 3.4.1, as my work in Sec. 5.2 consists in improving and monitoring the stability of this procedure.

Finally, in Sec. 3.5 I will describe the Borexino Monte Carlo (MC) simulation, an important parameter of which, the effective quantum efficiency of the PMTs mentioned above, has been comprehensively studied by me, and improved through my work that I will present later in Chapter 6.

### 3.1 Introduction

Borexino is a liquid scintillator detector located in Italy in the Hall C of *Laboratori Nazionali del Gran Sasso*, the largest underground laboratory in the world, under about 1400 m of rock with shielding capacity of 3800 m water equivalent. At this laboratory, the muon flux is suppressed by a factor of  $10^6$  with respect to the surface, which greatly reduces muon and muon induced backgrounds (more in Sec. 4.5).

The liquid scintillator technique that Borexino employs is the best for measuring solar neutrinos, since it allows for a low energy threshold due to high light yield, unlike, for example, Cherenkov experiments; and real-time measurement, as opposed to radiochemical experiments. Because of this, Borexino is the only detector able to perform spectroscopic detection of solar neutrinos of energies as low as 150 keV, with good energy resolution. Furthermore, the organic nature of liquid scintillators implies low solubility of ions and metal impurities, which allows for thorough purification of the material.

The Borexino experiment started data taking in May 2007, and has been running continuously since then. The original goal of the experiment was to measure the  ${}^7\text{Be}$  solar neutrinos (more on solar neutrinos in Sec. 2.1), the first results on which appeared already after two months of data taking between May and July 2007 [84]. Since then, however, Borexino exceeded expectations and expanded its scientific scope.

The first three years of data taking between May 2007 and May 2010 constitute the Borexino **Phase-I**. It resulted in the first precision measurement of the  ${}^7\text{Be}$  neutrino flux [85] and its day-night asymmetry [86], the first evidence of *pep* neutrinos [87], and the measurement of  ${}^8\text{B}$  neutrinos with the lowest threshold of 3 MeV [88].

End of Phase-I was followed by an extensive purification campaign, which I will talk about in Sec. 3.2.6. This resulted in an unprecedented level of radiopurity which characterizes **Phase-II** of data taking. This phase lasted from December 2011 to May 2016, and resulted in the first direct observation of *pp* neutrinos [89], and a comprehensive measurement of all *pp*-chain solar neutrinos that also included the first discovery of *pep* neutrinos [72].

July 2017 marks the beginning of **Phase-III**, characterized by thermal stability after the thermal insulation campaign (more in Sec. 3.2.7) and the goal of detecting CNO neutrinos.

The Phases and important detector operations are depicted visually in Fig. 3.1.

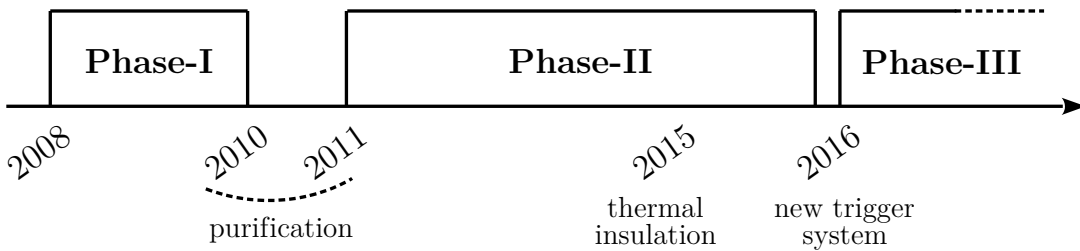


FIGURE 3.1: A diagram representing the Phases of Borexino as well as other important operations.

Aside from solar neutrinos, the scientific program of the Borexino experiment includes geoneutrinos (antineutrinos coming from the Earth) [34–37, 90], neutrino magnetic moment [91], seasonal variations of the muon flux [92], and nonstandard interactions [93].

## 3.2 The Borexino detector

### 3.2.1 Detector structure

The structure of the Borexino detector is based on the principle of *graded shielding*. The core of the detector is surrounded by concentric layers of material of increasing radiopurity [94]. This design is aimed at shielding the liquid scintillator from the external background coming from the surrounding rock. A schematic representation and a three-dimensional rendering of the detector are presented in Fig. 3.2.

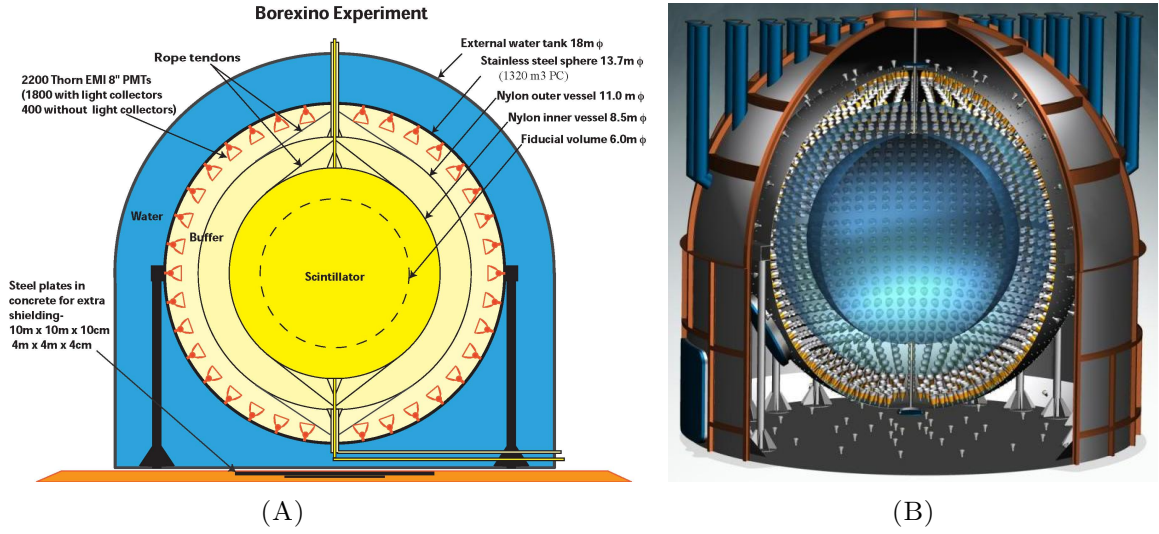


FIGURE 3.2: A schematic 2D representation (A) and a 3D rendering (B) of the Borexino detector, depicting various layers of its internal structure [72, 95].

Borexino can be divided into the Inner and the Outer Detector.

The **Inner Detector** (ID) is equipped with 2212 8-inch inward-facing photomultiplier tubes (PMTs), mounted on the **Stainless Steel Sphere** (SSS) of radius 6.85 m, which collect scintillation light and convert it to current signals. The Borexino PMTs will be discussed more in detail in Sec. 3.2.3. Inside the SSS are two transparent **nylon vessels** (NV): Outer Vessel (OV) and Inner Vessel (IV) of thickness  $125\ \mu\text{m}$  and radii 5.5 m and 4.25 m, respectively. The function of the OV is to stop the  $^{222}\text{Rn}$  contamination from the SSS and the PMTs to enter the center of the ID.

The IV is filled with 278 tons of liquid scintillator (LS) solution: *pseudocumene*,  $\text{C}_6\text{H}_3(\text{CH}_3)_3$ , as a solvent, and *fluor PPO*,  $\text{C}_{15}\text{H}_{11}\text{NO}$ , as a solute. The resulting effective light yield is around 500 p.e. detected per MeV of deposited energy, with energy resolution of 50 keV at 1 MeV (5%). Another function of the NVs is to form a nonscintillating **buffer**. The first shell (between the SSS and the OV), and the second shell (between the OV and the IV) are filled with LS diluted by *dimethylphthalate* (DMP),  $\text{C}_6\text{H}_4(\text{COOCH}_3)_2$ , with density similar to that of the LS in the IV. This is done to suppress scintillation due to external backgrounds and muons, and avoid oversaturation of the PMTs and the acquisition system.

The **Outer Detector** (OD) is equipped with 208 outward-facing PMTs, mounted on the SSS and the floor; and a **water tank** containing 2.4 kt of ultrapure water. The water tank shields the detector from external  $\gamma$ 's and neutrons, and serves as an active muon veto, allowing for detection of muon-induced Cherenkov light by the OD PMTs. More on muon detection will be described in Sec. 4.5. On the outside, the water tank is wrapped in a layer of thermal insulation (since 2015), which will be discussed in Sec. 3.2.7.

### 3.2.2 Particle detection principles

Particles are detected in Borexino by exploiting the phenomenon of *scintillation*. After a particle interacts in the detector, the LS molecules are excited directly by that particle, or indirectly through a secondary particle, after which they emit light during deexcitation. This light is then detected by the PMTs, which convert the detected photons to photoelectrons, i.e. electric current. The emission of light is isotropic, and information about the initial direction of particles is lost. This makes neutrinos indistinguishable from other particles on event-by-event basis. Depending on the analysis, they have to be evaluated through different methods (more in Sec. 4.4), removed during data selection (Sec. 7.1.1), or included in data fitting procedure (Sec. 7.1.2). In particular, to suppress  $\beta$  and  $\gamma$  backgrounds, high radiopurity and good data selection are needed. Different types of particles interact in the LS through different processes, which I will shortly describe below, and some of which can be exploited to identify them.

#### Neutrinos

In the Borexino detector, neutrinos are detected through elastic scattering off electrons in the LS:

$$\nu_{e,\mu,\tau} + e^- \rightarrow \nu_{e,\mu,\tau} + e^-. \quad (3.1)$$

This detection method has been described in detail in Sec. 1.5. The scattered electron carries a fraction of the neutrino energy, and then transfers it to the scintillator molecules. Due to this, the spectrum of the recoiled electrons is always continuous, even if the neutrino was monoenergetic. Note that in all the Borexino analysis, data is analyzed in terms of the energy of these recoiled electrons, not the neutrino energy.

#### Electrons

Some internal backgrounds, such as  $^{14}\text{C}$ ,  $^{210}\text{Bi}$  and  $^{85}\text{Kr}$ , are  $\beta$  emitters. The  $e^-$  produced in their decay excite the LS molecules in exactly the same fashion as described above for electrons recoiled after neutrino scattering. This makes  $e^-$  indistinguishable from neutrino signals, making  $\beta$  emitters the main background in the solar neutrino analysis of Chapter 7.

#### Gammas

The interaction of MeV  $\gamma$ 's, before they lose energy and get absorbed by the LS molecules, is dominated by *Compton scattering* with electrons, which in turn interact with the LS as described above.

#### Positrons

Positrons interact with LS molecules until they lose kinetic energy and, shortly after being emitted, annihilate with electrons present in the LS, since annihilation cross section decreases

with energy. The annihilation produces two 511 keV  $\gamma$ 's which in turn interact in the LS as described above. Another possibility is the formation of *orthopositronium*, an unstable compound of an electron and a positron [96]. The annihilation signal in such case has a characteristic delay in comparison with electrons, and this is exploited for  $e^-/e^+$  discrimination which I will describe later in Sec. 3.4.5. The only source of positrons in Borexino is the cosmogenic  $^{11}\text{C}$ , which is important for the measurement of solar neutrinos, and will be examined in Sec. 4.5.3.

### Muons

Muons spallate on LS molecules directly, exciting multiple molecules as they pass through the detector and producing large amounts of light, as well as Cherenkov radiation. Apart from being registered by the ID through scintillation light, muons are also detected in the OD through Cherenkov light produced as they travel through the water tank. The combination of both ID and OD results allows us to determine whether the particle is a muon, and whether it passed through only OD or both OD and ID. Due to this, muons are determined with high accuracy and do not constitute a big problem; however, muon spallation produces multiple secondary particles which constitute the so-called cosmogenic background which will be discussed in Sec. 4.5, as well as more details on muon detection in Borexino.

### Neutrons

Neutrons are detected via capture on protons, which happens within 250  $\mu\text{s}$  [97]:

$$n + p \rightarrow D + \gamma, \quad (3.2)$$

through the resulting  $\gamma$  of 2.2 MeV. With a much smaller probability, capture can happen on  $^{12}\text{C}$  nuclei.

### Alpha particles

Alpha particles are strongly affected by *quenching*, an effect that happens due to energy deposition density. Since  $\alpha$ 's are heavy, the LS molecules that they excite are distributed more densely, which results in energy loss before scintillation light is emitted. In Borexino, the dominant source of  $\alpha$ 's is  $^{210}\text{Po}$ , which will be discussed in Sec. 4.4.4. In principle,  $\alpha$ 's emitted by  $^{210}\text{Po}$  are not in the energy region of solar neutrinos; however, due to factor  $\sim 10$  quenching, their reconstructed energy falls in the sub-MeV part of the spectrum. The effect of quenching brings the energy of the emitted  $\alpha$ 's down to the energy region of solar neutrinos, which otherwise would not be expected to appear.

## 3.2.3 The Borexino photomultiplier tubes

The Borexino PMTs are 8-inch ETL-9351 PMTs, the role of which is to collect scintillation light and convert photons to photoelectrons (p.e.) through photoelectric effect, producing electric current signal. The number of p.e. collected by the PMTs is approximately 500 per MeV. In the energy region of solar neutrinos, Borexino PMTs work mostly in single p.e. regime. The PMTs have been designed to accommodate to the conditions of the Borexino detector, namely, to withstand the corrosion induced by the LS they are submerged in, and to meet the radiopurity requirements of the experiment (low radioactivity PMT glass).



All but 371 PMTs are equipped with cone-shaped aluminum *concentrators* (seen in Fig. 3.3A) to avoid registering photons coming from the outside of the active scintillator volume and increase light collection efficiency. This is an important factor that will be considered in Sec. 6.3.2, where the effect of the concentrators must be disentangled from that of the quantum efficiency of the PMTs, since the concentrators are accounted for in the Borexino Monte Carlo simulation (Sec. 3.5). Figure 3.3B demonstrates PMTs mounted on the SSS, illustrating both kinds of PMTs, with and without concentrators.

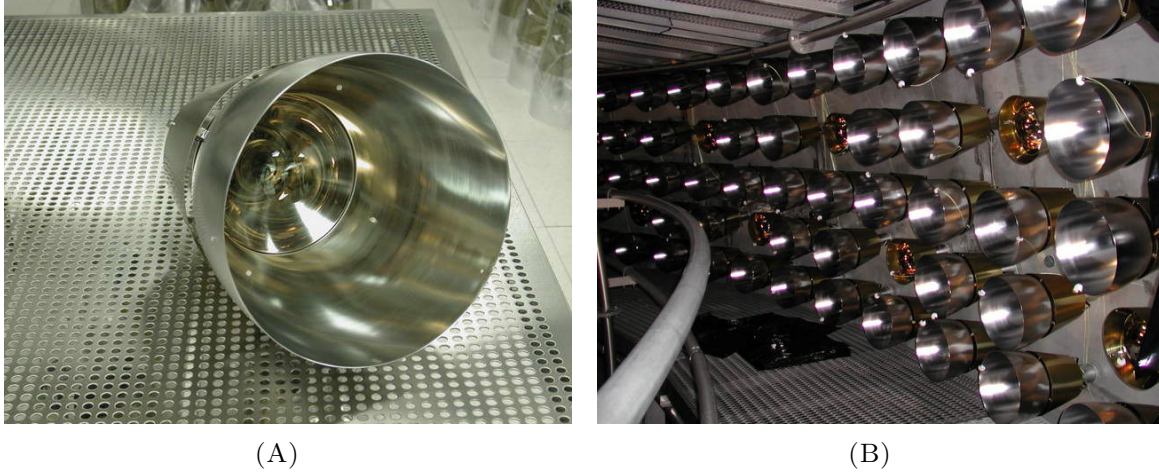


FIGURE 3.3: (A) A sealed Borexino PMT fully assembled with a light concentrator and mu-metal magnetic shield [98]. (B) PMTs during installation on the Borexino SSS [98].

The signals sent by PMTs are analyzed by different electronic circuits, which will be described in Sec. 3.3.1, to extract information about charge and photon arrival time. In order to correctly analyze the hit times on PMTs, the so-called *laser timing calibration* is performed, aimed at aligning the response of different PMTs. I will describe the system designed to perform this calibration in Sec. 3.2.4 below, and elaborate on the alignment procedure in Sec. 3.4.1.

Another crucial aspect of the PMTs is their quality, which can be measured by different parameters, such as quantum efficiency or dark noise rate. PMT quality directly influences their light collection properties, and ultimately, the energy estimators used to analyze the data (more in Sec. 3.4.3). I will discuss PMT quality in more detail in Sec. 5.3, where I will present my studies related to this topic.

### 3.2.4 The laser calibration system

As mentioned in Sec. 3.2.3 above, in order to calibrate the PMTs, a *laser calibration system* exists in Borexino. A schematic representation of the system is shown in Fig. 3.4A. A diode laser emits a light pulse, which is first distributed among 35 fibers that reach optical feedthroughs on the SSS (marked blue), and then further split into 90 fibers that couple to each PMT (marked orange). Figure 3.4B shows a PMT installed in the detector, depicting an optical fiber attached to the concentrator.

The laser calibration system is used for several purposes. It is used at the beginning of each data acquisition run for coarse time alignment on 50-ns scale, as well as throughout the run,

which I will mention in Sec. 3.3.1. In addition, it is used for weekly timing calibration on a scale of several ns, aimed at the fine time alignment of PMTs, which will be described in Sec. 3.4.1.

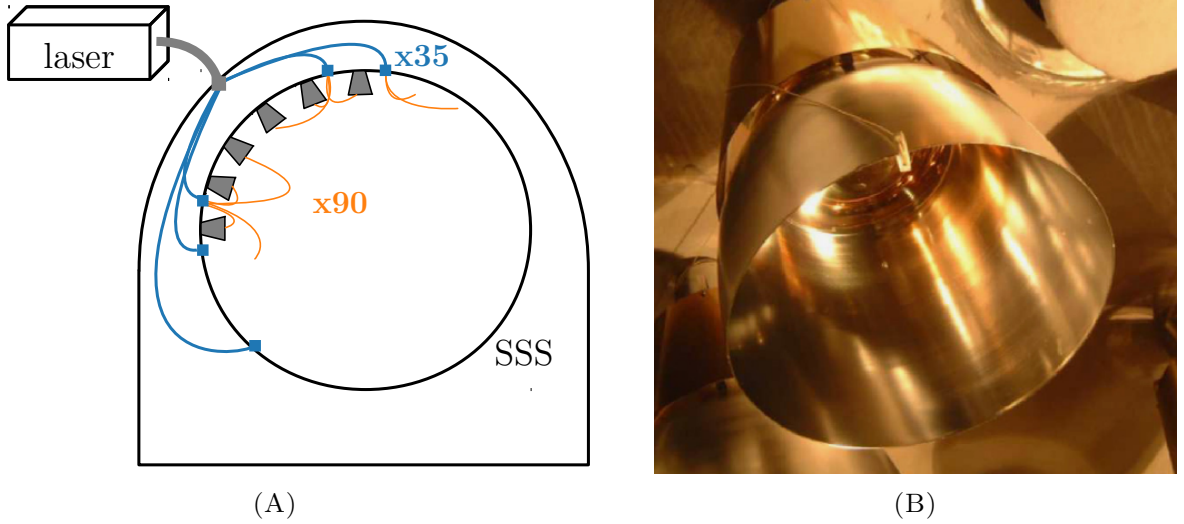


FIGURE 3.4: (A) A schematic representation of the laser calibration system. (B) Optical fiber connected to an installed PMT [94].

### 3.2.5 Source calibrations

In order to achieve high precision in the determination of neutrino rates, the detector design includes an internal and an external source calibration system [99, 100]. Several calibration campaigns took place between 2008 and 2011: the campaigns with internal *radioactive sources* inserted in the scintillator in October 2008, and January, June and July 2009; while that with an external  $\gamma$  source located in a special niche in the outer buffer region was performed in July 2010 and December 2011. The positions of various sources are shown in Fig. 3.5A.

The calibrations provided a detailed understanding of the detector response and led to a significant reduction of the systematic uncertainties in the Borexino measurements. In particular,  $^{222}\text{Rn} + ^{14}\text{C}$  sources were used for position reconstruction.

The calibration runs are used to tune the MC simulation, which will be described in Sec. 3.5. In particular, the custom made  $^{228}\text{Th}$  source was used to study the energy and radial distribution of the  $\gamma$ 's coming from the decays of  $^{208}\text{Tl}$ , its daughter.

### 3.2.6 Purification campaign

The Borexino LS already had extremely low radioactivity levels in 2002 [103]. In addition to that, a purification campaign was performed, leading to a further reduction of the remaining radioactive contaminants in Phase-II. It consisted of several cycles of water extraction during one year between the summer of 2010 and 2011. As a result, Borexino is the most radiopure LS detector in the world.

One of the main improvements of the purification was to greatly reduce  $^{210}\text{Bi}$  and  $^{85}\text{Kr}$  contamination. The rate of  $^{210}\text{Bi}$  has been reduced by a factor of 3.5 to the level of  $\leq 20$  cpd/100 t;

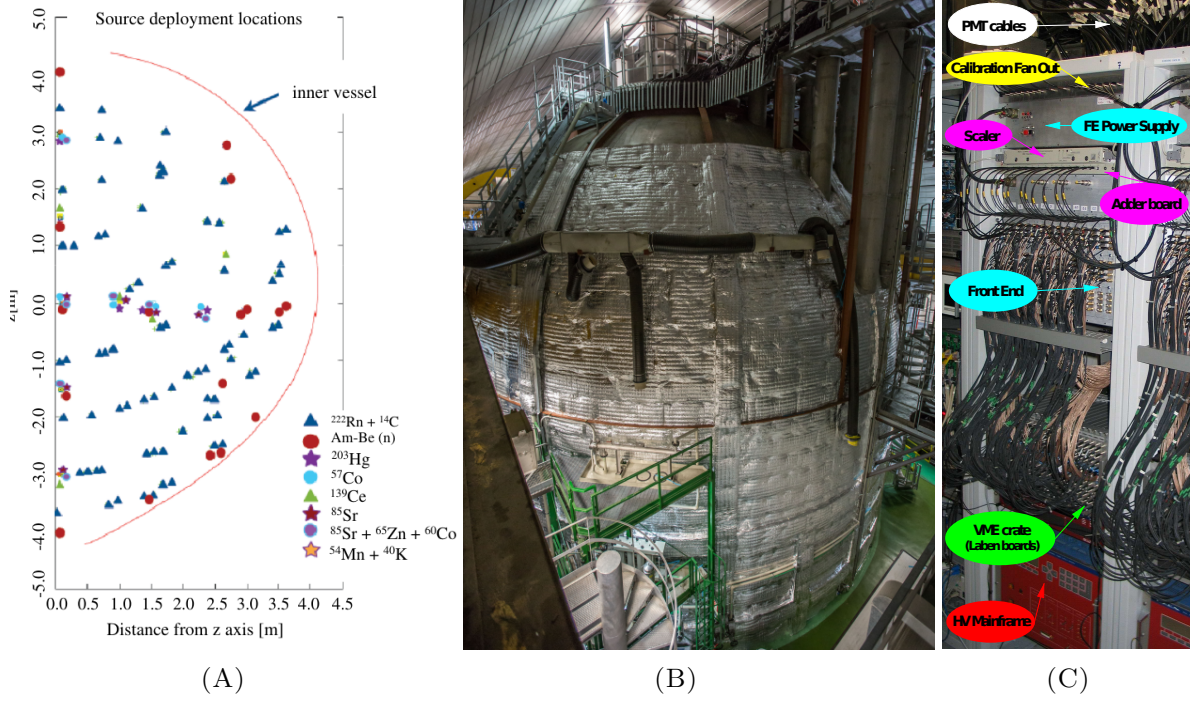


FIGURE 3.5: (A) Position of the various radioactive sources deployed in the scintillator and used to calibrate the Borexino detector [95]. (B) The detector after the thermal insulation (see text) [101]. (C) A Laben electronics rack, containing (bottom to top): the high voltage board, digital electronics crate, front-end electronics, analog adder, and low voltage power supplies (see text) [95, 102].

while  $^{85}\text{Kr}$  decreased from 20-30 cpd/100 t in Phase-I to impressive  $\leq 4.7$  cpd/100 t at 95% C.L. The levels of natural long-lived radioactive contaminants were reduced to negligible levels, including  $< 9.5 \cdot 10^{-20}$  g/g (95% C.L.) for  $^{238}\text{U}$ , and  $< 7.2 \cdot 10^{-19}$  g/g (95% C.L.) for  $^{232}\text{Th}$ , compared to the levels being of order 10 Bq/kg in nature (in water). More on Borexino radiopurity will be discussed in Sec. 4.2.

### 3.2.7 Thermal insulation

Assuring stable temperature inside the detector is important for reducing the contamination of the inner scintillator volume by radioactive materials present on the surface of the IV, for example,  $^{210}\text{Po}$ . In order to reduce convection movements that can wash off the contaminants, temperature gradient is maintained between the top and the bottom of the detector. The bottom is in contact with a *heat sink* provided by the rock (7-8 °C), and the top is in contact with the air of Hall C (around 15 °C). However, with this setup, the top temperature is still highly influenced by air currents in Hall C. To avoid this effect, a *thermal insulation* campaign was conducted in the summer of 2015. During the campaign, the detector was covered with two layers of 10-cm-thick mineral wool, which can be seen in Fig. 3.5B. The water tank and the insulation layers were equipped with temperature sensors to monitor the changes that propagate into the system from the outside. Thermal insulation was crucial for studies of  $^{210}\text{Po}$ , which is important for CNO neutrino detection, which will be described in Sec. 8.3.1.



### 3.3 Data acquisition

#### 3.3.1 The electronics layout

The Borexino PMTs are connected to several electronic circuits that read and process their signals [104]. The complete system is composed of the following stages:

- the **front-end** stage, which performs an analog processing of the PMT signals, including noise filtering, preamplification, shaping, and integration;
- the **readout** stage, used to digitize and store the signal processed by the front end and to measure the arrival time and the charge of each light pulse coming from the detector;
- the **trigger** stage, used to identify which events are worth considering.

These stages are shown schematically in Fig. 3.6. The ID and OD PMTs have independent electronics systems both of which are connected to the trigger system [105]. The readout electronics of the ID PMTs is called **Laben** and will be described below, as well as the additional **FADC** system which also processes ID information.

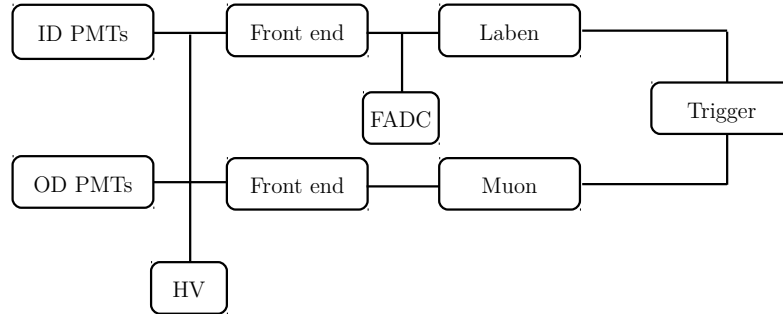


FIGURE 3.6: A simplified schematic representation of the Borexino electronics layout.

#### Laben

The digital electronics of the ID called Laben is contained in 14 racks, each managing 160 front-end channels. One electronic rack contains 20 boards plugged into a VME crate, each board housing electronics for 8 front-end channels (see Fig. 3.5C). Each channel has two inputs coming from the front end stage: a fast signal containing time information, and an integrated signal containing energy information. Each VME crate has an interface to the trigger system, which I will elaborate on in Sec. 3.3.2. The channels are also connected to a *pulser system*, which is used to monitor them during data acquisition (DAQ), the procedure of which will be described later in Sec. 3.3.3. Some channels are assigned special status for monitoring purposes of the *service triggers*, which will be described in Sec. 3.3.2. The so-called *laser reference* and *trigger reference* channels are used for reference time for signals sent by the laser calibration system, described in Sec. 3.2.4, and the pulser system mentioned above. The charge information from these channels is used in a part of the DAQ validation procedure, as shown in p. 156, which will be discussed in detail in Sec. 5.1.

## FADC

The main acquisition system of Borexino has been designed and optimized for the detection of low energy solar neutrinos in the sub-MeV range. In addition, a separate fast waveform digitizer system called FADC (flash-ADC) is explicitly dedicated to the higher energy range [94].

This system is based on the idea that at high energies it is possible to retain precision while not having to record each of the 2212 channels individually. The PMT signals are therefore grouped by solid angle sectors, thus reducing the number of acquisition channels. This way, the dynamic range of the system is extended to up to 30 MeV, as each front-end board provides an analog sum of 12 linear output signals.

### 3.3.2 The trigger system

As scintillation is recorded by PMTs, the information is stored temporarily and kept if the event meets certain *trigger conditions*, otherwise it is lost, and the memory is overwritten. A PMT is considered “hit” if it detected at least one photon. When the total number of PMTs hit within 90 ns exceeds the programmed threshold, the trigger sequence begins: the trigger signals are issued, the absolute time is read from the GPS clock, and a trigger record is written in the VME readable memory.

All the data in the  $16\text{-}\mu\text{s}$  window before the trigger time, called the **DAQ gate**, is recorded, followed by a  $4\text{-}\mu\text{s}$  dead time, after which DAQ is ready to release a new trigger, which is depicted schematically in Fig. 3.7. This way, each event is labeled with its unique 16-bit event number. The trigger threshold for the data taken until 2013 is equal to 25 PMTs. After that, it was lowered to 20 PMTs due to PMT loss.

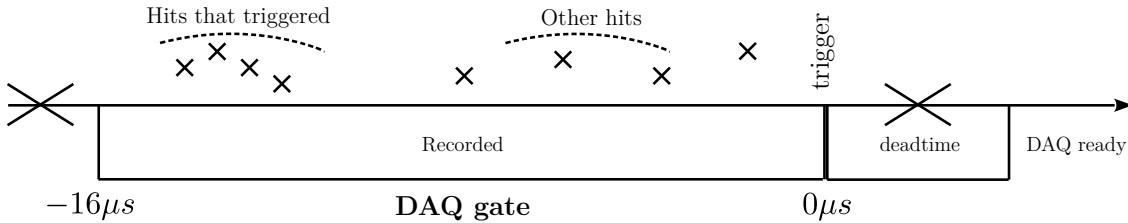


FIGURE 3.7: A schematic representation of the DAQ gate and the trigger procedure.

Depending on the combination of ID and OD information, as well as other factors that will be described below, a **trigger type** is assigned to each event. In addition, a bit in the “trigger word” of the Borexino Trigger Board (BTB) can be raised to indicate that the OD PMTs triggered during the event. Table 3.1 lists different trigger types in Borexino, combined with BTB information.

The trigger types assigned to physical events are **neutrino** (TT1 BTB0), **internal muon** (TT1 BTB4), and **external muon** (TT2). These triggers are assigned based on whether the ID and the OD PMTs have triggered, as visually explained in Table 3.2. The OD is considered to be triggered if 6 OD PMTs are hit within 150 ns. After each event marked as internal muon, a **neutron trigger** (TT128) is issued automatically in order to record neutrons produced by muon spallation, opening a longer 1.6-ms acquisition gate (five times neutron capture time).

Priority	TT	BTB input	Name
1	1	0	Neutrino
	1	4	Internal muon
2	2	4	External muon
3	128	8	Neutron
6	32	64	Pulser
7	8	64	Laser
8	64	64	Random

TABLE 3.1: Trigger types used in Borexino [102].

	ID ✓	ID ✗
OD ✓	TT1 BTB4	TT2
OD ✗	TT1 BTB0	-

TABLE 3.2: Trigger types of internal and external events (✓= triggered, ✗= did not trigger).

In addition to these physical triggers, there are three so-called *service triggers* implemented in the trigger system. The **laser trigger** (TT8) marks the events coming from the signal in the PMTs sent by the laser calibration system, described in Sec. 3.2.4. The **pulser trigger** (TT32) is issued to record digital pulses sent to the Laben channels for the purpose of channel monitoring, which I will elaborate on in Sec. 3.3.3. The **random trigger** (TT64) is issued with a frequency of 0.5 Hz, opening a DAQ gate unprompted, and recording all hits that would fall into the gate. The main purpose of this trigger type is to study the *dark noise* rate, as will be mentioned in Sec. 3.4.4.

Note that the “event” corresponding to one trigger may contain more than one *physical* event. In order to determine how many different physical events contribute to one “trigger event”, a *clustering algorithm* is used to separate the hits in its DAQ gate into groups (clusters) that represent scintillation events. Types of events depending on their trigger type and clustering, as well as their physical meaning, will be discussed in Sec. 3.4.4.

### 3.3.3 Main DAQ procedure

The Borexino data taking is split into 6-hour-long runs which are later grouped on a weekly basis. Before each run, two short procedures are performed to check the status of the PMTs and logical channels. The first check is the so-called **precalibration** aimed at probing the digital Laben channels. During the precalibration, the pulser system is used to send controlled electronic signals to each logical channel (TT32, as mentioned in Sec. 3.3.2). Depending on the response of a channel to the pulser, it may be marked as nonresponsive and will be disabled in this run during data processing which I will elaborate on in Sec. 3.4.2. Next, the so-called **calibration** part of the run is performed, consisting in a short period of laser pulses (TT8) sent by the laser calibration system, described in Sec. 3.2.4, with the purpose of time alignment of the Laben channels.

Laser and pulser events continue to be sent throughout the run with lower frequencies, 0.5 and 0.1 Hz, respectively; as well as the random trigger described in the previous section. The TT1 trigger is dominated by  $^{14}\text{C}$  events (99%) resulting in the total trigger rate of 20-30 Hz.

As mentioned in Sec. 3.3.1, among the Laben channels, there are special types of channels called **trigger** and **laser reference channels** corresponding, receiving information from the pulser and laser system, respectively. The charge and timing information received by these channels is used as a reference for TT32 and TT8 events.

Each run is validated using the so-called *RunValidation* procedure which will be described in detail in Sec. 5.1, using the results of precalibration and calibration, and multiple other parameters.

### 3.3.4 Complementary FADC DAQ procedure

The FADC DAQ procedure follows similar steps as the one of Laben, and has a separate validation procedure based on various parameters, for example, the counting rate shown in Fig. 3.8.

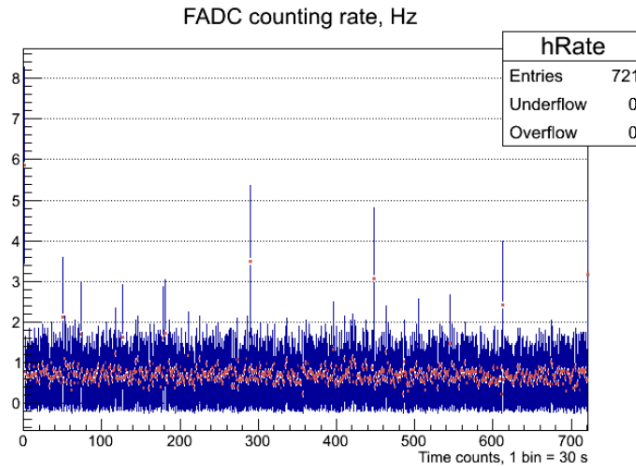


FIGURE 3.8: Overall FADC counting rate in units of Hz as a function of time in one DAQ run.

The FADC system, DAQ and runs are completely independent from those of Laben. However, the data from FADC and Laben events is saved in such a way that it is possible to access events corresponding to each other, and cross-check the information from the two DAQ systems. The main differences in the FADC DAQ compared to the main one lies in that it uses the FADC pulse shape to tag muons, as opposed to ID pulse shape in Laben; the energy is measured in MeV as opposed to p.e.; and the gate length is only  $1\ \mu\text{s}$ , compared to  $16\ \mu\text{s}$  in Laben.

As mentioned before, the FADC system is dedicated to the analysis in higher energy range, and used for different purposes like muon tagging and supernova studies. In particular, I have used the FADC information in my contribution to the *hep* neutrino counting analysis that I will describe in Sec. 7.2.3.

## 3.4 Data processing and data structure

### 3.4.1 Laser timing calibration

*Laser timing calibration* (LC) is a procedure aimed at time alignment of the Laben channels using the laser calibration system described in Sec. 3.2.4. It is performed by a shifter on a weekly basis, consisting of a special three-hour run with a more than usual frequency of laser pulses sent to the PMTs. Figure 3.9A shows the resulting distribution of hits in Laben channels.

As one can see, the hits from the same laser pulse that is sent to all channels at the same time are not measured simultaneously by all channels due to them having different response time. The channels are aligned in order to account for this effect, which is important for the determination of the time of the event, and ultimately for position reconstruction and particle identification based on pulse shape. The alignment is shown in Fig. 3.9B and is done by subtracting the offset of each channel relative to the average of all channels.

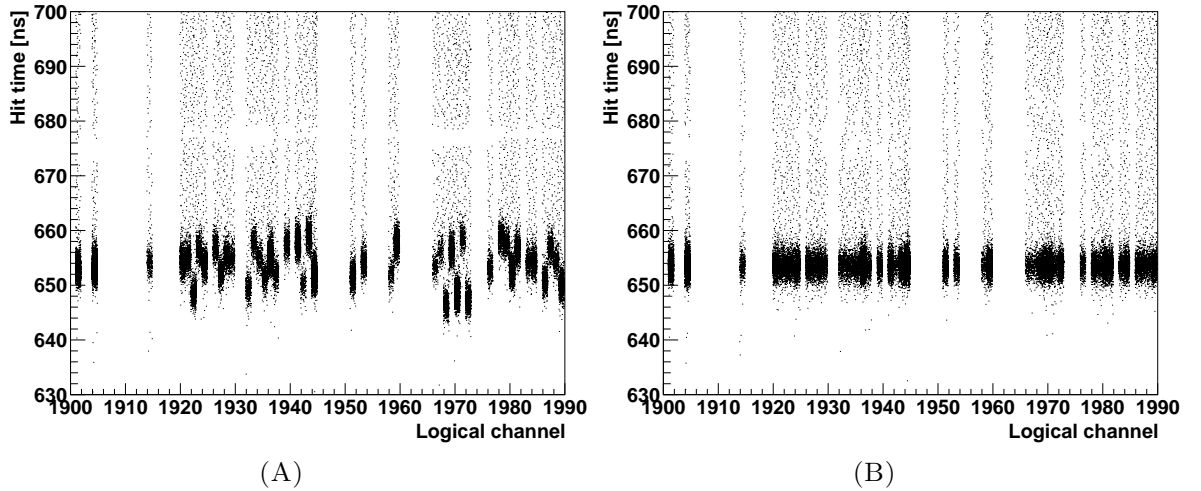


FIGURE 3.9: Distribution of hit times in response to a laser signal in each logical channel before (A) and after (B) alignment, with respect to laser reference time.

In Sec. 5.2.1 I will give more details about this procedure, as well as my work on an updated method for obtaining channel time offset. The channel offset correction resulting from one LC run is applied during the processing of all the DAQ runs during the week following that LC run. Figure 3.10A shows the distribution of hits coming from the laser in a single channel. Apart from the position of this peak which determines the offset of the channel, we are also interested in its width, which represents the sharpness of the channel response, influenced by the PMT transit time jitter. Figure 3.10B demonstrates the total contribution from all channels after alignment (blue), compared to the distribution before alignment (black).

My studies related to monitoring the channel response sharpness, as well as other work related to the quality of the laser timing calibration, will be shown in Sec. 5.2.

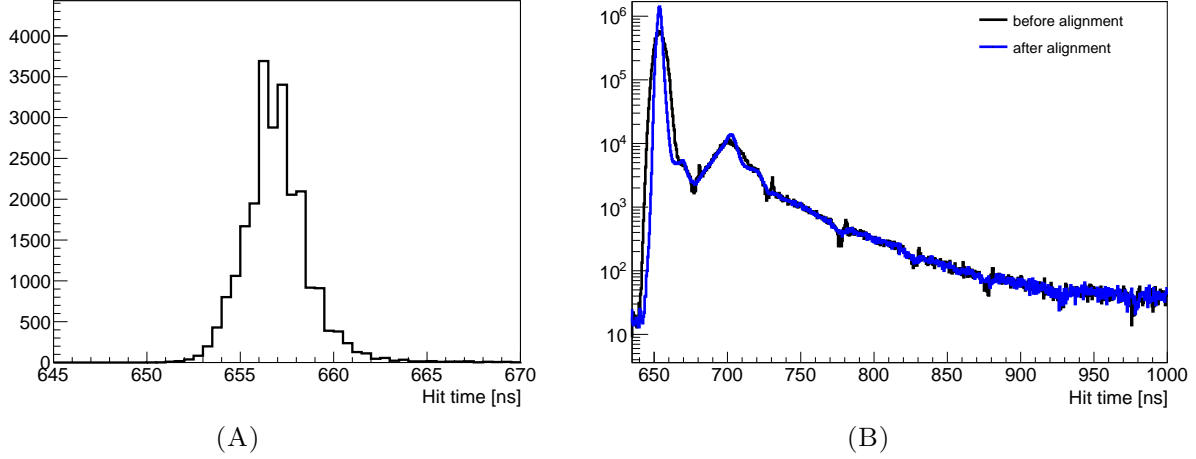


FIGURE 3.10: (A) Distribution of hit times in lg 350 in one LC run. (B) Projections of the histograms from Fig. 3.9.

### 3.4.2 Physical interpretation of raw data

In order to obtain physically meaningful parameters that quantify the recorded events, ROOT based software called **Echidna** has been developed by the Borexino collaboration. This software is used to process every DAQ run, analyzing the digital Laben information. Echidna also accounts for the digital Laben channels that should be ignored based on precalibration or the behavior of PMTs during the run.

During this procedure, a **clustering algorithm** is applied to each event. Its purpose is to disentangle hits coming from scintillation light and dark noise on an event-by-event basis, as well as distinguish separate physical events belonging to the same DAQ gate. A **cluster** represents a scintillation event that happens during a DAQ gate, be it the one that triggered it, or a coincidence, as shown in Fig. 3.7. A cluster of hits is defined as an aggregation of hits correlated closely in time. One event may have zero, one or several clusters. The algorithm goes through the *raw hits* recorded by Laben and removes bad unphysical hits using a decoding algorithm; these *decoded hits* are then grouped into clusters. Typical cluster duration in Borexino is  $1.5 \mu\text{s}$ .

Clusters are the primary objects of the Borexino analysis, as they represent physical scintillation events. Sometimes two clusters happen so close in time that the clustering algorithm cannot disentangle them. This creates a type of background called *pileup* which will be addressed in Sec. 4.6. Event types depending on trigger type and clustering will be summarized in Sec. 3.4.4.

As soon as a cluster is found:

- *energy reconstruction* is applied by Echidna to processes Laben signals and obtain *energy estimators* of the given event (more in Sec. 3.4.3);
- its position is determined using the **position reconstruction** algorithm;
- *particle identification* methods are used to determine the type of event (Sec. 3.4.5).

The position of an event is calculated by maximizing the likelihood of the measured distribution of hit times on PMTs, accounting for PMT positions [106]. The algorithm has been validated with Monte Carlo studies based on  $^{222}\text{Rn} + ^{14}\text{C}$  source calibrations mentioned in Sec. 3.2.5, by comparing the reconstructed positions with the true ones known with precision of 2 cm. The current reconstruction precision is 10 cm, which gives satisfactory accuracy for the solar neutrino analysis; worsening to up to 20 cm for lower energies.

Apart from data, the *Echidna* software is applied to events simulated with Monte Carlo (see Sec. 3.5). This way, both data and simulation are processed the same way and have the same data structure.

The software is updated regularly to improve its existing functions and add new ones. For example, as a result of my study on the quality of PMTs, which I will describe in Sec. 5.3, a new energy estimator has been implemented in *Echidna*, which will be mentioned in Sec. ?? as part of my study on the quantum efficiency of the Borexino PMTs.

### 3.4.3 Energy estimators

In order to estimate the energy of the event, *Echidna* exploits the information from *Laben* channels, namely, the time and *Laben* digital signal of each hit PMT. Depending on the analysis, we use different **energy estimators** for each *cluster*, summarized in Table 3.3 below.

Variable name	Definition
$N_p$	Number of hit PMTs.
$N_h$	Number of hits, including multiple hits on the same PMT.
$N_{pe}$	Charge collected in all PMTs.
$N_p^{dt1}$	Number of PMTs hit within the first 230 ns after the start of the cluster.
$N_p^{dt2}$	Number of PMTs hit within the first 400 ns after the start of the cluster.

TABLE 3.3: Energy estimators used in Borexino. The number of hits or PMTs is considered in terms of hits within a cluster.

For low-energy events single photoelectron regime is observed, which means  $N_p \approx N_h \approx N_{pe}$ . I will use the variable notation in Table 3.3 in further chapters.

Not all the PMTs are active at all times during the data taking. In some cases, the information from some PMTs is temporarily ignored in cases of malfunction. In other cases, a PMT might die permanently due to aging or issues with current and voltage. Due to the changing number of active PMTs, energy estimators need to be **normalized** to account for different number of PMTs contributing to the event:

$$N_x^{norm}(e) = \frac{2000}{N_{live}(e)} N_x(e), \quad (3.3)$$

where  $x = p, h$ , and  $N_{live}(e)$  is the number of live PMTs during the event  $e$ . The normalization for  $N_{pe}$  is similar, but  $N_{live}$  is calculated differently, using not only timing, like for  $N_h$  and  $N_p$ , but also charge information of each *Laben* channel. Normalization for  $N_p^{dt1(2)}$ , as well as normalized energy variables for separate clusters are defined in a similar fashion.

However, if an event happens off-center, PMTs closer to it will register more hits simply due to their proximity, and some less due to them being farther away, which does not represent the energy of the event correctly. In order to compensate for this and bring events happening off-center to equal footing with those happening in the center of the detector seen equally by all PMTs, one needs to account for the solid angle of each PMT contributing to the event, which is the purpose of the **geonormalized variables**, defined for each cluster. This is done by normalizing  $N_x$  by a geometrically weighted number of live PMTs  $N_{live}^{geo}(c)$  for the given cluster  $c$ , rather than the total number of live PMTs, as in Eq. 3.3:

$$N_x^{geo}(c) = \frac{2000}{N_{live}^{geo}(c)} N_x(c), \quad (3.4)$$

where  $x = p$  or  $x = h$ . The geometrically corrected number of live PMTs  $N_{live}^{geo}(c)$  is calculated for a given cluster  $c$  by giving each live PMT  $p$  a weight  $\alpha_p$ , which corresponds to the solid angle relative to the cluster  $c$  of interest:

$$N_{live}^{geo}(c) = \sum_p \alpha_p / 0.00298623, \quad (3.5)$$

where the sum is going over each live PMT  $p$ . The angle  $\alpha_p$  can be obtained the following way:

$$\alpha_p = \frac{\pi R_p^2}{d_p^2} \frac{x_p(x_p - x) + y_p(y_p - y) + z_p(z_p - z)}{d_p D_p}, \quad (3.6)$$

where  $R_p = 0.203$  m is the PMT radius;  $x_p, y_p$  and  $z_p$  are the coordinates of the PMT  $p$ ; and  $x, y$  and  $z$  are the the reconstructed coordinates of the cluster;  $d_p$  is the distance of the PMT  $p$  to the cluster position:

$$d_p = \sqrt{(x_p - x)^2 + (y_p - y)^2 + (z_p - z)^2}, \quad (3.7)$$

and  $D_p$  is the distance of the PMT  $p$  to the center of the detector:

$$D_p = \sqrt{x_p^2 + y_p^2 + z_p^2}. \quad (3.8)$$

Different energy estimators are used in different studies. For example, since the Borexino Monte Carlo (MC) simulation (Sec. 3.5) is better at simulating the hits on PMTs,  $N_h$  is often used in MC-based studies; while  $N_p$  is used to analyze data. The solar neutrino analysis that will be described in Chapter 7 is done using normalized estimator  $N_p^{dt1}$ ; while the work on PMT quantum efficiency presented in Chapter 6 has been done with geonormalized  $N_h$ .

### 3.4.4 Event types

Based on the trigger types and BTB input summarized in Table 3.1, and the number of clusters in the event, I summarize several *event types* from physics point of view in Table 3.4.

These event types are crucial for data selection and analysis. For example, only TT1 BTB0 C1 events, and first clusters of TT1 BTB0 C2 events are considered in the data selection for the solar neutrino analysis (Sec. 7.1.1), and events of type TT1 BTB4 must be considered to apply dead time after each muon event. TT1 BTB0 C2 events are used to study the rate of  $^{14}\text{C}$  events



(Sec. 4.4.1), since effect of the trigger threshold is not present in the second cluster falling into a DAQ gate opened by another event. TT64 C0 events are used to study dark noise rate, as will be shown in Sec. 6.3.1. In the sections to come, I will use the TTX BTBX CX notation to refer to different event types.

Event type	Meaning
<b>TT1</b>	Events that pass through the ID. Includes internal events and muons passing through the ID.
<b>TT1 BTB0</b>	Events that happen in the ID but not the OD. Only internal events.
<b>TT1 BTB0 C1</b>	TT1 BTB0 events with one cluster i.e. single physical events that trigger the ID ( <i>internal and external background, neutrino</i> ).
<b>TT1 BTB0 CN</b>	TT1 BTB0 events with $N$ clusters e.g. coincidences of $N$ physical events in one DAQ window. Dominated by coincidences of two $^{14}\text{C}$ events (i.e. TT1 BTB0 C2).
<b>TT1 BTB4</b>	Events that pass through the ID and the OD ( <i>muons</i> ).
<b>TT2</b>	Events that happen only in the OD and do not pass through the ID ( <i>muons, light leak</i> ).
<b>TT64</b>	Whatever events happen to fall into the DAQ gate (untriggered).
<b>TT64 C0</b>	TT64 events that have zero clusters i.e. <i>dark noise</i> events.
<b>TT128</b>	<i>neutrons</i> , coincidences with other events (dominated by $^{14}\text{C}$ ).

TABLE 3.4: Event types depending on trigger type, BTB input and clutsering.

### 3.4.5 Particle identification

As mentioned in Sec. 3.2.2, some particles produce different hit time patterns in the ID due to the nature of their interaction with the LS, which allows us to distinguish some particle types like  $e^-$ ,  $e^+$  and  $\alpha$ . In this section I will describe the techniques used for differentiation of  $e^-/e^+$  and  $\alpha/\beta$  events in Borexino, relevant for the low-energy solar analysis of Sec. 7.1.

#### $e^-/e^+$ discrimination

As mentioned in Sec. 3.2.2, it is possible to distinguish  $e^-$  and  $e^+$  events exploiting the phenomenon of orthopositronium formation. This is done on statistical basis by using a discrimination parameter based on pulse shape by studying the characteristic delay of the  $e^+$  that form orthopositronium and calculating the fraction of such events. This technique was successfully used in Borexino and other liquid scintillator experiments [87, 107].

Currently, we use an improved particle identification parameter based on the likelihood of position reconstruction ( $\text{PS-}\mathcal{L}_{\text{PR}}$ ), shown in Fig. 3.11. Figure 3.11A shows the distribution of  $\text{PS-}\mathcal{L}_{\text{PR}}$  of simulated  $e^-$  events (black) and that of high purity strict sample of cosmogenic  $^{11}\text{C}$  background (red). This sample was selected with optimized *threefold coincidence* (TFC) algorithm, which searches for coincidences of the parent muon and capture of a neutron produced together with  $^{11}\text{C}$  [87, 95]. More on cosmogenic  $^{11}\text{C}$  and the TFC algorithm will be described

in Sec. 4.5.3. Figure 3.11B shows the distribution of  $\text{PS-}\mathcal{L}_{\text{PR}}$  and  $N_p^{dt_1}$  energy estimator in the dataset surviving solar neutrino analysis cuts which will be listed in Sec. 7.1.1.

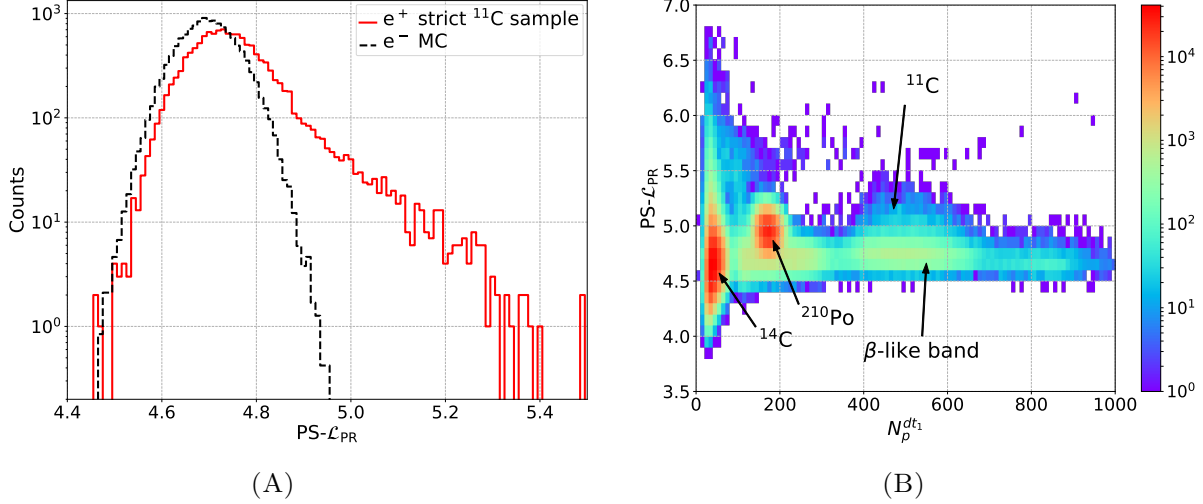


FIGURE 3.11: (A) Comparison of the distributions of the  $\text{PS-}\mathcal{L}_{\text{PR}}$  variable for MC-generated  $e^-$  events (black) and  $e^+$  events selected from data (green). (B) Distribution of  $\text{PS-}\mathcal{L}_{\text{PR}}$  as a function of  $N_p^{dt_1}$ .

An important use of the  $\text{PS-}\mathcal{L}_{\text{PR}}$  variable is to disentangle the residual  $^{11}\text{C}$  in the multivariate fit of the solar neutrino analysis that will be described in detail in Sec. 7.1.2.

#### $\alpha/\beta$ discrimination

As mentioned in Sec. 3.2.2,  $\alpha$ 's are strongly affected by ionization quenching, which influences the time distribution of the scintillation photons. In particular, the energy loss in space,  $dE/dx$ , is very different between  $\alpha$  and  $\beta$  particles. By selecting  $^{214}\text{Bi}$ - $^{214}\text{Po}$  coincidences originating from  $^{222}\text{Rn}$  decay (see Sec. 4.2.1), one can observe the difference in hit times, as shown in Fig. 3.12A. This difference can be exploited to construct  $\alpha/\beta$  discrimination variables.

A commonly used parameter is the **Gatti optimal filter**, which is a linear discrimination technique that allows to separate events in two classes based on their pulse shapes [108]. First, the probabilities  $P_\alpha(t)$  and  $P_\beta(t)$  that a photoelectron is detected between  $t$  and  $t + dt$  are calculated for  $\alpha$  and  $\beta$  events, respectively. Then, the Gatti parameter is calculated as follows:

$$G = \sum_n f(t_n)w(t_n), \quad (3.9)$$

where  $f(t_n)$  is the measured binned hit time distribution of an event,  $n$  being the number of bins; and  $w(t_n)$  stands for:

$$w(t_n) = \frac{P_\alpha(t_n) - P_\beta(t_n)}{P_\alpha(t_n) + P_\beta(t_n)}. \quad (3.10)$$

The distribution of  $G$  for the selection of  $^{214}\text{Bi}$  and  $^{214}\text{Po}$  events is shown in Fig. 3.12B.

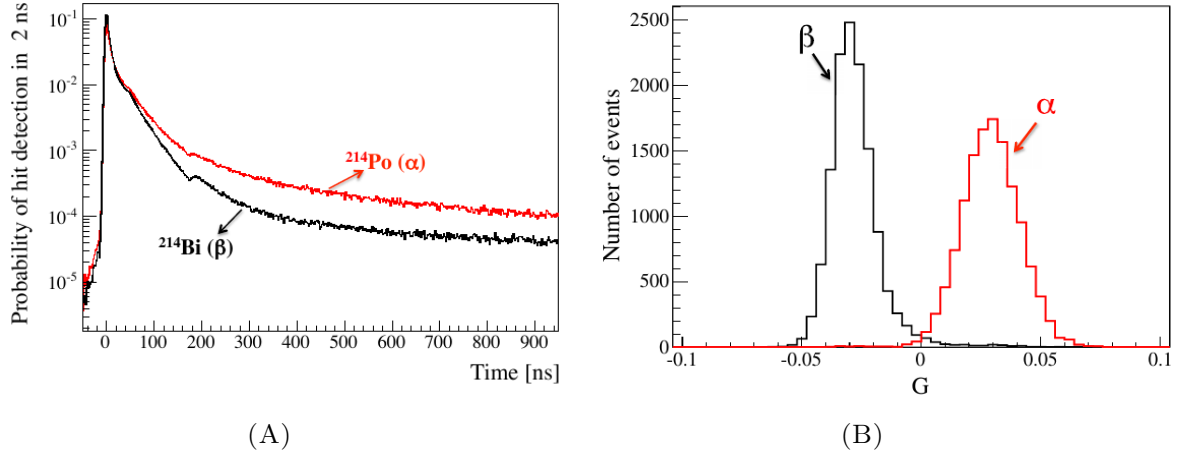


FIGURE 3.12: (A) Hit time distribution and (B) the Gatti parameter for  $^{214}\text{Po } \alpha$  (red) and  $^{214}\text{Bi } \beta$  (black) events [109].

Alternatively, a class of neural networks called **multilayer perceptron** (MLP) has been implemented in Borexino to build a binary classifier, shown in Fig. 3.13. This technique uses not only the hit time information, but many other inputs, and was trained on the same  $^{214}\text{Bi}$ - $^{214}\text{Po}$  data sample as the Gatti parameter, selected from the water extraction (WE) period (Sec. 3.2.6) where the  $^{222}\text{Rn}$  contamination was high; as well as the MC simulation of such events. The MLP classifier is used to select  $^{210}\text{Po}$  events from the data as part of the CNO neutrino detection strategy that will be outlined in Sec. 8.3.

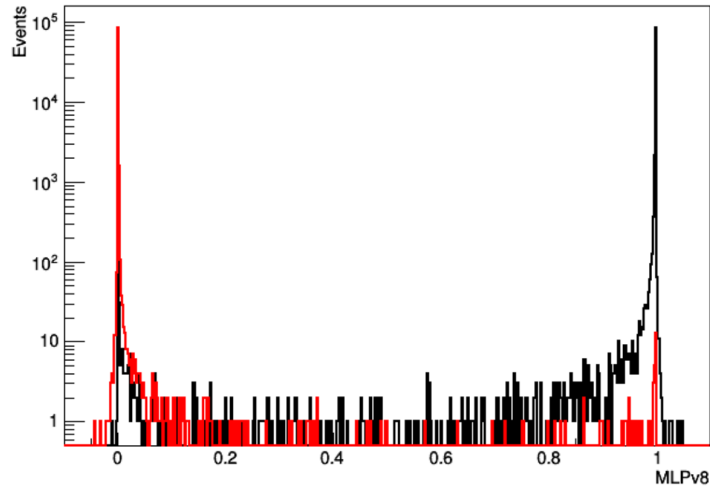


FIGURE 3.13: MLP classifier distribution of  $^{214}\text{Po } \alpha$  (red) and  $^{214}\text{Bi } \beta$  (black) events from the WE period data sample [110].

## 3.5 The Borexino detector simulation

The simulation of Borexino data consists of three stages:

- creation of physical events and light tracking using Monte Carlo Geant4 software (`g4bx2`);
- simulation of the Borexino electronics (`bx_elc`);
- processing of the resulting events with the `Echidna` reconstruction software, as mentioned in Sec. 3.4.2.

The latter means that the simulated and the real data undergo the same processing and have the same data structure, which allows us to use the exact same analysis software to analyze both. The simulation is used for validating different approaches, as mentioned before in this chapter, and for constructing reference shapes for the multivariate fit of the solar neutrino data (see Sec. 7.1.3).

### 3.5.1 Geant4 based Monte Carlo simulation

The Borexino Geant4-based Monte Carlo simulation, which we denote as `g4bx2`, is designed to model the physical processes associated with the energy deposition of a particle in the liquid scintillator (LS) [111].

The `g4bx2` procedure consists of the following steps.

1. Generation of the interaction of a particle (solar neutrino, radioactive decay, calibration source event etc.).
2. Simulation of the particle energy loss in the medium (LS, buffer, water).
3. Generation of scintillation and Cherenkov photons considering steps (1) and (2).
4. Tracking of each optical photon and its interactions with the medium until it is absorbed by a PMT.

In order to reproduce these processes, a comprehensive simulation of the detector geometry is performed. Examples of the 3D detector models are shown in Fig. 3.14. The data from the calibration campaign mentioned in Sec. 3.2.5 was used to tune `g4bx2`. This is one of the main advantages of using `g4bx2` in the analysis, since it has been validated on data independent from the one analyzed.

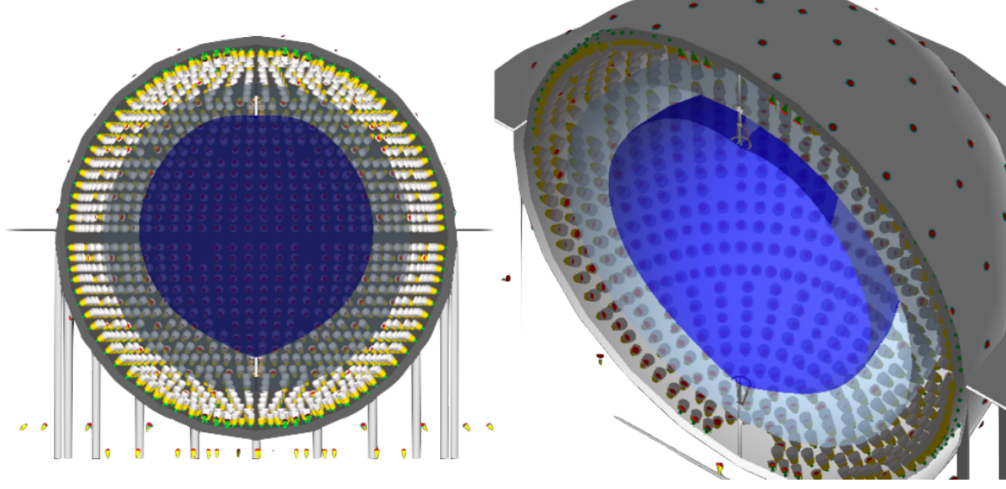


FIGURE 3.14: Simulation of the Borexino ID in g4bx2 [111].

### 3.5.2 Electronics simulation

The next stage after g4bx2 generation of physical events and processes, is the electronics simulation `bx_elec`, which models the response of PMTs and the electronic chain.

It consists of the following steps.

1. Generation of the PMT response to the photon(s) considering its quantum efficiency.
2. Generation of the PMT pulse based on the design of the electronics chain of Borexino described in Sec. 3.3.1.
3. Generation of the trigger and saving of the information for events that triggered.
4. Production of the raw data output in the same format as that of the Borexino DAQ system.

The last step after g4bx2 and `bx_elec` is the application of the Echidna reconstruction software, possible due to step (4) above, as mentioned before. Before, step (1) used to be done at the stage of g4bx2, however, in the recent update connected to the improved method of monitoring the quantum efficiency of the PMTs that I will described in Chapter 6, this stage has become part of the `bx_elec` procedure.



## Chapter 4

# Backgrounds in Borexino

In this Chapter, I am going to describe various sources of events in the Borexino detector that constitute background for the low energy region (LER) solar neutrino analysis, presented in Sec. 7.1 and Chapter 8. I am going to talk about the properties of the isotopes in discussion, as well as the methods of estimating the contamination, and tagging and removing these backgrounds from the dataset, if possible. As most of the backgrounds are discussed in the context of **fiducial volume** (FV), in which the LER analysis is performed, I will discuss the spatial distribution of background and the selection of FV in Sec. 4.1. In Sec. 4.2, I am going to talk about the unprecedented radiopurity levels of Borexino, and relevant isotopes belonging to the radioactive decay chain of  $^{238}\text{U}$ . The LER backgrounds are classified as follows:

- **external**, originating outside the liquid scintillator (LS);
- **surface**, contaminants on the nylon Inner Vessel (IV);
- **internal**, isotopes contaminating the LS itself;
- **cosmogenic**, cosmic muons and muon-induced isotopes.

These types of backgrounds will be discussed in Sec. 4.3 - 4.5. A summary of the isotopes relevant for the LER solar neutrino analysis is shown in Table 4.1.

Isotope	Type	Origin	Particle	Energy (keV)
$^{208}\text{Tl}$	E	PMTs, SSS, cones	$\gamma$	2614
$^{40}\text{K}$	E	PMTs, SSS, cones	$\gamma$	1460
$^{214}\text{Bi}$	E	PMTs, SSS, cones	$\gamma$	< 1764
	S	$^{222}\text{Rn}$ on IV	$e^-$	3272
$^{214}\text{Po}$	S	$^{214}\text{Bi}$ on IV	$\alpha$	7686
$^{210}\text{Bi}$	S	$^{222}\text{Rn}$ and $^{210}\text{Pb}$ on IV	$e^-$	1160
	I	$^{210}\text{Pb}$ in the LS		
$^{210}\text{Po}$	S	$^{210}\text{Bi}$ on IV, separate IV contamination	$\alpha$	5300
	I	$^{210}\text{Bi}$ in LS, convection from IV		
$^{14}\text{C}$	I	intrinsic part of LS	$e^-$	156
$^{85}\text{Kr}$	I	air	$e^-$	687
$^{11}\text{C}$	C	$\mu$ spallation	$e^+$	$960 + 2 \times 511$

TABLE 4.1: External (E), surface (S), internal (I) and cosmogenic (C) isotopes that constitute main backgrounds for the Borexino LER solar neutrino analysis. For  $\beta$ -emitters, column Energy reports  $Q$ -values.

In addition, there is a contribution from the so-called **pileup** of two or more events, which happens when the events occur too close in time for the Echidna software to distinguish them as separate clusters. This special type of background is important for the measurement of the  $pp$  neutrinos, and will be discussed in Sec. 4.6.

## 4.1 Fiducial volume

Figure 4.1 shows the spatial distribution of reconstructed events in main energy regions of interest: Fig. 4.1A  $^{210}\text{Po}$  peak, Fig. 4.1B  $^7\text{Be}$  neutrinos, Fig. 4.1C  $^{11}\text{C}$ , and Fig. 4.1D  $^{208}\text{Tl}$ . The lowest energy region, represented in Fig. 4.1B, most clearly shows events inside the IV and near the top end cap, occurring due to a small leak of LS from the IV to the buffer region.

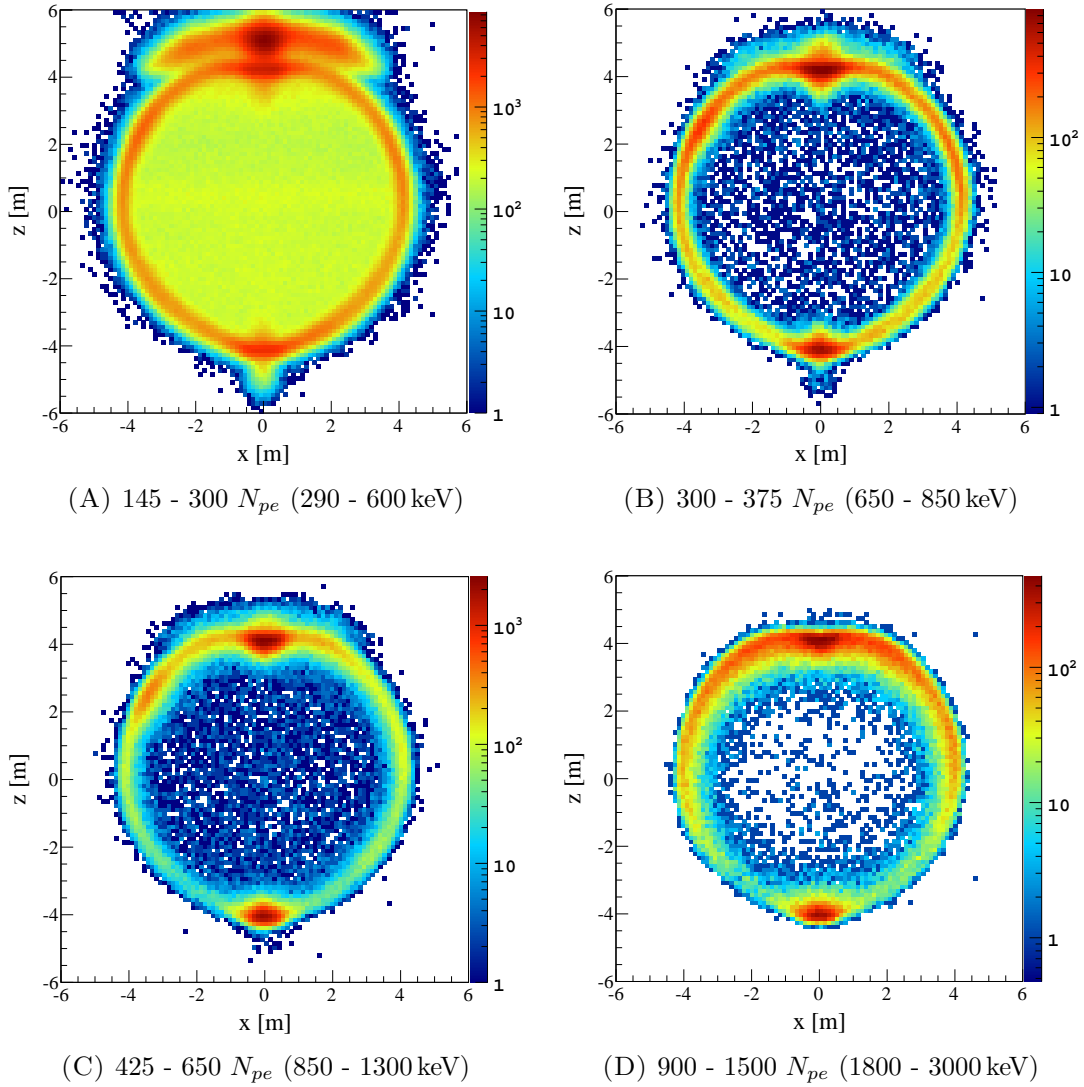


FIGURE 4.1: Distribution of all reconstructed events except muons in the  $x$ - $z$  plane with  $|y| < 0.5$  m in different energy regions. The color axis represents number of events per  $0.0144 \text{ m}^3$  [95].



The rate of background events is discussed in the units of counts per day per 100 t of LS (cpd/100t), or counts per second per 100 t (Bq/100t) for  $^{14}\text{C}$  which has a much higher rate than other isotopes.

The LER analysis is performed inside the **fiducial volume** (FV). It is defined as  $r < 2.8\text{ m}$  and  $-1.8\text{ m} < z < 2.2\text{ m}$ , where  $r$  is the radius and  $z$  is the vertical coordinate inside the IV based on reconstructed event position,  $(0,0,0)$  being the center of the detector. A two dimensional  $xz$ -plane slice (at  $y = 0$ ) of the FV is demonstrated in Fig. 4.2A. The FV is chosen this way in order to maximize the active volume while minimizing radioactive background coming from the nylon spheres, PMTs, the SSS, and the endcaps. It is asymmetrical in  $z$  coordinate to account for the higher levels of radioactive contamination in the lower hemisphere.

In the context of the LER solar neutrino analysis, which is the focal point of Sec. 7.1 and Chapter 8, some backgrounds, like muons and muon daughters or  $^{214}\text{Bi}$ - $^{214}\text{Po}$  coincidences, can be removed via *data selection cuts*, which will be described in Sec. 4.5.1 and Sec. 4.2.2, respectively. Note that the FV cut is one of the most important selection cuts, as it greatly reduces backgrounds, which is demonstrated in Fig. 4.2B.

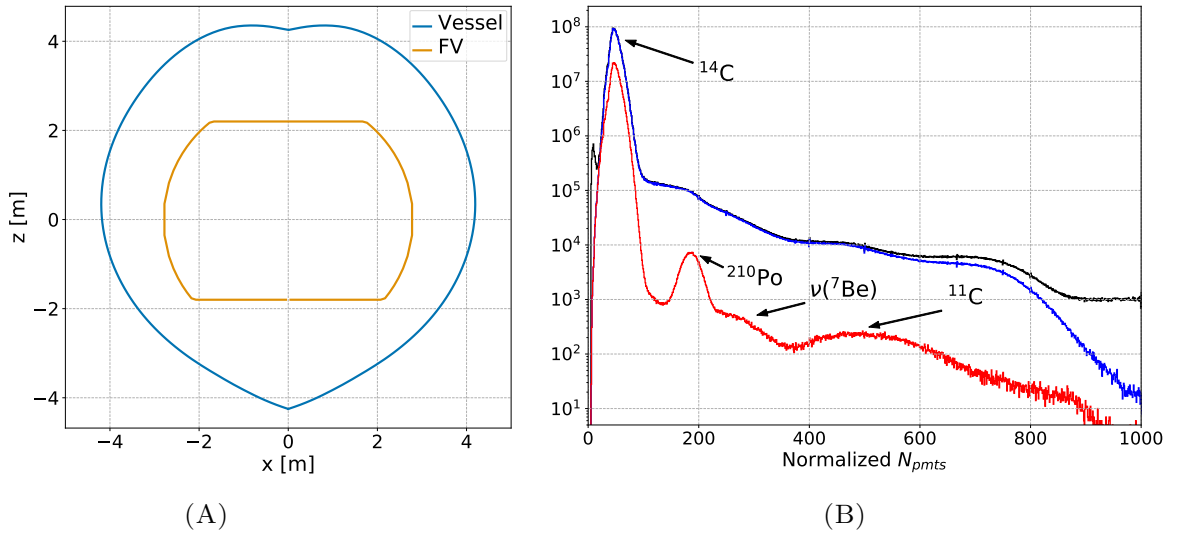


FIGURE 4.2: (A) Fiducial volume shape (orange) compared to the Borexino vessel (blue). (B) Phase-II energy spectrum of all Borexino internal events (black) and the spectra after applying the muon and muon daughter cut (blue) and FV cut (red), in this order. Major backgrounds visible after the FV cut are marked in the plot.

The backgrounds that cannot be removed are indistinguishable from neutrino events, and their contribution is disentangled performing a *multivariate fit* that will be discussed in Sec. 7.1.2, using, among other inputs, the **spectral shapes** of the backgrounds and signals. In Fig. 4.3, I demonstrate the spectral shapes of the background isotopes listed in Table 4.1, constructed using the so-called *analytical approach* for all isotopes except the external backgrounds (for which the *Monte Carlo approach* is used), after performing a fit on Phase-II data after aforementioned selection cuts, the main effect of which comes from the FV cut. Both approaches will be explained in detail in Sec. 7.1.3. In this Chapter, without focusing on the details of the spectral shape construction and the fitting procedure, I will use Fig. 4.1 for the purpose of demonstrating the energy spectra of the background isotopes in the FV.

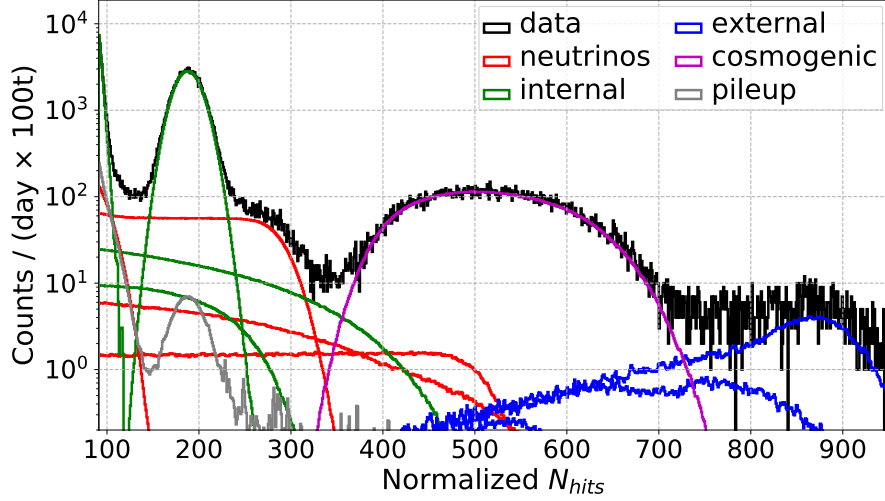


FIGURE 4.3: Borexino Phase-II data in FV after LER solar neutrino analysis cuts (black) with spectral shapes of solar neutrinos and background isotopes, constructed using the MC approach, after performing a multivariate fit.

## 4.2 Isotopes from the $^{238}\text{U}$ chain

The radioactive **uranium** and **thorium** isotopes,  $^{238}\text{U}$  and  $^{232}\text{Th}$ , are commonly present in the surrounding dust and rocks, and can thus easily contaminate the detector during construction. Contamination by these isotopes would be very problematic, as they are extremely long-lived, with half life of  $4.468 \cdot 10^9$  years and  $1.406 \cdot 10^{10}$  years for  $^{238}\text{U}$  and  $^{232}\text{Th}$ , respectively. Figure 4.4A depicts a schematic representation of part of the  $^{238}\text{U}$  decay chain, starting from one of its daughters,  $^{222}\text{Rn}$ . The decay sequences  $^{214}\text{Bi}$ - $^{214}\text{Po}$  and  $^{210}\text{Bi}$ - $^{210}\text{Po}$  (highlighted in red) are the main isotopes belonging to this chain that constitute background for the LER solar neutrino analysis.

Thanks to the purification campaign, described in Sec. 3.2.6, the levels of uranium and thorium were reduced to negligible levels, measured to be  $(5.3 \pm 0.5) \cdot 10^{-18}$  g/g for  $^{238}\text{U}$  and  $(3.8 \pm 0.8) \cdot 10^{-18}$  g/g for  $^{232}\text{Th}$  in FV [95]. The contamination has been estimated using  $^{214}\text{Bi}$ - $^{214}\text{Po}$  and  $^{212}\text{Bi}$ - $^{212}\text{Po}$  event coincidences, respectively, which appear in the decay chains of these isotopes. For comparison, the natural levels of  $^{238}\text{U}$  are reported to be  $3 \cdot 10^{-7} - 11.7 \cdot 10^{-6}$  g/g in soil and  $3.3 \cdot 10^{-9}$  g/g in seawater. This makes the Borexino scintillator to be 9-10 orders of magnitude less radioactive than materials in nature. Furthermore, the contamination levels are substantially lower than the initial design goals, which were  $< 10^{-16}$  g/g for  $^{238}\text{U}$  and  $< 10^{-14}$  g/g for  $^{232}\text{Th}$  [103]. This unprecedented radiopurity of Borexino is a milestone achievement that made the solar neutrino analysis possible.

Some elements belonging to the  $^{238}\text{U}$  chain, namely,  $^{222}\text{Rn}$  and possibly  $^{210}\text{Pb}$  (highlighted in blue in Fig. 4.4A), are or may be present in the detector through independent contamination not in equilibrium with  $^{238}\text{U}$ , and comprise a separate source of bismuth and polonium isotopes mentioned above. I will discuss the properties and roles of the isotopes belonging to the  $^{238}\text{U}$  chain in Sec. 4.2.1 below.

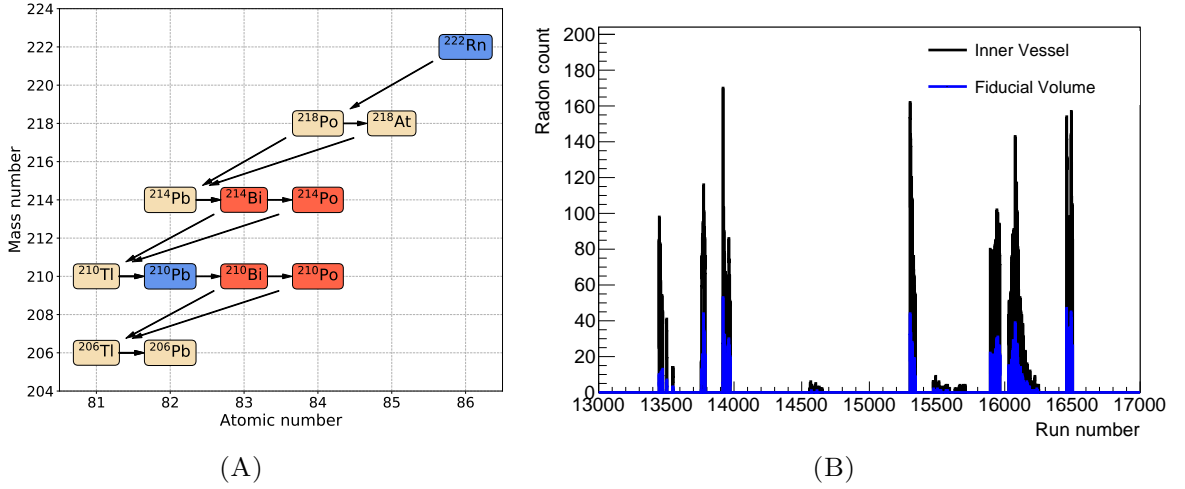


FIGURE 4.4: (A) A schematic representation of the  $^{222}\text{Rn}$  chain. Main isotopes constituting background for the LER analysis are highlighted in red. Main sources of said isotopes in Borexino are highlighted in blue. (B) Amount of  $^{222}\text{Rn}$  events in the IV (blue) and FV (red) in each run belonging to the purification period (see text).

#### 4.2.1 Radon isotope $^{222}\text{Rn}$

During the purification campaign, discussed in Sec. 3.2.6, the IV was contaminated with  $^{222}\text{Rn}$ . As it has a half life of 3.8 days, much shorter than  $^{238}\text{U}$ , most of it decayed since the campaign in 2011, and does not pose a large source of background. Figure 4.4B shows the amount of  $^{222}\text{Rn}$  events in each run belonging to the period of purification, determined using  $^{214}\text{Bi}$ - $^{214}\text{Po}$  tagging, which will be described in Sec. 4.2.2. Six main spikes in the  $^{222}\text{Rn}$  activity correspond to the six water extraction periods during the campaign. Current data shows little to no  $^{222}\text{Rn}$  (see p. 154 of App. A), and virtually none in the FV used in the LER analysis.

#### 4.2.2 Pairs of $^{214}\text{Bi}$ - $^{214}\text{Po}$

The fast decay sequence  $^{214}\text{Bi}$ - $^{214}\text{Po}$  highlighted in Fig. 4.4A starts with the  $\beta$  decay of  $^{214}\text{Bi}$ , followed by the prompt  $\alpha$  decay of  $^{214}\text{Po}$ :

$$^{214}\text{Bi} \rightarrow ^{214}\text{Po} + e^- + \bar{\nu}_e, \quad \tau = 28.7 \text{ m}, \quad (4.1)$$

$$^{214}\text{Po} \rightarrow ^{210}\text{Pb} + \alpha, \quad \tau = 237 \mu\text{s}. \quad (4.2)$$

Tagging of  $^{214}\text{Bi}$ - $^{214}\text{Po}$  sequences via searching for candidate coincidences in the Borexino data is based on the following selection:

- the time difference of the two events is  $\Delta t < 1.5 \text{ ms}$ , based on  $\tau$  from Eq. 4.2;
- the distance between the two events is  $d < 1 \text{ m}$ , to account for diffusion and position reconstruction accuracy;
- the energy ranges of the two reactions are well known:  $90 < N_{pe} < 1800$  for  $^{214}\text{Bi}$ , and  $200 < N_{pe} < 500$  for  $^{214}\text{Po}$  decay.

These conditions provide very specific signature for such coincidences, which allows us to select high purity  $^{214}\text{Bi}$ - $^{214}\text{Po}$  samples that can be used to monitor  $^{222}\text{Rn}$  concentration, set limits on the  $^{238}\text{U}$  contamination, and apply an efficient cut to remove the few background events originating from  $^{214}\text{Bi}$  and  $^{214}\text{Po}$  in the solar neutrino analysis data selection. Moreover, since this tagging approach does not rely on any particle identification, it can be used to produce a data sample of  $e^-$ - $\alpha$  events that can be used to benefit the analysis. Indeed, we exploit the high amount of  $^{222}\text{Rn}$  events, observed during the WE period mentioned above, to create a  $^{214}\text{Bi}$ - $^{214}\text{Po}$  sample which we use to develop  $\alpha/\beta$  discrimination methods, as described in Sec. 3.4.5.

### 4.2.3 Bismuth isotope $^{210}\text{Bi}$

Further down the  $^{222}\text{Rn}$  decay chain, shown in Fig. 4.4A, one finds the  $\beta$  emitter  $^{210}\text{Bi}$ :

$$^{210}\text{Bi} \rightarrow ^{210}\text{Po} + e^- + \bar{\nu}_e, \quad \tau = 7.23 \text{ d.} \quad (4.3)$$

Since  $^{214}\text{Bi}$ - $^{214}\text{Po}$  tagging confirms little to no radon contamination on the IV, one would expect few  $^{210}\text{Bi}$  events originating from the IV. However, there is  $^{210}\text{Bi}$  contamination present on the IV long after the radon introduced in 2010-2011 has decayed. Moreover,  $^{210}\text{Bi}$  events are present in the FV, where there are little to no  $^{214}\text{Bi}$ - $^{214}\text{Po}$  coincidences. This implies a separate source of  $^{210}\text{Bi}$ , not in equilibrium with  $^{222}\text{Rn}$ ; for instance, an independent  $^{210}\text{Bi}$  contamination of the IV.

The most likely source is the lead isotope  $^{210}\text{Pb}$ , highlighted in blue in Fig. 4.4A, which may be contaminating the IV and the scintillator, as it is a very commonly seen element. Its Q value is below  $^{14}\text{C}$ , making  $^{120}\text{Pb}$  itself not visible in the event spectrum. The treatment of  $^{210}\text{Bi}$  as surface and internal background will be discussed in Sec. 4.3 and Sec. 4.4.3, respectively.

### 4.2.4 Polonium isotope $^{210}\text{Po}$

The decay of  $^{210}\text{Bi}$ , shown in Eq. 4.3, is followed by the  $\alpha$  decay of its daughter,  $^{210}\text{Po}$ :

$$^{210}\text{Po} \rightarrow ^{206}\text{Pb} + \alpha, \quad \tau = 199.7 \text{ d.} \quad (4.4)$$

The main difference in the  $^{210}\text{Bi}$ - $^{210}\text{Po}$  sequence compared to  $^{214}\text{Bi}$ - $^{214}\text{Po}$  is the very long time separation between the  $e^-$  and the  $\alpha$  events, governed by the mean lifetime of  $^{210}\text{Po}$  shown in Eq. 4.4, namely, 199.7 d. Due to this, it is not possible to remove  $^{210}\text{Bi}$  and  $^{210}\text{Po}$  events from the Borexino data using a selection cut in the same way as  $^{214}\text{Bi}$ - $^{214}\text{Po}$ . For this reason,  $^{210}\text{Bi}$  and  $^{210}\text{Po}$  events that are present in the inner volume of the scintillator have to be treated as internal backgrounds, and will be discussed in Sec. 4.4.3 and Sec. 4.4.4, respectively.

Moreover, apart from the daughter  $^{210}\text{Po}$  events, which are in equilibrium with  $^{210}\text{Bi}$ , there is a separate additional contamination of  $^{210}\text{Po}$  on the IV, which will be discussed in Sec. 4.3.

### 4.3 Surface and external backgrounds

The contaminants on the nylon IV constitute the so-called **surface backgrounds**. They include isotopes shown in the decay chain of  $^{222}\text{Rn}$  in Fig. 4.4A, discussed in Sec. 4.2, namely,  $^{222}\text{Rn}$ , which produces  $^{214}\text{Bi}$ - $^{214}\text{Po}$  events;  $^{210}\text{Bi}$ , and/or possibly  $^{210}\text{Pb}$  (which could be the source of  $^{210}\text{Bi}$  events); and a  $^{210}\text{Po}$  contamination out of equilibrium with  $^{210}\text{Bi}$ .

Some surface contaminants can be exploited to benefit the analysis. In particular,  $^{210}\text{Bi}$  residing on the IV surface is used for dynamic vessel shape reconstruction [95]. Indeed, higher activity at the radius  $R \approx 4.25\text{ m}$  seen in Fig. 4.1 comes from the radioactive isotopes residing on the IV.

Other surface contaminants may be washed off into the inner volume by LS currents, becoming a source of *internal background*. For example,  $^{222}\text{Rn}$  can enter the scintillator, but since its diffusion rate is only 1 cm/day, while its half life is 3.8 days,  $^{214}\text{Bi}$ - $^{214}\text{Po}$  events coming from  $^{222}\text{Rn}$  are concentrated only on the IV or close to it.

Unlike  $^{222}\text{Rn}$ ,  $^{210}\text{Po}$  is easily washed off the IV surface by currents, and enters the FV. Figure 4.5A illustrates  $^{210}\text{Po}$  contamination (purple) on the IV penetrating the inner volume (black dashed region) due to convection. As a result, a high amount of internal  $^{210}\text{Po}$  events is observed, as can be seen on the left side of Fig. 4.5B, where the  $^{210}\text{Po}$  rate is shown, measured in cubic volumes uniformly dividing the inner vessel. After the detector thermal insulation campaign conducted in 2015 (Sec. 3.2.7), the influence of the changing temperature of the Hall C, where Borexino is located, was minimized. The resulting thermal stabilization created a more uniform temperature profile, reducing turbulent currents, and preventing more  $^{210}\text{Po}$  events from entering the inner volume. This effect can be seen on the right part of Fig. 4.5B, where  $^{210}\text{Po}$  rate is much lower and stable. More on the treatment of the internal  $^{210}\text{Po}$  background will be discussed in Sec. 4.4.4.

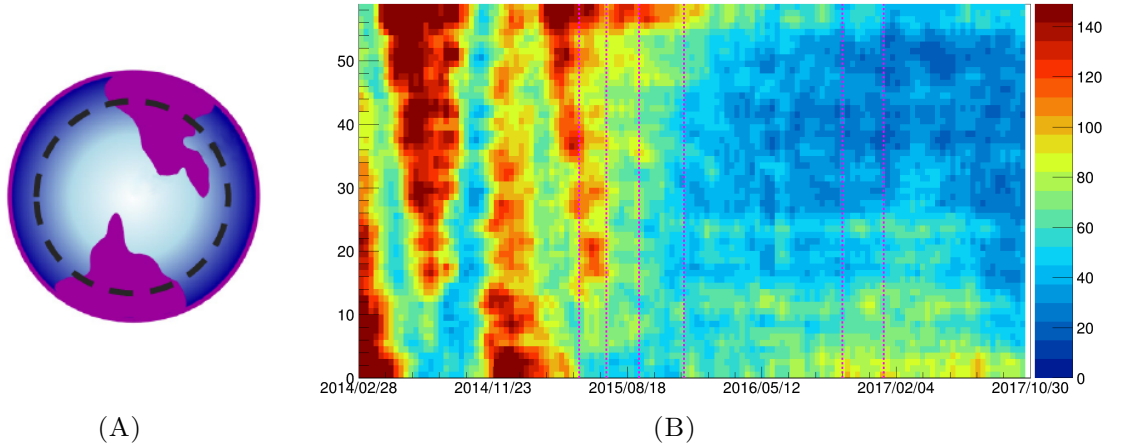


FIGURE 4.5: (A) Graphic illustration of  $^{210}\text{Po}$  contamination (purple) on the IV penetrating the inner volume (black dashed region) due to convection [112]. (B) The dependence of  $^{210}\text{Po}$  rate in cubic volumes on time [112, 113].

The main source of **external background** in Borexino is the radioactivity of the materials surrounding the LS, such as vessel support structure (Sec. 3.2.1), PMTs and their light concentrators (Sec. 3.2.3). The main external backgrounds are  $^{208}\text{Tl}$ ,  $^{214}\text{Bi}$ , and  $^{40}\text{K}$ , relevant for the

detection of  $pep$  and CNO neutrinos. A small contamination of these isotopes is present also on the nylon vessels. Consequently,  $\gamma$ 's are the only particles resulting from these radioactive decays that can reach the inner volume of the scintillator. The principle of  $\gamma$  particle detection in Borexino has been described in Sec. 3.2.2. The spatial distribution of external background can be seen in Fig. 4.1D, which depicts the energy region dominated by  $^{208}\text{Tl}$ . As can be seen, the activity decreases with radius, and is much smaller in the center of the IV.

The spectral shapes of the external isotopes are constructed using the Borexino Monte Carlo (MC) simulation (Sec. 3.5), based on the measured contamination of the PMTs [103], as well as SSS and concentrators [114]. The resulting reference shapes are shown in Fig. 4.6, highlighted in blue. Apart from the energy spectra, MC reference shapes for radial distribution are used to exploit the different distribution of external events as compared to the uniform internal background components. The simulation has been validated using  $^{208}\text{Tl}$  data from the  $^{228}\text{Th}$  calibration source (Sec. 3.2.5).

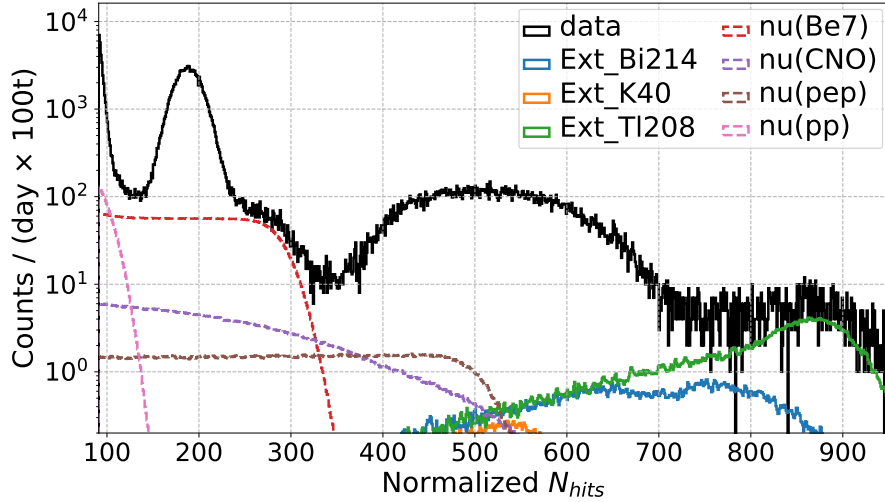


FIGURE 4.6: Spectral shapes of the external background isotopes compared to those of the solar neutrinos, refer to Fig. 4.3.

## 4.4 Internal background

The main **internal background** components are shown in Fig. 4.7:  $^{14}\text{C}$  (blue),  $^{85}\text{Kr}$  (orange),  $^{210}\text{Bi}$  (green), and  $^{210}\text{Po}$  (red). As can be seen, internal backgrounds cover the energy range relevant for the measurement of all solar neutrinos.

In particular, the knowledge of the very low-energy  $^{14}\text{C}$  background is crucial for  $pp$  neutrinos, as well as  $^{210}\text{Po}$ , contributing with the number of events several orders of magnitude above solar neutrinos. As can be seen in Fig. 4.7, the major part of the  $^{14}\text{C}$  spectrum is not included in the LER analysis, and the full spectrum will be shown later in Fig. 4.8A. The  $^{85}\text{Kr}$  background events are the most relevant for  $^7\text{Be}$  neutrino measurement; while  $^{210}\text{Bi}$  is the most challenging for the CNO neutrino detection, especially as their spectral shapes are extremely similar. More details about each isotope will be discussed in the subsections below.

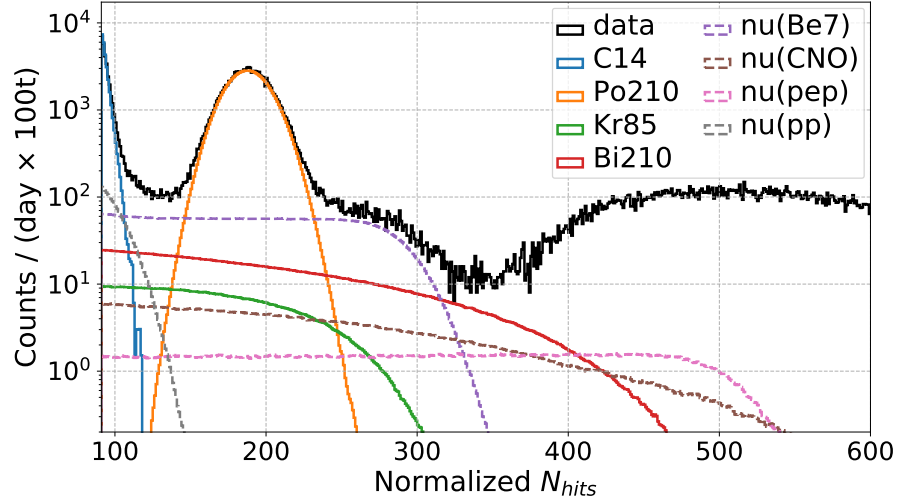


FIGURE 4.7: Spectral shapes of the internal background isotopes compared to those of the solar neutrinos, refer to Fig. 4.3.

#### 4.4.1 Carbon isotope $^{14}\text{C}$

The most dominant background in the Borexino detector is the  $\beta$  emitter  $^{14}\text{C}$  (mean lifetime  $\tau = 8270$  years), the decay of which is shown in Eq. 4.5:



As can be seen from Fig. 4.8A, the rate of  $^{14}\text{C}$  is several orders of magnitude higher than any other background or signal.

This isotope is present in the detector because it is chemically indistinguishable from  $^{12}\text{C}$ , which is a natural component of organic liquid scintillators. The rate of  $^{14}\text{C}$  decays in the scintillator is around 40 Bq/100t, and can be estimated by studying the second clusters in the DAQ gate of two-cluster TT1 BTB0 C2 events (Sec. 3.4.4), the spectrum of which is not influenced by the trigger PMT threshold. The absolute majority of such coincidences is dominated by  $^{14}\text{C}$  events. The resulting spectra are shown in Fig. 4.8B.

Due to the trigger threshold, and limitations of energy reconstruction, the minimum energy of the the LER analysis spectral fit is chosen to be  $N_p = 85$ . The end point of the  $^{14}\text{C}$  spectrum (156 keV), which corresponds to roughly  $N_p = 120$ , is well above this minimum, which allows us to include  $^{14}\text{C}$  in the fit, as seen in Fig. 4.7. Nevertheless, placing a constraint on  $^{14}\text{C}$  rate, based on the studies described above, is important for the  $pp$  neutrino measurement. Apart from being a source of background events,  $^{14}\text{C}$  can be used for various studies, since it has a stable energy range and a lot of statistics. For example, it is used to study the effective quantum efficiency of the PMTs, as will be shown in Chapter 6.



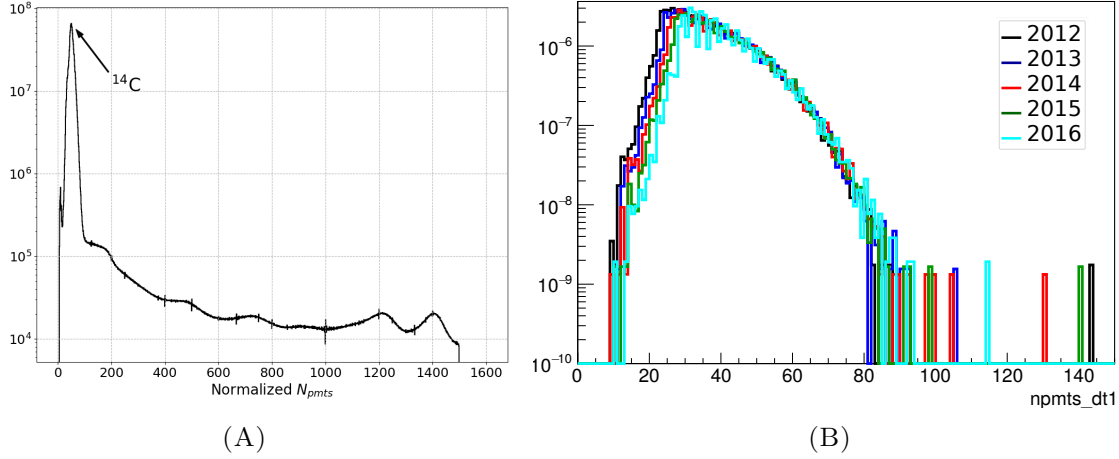


FIGURE 4.8: (A) The spectrum of all internal Borexino events (TT1) from Phase-II demonstrating the prominent  $^{14}\text{C}$  peak. (B) Spectral shapes of untriggered second cluster events for different years.

#### 4.4.2 Krypton isotope $^{85}\text{Kr}$

The **krypton** isotope  $^{85}\text{Kr}$  is a  $\beta$  emitter with a mean lifetime  $\tau = 15.4$  years. The spectral shape of the electrons coming from  $^{85}\text{Kr}$ , shown in Fig. 4.7, is very similar to the one of the electrons recoiled from  $^7\text{Be}$  neutrinos, making  $^{85}\text{Kr}$  one of the most important backgrounds for the  $^7\text{Be}$  neutrino measurement. It is present in the air with a concentration of  $\sim 1 \text{ Bq/m}^3$ , therefore even small air exposures during the filling of the LS result in its contamination. With a branching ratio of 43%,  $^{85}\text{Kr}$  decays into a metastable  $^{85}\text{Rb}^*$  state, emitting a  $\beta$  particle with energy  $E < 173 \text{ keV}$ , which in turn decays into the ground state  $^{85}\text{Rb}$  (mean lifetime  $\tau = 2.06 \mu\text{s}$ ), emitting a  $\gamma$  particle of  $514 \text{ keV}$ :



The resulting  $\beta - \gamma$  coincidences are exploited to estimate the rate of  $^{85}\text{Kr}$ , which is measured to be  $< 7.5 \text{ cpd}/100 \text{ t}$  in Phase-II, and  $< 9.8 \text{ cpd}/100 \text{ t}$  in Phase-III at 95% C.L. [95, 115]. This independent estimation can be used to place a constraint on the  $^{85}\text{Kr}$  rate in the multivariate fit of the LER analysis, which will be described in Sec. 7.1.2.

#### 4.4.3 Bismuth isotope $^{210}\text{Bi}$

The **bismuth** isotope  $^{210}\text{Bi}$  is a  $\beta$  emitter with mean lifetime  $\tau = 7.23$  days (see Eq. 4.3), its spectrum covering the energy regions of  $^7\text{Be}$ ,  $pep$ , and CNO neutrinos, as can be seen in Fig. 4.7. As explained in Sec. 4.2.3, the most likely source of  $^{210}\text{Bi}$  in the inner volume is the  $^{210}\text{Pb}$  contamination of the LS. Bismuth does not have a specific signature by which it can be determined, except its spectrum, making it a background for LER solar neutrino analysis that is not possible to disentangle. For the measurement of  $pp$ -chain neutrinos, it is enough to include the spectral shape of  $^{210}\text{Bi}$  into the multivariate fit procedure (Sec. 7.1.2).



However, it poses a bigger issue for the CNO neutrinos, the spectral shape of which is very similar to that of  $^{210}\text{Bi}$ . More details on the CNO neutrino detection strategy regarding this challenge will be presented in Sec. 8.3.1.

#### 4.4.4 Polonium isotope $^{210}\text{Po}$

The **polonium** isotope  $^{210}\text{Po}$  is the most abundant internal background after  $^{14}\text{C}$ , as can be seen in Fig. 4.7. It is an  $\alpha$  emitter with mean lifetime  $\tau \approx 200$  days (see Eq. 4.4). As described in Sec. 3.2.2,  $\alpha$ 's are subject to strong *quenching* effect, which moves the  $^{210}\text{Po}$  energy down by factor 10, bringing it into the lower energy range, affecting  $pp$ ,  $^7\text{Be}$ ,  $pep$ , and CNO neutrinos.

Since  $^{214}\text{Po}$  events can be removed exploiting fast coincidence signature  $^{214}\text{Bi}$ - $^{214}\text{Po}$ , described in Sec. 4.2.2, it leaves  $^{210}\text{Po}$  to be the only  $\alpha$  background relevant for the LER analysis. Using  $\alpha/\beta$  discrimination techniques, mentioned in Sec. 3.4.5, a  $^{210}\text{Po}$  sample can be selected from data. Such samples are used for different studies, like that of effective light yield, which will be mentioned in Sec. 6.7.2.

Another use of such a sample is for measurement of  $^{210}\text{Po}$  rate, which we can use to determine the rate of  $^{210}\text{Bi}$ , important for the CNO neutrino detection. As mentioned before,  $^{210}\text{Po}$  events that constitute internal background come from two sources: decay of  $^{210}\text{Bi}$  isotopes present in the inner volume, and a separate IV contamination that enters the inner volume due to currents and convection, discussed in Sec. 4.3. Methods for disentangling this additional contribution from the one in equilibrium with  $^{210}\text{Bi}$  will be discussed in Sec. 8.3.1.

## 4.5 Cosmogenic background

The muon flux passing through the Borexino detector is strongly reduced compared to the flux at sea level, which amounts to around  $6.5 \cdot 10^5 \text{ m}^{-2}\text{hr}^{-1}$  [116]. This is achieved because of the shielding provided by the rock above the underground laboratory where Borexino is located. The average muon rate in Borexino is around  $1.2 \text{ m}^{-2}\text{hr}^{-1}$  [97], which amounts to  $\sim 6000$  events per day, half of them crossing only the outer detector (OD), which we call *external muons*; and half crossing both outer and inner (ID) detector, called *internal muons*. Muons are detected in Borexino using the OD, which serves as an active muon veto; and additional information from the ID. More on muon detection will be described in Sec. 4.5.1.

Not only do muons themselves constitute background for various analyses, they also create so-called “*muon daughters*”, products of muon spallation on the LS molecules. Most of the muon daughters are removed after the 300 ms veto applied after each detected internal muon. The residual rates of various daughters span from  $< 5 \cdot 10^{-5} \text{ cpd}/100 \text{ t}$  (e.g.  $^{12}\text{N}$ ,  $^{13}\text{B}$ ) to  $\sim 0.54 \text{ cpd}/100 \text{ t}$  ( $^{10}\text{C}$ ). One of the muon daughters are cosmogenic neutrons, on which I will focus in Sec. 4.5.2.

An important isotope that does not get removed by the muon veto is  $^{11}\text{C}$ . It presents background for the LER analysis, and requires special treatment due to its long mean lifetime. I will discuss  $^{11}\text{C}$  in more detail in Sec. 4.5.3.

### 4.5.1 Muon detection

In Borexino, three methods are employed for muon identification [97]: Muon Trigger Flag (MTF), Muon Cluster Flag (MCF), and Inner Detector Flag (IDF). Details about each method are summarized in Table 4.2.

Method	Efficiency	Information	Condition
MTF	0.9925	OD	Trigger in the OD.
MCF	0.9928	OD	There is at least one cluster in the event registered by the OD, as identified by Echidna (Sec. 3.4.2).
IDF	0.9890	ID	Identified by number of clustered hits and pulse shape in TT1 events.

TABLE 4.2: Muon identification methods in Borexino.

MTF and MCF can identify both internal and external muons, for which an extra condition must be included to determine whether the muon passed through the ID (TT1BTB4), or happened only in the OD (TT2); while IDF identifies internal muons only. More on trigger types and event types in the detector can be found in Sec. 3.3.2 and Sec. 3.4.4, respectively.

Figure 4.9 shows internal and external muon rate as determined by the three methods.

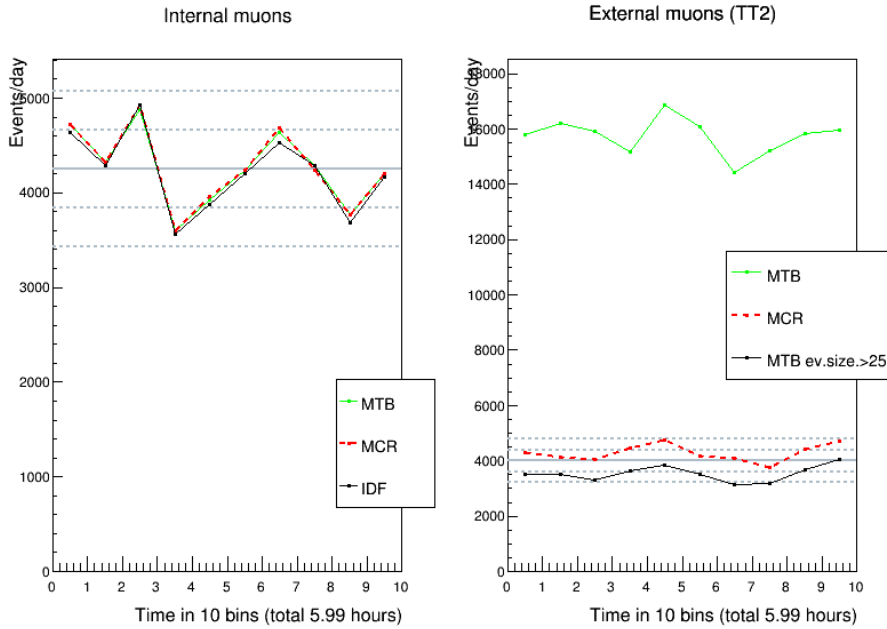


FIGURE 4.9: Internal (left) and external (right) muon rate, calculated using MTF (green), MCF (red) and IDF (black in the left plot). For external muons, MTF method with an extra condition for number of decoded hits to be  $> 25$  is shown in black.

As can be seen in the second plot, without additional conditions, MTF shows a much higher rate of external muons. This happens due to the so-called *light leak*, which constitutes 70% of the OD triggers. This contribution is removed with an additional condition on the number of hits, since scattered light signals never produce many (shown in black in the second plot of

Fig. 4.9). Figure 4.9 is a figure that belongs to the so-called *RunValidation* procedure, used for data quality monitoring. More about this will be discussed in Sec. 5.1, and on p. 164 in App. A, for the discussion of this plot in particular.

Tagging of internal muons and the consequent muon daughter veto are important for the solar neutrino analysis. The so-called muon and muon daughter cut constitutes one of the steps of data selection for low- (Sec. 7.1.1) and high-energy (Sec. 7.2.1) neutrino analysis. In addition, muon samples are used to study seasonal variations in muon flux [92].

#### 4.5.2 Cosmogenic neutrons

Fast neutrons are one of the products of muon spallation on  $^{12}\text{C}$  in the LS. In the 99% of the cases they are captured on hydrogen:

$$n + p \rightarrow \text{D} + \gamma, \quad (4.7)$$

producing a gamma ray of 2.2 MeV; or on  $^{12}\text{C}$ . Average neutron capture time in the scintillator has been measured to be  $\sim 250 \mu\text{s}$  [116].

As mentioned before, in the LER solar neutrino analysis, which will be described in Sec. 7.1, a 300 ms veto after each detected muon is applied, removing the vast majority of neutrons, with the resulting residual rate of  $< 0.005 \text{ cpd}/100 \text{ t}$  [95]. For the high energy neutrino analysis, neutron captures on  $^{12}\text{C}$  are removed with a smaller 2 ms veto (see Sec. 7.2.1). For the analysis of neutrons themselves, a DAQ gate is opened after every TT1BTB4 event, and assigned a special trigger type TT128 (Sec. 3.3.2). Neutron detection is also important for the threefold coincidence technique, aimed at tagging cosmogenic  $^{11}\text{C}$  events, which I will discuss in Sec. 4.5.3.

#### 4.5.3 Carbon isotope $^{11}\text{C}$

The cosmogenic  $^{11}\text{C}$  atoms decay emitting positrons:

$$^{11}\text{C} \rightarrow ^{11}\text{B} + e^+ + \nu_e, \quad (4.8)$$

which constitute background for the measurement of *pep* and CNO neutrinos, as can be seen in Fig. 4.10.

This isotope is a product of muon spallation on  $^{12}\text{C}$  present in the LS:

$$\mu + ^{12}\text{C} \rightarrow \mu + ^{11}\text{C} + n. \quad (4.9)$$

It is not possible to tag  $^{11}\text{C}$  via coincidences with the parent muon due to the long mean lifetime of  $^{11}\text{C}$ ,  $\tau = 29.4 \text{ min}$ . For that reason, in the LER analysis we do not search for  $^{11}\text{C}$  events and remove them from the dataset, but rather split the exposure into two parts that have higher and lower expected amounts of  $^{11}\text{C}$ , respectively, using the **threefold coincidence algorithm** (TFC).

The TFC technique exploits the fact that  $^{11}\text{C}$  is often produced with a neutron or a burst of neutrons (see Eq. 4.9). The algorithm looks for coincidences of a muon and a neutron event

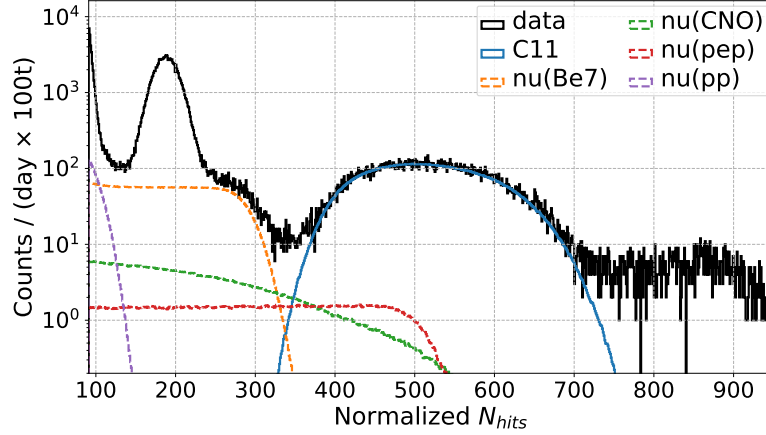


FIGURE 4.10: Spectral shape of the cosmogenic  $^{11}\text{C}$  (blue) compared to those of the solar neutrinos, refer to Fig. 4.3.

within 1.6 ms after the muon (five times the average time for neutron capture, see Eq. 4.7). After such a coincidence is found, a cylindrical volume of  $r = 3\text{ m}$  following the muon path is constructed, as demonstrated in Fig. 4.11A. Most of the positron events resulting from  $^{11}\text{C}$  decays are expected to happen inside that volume. The effect of  $^{11}\text{C}$  removal using the TFC technique is shown in Fig. 4.11B by the red energy spectrum.

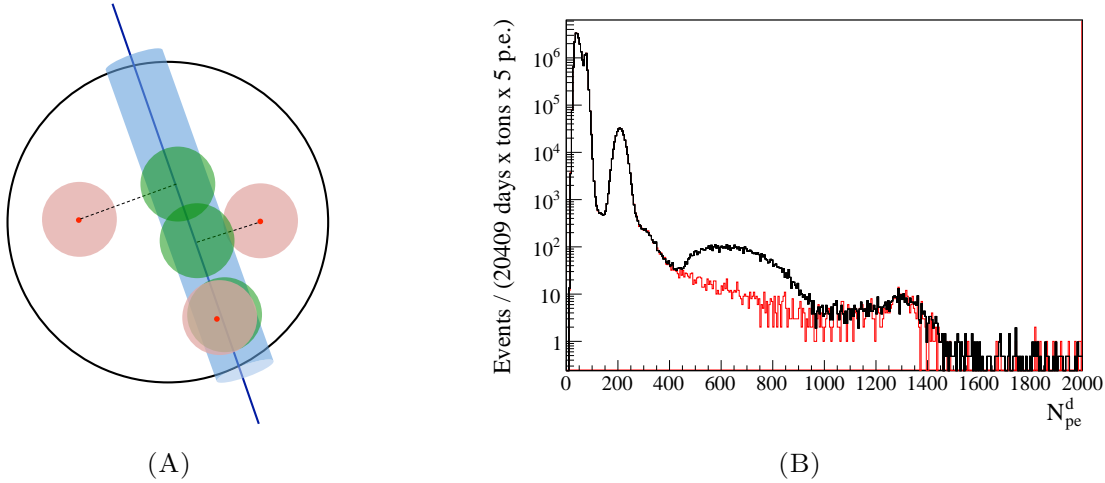


FIGURE 4.11: (A) Schematic representation of a muon track passing through the detector (blue line) and the region around it where high amounts of  $^{11}\text{C}$  are expected to be found (light blue cylinder). The figure shows examples of regions where  $\gamma$ 's following neutron captures are reconstructed (red), and their projections along the muon track (green). (B) Energy spectra after LER analysis selection cuts in the  $N_{pe}$  estimator before (black) and after (red) the application of the TFC selection (both normalized to the same exposure) [95].

In the LER solar neutrino analysis, the standard TFC method is not used to remove  $^{11}\text{C}$  events, but rather provides with the possibility to split the data sample into two classes: *TFC-tagged*, where most of the  $^{11}\text{C}$  events are expected to end up; and *TFC-subtracted*, with the rest of the exposure. Both spectra are then used in the multivariate fit, which will be described in Sec. 7.1.2. A version of the TFC method tuned to provide a  $^{11}\text{C}$  selection with the highest

purity is used to create so-called “strict samples” of  $^{11}\text{C}$  data. Such samples can be used for other analysis purposes, for example, to tune the MC generated  $^{11}\text{C}$  reference shape, as will be mentioned in Sec. 8.6.2.

In the improved TFC method,  $^{11}\text{C}$  tagging is based on a new algorithm that evaluates the likelihood  $\mathcal{L}_{\text{TFC}}$  that an event is a  $^{11}\text{C}$  candidate, rather than simply selecting all events in the cylinder as  $^{11}\text{C}$ . The algorithm considers multiple relevant parameters such as the distance of the candidate to the parent muon in space and time, distance from the coincident neutron, neutron multiplicity, and muon  $dE/dx$ . As can be seen from Fig. 4.12, the majority of the  $\log\mathcal{L}_{\text{TFC}}$  distribution for  $^{11}\text{C}$  is above the threshold used to tag candidates, marked by the green line.

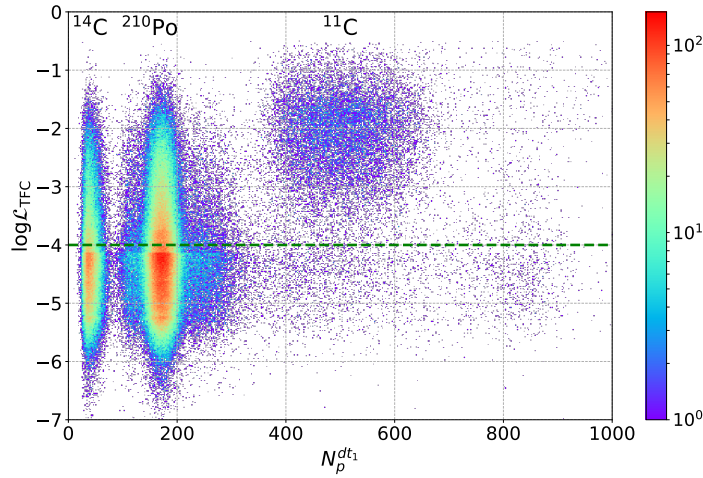


FIGURE 4.12: Distribution of events as a function of  $\log\mathcal{L}_{\text{TFC}}$  and  $N_p^{dt1}$ . The green line represents the  $\log\mathcal{L}_{\text{TFC}}$  threshold.

## 4.6 Event pileup

A so-called **event pileup** happens when several scintillation events occur too close in time (few hundreds of ns) within the same DAQ gate for the clustering algorithm of the Echidna software to distinguish them as separate clusters. The probability of three or more events happening so close in time is extremely small, thus it is safe to consider only the contribution coming from pileup of two events. Such pileup events are dominated by coincidences of two  $^{14}\text{C}$  events, as  $^{14}\text{C}$  is the most abundant background isotope (Sec. 4.4.1).

This fact can be used to estimate the rate of pileup events from the  $^{14}\text{C}$ - $^{14}\text{C}$  contribution. Considering that the rate of  $^{14}\text{C}$  is  $\sim 40$  Bq/100 t [89], and the mass of the LS in the IV is  $\sim 280$  t, this yields the total  $^{14}\text{C}$  rate as  $r_c = 112$  Bq. Based on the efficiency of Echidna clustering, given the time window of overlap to be  $t = 230$  ns (which corresponds also to the  $N_p^{dt1}$  variable), the resulting rate of pileup is  $r_p = r_c^2 t = 0.00289$  Bq  $\approx 250$  cpd. Finally, this gives us a rate of  $\sim 90$  cpd/100 t in the IV. Considering that the rate of  $pp$  neutrinos is around 130 cpd/100 t [72], and their spectral shape is very similar to the convolution of  $^{14}\text{C}$  shape with itself, pileup presents the most crucial background for the measurement of  $pp$  neutrinos. Two other main contributions to pileup are  $^{14}\text{C}$  with external background (EB), or  $^{14}\text{C}$ -EB; and  $^{210}\text{Po}$ -EB.

We employ three methods of modelling the effect of pileup on the energy spectrum, two of which are based on data and are applied to the spectral shapes constructed using the analytical approach; and one method based on the MC simulation, which will be described in sections Sec. 4.6.1 and Sec. 4.6.2, respectively.

### 4.6.1 Data driven methods

#### Convolution with untriggered spectrum

In order to model the effect of pileup, the spectral shapes of each background and signal component of interest, constructed based on the analytical approach, are convoluted with the spectrum resulting from untriggered data, collected exploiting the TT64 events (Sec. 3.3.2), which is demonstrated in Fig. 4.14A. This convolution distorts the original spectra, effectively accounting for the case when the events from the component of interest get piled up with other events happening in the detector.

#### Synthetic pileup

The second approach is the so-called **synthetic pileup**, in which untriggered random data is used to construct the spectrum of pileup events, shown in Fig. 4.14B. This approach consists in analyzing TT1 events, and overlapping the hits from the last  $6\mu s$  with the beginning of the DAQ gate, as demonstrated in Fig. 4.13.

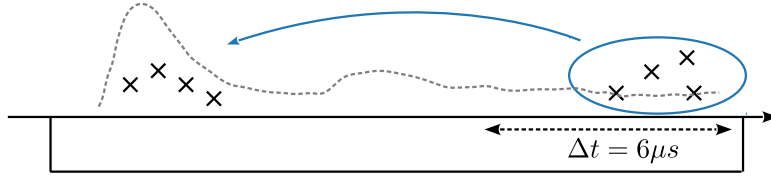


FIGURE 4.13: Schematic representation of the synthetic pileup approach. The grey dashed line represents the probability distribution of hits from a TT1 event in a DAQ gate.

This is done with the consideration that the probability of pileup, i.e. a second cluster falling into the beginning of the DAQ gate, close to the first cluster, is the same as a second cluster falling into the last part of the DAQ gate. Therefore, when adding the hits from the end of the DAQ gate to its beginning, if the resulting total hits overcome a certain threshold  $E_{min}$  that accounts for dark noise level, the created *synthetic* event is considered to represent pileup.

Both methods are used in the multivariate fit of the LER solar neutrino analysis, which will be discussed in Sec. 7.1.2. Example of the fit where pileup is modeled via convolution is shown in Fig. 4.15A. Note the deformation of the reference shape for  $pp$  neutrinos seen in Fig. 4.15A, after the convolution method, compared to the original shape in Fig. 4.15B, which demonstrates the synthetic pileup approach (the spectrum included is after the solar neutrino selection cuts). In the latter, the contribution of pileup is treated as a separate background component (spectrum in dashed pink), rather than an effect on other spectral shapes.

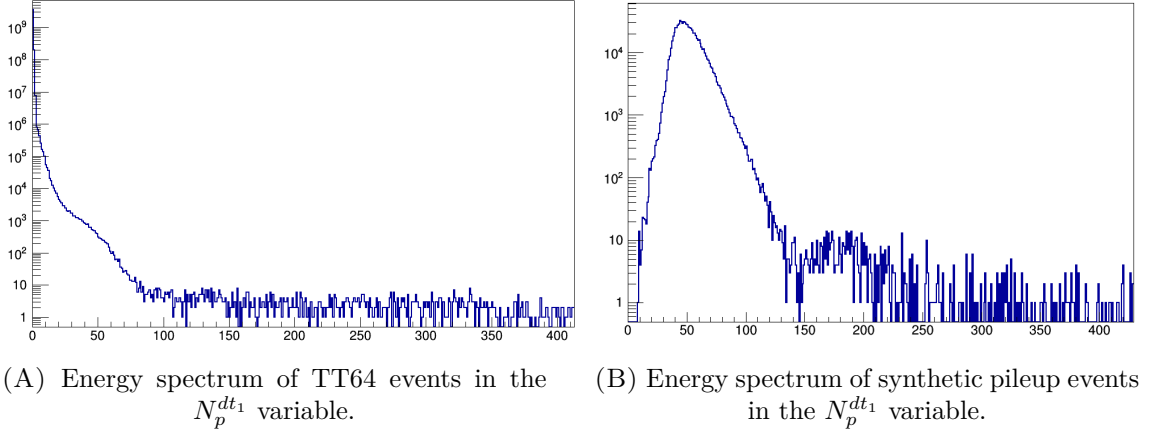


FIGURE 4.14

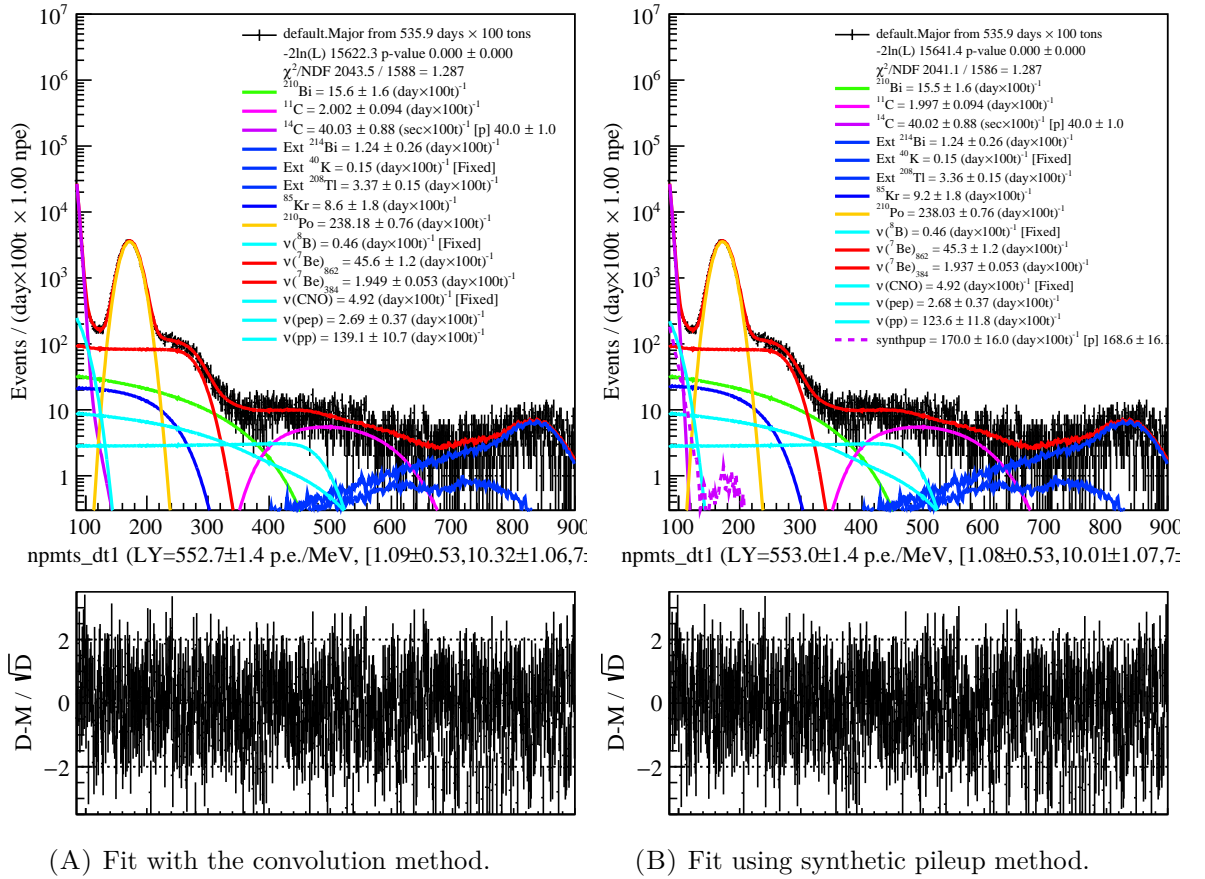


FIGURE 4.15: The effect of the convolution with the spectrum shown in Fig. 4.14A can be best seen through the deformation the  $^{14}\text{C}$  spectral shape (purple) in (A), compared to the one in (B), where synthetic pileup is shown by the purple dashed curve.

### 4.6.2 Monte Carlo simulation of pileup

In the MC approach, pileup events are constructed by overlapping g4bx2 generated physical events, to simulate exactly what happens in real data. First, sets of  $^{14}\text{C}$ , EB, and  $^{210}\text{Po}$  events are generated with g4bx2. Next, for each combination,  $^{14}\text{C}$ - $^{14}\text{C}$ ,  $^{14}\text{C}$ -EB, and  $^{210}\text{Po}$ -EB, a g4bx2 event for the first component is taken, and one for the second component is added to it with a randomly picked  $\Delta t$  up to  $2\mu\text{s}$  relative to the start of the first cluster. This is shown schematically in Fig. 4.16A.

Then, the resulting g4bx2 events are processed through the electronics simulation bx\_elec (Sec. 3.5.2), and Echidna. If Echidna fails to distinguish two clusters in the g4bx2 events, these events are considered to be pileup. The resulting spectral shapes for  $^{14}\text{C}$ - $^{14}\text{C}$ ,  $^{14}\text{C}$ -EB, and  $^{210}\text{Po}$ -EB are shown in Fig. 4.16B.

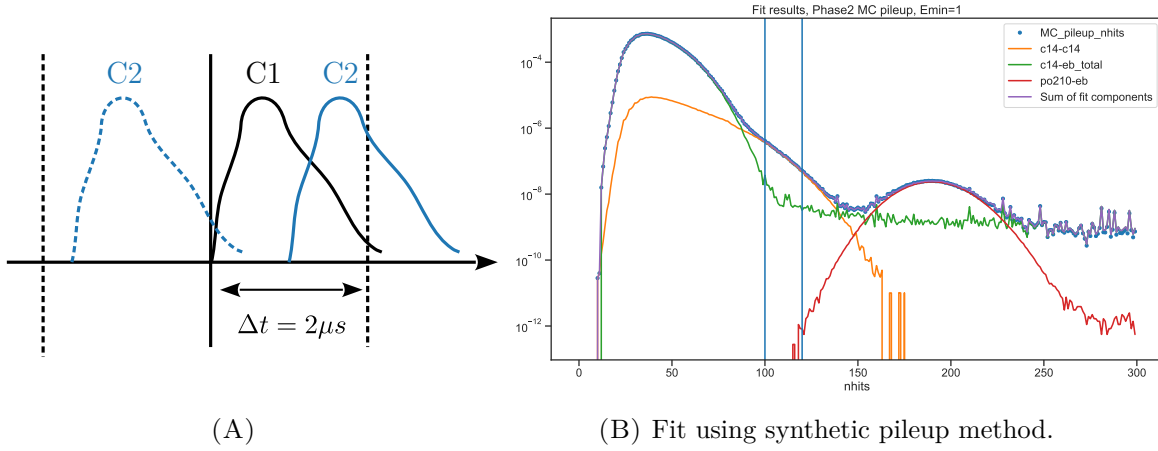


FIGURE 4.16: (A) Schematic representation of the MC approach to pileup. The first g4bx2 event, C1 (black), and the second one, C2 (blue), are overlapped with a  $\Delta t$  up to  $2\mu\text{s}$ . (B) Fit of major pileup components to the synthetic pileup spectrum (*credit: Alexandre Göttel*).

In order to construct the total pileup spectrum, based on these three major contributions, their shapes are fit to that of synthetic pileup to determine the relative weights of each shape, as shown in Fig. 4.16B.



## Chapter 5

# Detector stability and data quality

Before performing analysis on data, one needs to make sure that the data does not contain any false unphysical information and guarantee it is reliable. To assure data quality, we monitor the behavior of the detector and its various electronics parts (described in Sec. 3.3), and reject data taken while the data acquisition (DAQ) system (Sec. 3.3.1) was not working correctly, or when detector electronics was faulty. To achieve this, DAQ monitoring and **data validation** procedures exist in Borexino, which I will talk about in Sec. 5.1 and describe my work related to data quality.

As described in Sec. 3.4.2, the raw data which we collect during DAQ is analyzed by `Echidna` to produce physical parameters like time, position, and energy of the event, etc. One of the important procedures to guarantee correct physical interpretation of raw data is the **laser PMT timing calibration**, mentioned in Sec. 3.4.1. In Sec. 5.2 I will explain the importance of this procedure, and describe the work I have done to assure the calibration works correctly.

Lastly, working on detector stability has brought me to study the Borexino PMTs. In Sec. 5.3 I will look into **behavior of the PMTs in time**, and several features related to their quality that I have discovered in my work.

### 5.1 DAQ monitoring and data validation

The quality and stability of the Borexino data taking relies on the function of the the trigger system, the inner and outer detector (ID and OD, respectively), the PMTs and their corresponding digital Laben channels, etc. The so-called **RunValidation** C++/ROOT framework exists in order to test the work of all these components by judging the performance of the electronics and hardware directly through specifically designed procedures, and indirectly through different physical quantities. As discussed in Sec. 3.3.2 and Sec. 3.4.4, we use the so-called **service triggers** to check the performance of the PMTs and Laben channels. At the beginning of each DAQ run, short precalibration is performed (discussed in Sec. ??), which is used to align channels in time, and possibly disable the ones for which this alignment did not work. We can also exploit this data in RunValidation for monitoring purposes. In addition to that, we use well known internal backgrounds and muons (refer to Chapter 4) to check the behavior of the detector and its stability, since its response to such events should not change in time. During each week there is a person (called the **shifter**) responsible for monitoring DAQ and validating the data using the RunValidation procedure to accept or reject each run. In these checks, we always look at different **trigger types (TTs)**, which were summarized in Table 3.1, and discussed in Sec. 3.3.2.

Below is a short summary of the key parameters that are checked:

- **Stability check:** compare current run to previous runs
  1. The rate of external and internal muons, dark noise
  2. Number of live PMTs
  3. Event rates in the energy regions corresponding to major internal backgrounds ( $^{14}\text{C}$ ,  $^{210}\text{Po}$ ,  $^{11}\text{C}$ ) in IV and FV
  4.  $^{222}\text{Rn}$  events in IV and FV
  5. Cluster start time relative to the issued trigger
- **Precalibration (time alignment):** charge and time response in Laben channels
- **Calibration (special channels):**
  1. Trigger reference channels
  2. Laser reference channels
- **Main validation procedure**
  1. Inner and Outer Detector
    - Event size (decoded hits) of different TTs in each event
    - Event size of different TTs in each channel
    - Trigger rates
    - Hit time distribution of different TTs
    - Muon rate and hit distribution
  2. Inner detector
    - Cluster start time relative to the issued trigger
    - Number of clusters and hits in TT1BTB0 events
    - TT128 (neutron trigger) behavior

The procedure is organized into **canvases**, where each canvas displays one or more plots related to a certain topic. One example of a canvas is shown in Fig. 5.1, where one can see the the distribution of hit times relative to trigger time of events of various TTs. The job of the shifter is to analyze the canvas and note if there are deviations from normal behavior, or unusual structures. There are 4 canvases related to stability in time, 3 related to calibration and precalibration, and 17 related to the main DAQ check.

My work related to this framework consists of the following.

1. **Debugging** and improvement of the code (Sec. 5.1.1).
2. Introduction of **new quality checks** (Sec. 5.1.2).
3. Creating a comprehensive **manual** for shifters.

All the canvases are shown and explained in detail in App. A, where I attach the aforementioned manual for shifters. Each page shows an example of one canvas, a detailed description of what is shown in each plot, why it is important to make a check using the canvas information, and what the shifter needs look out for, as well as several examples of bad performance cases.

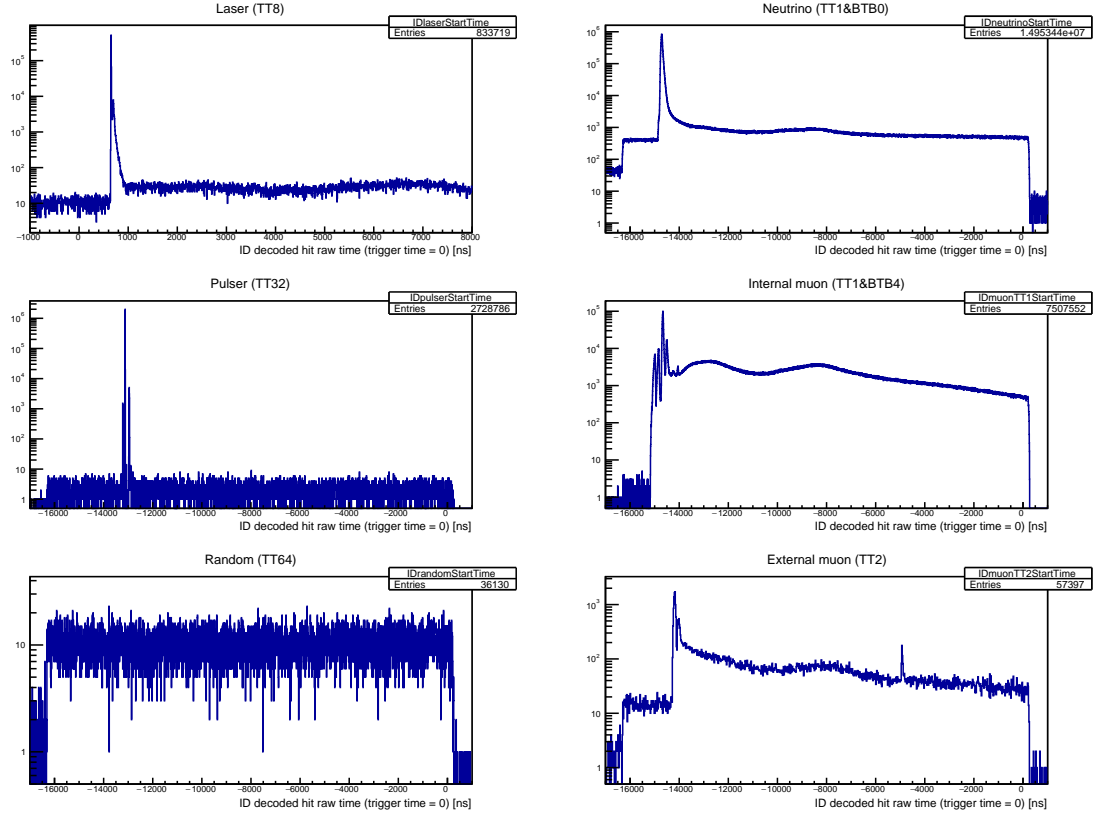


FIGURE 5.1: RunValidation *canvas 05*: Hit time distribution of events of different TTs in the ID. See more on p. 161.

### 5.1.1 Debugging

One example of an important bug fix is demonstrated in Fig. 5.2, where rate of muon events in the ID is calculated.

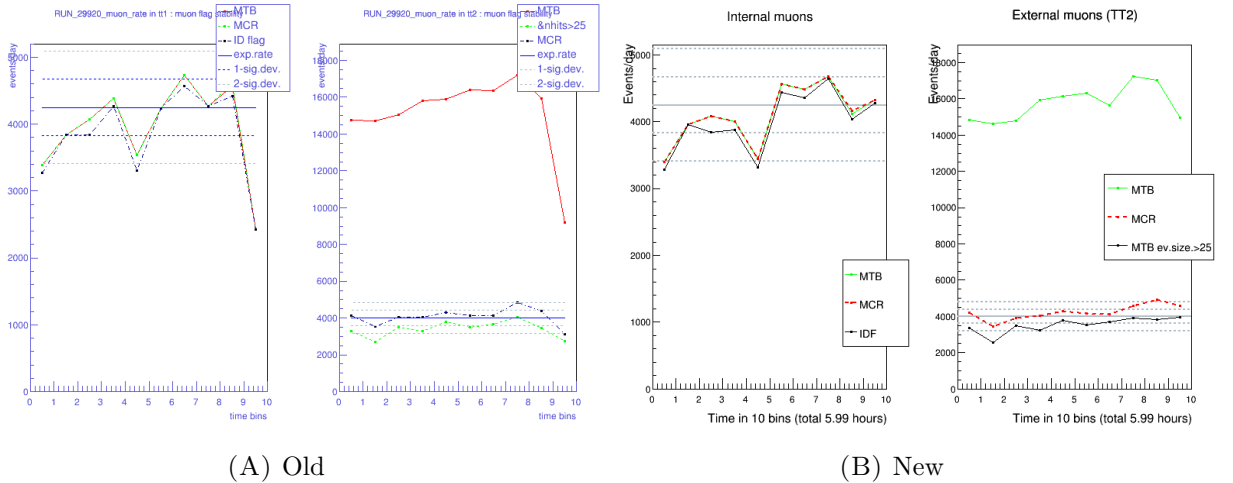


FIGURE 5.2: Rates of muon events, calculated the with old and new code. Left and right plots in each subfigure show the rates of internal and external muons, respectively. Note the erroneous drop in the rates in (A).

In the old RunValidation code, the calculation of the muon rate in the ID was wrong, and the plot always showed a large ( $> 2\sigma$ ) drop at the end of each DAQ run, as seen in Fig. 5.2A; while in the new code this bug is fixed, which allows correct monitoring of the rate shown in Fig. 5.2B. These plots belong to *canvas 08* of the Run Validation procedure, and are explained in more detail on p. 164 in App. A. The same bug has been fixed in the calculation of the OD muon rate (*canvas 11*, p. 167) and TT1 trigger rate (*canvas 03*, p. 159).

### 5.1.2 New quality checks

After becoming deeply familiar with the DAQ procedure and the old RunValidation framework, I have proceeded to add several new canvases and plots, introducing new quality checks that have not been present in the old procedure. For example, *canvas 0*, demonstrated in Fig. 5.3, is aimed at checking the performance of the precalibration at the beginning of each DAQ run, aimed at coarse 50-ns time alignment of Laben channels.

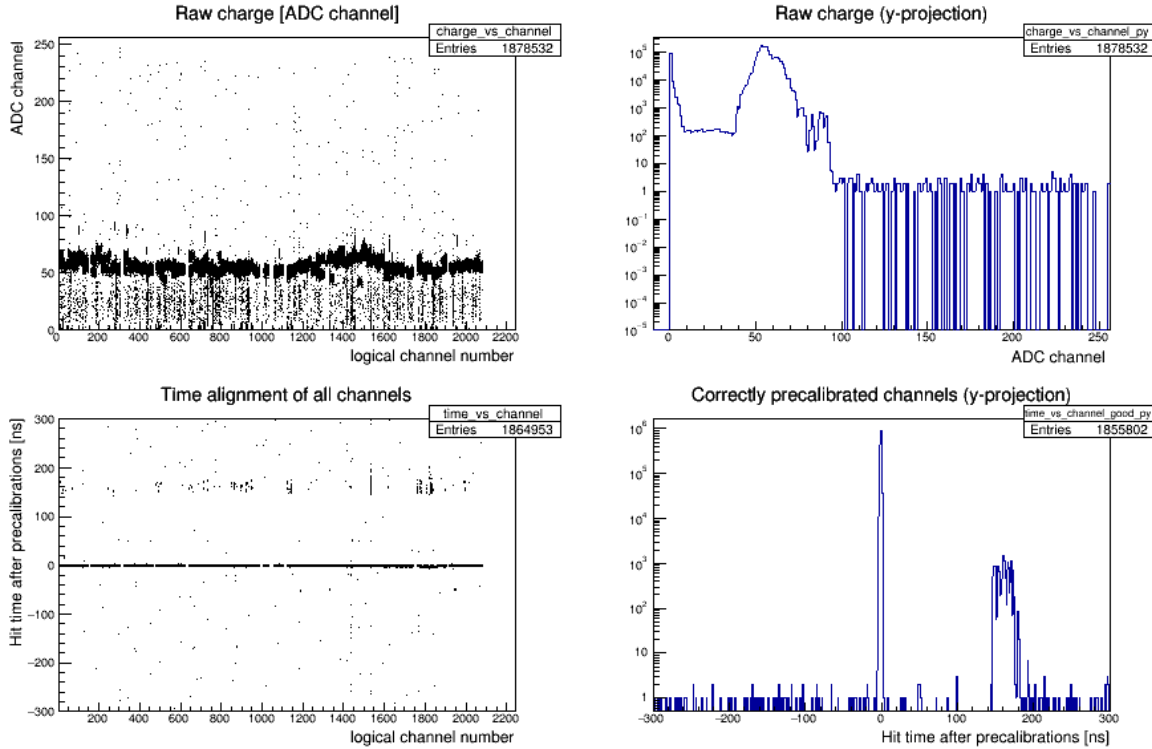


FIGURE 5.3: *Canvas 0*: precalibration.

The detailed description of this canvas and the meaning of each plot can be found on p. 155 of App. A in the attached manual for shifters. All the new canvases and plots introduced by me are marked as “**new**” in the manual.

## 5.2 Laser PMT timing calibration

As described in Sec. 3.4.1, **laser PMT timing calibration** (LC for short) is performed on a weekly basis with the purpose of time alignment of Laben channels. It utilizes the laser calibration system, introduced in Sec. 3.2.4. Some of my work related to LC includes improvement in the calculation of the channel **time offset** used for the time alignment, which I will elaborate on in Sec. 5.2.1.

Complementary to the RunValidation framework, discussed in Sec. 5.1, I have designed and implemented a novel **LaserValidation** procedure, aimed at monitoring the performance and stability of the laser system used in LC. It is composed of three canvases designed to check three important parameters:

- **goodness of channel alignment** after the LC (Sec. 5.2.2);
- **channel peak deformation**, which might result from bad clock signal pickup (Sec. 5.2.3);
- **laser intensity** (Sec. 5.2.4).

I will describe each check in more detail in the corresponding subsections below.

### 5.2.1 Channel time offset

For more information on the LC procedure, see Sec. 3.4.1. In order to align the channels in time, we analyze the peaks resulting from the laser pulses in each channel. The peak position of each Laben channel or *logical channel* is then corrected by a time offset  $\Delta T_{lg}$  relative to the mean of the distribution all channels. Figure 5.4 demonstrates the distribution of hits in each channel before and after they are aligned by applying the  $\Delta T_{lg}$  correction.

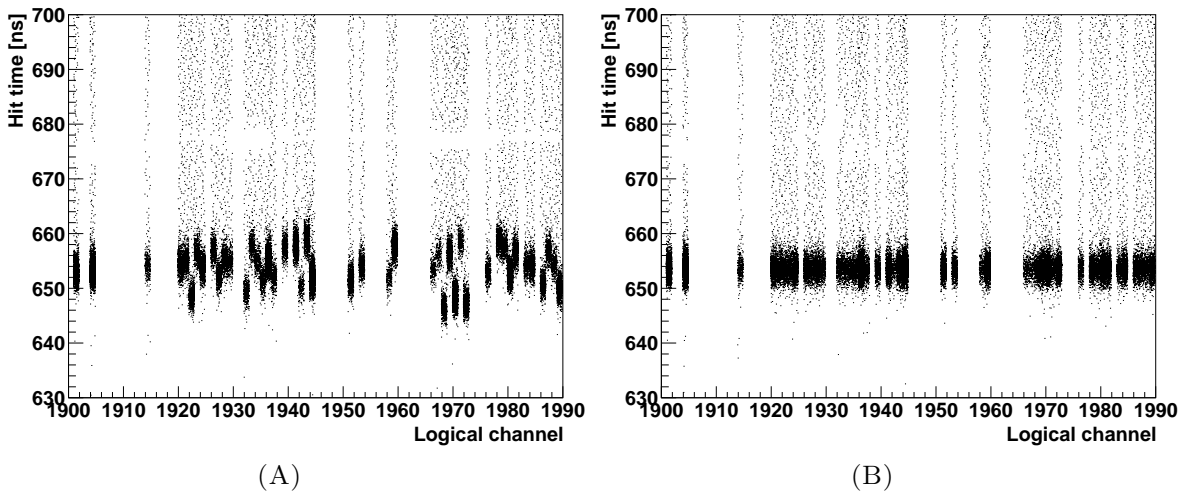


FIGURE 5.4: Distribution of hit times in response to a laser signal in each logical channel before (A) and after (B) alignment, with respect to laser reference time.

The offset  $\Delta T_{lg}$  is calculated as follows:

$$\Delta T_{lg} = T_{lg} - T_{all}, \quad (5.1)$$

where  $T_{lg}$  is the position of the peak in the logical channel  $lg$ , and  $T_{all}$  is the mean value of all hit times in all channels.

Using the old method,  $T_{all}$  was calculated as the bin center of the bin with the highest content (“maximum bin”), while  $T_{lg}$  was extracted from a Gauss fit of the peak of the given channel. However, my studies have shown that the peak does not exhibit Gaussian behavior, especially after the installation of a new trigger system, which reduced previously present Gaussian fluctuations. This makes a Gaus fit not suitable for the determination of peak position. Furthermore, since the statistics in one channel is not so large, in cases of fluctuations, the fit performs poorly.

I have studied several methods including asymmetrical Gauss fit, different binning and dithering, and arrived at the conclusion that a simple mean is actually superior to more complex methods. Using the maximum bin, or the simple mean for  $T_{all}$  gives compatible results due to high statistics. Figure 5.5 demonstrates the improvement brought by the usage of simple mean (orange) for  $T_{lg}$  compared to the Gauss fit (blue) using one logical channel as an example.

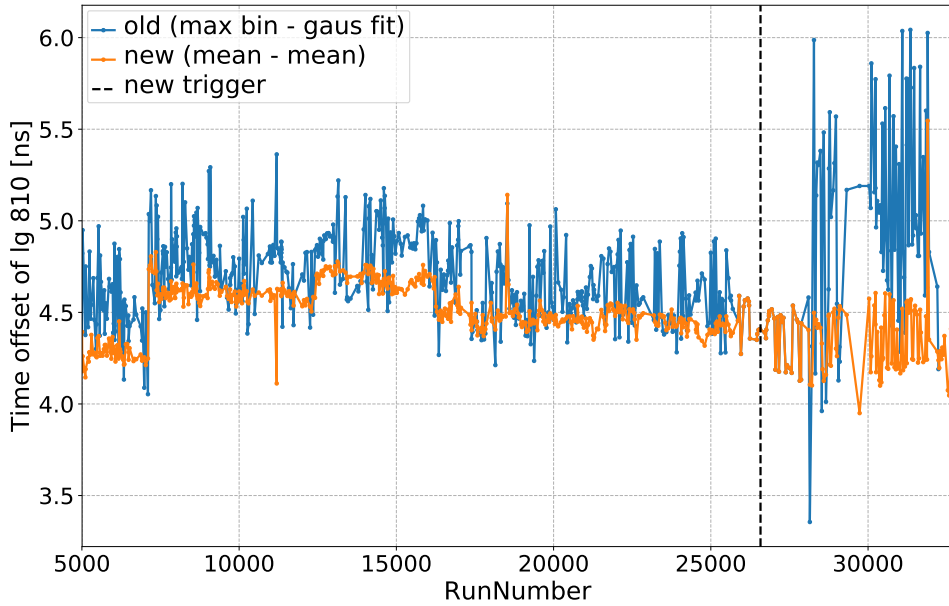


FIGURE 5.5: Channel offset  $\Delta T_{lg}$  as in Eq. 5.1 for one Laben channel, obtained with the old (blue) and new (orange) method, in the time span from around May 2007 (run 5000) to February 2019 (run 32238). The installation of the new trigger system is marked by the black dashed line.

### 5.2.2 Goodness of alignment

The effect of the time alignment of the Laben channels is demonstrated in Fig. 5.6, which shows the projection of the hits in all channels shown previously in Fig. 5.4. As can be seen, the resulting total peak after the alignment (blue) is considerably narrower than before (black).

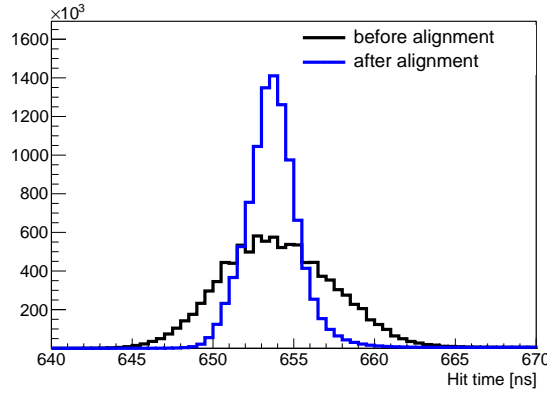


FIGURE 5.6: Total projection of hits from all channels before alignment (Fig. 5.4A, blue) and after alignment (Fig. 5.4B, black).

To quantify the width of the total peak, shown in Fig. 5.6 in blue, we obtain  $\sigma$  from a Gauss fit, as well as an *asymmetrical Gauss* fit ( $\sigma_L$  and  $\sigma_R$ ), using the following function:

$$f_{A-Gauss}(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{avg}} e^{-(x-\mu)^2/2\sigma_L^2} & \text{for } x \leq \mu \\ \frac{1}{\sqrt{2\pi}\sigma_{avg}} e^{-(x-\mu)^2/2\sigma_R^2} & \text{otherwise,} \end{cases} \quad (5.2)$$

where  $\sigma_{avg} = (\sigma_L + \sigma_R)/2$ .

Note that since I am using all channels to perform the fit, the values of  $\sigma$ ,  $\sigma_L$ , and  $\sigma_R$  represent the width of the total projection, and are influenced by two factors: successful alignment of channels, which depends on the procedure described in Sec. 5.2.1; and the  $\sigma$  of each individual channel, influenced by transit time jitter of laser reference channels. Therefore, monitoring this parameter allows us to see if there are any changes in either factor. Figure 5.7 demonstrates the time evolution of the three different peak width estimates from a standard Gauss fit, and the asymmetrical Gauss function, shown in Eq. 5.2. This figure constitutes **canvas 1** of the LaserValidation procedure.

As seen from Fig. 5.7, the implementation of the new trigger system (marked in by the blue dashed line) greatly reduced the aforementioned transit time jitter, which results in not only smaller, but also more stable peak spread in each Laben channel. In addition, the peak width can be influenced by different maintenance procedures. For example, several so-called *crate recabling* procedures, during which some PMTs have been reassigned to different Laben channels, disturbed the stability of the peak spread. With the implementation of this canvas, the shifter who performs LaserValidation after the weekly LC can notify the collaboration about the changes they see, and the reason and the impact of this change can be investigated.

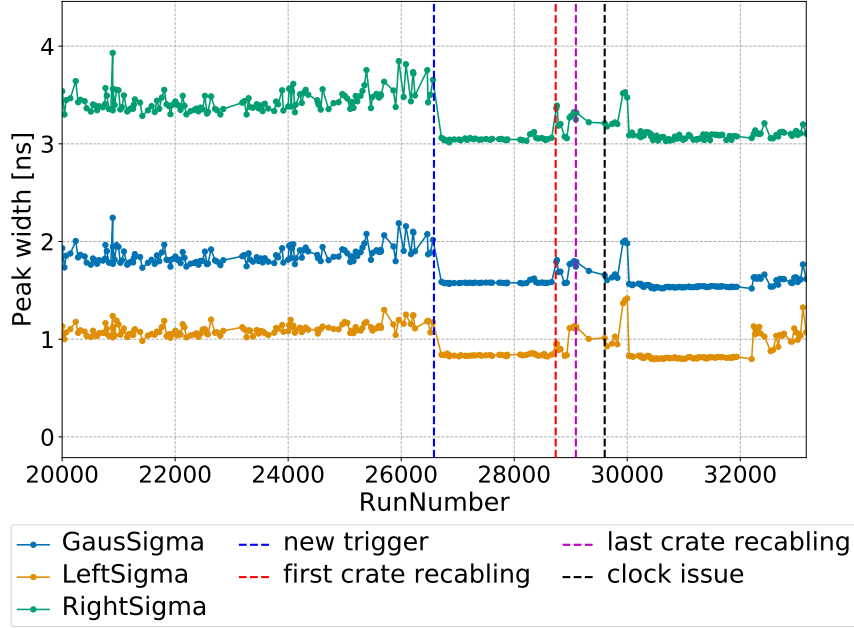


FIGURE 5.7: Width of the total peak after channel alignment calculated as  $\sigma$  from a Gaus fit (blue) and as  $\sigma_L$  (orange) and  $\sigma_R$  (green) from an asymmetrical Gaus fit.

### 5.2.3 Single channel peak deformation

The second check of the LaserValidation procedure is focused on a specific type of deformation of the peaks in channels, the so-called “**split channels**”. This type of deformation was first discovered in September 2018 after an issue with the GPS signal pickup. Due to the nature of the clock signal, based on triangular and square waveforms with a 50-ns period, hits arriving in a certain range could not be registered and were lost. This issue is shown in Fig. 5.8C where one can see an empty band around 648-650 ns, not seen normally in a good LC run shown in Fig. 5.8A.

As a result, the calculation of the time offsets of each channel, described in Sec. 5.2.1, did not work correctly, and the channels were not aligned correctly, as can be seen from Fig. 5.8D, compared to the good behavior in Fig. 5.8B. In order to find which LC runs have been affected by the timing issue, I have designed an algorithm that looks for split channels and distinguishes them from ordinary unaffected channels. The resulting efficiency of the algorithm is satisfactory (more than 90%), while around five good channels per normal LC run are falsely tagged as split, which is sufficient for monitoring the channels to prevent issues. Figure 5.9 shows the number of split channels in each LC run.

The unacceptably high amounts of split channels are marked by red points. There were in total eight LC runs heavily affected by the timing issue (marked with the black dashed line) which had to be removed from usage. After the clock issue was fixed, no abnormally high amount has been observed. Now this is a standard weekly check which belongs to **canvas 2** of LaserValidation.



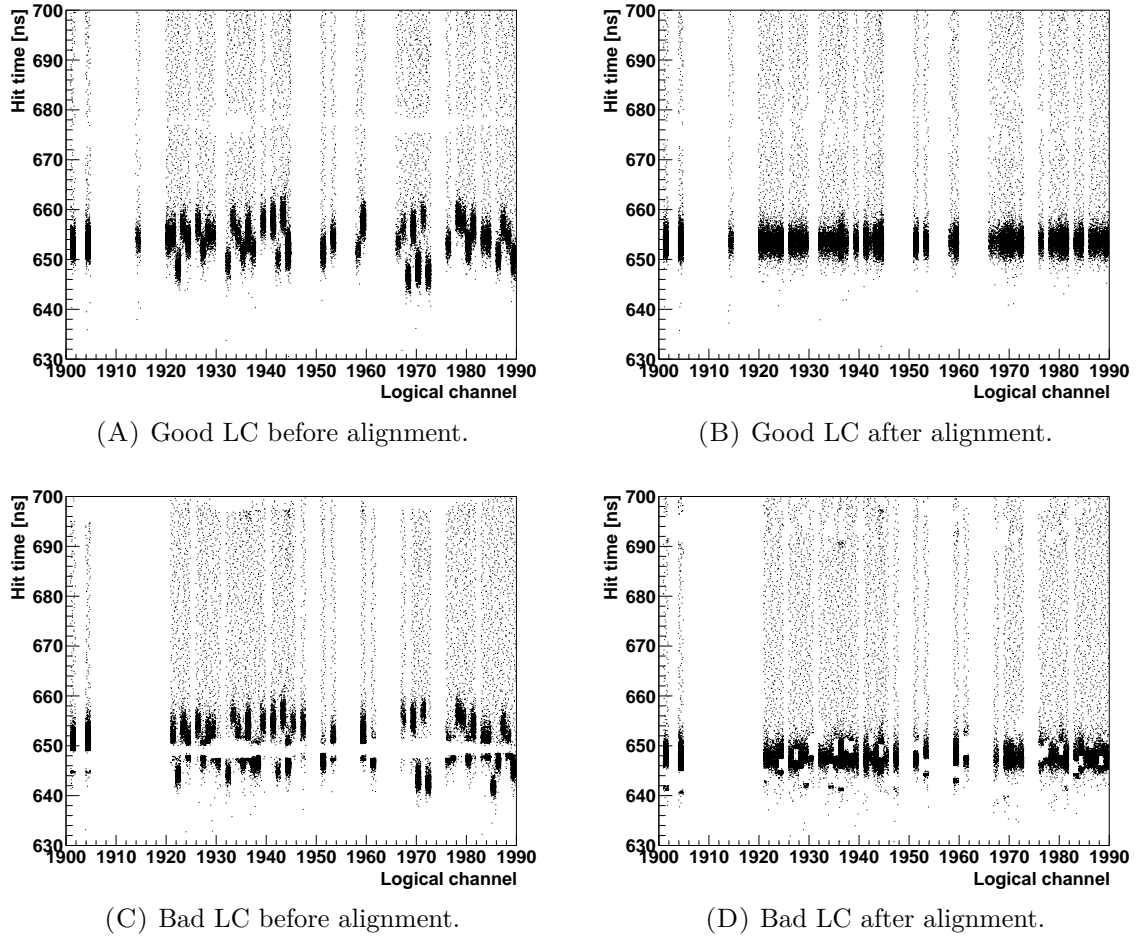


FIGURE 5.8: Distribution of hit times in response to a laser signal in each channel before (left) and after (right) alignment in a good (top) and bad (bottom) run.

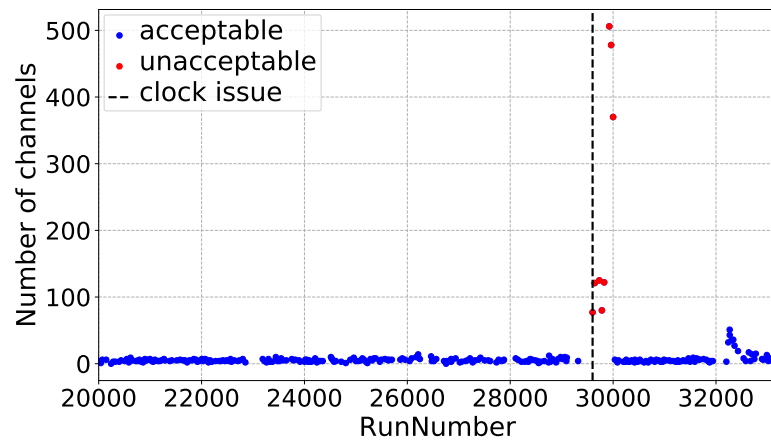


FIGURE 5.9: Number of split channels in each LC run. Unacceptably high values are marked red. The black dashed line denotes the time when the timing issues started happening.

### 5.2.4 Laser intensity

The last check belonging to LaserValidation is that of laser intensity. The intensity of the laser determines how much statistics will be collected during LS run. We are interested in keeping it sufficiently high, as well as in monitoring the quality and aging of the laser itself. I calculate laser intensity in two ways. The first way is to simply take the total hits in all channels after the LC run  $N_{total}$ , normalizing by the number of live channels  $N_{live}$  and the total number of TT8 (i.e. laser) events during the LC run  $N_{TT8}$ :

$$I_1 = \frac{N_{total}}{N_{live}N_{TT8}}. \quad (5.3)$$

The intensity calculated this way is shown in Fig. 5.10 by the blue points. The second way is to only use the hits in the main peak coming from laser events, shown in Fig. 5.6. For this, a Gauss fit is performed, and the entries in the peak are calculated as the area of the Gauss distribution:

$$I_2 = \frac{A\sigma\sqrt{2\pi}}{N_{live}N_{TT8}}, \quad (5.4)$$

where  $A$  (normalization) and  $\sigma$  are results of the fit. Intensity calculated this way is shown in Fig. 5.10 by the orange points, and displays the same trend as the one obtained with the first method. We use both methods as complementary to judge the status of the laser intensity.

Using these methods, I noticed that the intensity of the laser decreased noticeably compared to its original level. A decision has been made to manually increase the laser intensity, and this maintenance operation is shown in Fig. 5.10 by the black dashed line. As can be seen, the intensity was successfully brought back to its previous level. I implemented this check in the weekly LaserValidation procedure as **canvas 3**.

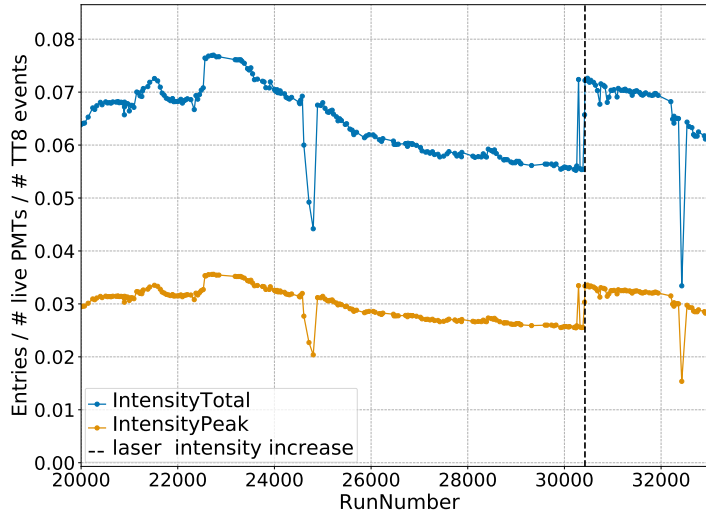


FIGURE 5.10: Intensity of the laser estimated using the total number of hits in all channels (blue), as in Eq. 5.3; and the number of hits in the peak (orange), as in Eq. 5.4. The black dashed line marks the time when the intensity was manually increased.

## 5.3 PMT quality

### 5.3.1 Correlation between PMT quality and livetime

One of the important discoveries resulting from my work is the phenomenon observed among the Borexino PMTs that consists of the three following properties.

1. PMTs that live longer tend to be of better quality.
2. PMTs that die tend to have been of worse quality from the very beginning.
3. PMTs that die tend to worsen before dying.

The first two properties effectively split the Borexino PMTs into two categories: the better quality PMTs ones that tend to live, and the worse quality ones that have a higher probability to die. This effect can be seen in Fig. 5.11, where several parameters related to PMT quality are shown. In order to demonstrate this finding, I look at the behavior of the set of PMTs that were live at the moment this study was conducted in July 2019, which consists of 1160 PMTs and which I will call **L2019**. At each time point, I separate the Borexino PMTs into two groups: L2019 (shown in green) and the rest of live PMTs (shown in red). Thus, the number of PMTs contributing to the green points is always roughly the same, while the ones in the red curve gradually decrease to zero, until the black and the green curves meet, since L2019 constitute all live PMTs in July 2019.

Figure 5.11A shows the average dark noise rate in each set of PMTs, exhibiting a clear trend mentioned above: L2019 show lower dark noise rate already from the very run 5000, corresponding to the very first data taken in 2007, while the rest of the PMTs have been worse from the beginning. The same picture is seen in Fig. 5.11B, where the sharpness of the PMT response to  $^{14}\text{C}$  events is shown looking at RMS relative to the mean of the charge distribution resulting from those events. In addition, one can see the effect of L2019 PMTs staying stable in time and the rest steadily worsening. Furthermore, a noticeable difference between L2019 and the rest of the PMTs emerged after the installation of the new trigger system, clearly seen in Fig. 5.11C. The new system resulted in the decrease of the laser timing calibration peak width (described earlier in Sec. 5.2.2), after which the difference between L2019 and non-L2019 PMTs emerged, showing visibly worse response of the latter compared to L2019.

Finally, the average *effective quantum efficiency*, shown in Fig. 5.11D, is higher and more stable among L2019 than that of the worsening non-L2019 PMTs. This feature plays a key role in calculating EQE, which I will discuss in Sec. 5.3.2 and more in detail in Chapter 6, where I focus on my work related to the procedure for obtaining EQE of Borexino PMTs needed for the Borexino Monte Carlo simulation.

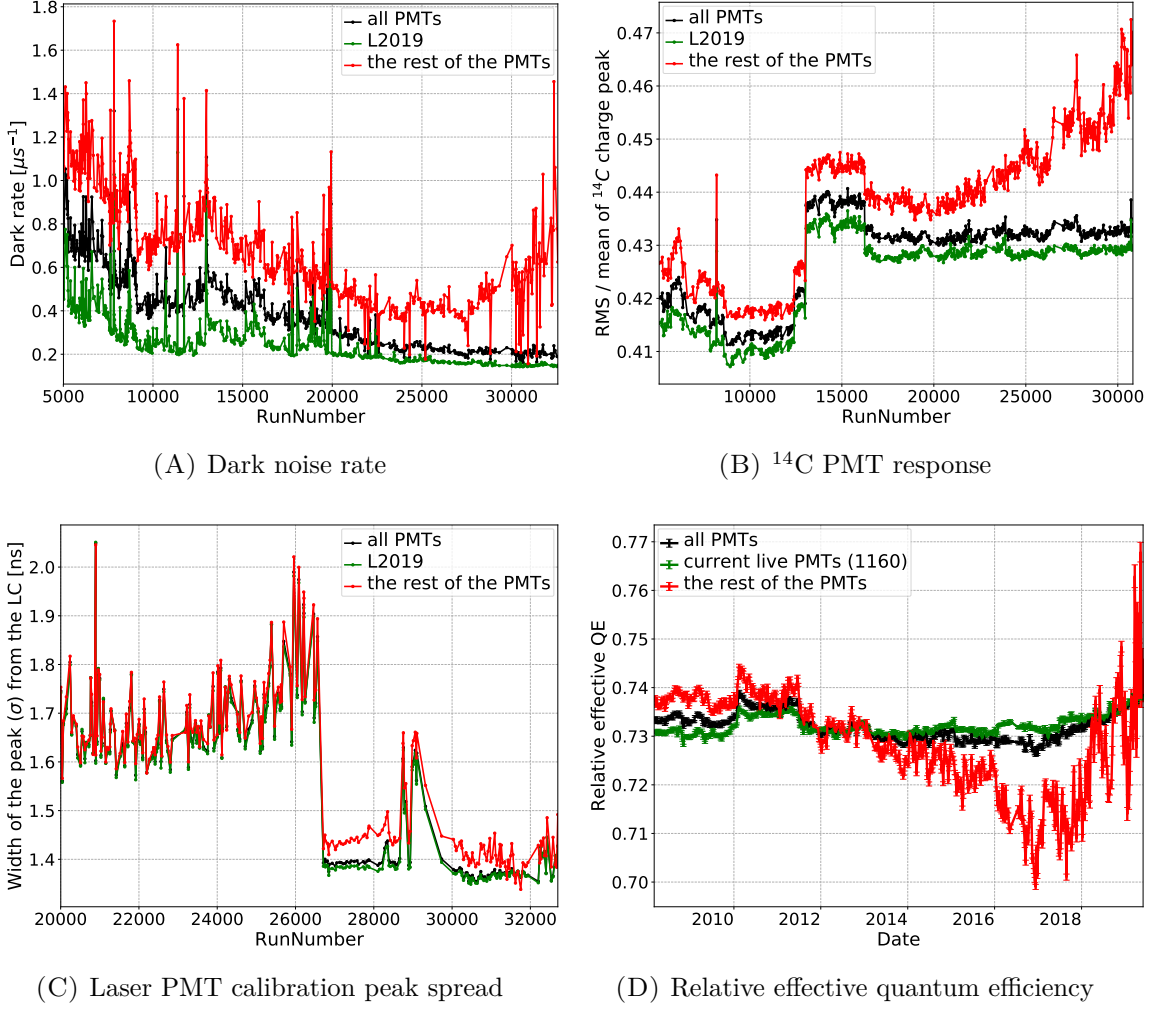


FIGURE 5.11: Each subfigure shows an average value of a certain parameter taken among all enabled PMTs at a given time point (black), enabled L2019 PMTs (green), and the remaining PMTs (red).

### 5.3.2 Stable PMT subsets

The discovery of the changing PMT quality and its correlation with PMT livetime (LT) has a crucial consequence for the calculation of effective quantum efficiency (EQE) of PMTs. We use  $^{14}\text{C}$  events selected from Borexino data for that calculation, and since the EQE of PMTs directly influences the energy of the selected  $^{14}\text{C}$  events that we use to judge the EQE of the PMTs, we end up with a bad recursive dependence sensitive to the changing PMT quality. To guarantee a stable selection of  $^{14}\text{C}$  events, and for other studies related to detector stability, I have studied different subsets of PMTs selected based on LT and parameters that represent quality.

Before my study, a selection of PMTs based on LT, dark noise (DN), and EQE has already been done on Phase-II. The resulting subset is called **A1000**, and was originally meant to be used for light yield studies [113]. This method of studying the detector using only a stable subset of PMTs has proven to be successful, and it has been proposed for other studies including EQE [117]. However, since A1000 is only based on Phase-II information, I have constructed subsets of

PMTs based on their behavior in Phase-I and Phase-II combined (time period May 2007 to July 2018), similarly to the selection of A1000 PMTs. These subsets are summarized in Table 5.1.

PMT set	Description	$LT_{min}$ (%)	Dark noise	N/S
<b>L2019</b>	Currently live PMTs (1160)	1.6e-3	0.004 - 24.5	1.52
<b>M1000</b>	Top 1000 by LT	94.5	0.05 - 2.23	1.65
<b>B1000</b>	Top 1000 PMTs by DN among those with $LT > 90\%$	90	0.05 - 0.36	1.49
<b>M900</b>	Top 900 by LT	96.5	0.05 - 2.13	1.72
<b>B900</b>	Top 900 by DN among M1000	94.5	0.05 - 0.39	1.61
<b>B800</b>	Top 800 by DN among M900	96.5	0.05 - 0.33	1.66
<b>M800</b>	Top 800 by LT	97.7	0.05 - 2.13	1.68

TABLE 5.1: Subsets of PMTs and their parameters. The third column reports minimum relative LT in Phase-I + Phase-II among the PMTs belonging to the given set. The last column reports the ratio of the number of PMTs in the north (N) and south (S) hemisphere of the detector.

The selection of a good subset of PMTs revolves around the trade-off between livetime/quality and resolution. The smaller the subset, the higher the livetime and the quality, since it correlates with livetime, as shown in Sec. 5.3.1 above. However, fewer PMTs mean larger statistical error. Additionally, due to the fact that more PMTs die in the south hemisphere of the detector, a stricter selection of PMTs results in a larger N/S asymmetry in terms of the number of PMTs, which can give a bias towards the north PMTs. Thus, as the final choice for the EQE procedure we have selected the **B900** subset, which gives a good balance between these factors. In the next chapter, Chapter 6, I will elaborate on how exactly B900 comes into play in the EQE calculation.

Energy estimators based on the B900 PMTs have been implemented into the Echidna framework (Sec. 3.4.2), enabling other studies based on this set of good stable PMTs, which I will discuss in the Outlook.



## Chapter 6

# Effective quantum efficiency of the Borexino photomultiplier tubes

In this chapter, I will talk about a parameter called **effective quantum efficiency**, which represents the light collection ability of a photomultiplier tube (PMT).

I will start this chapter by explaining the meaning of the terms *quantum efficiency* (QE), *effective quantum efficiency* (EQE), and *relative effective quantum efficiency* (REQE), and the role of these parameters in the Monte Carlo simulation of the Borexino detector (Sec. 6.1). Then, I will describe the improvements my work brought to the procedure of selecting the data necessary for the EQE calculation (Sec. 6.2), and the calculation procedure itself (Sec. 6.3). In Sec. 6.4, I will describe the new approach to weekly monitoring of EQE, first implemented after this work, which is related to statistics and PMT behavior. Finally, in Sec. 6.5, I will show the new improved method of converting EQE to REQE. After showing the results (Sec. 6.6) and examples of the validation of the new approach and its impact on the Borexino MC (Sec. 6.7), I will conclude this chapter with a summary of the differences between the old and the new approach, and improvements my work brought to the EQE calculation procedure and monitoring (Sec. 6.8).

### 6.1 Motivation and the general approach

The Borexino Monte Carlo (MC) simulation (Sec. 3.5) is employed in the solar neutrino analysis, which will be presented in Chapter 7 and Chapter 8, to construct spectral shapes of neutrino and background events. The events themselves are simulated using Geant4 based simulation `g4bx2`, and are later processed by `bx_elec`, the the Borexino electronics simulation. One of the inputs to the latter, is the so-called **effective quantum efficiency**, which represents the light conversion ability of the photomultiplier tubes (PMTs). This is an important parameter of `bx_elec`, as the energy spectra of the simulated events directly depend on the amount of light a PMT can register.

The **quantum efficiency** (QE) of a PMT is a measure of its ability to convert light to photoelectrons (p.e.). QE is calculated as the fraction of photons converted to p.e. relative to the number of incident photons on a PMT (more on PMTs in Sec. 3.2.3). This quantity is measured in the laboratory before the PMT is installed, however, it might decrease due to aging or worsening. After the PMT is installed in the detector, it is not possible to directly monitor the changes of its QE. We can, however, monitor the so-called *effective quantum efficiency* (EQE). Apart from the QE of the PMT, other effects, such as changes of the light yield (LY) of the liquid scintillator or tuning of the applied voltage, might influence the number of p.e. (thus the

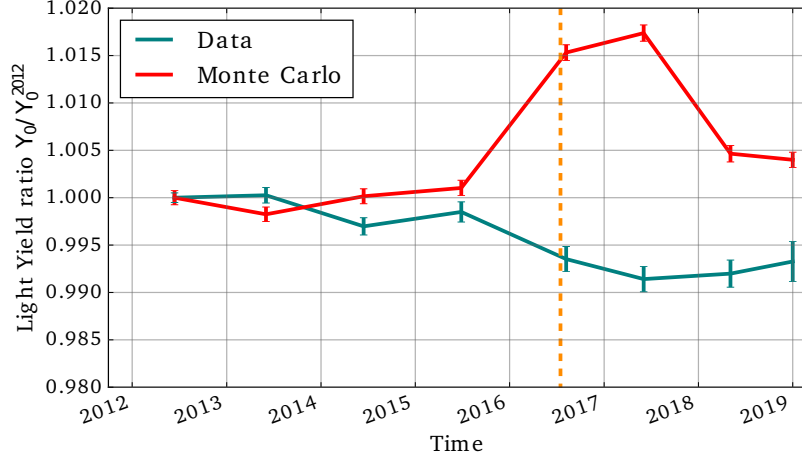


FIGURE 6.1: Relative changes of effective LY, calculated using  $^{210}\text{Po}$  data (green) and MC simulation with old REQE inputs (red) [101]. The beginning of Phase-III is marked with the orange dashed line.

term “effective”). Ultimately, this is the quantity we want to monitor to be able to simulate the detector and account for all the sources of change in the PMT light collection. EQE of a PMT  $p$ , which I will denote as  $Q(p)$ , is calculated as the fraction of hits the given PMT sees in response to the so-called *candle events*:

$$Q(p) = \frac{N_h^c(p)}{N_c}, \quad (6.1)$$

where  $N_h^c(p)$  is the total number of hits registered by PMT  $p$  coming from  $N_c$  candle events  $c$ .

A **candle event** must have low energy (to guarantee one p.e. per photon), a stable energy range (ideally monoenergetic) and equal probability to be seen by any PMT (ideally in the center of the detector). This way, changes in  $Q(p)$  directly represent effective properties of light collection described above, as the candle events do not change in time.

The final parameter used in the Borexino MC is the so-called **relative effective quantum efficiency** (REQE) which takes values from 0% to 100% by design of the simulation, and represents the EQE of a PMT relative to maximum possible QE, which has been measured to be 32%. The general requirements for the REQE calculation are:

1. A reliable source of candle events and an algorithm to select them from data (Sec. 6.2)
2. EQE calculation procedure using the selected events (Sec. 6.3)
3. Choice of statistics to guarantee good precision and monitor EQE changes in time (Sec. 6.4)
4. Conversion method of EQE to REQE (Sec. 6.5)

The motivation to look into the REQE calculation method was brought forth by the issue with the Monte Carlo approach (Sec. 7.1.3) to fitting the energy spectra constructed using Phase-III data. Hints that the issue has to do specifically with REQE inputs in the simulation come from the fact that independent studies on  $^{210}\text{Po}$  events conclude that effective LY is stable or



decreasing by no more than 1%, while the same analysis performed on MC simulation shows significant disagreement starting from mid 2015, as shown in Fig. 6.1 [101]. More on the details of the  $^{210}\text{Po}$  studies will be presented in Sec. 6.7.2. Further inspection has shown that, indeed, the method that was used to calculate REQE was not designed to monitor effective LY changes, which were small in Phase-II, but grew too large in Phase-III for MC to be accurate. My studies have concluded that the old approach needed changes in steps (1) and (4) above, which are now developed in the new approach.

In the following sections I will describe the details of the old and new approaches, and highlight the improvements and new features brought by my work.

## 6.2 Selection of candle events

### 6.2.1 Source of candle events

As mentioned in the previous section, an ideal candle event is monoenergetic and located in the center of the detector. The best candle event candidates in Borexino are **cosmogenic neutrons** (Sec. 4.5.2), since they are monoenergetic and uniformly distributed in the scintillator volume. Each neutron event will be equally seen by any PMT, provided a selection of such events in a sphere with a small radius in the center of the detector. However, neutrons do not happen frequently enough to provide sufficient statistics for an accurate calculation. I studied neutrons as candle events, and my results show that one needs to collect half a year of neutron statistics in order to achieve reasonable precision.

The second best option is  $^{14}\text{C}$ , which has a well known stable energy range, and is uniformly distributed in the detector. As seen from the spectrum demonstrated in Fig. 6.2,  $^{14}\text{C}$  is the most abundant isotope in the detector (see 4.4.1). It composes 99% of the Borexino data, and provides enough statistics to be able to monitor EQE changes on a weekly basis, as will be shown later in Sec. 6.4. The next step is to find a good method of selecting  $^{14}\text{C}$  events in a stable energy range.

### 6.2.2 Untriggered $^{14}\text{C}$ events

The best way to avoid the effect of a trigger threshold and the changing number of live PMTs, is to use the so-called “untriggered” events, i.e. the ones that appear as a second cluster of two-cluster events (mentioned in Sec. 4.4.1). I have performed studies on  $^{14}\text{C}$  events falling into the TT128 DAQ gate, as well as the “neutrino-like events” TT1 BTB0 C2 (more on event types in Sec. 3.4.4). However, these events do not yield enough statistics for regular monitoring of EQE that we require for the MC simulation to follow the changes of the PMTs. The resulting EQE is shown in Fig. 6.3. Same as with neutrons, sufficient statistics is only achieved when analyzing data samples on a yearly basis. The only choice left is to use events that trigger the detector, which is the standard approach of selecting  $^{14}\text{C}$  data.

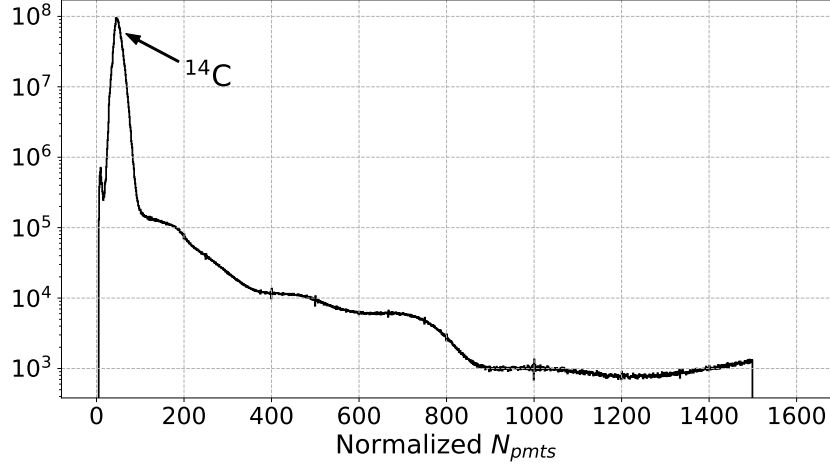


FIGURE 6.2: The spectrum of all Borexino TT1 data from Phase-II without any cuts, demonstrating the prominent  $^{14}\text{C}$  peak.

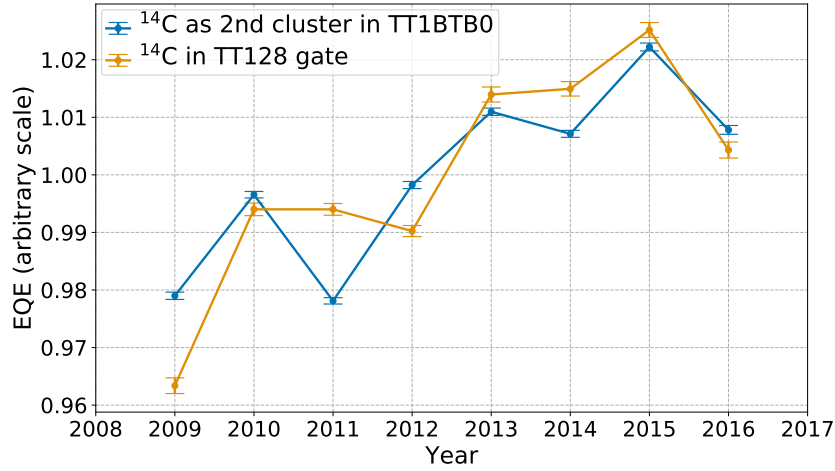


FIGURE 6.3: EQE scaled by the average over all years calculated using TT1 BTB0 C2  $^{14}\text{C}$  events (blue) and  $^{14}\text{C}$  falling into TT128 DAQ gate (orange).

### 6.2.3 Standard approach

The standard approach of selecting  $^{14}\text{C}$  events is to apply the following conditions:

- event type: internal one-cluster event (TT1 BTB0 C1);
- energy range:  $50 < N_h^{geo} < 100$ .

The variable  $N_h^{geo}$  is the *geonormalized* number of hits (see Sec. 3.4.3). Using a geonormalized variable ensures equal treatment of all PMTs even for events that are happening at a larger radius, as it accounts for the the solid angle of each PMT. However, this approach does not result in a selection of events with a constant energy range as demonstrated in Fig. 6.4 by the blue curve. The visible energy of the selected events is steadily increasing in time. This happens

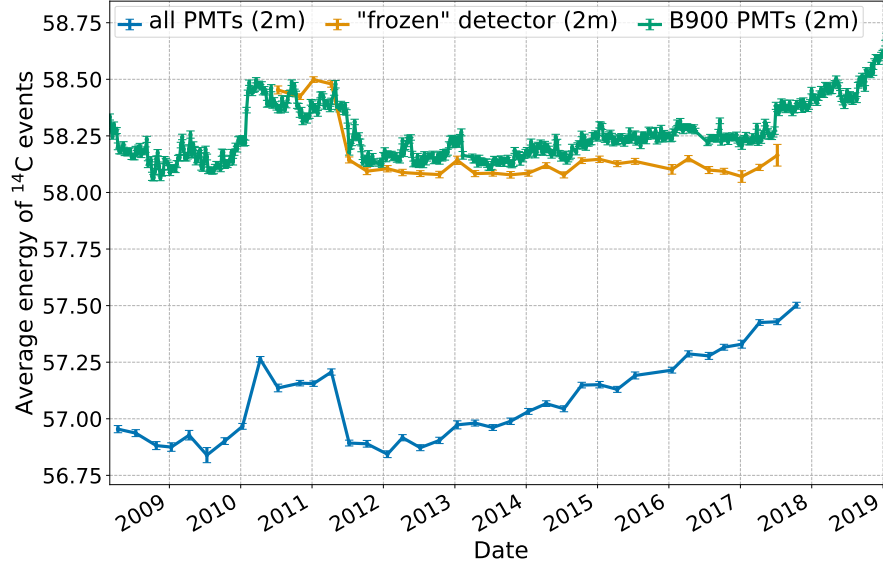


FIGURE 6.4: Average  $^{14}\text{C}$  energy resulting from selections based on  $N_h^{geo}$  (blue),  $N_h^{geo}$  in a “frozen” detector (orange), and  $N_h^{geo}$  (B900) in the full detector (green), measured in the corresponding energy estimator.

due to the fact that the sample of live PMTs in different time points has different quality, as the PMTs that die worsen before dying or are of worse quality to begin with; and PMTs that are of good quality tend to stay live longer. This phenomenon was described in detail in Sec. 5.3.1 and makes the task of selecting  $^{14}\text{C}$  nontrivial. Varying PMT quality directly affects the energy estimators used to select candle events that are supposed to represent the PMT quality itself, which results in a bad recursive dependence.

In order to test the hypothesis that the issue stems from this phenomenon, studies have been done on Phase-II data based on a “frozen” detector. The principle of the “frozen” detector approach is to “pretend” that in Phase-II only the stable subset of **A1000** PMTs, mentioned in Sec. 5.3.2, period exist in the detector. This is done simply by disabling the rest of the PMTs on the software level in Echidna (Sec. 3.4.2), so that energy and position reconstruction is done using only the A1000 PMTs. The orange curve in Fig. 6.4 shows the resulting average  $^{14}\text{C}$  energy with a subpercent stability as opposed to the steady 1.5% increase observed when using all PMTs. This study confirmed the hypothesis and inspired the following improved method of obtaining candle events correctly.

#### 6.2.4 Improved approach

Based on the above, I conclude that to correctly estimate the energy of  $^{14}\text{C}$  events, I must use a set of PMTs the number and quality of which does not vary significantly in time. The so-called B900 PMTs (Table 5.1) have been chosen for the  $^{14}\text{C}$  selection since among all other PMT subsets, the B900 set shows the best performance in terms of stability, dark noise, and detector top-bottom (referred also as north-south) symmetry. I obtain geonormalized hits in

B900 PMTs similarly to geonormalized hits in all PMTs as defined in Eq. 3.4:

$$N_h^{geo}(\text{B900})(c) = \frac{2000}{N_{\text{B900}}^{geo}(c)} N_x(c). \quad (6.2)$$

In this equation,  $N_{\text{B900}}^{geo}(c)$  is the geometrically weighted number of live B900 PMTs for each cluster  $c$ , obtained similarly to Eq. 3.5:

$$N_{\text{B900}}^{geo}(c) = \sum_{p \in \text{B900}} \alpha_p / 0.00298623, \quad (6.3)$$

where  $\alpha_p$  is the solid angle of the given PMT  $p$ , calculated as in Eq. 3.6, the sum going over each PMT  $p$  belonging to the B900 set. The results obtained by changing the energy condition to  $50 < N_h^{geo}(\text{B900}) < 100$ , demonstrated in Fig. 6.4 by the green curve, are compatible with the ideal “frozen detector” study (orange). However, the new results show a slight increase of 0.8% in the estimated  $^{14}\text{C}$  energy, not present in the is observed towards the later years. My hypothesis is that this is happening because of the same issue, the changing quality of the B900 set due to dying PMTs. To avoid this, a set of a smaller number of PMTs expected to live longer could have been chosen, but firstly, that would have worsened the precision, and secondly, such a selection would have been difficult due to the inability to predict when and which PMT would die. Note that despite the slightly imperfect stability, the B900 set has been selected for a larger time period (2007-2017) than A1000 (2012-2016).

### 6.3 EQE calculation

As mentioned previously, EQE is calculated as the fraction of hits seen by a PMT relative to the number of candle events that happened during a certain time window (Eq. 6.1). However, to obtain the hits  $N_h^c(p)$  coming from candle events is not entirely trivial, and requires three steps that involve specific corrections, which will be described in the subsections below.

Firstly, the number of hits registered by the logical Laben channels, to which the PMTs are connected, includes hits coming from **dark noise**, in addition to the ones coming from the candle events themselves. These hits need to be subtracted, resulting in the first stage of EQE calculation, which I will call  $Q^{\text{I}}$  (Sec. 6.3.1).

Secondly, some PMTs register more hits due to having **conic concentrators** (Sec. 3.2.3), and this effect has to be accounted for to avoid double counting, since the concentrators are simulated in g4bx2. This results in  $Q^{\text{II}}$  (Sec. 6.3.2).

Finally, one needs to correct for the so-called **chosen PMT bias**, newly introduced by the usage of the energy estimator  $N_h^{geo}(\text{B900})$ , defined in Eq. 6.2. This yields  $Q^{\text{III}}$  (Sec. 6.3.3).

#### 6.3.1 Dark noise correction

In Eq. 6.1,  $N_h^c$  stands for hits resulting from the candle events  $c$ . However, when an event is registered by the detector electronics, part of the hits come from dark noise. Therefore, in order

to obtain  $N_h^c$  I must subtract estimated dark noise hits:

$$N_h^c(p) = N_h^{ev}(p) - N_h^{dark}(p), \quad (6.4)$$

where  $N_h^{ev}$  are hits registered in Laben, and  $N_h^{dark}$  is our estimation of the number of dark hits expected to happen in the corresponding time window. It can be estimated as follows:

$$N_h^{dark}(p) = R_{dark}(p)t_{ev}, \quad (6.5)$$

where  $t_{ev}$  is the sum of the durations of each registered physical candle event (i.e. *cluster*, see Sec. 3.4.2), and  $R_{dark}(p)$  is the approximate average dark noise rate in a given PMT  $p$ . It, in turn, can be calculated as:

$$R_{dark}(p) = \frac{N_h^{TT64}(p)}{N_{TT64}t_{gate}}. \quad (6.6)$$

In the formula above,  $N_h^{TT64}$  is the number of hits on PMT  $p$  coming from noise events (TT64 C0) collected in the same time window as the candle event statistics,  $N_{TT64}$  is the number of such events, and  $t_{gate}$  is the DAQ gate duration. The rate of dark noise is estimated to be of the order of  $1 \mu s^{-1}$  (see Fig. 5.11A); taking into account that the typical cluster duration is  $\sim 1.5 \mu s$ , the resulting contribution from dark hits is of the order of 1-2 hits per cluster. After the dark hits are subtracted, I calculate the first-stage EQE as follows:

$$Q^I(p) = \frac{N_h^c(p)}{N_c(p)}. \quad (6.7)$$

The new approach for EQE calculation accounts for the  $^{14}C$  event duration  $t_{ev}$  in Eq. 6.5 ( $\sim 1.5 \mu s$ ), while in the old approach the whole DAQ gate duration  $t_{gate}$  ( $16 \mu s$ ) was used instead, resulting in an overestimation of the dark noise hits. The parameter  $Q^I$  will be later scaled to account for other effects described in the subsections below.

The uncertainty on the value of EQE is calculated assuming a Poisson fluctuation for number of hits  $\sigma(N_h) = \sqrt{N_h}$  for  $N_h^{ev}(p)$ ,  $N_h^{TT64}(p)$ , and other  $N_h$  parameters that are later involved in the calculation of  $Q^{II}$  and  $Q^{III}$ ; and performing corresponding propagation of uncertainty.

### 6.3.2 Correction for conic concentrators

All but 371 PMTs are equipped with cone-shaped concentrators (or “cones”) that assist light collection (see Sec. 3.2.3). The cones make these PMTs register more hits coming from the fiducial volume, and this effect should be decoupled and distinguished from the PMTs having higher EQE, as conic concentrators are simulated in g4bx2. We introduce the cone correction factor  $f_C$  that is used to scale  $Q^I(p)$ , obtained in Eq. 6.7:

$$Q^{II}(p) = \begin{cases} Q^I(p)f_C, & p \in P_C \\ Q^I(p), & p \notin P_C \end{cases}, \quad (6.8)$$

where  $P_C$  is the subset of PMTs that are equipped with cones. After this correction, the PMTs with cones are “brought to the same footing” as the ones that do not. In the old approach,  $f_C$

was calculated as the ratio of the *average* EQE values  $\hat{Q}^I$  in PMTs without and with cones in each time window of the data sample:

$$f_C = \frac{\hat{Q}^I(p \notin P_C)}{\hat{Q}^I(p \in P_C)}. \quad (6.9)$$

However, this ratio is time dependent which is demonstrated in Fig. 6.5A by the blue curve. To cross-check this dependence, I also calculate  $f_C$  in a similar way, but using average EQE among only B900 PMTs with and without cones (orange curve), which show better stability. Since there is no reason for the geometrical effect of the cone to be dependent on time, and since B900 PMTs yet again show more stable behavior, I conclude that the time variations must be introduced by other effects which should not contribute to  $f_C$ .

For the new approach, a study using MC simulation of monoenergetic electrons was conducted to obtain the cone factor independent from other effects, since, as mentioned above, g4b $\times$ 2 simulation includes the conic concentrators. The electrons have been generated in the energy range 50 - 200 keV (range of  $^{14}\text{C}$ ), distributed uniformly in a sphere with a radius of 2 m in an “ideal” detector, in which all PMTs are alive and enabled. In the simulation, all PMTs have been set to have equal QE and no dark noise. The resulting distribution of hits in the PMTs with and without cones is depicted in Fig. 6.5B. The factor  $f_C$  is then calculated as the ratio of the mean of the two distributions and was found to be  $f_C = 1.4891 \pm 0.0008$ .

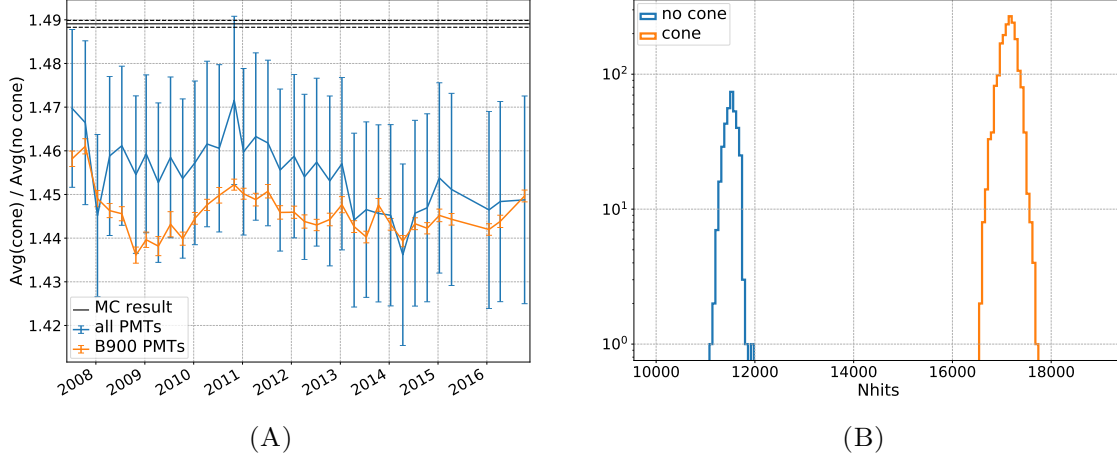


FIGURE 6.5: (A) The factor  $f_C$  calculated from the MC study (black) compared to the one calculated in Eq. 6.9 for all PMTs (blue) and B900 PMTs (orange). (B) Distribution of hits from the simulated monoenergetic  $e^-$  in PMTs with (orange) and without cones (blue).

### 6.3.3 Chosen PMT correction

Using the energy estimator based on the B900 PMTs  $N_h^{geo}(\text{B900})$  from Eq. 6.2 introduces what we call a **Poissonian bias** in the number of hits collected by the B900 PMTs. Since the condition  $50 < N_h^{geo}(\text{B900}) < 100$  is applied to guarantee a certain number of hits collected in the B900 PMTs, but not in the rest of the PMTs, they tend to have 20-30% more hits on

average, as can be seen in Fig. 6.6. In this figure, I define the ratios  $R_{true}$  as follows:

$$R_{true} = \frac{\hat{Q}_{all}^{\Pi}(p \in \text{B900})}{\hat{Q}_{all}^{\Pi}(p \notin \text{B900})}, \quad (6.10)$$

where  $Q_{all}$  is the EQE calculated with the condition  $50 < N_h^{geo} < 100$ , i.e. using *all* PMTs to select  $^{14}\text{C}$  events, rather than only B900; while  $R_{biased}$  is defined similarly as:

$$R_{biased} = \frac{\hat{Q}^{\Pi}(p \in \text{B900})}{\hat{Q}^{\Pi}(p \notin \text{B900})}, \quad (6.11)$$

for which the condition  $50 < N_h^{geo}(\text{B900}) < 100$  is used. As mentioned before, we do not trust the EQE calculated using  $N_h^{geo}$  since it introduces an unphysical increase seen in Fig. 6.4, and discussed in detail in Sec. 6.2.3. However, since this effect influences B900 and non-B900 PMTs equally, the ratio between their average EQEs is unaffected. As one can see,  $R_{biased}$  is systematically larger than  $R_{true}$ , while following the same trend.

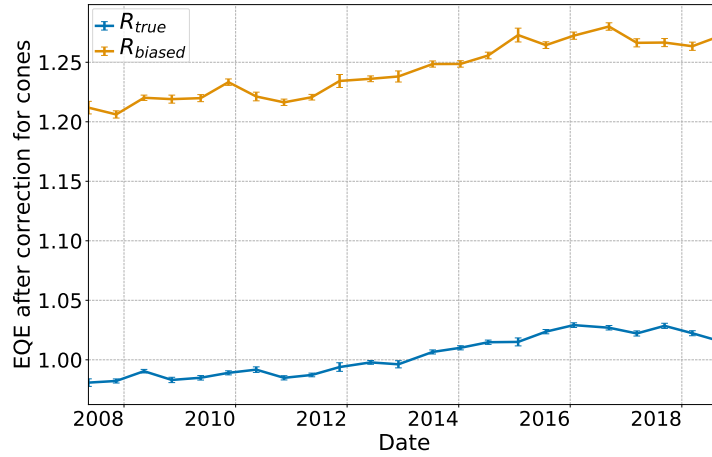


FIGURE 6.6: Ratios  $R_{true}$  (orange) and  $R_{biased}$  (blue) from Eq. 6.10 and Eq. 6.11, respectively (see text).

### Studies B900 subsets

In order to analyze the chosen PMT bias effect better, and prove that it does not come from other effects, I performed a study based on subsets of B900 PMTs. For this purpose, I split the B900 PMT set into two arbitrary halves, **B450a** and **B450b**. I performed such a splitting three times resulting in three different B450 pairs. Each time, I took turns to use each of the two subsets to select  $^{14}\text{C}$  events with an energy estimator  $N_h^{geo}(\text{B450a})$  and  $N_h^{geo}(\text{B450b})$ , defined similarly to  $N_h^{geo}(\text{B900})$ . As can be seen from Fig. 6.7A, when B450a is used to select  $^{14}\text{C}$  events, the distribution of hits in the B450a PMTs represents the tail of the  $^{14}\text{C}$  spectrum, as expected (blue histogram). However, the distribution of B450b PMTs shows an entirely different behavior. The same picture is observed when the roles of B450a and B450b are reversed. However, when B900 are used for  $^{14}\text{C}$  selection, the two subsets do not show any difference, as shown in Fig. 6.7B.

To understand the effect seen in Fig. 6.7A, I have looked at a single bin of the B450a distribution in Fig. 6.7A, and the corresponding hits in B450b. This study is illustrated in Fig. 6.8.

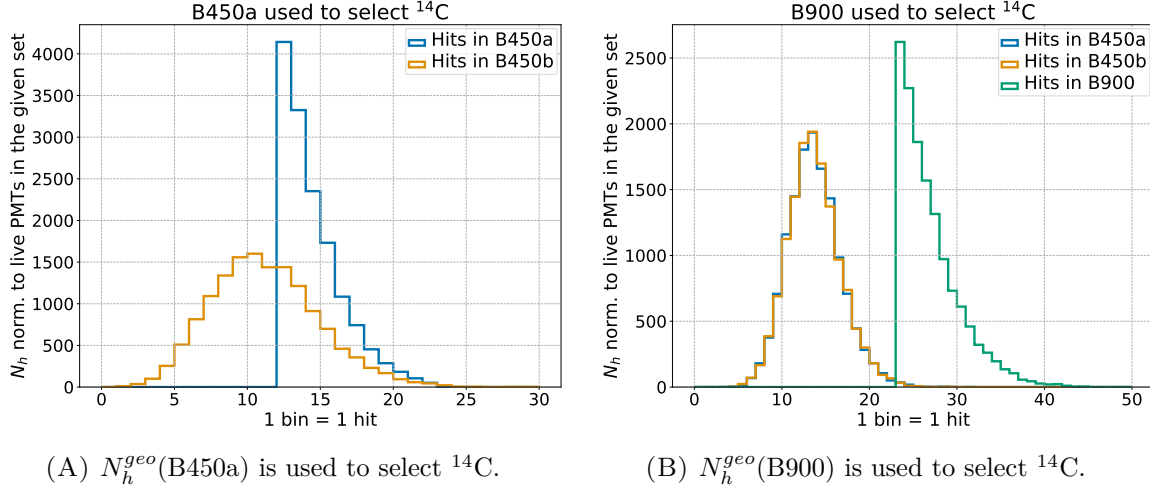


FIGURE 6.7: Distribution of hits in B450a (blue), B450b (orange) and B900 (green) PMTs using different PMT sets to select  $^{14}C$ .

When the B450a PMTs collectively register  $N$  hits, the B450b set has a certain probability to register  $N$ , more, or less hits, due to statistical fluctuations. Poisson distributions with  $\lambda = N$  are shown in Fig. 6.8A and Fig. 6.8B for  $N = 12$  and  $N = 15$ , respectively, taken as an example. The Poisson distributions are shifted with respect to the histograms of B450b, which suggests additional effects. However, we conclude that the nature of the chosen PMT bias lies in statistical fluctuations, and not the properties of the set of chosen PMTs, and this bias has to be corrected.

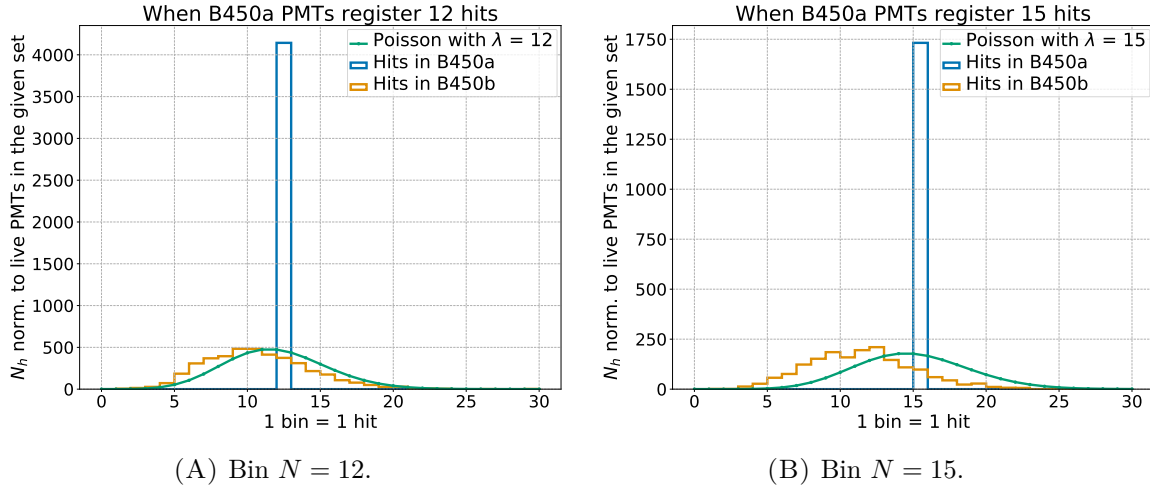


FIGURE 6.8: Selected bin of spectrum in Fig. 6.7A (blue) compared to hits in B450b (orange) and Poissonian distribution centered around the selected bin value (green).



### Correction factor

To correct for the chosen PMT bias, the EQE  $Q^{\text{II}}$  calculated in the previous stage in Eq. 6.8 for the non B900 PMTs is scaled up to be on the same level as the EQE of B900 PMTs with a factor  $f_B$ :

$$Q_h^{\text{III}}(p) = \begin{cases} Q_h^{\text{II}}(p)f_B, & p \notin \text{B900} \\ Q_h^{\text{II}}(p), & p \in \text{B900}. \end{cases} \quad (6.12)$$

Due to unknown effects that make the distributions in Fig. 6.8 not be exactly Poissonian, we have abandoned the theoretical approach to the correction of the statistical effect described above (for more, see Outlook). Instead, we base the correction factor on a more practical approach and define it as follows:

$$f_B = R_{\text{biased}}/R_{\text{true}}, \quad (6.13)$$

where  $R_{\text{biased}}$  and  $R_{\text{true}}$  are the ratios defined in Eq. 6.11 and Eq. 6.10, respectively.

We consider  $Q^{\text{I}}$  of B900 PMTs, calculated using  $N_h^{\text{geo}}(\text{B900})$ , to be affected by a bias towards the B900 PMTs that play the role of the “chosen PMTs”, same as the B450a PMTs are affected compared to B450b when B450 are “chosen” (Fig. 6.7A). I remind that these arbitrarily chosen subsets have no difference in PMT quality or other parameters. The ratio  $R_{\text{biased}}$  quantifies this bias. On the other hand, when EQE is calculated using all PMTs, the B900 PMTs are unaffected by this bias, same as B450a PMTs are not affected in the case when B900 is used for  $^{14}\text{C}$  selection (Fig. 6.7B). The ratio  $R_{\text{true}}$  represents the actual difference in EQE of B900 compared to other PMTs, free of the Poissonian bias.

## 6.4 Collecting sufficient statistics

### 6.4.1 Time window and radius trade-off

Apart from the conditions that define a  $^{14}\text{C}$  event, described in Sec. 6.2, two more parameters come into play when selecting candle events: **time window** of the data sample for regular monitoring of EQE, and the **radius** of the sphere in which I select such events. There is a trade-off in the choice of these two parameters as both of them influence statistics:

- larger radius yields more statistics for EQE calculation, and, therefore, results in a smaller uncertainty; but too large of a radius might introduce a bias in the PMTs in the top hemisphere of the detector (“north PMTs”), which outnumber the ones in the bottom one (“south PMTs”), due to events happening too close to the edge, even when using a geonormalized variable;
- a smaller time window allows us to monitor EQE changes more frequently, but reduces statistics.

One can optimize the selection of the radius and the time window, tuning it based on the resulting uncertainty, and the north/south bias. I approximate the uncertainty as a function of time window and radius with two assumptions, that the number of collected  $^{14}\text{C}$  events is

proportional to  $R^3$ , and to the number of weeks:

$$N_{ev}(N_{weeks}, R) = N_{weeks} R^3 N_{ev}(1 \text{ week}, 1 \text{ m}). \quad (6.14)$$

Taking the simple form of EQE ( $Q$ ) from Eq. 6.1 gives:

$$\sigma_Q = \frac{\sqrt{N_h^c(p)}}{N_{ev}}. \quad (6.15)$$

Using Eq. 6.1, Eq. 6.14, and Eq. 6.15, I derive the uncertainty  $U_Q$ :

$$U_Q(N_{weeks}, R) = \frac{\sigma_Q}{Q} = \frac{1}{\sqrt{N_{weeks} R^3}} U_Q(1 \text{ week}, 1 \text{ m}). \quad (6.16)$$

The value of  $U_Q(1 \text{ week}, 1 \text{ m})$  is calculated as  $1/\sqrt{\hat{N}_h^c(1 \text{ week}, 1 \text{ m})}$ , where  $\hat{N}_h^c(1 \text{ week}, 1 \text{ m})$  is the average number of hits collected by a PMT from 1 week of  $^{14}\text{C}$  statistics selected with  $R = 1 \text{ m}$ . The value  $U_Q$  in % is shown in Fig. 6.9. The decision has been made that EQE changes have to be monitored on a weekly basis, to follow the Borexino data which is grouped by week, and weekly electronics maintenance which can also influence EQE. From this study I conclude that in order to achieve an uncertainty of 1.5% using a time window of one week, I have to use  $R = 2 \text{ m}$ .

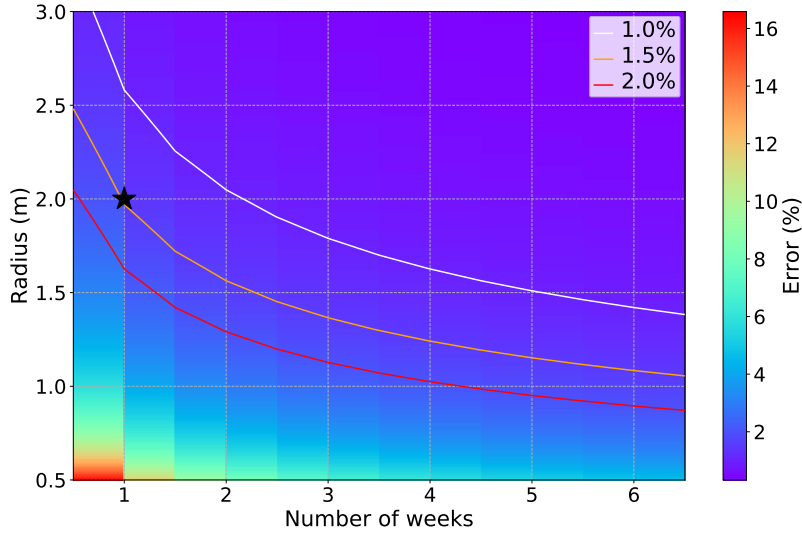


FIGURE 6.9: Estimated EQE uncertainty  $U_Q$  (Eq. 6.16) as a function of the time window and radius used to collect  $^{14}\text{C}$  events.

As the next step, I have verified that the radius of 2 m is not too large to cause any issues or biases. Figure 6.10A average EQE in B900 PMTs, calculated using with different radii. As one can see, all radii  $R \geq 1.5 \text{ m}$  give virtually the same results, and are compatible with those from  $R < 1.5 \text{ m}$ . Moreover, larger radii give a better more stable trend due to more statistics that grows rapidly with  $R^3$ .

Furthermore, I have looked at the ratio of the average EQE in the north (top hemisphere) and south (bottom hemisphere) B900 PMTs to probe for biases described above. Since there are more live B900 PMTs in the north than in the south (with the ratio of 1.6, reported in Table 5.1), it could lead to a “preference” to  $^{14}\text{C}$  events happening in the top half of the detector, or result in an incorrectly higher EQE of north PMTs. Exactly this bias can be seen when using the normalized variable  $N_h^{norm}(\text{B900})$ , as demonstrated in Fig. 6.10B by the purple curve and the ones above it, while no such bias is introduced when using  $N_h^{geo}(\text{B900})$  (blue, orange, green and red curves). Note that I use average EQE of B900 PMTs to perform both checks in order to avoid other effects coming into play, since we know that the EQE of other PMTs of potentially unstable quality may not represent effective LY changes.

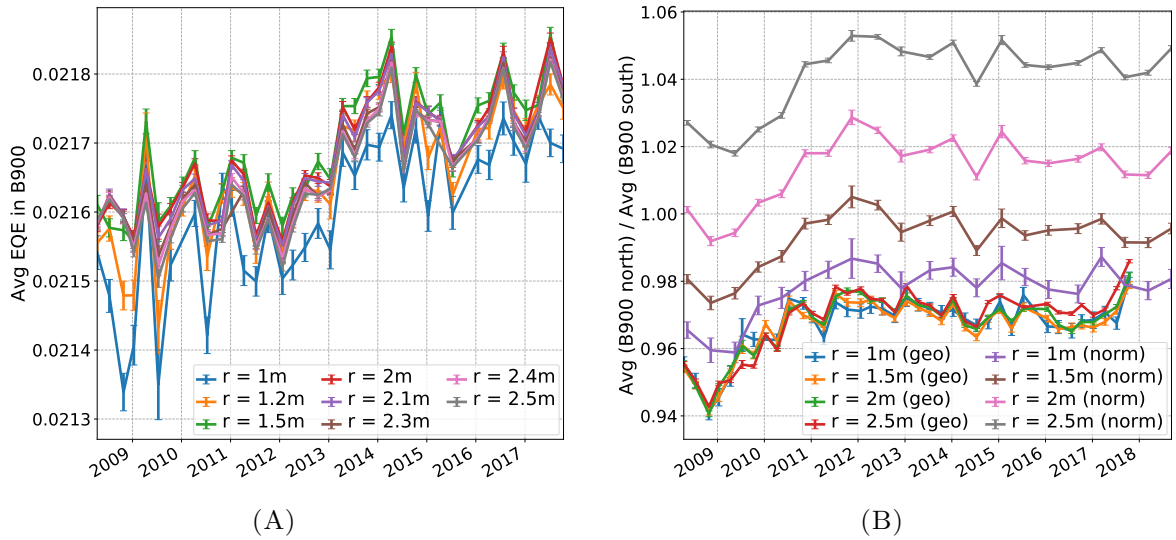


FIGURE 6.10: (A) Weekly average EQE in B900 PMTs based on selections with different radii. (B) Ratio of the average EQE in north and south B900 PMTs for different selection radii and energy estimators  $N_h^{norm}(\text{B900})$  and  $N_h^{geo}(\text{B900})$ .

Another important conclusion is that while the north/south ratio of the  $R = 2\text{m}$  selection is similar to those of  $R < 2\text{m}$ , choosing  $R = 2.5\text{m}$  shows a small but more pronounced deviation, as shown by the red curve in Fig. 6.10B. Therefore, since  $R = 2\text{m}$  is enough to satisfy our requirement for a sufficiently small uncertainty, we prefer to sacrifice an even smaller uncertainty to guarantee no N/S biases, and choose  $R = 2\text{m}$  as the final selection radius.

#### 6.4.2 Dealing with low statistics

In spite of the tuning of the radius  $R$  for the selection of candle events, there are cases of low statistics, that may occur for different reasons. I have designed measures to be taken to deal with such cases, in order to guarantee the desired uncertainty of 1.5% or lower.

##### Low statistics weeks

Some weeks have DAQ issues and the resulting statistics is not sufficient to attain the required precision. In such cases, if the livetime of a week is less than a certain threshold, the week is “stretched”: more data is added to it run by run from the previous and next weeks, until the

livetime reaches the threshold. The effect of stretching is demonstrated by the orange points in Fig. 6.11, which shows the average uncertainty among enabled PMTs as a function of the number of  $^{14}\text{C}$  events, each point representing one week of statistics. While the initial calculation includes “unlucky” weeks with very low amount of  $^{14}\text{C}$  events (blue), after stretching these weeks are shifted above the threshold, decreasing the average uncertainty. Figure 6.11 was made using a study with a threshold of 150000  $^{14}\text{C}$  events. It was later decided to move the threshold up to 300000 events, corresponding to roughly 75% of a complete week of data taking.

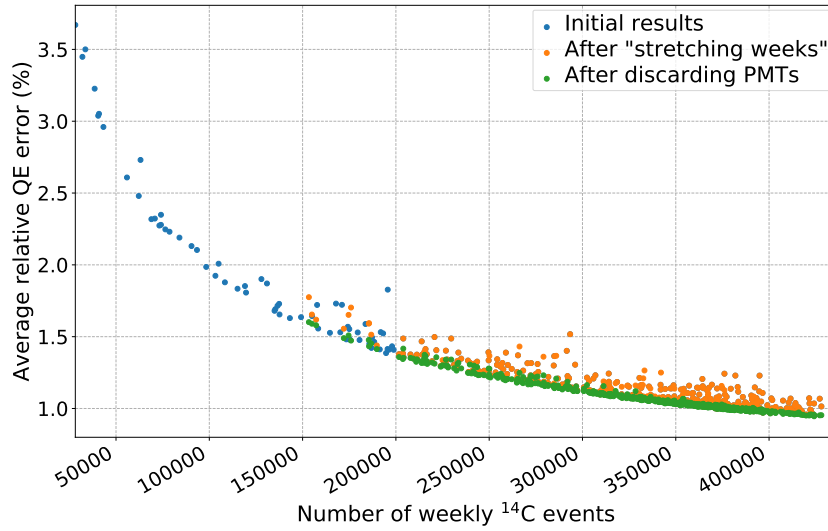


FIGURE 6.11: Relative uncertainty on EQE depending on the number of  $^{14}\text{C}$  events in a week after the initial EQE calculation (blue), after “stretching” (orange), and after discarding low statistics PMTs (green). Each dot represents one week of  $^{14}\text{C}$  statistics.

### Low statistics PMTs

A PMT will have a large uncertainty on EQE if it registered relatively few hits from  $^{14}\text{C}$  events. This can happen for two reasons: its EQE is low, so it registers fewer hits than other PMTs; or it was partially disabled during the week and registered hits from fewer  $^{14}\text{C}$  events than other PMTs. In the first case, no matter how large the error is, I keep the EQE information of the PMT. In the second case, I discard the value computed using the data of that week, because we do not trust its accuracy.

To differentiate between these two cases, I first check how long a PMT was enabled during the week. If it registered more than 150 000  $^{14}\text{C}$  events, its value is kept even if the error is large, because it is so not due to lack of statistics, but low EQE. If it registered less than that, and still has a relatively small error, its EQE value is kept as well; however, if the relative error on EQE is  $> 2\%$ , this PMT is “discarded” and treated the same way as PMTs that were disabled during the whole week.

The effect of discarding is demonstrated in Fig. 6.11 by the green points. Note how the already small average error after “stretching” gets even more refined as “unlucky” PMTs are discarded. The final results demonstrate uncertainty predominantly smaller than 1.5%, as estimated in Sec. 6.4.1.

### Dealing with discarded and disabled PMTs

Disabled and discarded PMTs are assigned the EQE value from the previous week. The EQE value of disabled PMTs is relevant when simulating an “ideal” detector in which all nominal PMTs are always enabled. Note the new features compared to the old EQE approach: 1) the concept of discarded PMTs, and 2) in the old approach, disabled PMTs were assigned a constant value corresponding to average EQE of the week 2009-Jun-14, determined based on source calibrations conducted during that week (Sec. 3.2.5).

## 6.5 Conversion of EQE to REQE

As mentioned in Sec. 6.1, the final parameter used in `bx_elec` is REQE, which takes values from 0% to 100% by design of the simulation, and represents the EQE of a PMT relative to the maximum possible QE that has been measured to be  $\sim 32\%$ .

### Old approach

The old approach to converting EQE to REQE relied on a wrong assumption that the average EQE of all PMTs is stable in time. The first step of the old method was to obtain the distribution of hits in enabled PMTs resulting from the 2009 calibrations using  $^{203}\text{Hg}$  source (Sec. ??), and scale the values as  $Q_{\text{Hg}}(p) = N_h(p)/N_h^{\text{max}}(p_{\text{max}})$  such that the PMT with the highest number of hits has  $Q_{\text{Hg}}(p_{\text{max}}) = 1$ . Then, the mean value of this distribution (calculated to be 0.748) was taken and used to scale EQE (denoted as  $Q$ ) such that the average REQE in a given time sample is always the same:

$$Q_R(p) = Q(p) \cdot 0.748 / \hat{Q}(\text{all } p). \quad (6.17)$$

There are two big flaws with this approach. The first one is that it is not designed to follow LY and other effective changes, since it is forcing the average EQE of the live PMTs to be the same at any time point. This is not a guarantee, and not only that, the purpose of monitoring EQE is precisely to account for possible changes like that. Moreover, even if each single PMT has a perfectly stable EQE, the average EQE may change due to dying PMTs.

The second flaw is that due to this scaling factor, many PMTs ended up having  $Q_R(p) > 1.0$ , possibly due to wrong calculation of the factor 0.748. Since such input is not possible in `bx_elec`, the values  $Q_R(p) > 1.0$  were set to  $Q_R(p) = 1.0$ . The effects of this “saturation” can be seen in Fig. 6.12A, where 114 PMTs in the week of 2009-Jun-14 (orange histogram) have  $Q_R(p) = 1.0$ . This bad scale distorted the true distribution of EQE and did not represent the true picture of the detector.

### Improved approach

In the improved approach, I seek an **absolute scale** of converting EQE to REQE, without relying on average EQE, which might change in time. In order to avoid retuning of the LY parameter for the MC simulation (determined from the calibrations), I set the average REQE in the week of 2009-Jun-14 to be the same as the average of the REQE calculated using the old approach (found to be 0.7326 after excluding disabled PMTs). Thus, the average EQE of that week (calculated to be 0.021) is used as a scale for all further weeks, i.e. EQE of 21 hits per

1000 candle events is set to correspond to REQE of 74.8%:

$$Q_R(p) = Q^{\text{III}}(p) \cdot 0.748 / \hat{Q}^{\text{III}}(2009\text{-Jun-14}), \quad (6.18)$$

where  $Q^{\text{III}}$  is obtained as in Eq. 6.12.

The resulting distribution is shown in Fig. 6.12A by the blue histogram. As you can see, while the average REQE is the same as in the old sample, the issue with saturated PMTs is avoided (only 2 PMTs have REQE > 1.0). Figure 6.12B shows the evolution of the number of saturated PMTs in time, and as you can see, in the old approach, this number is not only high, but also very unstable in time, since it relies on the changing average EQE; while in the new approach it is well below 10 PMTs and stable.

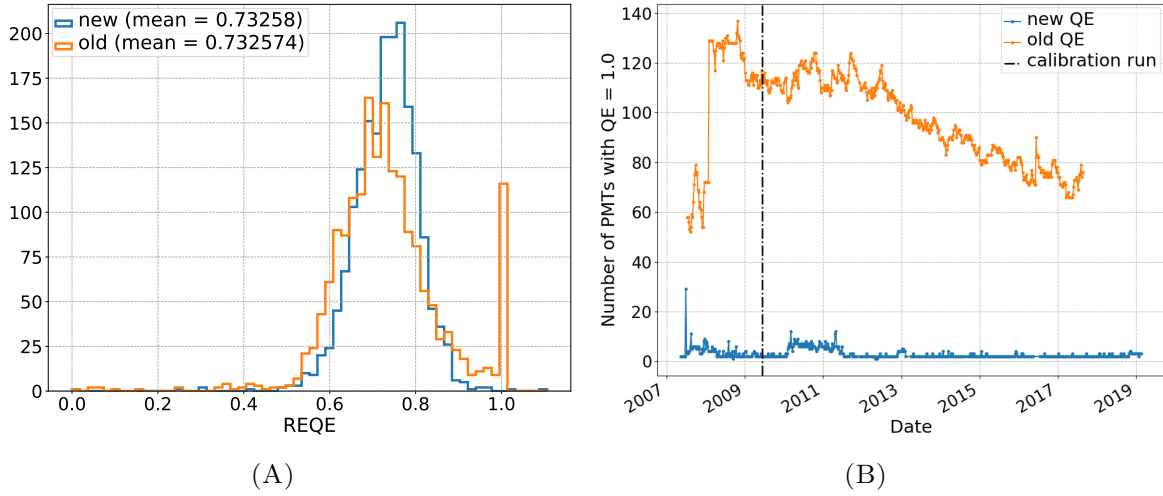


FIGURE 6.12: (A) Distribution of the old (orange) and new (blue) REQE in enabled PMTs in the week of 2009-Jun-14. (B) Evolution of the number of saturated PMTs in time.

## 6.6 Results

The resulting REQE in each week is shown in Fig. 6.13. For comparison, Fig. 6.14 shows the time evolution of the REQE calculated with the old method. Note that the old REQE does not have information about the uncertainty on the calculated REQE value. Several examples of the time evolution of REQE of single PMTs is shown in App. B.1.

The green points represent the average REQE in the PMTs that were live at the moment the study was conducted, in July 2019 (**L2019**). The black curve shows the average in all enabled PMTs, and the red shows the average in the remaining enabled PMTs in that data sample (non-L2019). The number of L2019 PMTs stays more or less constant in time, since none of them die before July 2019, and the only deviations come from temporarily disabled PMTs. The number of PMTs belonging to non-L2019 is decreasing until it reaches zero in July 2019, where “all” = “L2019”. The REQE results are analyzed this way since simply looking at the average in all enabled PMTs only shows the general picture of the detector, but does not contain information about the effective LY or the effect of dying PMTs. Also, as mentioned before, survivor PMTs

and the rest of the PMTs tend to show a systematic difference in quality, as demonstrated in Sec. 5.3.1.

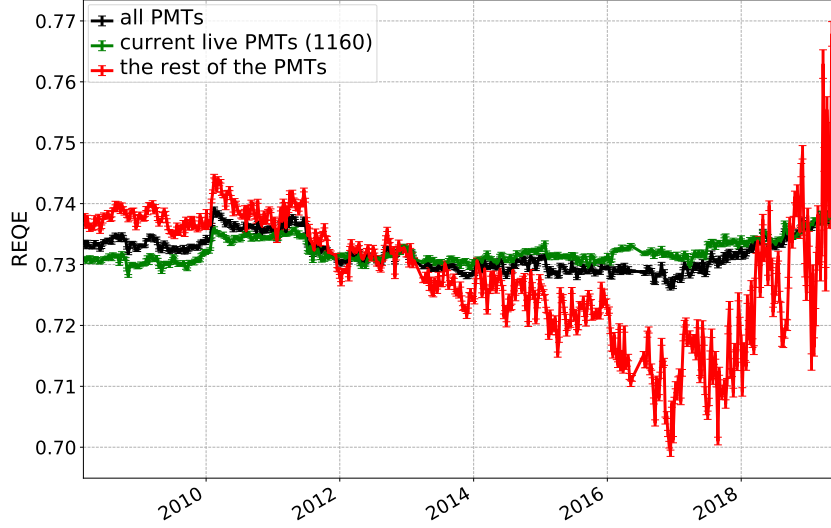


FIGURE 6.13: Average REQE among all live PMTs in a given week (black), L2019 PMTs (green) and the remaining PMTs in each week (red).

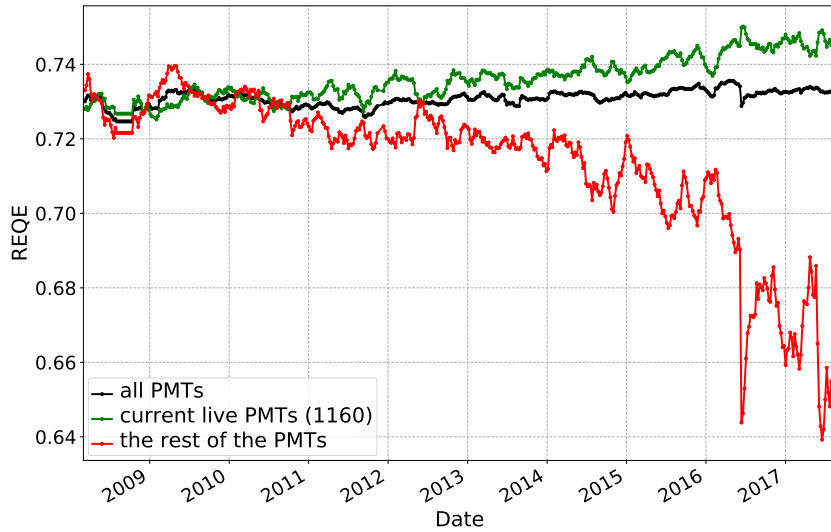


FIGURE 6.14: Average REQE among all live PMTs in a given week (black), L2019 PMTs (green) and the remaining PMTs in each week (red), calculated with the old method

A bump of higher REQE values is observed around years 2010-2011 in Fig. 6.13. This happens due to the  $^{222}\text{Rn}$  contamination during the water extraction period of the purification campaign (Sec. 3.2.6). These REQE values are not valid, but they are not used in any analysis, since Phase-I ends before, and Phase-II starts after that period. Note that while the water extraction is very clearly visible in the REQE calculated with the new method, the REQE based on the old method, shown in Fig. 6.14, does not illustrate such an effect. This happens because the average REQE in each time window was forced to be the same, as explained in Sec. 6.5. For the



same reason, the average of all PMTs, shown in black, shows a false stability of  $< 0.5\%$ , as this same average was simply set to be the same in each dataset.

Starting from the beginning Phase-II (Dec 2011), PMTs start dying with a constant rate. This fact alone does not explain the resulting decrease in non-L2019 REQE followed by an increase starting from around 2017. My interpretation is that the PMTs that die tend to also worsen before dying, while stable high-QE PMTs tend to survive, so after all the worst ones have died, the average goes up again, as the ones with worst REQE do not contribute anymore. The effect of single PMTs worsening before dying, or being worse to begin with, can be observed in the plots in App. B.1. This conclusion is also supported by other studies on PMT quality and its correlation with their livetime shown in Sec. 5.3.1. Comparison of REQE in each set of PMTs in the old and new method is shown in Fig. 6.15 - 6.17.

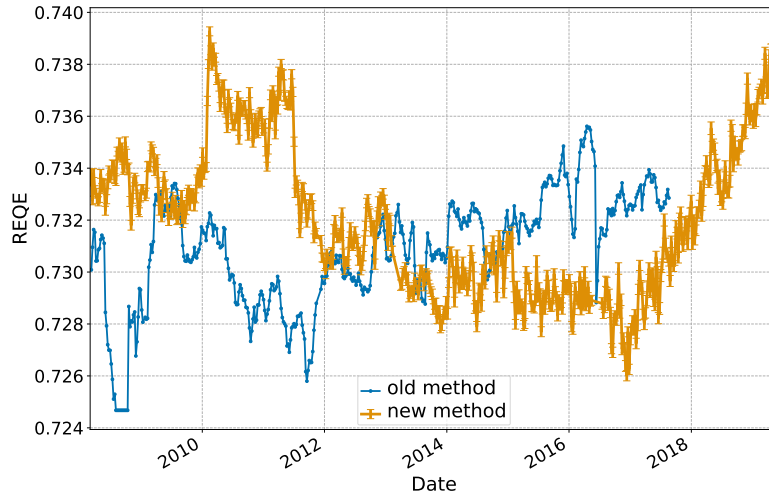


FIGURE 6.15: Average REQE of all enabled PMTs.

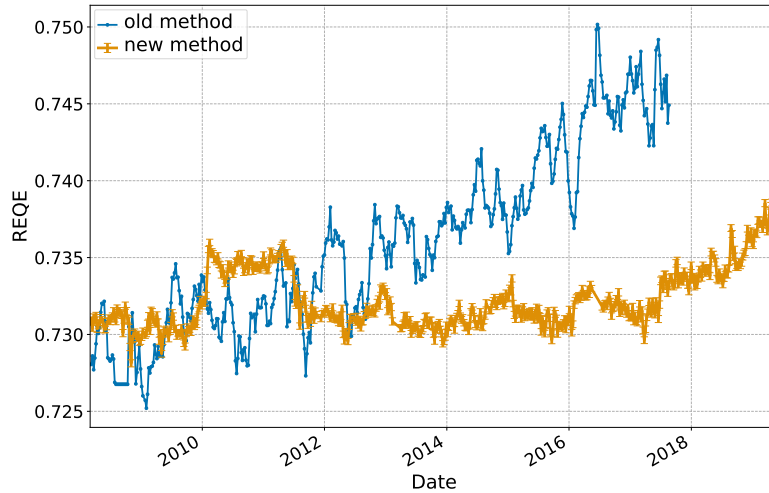


FIGURE 6.16: Average REQE in enabled L2019 PMTs.



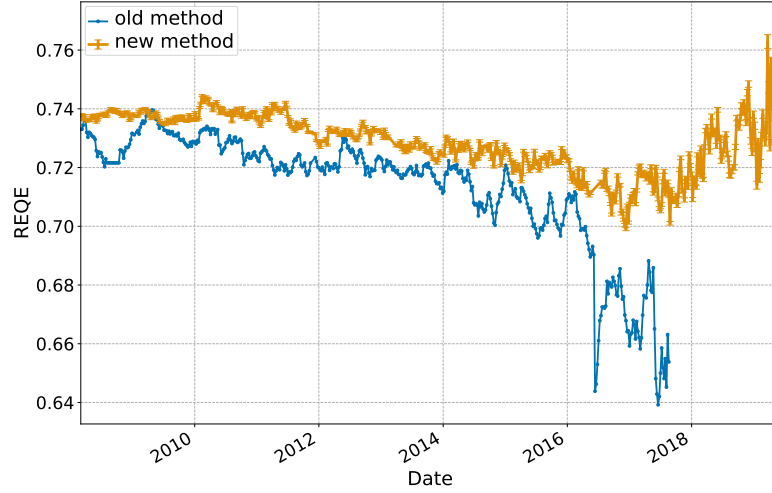


FIGURE 6.17: Average REQE in non-L2019 PMTs

## 6.7 Validation of the new REQE

In order to verify that the new REQE inputs to `bx_elec` obtained with the improved method improve the MC simulation, and eliminate the mismatch with the data, extensive checks have been performed using gamma calibrations (Sec. 6.7.1), and effective LY studies based on  $^{210}\text{Po}$  data (Sec. 6.7.2) [101]. Furthermore, I performed additional checks with multivariate fits on Phase-II using MC reference spectral shapes constructed with old and new REQE inputs (Sec. 6.7.3), and the fit results on the position of the  $^{210}\text{Po}$  peak (Sec. 6.7.4) in Phase-II and Phase-III data.

### 6.7.1 Monte Carlo simulation of source calibrations

The calibration run 10396 with three gamma sources ( $^{85}\text{Sr}$ ,  $^{65}\text{Zn}$ ,  $^{60}\text{Co}$ ) has been simulated using old and new REQE values as input. Then, for data itself, and for the two versions of MC simulation with old and new REQE, the fraction of hits  $F$  in a PMT  $p$  was defined the following way:

$$F(p) = \frac{N_h(p)}{N_h(all)} \cdot 1000, \quad (6.19)$$

representing the light collection of the PMT [101]. The distribution of  $F(p)$  calculated for each PMT is shown in Fig. 6.18, for PMTs with and without conic concentrators (“cones”) separately.

As one can see, new REQE results in a distribution much more similar to the one of data, as opposed to the old REQE inputs. The mean values of the distributions are reported in Table 6.1.

Additionally, the data from multiple runs with different sources, covering different positions and energy ranges, has been simulated using old and new REQE values as input. Different physical quantities, describing the detector response and values of energy estimators, have been obtained from simulations based on old and new REQE inputs, and compared with data using the  $\chi^2$  test [101]. Figure 6.19 demonstrates global distributions of  $\chi^2/N_{dof}$  across all the considered

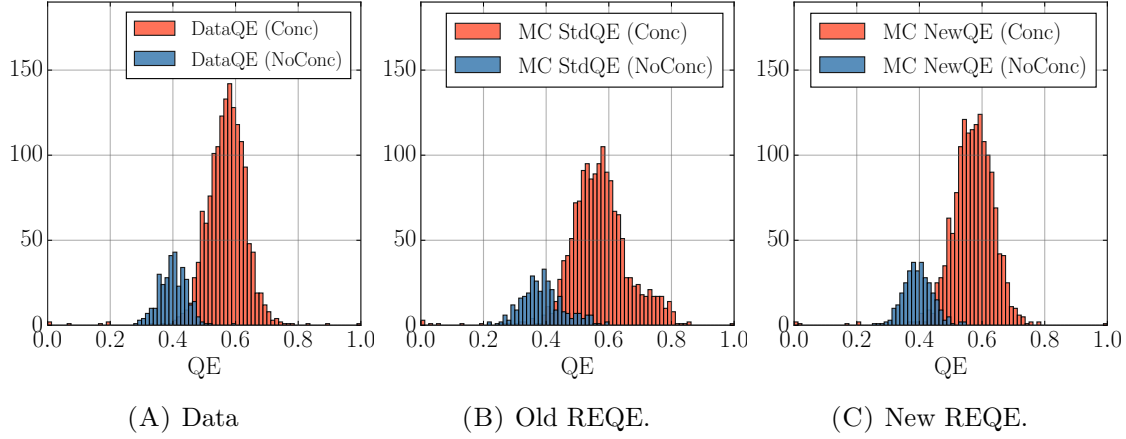


FIGURE 6.18: Distributions of  $F(p)$ , obtained as in Eq. 6.19, for each PMT  $p$  with (orange) and without (blue) conic concentrators based on data and MC simulation [101].

	Data	MC with old REQE	MC with new REQE
<b>Cone</b>	0.5706	0.5726	0.5716
<b>No cone</b>	0.4005	0.3906	0.3957

TABLE 6.1: Mean values of the distributions of  $F$ , shown in Fig. 6.18, obtained based on calibration data, and MC simulation with new and old REQE.

variables. More details on this study can be found in [101]. One can see that the distributions with old and new REQE inputs are compatible, and the mean values of the two are 1.38 and 1.31, respectively. We conclude that for the calibration time period, the new REQE inputs perform similarly or better compared to the old ones. This is what we aimed to achieve, since the absolute scale for obtaining REQE from EQE has been tuned to agree with the old one (Sec. 6.5), and the result confirms that no retuning of MC LY parameter is needed to be able to use new REQE inputs.

### 6.7.2 Effective light yield

In order to monitor effective LY changes, strict  $^{210}\text{Po}$  data sample selected using the MLP discriminator (Sec. 3.4.5) in a sphere of  $r = 3\text{ m}$  was used to measure the detected hits in all PMTs and obtain the position of the  $^{210}\text{Po}$  peak. The resulting relative trend of the peak position compared to 2012, has been shown in Fig. 6.1 and is also demonstrated here in Fig. 6.20 by the green curve. To check whether the MC follows this trend, a sample of  $^{210}\text{Po}$  was simulated based on realistic spatial distribution with the same radius, using MC with old and new REQE inputs. The same calculation done on data was performed on these MC events [101].

As we expect, the REQE inputs in `bx_elec` directly influence the number of hits registered by the PMTs in the simulation, which, in turn, has an impact on the energy estimators, and the position of the  $^{210}\text{Po}$  peak. As I showed in Sec. 6.1, the old REQE displayed an increasing trend, opposite to the effective LY trend from  $^{210}\text{Po}$  studies, which was our motivation for the development of the new improved REQE method. As one can see now in Fig. 6.20, the effective LY calculated from MC with new REQE inputs correctly follows the  $^{210}\text{Po}$  peak trend seen in

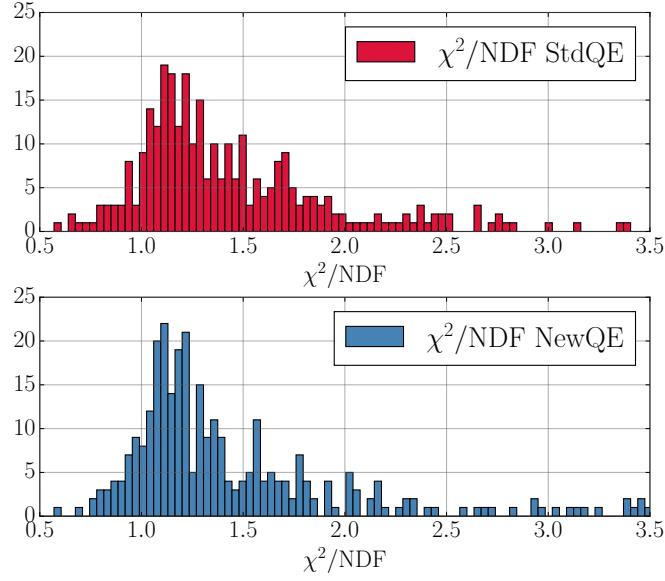


FIGURE 6.19:  $\chi^2/N_{dof}$  comparing data and MC using old (red) and new (blue) REQE [101].

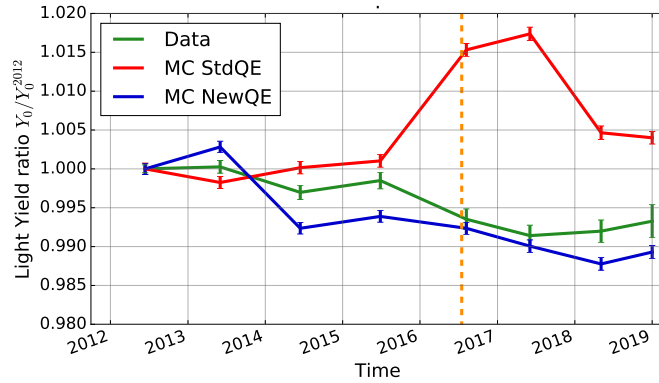


FIGURE 6.20: Relative variations of the  $^{210}\text{Po}$  peak position from data (green) and  $^{210}\text{Po}$  MC with old (red) and new (blue) REQE inputs. The beginning of Phase-III is marked with the orange dashed line [101].

data, as opposed to the old REQE. Note how the largest discrepancy between data and MC with the old REQE inputs is observed in Phase-III, which was the reason behind the large mismatch causing inability to perform the MCA multivariate fit.

### 6.7.3 Phase-II multivariate fit with MCA based on new REQE

In order to test the behavior of the new REQE inputs, I have performed multivariate fit with Monte Carlo Approach (MCA, Sec. 7.1.3) on Phase-II data with MC reference shapes generated using old and new REQE inputs, while keeping the rest of the fit configuration the same. The fit is performed in the so-called “**high setting**”, used in the Phase-III CNO neutrino analysis (see Sec. 8.5). The “high setting” considers data in the energy range  $140 < N_h < 950$ , excluding  $pp$  neutrinos,  $^{14}\text{C}$ , and pileup. In this study, the CNO neutrino rate is constrained to the value predicted by HM-SSM, similar to the configuration of the fit LER (Sec. 7.1.2). The results are

shown in Fig. 6.21. As one can see, old and new REQE inputs give compatible results. This check gave us the green light in using the new REQE in the analysis of Phase-III data.

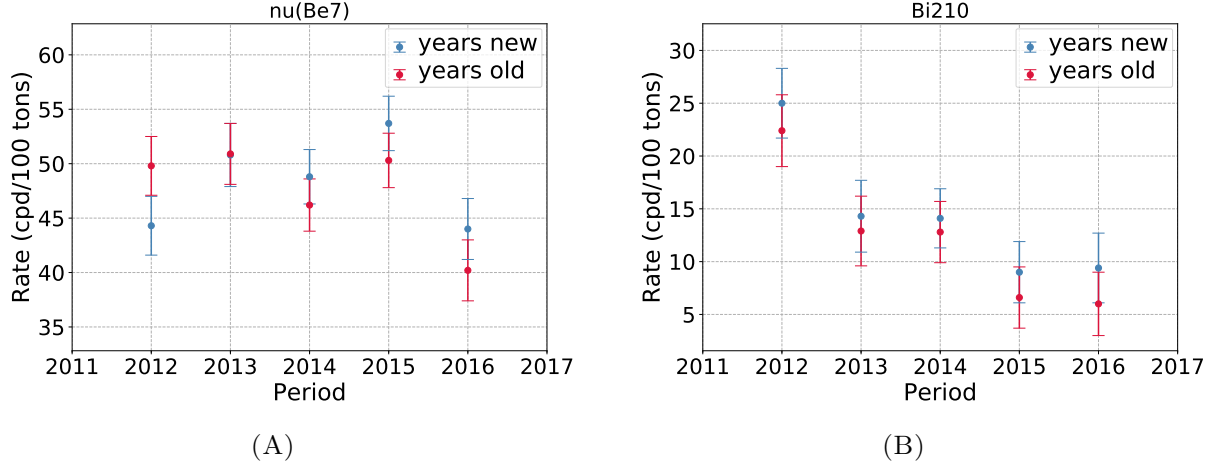


FIGURE 6.21: Phase-II fit results on the rates of  ${}^7\text{Be}$  neutrinos (A) and  ${}^{210}\text{Bi}$  (B), using MC reference shapes constructed using new (blue) and old (red) REQE inputs.

#### 6.7.4 ${}^{210}\text{Po}$ shift with new REQE

The MC reference shapes for  ${}^{210}\text{Po}$  constructed with old REQE inputs had an issue that the position of the peak did not correspond to the correct energy observed in data. At the time, solution was not found, and instead, a workaround was implemented which consisted in shifting the reference shape to adjust the peak position. The resulting shift was of around 1.5%. I performed a MC fit in the “high setting” described above in Sec. 6.7.3, with the peak position of  ${}^{210}\text{Po}$  being a free parameter in the fit. The resulting  ${}^{210}\text{Po}$  shift is less than 0.5% in any year, as demonstrated in Fig. 6.22. Indeed, further studies showed that the impact of the shift of the  ${}^{210}\text{Po}$  shape is negligible in the multivariate fit.

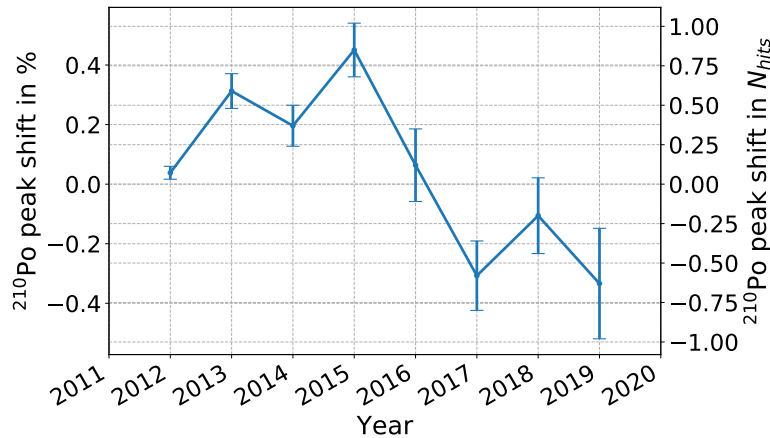


FIGURE 6.22: Absolute and relative shift of the  ${}^{210}\text{Po}$  reference shape peak position, when left free in the fit.

## 6.8 Summary

Inaccuracy of the Borexino MC in representing the latest data after 2016 motivated us to look into REQE inputs, which determine light collection in PMTs and directly influence simulation results. A new method of obtaining REQE has been implemented, which improved two major sources of discrepancy for REQE: the approach to selecting candle events (Sec. 6.2.4), and the conversion of EQE to REQE (Sec. 6.5). Other improvements have been introduced, in particular, better choice of the time window for monitoring REQE time changes (Sec. 6.4.1), and methods of dealing with low statistics cases which guarantee small uncertainty on the resulting REQE values (Sec. 6.4.2). The improved REQE shows a more physically meaningful evolution in time (Sec. 6.6). Extensive validation has been performed to make sure the new inputs perform the same and better as the old ones (Sec. 6.7).

Table 6.2 presents a summary of the differences between the old and the new method. Apart from the major improvements, some small changes have been implemented. Previously, the information about REQE was stored in text files and read by g4bx2. After some consideration, it was decided to store the REQE values in a PSQL database, same as many other detector parameters. Also, since REQE is irrelevant at the level of g4bx2, where events are generated, its processing was moved to bx\_elec where the PMT properties come into effect before the electronics is simulated.

With the new inputs, Phase-III multivariate fit using MC reference shapes is now possible. Fits on Phase-III data have already been shown in Sec. 6.7.4 with CNO neutrino rate constrained to the value predicted by HM-SSM, similarly to the fit configuration in Sec. 7.1.2. In addition to that, preliminary results of Phase-III fit with the goal of detecting CNO neutrinos, possible due to the improved REQE inputs, will be discussed in detail in Sec. 8.5.2.

	Standard approach	New approach
<b>Energy condition</b>	$50 < N_h^{geo}(all) < 100$	$50 < N_h^{geo}(B900) < 100$
<b>Radius</b>	1 m	2 m
<b>Time window</b>	sliding 3-week window	1 week
<b>Cone correction</b>	data; time dependent	MC; fixed
<b>EQE <math>\rightarrow</math> REQE</b>	scale relative to avg QE	independent absolute scale
<b>REQE error</b>	unknown	$\leq 1.5\%$
<b>Monitoring of eff. LY changes</b>	no	yes
<b>Low statistics PMTs</b>	no treatment	treated as disabled
<b>Value assigned to disabled PMTs</b>	constant	value from last available week
<b>Storage of information</b>	txt files	PSQL database
<b>Stage of usage in MC</b>	g4bx2	bx_elec

TABLE 6.2: Summary of the similarities and differences between the standard and the new approaches.



## Chapter 7

# Comprehensive measurement of the proton-proton chain neutrinos

The original goal of the Borexino experiment was to measure the  ${}^7\text{Be}$  solar neutrinos [118]. Since its construction, Borexino exceeded the expectations and measured all the  $pp$ -chain neutrinos comprehensively [72]. In this chapter, I am going to introduce you to the Borexino solar neutrino analysis techniques, my contributions to the Phase-II analysis, and finally, the published results and their implications in neutrino and solar physics.

The Phase-II solar neutrino analysis is divided into two parts with different approaches: the Low Energy Region (LER, Sec. 7.1) and High Energy Region (HER, Sec. 7.2). The LER analysis results in the measurement of the interaction rates of  $pp$ ,  ${}^7\text{Be}$ , and  $pep$  neutrinos, as well as an upper limit on CNO neutrinos; while  ${}^8\text{B}$  neutrinos measurement is conducted in the HER analysis, with an additional study on the rare  $hep$  neutrinos.

My main contribution to the LER analysis lays in performing the *multivariate fit*, which I will explain in Sec. 7.1.2), from which we obtain the final results on the rates of neutrinos and background isotopes. My work focused on the so-called analytical approach, which I will elaborate on in Sec. 7.1.3. My largest contribution consisted in the evaluation of the systematic uncertainties, which will be described in Sec. 7.1.4. I also participated in the counting analysis of the HER, resulting in an upper limit on the  $hep$  neutrinos, which will appear in Sec. 7.2.3. The results and implications of both LER and HER analyses will be reported in Sec. 7.3.

### 7.1 Low energy region analysis

The energy range of the LER analysis spans from 0.19 to 2.93 MeV, to include  $pp$ ,  ${}^7\text{Be}$ ,  $pep$ , and CNO neutrinos (refer to Fig. 2.4). The first step of the analysis is to obtain a spectrum of neutrino-like events using selection cuts that will be described in Sec. 7.1.1. After this there is no way for us to distinguish which events in that sample are neutrinos, and which come from backgrounds that mimic neutrino interactions. In order to extract the neutrino rates, a multivariate (MV) fit is performed. The details of the MV fit techniques will be described in Sec. 7.1.2.

To perform such a fit on the data, reference spectral shapes of the background and signal contributions have to be used. These reference shapes are obtained in two ways, the so-called Analytical and Monte Carlo Approaches. In the Analytical Approach (AA), the detector response is modeled using a mathematical formulation of the parameters for various detector characteristics, regarding the energy scale and resolution; and physical phenomena such as quenching and

Cherenkov radiation. In the Monte Carlo Approach (MCA), the reference shapes are obtained using `g4bx2`, the Geant4 based Borexino Monte Carlo simulation (Sec. 3.5). Each approach has its own advantages and disadvantages, a slightly different treatment in the MV fit, which I will talk about in Sec. 7.1.3, as well as different treatment regarding the systematic uncertainties.

The resulting differences coming from using one or the other approach contribute to the total systematic uncertainties which will be described in Sec. 7.1.4. The final results of the LER analysis will be presented in Sec. 7.3 together with those of the HER analysis, as well the neutrino and solar physics implications based on both analyses. The Phase-II upper limit results on the CNO neutrinos will be presented and discussed separately in Chapter 8, together with the analysis conducted on Borexino Phase-III data aimed at CNO neutrino detection.

### 7.1.1 Data selection

The data used in this analysis covers the so-called *Phase-II* of the Borexino dataset (see Sec. 3.1) which spans from December 2011 to May 2016 and covers the exposure of  $1291.51 \text{ days} \times 71.3 \text{ t}$  [115]. In order to obtain a sample of neutrino-like events, software selection is used to analyze each candidate event from the Phase-II Borexino dataset. The following conditions (the so-called “cuts”) are applied to remove unwanted events from the data sample:

- **muon cut:** removes muon events. The efficiency of this cut is  $\sim 99.992\%$  (see Sec. 4.5.1);
- **muon daughter cut:** removes the products of muon spallation in the LS (see Sec. 4.5) and effects of electronics saturation, via applying a 300 ms veto after each muon event;
- **trigger cut:** neutrino-like events are events of the type TT1 BTB0 i.e. internal events (see Sec. 3.4.4 on trigger types); all the other trigger types are removed by this cut;
- **cluster cut:** a neutrino-like event is selected as 1) an event with one cluster, or 2) first cluster of a two-cluster event (see Sec. 3.4.2 on clusters); all other events are removed by this cut;
- **$^{214}\text{Bi}$ - $^{214}\text{Po}$  cut:** removes coincidence pairs of  $^{214}\text{Bi}$  and  $^{214}\text{Po}$ , daughters of the radioactive decay of  $^{222}\text{Rn}$  (see Sec. 4.2.1). The efficiency of this cut is  $\sim 92\%$ ;
- **start time cut:** removes two cases of events with an unnatural cluster start time relative to the start of the DAQ gate (see Sec. 3.3.2), which might originate from a faulty trigger: 1) cluster start time is out of gate, 2) cluster starts too late after the beginning of the gate and should have triggered to open its own gate, but did not;
- **fiducial volume cut:** removes events far from the center of the ID. More details about the selection of the fiducial volume for the solar analysis will be presented later in this section;
- **reconstructed charge cut:** removes events that have a mismatch between the number of fired PMTs and the reconstructed charge;
- **crate fraction cut:** removes events attributed to crate noise i.e. if 75% of the fired PMTs belong to the same Laben crate (see Sec. 3.3.1);



The **fiducial volume** (FV) used in this analysis is defined as  $r < 2.8$  m and  $-1.8$  m  $< z < 2.2$  m, where  $r$  is the radius and  $z$  is the vertical coordinate inside the IV based on reconstructed event position,  $(0,0,0)$  being the center of the detector. A two dimensional  $xz$ -plane slice (at  $y = 0$ ) of the FV is demonstrated in Fig. 7.1A. The FV is chosen this way in order to maximize the active volume while minimizing radioactive background coming from the nylon spheres, PMTs, the SSS, and the endcaps. It is asymmetrical in  $z$  coordinate to account for the higher levels of radioactive contamination in the lower hemisphere. More details on the FV can be found in Sec. 4.1. Uncertainties in the calculation of the FV mass contribute to the systematic uncertainties which will be discussed in Sec. 7.1.4.

Several prominent selection cuts are demonstrated in Fig. 7.1B. Note how after the FV cut (red) three peaks from the main internal background contributions become clearly visible:  $^{14}\text{C}$ ,  $^{210}\text{Po}$ , and  $^{11}\text{C}$ , from left to right. One can even see the Compton shoulder of  $^7\text{Be}$  neutrinos following the  $^{210}\text{Po}$  peak.

Apart from these cuts, a special **threefold coincidence** (TFC) method is applied, which does not remove any events from the sample, but rather divides the sample into two categories, enriched and depleted in  $^{11}\text{C}$ . Its effect can be seen in Fig. 7.1B demonstrated by the green spectrum, with notably reduced  $^{11}\text{C}$  peak. The details of this technique were described in Sec. 4.5.3, and its usage in fitting the data will be described in Sec. 7.1.2.

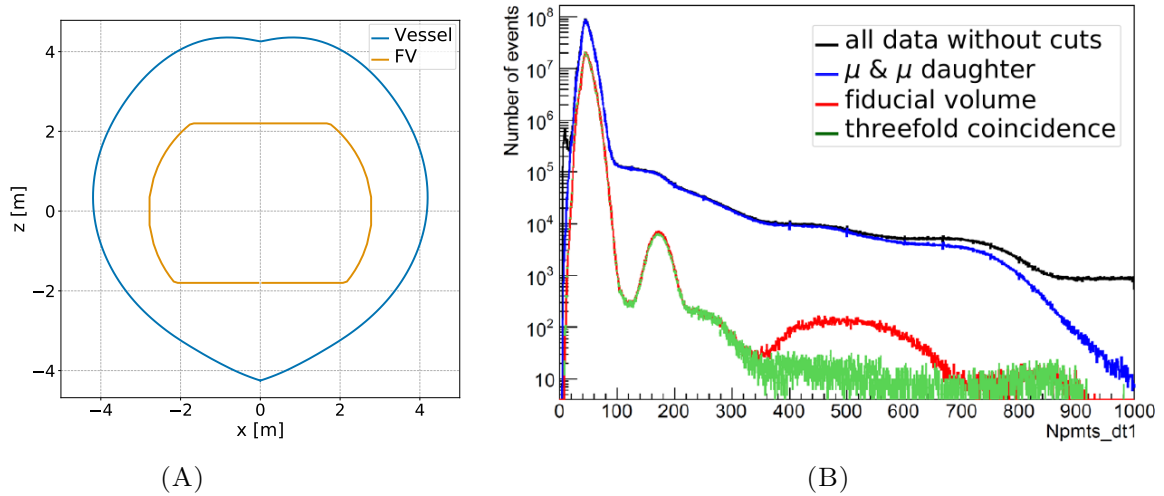


FIGURE 7.1: (A) FV shape (orange) compared to the Borexino vessel (blue). (B) Phase-II energy spectrum of all Borexino internal events (black) and the spectra after applying the muon and muon daughter cut (blue), FV cut (red) and TFC cut (green) in this order.

## 7.1.2 Fitting method

### Energy fit

To understand the multivariate fit technique, I first have to demonstrate the principle of a simple fit of a single energy spectrum. The basic idea of the fit is to take the reference shapes of all the contributing components (example in Fig. 7.2A) and fit them to the energy spectrum of the Borexino events from the chosen data sample selected as described in Sec. 7.1.1, using an energy estimator of choice:  $N_p$ ,  $N_p^{dt1(2)}$ ,  $N_h$ , or  $N_{pe}$  (more in Sec. 3.4.3). The goal of the fit is to

find the values of the weights assigned to each spectral component that result in the combination of these shapes that best describes the data. To achieve this, a likelihood function based on the set of parameters  $\vec{\theta}$  that define the reference shapes is constructed and a fit procedure is written to find the parameters that maximize it. Here is an example of a standard Poissonian likelihood used in the energy fit:

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^N \frac{\lambda_i(\vec{\theta})^{k_i}}{k_i!} e^{-\lambda_i(\vec{\theta})}. \quad (7.1)$$

The product goes over each bin  $i$  of the  $N$  bins in the spectrum, and  $k_i$  represents the number of events falling into that bin. The parameter  $\lambda_i$  is the expected value for the bin  $i$  based on the weighted reference shape constructed using AA or MCA. The set of variables  $\vec{\theta}$  depends on the approach: for MCA it is only the aforementioned weights, while AA includes other parameters of choice defining the reference shape (more in Sec. 7.1.3). The resulting values of the weights that produce maximum likelihood are used to calculate the rates of backgrounds and signals expressed in counts per day per 100 t of the scintillator (cpd/100 t), accounting for the exposure of the used data sample. An example of a fit done on the TFC-subtracted energy spectrum using the MC approach is shown in Fig. 7.2B.

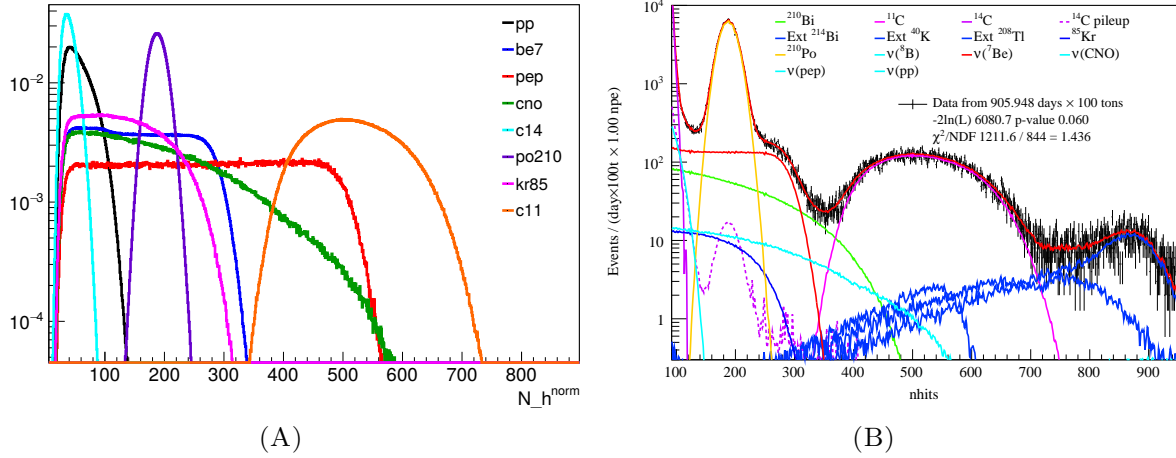


FIGURE 7.2: (A) Reference shapes of solar neutrino and several main internal background components constructed for the  $N_h$  energy estimator (number of hits) with the MC approach (area normalized to 1). (B) Example of an energy fit using the MC shapes.

### Multivariate likelihood

The principle of the MV fit is to simultaneously use multiple likelihood functions that depend on the same parameters but describe different spectra. In the Borexino LER analysis, four spectra contribute to the MV likelihood: the **TFC-tagged** and **TFC-subtracted** energy spectra, rich and depleted in  $^{11}\text{C}$ ; the **radial distribution** of the selected neutrino-like events, as well as the so-called **pulse shape** variable distribution. The meaning of these components will be discussed below.

The resulting total likelihood function is simply a product of the functions corresponding to each spectrum:

$$\mathcal{L}_{\text{MV}}(\vec{\theta}) = \mathcal{L}_{\text{TFC-tag}}(\vec{\theta}) \cdot \mathcal{L}_{\text{TFC-sub}}(\vec{\theta}) \cdot \mathcal{L}_{\text{PS}}(\vec{\theta}) \cdot \mathcal{L}_{\text{RD}}(\vec{\theta}). \quad (7.2)$$

The new TFC technique, described in detail in Sec. 4.5.3, has an efficiency of 92%, which allows us to divide the whole exposure into two sets, as mentioned above:  $\sim 36\%$  of exposure containing the majority of  $^{11}\text{C}$  (**TFC-tagged**), and  $\sim 64\%$  containing the remaining 8% of  $^{11}\text{C}$  events (**TFC-subtracted**). The likelihood functions for  $\mathcal{L}_{\text{TFC-tag}}$  and  $\mathcal{L}_{\text{TFC-sub}}$  in Eq. 7.2 are simple Poissonian likelihoods as in Eq. 7.1, representing the counts in each bin of each energy spectrum. Examples of these spectra and the fit results are shown in Fig. 7.3A and Fig. 7.3B, respectively..

An example spectrum and fit of the **radial distribution** (RD) can be seen in Fig. 7.3D. The goal of this fit is to separate the homogeneous internal, and external event contributions. There are two reference shapes in this fit: that of the internal uniform component (internal background distributed uniformly within the IV), and that of the external component (events originating on the surface of the IV).

The **pulse shape** variable, denoted as  $\text{PS-}\mathcal{L}_{\text{PR}}$ , is a parameter tuned to distinguish electron and positron events, shown in Fig. 3.11A and discussed in detail in Sec. 3.4.5. It is used to disentangle positron events originating from the decays of the remaining  $^{11}\text{C}$  in the TFC-subtracted spectrum. An example fit is shown in Fig. 7.3C. Two reference shapes for electrons and positrons are used to fit the data.

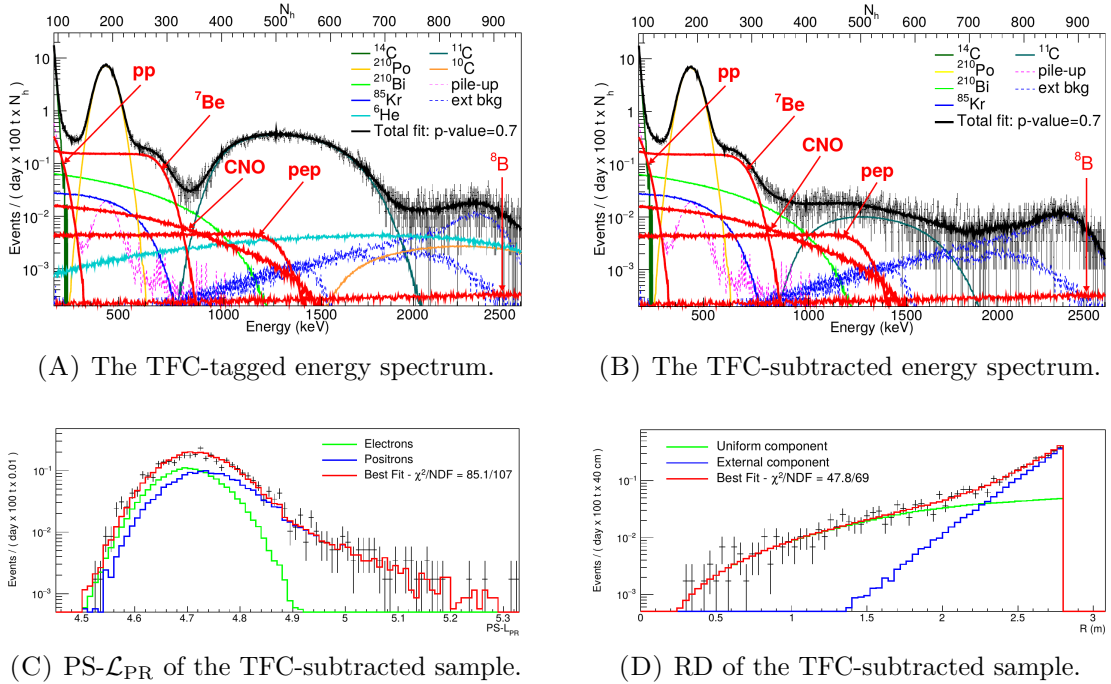


FIGURE 7.3: Distributions and fits of the components of the MV fit using the MCA. In each subfigure, black points with error bars show the data sample. In (A) and (B), the total fit is represented by a black curve, while in (C) and (D) by a red one [115].

The likelihood function describing both  $\mathcal{L}_{\text{PS}}$  and  $\mathcal{L}_{\text{RD}}$  has the following form:

$$\mathcal{L}_{\text{PS/RD}}(\vec{\theta}) = \prod_{i=1}^{N_{\text{PS/RD}}} \frac{a\lambda_i^{k_i}(\vec{\theta})e^{-a\lambda_i(\vec{\theta})}}{k_i!}. \quad (7.3)$$

In Eq. 7.3,  $a$  is a scaling parameter obtained through normalization to the total number of entries in the data histogram  $N_{entries}$ :

$$N_{entries} = a \sum_{i=1}^{N_{PS/RD}} \lambda_i(\vec{\theta}). \quad (7.4)$$

This normalization is introduced since there is a correlation between the number of entries in different histograms, since the events contributing to the energy spectrum also contribute to  $\mathcal{L}_{PS/RD}$ .

In the newer version of MCA, a 3D likelihood function, constructed for TFC-subtracted and TFC-tagged data sets, is defined as:

$$\mathcal{L}_{3D}(\vec{\theta}) = \mathcal{L}_{sub}(\vec{\theta}) \cdot \mathcal{L}_{tag}(\vec{\theta}), \quad (7.5)$$

where

$$\mathcal{L}_{sub}(\vec{\theta}) = \prod_{i,j,l}^{N_{E,RD,PS}} \frac{\lambda_{ijl}^{k_{ijl}}(\vec{\theta})}{k_{ijl}!} e^{-\lambda_{ijl}(\vec{\theta})}, \quad (7.6)$$

and

$$\mathcal{L}_{tag}(\vec{\theta}) = \prod_{i,j}^{N_{E,RD}} \frac{\lambda_{ij}^{k_{ij}}(\vec{\theta})}{k_{ij}!} e^{-\lambda_{ij}(\vec{\theta})}. \quad (7.7)$$

Same as in the approximated version of this likelihood in Eq. 7.2,  $k_{ijl}$  represents the data points falling into the three-dimensional bin  $ijl$ , where  $i$ ,  $j$ , and  $l$  are the bins in the distribution of energy, position, the pulse shape parameter, respectively. Note that the difference between Eq. 7.6 and Eq. 7.7 is the latter one does not consider the pulse shape parameter, since it is used only in  $\mathcal{L}_{sub}(\vec{\theta})$  to disentangle the remaining  $^{11}\text{C}$  contribution.

### Fit configuration

Depending on the analysis or approach, the parameters of the MV likelihood from Eq. 7.2 that govern the rate of background and signal components may be free, constrained or fixed. If the rate is **free**, the fit procedure finds the best value in a given allowed range for that rate that maximizes the likelihood. If the rate is **fixed**, it is not treated as a variable in the fit but is simply a constant. If it is **constrained**, a so-called penalty term is added to the likelihood to give preference to a certain value. Here is an example of a Gaussian penalty term:

$$\mathcal{L}(\vec{\theta}) = \mathcal{L}_0(\vec{\theta}) \frac{1}{2\pi\sigma} e^{\frac{-(p-\mu)^2}{2\sigma^2}}, \quad (7.8)$$

where  $\mathcal{L}_0$  is the original likelihood function e.g. as in Eq. 7.1,  $p$  is the value of the parameter of interest, while  $\mu$  and  $\sigma$  represent the constraint value and error, obtained from independent studies. In some cases Poissonian penalty is used.

In order to extract the  $pp$ ,  $pep$  and  $^7\text{Be}$  neutrino rates, the CNO neutrino rate in the fit is constrained based on the HM-SSM, and then separately the LM-SSM predictions (see Sec. 2.3 for solar models and solar metallicity). The two different constraints do not influence the resulting neutrino rates, except for  $pep$  neutrinos, the rates for which will be listed separately in Table 7.2.

In both cases, the rate of  $^8\text{B}$  neutrinos is fixed to the value obtained from the HER analysis, as their contribution to the LER is very small. The rate of  $^{14}\text{C}$  is constrained based on the rate obtained with the second cluster approach, described in Sec. 4.4.1. The inclusion of the constraint on the rate of  $^{85}\text{Kr}$ , estimated independently from the fit as described in Sec. 4.4.2, is considered in the calculation of the systematic uncertainties (Sec. 7.1.4).

### 7.1.3 Analytical and Monte Carlo approaches

#### Construction of reference shapes

The goal of modeling the detector analytically is to obtain a formula which would convert the deposited energy of the particle to an energy estimator. The first step of the AA is to obtain the number of photoelectrons ( $N_{pe}$ ) produced in the PMTs after an event of a certain energy. The next step is to convert  $N_{pe}$  to the number of fired PMTs ( $N_p$ ). The detailed description of the analytical formulas that do this conversion can be found in App. C. Using Eqs. C.6 and C.7, one can start from the theoretical spectra of recoiled electrons resulting from scattering with solar neutrinos, as well as background components, to construct the corresponding distributions in a given energy estimator used to fit the data energy spectra.

The benefit of using the AA is that the resulting spectral shape depends on various parameters representing the light yield, the Cherenkov radiation contribution, and other detector response parameters which are described in detail in Table C.1. These parameters describe properties of the detector that may vary in time, and can be left free to vary in the fit, allowing us to follow their changes, unlike the MCA, where parameters like LY have to be measured independently and fixed at the level of  $g4bx2$ . However, having more free parameters in the fit has the drawback of introducing more correlations between various parameters.

To construct reference shapes using the MCA, `g4bx2` is used to model all the physical processes from energy deployment, and `bx_elec` is used to simulate the behavior of the Borexino electronics (see Sec. 3.5.2), which includes detector properties like effective quantum efficiency of the PMTs discussed in Chapter 6. The disadvantages of the MCA include the inability to simulate a sufficient amount of  $^{14}\text{C}$  events due to their abundance in the data; while the biggest advantage is the fact that the simulation is based on the calibration data, completely independent from the data used in this analysis.

#### Fit configuration

The following are the main differences between AA and MCA regarding the fitting procedure:

- **likelihood function:** in the AA (as well as the earlier version of MCA) the approximated MV likelihood function in Eq. 7.2 is used, while MCA implements the 3D likelihood from Eq. 7.5;
- **likelihood parameters:** the AA has more parameters  $\vec{\theta}$  contributing to the likelihood mentioned above, including detector response parameters (Table C.1);
- **pileup modeling:** there are three main methods of including the pileup effect into the picture. Two methods are used in the AA:
  - 1) convolution of the spectra of all fit components with untriggered data spectrum; and

2) construction of the so-called synthetic pileup.

One method is used in the MCA: 3) using the Borexino MC to construct the spectrum of pileup events. More details on these methods in Sec. 4.6. The difference between methods (1) and (2) contributes to the systematic uncertainties. In methods (2) and (3), the pileup is effectively modeled as a separate background component, and the rate of the component has to be constrained in the fit. Fits with pileup modeling approaches (1), (2) and (3) are demonstrated in Fig. 7.4A, Fig. 7.4B, and Fig. 7.3B, respectively, using the TFC-subtracted spectra as an example.

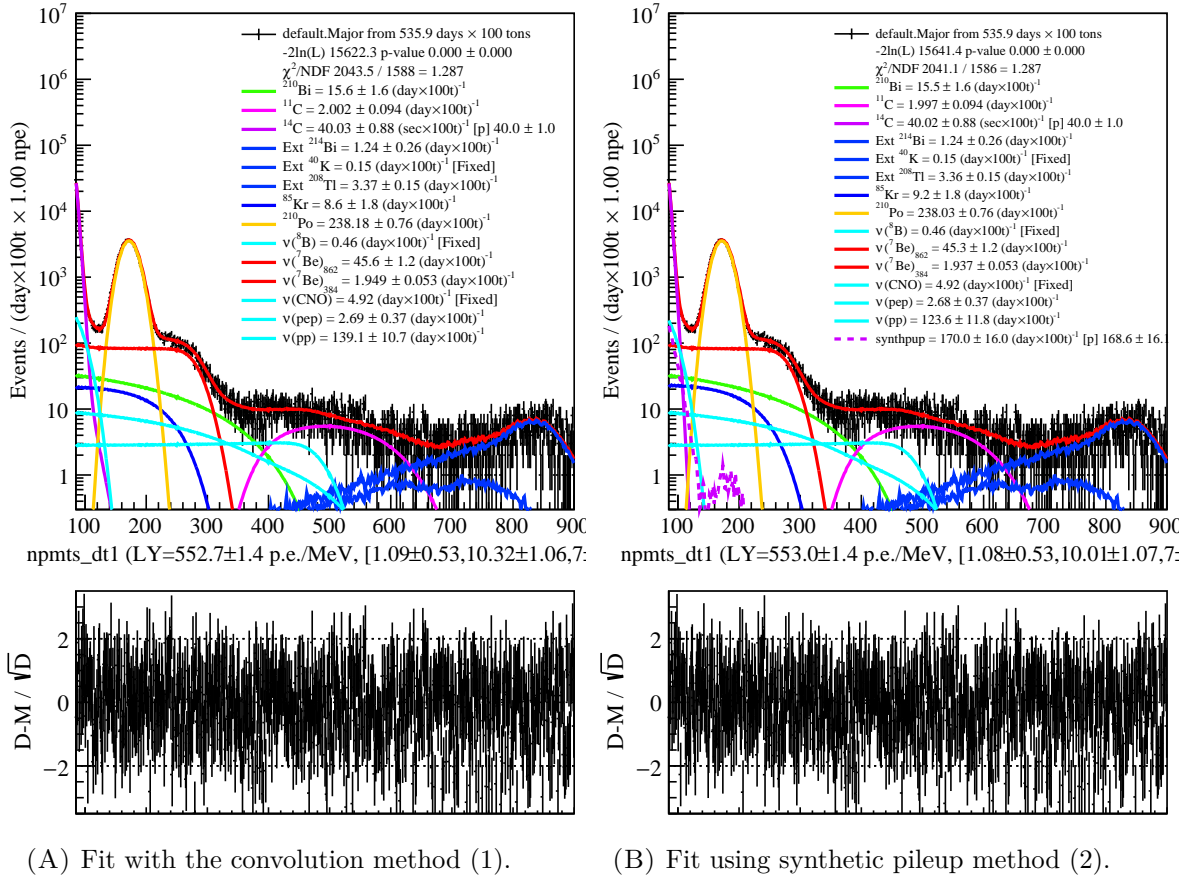


FIGURE 7.4: The effect of the dark noise convolution in (A) can be best seen through the deformation the  $^{14}\text{C}$  spectral shape (purple) compared to the one in (B), where synthetic pileup is shown by the purple dashed curve.

- **Energy estimator:** The fit using the MCA is done using the  $N_h$  energy estimator, while the MCA uses  $N_p^{dt1}$  and  $N_p^{dt2}$ , the differences between which contribute to the systematic error.
- **Fit range:** The fit with the MCA starts from  $N_h = 92$ , while the fit using the AA starts from  $N_p^{dt1(2)} = 85$  (in this energy range, 1 hit  $\approx$  1 fired PMT). The higher energy range of MCA is due to the inability of MC to simulate  $^{14}\text{C}$  very well.

### 7.1.4 Systematic uncertainties

An extensive study of the systematic uncertainties has been performed. The relevant sources of uncertainty and their contributions are summarized in Table 7.1. For this analysis, I have created a Python based framework to perform numerous fits with different conditions and collect the resulting values. My main contribution to this study lies in performing and analyzing these fits to obtain the average result for the rates of different components, and calculate uncertainties related to the fitting procedure.

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

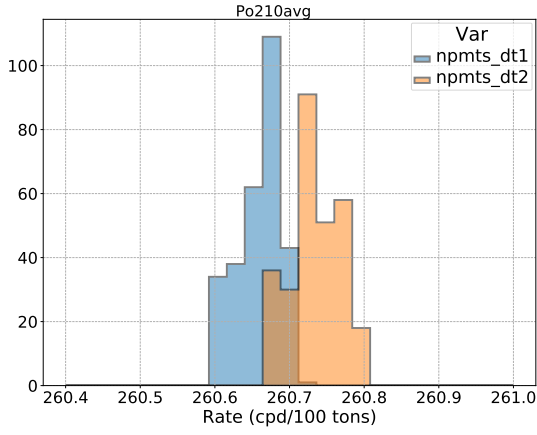
TABLE 7.1: Relevant sources of systematic uncertainties and their contributions relative to the measured neutrino rates [115].

A short summary of the uncertainty sources is presented below, the ones to which I contributed directly marked with stars.

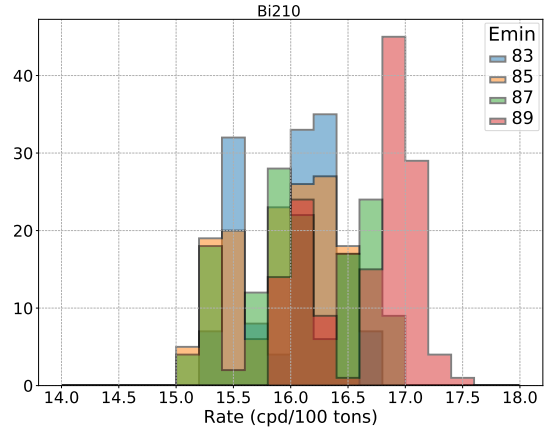
- **Fit method\***: the differences between the results based on the analytical and Monte Carlo approaches described in Sec. 7.1.3.
- **Choice of energy estimator\***: differences introduced by using  $N_p^{dt1}$  or  $N_p^{dt2}$  as the energy variable in the AA fit. Example distributions demonstrating these differences are shown in Fig. 7.5A.
- **Pileup modeling**: differences introduced by methods (1) and (2) of modeling pileup in the AA that were described above in Sec. 7.1.3.
- **Fit range and binning\***: differences coming from small deviations in the ranges and binning of the histograms of the different MV fit components (see Fig. 7.3). Example distributions can be seen in Fig. 7.5B and Fig. 7.5C.



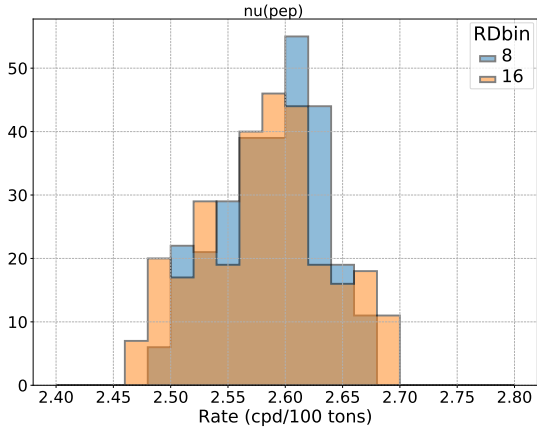
- **Fit models:** uncertainties originating from the imprecision in detector response modeling. In order to study this effect, the fit has been performed using shapes with small nonlinear deformations within uncertainties based on calibration data, to test its robustness to possible lack of knowledge about the detector response function.
- **Inclusion of  $^{85}\text{Kr}$  constraint\*:** differences that we observe when the  $^{85}\text{Kr}$  rate in the fit is free or constrained based on the upper limit described in Sec. 4.4.2 (Fig. 7.5D).
- **Exposure related uncertainties:** these include the uncertainty on live time, LS density, and FV (coming from position reconstruction).



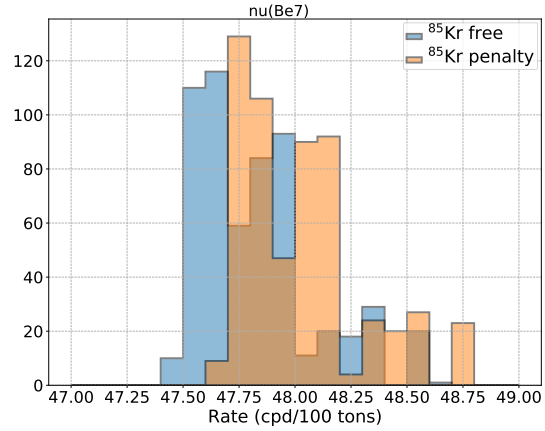
(A)  $^{210}\text{Po}$  rate results from fits with different energy estimators.



(B)  $^{210}\text{Bi}$  rate results with different minimum energy of the fit.



(C) The rate of *pep* neutrinos resulting from fits with different bin width of the RD histogram.



(D)  $^7\text{Be}$  neutrino rate resulting from fits with free (blue) and constrained (orange)  $^{85}\text{Kr}$  rate.

FIGURE 7.5: Distributions of rates of different fit components resulting from numerous fits with different settings.



## 7.2 High energy region analysis

As mentioned at the beginning of Chapter 7, the Borexino solar neutrino analysis is divided into two energy regions, and the analysis of the high energy region (HER) results in the measurement of the rate of  $^8\text{B}$  neutrinos, and an upper limit on the *hep* neutrino rate [119]. The HER analysis covers the energy range of 3.2-16 MeV. In this section I will briefly describe the approach of this method, focusing on the *hep* neutrino counting analysis in which I took part. The results of the HER analysis will be presented in Sec. 7.3 together with those of the LER analysis.

Unlike LER, HER is not affected by the long-lived radioactive background, which allows us to use a larger dataset between January 2008 and December 2016, covering both Phase-I and Phase-II. The HER itself is divided into two subregions, HER-I and HER-II, below and above 5.7 MeV. The main backgrounds in HER-I and HER-II are natural radioactivity and external gamma rays following neutron capture processes on the SSS, respectively. The value of 5.7 MeV has been chosen this way because the natural radioactive background never exceeds 5 MeV (the Q value of the  $\beta$  decay of  $^{208}\text{Tl}$ ).

The HER-I analysis requires a *z*-cut, which will be described in Sec. 7.2.1 below, resulting in the total mass of 227.8 t, while HER-II uses the full IV volume of 266 t. This results in the total exposure of 2062.4 days  $\times$  227.8 t (266.0 t) for HER-I (HER-II).

### 7.2.1 Data selection

The data selection method in the HER analysis is similar to the one of LER described in Sec. 7.1.1 and consists of applying selection cuts to remove unwanted events. The main difference comes from the fact that LER and HER are affected by different types of backgrounds which are dominant in the corresponding energy ranges. The selection cuts of the HER analysis consist of the:

- **Muon cut:** same as in LER (Sec. 7.1.1).
- **Neutron cut:** remove cosmogenic neutron captures on  $^{12}\text{C}$  by applying a 2 ms veto after all muons.
- **Fast cosmogenics cut:** remove cosmogenic isotopes by applying a 6 s veto after each internal muon.
- **Run start/break cut:** remove fast cosmogenic isotopes from potential muons missed during run restart.
- **$^{10}\text{C}$  cut:** remove cosmogenic  $^{10}\text{C}$  by vetoing a spherical volume of radius 0.8 m around all muon-induced neutron captures.
- **$^{\text{c}}\text{Bi}$ - $^{\text{c}}\text{Po}$  cut:** the same as in LER (Sec. 7.1.1).
- ***z*-cut:**  $z < 2.5$  m to suppress the background events in the top layer of the LS resulting from a small pinhole in the IV; applied only in HER-I.

### 7.2.2 Fit technique of the $^8\text{B}$ neutrino analysis

The  $^8\text{B}$  neutrinos are the only neutrinos (apart from the extremely rare *hep* process) in the HER as can be seen from Fig. 2.4, and there are much fewer sources of background contributions than in LER, especially in HER-II where only neutron captures contribute. Most of the background being external makes it possible to perform a fit without needing the  $^8\text{B}$  neutrino energy spectrum: only the radial distribution of events is used. This makes it possible to conduct this analysis without any assumption on the energy of  $^8\text{B}$  neutrinos, and therefore, the neutrino oscillation mechanism. This allows us to probe for deviations from the MSW-LMA theory and nonstandard interactions.

The fit procedure is similar to the one of the LER analysis described in Sec. 7.1.2, but it uses only a simple likelihood based on the radial distribution of events. The HER analysis fit is shown in Fig. 7.6.

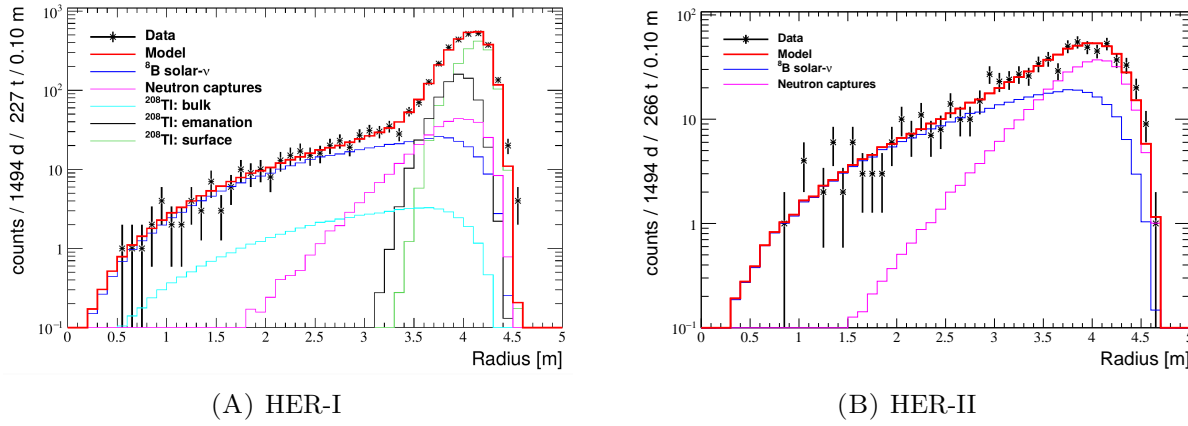


FIGURE 7.6: The data sample (black stars), fit results (red) and reference shapes (see legends) of the HER-I and HER-II analyses [119].

### 7.2.3 The *hep* neutrino counting analysis

As mentioned above, due to the scarcity of the *hep* process depicted in Fig. 2.1A, it is not possible to perform a complex study of *hep* neutrinos using a spectral fit similar to the LER and HER analyses. The only possible approach is that of a simple counting analysis that allows us to set an upper limit on the *hep* neutrinos which will be reported in Sec. 7.3 together with the other results from the LER and HER analyses.

This analysis has been done in the energy range 11–20 MeV using the Borexino FADC DAQ system (see Sec. 3.3.4) as well as the Laben DAQ system (Sec. 3.3.1) to perform a cross-check. Due to the overefficiency of muon tagging at high energies in both Laben and FADC, only few events are expected in the whole statistics. The main backgrounds of this analysis are cosmogenics and atmospheric neutrinos.

In order to select potential candidates, the following conditions are applied:

- a 6.5 s veto after internal muons,
- a 2 ms veto after external muons,
- $r < R_{IV}(\theta, \phi) - 0.25$  m, where  $r$  is the radial position of the candidate event, and  $R_{IV}(\theta, \phi)$  is the IV radius in the direction of the event,
- energy threshold  $> 11$  MeV.

My contribution to this analysis consists in performing the aforementioned cross-check of the FADC results using the Laben system for muon tagging. The additional selection conditions applied in the Laben analysis are the event type conditions TT1 BTB0 CL1 to select neutrino-like events. The energy condition in Laben corresponding to  $E > 11$  MeV is  $Q > 5000$  p.e.

The two systems both use the OD information, but rely on different sources for muon tagging, namely, FADC pulse shape and ID pulse shape for the FADC and Laben system, respectively. They also have different gate length,  $1 \mu\text{s}$  and  $16 \mu\text{s}$ , respectively. Because of this they have independent inefficiencies and malfunctions, and can complement each other in terms of muon tagging. The counting analysis I performed based on Laben has confirmed 10 out of the 12 *hep* neutrino candidates found by the FADC analysis, rejecting two that were identified as muons; and finding an additional one, resulting in the total of 11 candidates.

The limit on the *hep* neutrino flux is set assuming their elastic scattering on electrons and neutral current mediated inelastic scattering on carbon nuclei:

$$\Phi = \frac{S}{(N_e \sigma_e + N_C \sigma_C) T \epsilon} \text{cm}^{-2} \text{s}^{-1}, \quad (7.9)$$

where

$S = 3.95$  is the maximum allowed number of events at 90% C.L. for  $N_{hep} = 11$  and number of expected background events  $N_{bkg} = 13$  in 11-20 MeV,

$N_e \sigma_e + N_C \sigma_C = 2.441 \cdot 10^{-13} \text{cm}^2$  is the total target cross section with  $N_e(\sigma_e)$  and  $N_C(\sigma_C)$  being the total number (cross section) of the electrons and carbon nuclei in the LS, respectively,

$T = 1259$  days is the combined livetime of the FADC and Laben systems,

and  $\epsilon = 0.928$  is the total detection efficiency in 11-20 MeV obtained from MC.

The resulting *hep* neutrino limit is reported in Table 7.2 together with other *pp*-chain neutrino results.

A more sophisticated analysis was performed following this study [119]. It was done exploiting the data selection cuts described in Sec. 7.2.1 and using Monte Carlo simulation. The analysis found total of 10 candidates out of  $12.8 \pm 2.3$  expected, demonstrated in Fig. 7.7, compatible with the simple analysis conducted in my work.

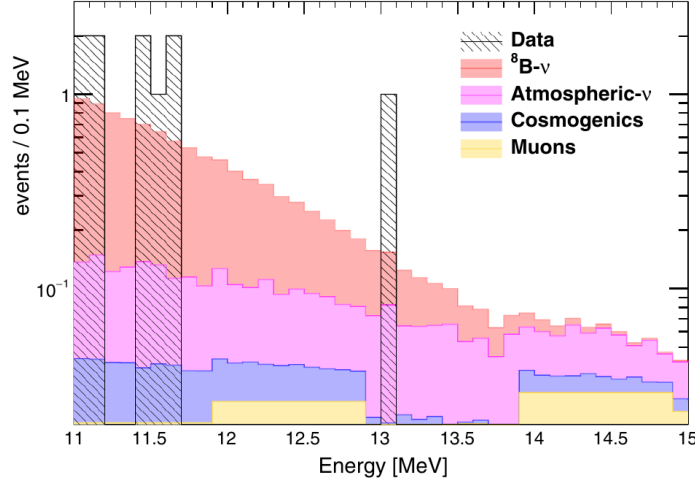


FIGURE 7.7: FADC energy spectrum of selected events above 11 MeV, compared with the expected background spectrum [119].

### 7.3 Results and Implications

The results of the HER and LER analyses, the details of which I have discussed in Sec. 7.1 and Sec. 7.2, respectively, are presented in Table 7.2.

Solar $\nu$	Rate (cpd/100 t)	Flux ( $\text{cm}^{-2}\text{s}^{-1}$ )	Flux SSM predictions ( $\text{cm}^{-2}\text{s}^{-1}$ )	
$pp$	$134 \pm 10^{+6}_{-10}$	$(6.1 \pm 0.5^{+0.3}_{-0.5}) \cdot 10^{10}$	$5.98(1.0 \pm 0.006) \cdot 10^{10}$	(HM)
			$6.03(1.0 \pm 0.005) \cdot 10^{10}$	(LM)
${}^7\text{Be}$	$48.3 \pm 1.1^{+0.4}_{-0.7}$	$(4.99 \pm 0.11^{+0.06}_{-0.08}) \cdot 10^9$	$4.93(1.0 \pm 0.006) \cdot 10^9$	(HM)
			$4.50(1.0 \pm 0.006) \cdot 10^9$	(LM)
$pep$ (HM)	$2.43 \pm 0.36^{+0.15}_{-0.22}$	$(1.27 \pm 0.19^{+0.08}_{-0.12}) \cdot 10^8$	$1.44(1.0 \pm 0.009) \cdot 10^8$	(HM)
			$1.46(1.0 \pm 0.009) \cdot 10^8$	(LM)
$pep$ (LM)	$2.65 \pm 0.36^{+0.15}_{-0.24}$	$(1.39 \pm 0.19^{+0.08}_{-0.13}) \cdot 10^8$	$1.44(1.0 \pm 0.009) \cdot 10^8$	(HM)
			$1.46(1.0 \pm 0.009) \cdot 10^8$	(LM)
${}^8\text{B}_{\text{HER-I}}$	$0.136^{+0.013+0.003}_{-0.013-0.003}$	$(5.77^{+0.56+0.15}_{-0.56-0.15}) \cdot 10^6$	$5.46(1.0 \pm 0.12) \cdot 10^6$	(HM)
			$4.50(1.0 \pm 0.12) \cdot 10^6$	(LM)
${}^8\text{B}_{\text{HER-II}}$	$0.087^{+0.080+0.005}_{-0.010-0.005}$	$(5.56^{+0.52+0.33}_{-0.64-0.33}) \cdot 10^6$	$5.46(1.0 \pm 0.12) \cdot 10^6$	(HM)
			$4.50(1.0 \pm 0.12) \cdot 10^6$	(LM)
${}^8\text{B}_{\text{HER}}$	$0.223^{+0.015+0.006}_{-0.016-0.006}$	$(5.68^{+0.39+0.03}_{-0.41-0.03}) \cdot 10^6$	$5.46(1.0 \pm 0.12) \cdot 10^6$	(HM)
			$4.50(1.0 \pm 0.12) \cdot 10^6$	(LM)
$hep$	$< 0.002$ (90% C.L.)	$< 2.2 \cdot 10^5$ (90% C.L.)	$7.98(1.0 \pm 0.30) \cdot 10^3$	(HM)
			$8.25(1.0 \pm 0.12) \cdot 10^3$	(LM)

TABLE 7.2: Results of the Borexino Phase-II solar neutrino analysis. The rates and fluxes are integral values without any threshold [72, 115].

As mentioned before, the assumption on the low or high metallicity model for the constraint of the CNO neutrino rate influences only the *pep* neutrino rate results which are reported separately. Both models yield a result of  $> 5\sigma$  significance, as shown later in Fig. 7.8. The third column of Table 7.2 translates the rates in cpd/100t of neutrino events in the scintillator to fluxes in  $\text{cm}^{-2}\text{s}^{-1}$  using the flavor composition predicted by the MSW-LMA phenomenon [120] and the well known electron, muon and tau neutrino cross sections [109]. The Phase-II results on the CNO neutrino rate upper limit will be discussed in Chapter 8, together with further studies towards the CNO neutrino detection with Borexino Phase-III data.

In order to perform a test for signal discovery, the so called **likelihood profile** of a neutrino rate is constructed. The fit is performed several times with the neutrino rate fixed to different values. From this ensemble of fits, the  $\Delta\chi^2$  parameter is calculated based on the resulting maximized likelihood as  $\Delta\chi^2 = -2\ln \Delta\mathcal{L}$ , where  $\Delta\mathcal{L}$  is the difference between the likelihood of the given fit, and the lowest likelihood in the fit ensemble.

An example of a profile likelihood scan as a function of *pep* neutrino rate is shown in Fig. 7.8. As one can see, the null hypothesis is rejected with a significance of  $\sim 6.5\sigma$  and  $\sim 7.1\sigma$  for HM- and LM-SSM, respectively.

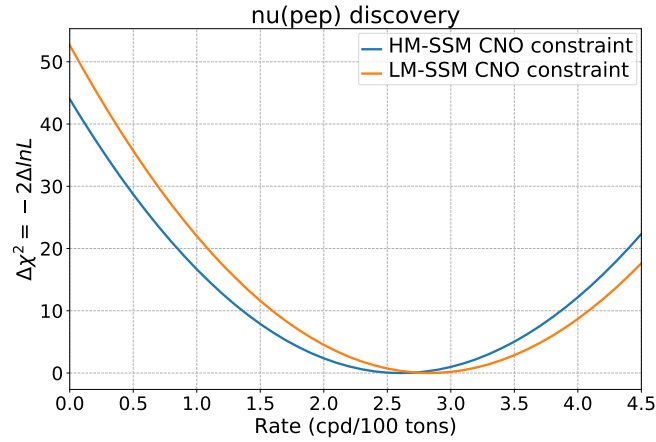


FIGURE 7.8: Scan of the profile likelihood as a function of the *pep* neutrino rate, resulting from the MV fit of the Phase-II data.

Using the well known models of neutrino oscillation, we can test the predictions of the LM- and HM-SSMs, which are shown in the last column of Table 7.2; and vice versa, by taking assumptions of SMMs, we can probe the MSW-LMA model (see Sec. 1.4.2). Based on these results, one can draw several conclusions in terms of solar and neutrino physics which are summarized below.

### 7.3.1 Thermodynamic equilibrium of the Sun

Using the neutrino rates from Table 7.2 and the nuclear reactions from Fig. 2.1A, one can calculate the total power to be  $L = (3.89 \pm 0.42) \cdot 10^{33} \text{ erg s}^{-1}$  which is compatible with what we observe from the photon output,  $L = (3.846 \pm 0.015) \cdot 10^{33} \text{ erg s}^{-1}$  (more in Sec. 2.4.1).

This finding is an experimental confirmation of the nuclear origin of the solar power. Since Borexino gives a real-time picture of the core of the Sun, and considering that the time it takes

for radiation to flow from the energy-producing region to the surface of the Sun is of the order of  $10^5$  years, the finding also proves the thermodynamic equilibrium of the Sun over this time scale. Moreover, this latest Borexino result holds the best precision obtained by a single solar neutrino experiment.

### 7.3.2 Probing solar metallicity

As discussed in Sec. 2.4.3, the fluxes of  $^7\text{Be}$  and  $^8\text{B}$  neutrinos are the most influenced by the solar metallicity assumed by LM- and HM-SSMs. Combining the LER and HER results on  $^7\text{Be}$  and  $^8\text{B}$  neutrino fluxes, one can construct an allowed region in the parameter space of these two variables based on the best fit value and the uncertainties. As can be seen from Fig. 7.9, the Borexino measurement gives a weak hint that the temperature profile in the Sun is more compatible with HM-SSM. In a global fit together with all solar experiments and KamLAND, the hint further weakens. In this fit, the oscillation parameters  $\theta_{12}$  and  $\Delta m_{12}^2$  are left free to vary, and the results returned by the fit are consistent with the ones obtained in [120].

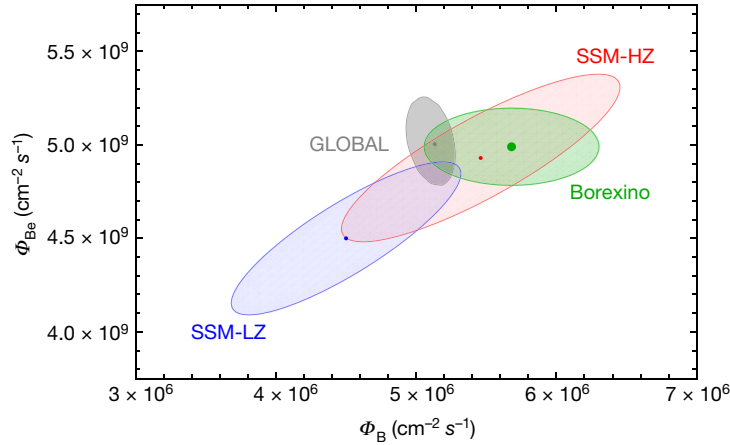


FIGURE 7.9: Allowed  $1\sigma$  theoretical regions of the  $^8\text{B}$  and  $^7\text{Be}$  neutrino fluxes according to LM-SSM (blue) and HM-SSM (red), compared to the Borexino measurement (green) and a global fit of all solar neutrino experiments and KamLAND (grey) [72].

### 7.3.3 Studying the $pp$ chain terminations

Using the LER results on  $pp$  and  $^7\text{Be}$  neutrinos, one can calculate the relative intensity of the  $pp$ -I and  $pp$ -II terminations of the  $pp$  chain, depicted in Fig. 2.1A:

$$R_{\text{II/I}} = \frac{2\phi(^7\text{Be})}{\phi(pp) - \phi(^7\text{Be})}, \quad (7.10)$$

where  $\phi$  denotes the neutrino flux of the given type (more in Sec. 2.4.3). The theoretical predictions are  $R_{\text{II/I}} = 0.180 \pm 0.011$  and  $R_{\text{II/I}} = 0.161 \pm 0.010$  for HM- and LM-SSM, respectively, as discussed in Sec. 2.4.4. The experimental result is  $R_{\text{II/I}} = 0.178 \pm 0.027$ , which is compatible with the expected values.

### 7.3.4 Probing flavor transformation in matter

Apart from testing the validity of SSMs using our knowledge about neutrino physics, one can also perform the reverse study of probing neutrino flavor transformation in the matter of the Sun by assuming a certain solar model. Flavor transformation of solar neutrinos occurs via *matter-enhanced oscillation* based on the MSW-LMA theory, introduced in Sec. 1.4.2, and discussed in Sec. 2.4.5 in the context of solar neutrinos. Figure 7.10 demonstrates the electron neutrino survival probability  $P_{ee}$ , calculated based on the Borexino measurements of the interaction rates of different  $pp$ -chain neutrinos, and initial fluxes predicted by the HM-SSM. The Borexino results are represented by points, while the theoretical predictions according to the vacuum oscillation model and the MSW-LMA model are represented by the grey and the pink bands, respectively. Borexino is the only experiment that can probe  $P_{ee}$  in both vacuum and matter-effect-dominated regions. The results disfavor vacuum oscillation at 95% CL.

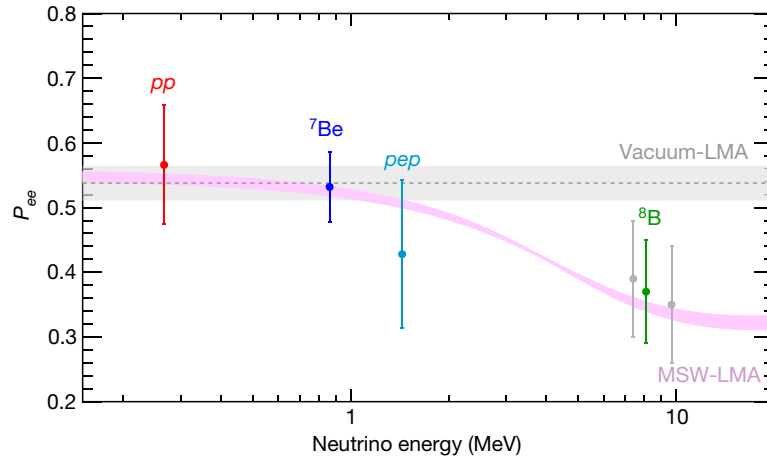


FIGURE 7.10: Electron survival probability  $P_{ee}$  in different energy regions, calculated using the Borexino measurements and the HM-SSM flux predictions, compared to the vacuum oscillation model (grey) and the MSW-LMA model (pink) [72].

Some theories of physics beyond the Standard Model (SM) consider the existence of non-standard interactions (NSIs) of neutrinos with other SM fermions [83]. Based on the Borexino Phase-II data, no indication of new physics was found at the level of sensitivity of the detector, and constraints on the NSI parameters were placed [93].





## Chapter 8

# First evidence of solar neutrinos from the CNO cycle

The carbon-nitrogen-oxygen (CNO) catalytic cycle, discussed in Sec. 2.2 and shown in Fig. 2.2, has never been observed until now, and is hypothesized to be the dominant hydrogen-to-helium fusion process in stars heavier than the Sun. The measurement of CNO neutrinos could bring insights into cosmology and the physics of heavier stars. It can also potentially solve the solar metallicity puzzle (see Sec. 2.3 for solar models and solar metallicity), since it is the most influenced by the abundance of carbon, nitrogen and oxygen, governed by this parameter. In this chapter, I will present my contribution to the CNO neutrino analysis with Borexino.

I will start with discussing the challenges of the CNO neutrino measurement in the Borexino experiment in Sec. 8.1. Then, in Sec. 8.2, I will present the results on the upper limit of the CNO neutrino flux, which we obtain from the low energy region (LER) analysis on Phase-II data, described in detail in Sec. 7.1.

I will proceed to the Phase-III analysis, aimed at observation of the CNO neutrinos, by discussing the strategy of performing such analysis in Sec. 8.3. The confirmation that this strategy works is shown by sensitivity studies in Sec. 8.4. In Sec. 8.5, I will discuss the multivariate fit of Phase-III data with the “high setting” mentioned in Sec. 6.7.3, using the Monte Carlo Approach and the Analytical Approach. In Sec. 8.6 I will elaborate on the studies of systematic effects on the CNO neutrino discovery power, including those originating from the fit parameters (Sec. 8.6.1), which is my main contribution to this analysis. I will report the results of this analysis and their implications in Sec. 8.7.

### 8.1 Challenges of the CNO neutrino detection

At the moment, all the neutrinos originating from the  $pp$  chain have been not only observed but also measured comprehensively [72]. However, evidence of the hypothesized CNO mechanism is still missing. Currently one of the main goals in the field of solar neutrinos is to probe its existence, as it carries implications in astrophysics and for Standard Solar Models. The Borexino experiment has the biggest potential to observe CNO neutrinos due to its sensitivity at low energies.

The main challenge of the CNO neutrino measurement in Borexino is the extreme similarity of the CNO neutrino spectral shape to that of  $^{210}\text{Bi}$ , and  $pep$  neutrinos, shown in Fig. 8.1. Due to this similarity, the three contributions cannot be disentangled in the multivariate (MV) fit of the low-energy solar analysis, described in Sec. 7.1.2.

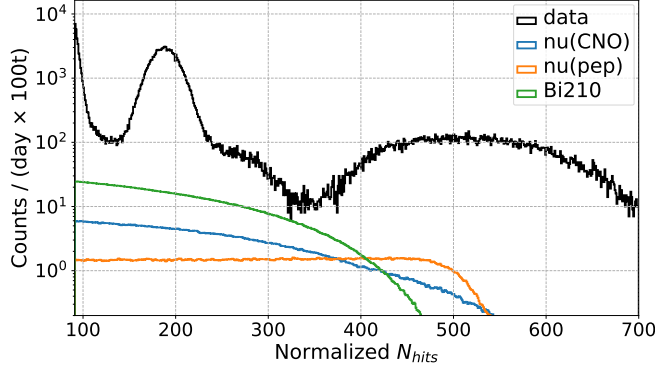


FIGURE 8.1: The Borexino neutrino data (black) and the reference spectral shapes of CNO (blue) and *pep* (orange) neutrinos, and  $^{210}\text{Bi}$  (green).

For this reason, only an upper limit has been established so far, and CNO neutrino measurement has never been possible yet. The Phase-II results on the CNO analysis, including the upper limit, will be discussed in Sec. 8.2 below.

## 8.2 Phase-II CNO analysis

As described in Sec. 7.1.2, in order to extract the *pp*, *pep*, and  $^7\text{Be}$  rates in the Phase-II LER analysis (Sec. 7.1), the CNO rate in the MV fit is constrained based on the HM-SSM, and then separately the LM-SSM predictions. Complementary to this analysis of the *pp*-chain neutrinos, we perform MV fits with CNO neutrino rate left free in the fit, with an indirect constraint placed on the ratio of *pp* and *pep* neutrino rates, predicted by HM-SSMs and LM-SSMs to be  $47.8 \pm 0.8$  and  $47.5 \pm 0.8$ , respectively (see Sec. 2.4.2). The rest of the fit settings is the same as in Sec. 7.1.2. It is, unfortunately, only possible to obtain an upper limit, and not a measurement, for the CNO neutrino rate based on Phase-II analysis. This fact is demonstrated in Fig. 8.2, where the scan of the profile likelihood depending on the CNO rate is shown.

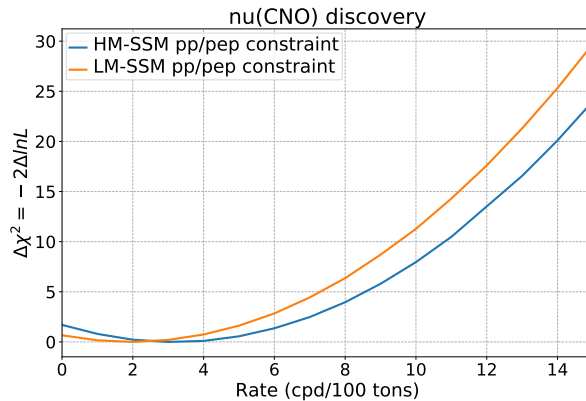


FIGURE 8.2: Scan of the profile likelihood as a function of the CNO neutrino rate, based on the MV fit on the Phase-II data with a constraint placed on the ratio of *pp* and *pep* neutrino fluxes based on the predictions of HM- (blue) and LM-SSM (orange).

As one can see, the  $\Delta\chi^2$  corresponding to the null hypothesis is less than  $1\sigma$  away from the result which gives us maximum likelihood. The resulting limit on the CNO neutrino rate and flux with 95% C.L. is reported in Table 8.1.

Solar $\nu$	Rate (cpd/100 t)	Flux ( $\text{cm}^{-2}\text{s}^{-1}$ )	Flux SSM predictions
CNO	$< 8.1$	$< 7.9 \cdot 10^8$	$4.88(1.0 \pm 0.11) \cdot 10^8$ (HM)
			$3.51(1.0 \pm 0.10) \cdot 10^8$ (LM)

TABLE 8.1: Upper limit on CNO neutrino rate and flux resulting from Phase-III LER analysis (95% C.L.), compared to the theoretical predictions. [67, 72]

A separate approach is being developed in order to measure the CNO rate using Phase-III data, and it will be described in Sec. 8.3 below.

### 8.3 CNO neutrino measurement strategy

The measurement of the CNO neutrino interaction rate in Borexino is based on two approaches:

- *rate+shape* analysis, employing the same multivariate fit technique that was used in the low-energy Phase-II analysis, described in detail in Sec. 7.1;
- *rate only* analysis, or *counting analysis*, that does not rely on the spectral information.

In both approaches, the knowledge of the rates of  $^{210}\text{Bi}$  and the *pep* neutrinos is crucial, as they present the largest background for CNO neutrinos, and, in case of the rate+shape analysis, due to the similarity of their spectral shapes, as discussed in Sec. 8.1.

It is now possible to determine an upper limit on the rate of  $^{210}\text{Bi}$ , thanks to the temperature stabilization in the detector after the thermal insulation campaign (Sec. 3.2.7) preceding the beginning of Phase-III. This is the main reason for the CNO neutrino analysis to be performed on Phase-III data. I will elaborate on the approach of determining the  $^{210}\text{Bi}$  upper limit, and its current results in Sec. 8.3.1. As for *pep* neutrinos, their interaction rate can be constrained based on the SSM prediction, which I will discuss in Sec. 8.3.2.

The MV fit of the Phase-III CNO analysis relies on the same technique, as described in Sec. 7.1.2; while the counting analysis will be discussed in Sec. 8.3.3. Both approaches employ the same data selection cuts as in Sec. 7.1.1.

#### 8.3.1 Upper limit on the rate of $^{210}\text{Bi}$

The rate of  $^{210}\text{Bi}$  can be estimated by measuring the events coming from its daughter,  $^{210}\text{Po}$ , in the decay chain of  $^{210}\text{Pb}$ :

$$^{210}\text{Pb} \xrightarrow{32\text{y}} ^{210}\text{Bi} \xrightarrow{7.23\text{d}} ^{210}\text{Po} \xrightarrow{199.1\text{d}} ^{206}\text{Pb}. \quad (8.1)$$

The values above the arrows correspond to the mean lifetimes of the parent isotopes. The evolution of the rate of  $^{210}\text{Po}$  events in time can be expressed as follows:

$$R_{\text{Po}}^{\text{tot}}(t) = R_{\text{Po}}^u e^{-t/\tau_{\text{Po}}} + R_{\text{Po}}^m(t) + R_{\text{Po}}^s, \quad (8.2)$$

where  $R_{\text{Po}}^u$  is the so-called *unsupported term*, corresponding to the  $^{210}\text{Po}$  contribution out of equilibrium with  $^{210}\text{Bi}$ , but uniformly distributed within the fiducial volume (FV); while  $R_{\text{Po}}^s$  is the *supported term*, providing a direct measurement of the rate of the parent  $^{210}\text{Bi}$  [121, 122]. The parameter  $R_{\text{Po}}^m(t)$  stands for a possible extra  $^{210}\text{Po}$  source due to *migration* via convection and diffusion. An important part of this estimation is to disentangle this migratory contamination of  $^{210}\text{Po}$ , brought to the FV by convection from the IV, discussed in Sec. 4.3, and graphically illustrated in Fig. 8.3A. This is made possible due to the detector thermal insulation campaign conducted in 2015 (Sec. 3.2.7). The resulting thermal stabilization can be seen in Fig. 8.3B, where the  $^{210}\text{Po}$  rate is measured in cubic volumes uniformly dividing the inner vessel.

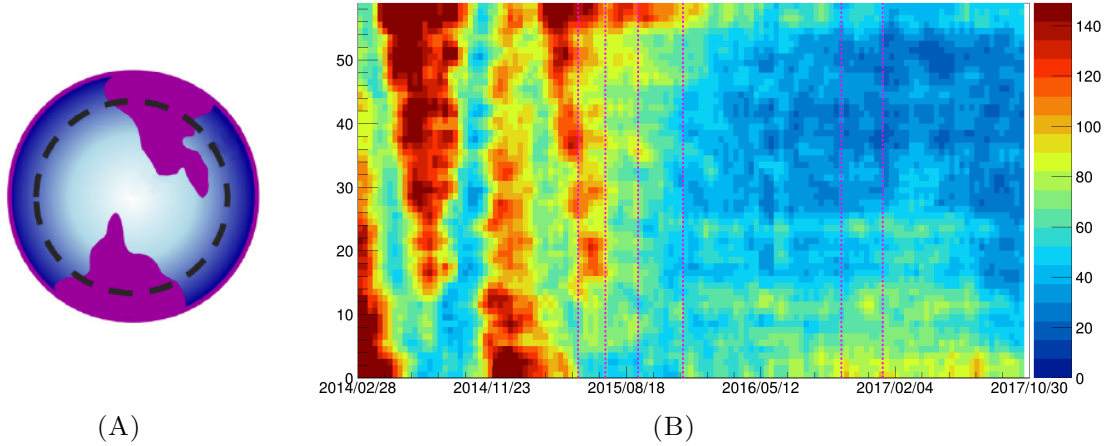


FIGURE 8.3: (A) Graphic illustration of  $^{210}\text{Po}$  contamination (purple) on the IV penetrating the inner volume (black dashed region) due to convection [112]. (B) The dependence of  $^{210}\text{Po}$  rate in cubic volumes on time [113].

Due to the stabilization, there is no  $^{210}\text{Po}$  contamination introduced into the FV by convection in Phase-III, which allows us to consider the migratory contribution of  $^{210}\text{Po}$  to be negligible. Figure 8.4 shows the time evolution of the event rates from the isotopes shown in Eq. 8.1, assuming no independent sources of  $^{210}\text{Bi}$  and  $^{210}\text{Po}$ , and the initial rate of  $^{210}\text{Pb}$  at  $t = 0$  to be 20 cpd/100 t [101].

This demonstrates that the activity of  $^{210}\text{Po}$  reaches equilibrium with  $^{210}\text{Pb}$  in  $\sim 2$  years. After it is reached, the rates of  $^{210}\text{Bi}$  and  $^{210}\text{Po}$  are approximately equal,  $R(^{210}\text{Bi})/R(^{210}\text{Po}) \approx 1.015$ . This fact can be used to estimate an upper limit, or possibly a constraint on the rate of  $^{210}\text{Bi}$  by measuring the rate of  $^{210}\text{Po}$ .

As mentioned in Sec. 4.4.4,  $^{210}\text{Po}$  is the only  $\alpha$  background after the removal of  $^{214}\text{Bi}$ - $^{214}\text{Po}$  coincidences (Sec. 4.2.2), and can be measured on an event-by-event basis using pulse shape information. The distribution of  $^{210}\text{Po}$  events selected using the MLP  $\alpha/\beta$  classifier (see Sec. 3.4.5) is shown in Fig. 8.5A as a function of the vertical position  $z_0$ , and  $\rho^2$ , defined as  $\rho^2 = x^2 + y^2$ .

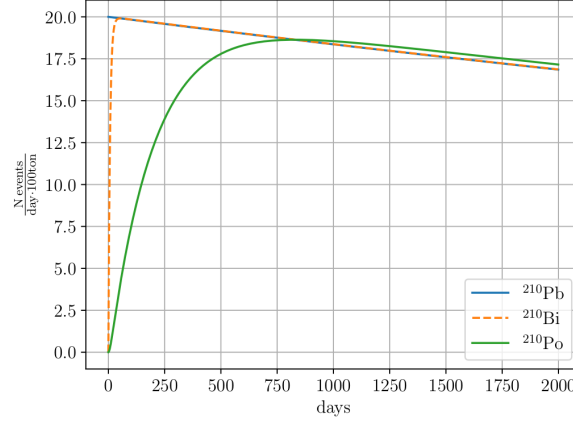


FIGURE 8.4: Time evolution of the rates of  $^{210}\text{Pb}$  (blue),  $^{210}\text{Bi}$  (green), and  $^{210}\text{Po}$  (orange) in the absence of independent sources of  $^{210}\text{Bi}$  and  $^{210}\text{Po}$ , assuming the initial rate of  $^{210}\text{Pb}$  to be 20 cpd/100 t [101].

Assuming rotational symmetry, it can be approximated with a paraboloid function:

$$R(\rho^2, z) = M \cdot T \cdot \left( N_0 + \frac{\rho^2}{a^2} + \frac{(z - z_0)^2}{b^2} \right), \quad (8.3)$$

where  $M \cdot T$  is the exposure. The parameters  $a$ ,  $b$  (shape) and  $z_0$  (position), governing the paraboloid, are obtained via fitting the function in Eq. 8.3 to the  $^{210}\text{Po}$  data [113, 121]. The result of the fit is demonstrated in Fig. 8.5A by the mesh grid. Figure 8.5B shows the temporal evolution of the position  $z_0$  of the distribution, and the corresponding  $^{210}\text{Po}$  rate in each time point. Using the time period in which the  $^{210}\text{Po}$  rate is the most stable (after 2018, as seen in Fig. 8.5B), the upper limit on the  $^{210}\text{Bi}$  rate has been estimated to be  $12.5 \pm 1.5$  cpd/100 t [122, 123].

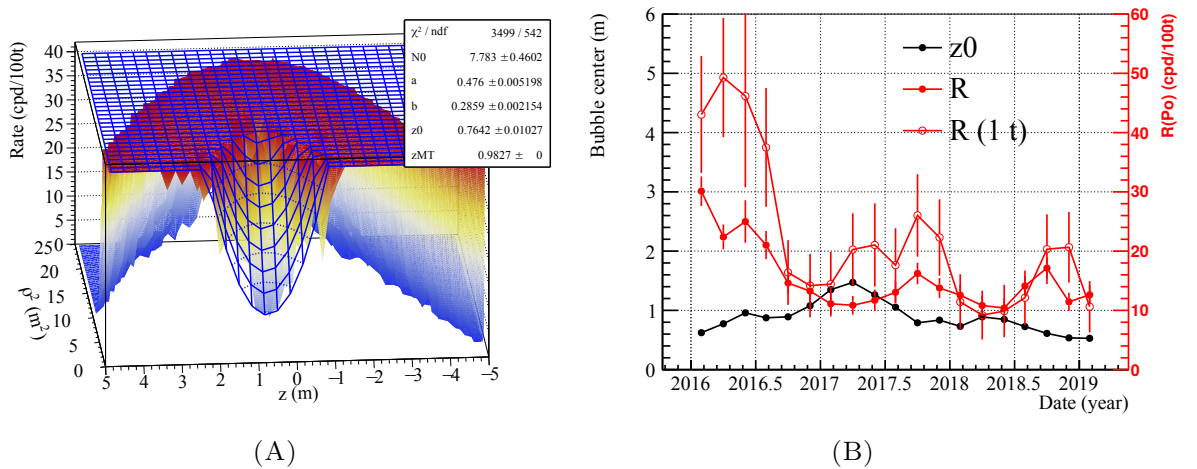


FIGURE 8.5: (A) Fit of the paraboloid function in Eq. 8.3 (mesh grid) to  $^{210}\text{Po}$  data (colored) [113]. (B) Temporal evolution of the vertical position  $z_0$  of the  $^{210}\text{Po}$  distribution in units of m (black), and the corresponding  $^{210}\text{Po}$  rate (red) (filled marker) and the innermost 1-t volume (hollow marker) [113].

### 8.3.2 Constraint on the rate of *pep* neutrinos

The electron capture reaction producing *pep* neutrinos (Eq. 2.3) and that of the  $\beta$  decay process resulting in *pp* (Eq. 2.2) depend on the same theoretical nuclear matrix element, making the ratio of their fluxes, and therefore interaction rates, independent from nuclear physics parameters. This allows to make a robust prediction on this ratio with 1% precision based on SSMs [80, 81, 124], which has already been used in Borexino in the Phase-II analysis in order to obtain an upper limit on the CNO neutrino rate, as mentioned in Sec. 8.2. Using the *pp/pep* approach for the Phase-III analysis, the SSM prediction on the ratio, combined with the Borexino results on the *pp* neutrino rate resulting from the Phase-II analysis (reported in Table 7.2), can be used to place a constraint on the rate of *pep* neutrinos with a precision of about 10%. Such a constraint motivated the sensitivity studies that will be shown in Sec. 8.4 that assume  $\sigma_{pep} = 0.28$  cpd/100 t, and translates into a statistical uncertainty on CNO neutrinos of 1-3 cpd/100 t, depending on the uncertainty on the  $^{210}\text{Bi}$  rate.

The accuracy of the *pep* neutrino rate value can be further improved by employing the *solar luminosity constraint* [77, 78], discussed in Sec. 2.4.1. In order to obtain the constraint, a global fit using data from all existing solar neutrino experiments is performed [79], applying the luminosity constraint, and considering the most recent uncertainties on the oscillation parameters [54]. This approach results in a constraint of  $2.74 \pm 0.04$  cpd/100 t (1% precision), motivating the counting analysis in Sec. 8.3.3 with  $\sigma_{pep} = 0.04$  cpd/100 t.

### 8.3.3 Counting analysis

As mentioned earlier, apart from the *rate + shape* analysis using the MV fit, sensitive to the similarity between CNO,  $^{210}\text{Bi}$  and *pep*, a rate only analysis is performed treating Borexino as a *counting experiment*. In fact, because it is hard to disentangle the three contributions in the fit, including the shape does not greatly improve CNO sensitivity.

In order to determine the energy region, in which Borexino would be the most sensitive to CNO neutrinos, the ratio  $f$  has been defined as:

$$f = \frac{R_{\text{Bi}} + R_{\nu(\text{CNO})} + R_{\nu(\text{pep})}}{R_{\text{tot}}}, \quad (8.4)$$

between the sum of the rates of  $^{210}\text{Bi}$ , and CNO and *pep* neutrinos, and the total event rate in a given region. Setting the condition  $f > 70\%$ , such a region has been determined to be 0.8-1.0 MeV, called **region of interest** (ROI) [125].

The selection of the ROI allow us to measure the rate of residual background components in the ROI independently, and discard negligible ones. Figure 8.6A shows the contributions of  $^{210}\text{Bi}$ , and CNO and *pep* neutrinos in the ROI, as well as the residual background (shown in grey), dominated by  $^{11}\text{C}$ ,  $^7\text{Be}$ , and external backgrounds. In this plot the rate of  $^{210}\text{Bi}$  to be 17.5 cpd/100 t is assumed, based on Phase-II results [72, 115]; and the rate of *pep* and CNO neutrinos to be 2.8 and 4.9 cpd/100 t, respectively, corresponding to the flux predicted by HM-SSM, shown in Table 2.3.

The CNO neutrino rate and uncertainty are obtained from the counting analysis the following way:

$$R_{\text{CNO}} = \frac{1}{\epsilon_{\text{CNO}}} (r_{\text{tot}} - \epsilon_{\text{Bi}} R_{\text{Bi}} - \epsilon_{\text{pep}} R_{\text{pep}} - r_{\text{other}}), \quad (8.5)$$

$$\sigma_{\text{CNO}} = \frac{1}{\epsilon_{\text{CNO}}} (\sigma_{\text{tot}} \oplus \epsilon_{\text{Bi}} \sigma_{\text{Bi}} \oplus \epsilon_{\text{pep}} \sigma_{\text{pep}} \oplus \sigma_{\text{other}}), \quad (8.6)$$

where  $R_{\text{CNO}}$  and  $\sigma_{\text{CNO}}$  are the value and statistical uncertainty of the measured CNO interaction rate,  $r_{\text{tot}}$  and  $\sigma_{\text{tot}}$  mark the total rate in the ROI and its uncertainty, while  $R$  and  $\sigma$  for  $^{210}\text{Bi}$  (Bi) and  $\text{pep}$  stand for the rate of those components determined independently in the full energy range. Corresponding to those,  $\epsilon$  is the fraction of events coming from the given component falling into the ROI, and  $r_{\text{other}}$  is the total event rate of all other components in the ROI. Here  $\oplus$  stands for  $A \oplus B = \sqrt{A^2 + B^2}$ . As mentioned before,  $r_{\text{other}}$  is negligible, and possibly determined independently based on their spectral features outside of ROI.

The proportion of the contributions in the ROI  $r_{\text{tot}} : R_{\text{Bi}} : R_{\text{pep}} : r_{\text{other}}$  is 9 : 4 : 2 : 1, the biggest impact on  $\sigma_{\text{CNO}}$  coming from  $^{210}\text{Bi}$ , and  $\text{pep}$  neutrinos. The uncertainty on CNO neutrino rate as a function of the two components is shown in Fig. 8.6B.

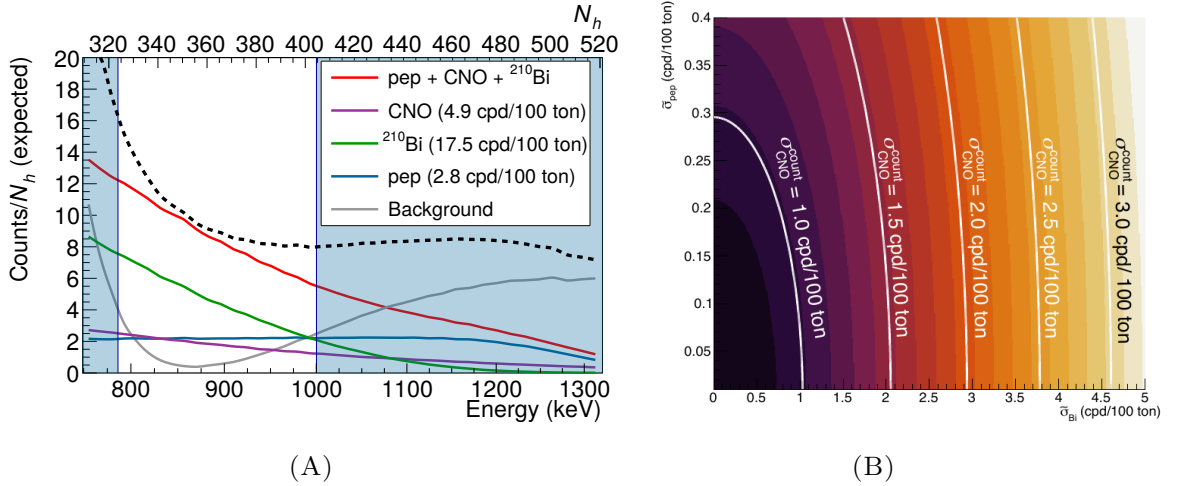


FIGURE 8.6: (A) Expected energy spectrum of Borexino events (black dashed line) after applying the TFC cut, assuming  $\text{pep}$  and CNO neutrino rates predicted by HM-SSM, and  $^{210}\text{Bi}$  rate based on Borexino Phase-II results [72, 115]. The ROI is highlighted with a white band [125]. (B) Uncertainty  $\sigma_{\text{CNO}}$  on the CNO neutrino rate, as a function of uncertainties on  $^{210}\text{Bi}$  and  $\text{pep}$  neutrino rates (see Eq. 8.6) [125].

The predictions on the CNO neutrino rate based on the HM- and LM-SSMs, shown in Table 2.3, translate to the rates of 3.52 and 4.91 cpd/100 t, respectively, as measured by Borexino. Therefore, in order to be able to claim a  $3\sigma$  evidence of CNO neutrinos, or distinguish the two predictions, the uncertainty on the measured CNO neutrino rate needs to be less than 1.5 cpd/100 t.



## 8.4 Sensitivity studies

The sensitivity study presented in this section is aimed at the evaluation of the potential of Borexino to observe a non-zero CNO signal, i.e. reject the *null hypothesis* (denoted  $H_0$ ), based on the multivariate fit. To evaluate the discover power, a *frequentist hypothesis test* has been performed.

To test a certain hypothesis  $H$ , one first constructs the so-called *profile likelihood ratio*  $\lambda$ , defined as:

$$\lambda(\mu_H) = \frac{\mathcal{L}(\mu_H, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \quad (8.7)$$

In Eq. 8.7,  $\hat{\boldsymbol{\theta}}_\mu$  denotes the values of the variables  $\boldsymbol{\theta}$  that maximize the likelihood  $\mathcal{L}$  for a specific value of  $\mu_H$  predicted by hypothesis  $H$ , called *conditional maximum likelihood* (ML). In the denominator is the *unconditional* ML, in which both  $\mu$  and  $\boldsymbol{\theta}$  are ML estimators, and  $\hat{\mu}$  and  $\hat{\boldsymbol{\theta}}$  are the values that maximize it. In other words,  $\lambda$  is the ratio of the maximum likelihood in case we force it to consider the hypothesis  $H$ , and the likelihood in case it can decide the best value  $\hat{\mu}$ .

In our case,  $\mu$  is the rate of CNO neutrinos, and  $\boldsymbol{\theta}$  are the rates of other backgrounds and neutrinos in the multivariate fit, which computes the value of  $\mathcal{L}$ . The hypothesis  $H$  we want to test is the null hypothesis  $H_0$  that predicts  $\mu_H = 0$ . The value  $\mathcal{L}(0, \hat{\boldsymbol{\theta}}_0)$  is obtained from the fit performed when assuming the null hypothesis, i.e. fixing the rate of CNO neutrinos to zero; while  $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})$  is obtained by leaving it free in the fit.

In the specific case when the hypothesis being tested is  $H_0$ , we use the *test statistic for discovery of positive signal*  $q_0$  [126]. It is constructed as follows:

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0. \end{cases} \quad (8.8)$$

To quantify the level of disagreement between  $H_0$  and the observed CNO rate in the data, we need to know the probability density function (PDF)  $f(q_0|H_0)$ . It has been constructed by generating an ensemble of Monte Carlo datasets with an injected CNO rate  $R_{\text{CNO}} = 0$ , and performing MV fits on them to obtain the conditional and unconditional ML, described above [125]. The simulations have been done assuming the exposure of 860 days  $\times$  70 tons, similar to Phase-II, and the values for other components obtained in the Phase-II analysis [72, 115]. The resulting distribution of  $q_0$  is shown in Fig. 8.7 in red.

Then, depending on the actual observed value of  $q_0$ , denoted as  $q_{0,obs}$ , we use the constructed PDF of  $q_0|H_0$  to compute the *p-value*:

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|H_0) dq_0, \quad (8.9)$$

which determines the power of null hypothesis rejection.



In order to evaluate the sensitivity, we first consider *pseudo-experiments* generated based on the same Monte Carlo simulation as described above, with *non-zero* injected rates of CNO neutrinos, from which we obtain the “observed” value  $q_{0,obs}$ . We consider two hypotheses  $H_{HM}$  and  $H_{LM}$ , corresponding to HM- and LM-SSM predictions on CNO neutrino rate  $R_{CNO} = 4.92 \pm 0.55$  cpd/100 t and  $R_{CNO} = 3.52 \pm 0.37$  cpd/100 t, respectively. Fitting these ensembles of “fake data”, we obtain distributions  $f(q_0|H_{HM})$  and  $f(q_0|H_{LM})$ , shown in Fig. 8.7 in blue and green, respectively. My contribution to this study consisted in performing MV fits on the simulated data, collecting the fit results, and setting up a framework that would allow others to do so.

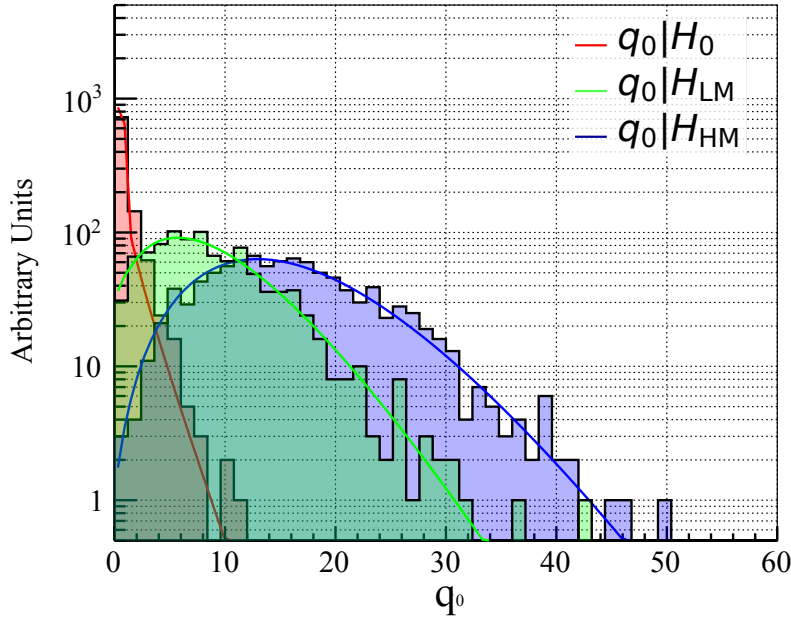


FIGURE 8.7: Distributions of the test statistic  $q_0$  for CNO neutrino discovery assuming the null hypothesis (red), and the presence of CNO neutrino signal, as predicted by the HM- (blue) and LM-SSMs (green) [113].

The distributions  $f(q_0|H_{HM/LM})$  show us where an actual observed value  $q_{0,obs}$  may appear when we perform the analysis on Borexino data. In case we are lucky, the observed value might fall in the very tail of  $q_0|H_{HM}$  or  $q_0|H_{LM}$ , leading to discovery. In case we are unlucky, it may fall within the range of  $f(q_0|H_0)$ , resulting in a large integral in Eq. 8.9. To evaluate the sensitivity, we take the *median* of each distribution as the most likely scenario, and compute the *median p-value* based on Eq. 8.9.

In the fits, the values for  $^{210}\text{Bi}$  and *pep* neutrinos have been constrained to 17.5 and 2.8 cpd/100 t, respectively. The accuracies of the two constraints are the main parameters influencing the p-values i.e. the sensitivity. The obtained median p-values for the cases of LM- and HM-SSM hypotheses, resulting from fits performed with different constraint accuracies, are shown in Fig. 8.8. For comparison, p-values computed based on the counting analysis, described in Sec. 8.3.3, are shown by hollow markers. As can be seen, the difference between the median p-values obtained with the counting analysis and the MV fit decreases as the precision on the  $^{210}\text{Bi}$  and *pep* constraint increases.

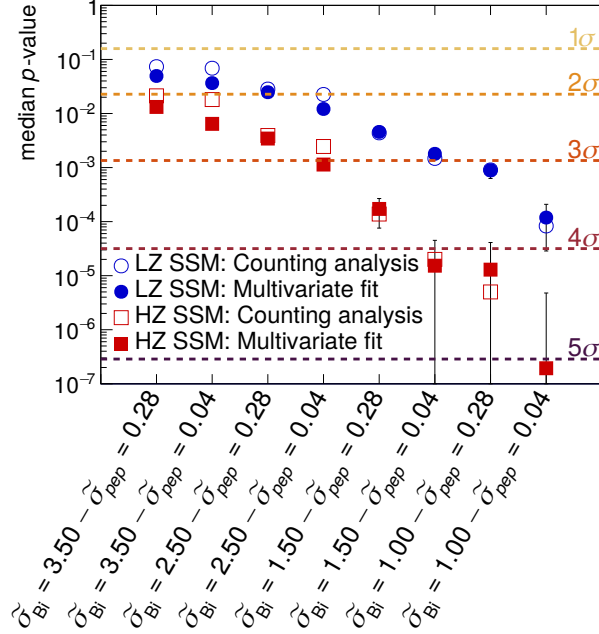


FIGURE 8.8: Discovery power of a CNO neutrino signal depending on the accuracy of the constraint placed on  $^{210}\text{Bi}$  and  $pep$  neutrino rates, as predicted by HM- (red) and LM-SSM (blue). Empty and filled markers denote the result of the counting analysis (rate) and the multivariate fit (rate + shape), respectively [125].

The current constraint placed on the  $^{210}\text{Bi}$  rate has an uncertainty  $\sigma_{\text{Bi}} = 1.1 \text{ cpd}/100 \text{ t}$ , as resulting from studies described in Sec. 8.3.1; while applying the luminosity constraint allows us to achieve a precision on the  $pep$  neutrino rate of  $\sigma_{pep} = 0.04 \text{ cpd}/100 \text{ t}$ , as mentioned in Sec. 8.3.2. Figure 8.8 demonstrates that in this case, Borexino has the potential to reject absence of CNO neutrinos with a significance of  $\sim 3\text{--}4\sigma$ , depending on whether in reality solar metallicity is as predicted by LM- or HM-SSM, respectively.

## 8.5 Phase-III “high setting” multivariate fit

In the multivariate fit of the low-energy solar neutrino analysis, described in Sec. 7.1.2, there are two ways of constructing signal and background reference shapes, namely, using an analytical detector response function, called *analytical approach*, or AA; and using g4b $\times$ 2, the Borexino MC simulation, called the *Monte Carlo approach*, or MCA. Both have been discussed in more detail in Sec. 7.1.3.

Fits with both MCA and AA are performed in the so-called “**high setting**”, mentioned in Sec. 6.7.4, which consists in performing the fit in the energy range above 140 (in the  $N_p^{dt_{1(2)}}$  or  $N_h$  variable, see Sec. 3.4.3 for more on energy estimators), excluding  $pp$  neutrinos,  $^{14}\text{C}$ , and pileup, as the lower energy region requires understanding and additional studies. The data sample used in this analysis spans from July 2016 to February 2020 with the exposure of 1071.95 days  $\times$  71.3 tons. It is obtained using the same data selection cuts, as described in Sec. 7.1.1.

The fits are performed using the GPU-based MV fit framework `GoStats` [127], setting an upper limit on the rate of  $^{210}\text{Bi}$  of  $12.5 \pm 1.1$  cpd/100t, obtained from the analysis described in Sec. 8.3.1; and an absolute constraint of  $2.74 \pm 0.04$  cpd/100t on *pep* neutrinos, based on the luminosity constraint, discussed in Sec. 8.3.2. The rates of other background and neutrino components, including CNO neutrinos, are free in the fit. I will discuss the studies based on the AA and MCA in Sec. 8.5.1 and Sec. 8.5.2, respectively.

### 8.5.1 Multivariate fit with the Analytical Approach

The AA fit was performed in the  $N_p^{dt1}$  variable, in the “high setting” energy range  $140 < N_p^{dt1} < 900$ . Figure 8.9 demonstrates an example of the output of the AA MV fit.

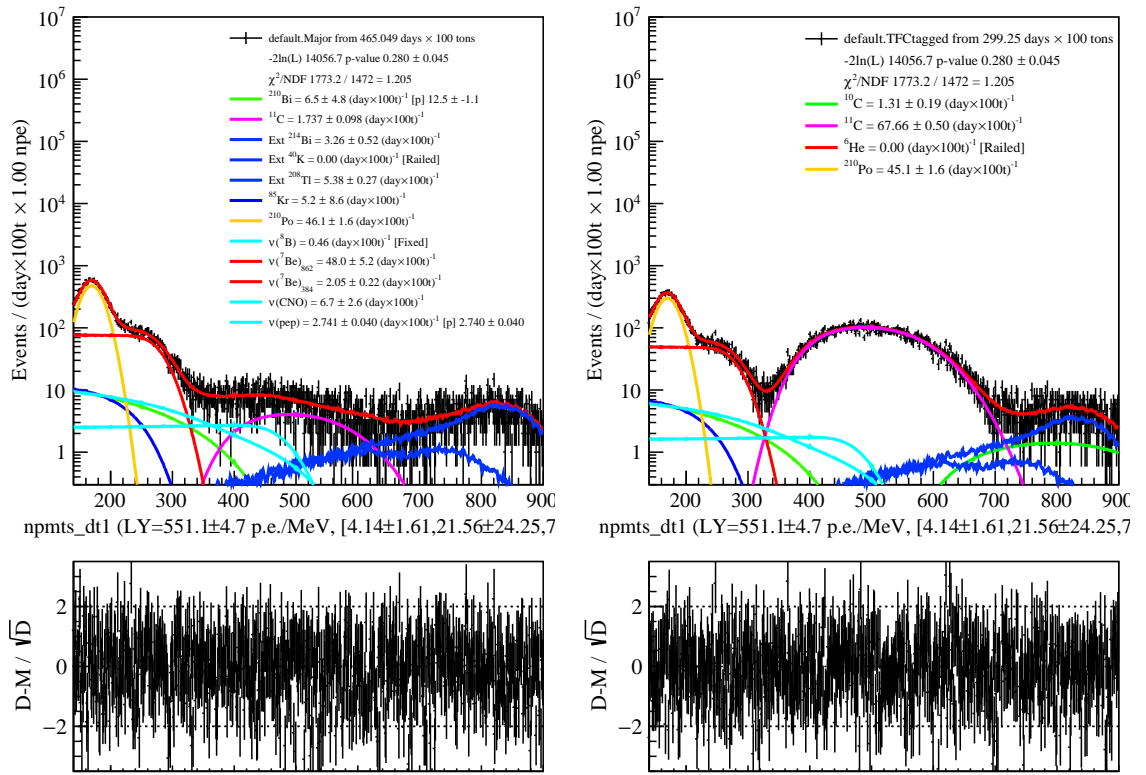


FIGURE 8.9: Example of a “high setting” MV fit of the Phase-III data with the AA.

The advantage of the AA compared to the MCA is its ability to follow the changes of the detector, since parameters governing detector response and the energy scale can be left free in the fit (see Sec. 7.1.3 for more details). The downside that comes with it is increased correlations between fit results, and decreased robustness of the fit compared to the MCA, where parameters such as light yield are tuned based on the calibration data. This makes the MV fit based on the AA more sensitive the energy range and the amount of statistics.

The MV fit is sensitive to the valley between the peaks of  $^{14}\text{C}$  and  $^{210}\text{Po}$ , important for the determination of the rate of  $^{85}\text{Kr}$ . This especially applies to the fit based on the AA. In the high setting, this energy region is excluded. As a result, we observe unstable behavior in the

high-setting AA fit of the Phase-III data, seen by the rate of  $^{85}\text{Kr}$  being compatible with zero, and a large error on the  $^{210}\text{Bi}$  rate.

In the Phase-II analysis, an independent estimation of the  $^{85}\text{Kr}$  contamination determined the rate to be  $\leq 7.5$  cpd/100 t at 95% C.L. (see Sec. 4.4.2). The MV fit results on Phase-II resulted in  $^{85}\text{Kr}$  rate of  $6.8 \pm 1.8$  cpd/100 t, regardless of whether the information from the independent  $^{85}\text{Kr}$  measurement was used or not. This suggests that studying the rate of  $^{85}\text{Kr}$  in Phase-III using the same method of estimation might help stabilize the fit results. In the meantime, the MCA is used for all further studies.

### 8.5.2 Multivariate fit with the Monte Carlo Approach

As mentioned in Sec. 6.1, MCA fit on data after 2016 was not successful before, as the changes in the detector in the recent years became larger compared to Phase-II. As a result, even though the MC reference shapes used to be accurate enough for Phase-II, they did not represent recent Phase-III data well. This has been resolved with the new algorithm for the calculation of the effective quantum efficiency of the Borexino PMTs, which is an important input to `bx_elec`, the Borexino electronics simulation (Sec. 3.5.2), which has been discussed extensively in Chapter 6. My work on the improvement of the MC references shapes found its natural continuation in the performance Phase-III MCA MV fit and related studies.

The fit is performed in the  $N_h$  variable, in the energy range  $140 < N_h < 950$ , and is shown in Fig. 8.10.

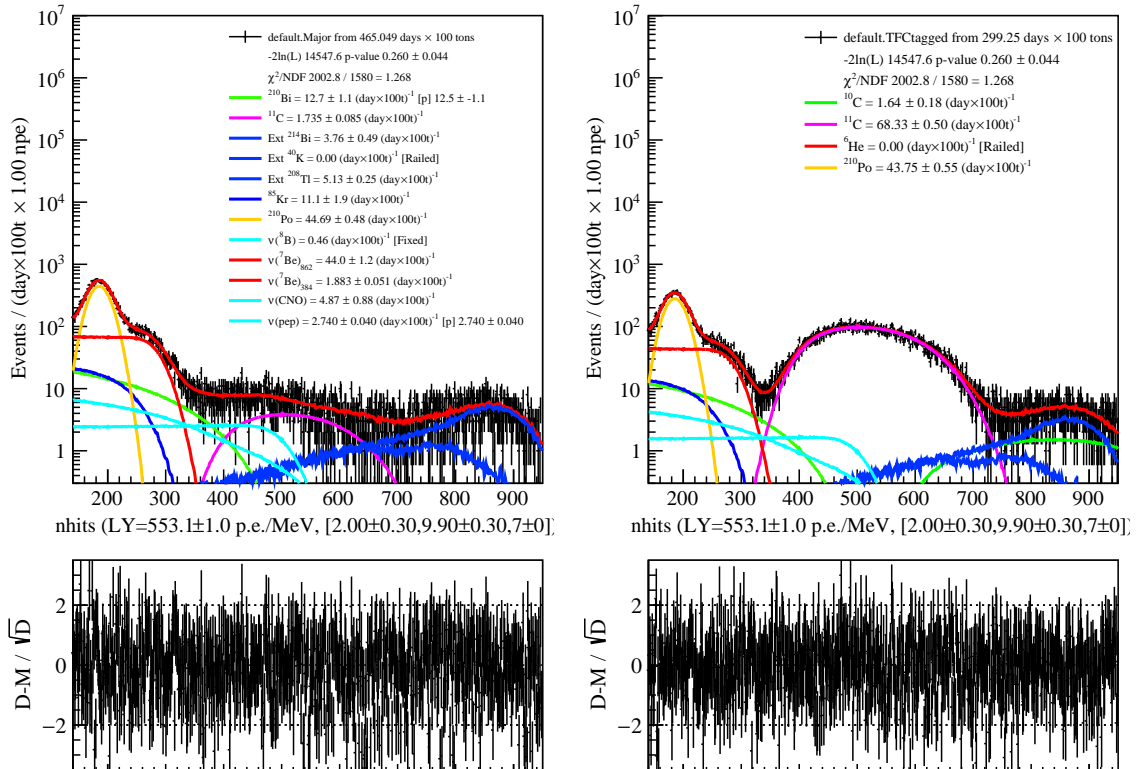


FIGURE 8.10: Example of a “high setting” MV fit of the Phase-III data with the MCA.

One of the important steps of the low-energy neutrino data selection, described in Sec. 7.1.1, is the tagging of the cosmogenic  $^{11}\text{C}$  background. It relies on a technique called *three fold coincidence* (TFC), explained in detail in Sec. 4.5.3. Using TFC, the data is separated into two groups: *TFC-tagged* (rich in  $^{11}\text{C}$ ) and *TFC-subtracted* (with little to no  $^{11}\text{C}$ ). The TFC-tagged and TFC-subtracted energy spectra are fit simultaneously (see Fig. 8.10), the resulting likelihoods being the components of the MV fit, as explained in Sec. 7.1.2.

The Borexino experiment employs two different variants of the TFC method, called TFC-*Milano* (TFC-MI) and TFC-*Mainz* (TFC-MZ), named so according to the institutions that are responsible for their development. In the Phase-II analysis, both techniques were fully compatible and showed virtually no differences. However, when using TFC-MI and TFC-MZ in the Phase-III MV fit, unexpected differences have emerged. I have performed multiple MV fits with variations on the parameters of the fit (energy range, binning etc.) for the purpose of the analysis of systematic effects, which will be presented in Sec. 8.6. The resulting values of the CNO neutrino rate and error are shown in Fig. 8.11.

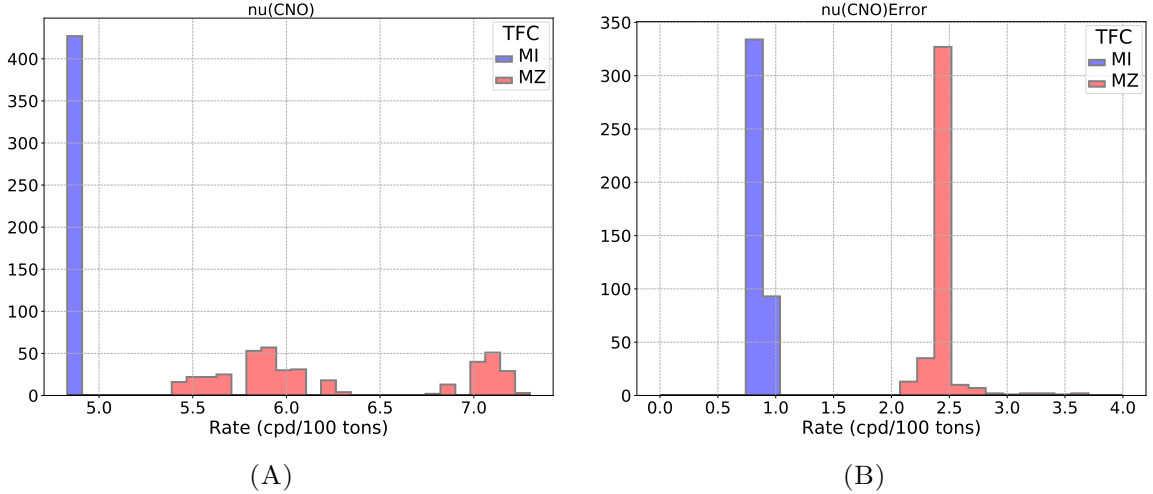


FIGURE 8.11: CNO neutrino rate (A) and error (B), resulting from multiple MV fits with TFC-MI (blue) and TFC-MZ (red).

As can be seen in Fig. 8.11A, there is a very prominent large difference between TFC-MI (blue) and TFC-MZ (red). The results using TFC-MI show extreme robustness to the variation of fit parameters, and the absolute variation on CNO neutrino rate of  $< 1$  cpd/100 t; while the results based on TFC-MZ show a much broader distribution. Moreover, Fig. 8.11B demonstrates that the uncertainty on the resulting CNO value is always  $> 2$  cpd/100 t in the case of TFC-MZ. The disagreement between the two methods compels us to perform more studies to find out the reasons behind the differences. As the reasons have not been discovered yet, I perform my studies for TFC-MI and TFC-MZ separately.

## 8.6 Studies of the systematic effects

In order to evaluate the rejection of the null hypothesis  $H_0$  based on data, we use the method based on the *test statistic for discovery of positive signal*  $q_0$ , described in detail in Sec. 8.4. For that we need to obtain  $q_{0,obs}$  (Eq. 8.10), based on the results of the MV fit of the Phase-III data, and calculate the value of  $p_0$  as in Eq. 8.9.

Based on the multivariate fit, this value is computed in the following way:

$$q_{0,obs} = -2 \ln \frac{\mathcal{L}(0, \hat{\theta}_0)}{\mathcal{L}(\hat{\mu}, \hat{\theta})}. \quad (8.10)$$

The maximized likelihood  $\mathcal{L}(0, \hat{\theta}_0)$  is obtained from performing the MV MCA fit on the Phase-III data sample when fixing the rate of CNO neutrinos to zero; while  $\mathcal{L}(\hat{\mu}, \hat{\theta})$  is obtained from the fit where the rate is free.

We perform two studies to account for the systematic effects on the resulting  $p$ -value:

- variation of  $q_{0,obs}$  based on multiple MV fits with systematic deviations, discussed in Sec. 8.6.1;
- distortion of  $f(q_0|H_0)$ , shown in Fig. 8.7, due to systematic effects, reported in Sec. 8.6.2 (my contribution).

### 8.6.1 Systematic effects in data originating from the MV fit parameters

The MV fit of the low energy solar neutrino analysis maximizes the total multivariate likelihood:

$$\mathcal{L}_{MV}(\vec{\theta}) = \mathcal{L}_{TFC-tag}(\vec{\theta}) \cdot \mathcal{L}_{TFC-sub}(\vec{\theta}) \cdot \mathcal{L}_{PS}(\vec{\theta}) \cdot \mathcal{L}_{RD}(\vec{\theta}), \quad (8.11)$$

based on fits of four distributions: TFC-tagged and TFC-subtracted energy spectra, the distribution of the pulse shape (PS) parameter  $\mathcal{L}_{PS}$ , and the radial distribution (RD). More details on the procedure can be found in Sec. 7.1.2.

In this study, I have performed  $\sim 400$  fits with different variations of the following fit parameters, related to the fitted distributions:

- energy range ( $E_{min} - E_{max}$ ) of the TFC-tagged and TFC-subtracted spectra;
- the minimum and maximum energy of events that contribute to the PS distribution ( $E_{min}^{PS} - E_{max}^{PS}$ );
- the minimum and maximum energy of events that contribute to the RD ( $E_{min}^{RD} - E_{max}^{RD}$ ), as well as its binning ( $\Delta E^{RD}$ ).

Examples of the energy spectra, the RD, and the PS distribution can be found in Fig. 7.3. The energy estimator in consideration is  $N_h^{norm}$  (see Sec. 3.4.3), used in the MCA. The values of the parameters listed above are reported in Table 8.2.

Parameter	Values
$E_{min}$	136, <b>140</b> , 144
$E_{max}$	946, <b>950</b> , 954
$E_{min}^{RD}$	484, <b>500</b> , 516
$E_{max}^{RD}$	884, <b>900</b>
$\Delta E^{RD}$	8, <b>16</b>
$E_{min}^{PS}$	398, <b>400</b>
$E_{max}^{PS}$	648, <b>650</b>
Total	432 combinations

TABLE 8.2: Variation of different energy parameters of the multivariate fit used in the study of systematic effects (see text).

Figure 8.12 shows the resulting distributions of  $q_{0,obs}$  in separate cases of using TFC-MI and TFC-MZ. The distribution based on TFC-MI is much narrower, and shows higher values compared to TFC-MZ. The mean values  $q_{0,obs}$  resulting from the fits with TFC-MI and TFC-MZ are 29.5 and 26.1, respectively. These values are used to compute the p-values, which will be reported in Sec. 8.7.

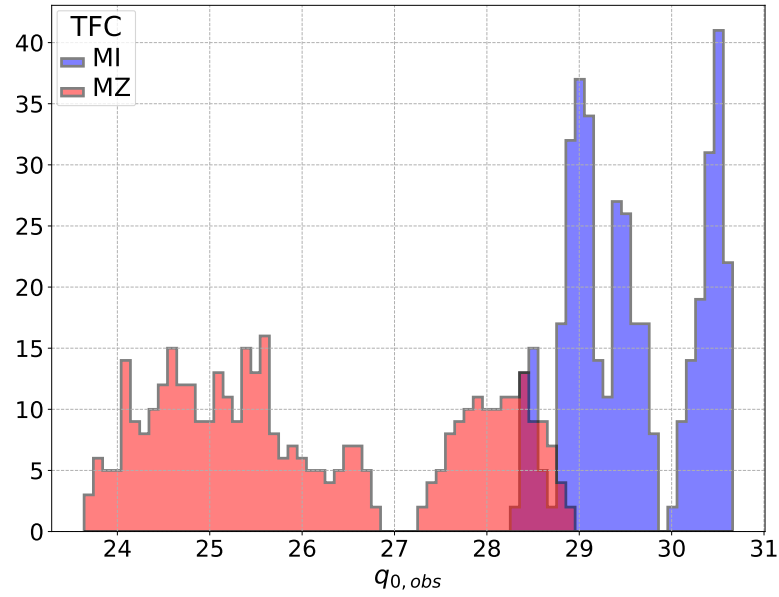


FIGURE 8.12: Distribution of  $q_{0,obs}$ , calculated as in Eq. 8.10 based on the MV fit of Phase-III data with TFC-MI (blue) and TFC-MZ (red).

The main systematic effect on the value of  $q_{0,obs}$  comes from the parameter  $E_{min}^{RD}$ , explained above, and is shown in Fig. 8.13 for the cases of TFC-MI and TFC-MZ.

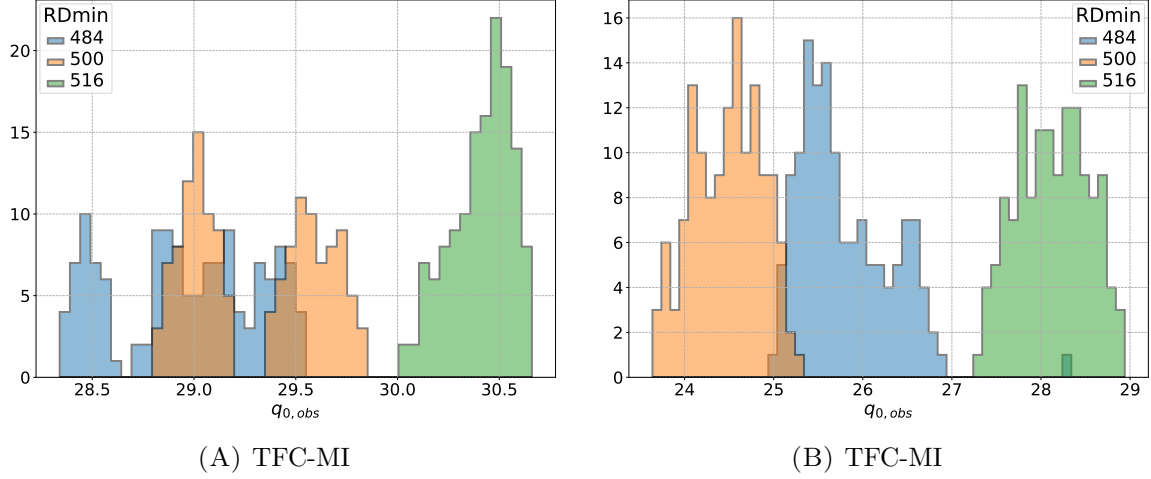


FIGURE 8.13: Distributions of  $q_{0,obs}$  depending on the value of  $E_{min}^{RD}$  for different TFC methods.

### 8.6.2 Systematic effects based on Monte Carlo datasets

This study was aimed at evaluating the influence of systematic effects on the PDF  $f(q_0|H_0)$ . It was based on Monte Carlo generated datasets, in which certain distortions have been applied to different parameters defining the reference shapes of neutrinos and backgrounds [128]. Then,  $\sim 10000$  fits have been performed on these MC datasets, and the resulting distribution of the CNO neutrino rate was compared to the case of no distortion.

The major sources of systematic uncertainties related to the MCA include.

1.  **$^{210}\text{Bi}$  reference shape (18%)**: the impact of the choice of the energy spectrum of electrons originating from  $^{210}\text{Bi}$ , constructed based on different measurements [129–133].
2.  **$^{11}\text{C}$  reference shape (15%)**: while the MC-constructed spectral shape of  $e^+$  emitted by  $^{11}\text{C}$  is quite accurate, its energy scale has an uncertainty. This impacts the CNO neutrino rate, as the energy range  $^{11}\text{C}$  has a significant overlap with that of CNO neutrinos. This study has been performed based on a strict TFC selection (Sec. 4.5.3) of  $^{11}\text{C}$  in Phase-III.
3. **Nonlinearity of the energy scale (0.4%)**: evaluated by comparing calibration data with its MC simulation.
4. **MC LY parameter (0.32%)**: in the MC simulation of the Borexino detector, there is a fixed parameter defining the light yield (LY) in p.e./MeV, imprecision on which would influence the reference shapes used in the MCA fit, and translate into uncertainty the CNO neutrino rate. Its impact has been evaluated by studying the trend of the  $^{210}\text{Po}$  peak position in data and MC, similar the the study described in Sec. 6.7.2.
5. **Spatial nonuniformity (0.28%)**: evaluated using data based on calibration sources in different positions.

The distribution of  $f(q_0|H_0)$ , constructed fitting the MC datasets that include aforementioned distortions, will be shown in Sec. 8.7 in Fig. 8.14 by the blue histogram.



## 8.7 Results and conclusions

To evaluate the significance of the null hypothesis  $H_0$  rejection, i.e. exclude absence of CNO neutrinos, we compute the p-value as shown in Eq. 8.9, by integrating the PDF  $f(q_0|H_0)$  above the observed value  $q_{0,obs}$ , based on the method explained in Sec. 8.4. The undistorted PDF  $f(q_0|H_0)$  is calculated theoretically [126]:

$$f(q_0|H_0) = \frac{1}{2}\delta(q_0) + \frac{1}{2\sqrt{2\pi q_0}}e^{-q_0/2}, \quad (8.12)$$

and is shown in Fig. 8.4 by the red curve. This distribution is broadened when accounting for systematic effects, studied in Sec. 8.6.2, as seen by the blue histogram. I obtained the distribution of  $q_{0,obs}$  values, extracted from the MV fit of the Phase-III data, based on the studies presented in Sec. 8.4. The distribution is shown in Fig. 8.12, and covers the range  $26 < q_{0,obs} < 30$ , depicted in Fig. 8.4 by the green band. The values observed in the data fall within the mean of the PDF  $f(q_0|H_{HM})$ , as predicted by the sensitivity studies presented in Sec. 8.4.

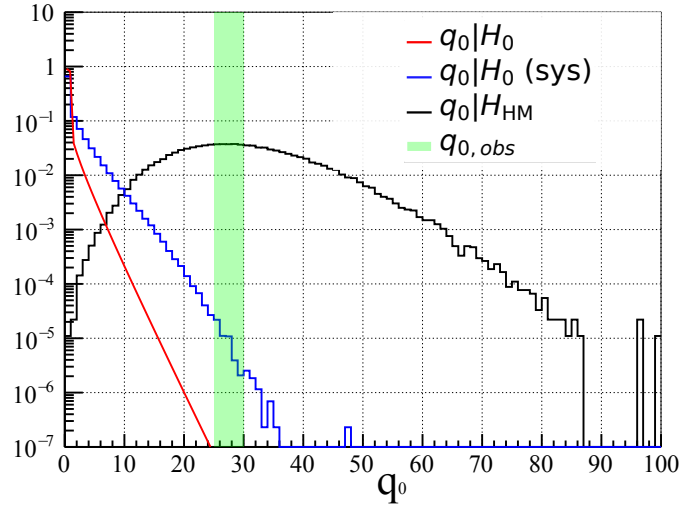


FIGURE 8.14: Initial distribution of  $f(q_0|H_0)$  (red), and the one distorted due to systematic effects (blue), resulting from the MV MCA fit of Monte Carlo generated datasets. The green band indicates the range of the values of  $q_{0,obs}$  obtained from fitting Phase-III data. For comparison,  $f(q_0|H_{HM})$  is shown in black [128].

We compute p-values by integrating both undistorted and distorted PDFs  $f(q_0|H_0)$ . The results are reported in Table 8.3. Even though the discovery power is diminished by the systematic effects, we still obtain  $> 3\sigma$  results, which present evidence of the existence of the CNO cycle in the Sun.

TFC method	Mean $q_{0,obs}$	No syst.		Syst.	
		p-value	significance	p-value	significance
TFC-MZ	26.1	$1.43 \cdot 10^{-7}$	$5.13\sigma$	$(5.92 \pm 0.37) \cdot 10^{-5}$	$(3.85 \pm 0.02)\sigma$
TFC-MI	29.5	$6.49 \cdot 10^{-9}$	$5.69\sigma$	$(4.4 \pm 1.0) \cdot 10^{-6}$	$(4.45 \pm 0.05)\sigma$

TABLE 8.3: Significance of observation of non-zero CNO neutrino signal [134].

The results of CNO neutrino analysis, presented in this chapter, are discussed in the context of the null hypothesis rejection, i.e. evidence of the existence of CNO neutrinos. An upper limit on the rate of  $^{210}\text{Bi}$  translates into a lower limit (LL) on the CNO neutrino interaction rate. Figure 8.15 demonstrates a scan of the profile likelihood depending on the CNO neutrino rate, based on the fit configuration that yields the mean value of the  $q_{0,obs}$  distributions for both TFC-MI (29.5) and TFC-MZ (26.1), shown previously in Fig. 8.12. Based on such a scan, one can determine a LL on the CNO neutrino rate with a certain confidence level. The preliminary result from this study for 99.7% C.L. ( $3\sigma$ , shown by the green dashed line) corresponds to a LL of 2.2 and 2.0 cpd/100t for TFC-MI and TFC-MZ, respectively (not including systematic effects).

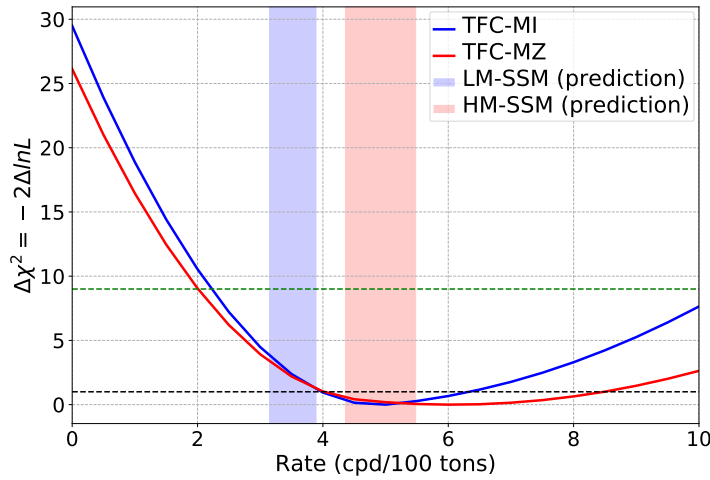


FIGURE 8.15: Scan of the profile likelihood as a function of the CNO neutrino rate based on the MCA fit of Phase-III data using TFC-MI (blue) and TFC-MZ (red).

## 8.8 Towards the absolute measurement of the CNO neutrino flux

The LL on the CNO neutrino rate is governed by the upper limit (UL) on the rate of  $^{210}\text{Bi}$ , obtained from studies reported in Sec. 8.3.1; while the UL on CNO neutrinos is determined by the intrinsic ability of the multivariate fit to distinguish the  $^{210}\text{Bi}$  and CNO neutrino spectral shapes. The  $1\sigma$  uncertainty region denoted by the dashed line in Fig. 8.15, as well as the distributions depicted in Fig. 8.11B, show that the uncertainty on the CNO neutrino rate is quite large for the case of TFC-MZ ( $\sim 2$  cpd/100t). In addition, the value of the UL depends on the TFC method, which is used to create the TFC-tagged and TFC-subtracted spectra used in the fit (see Eq. 8.11). In order to obtain a confidence interval on the CNO neutrino rate, TFC-MI and TFC-MZ are being examined to understand their differences, and further studies of systematic effects are being conducted.

Furthermore, a measurement of the CNO neutrino rate may be possible as well. In the future, the paraboloid region that is used to determine the minimum  $^{210}\text{Po}$  rate (see Sec. 8.3.1), might enlarge thanks to the thermal stability of the detector. Collecting more  $^{210}\text{Po}$  data might allow

us to obtain not just an UL, but an absolute measurement of the  $^{210}\text{Bi}$  contamination, which would translate into a measurement of the CNO neutrino rate.

Lastly, there have been proposals of methods that may allow the observation of CNO neutrinos without an upper limit on  $^{210}\text{Bi}$ , leaving it a free parameter in the MV fit. This may be possible by performing the analysis on an increased exposure, exploiting both Phase-II and Phase-III datasets. Further development of this approach may result in yet another robust probe of these neutrinos and the existence of the CNO cycle, as well as implications in Standard Solar Models.

This chapter concludes my thesis. In the section called [Conclusions](#), you can find the grand summary of Chapters [5-8](#) that present my research. The ideas about applications of my work for further studies, as well as discussion of the future of solar neutrino analysis, can be found in the [Outlook](#).



# Conclusions

The research conducted in this PhD thesis is focused on the detection of solar neutrinos with the Borexino detector. My work in this field encompasses three major directions:

- quality of the Borexino data, on which the solar neutrino measurement is based;
- accuracy of the Monte Carlo simulation, used in the solar neutrino analysis;
- the analysis itself, and the computational tools behind it.

I believe that these subjects belong to the main ingredients of robust analysis.

In Chapter 5 [Detector stability and data quality](#), I have presented my work on ensuring **stable performance of the detector**, and good **reliable data**. My contribution improved the data validation procedure, developed by the Borexino collaboration. It is performed by the so-called *shifters*, whose responsibility it is to monitor the behavior of the detector and spot instances of malfunction. In addition to improvements and introduction of new quality checks, I composed a comprehensive manual of the validation procedure, aimed at reducing mistakes originating from this human factor, and making sure shifters understand the structure and functions of the electronics and hardware. This experience allowed me to develop a novel validation framework for the weekly laser timing calibration of the Borexino photomultiplier tubes (PMTs). Its purpose is the time alignment of the PMT signals, governed by the response of the digital channels they are connected to. This procedure is crucial for position reconstruction, as well as particle identification, which relies on the pulse shape of Borexino events. Lastly, my studies of PMT performance and durability resulted in a discovery of the correlation between PMT livetime and quality. This work helped us to understand the time evolution of the detector, and the influence of the changing number of live PMTs in the context of their changing average quality. Furthermore, it allowed us to select a set of table good quality PMTs, which proved to be essential for the monitoring of the **effective quantum efficiency** (EQE) of the PMTs.

I have discussed the latter in great detail in Chapter 6 [Effective quantum efficiency of the Borexino photomultiplier tubes](#). This chapter reports my work that was prompted by discrepancies between data and its Monte Carlo simulation in the time period after 2016. Several studies have suggested that it stems from the inaccuracy of the simulation inputs of the EQE of the PMTs, which is a parameter governing the PMT light collection ability in the simulation. Utilizing the set of chosen PMTs based on the work described in Chapter 5, I developed an improved calculation procedure, based on a more stable selection of  $^{14}\text{C}$  events, which resulted in a more accurate measurement of EQE. In addition, the improved procedure included estimation of the error on the calculated EQE value, and methods of ensuring high precision in cases when the available  $^{14}\text{C}$  statistics is low. As a result of these studies, the aforementioned discrepancy was eliminated, and it became possible to perform solar neutrino analysis that utilizes the Monte Carlo simulation on the recent data. The analysis of this data, aimed at detection of CNO neutrinos, has been presented in Chapter 8.

The Borexino detector was designed with the primary goal of detecting solar  ${}^7\text{Be}$  neutrinos, originating in the nuclear fusion process called **proton-proton ( $pp$ ) chain**, the main energy production mechanism in the Sun, which converts hydrogen to helium. Since the experiment was established, neutrinos emitted in different stages of the  $pp$  chain have all been measured one by one. The grand finale of this endeavor was their simultaneous measurement, which has been presented in Chapter 7 [Comprehensive measurement of the proton-proton chain neutrinos](#), focused on the analysis of Borexino Phase-II data (2012-2016). I became a member of the Borexino collaboration half a year before the publication of these final results, and joined the efforts in the last stages of the analysis, namely, obtaining the final results, and evaluating the impact of the systematic uncertainties.

An alternative fusion process that converts hydrogen to helium is the so-called **carbon-nitrogen-oxygen (CNO) cycle**, which is hypothesized to be the dominant fusion process in heavier stars, but has evaded detection until now. The Borexino Phase-III analysis (2016-2020), presented in Chapter 8 [First evidence of solar neutrinos from the CNO cycle](#), resulted in the first evidence of its existence through CNO neutrino detection. My work on the improvement of the accuracy of the Monte Carlo (MC) simulation in Chapter 6 found its natural continuation in the Phase-III analysis based on the MC approach. It consisted in studying the Phase-III data, evaluating the significance with which the nonexistence of CNO neutrino signal can be rejected, as well as a potential lower limit and a measurement of the rate, which could provide valuable information in terms of Standard Solar Models, and astrophysics in general.

My work has brought significant contribution to the analysis of the solar  $pp$ -chain and CNO neutrinos through improvement of data quality and accuracy of Borexino simulation, as well as development and application of analysis tools. My results can also be put into use in various future studies, which I will elaborate on in [Outlook](#).

# Outlook

The work presented in this thesis opens new prospects of studying different effects and phenomena, and new directions of analysis that can be undertaken by the next generation of students. Below I present my examination of these open possibilities, my ideas on the potential they hold, and how they can be developed further.

## Chapter 5 Detector stability and data quality

Based on the work presented in Sec. 5.2 regarding the laser photomultiplier (PMT) timing calibration procedure, I have found out that the width in the response of each digital Laben channel improved significantly after the installation of a new trigger system, as can be seen in Fig. 5.7. This is not accounted for in the simulation of the Borexino electronics `bx_elec`, where the improvement and the residual imprecision on the PMT timing calibration is not simulated. The framework for the implementation of this effect has already been set up, and further studies of its influence could improve the accuracy of the Monte Carlo simulation.

My study of the PMT quality and durability and their correlation resulted in a selection of a stable set of chosen PMTs, as described in Sec. 5.3.2. This set has already been used to improve the selection of  $^{14}\text{C}$  events needed for the computation of the effective quantum efficiency of the PMTs. This set was chosen based on 2007-2017 data, and encompasses both Phase-I and Phase-II. It can be used for other studies involving stability in time, or comparison of energy estimators based on the chosen PMTs with the ones based on all PMTs. Furthermore, an updated chosen set can be made by extending the time period to include the more recent data, as well as by incorporating modern machine learning techniques to determine PMT subsets based on their livetime, efficiency, noise etc. This could result in a better selection of PMTs, as well as provide a method of predicting the livetime of a PMT based on its performance.

## Chapter 6 Effective quantum efficiency of the Borexino photomultiplier tubes

My work on the development of a new method of calculation of effective quantum efficiency (EQE), combined with the studies performed in Chapter 5, resulted not only in an improvement of the accuracy of the Monte Carlo simulation, but also in the conclusion that the Borexino PMTs exhibit correlations between light collection abilities and other behaviors. This effect can be seen in Fig. 6.13. It means that while one PMT of lower EQE registers one photon in response to an event, another PMT, given the same conditions, may register two photons, due to having higher EQE. With this situation, it is not correct to treat all PMTs as equal when calculating energy estimators based on their signals. Since the quality does not differ as much as in the exaggerated example above, the current analysis does not show any extreme

inconsistencies. However, by using the EQE of a PMT as a weight when considering the hits it registered, the energy estimators could be improved, and would better represent event spectra. This is especially true in the case of comparing energy spectra from 2007 and 2017, between which many PMTs stopped working. According to the conclusions of Chapter 5, the PMTs that tend to die also tend to have worse EQE, which means the average light collection ability of the recent PMT configuration is actually better than in 2007.

A “side effect” of my study regarding the radius of the  $^{14}\text{C}$  selection for the EQE calculation, showed that the PMTs in the top hemisphere show on average  $\sim 2\%$  higher EQE than the PMTs in the bottom one, seemingly from the very beginning of data taking (refer to Fig. 6.10B). Further studies of this phenomenon could shed light on correlations in PMT parameters, and help avoid biases in energy estimation.

The procedure of selecting  $^{14}\text{C}$  events, used for EQE calculation, relies on the chosen PMT set determined during my PMT quality studies. One effect emerging from my studies on  $^{14}\text{C}$  selection is that of a bias in the chosen PMT set, which comes from Poisson fluctuations, presented in Sec. 6.3.3. At the time, this phenomenon was not understood to its fullest, and gathering more insight about its mathematical formulation would not only confirm it, but could also provide us with a better way of correcting the bias, resulting in more precise and accurate EQE inputs for the MC simulation.

Finally, the EQE calculation procedure, as well as other potential studies, would benefit greatly from more studies on the behavior of single PMTs (see App. B), and the investigation of possible reasons behind it.

## Chapter 7 Comprehensive measurement of the proton-proton chain neutrinos

The analysis presented in this chapter was performed on Phase-II data which covers the range between December 2011 and May 2016. Phase-II is characterized by significantly lower levels of radioactive contaminants compared to Phase-I (2007-2010), achieved thanks to the purification campaign conducted in 2010-2011. Apart from natural long-lived radioactive materials such as  $^{238}\text{U}$  and  $^{232}\text{Th}$ , present in negligible levels already in Phase-II, and further reduced after the campaign, the contamination levels of  $^{210}\text{Bi}$  and  $^{85}\text{Kr}$  greatly decreased as well. As seen in Fig. 4.4, these isotopes present major background for the measurement of low-energy solar neutrinos. However, even though their levels in Phase-I are higher, it could be possible to conduct the analysis on the data of both Phase-I and Phase-II, which would benefit from the increased statistics.

As a part of my work on the systematic uncertainties of this analysis, I developed a Python based tool that allowed generating thousands of multivariate fits, collecting their output, obtaining the uncertainties, as well as graphically illustrating the distributions of neutrino rates, and profile likelihood scans. I later improved and enriched this framework, and applied it in the Phase-III analysis, presented in Chapter 8. This tool can be further applied for similar future studies, as well as expanded and customized for the analysis of systematic uncertainties



and discovery significance, resulting from other kinds of analyses, not necessarily based on a multivariate fit.

## Chapter 8 First evidence of solar neutrinos from the CNO cycle

The last chapter of my thesis demonstrated preliminary results based on the Borexino Phase-III CNO neutrino analysis. One of the findings that emerged from these studies are instabilities in fits with the analytical approach (AA) for constructing reference shapes of neutrinos and backgrounds. This results from the exclusion of the lower energy region in the “high setting” fit of Phase-III, which especially affects the the AA. For example, the results on  $^{85}\text{Kr}$  is affected, as the fit relies on the energy region between the peaks of  $^{14}\text{C}$  and  $^{210}\text{Po}$ , excluded in the high setting, for its determination. The rate of  $^{85}\text{Kr}$  can be estimated from independent studies, thanks to its specific signature, described in Sec. 4.4.2. Such estimation has already been done on Phase-II. Applying this method in Phase-III can help us understand the behavior of the MV fit, as well as the status of  $^{85}\text{Kr}$  in Phase-III.

The CNO analysis is performed on Phase-III data (2016-2020), because it is characterized by thermal stability, achieved after the thermal insulation campaign in 2015. This stability prevents migration of  $^{210}\text{Po}$  into the innermost volume of the detector, which allows us to determine its rate and place an upper limit on the rate of  $^{210}\text{Bi}$ , the largest background for CNO neutrino detection. The current  $^{210}\text{Bi}$  upper limit can be improved, which could result in a more stringent lower limit on the CNO neutrino rate.

The Phase-III analysis is performed in a higher energy range that excludes  $^{14}\text{C}$ ,  $pp$  neutrinos, and the contribution from *pileup* of events. In order to be able to include the lower energy components, we need to make a thorough evaluation of the rates of  $^{14}\text{C}$  and *pileup*, which need to be constrained in the MV fit. I have already participated in studies aimed at the evaluation of both of these backgrounds, and they can be further continued to allow a full energy range analysis of Phase-III with a simultaneous fit of all neutrino and background components.

The results reported in this chapter consider the rejection of the null hypothesis, i.e. nonexistence of CNO neutrinos, and a preliminary value on the lower limit of the CNO neutrino rate. A confidence level interval including an upper limit can be achieved after performing more studies on the AA and MCA multivariate fit behavior. Furthermore, a determination of the CNO neutrino flux may be achieved through an absolute measurement of the  $^{210}\text{Bi}$  rate in the future.

Lastly, there have been proposals of methods that may allow the observation of CNO neutrinos without an upper limit on  $^{210}\text{Bi}$ , leaving it a free parameter in the MV fit. This may be possible by performing the analysis on an increased exposure, exploiting both Phase-II and Phase-III datasets. Further development of this approach may result in yet another robust probe of these neutrinos and the existence of the CNO cycle, as well as implications in Standard Solar Models.



## Appendix A

# Comprehensive manual of the Run-Validation procedure

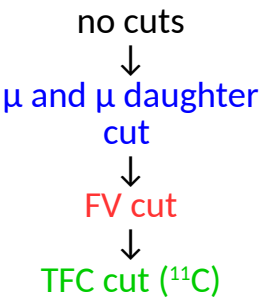
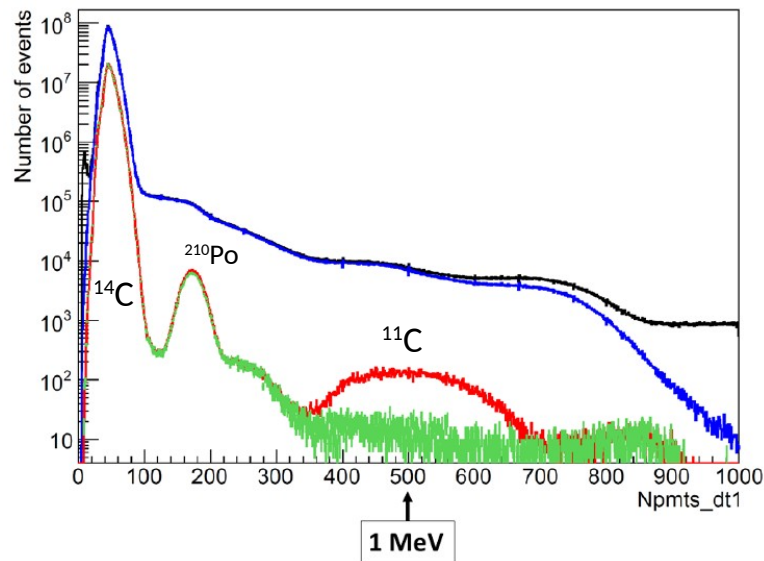
The full list of canvases of the RunValidation procedure discussed in Sec. 5.1:

- **Stability n.1:** muon and noise rates, number of live PMTs
- **Stability n.2:** rates of events in the range of  $^{14}\text{C}$  and the  $^{11}\text{C}$ - $^{210}\text{Po}$  valley in IV and FV
- **Stability n.3:** rates of events in the range of  $^{210}\text{Po}$  and  $^{11}\text{C}$  in IV and FV
- **Stability n.4:** rate of  $^{222}\text{Rn}$  events in IV and FV, spread of the cluster start times
- **Canvas 0:** precalibration
- **Canvas 1:** trigger reference channels
- **Canvas 2:** laser reference channels
- **Canvas 01-02:** event size VS event number (ID)
- **Canvas 03:** trigger rate (ID)
- **Canvas 04:** neutrino trigger
- **Canvas 05:** pulse shape (ID)
- **Canvas 06:** event size VS channel (ID)
- **Canvas 07:** hits in neutrino trigger
- **Canvas 08:** muon flag stability
- **Canvas 09:** event size VS event number (OD)
- **Canvas 11:** trigger rate (OD)
- **Canvas 12:** hits in muon triggers (ID)
- **Canvas 13:** hits in muon triggers (OD)
- **Canvas 14:** pulse shape (OD)
- **Canvas 15:** event size VS channel (OD)
- **Canvas 16-17:** neutron trigger

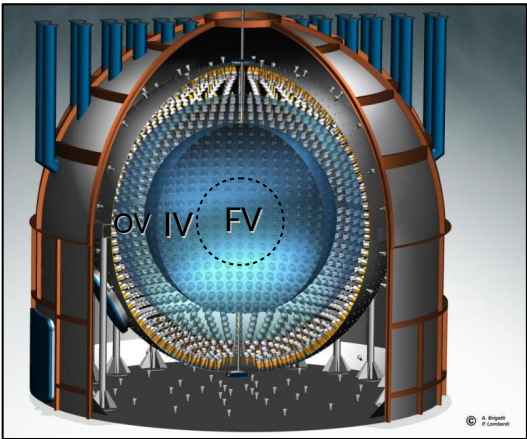
Below I attach the **Run Validation manual** designed by me for the shifters to follow the validation procedure.

General information

The Borexino spectrum



Detector structure



Types of events

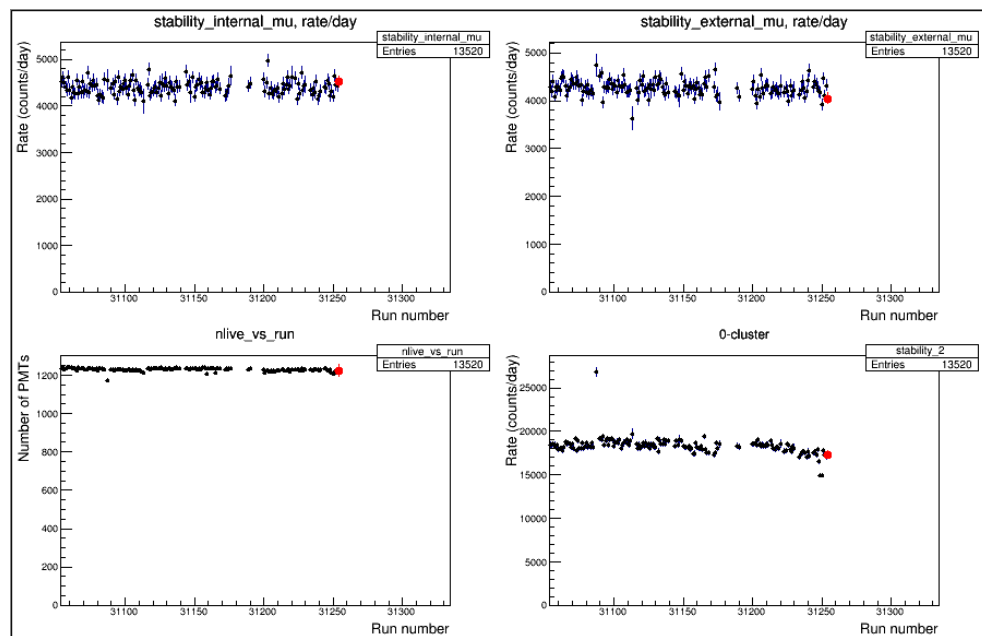
Laser	TT8	→ 0.5Hz
Pulser	TT32	→ 0.1 Hz
Random	TT64	→ 0.5 Hz
Internal event	TT1	
Neutrino	TT1&BTB0	→ 99% $^{14}\text{C}$
Internal muon	TT1&BTB4	
External muon	TT2	→ 75% light leak
Neutron	TT128	

## Stability n.1

Refer to the Borexino spectrum

Compare the current run (red dot) to the previous runs.  
Do not validate if the red dot is 5 own error bars away from the average.

**Internal/external muons** are expected to have a stable rate in each run. Otherwise it could mean electronics failure. The rate of internal/external muons can be also seen on canvases 08 and 11.



Number of **live PMTs** in the last event of the run.  
Expected to be stable.  
If **>40 PMTs below average**, check out canvases 01 and 02 to see how this number varies throughout the run. If the drop happened only in the end, the run can be truncated to avoid this.

**Zero-cluster** events are noise.  
Expected to be stable. Deviations might mean high noise → notify run coordinator. 0-cluster events can be also seen on canvas 04

## Stability n.2

*Refer to the Borexino spectrum*

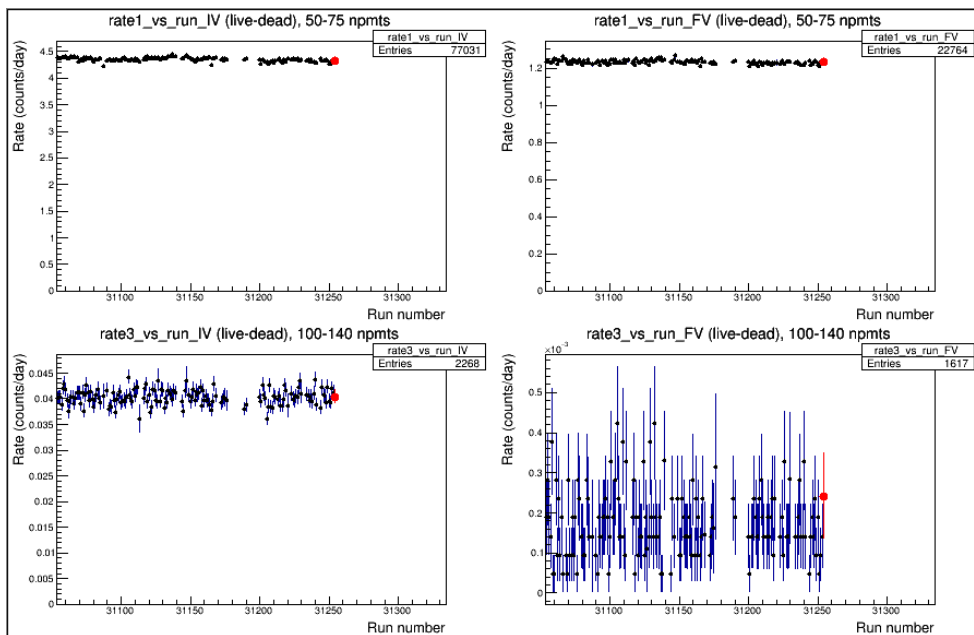
Compare the current run (red dot) to the previous runs.  
Do not validate if the red dot is 5 own error bars away from the average.

Energy region **50-75 PMTs:**  
 **$^{14}\text{C}$  peak**

(see Borexino spectrum)  
Low and stable rate is expected  
in the **inner vessel**.

Energy region **50-75 PMTs:**  
 **$^{14}\text{C}$  peak**

The same as the top left plot but  
in **fiducial volume**. The rate in the  
FV is even lower than the one in  
the IV, also expected to be stable.



Energy region **100-140 PMTs:**  
**valley between  $^{14}\text{C}$  and  $^{210}\text{Po}$**   
(see Borexino spectrum)  
Low and stable rate expected  
in the **inner vessel**.

Energy region **100-140 PMTs:**  
**valley between  $^{14}\text{C}$  and  $^{210}\text{Po}$**   
The same as the one the bottom left  
but in **fiducial volume**. The rate in  
the FV is even lower than the one in  
the IV, also expected to be stable.

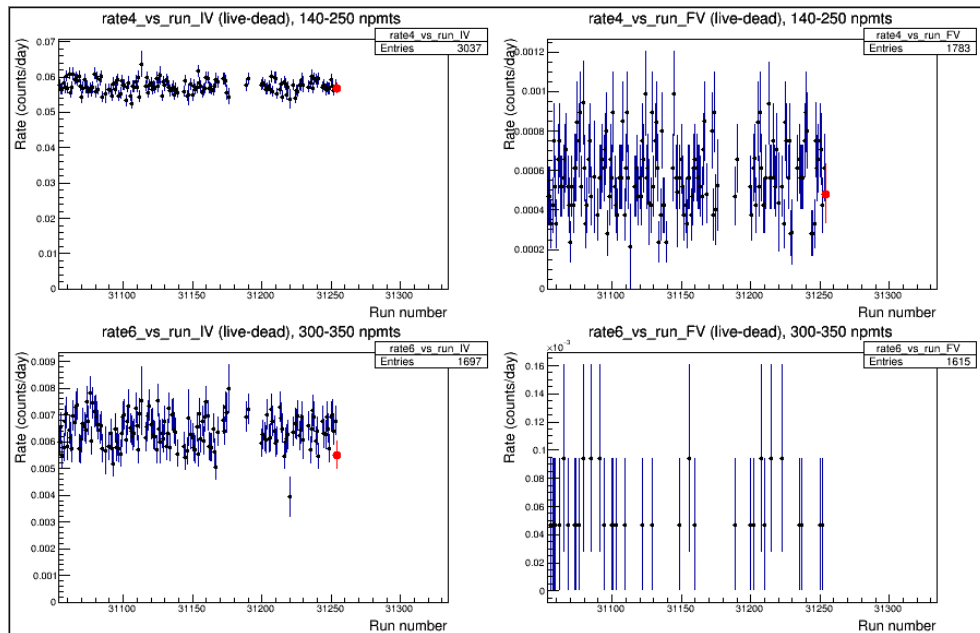
### Stability n.3

Refer to the Borexino spectrum

Compare the current run (red dot) to the previous runs.  
Do not validate if the red dot is 5 own error bars away from the average.

Energy region **140-250 PMTs:**  
 **$^{210}\text{Po}$  peak**  
(see Borexino spectrum)  
Low and stable rate is expected in the **inner vessel**.

Energy region **140-250 PMTs:**  
 **$^{210}\text{Po}$  peak**  
The same as the top left plot but in **fiducial volume**. The rate in the FV is even lower than the one in the IV, also expected to be stable.



Energy region **300-350 PMTs:**  
 **$^{11}\text{C}$  region**  
(see Borexino spectrum).  
**Very** low and stable rate is expected in the **inner vessel**.

Energy region **300-350 PMTs:**  
 **$^{11}\text{C}$  region**  
The same as the bottom left plot but in **fiducial volume**. The rate in the FV is even lower than the one in the IV, also expected to be stable.

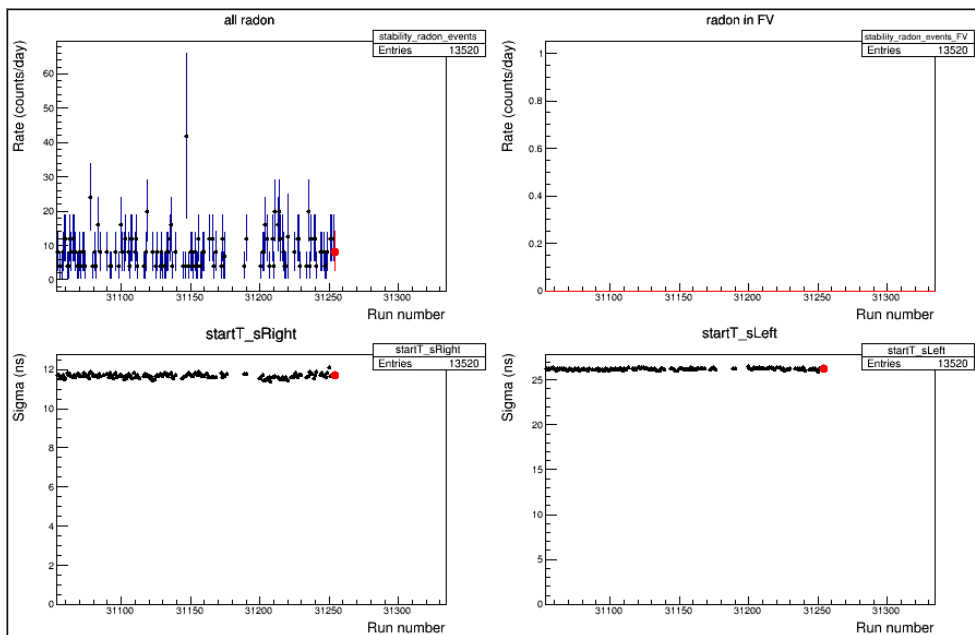
### Stability n.4

Related to canvas 04

Compare the current run (red dot) to the previous runs.  
Do not validate if the red dot is 5 own error bars away from the average.

Most of the **radon** is sitting on the nylon **inner vessel**. The scintillator **inner volume** has some Rn events coming from the NV itself → we see a **small amount of Rn events in the IV**.

**No radon events** are expected in the **fiducial volume** because Rn sits on the inner nylon sphere and its diffusion rate is only 1 cm per day → can see it in the inner vessel, but **not in the FV**.

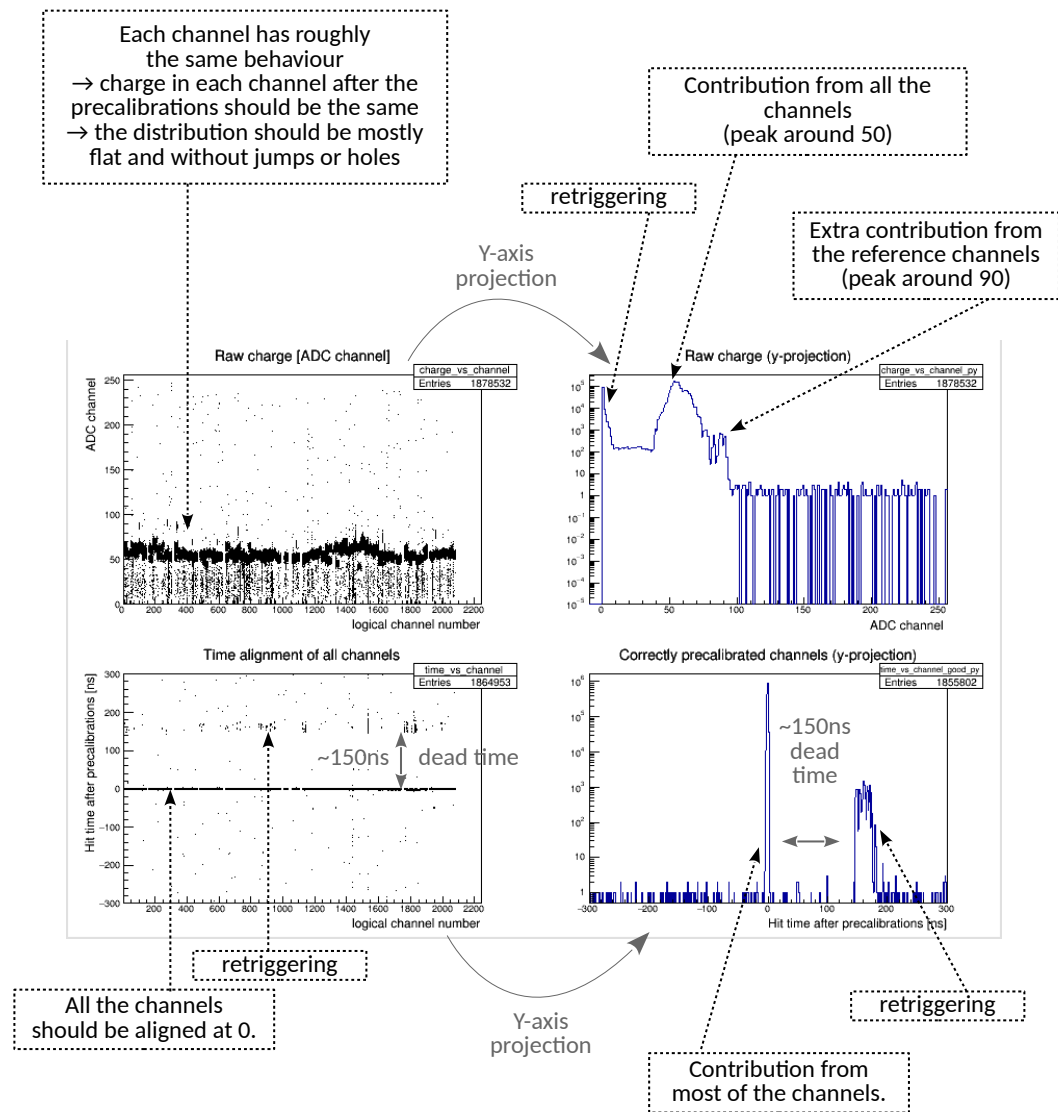


The left/right plot shows the left/right  $\sigma$  of the distribution of cluster start times in the ID (you can also see it in canvas 04).  
Expected to be stable throughout different tuns → stable electronics functioning. If deviating too much, check other canvases for weird behaviour, notify run coordinator etc.

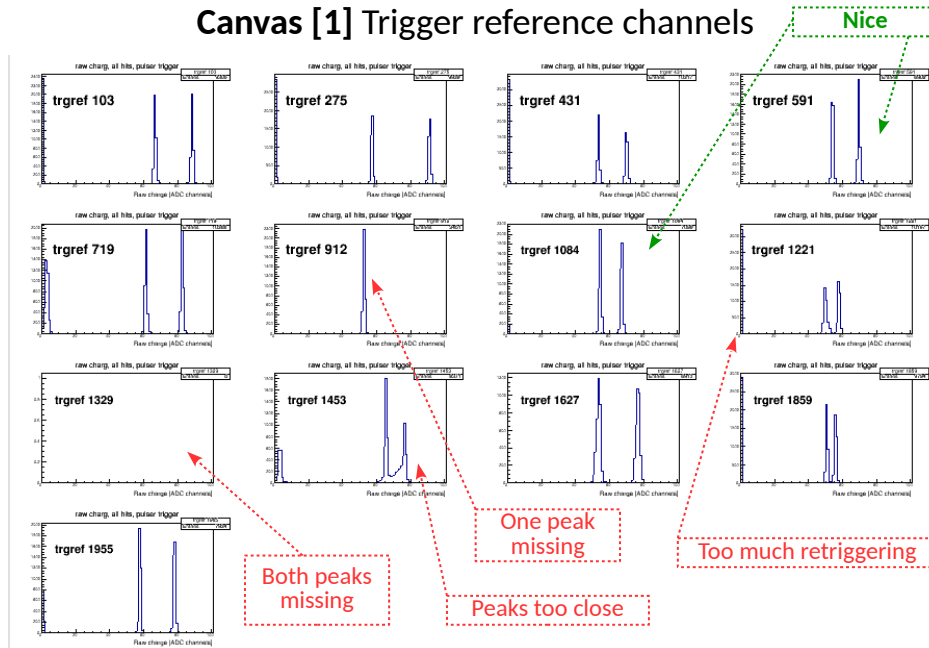


## Canvas [0] Precalibration

NEW

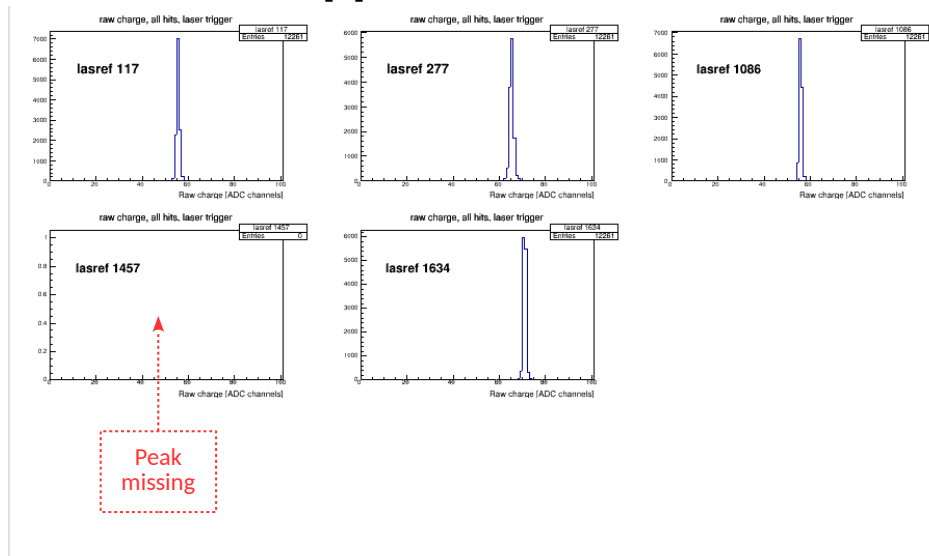


## Canvas [1] Trigger reference channels



Trigger reference channels should have two peaks: one at ~60 from the pulser (that also every channel gets) + one peak ~90 special for the reference channels. The peak at 0 is due to retriggering.

## Canvas [2] Laser reference channels



Laser reference channels should have one peak. Note 1457 missing (bad behaviour).

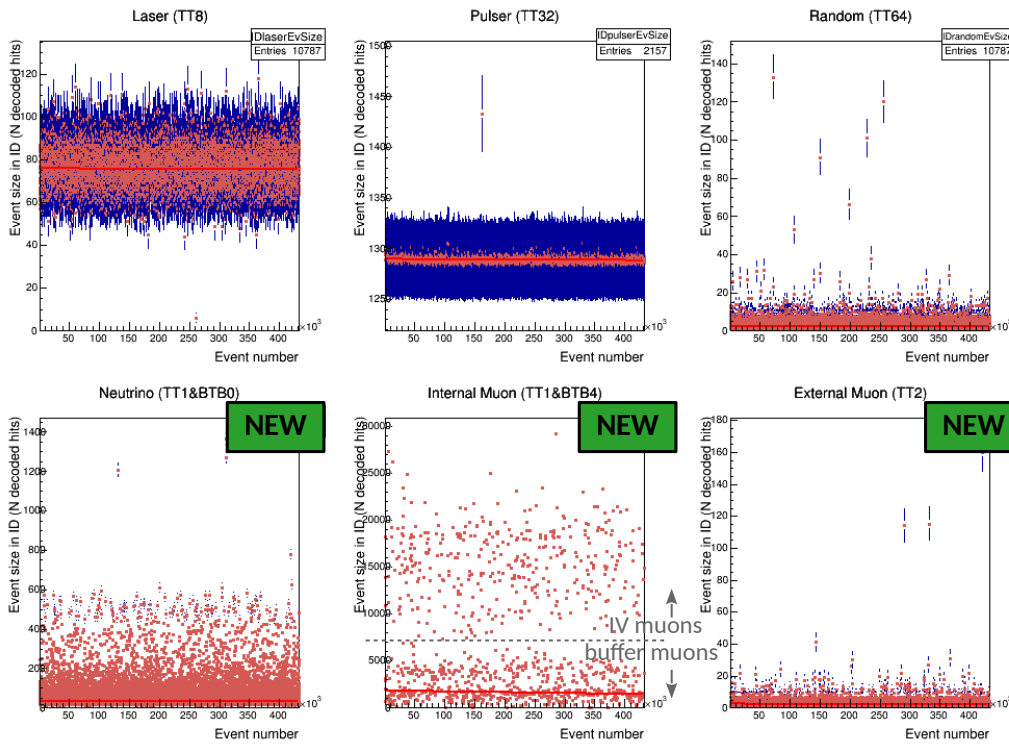
## Canvas [01] ID: event size VS evnum

Note: similar to canvas 09 for OD

**Laser** sends a signal to PMTs in such a way that the probability of one PMT seeing 1 hit is much smaller than 1 in order to ensure single p.e. mode (i.e. small chance to see 1 hit, but virtually zero chance to see 2 hits).  
The resulting event size is ~80 dec. hits  
(80/1300 PMTs ~ 6% chance)

Every **pulsar** event produces 1 hit in each channel. The event size in each event should be equal to the # of live channels:  
 $1 \text{ hit/event/channel} \times N \text{ ch.} = N \text{ hits/event}$   
A sudden drop of 24 ch. is a sign of a **tripper**. Compare to stability canvas n.3

Events in the **random** trigger are dominated by dark noise + some events of any other trigger type which happened to be in the gate randomly. Event size should be very small on average (represents dark noise).



**"Neutrino"** events are 99%  $^{14}\text{C}$  events and 1% everything else (including neutrinos).  
Typical  $^{14}\text{C}$  event gives 30-40 decoded hits.  
(refer to the Borexino spectrum and canvases 04 and 07)

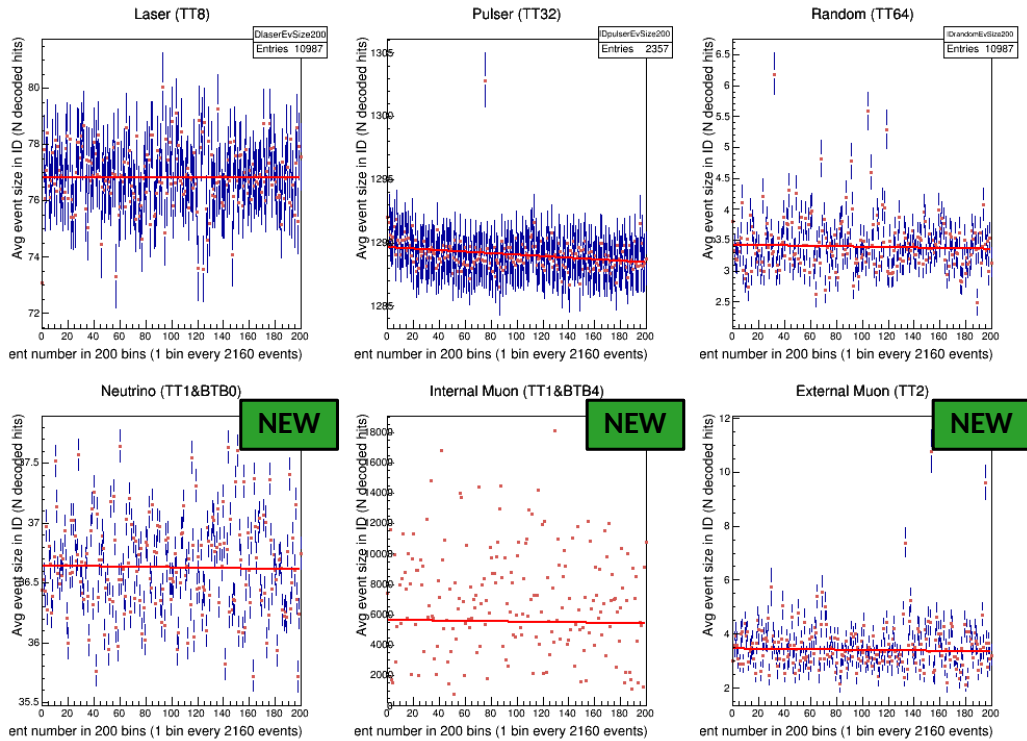
**Internal muons** are very energetic.  
1 PMT has much more than 1 hit in each event.  
Buffer muons have fewer hits because buffer does not have that much scintillator in it.

**External muons** do not pass through the ID. What is recorded in the ID is dark noise and random background. This plot should look similar to the one of the random trigger.

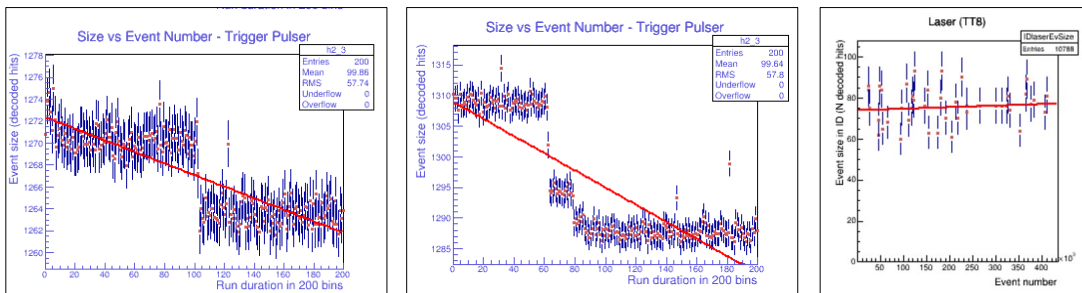
## Canvas [02] ID: event size VS evnum (200 bins)

*Note: this canvas is averaged out stuff from canvas 01*

*Note: similar to canvas 10 for OD*



### Some examples of bad behaviour



Single drop of 10 channels  
→ normal

Validate and make a note

A drop of ~24 channels  
→ tripper!

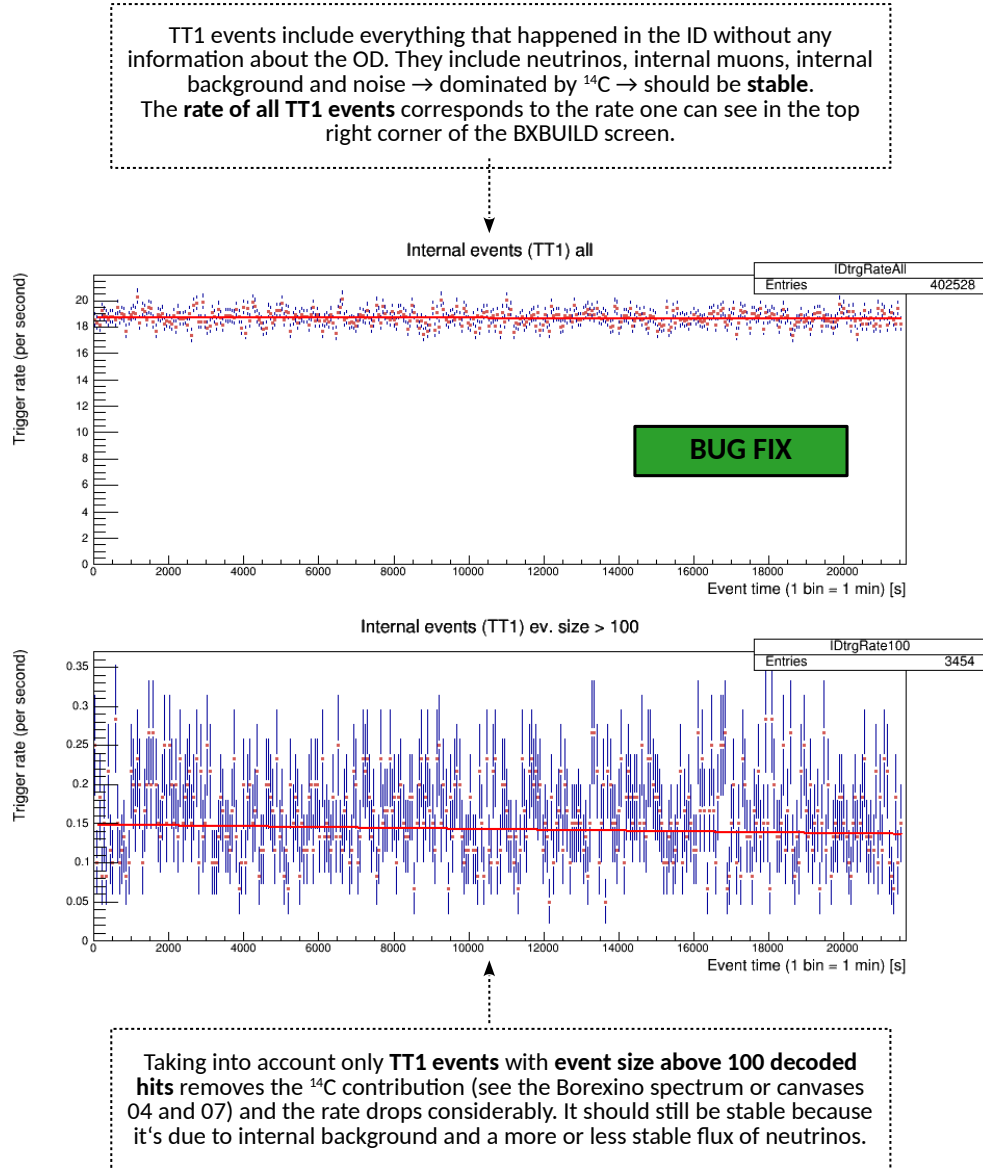
Notify and liquidate!  
(24 channels = 2 front end  
boards with 12 channels each = 1  
HV board that turned off)

Too few laser events  
→ laser was not  
working properly!

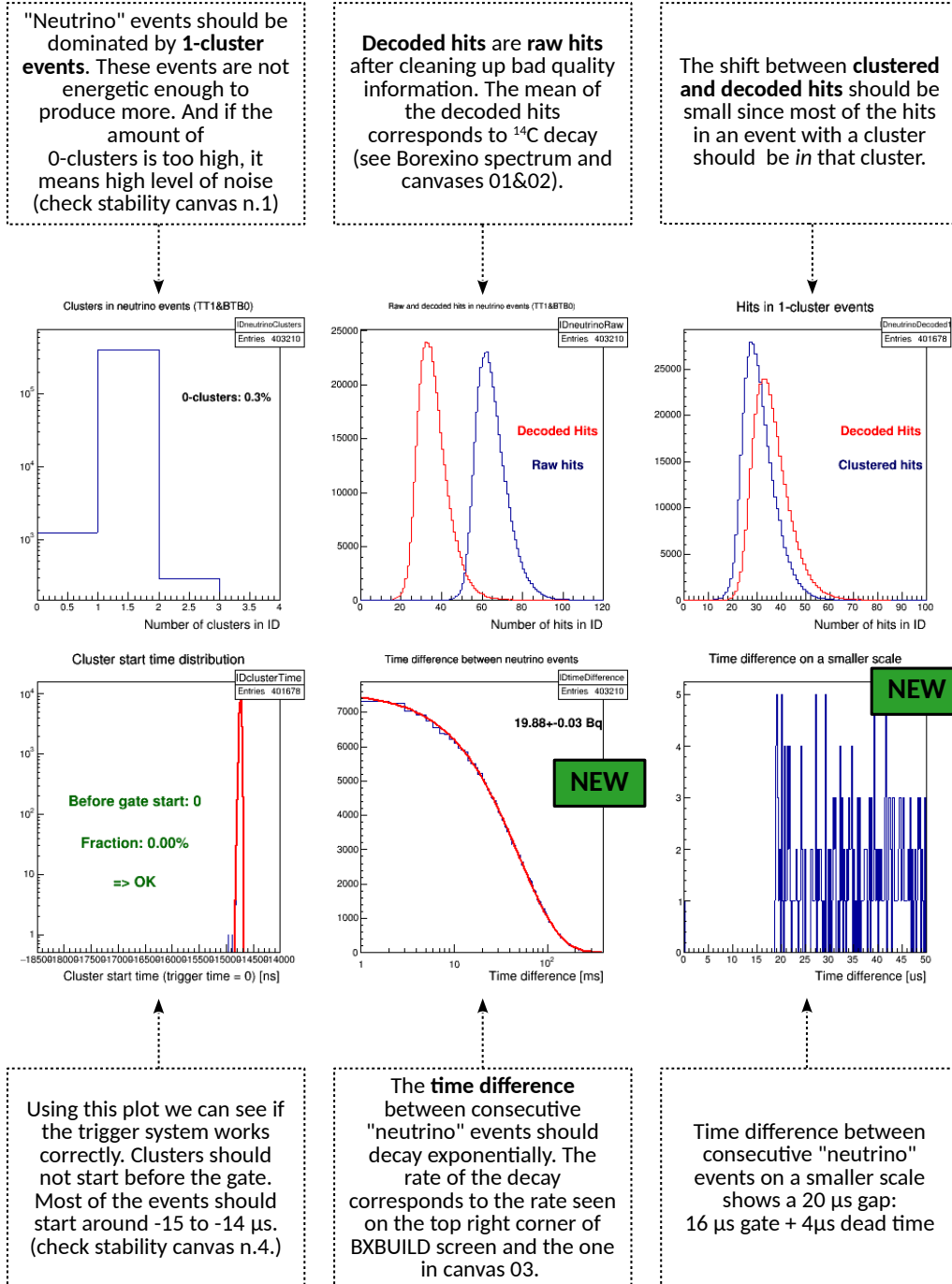
Validate and make a  
note, check if this  
reoccurs in the next  
run

### Canvas [03] ID: trigger rate

*Note: similar to canvas 11 for OD*



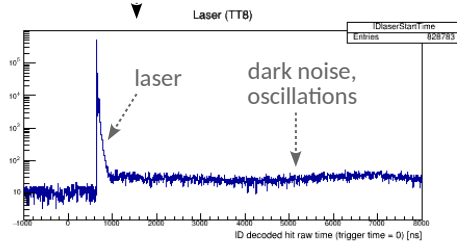
### Canvas [04] ID: neutrino trigger (TT1&BTB0)



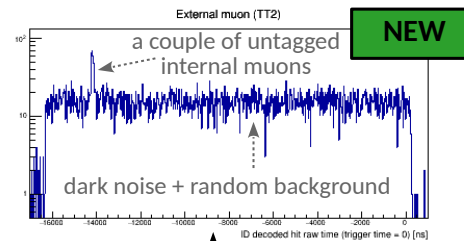
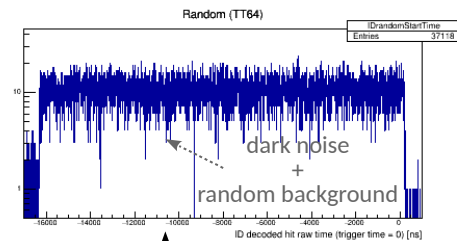
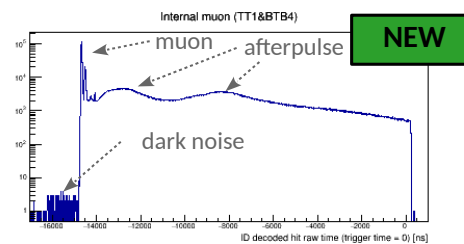
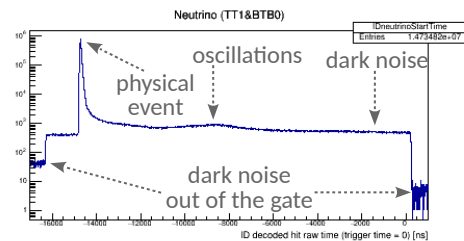
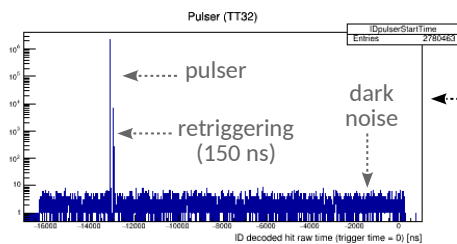
## Canvas [05] ID: pulse shape

Note: similar to canvas 14 for OD

Laser hits have a sharp peak smeared out due to PMT response.

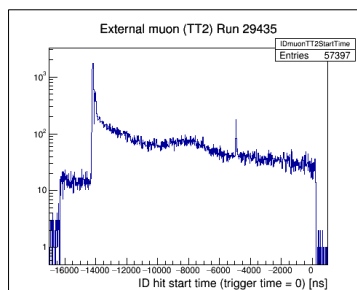


Pulsar sends the signal to the channel directly (not touching the PMT) → peak even sharper than that of the laser hits. The peak should be sharp meaning stable response.



Random trigger is issued with the frequency of 0.56 Hz to record dark noise and random background. The distribution should be flat since the hits have equal chance to happen any time during the gate.

External muons do not pass through the ID → what is recorded is the same as for the random trigger. The distribution should look flat for the same reason.



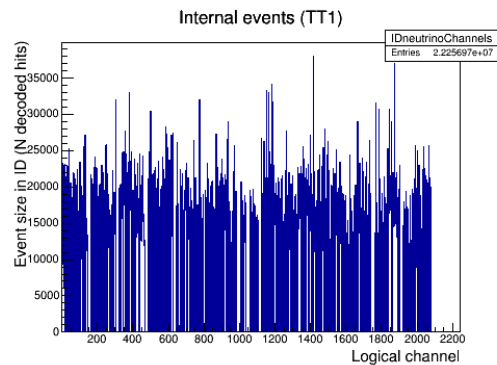
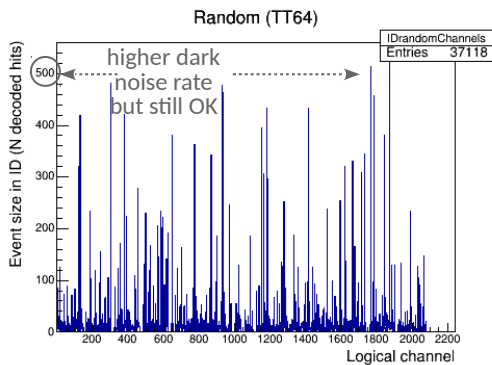
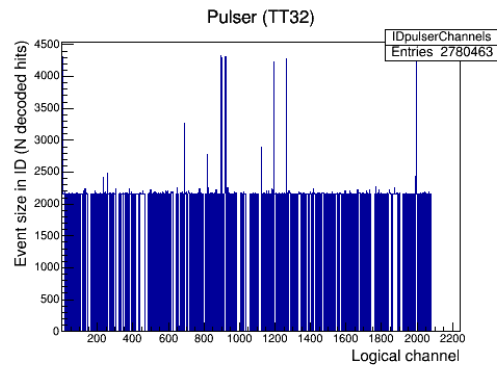
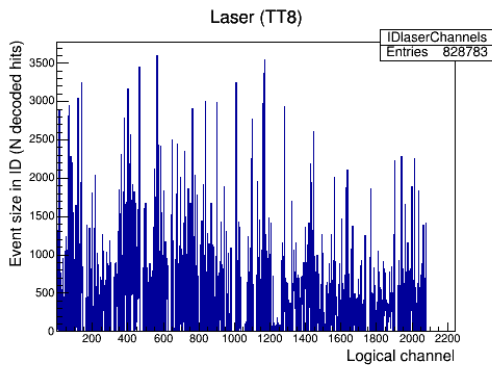
Example of bad behaviour:  
External muon hits have a pulse shape corresponding to a physical event rather than dark noise → the hits actually come from internal muons that get wrongly tagged as external.

## Canvas [06] ID: event size VS channels

Note: similar to canvas 15 for OD

**Laser** is issued once every 2 seconds. We can check the average event size from canvas 02. The amount of hits in all the TT8 events throughout the whole run:  
 $0.5 \text{ ev/s} * 80 \text{ hits/ev} * 6 * 60 * 60 \text{ s}$   
 $= 864000 \text{ hits}$   
 Spread out among ~1300 channels  
 → on average **665 hits** in one channel

**Pulser** is issued with a uniform frequency. It gives precisely 1 hit to each channel during one event.  
 The total amount can be estimated as:  
 $0.1 \text{ pulser events/s} * 1 \text{ hit/event/channel}$   
 $* 6 * 60 * 60 \text{ s} =$   
**2160 hits/channel**



**Random** trigger is issued with a uniform frequency. We can check the average event size from canvas 02. The amount of hits in all the TT64 events throughout the whole run:  
 $0.5 \text{ ev/s} * 3.5 \text{ hits/ev} * 6 * 60 * 60 \text{ s}$   
 $= 37800 \text{ hits}$   
 Spread out among ~1300 channels →  
 on average **29 hits** in one channel.

The frequency of **internal events** is ~20 Hz (canvas 03), and the event size is 54 dec. hits averaging out TT1&BTB0 (36) and TT1&BTB4 (6000) events (canvas 02).  
 $20 \text{ ev/s} * 54 \text{ hits/ev} * 6 * 60 * 60 \text{ s}$   
 $= 23328000 \text{ hits}$   
 Spread out among ~1300 channels →  
 on average **~18000 hits** in one channel.

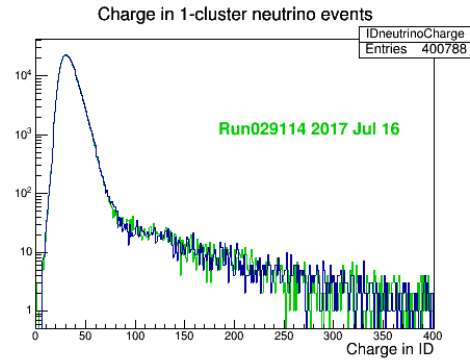
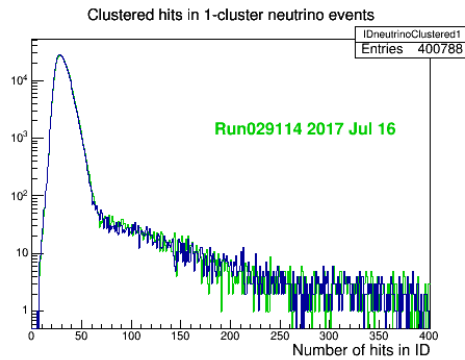
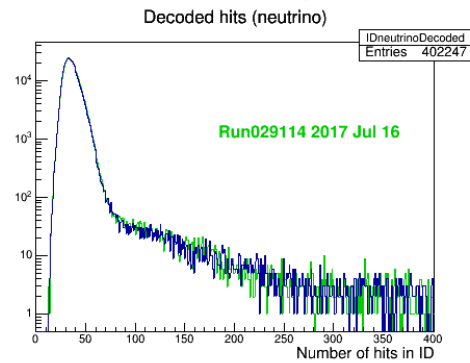
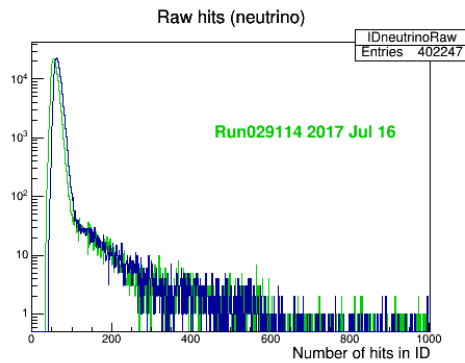


## Canvas [07] ID: hits in neutrino trigger

Compare the current run (blue) to a "good" run (green).  
All these histograms have been shown in canvas 04, but here they are in log scale.  
Recall: low energy peak corresponds to  $^{14}\text{C}$  events which dominate TT1&BTB0  
(see Borexino spectrum and canvases 01, 02 and 04)

The **raw hits** in "neutrino" events. A small shift of the peak is possible. If the shift is large it could mean oscillating boards.

The **decoded hits** in "neutrino" events are "cleaned up" raw hits (check canvas 04)



The **clustered hits** in "neutrino" events are decoded hits that form clusters (check canvas 04)

The **charge** is the number of photoelectrons which should correspond to the number of hits because we work in the single p.e. mode (check canvas 04)

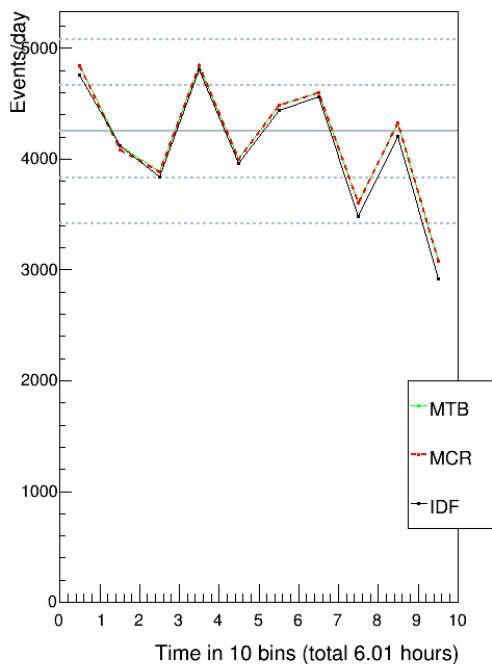
## Canvas [08] OD: muon flag stability

BUG FIX

**MTB:** TT1 & BTB4**MCR:** TT1 & cluster in the OD**IDF:** another way to determine if a TT1 event is an **internal muon** without looking at the OD (using ID pulse shape)

MTB and MCR should overlap perfectly, since all muons produce a cluster in the OD. IDF has a slightly worse efficiency.

Internal muons (TT1&amp;BTB4)

**MTB:** TT2 events which are 75% light leak

→ much higher rate

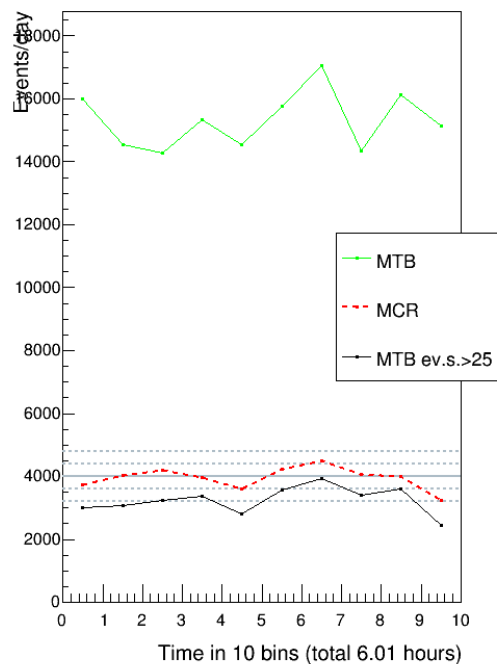
**MCR:** TT2 events that have a cluster

→ actually muons → lower rate

**MTB ev.s. > 25:** TT2 events with >25 dec. hits

→ even lower rate because with this cut we remove light leak and also some actual muons

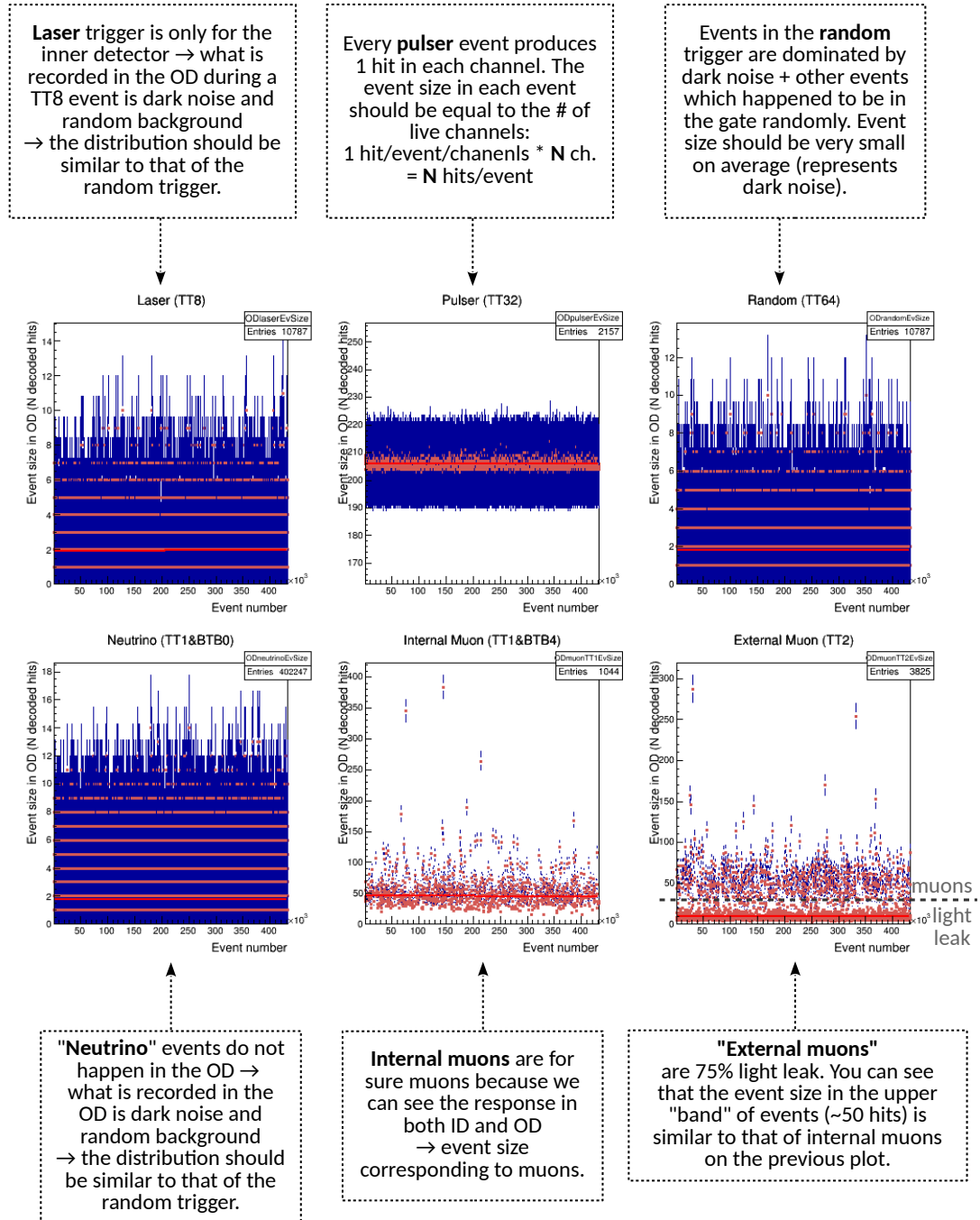
External muons (TT2)



If any of these rates is too low or too high compared to the expected muon rate (especially in the beginning or in the end), it might mean electronics failure. Muon rates are shown on canvas 11. Muon rates can also be checked by FADC. Light leak can be seen on canvas 13.

## Canvas [09] OD: event size VS evnum

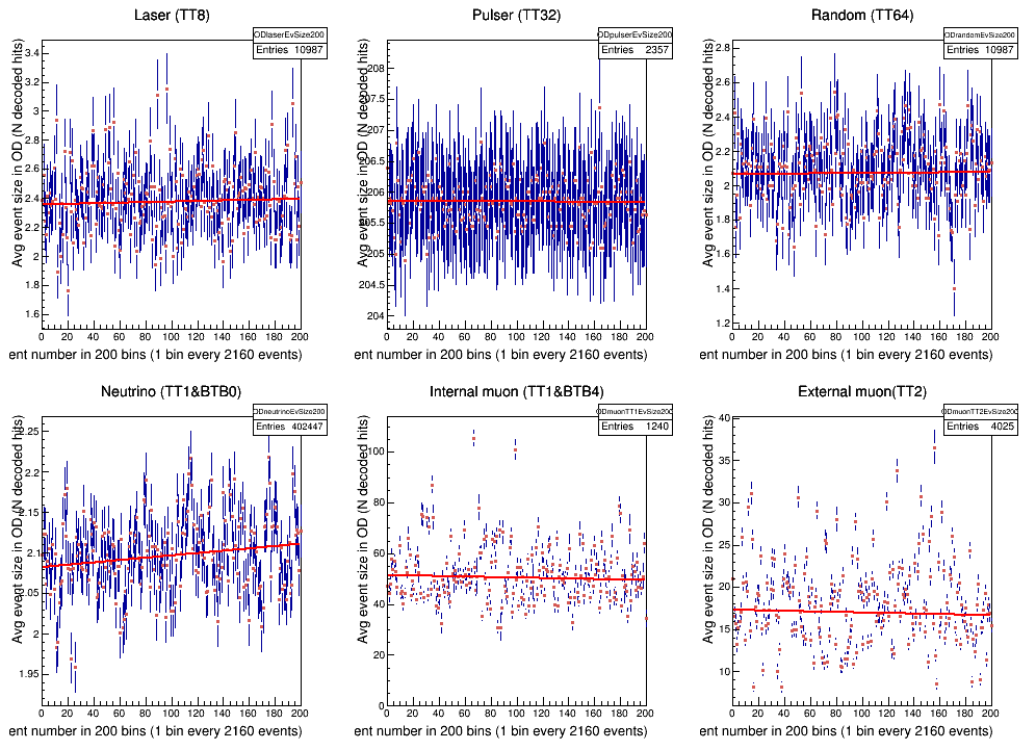
Note: similar to canvas 01 for ID



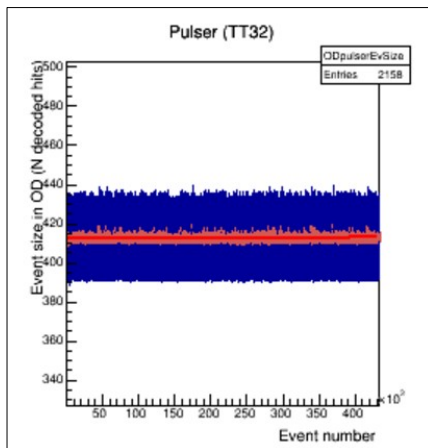
### Canvas [10] OD: event size VS evnum (200 bins)

*Note: this canvas is averaged out stuff from canvas 09*

*Note: similar to canvas 02 for ID*



Some examples of bad behaviour



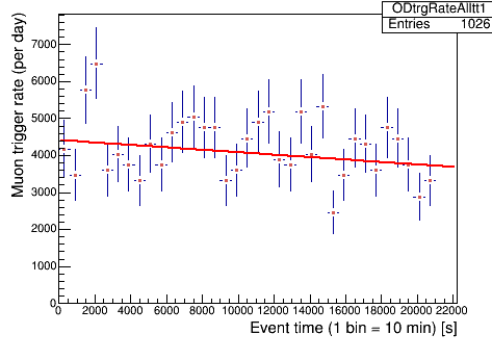
Pulser produced 2 pulses instead of 1, each giving 1 hit  
 → event size is twice larger than it should be.  
 → still OK, validate but make a note.  
 (this can also be seen from the pulse shape in OD on canvas 14)

**Canvas [11] OD: trigger rate***Note: similar to canvas 03 for ID***BUG FIX**

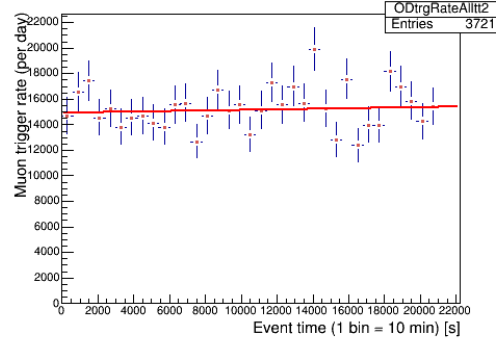
**Internal muons** are expected to have a rate between 2800 and 5800 events/day

**"External muon"** (TT2) events are 75% light leak, therefore we see a much higher rate than the one of internal muons.

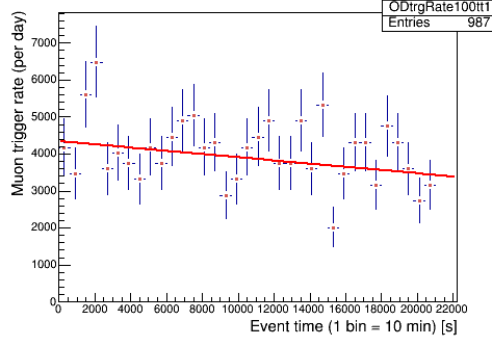
Internal muon (TT1&amp;BTB4) all



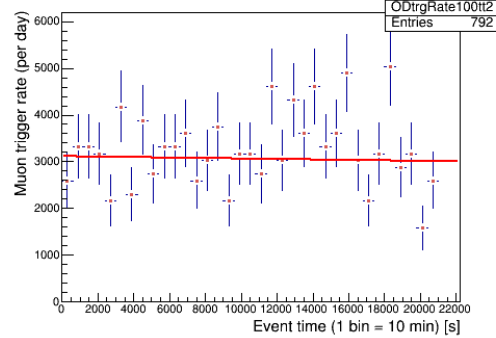
External muon (TT2) all



Internal muon (TT1&amp;BTB4) ev.size &gt; 25



External muon (TT2) ev.size &gt; 25



Most of the **internal muons** have **event size > 25 hits** in the OD. There is little to no difference between this plot and the one above.

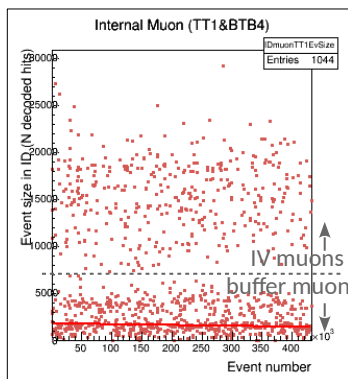
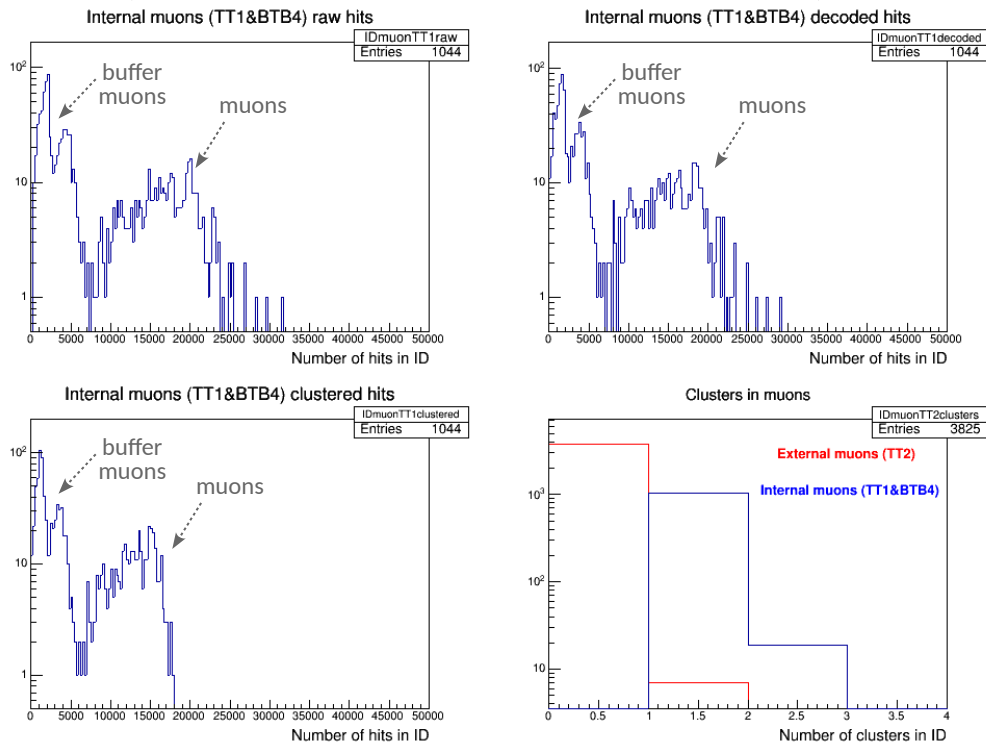
After selecting only **external muon** events with **event size > 25 hits** we remove the light leak and end up with actual muon events. One can see that the rate of such events is comparable to the one of internal muons.

If any of these rates is too low or too high compared to the expected muon rate (especially in the beginning or in the end), it might mean electronics failure. Muon rates are shown on canvas 08 as well. Muon rates can also be checked by FADC.

## Canvas [12] ID: hits in muon triggers

Note: similar to canvas 13 for OD

Check canvas 01: the event size of internal muon events has a distribution at around 15000 hits corresponding to the hits coming from the muon going through the inner vessel, and a peak at lower energy corresponding to the muons only passing through the buffer which does not have that much scintillator in it.



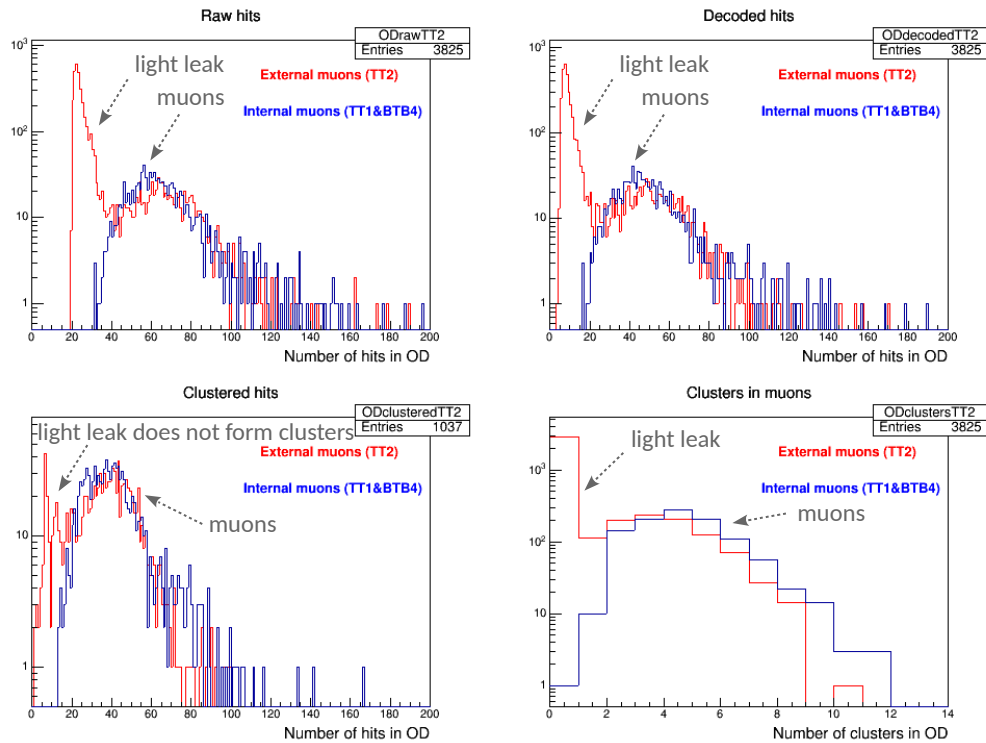
**External muons** do not pass through the ID → TT2 events are dominated by 0-clusters (dark noise) with several 1-cluster coincidental events.  
**Internal muons** are dominated by 1-cluster events with some very energetic 2-cluster events or coincidences.

A plot from canvas 01 for reference.

### Canvas [13] OD: hits in muon triggers

*Note: similar to canvas 12 for ID*

Note on all the canvases that TT2 events are 75% light leak. Light leak does not form clusters and does not produce a large number of hits. Ignoring the light leak you will see that the distributions for internal and external muons look the same → what one would expect, because those are the same kind of muons, just passing through the detector differently. And there is obviously no light leaking into the IV because there's the stainless steel sphere there. Also remember how we did a cut of  $ev.size > 25$  on canvas 11 to get rid of light leak? Here you can see why it was 25 decoded hits.

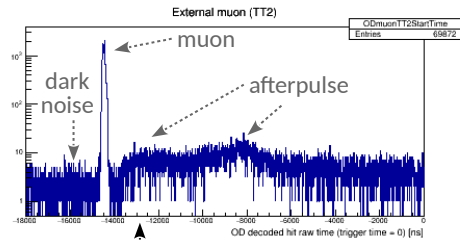
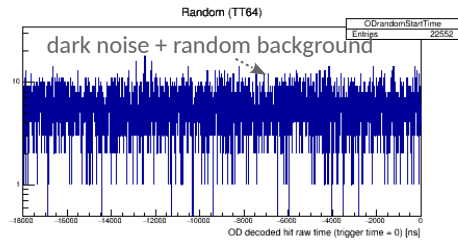
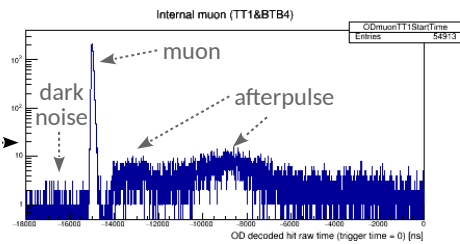
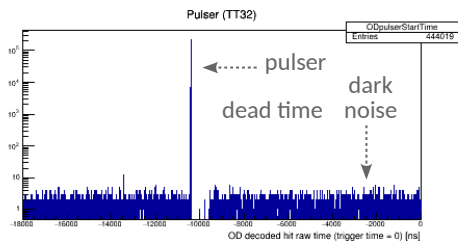
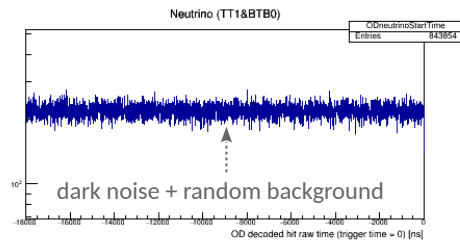
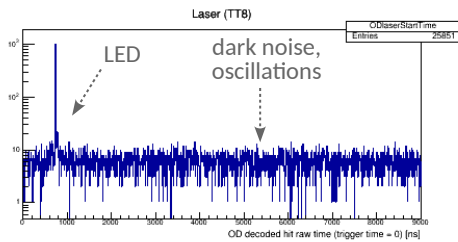


## Canvas [14] OD: pulse shape

Note: similar to canvas 05 for ID

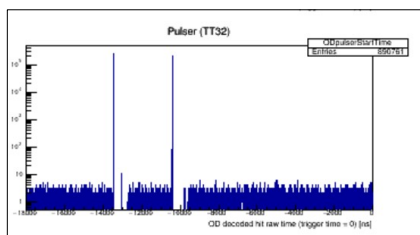
**Laser** events happen only in the ID, but the LEDs in the OD react to it as well → can see some pulse shape.

**Neutrino** events do not happen in the OD → what is recorded is the same as for the random trigger. The distribution should look flat for the same reason.



**External and internal muons** have a similar behaviour in the OD which is what we would expect.

### Some examples of bad behaviour



Pulser produced 2 pulses instead of 1, each giving 1 hit  
→ event size is twice larger than it should be.  
→ still OK, validate but make a note.  
(this can also be seen from the event size in OD on canvases 09 and 10)



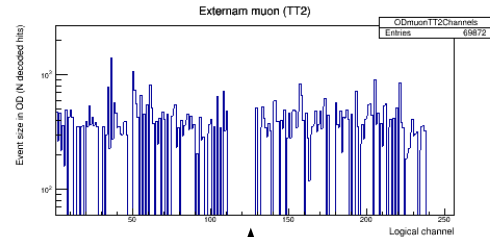
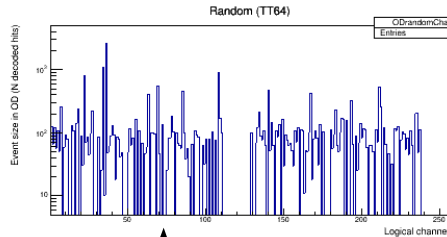
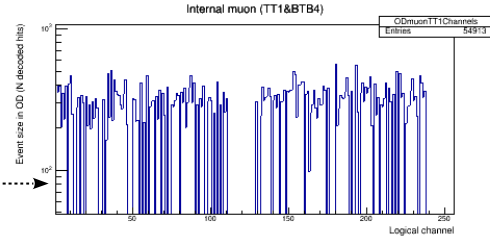
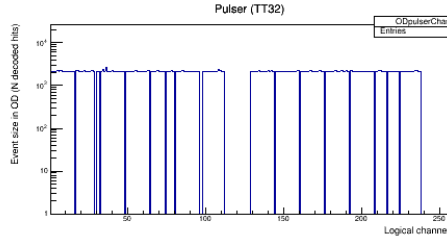
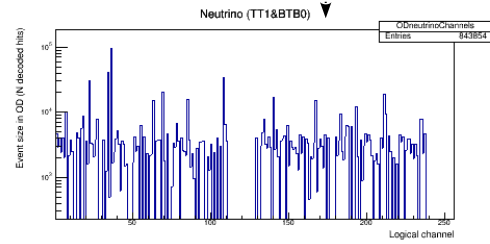
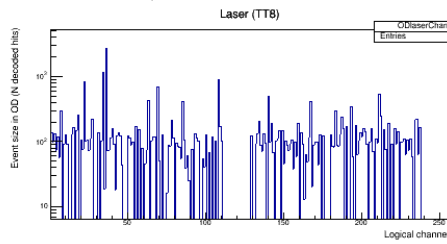
## Canvas [15] OD: event size VS channels

Note: similar to canvas 06 for ID

**Laser** is issued once every 2 seconds. The event size is ~2 dec. hits corresponding to dark noise (canvas 09). The amount of hits in all the TT8 events throughout the whole run:  
 $0.5 \text{ ev/s} * 2 \text{ hits/ev} * 6 * 60 * 60 \text{ s} = 21600 \text{ hits}$   
 Spread out among ~200 channels  
 → on average **108 hits** in one channel

**Pulser** is issued with a uniform frequency. It gives precisely 1 hit to each channel during one event. The total amount can be estimated as:  
 $0.1 \text{ pulser events/s} * 1 \text{ hit/event/channel} * 6 * 60 * 60 \text{ s} = 2160 \text{ hits/channel}$

The frequency of **neutrino events** is ~20 Hz (canvas 03), and the event size is ~2 dec. hits corresponding to dark noise (canvas 09).  
 $20 \text{ ev/s} * 2 \text{ hits/ev} * 6 * 60 * 60 \text{ s} = 864000 \text{ hits}$   
 Spread out among ~200 channels → on average **~4320 hits** in one channel.



**Random** trigger is issued with a uniform frequency. We can check the average event size from canvas 09. The amount of hits in all the TT64 events throughout the whole run:  
 $0.5 \text{ ev/s} * 2 \text{ hits/ev} * 6 * 60 * 60 \text{ s} = 21600 \text{ hits}$   
 Spread out among ~200 channels → on average **108 hits** in one channel.

**Internal muons** have a rate or ~4000 events/day and produce ~50 dec. hits per event in the OD (canvas 10). Hits accumulated throughout the run:  
 $4000 \text{ ev/day} * 50 \text{ hits/ev} * 6 \text{ hrs} * 1 \text{ day/24 hrs} = 50000 \text{ hits}$   
 Spread out among ~200 channels  
 → on average **250 hits** in one channel.

**"External muons"** (75% light leak) have a rate of ~16000 events/day and produce on average ~17 dec. hits per event in the OD (canvas 10). Hits accumulated throughout the run:  
 $16000 \text{ ev/day} * 17 \text{ hits/ev} * 6 \text{ hrs} * 1 \text{ day/24 hrs} = 68000 \text{ hits}$   
 Spread out among ~200 channels → on average **340 hits** in one channel.

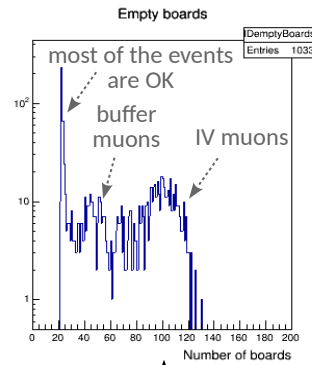
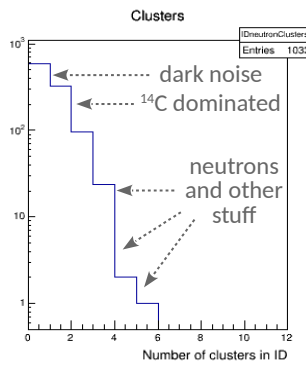
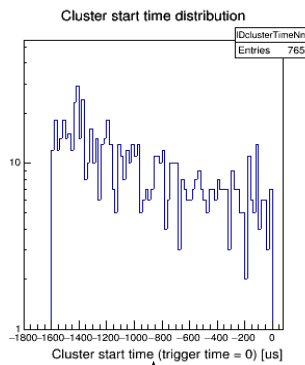
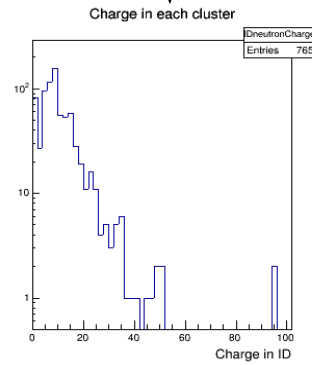
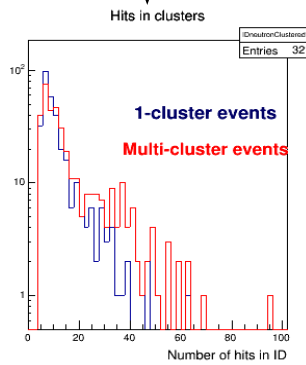
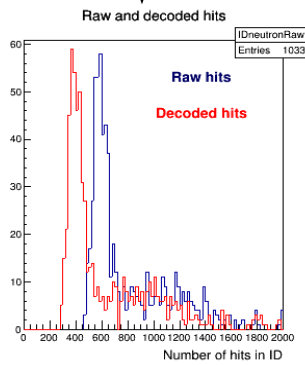
## Canvas [16] ID: neutron trigger (TT128)

NEW

Decoded hits are "cleaned up" raw hits → shift due to removing "bad" hits.

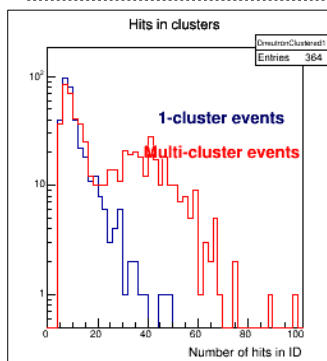
The distribution of hits in a cluster should not differ whether the cluster belongs to a one-cluster event or to a multi-cluster event.

Charge is comparable to number of hits since 1 photon = 1 hit = 1 photoelectron.



TT128 gate is opened after an internal muon i.e. TT1&BTB4.  
It is dominated by  $^{14}\text{C}$  events and dark noise, therefore clusters that fall into the gate don't have a particular start time relative to when we issue the trigger → happen uniformly throughout the gate.

We always have 20 empty boards because many PMTs died and the corresponding channels are not used. Higher amount of empty boards is caused by some muons "blinding" the channels by the vast amount of light muons produce.



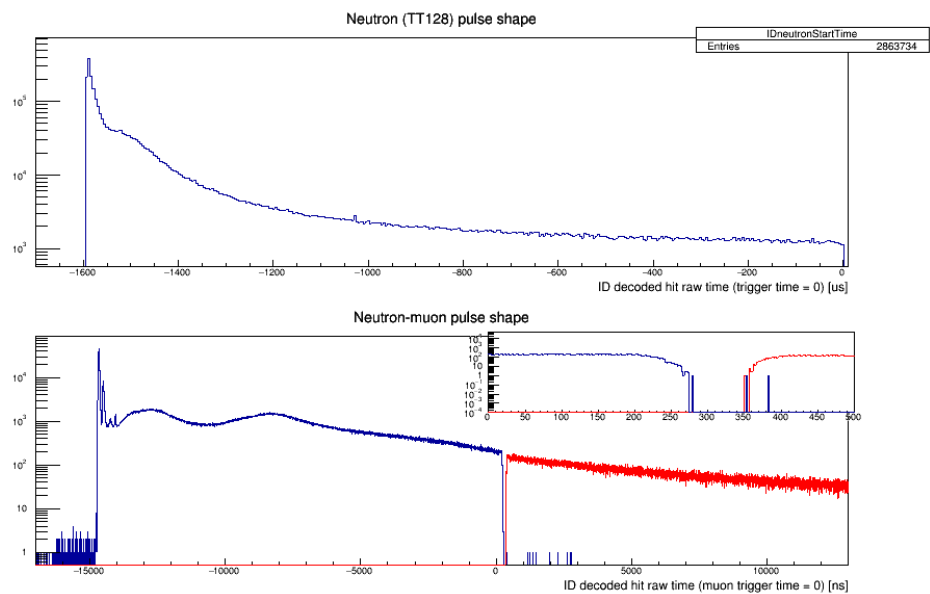
→ validate and leave a comment

Canvas [17] ID: neutron trigger (TT128) pulse shape

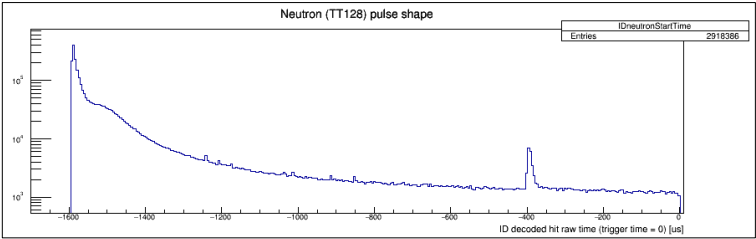
Note: similar to canvas 11 for OD

NEW

This canvas shows the raw time of decoded hits in TT128 events in the gate relative to trigger time and relative to the related internal muon event.



Caused by a second neutron falling into the gate  
→ validate and make a comment





## Appendix B

# Additional information on the EQE of the Borexino PMTs

This appendix contains additional information regarding my work on the **relative effective quantum efficiency** (REQE) of the Borexino PMTs, presented in Chapter 6.

In Sec. B.1, I will show the time evolution of the REQE of single PMTs, as opposed to the *average* REQE of all enabled PMTs or a subset of PMTs, shown in Sec. 6.6. Sec. B.2 will demonstrate the time evolution of the number of live PMTs in different PMT subsets, as obtained during the REQE calculation procedure.

### B.1 Time evolution of REQE of single PMTs

In Sec. 6.6, I have shown the REQE trends of subsets of PMTs, demonstrated here for reference by Fig. B.1. In this figure, **L2019** denotes the subset of PMTs that stayed live until July 2019.

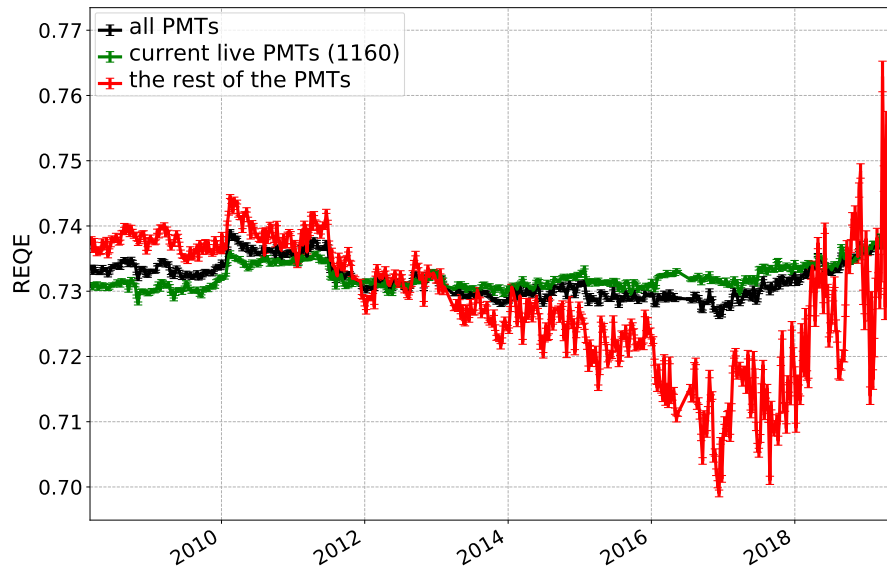


FIGURE B.1: Average REQE among all live PMTs in a given week (black), L2019 PMTs (green) and the remaining PMTs in each week (red).

The average REQE of the PMTs not belonging to L2019 (i.e. PMTs that died at some point) shows a steady decrease, followed by a turning point after which it starts increasing.

My contemplation regarding this trend is that:

- the quality of PMTs that are destined to die before July 2019 worsens in time, as demonstrated not only by REQE, but other parameters, as discussed in Sec. 5.3.1 (see Fig. 5.11);
- after such PMTs die, the average REQE of the remaining non-L2019 PMTs effectively increases, as the worst ones remove themselves from the set.

In order to test this idea, I have looked at the average REQE of single PMTs. The Borexino PMTs are labeled with a four-digit number based on their location on the stainless steel sphere, and a sign denoting whether they belong to the top (+, omitted) or bottom (-) hemisphere.

First, for reference, Fig. B.2 displays the time evolution of the REQE of several PMTs belonging to the **B900** set that were **live** as of July 2019. As one can see, their REQE is stable in time, apart from minor fluctuations. An exception can be seen in the behavior of the PMT -2106 (green), which showed rapidly worsening REQE, which as rapidly returned back to normal. This might have happened due to malfunction of the Laben channel, to which this PMT is connected, which was fixed.

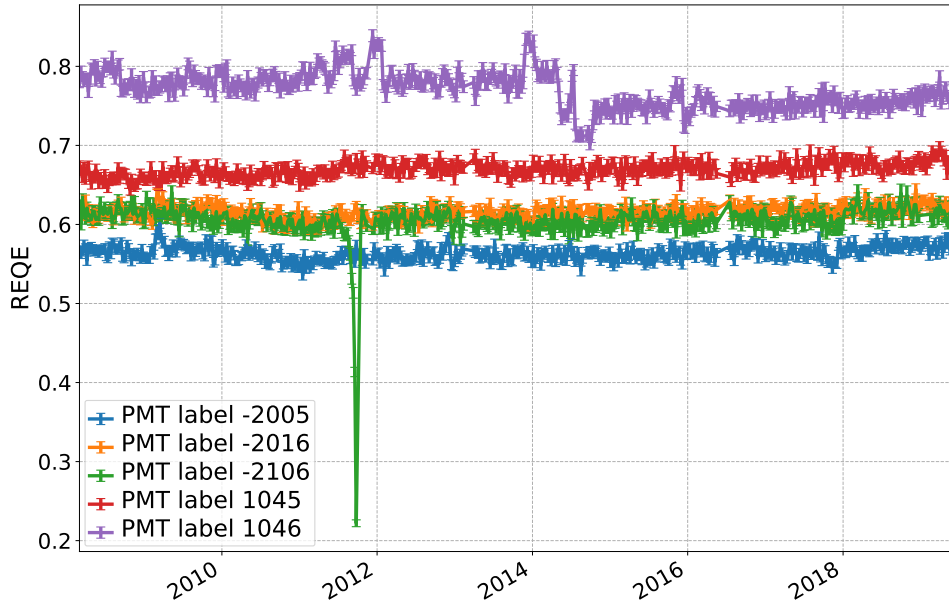


FIGURE B.2: REQE of live B900 PMTs.

In Fig. B.3, I show the behavior of **B900** PMTs that are **not live** anymore. PMTs -1514 (orange), 866 (red), and 738 (green) demonstrate “peaceful death”. Their REQE remained stable in time until the moment they were disconnected. This way of dying is typical for B900 PMTs. Some other PMTs, like -1134 (blue) and 832 (purple), show unstable behavior throughout their life and particularly close to the end. Ideally, these PMTs should not belong to the B900 subset, the purpose of which is to select only stable good quality ones. However, the selection method is very simple, and does not fully exploit possible correlations between PMT parameters and their quality or lifetime. Some ideas on how to improve this selection are presented in the [Outlook](#).

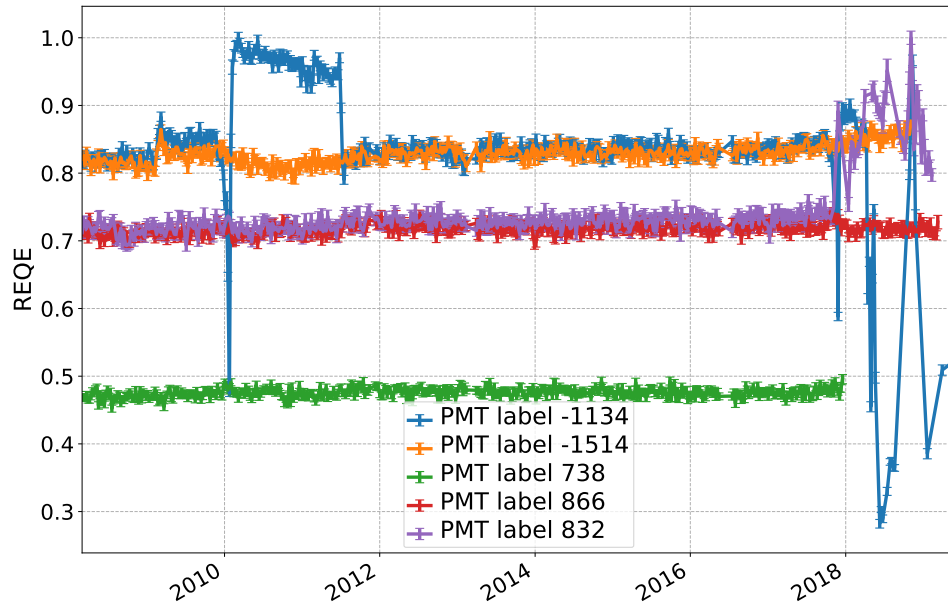


FIGURE B.3: REQE of B900 PMTs that are not live in July 2019.

Among the 1160 PMTs **live** in July 2019, 900 (minus the possibly temporarily disabled ones) belong to B900, and the remaining ones **do not**. Figure B.4 shows the behavior of several such PMTs. Some of them, like PMT 1618 (purple) and -1724 (orange) show perfectly stable behavior. Some, however, show excessively chaotic behavior, such as -2118 (blue) and -2117 (green). Missing points in Figure B.4 represent the weeks in which these PMTs were disabled, or discarded by the REQE procedure as not having enough statistics for REQE calculation (see Sec. 6.4.2).

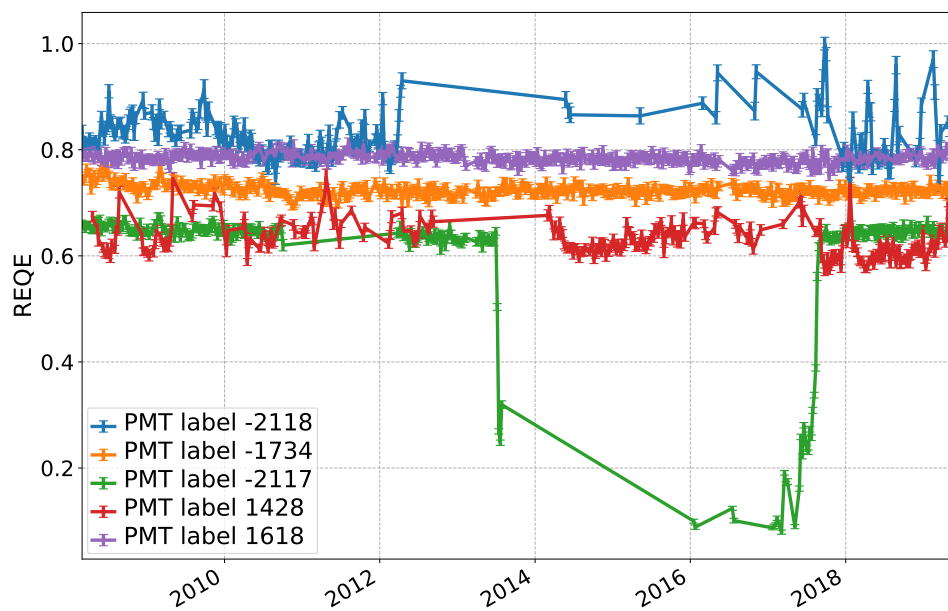


FIGURE B.4: REQE of live PMTs that do not belong to the B900 set.

Finally, PMTs that do **not** belong to the B900 set, and are **not live** anymore, display various trends as well, shown in Fig. B.5. While some show stable (-2107, red) or relatively stable (-2110, orange) behavior, and some show chaotic changes (-2114, blue), a new trend, not seen before, is visible in the death of PMT -2112 (green). Unlike the others, it worsens steadily and dies. The removal of the contribution of such PMTs is what could explain the trend change displayed by the red points in Fig. B.1.

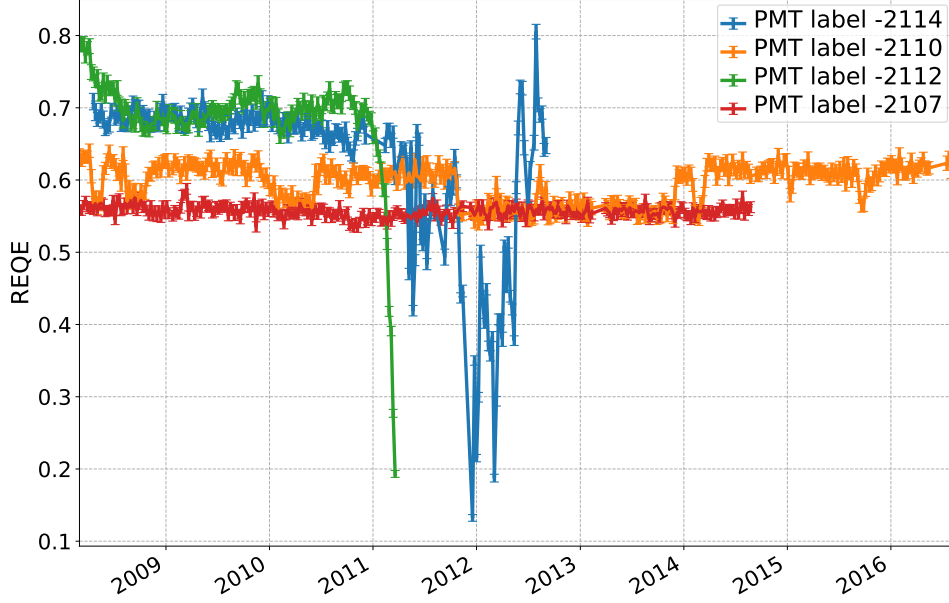


FIGURE B.5: REQE of PMTs that do not belong to the B900 set and are not live in July 2019.

Overall, the fact that during the same week some PMTs display perfectly stable behavior, and some excessively chaotic one, indicates that this behavior is not related to general changes in the detector, but to the individual PMT in consideration. Such changes may appear due to fluctuations in the electronics and hardware, or due to the aging of the PMT, as very often they are correlated with its death. The REQE calculation procedure, as well as other potential studies, would benefit greatly from more studies on the behavior of single PMTs and the investigation of possible reasons behind it; as well as an improved selection of a stable PMT set.



## B.2 Evolution of live PMTs

The REQE procedure obtains the information about Laben channels the information from which should be ignored from a database of disabled channels, filled during the DAQ procedure (*disabled* PMTs); as well as from judging their behavior based on the number of hits from  $^{14}\text{C}$  events in the data sample of the given week (*discarded* PMTs, see Sec. 6.4.2). In Fig. B.6, disabled PMTs are excluded from the sample, while discarded PMTs are assigned the REQE value calculated in the previous week. Figure B.6 presents a “consistency check” of the REQE calculation procedure: the number of all live PMTs in the given week (black) becomes equal to the number of L2019 PMTs in the week corresponding to the last analyzed week of July 2019, as “current” live PMTs become “all” PMTs in that time point. Note that the number of (enabled) L2019 PMTs is very stable, displaying only minor fluctuations, as few of them get disabled temporarily. The number of non-L2019 PMTs (red) slowly decreases and becomes zero in the last analyzed week, as no non-L2019 PMTs are live at that moment.

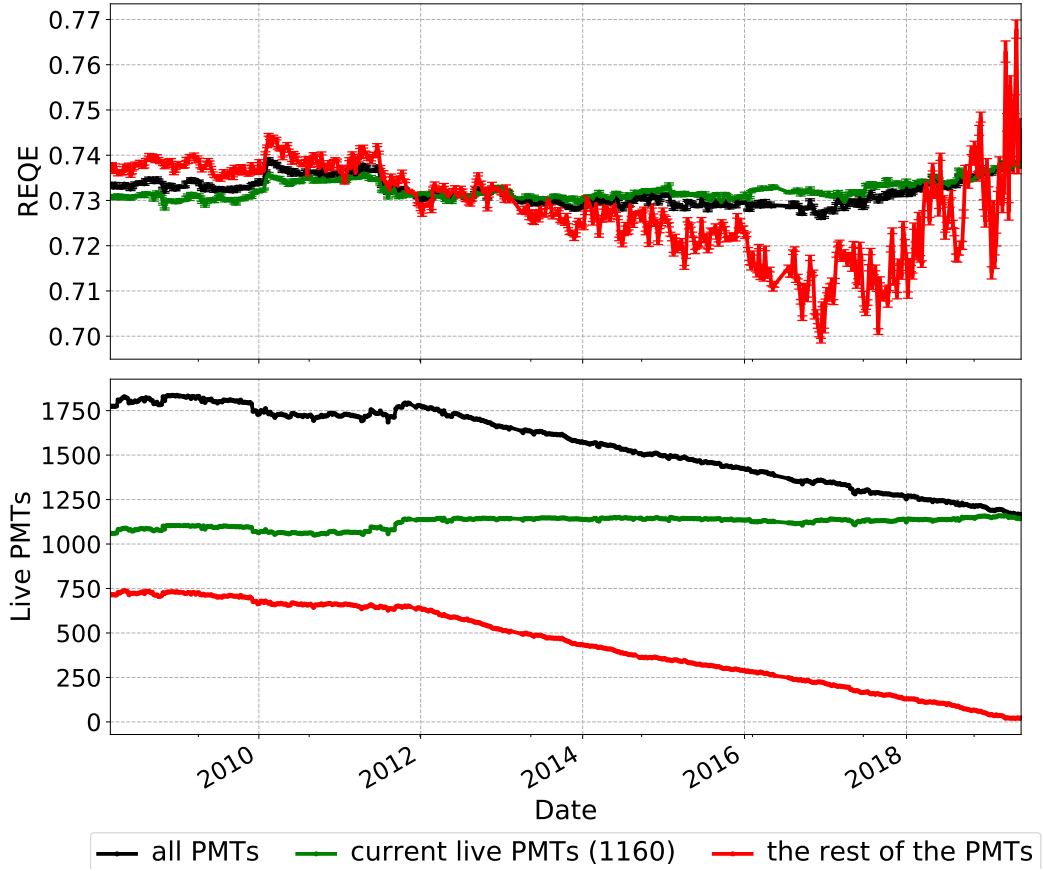


FIGURE B.6: Average REQE of all, L2019 and non-L2019 PMTs and the number of enabled PMTs among that group set in each time point.



## Appendix C

# Analytical approach to reference shape construction

As discussed in Sec. 7.1.3, the goal of modeling the detector analytically is to obtain a formula which would convert the deposited energy of the particle to an energy estimator. The first step in describing the physical processes following an event is obtaining the **mean number of photoelectrons** ( $N_{pe}$ ) after an event of a certain energy:

$$\hat{N}_{pe}(E) = \epsilon_{sc}N_{sc}(E) + \epsilon_{ch}N_{ch}(E), \quad (C.1)$$

where  $E$  is the energy of the particle in MeV,  $N_{sc}$  is the number of emitted **scintillation photons**,  $N_{ch}$  is the number of photons due to the **Cherenkov radiation**, and  $\epsilon_{sc}$  and  $\epsilon_{ch}$  are respective efficiencies of conversion of photons to photoelectrons that include absorption, reemission and PMT quantum efficiencies. The scintillation contribution is calculated as follows:

$$N_{sc}(E) = YEQ(E), \quad (C.2)$$

where  $\underline{Y}^1$  is the light yield in photoelectrons per MeV, and  $Q(E)$  is the so-called **quenching factor** which describes deviations from the linear relation  $N_{sc} = YE$ :

$$Q(E) = \frac{1}{E} \int_0^E \frac{dE}{1 + k_B \frac{dE}{dx}}. \quad (C.3)$$

In the formula above,  $k_B$  is the so-called Birks parameter which is a constant that depends on the considered particle and the medium it is traversing; and  $\frac{dE}{dx}$  is the average ionization particle energy loss.

The number of Cherenkov photons can be obtained as follows (valid for  $e^-$ ):

$$N_{ch}(E) = (A_0 + A_1x + A_2x^2 + A_3x^3)(1 + A_4E), \quad (C.4)$$

where  $x = \ln(1 + \frac{E}{E_0})$ ,  $E_0 = 0.165$  MeV (Cherenkov threshold). The coefficients  $A_i$  are fixed based on the MC simulation.

---

<sup>1</sup>Underlined variables are parameters in the multivariate likelihood function of the analytical approach that can be free, fixed or constrained in the fit. They will be summarized later in Table C.1

In order to absorb several parameters into one, a so-called Cherenkov tuning parameter  $f_{Cher}$  is introduced, representing the relative contribution compared to scintillation light:

$$f_{Cher} = \frac{\epsilon_{ch}}{\epsilon_{sc}Y}, \quad (C.5)$$

which reformulates Eq. C.1 in the following way:

$$\hat{N}_{pe}(E) = Y \left( Q(E)E + f_{Cher}N_{ch}(E) \right). \quad (C.6)$$

The next step is to obtain the **mean number of fired PMTs** ( $\hat{N}_p$ ) based on  $\hat{N}_{pe}$  calculated above:

$$\hat{N}_p = N_{tot}(1 - e^{-\mu} [1 + p_t\mu])(1 - g_C\mu), \quad (C.7)$$

where  $N_{tot}$  is the total number of PMTs (2000 for the normalized variable) and

$$\mu = \frac{\hat{N}_{pe}(E)}{N_{tot}}. \quad (C.8)$$

The parameters  $p_t$  and  $g_C$  are related to the single electron response and geometric correction, respectively. The LER energy response function is a scaled Poisson function of the energy estimator  $N_p$ :

$$f(N_p) = \frac{m^{N_p s}}{(N_p s)!} e^{-m}, \quad (C.9)$$

where

$$m = \frac{\hat{N}_p^2(E)}{\sigma_p^2}, \quad (C.10)$$

and

$$s = \frac{\sigma_p^2}{\hat{N}_p(E)}. \quad (C.11)$$

The two parameters  $m$  and  $s$  use the mean value  $\hat{N}_p(E)$  expressed above in Eq. C.7, and the variance  $\sigma_p^2$  of the  $N_p$  variable. The variance contains several parameters that contribute to the final response function:

$$\sigma_p^2 = \frac{N_{tot}}{N_{live}} [1 - (1 + v_1)p_1] \hat{N}_p(E) + v_T^0 \hat{N}_p^3(E) + v_T^q \left( \mu \frac{p_0}{p_1} \right)^2 \hat{N}_p^2(E) + v_N \hat{N}_p(E) + \sigma_d^2, \quad (C.12)$$

where  $N_{live}$  is the number of working PMTs in the selected time period,  $p_1 = 1 - e^{-\mu}$  is the probability of having a signal at any PMT,  $p_0 = e^{-\mu}$  is the probability of absence of the signal,  $v_1$ ,  $v_T^0$ ,  $v_T^q$ ,  $v_N$  are parameters related to resolution and uniformity of light collection, and  $\sigma_d$  is the PMT dark noise contribution.

The summary of the parameters of the analytical function is shown in Table C.1. Four of these parameters are free (LY and three energy resolution parameters), while the rest are obtained from measurements or Monte Carlo (MC) studies.

Parameter	Fixed/Free	Meaning
$Y$	free	Light yield
$f_{Cher}$	1.0	Relative weight of the scintillation and Cherenkov light, fixed using analytical fit studies.
$p_t$	0.12	Fraction of single p.e. charge spectrum below the electronics threshold obtained from calibrations.
$g_C$	0.101	Geometric correction factor tuned via a fit of Eq. C.7 using MC monoenergetic $e^-$ samples.
$v_1$	0.16	Relative variance of the probability that a PMT triggers due to events uniformly distributed in the IV. Calculated using MC.
$v_T^0$	free	Parameter representing spatial nonuniformity of $N_p$ .
$v_T^\alpha$	free	Parameter corresponding to $v_T^0$ , used in the simplified form of Eq. C.12 which involves only the first two terms, used for $\alpha$ particles ( $^{210}\text{Po}$ ).
$v_T^q$	7.0	Parameter representing nonuniformity of light collection, calculated using MC.
$v_N$	free	Intrinsic scintillator resolution parameter for $\beta$ s.
$\sigma_d$	0.23 (0.4) $N_p^{dt_1(2)}$	PMT dark noise contribution.

TABLE C.1: Parameters of the analytical response function that contribute to the multivariate likelihood [115].



# Bibliography

- [1] J. Chadwick, “Intensitätsverteilung im magnetischen Spectrum der  $\beta$ -Strahlen von radium B + C,” *Verhandl. Dtsc. Phys. Ges.*, vol. 16, p. 383, 1914. [Cited on pages 1 and 2.]
- [2] W. Pauli, “Brief an die Gruppe der Radioaktiven (1930),” *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg*, 1985. [Cited on pages 1 and 2.]
- [3] E. Fermi, “Tentativo di una teoria dell’emissione dei raggi beta,” *Ric. Sci.*, vol. 4, pp. 491–495, 1933. [Cited on pages 1 and 2.]
- [4] B. Pontecorvo, “Inverse beta process,” *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 1, pp. 25–31, 1991. [,97(1946)]. [Cited on pages 1 and 2.]
- [5] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, “Detection of the free neutrino: A Confirmation,” *Science*, vol. 124, pp. 103–104, 1956. [Cited on pages 1 and 2.]
- [6] “The Nobel Prize in Physics 1995.” <https://www.nobelprize.org/prizes/physics/1995/summary>, 1995. [Cited on pages 1 and 2.]
- [7] B. Pontecorvo, “Mesonium and anti-mesonium,” *Sov. Phys. JETP*, vol. 6, p. 429, 1957. [*Zh. Eksp. Teor. Fiz.*33,549(1957)]. [Cited on pages 1 and 2.]
- [8] B. Pontecorvo, “Inverse beta processes and nonconservation of lepton charge,” *Sov. Phys. JETP*, vol. 7, pp. 172–173, 1958. [*Zh. Eksp. Teor. Fiz.*34,247(1957)]. [Cited on pages 1 and 2.]
- [9] G. Danby, J. M. Gaillard, K. A. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz, and J. Steinberger, “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos,” *Phys. Rev. Lett.*, vol. 9, pp. 36–44, 1962. [Cited on pages 1 and 2.]
- [10] “The Nobel Prize in Physics 1998.” <https://www.nobelprize.org/prizes/physics/1988/summary>, 1995. [Cited on pages 1 and 2.]
- [11] Z. Maki, M. Nakagawa, and S. Sakata, “Remarks on the unified model of elementary particles,” *Prog. Theor. Phys.*, vol. 28, pp. 870–880, 1962. [,34(1962)]. [Cited on pages 1, 3, and 9.]
- [12] M. Nakagawa, H. Okonogi, S. Sakata, and A. Toyoda, “Possible existence of a neutrino with mass and partial conservation of muon charge,” *Prog. Theor. Phys.*, vol. 30, pp. 727–729, 1963. [,45(1963)]. [Cited on pages 1, 3, and 9.]
- [13] R. Davis, Jr., D. S. Harmer, and K. C. Hoffman, “Search for neutrinos from the sun,” *Phys. Rev. Lett.*, vol. 20, pp. 1205–1209, 1968. [Cited on pages 1, 2, and 20.]

- [14] “The Nobel Prize in Physics 2002.” <https://www.nobelprize.org/prizes/physics/2002/summary>, 1995. [Cited on pages 1 and 2.]
- [15] M. L. Perl *et al.*, “Evidence for Anomalous Lepton Production in  $e^+ - e^-$  Annihilation,” *Phys. Rev. Lett.*, vol. 35, pp. 1489–1492, 1975. [Cited on pages 1, 3, and 8.]
- [16] S. M. Bilenky and B. Pontecorvo, “REACTOR EXPERIMENTS AND SOLAR NEUTRINO PROBLEM,” *Lett. Nuovo Cim.*, vol. 40, p. 161, 1984. [Cited on pages 1, 3, and 9.]
- [17] Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998. [Cited on pages 1, 3, and 11.]
- [18] E. Kearns, T. Kajita, and Y. Totsuka, “Detecting massive neutrinos,” *Sci. Am.*, vol. 281N2, pp. 48–55, 1999. [Spektrum Wiss. Dossier2003N1,64(2003)]. [Cited on pages 1 and 3.]
- [19] K. Kodama *et al.*, “Observation of tau neutrino interactions,” *Phys. Lett.*, vol. B504, pp. 218–224, 2001. [Cited on pages 1 and 8.]
- [20] Q. R. Ahmad *et al.*, “Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory,” *Phys. Rev. Lett.*, vol. 89, p. 011301, Jun 2002. [Cited on pages 1 and 4.]
- [21] “The Nobel Prize in Physics 2015.” <https://www.nobelprize.org/prizes/physics/2015/summary>, 1995. [Cited on pages 1 and 11.]
- [22] M. Bardon, K. Lande, L. M. Lederman, and W. Chinowsky, “Long-lived neutral K mesons,” *Annals Phys.*, vol. 5, no. 2, pp. 156–181, 1958. [Cited on page 2.]
- [23] J. N. Abdurashitov *et al.*, “Results from SAGE,” *Phys. Lett.*, vol. B328, pp. 234–248, 1994. [Cited on page 3.]
- [24] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, “Measurement of the solar electron neutrino flux with the Homestake chlorine detector,” *Astrophys. J.*, vol. 496, pp. 505–526, 1998. [Cited on page 4.]
- [25] M. Altmann *et al.*, “Complete results for five years of GNO solar neutrino observations,” *Phys. Lett.*, vol. B616, pp. 174–190, 2005. [Not cited.]
- [26] J. N. Abdurashitov *et al.*, “Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period,” *Phys. Rev.*, vol. C80, p. 015807, 2009. [Not cited.]
- [27] K. Abe *et al.*, “Solar neutrino results in Super-Kamiokande-III,” *Phys. Rev.*, vol. D83, p. 052010, 2011. [Not cited.]
- [28] J. Maneira, “Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory,” in *Proceedings, 13th ICATPP Conference on Astroparticle, Particle, Space Physics and Detectors for Physics Applications (ICATPP 2011): Como, Italy, October 3-7, 2011*, pp. 360–366, 2012. [Cited on page 4.]



- [29] D. Waller, “Results from the Sudbury Neutrino Observatory,” *eConf*, vol. C040802, p. WET001, 2004. [Cited on page 4.]
- [30] U. F. Katz and C. Spiering, “High-Energy Neutrino Astrophysics: Status and Perspectives,” *Prog. Part. Nucl. Phys.*, vol. 67, pp. 651–704, 2012. [Cited on page 5.]
- [31] C. Spiering, “Towards High-Energy Neutrino Astronomy. A Historical Review,” *Eur. Phys. J. H*, vol. 37, pp. 515–565, 2012. [Cited on page 5.]
- [32] M. Aartsen *et al.*, “Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data,” *Phys. Rev. D*, vol. 98, no. 6, p. 062003, 2018. [Cited on page 5.]
- [33] A. Gando *et al.*, “Partial radiogenic heat model for Earth revealed by geoneutrino measurements,” *Nature Geo.*, vol. 4, pp. 647–651, 2011. [Cited on page 6.]
- [34] G. Bellini *et al.*, “Observation of Geo-Neutrinos,” *Phys. Lett.*, vol. B687, pp. 299–304, 2010. [Cited on pages 6 and 28.]
- [35] G. Bellini *et al.*, “Measurement of geo-neutrinos from 1353 days of Borexino,” *Phys. Lett.*, vol. B722, pp. 295–300, 2013. [Not cited.]
- [36] M. Agostini *et al.*, “Spectroscopy of geoneutrinos from 2056 days of Borexino data,” *Phys. Rev.*, vol. D92, no. 3, p. 031101, 2015. [Not cited.]
- [37] M. Agostini *et al.*, “Comprehensive geoneutrino analysis with Borexino,” *Phys. Rev.*, vol. D101, no. 1, p. 012009, 2020. [Cited on pages 6 and 28.]
- [38] P. Antonioli *et al.*, “SNEWS: The Supernova Early Warning System,” *New J. Phys.*, vol. 6, p. 114, 2004. [Cited on page 6.]
- [39] P. Alivisatos *et al.*, “KamLAND: A Liquid scintillator anti-neutrino detector at the Kamioka site,” 7 1998. [Cited on page 10.]
- [40] K. Eguchi *et al.*, “First results from KamLAND: Evidence for reactor anti-neutrino disappearance,” *Phys. Rev. Lett.*, vol. 90, p. 021802, 2003. [Cited on page 10.]
- [41] B. Aharmim *et al.*, “Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory,” *Phys. Rev. C*, vol. 72, p. 055502, 2005. [Cited on page 10.]
- [42] S. Abe *et al.*, “Precision Measurement of Neutrino Oscillation Parameters with KamLAND,” *Phys. Rev. Lett.*, vol. 100, p. 221803, 2008. [Cited on page 10.]
- [43] A. Gando *et al.*, “Constraints on  $\theta_{13}$  from A Three-Flavor Oscillation Analysis of Reactor Antineutrinos at KamLAND,” *Phys. Rev. D*, vol. 83, p. 052002, 2011. [Cited on page 10.]
- [44] M. Ahn *et al.*, “Indications of neutrino oscillation in a 250 km long baseline experiment,” *Phys. Rev. Lett.*, vol. 90, p. 041801, 2003. [Cited on page 11.]
- [45] E. Aliu *et al.*, “Evidence for muon neutrino oscillation in an accelerator-based experiment,” *Phys. Rev. Lett.*, vol. 94, p. 081802, 2005. [Cited on page 11.]

- [46] D. Michael *et al.*, “Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam,” *Phys. Rev. Lett.*, vol. 97, p. 191801, 2006. [Cited on page 11.]
- [47] X. Guo *et al.*, “A Precision measurement of the neutrino mixing angle  $\theta_{13}$  using reactor antineutrinos at Daya-Bay,” 1 2007. [Cited on page 11.]
- [48] F. An *et al.*, “Observation of electron-antineutrino disappearance at Daya Bay,” *Phys. Rev. Lett.*, vol. 108, p. 171803, 2012. [Cited on page 11.]
- [49] F. Ardellier *et al.*, “Double Chooz: A Search for the neutrino mixing angle  $\theta(13)$ ,” 6 2006. [Cited on page 11.]
- [50] H. de Kerret *et al.*, “First Double Chooz  $\theta_{13}$  Measurement via Total Neutron Capture Detection,” *Nature Phys.*, vol. 16, no. 5, pp. 558–564, 2020. [Cited on page 11.]
- [51] S.-B. Kim, “RENO for neutrino mixing angle  $\Theta(13)$ ,” *Prog. Part. Nucl. Phys.*, vol. 64, pp. 346–347, 2010. [Cited on page 11.]
- [52] J. Ahn *et al.*, “Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment,” *Phys. Rev. Lett.*, vol. 108, p. 191802, 2012. [Cited on page 11.]
- [53] F. Capozzi, E. Lisi, A. Marrone, and A. Palazzo, “Current unknowns in the three neutrino framework,” *Prog. Part. Nucl. Phys.*, vol. 102, pp. 48–72, 2018. [Cited on page 11.]
- [54] F. Capozzi, E. Lisi, A. Marrone, and A. Palazzo, “Global analysis of oscillation parameters,” *J. Phys. Conf. Ser.*, vol. 1312, no. 1, p. 012005, 2019. [Cited on pages 11 and 128.]
- [55] L. Wolfenstein, “Neutrino Oscillations in Matter,” *Phys. Rev.*, vol. D17, pp. 2369–2374, 1978. [294(1977)]. [Cited on page 11.]
- [56] S. P. Mikheyev and A. Yu. Smirnov, “Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos,” *Sov. J. Nucl. Phys.*, vol. 42, pp. 913–917, 1985. [305(1986)]. [Cited on pages 11 and 12.]
- [57] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics*. 2007. [Cited on page 12.]
- [58] A. Yu. Smirnov, “The MSW effect and solar neutrinos,” in *Neutrino telescopes. Proceedings, 10th International Workshop, Venice, Italy, March 11-14, 2003. Vol. 1+2*, pp. 23–43, 2003. [Cited on pages 12 and 24.]
- [59] A. W. Poon *et al.*, “Solar neutrino observations at the Sudbury Neutrino Observatory,” *eConf*, vol. C020805, p. TTH01, 2002. [Cited on page 12.]
- [60] A. Eddington, “The Internal Constitution of the Stars,” *Nature*, vol. 106, p. 14–20, 1920. [Cited on page 16.]
- [61] H. Bethe and C. Critchfield, “The formation of deuterons by proton combination,” *Phys. Rev.*, vol. 54, p. 248, 1938. [Cited on page 16.]
- [62] C. von Weizsäcker, “Über Elementumwandlungen in Innern der Sterne II,” *Physikalische Zeitschrift*, vol. 39, p. 633, 1938. [Cited on page 17.]

- [63] H. A. Bethe, “Energy production in stars,” *Phys. Rev.*, vol. 55, pp. 434–456, 1939. [Cited on page 17.]
- [64] I. Barnes, Thomas G. and F. N. Bash, eds., *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis in Honor of David L. Lambert: Proceedings, International Symposium, Austin, Texas, 17-19 Jun 2004*, vol. 336, 2005. [Cited on page 18.]
- [65] M. Asplund, N. Grevesse, A. J. Sauval, and P. Scott, “The chemical composition of the sun,” *Annual Review of Astronomy and Astrophysics*, vol. 47, p. 481–522, Sep 2009. [Cited on pages 18 and 19.]
- [66] E. Caffau, H.-G. Ludwig, M. Steffen, B. Freytag, and P. Bonifacio, “Solar Chemical Abundances Determined with a CO5BOLD 3D Model Atmosphere,” *Solar Phys.*, vol. 268, p. 255, 2011. [Cited on page 18.]
- [67] N. Vinyoles, A. M. Serenelli, F. L. Villante, S. Basu, J. Bergström, M. C. Gonzalez-Garcia, M. Maltoni, C. Peña-Garay, and N. Song, “A new Generation of Standard Solar Models,” *Astrophys. J.*, vol. 835, no. 2, p. 202, 2017. [Cited on pages 18, 19, 20, 21, 23, and 125.]
- [68] N. Grevesse and A. J. Sauval, “Standard Solar Composition,” *Space Sci. Rev.*, vol. 85, pp. 161–174, 1998. [Cited on pages 18 and 19.]
- [69] J. Christensen-Dalsgaard, “Helioseismology,” *Rev. Mod. Phys.*, vol. 74, pp. 1073–1129, 2003. [Cited on page 19.]
- [70] S. Basu and H. Antia, “Helioseismology and Solar Abundances,” *Phys. Rept.*, vol. 457, pp. 217–283, 2008. [Cited on page 19.]
- [71] A. M. Serenelli, W. Haxton, and C. Pena-Garay, “Solar models with accretion. I. Application to the solar abundance problem,” *Astrophys. J.*, vol. 743, p. 24, 2011. [Cited on page 20.]
- [72] M. Agostini *et al.*, “Comprehensive measurement of *pp*-chain solar neutrinos,” *Nature*, vol. 562, no. 7728, pp. 505–510, 2018. [Cited on pages 20, 28, 29, 63, 105, 118, 120, 121, 123, 125, 128, 129, and 130.]
- [73] J. N. Bahcall, A. M. Serenelli, and S. Basu, “New solar opacities, abundances, helioseismology, and neutrino fluxes,” *Astrophys. J. Lett.*, vol. 621, pp. L85–L88, 2005. [Cited on page 21.]
- [74] J. Bahcall, W. A. Fowler, J. Iben, I., and R. Sears, “Solar neutrino flux,” *Astrophys. J.*, vol. 137, pp. 344–346, 1963. [Cited on page 21.]
- [75] E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle, “Synthesis of the elements in stars,” *Rev. Mod. Phys.*, vol. 29, pp. 547–650, Oct 1957. [Cited on page 22.]
- [76] W. Fowler, “Experimental and theoretical nuclear astrophysics; the quest for the origin of the elements: Nobel prize lecture,” *Rev. Mod. Phys.*, vol. 56, p. 149, 1984. [Cited on page 22.]

- [77] V. Castellani, S. Degl’Innocenti, G. Fiorentini, M. Lissia, and B. Ricci, “Solar neutrinos: Beyond standard solar models,” *Phys. Rept.*, vol. 281, pp. 309–398, 1997. [Cited on pages [22](#) and [128](#).]
- [78] J. N. Bahcall, “The Luminosity constraint on solar neutrino fluxes,” *Phys. Rev.*, vol. C65, p. 025801, 2002. [Cited on pages [22](#) and [128](#).]
- [79] J. Bergstrom, M. Gonzalez-Garcia, M. Maltoni, C. Pena-Garay, A. M. Serenelli, and N. Song, “Updated determination of the solar neutrino fluxes from solar neutrino data,” *JHEP*, vol. 03, p. 132, 2016. [Cited on pages [22](#), [23](#), and [128](#).]
- [80] J. Bahcall and R. May, “The rate of the proton-proton reaction and some related reactions,” *Astrophysical Journal*, vol. 155, p. 501, 1969. [Cited on pages [22](#) and [128](#).]
- [81] J. N. Bahcall, “LINE VERSUS CONTINUUM SOLAR NEUTRINOS,” *Phys. Rev. D*, vol. 41, p. 2964, 1990. [Cited on pages [22](#) and [128](#).]
- [82] J. N. Bahcall and C. Pena-Garay, “Global analyses as a road map to solar neutrino fluxes and oscillation parameters,” *JHEP*, vol. 11, p. 004, 2003. [Cited on pages [24](#) and [25](#).]
- [83] S. Antusch, J. P. Baumann, and E. Fernandez-Martinez, “Non-Standard Neutrino Interactions with Matter from Physics Beyond the Standard Model,” *Nucl. Phys. B*, vol. 810, pp. 369–388, 2009. [Cited on pages [25](#) and [121](#).]
- [84] C. Arpesella *et al.*, “First real time detection of  ${}^7\text{Be}$  solar neutrinos by Borexino,” *Phys. Lett.*, vol. B658, pp. 101–108, 2008. [Cited on page [28](#).]
- [85] G. Bellini *et al.*, “Precision measurement of the  ${}^7\text{Be}$  solar neutrino interaction rate in Borexino,” *Phys. Rev. Lett.*, vol. 107, p. 141302, 2011. [Cited on page [28](#).]
- [86] G. Bellini *et al.*, “Absence of day–night asymmetry of 862 keV  ${}^7\text{Be}$  solar neutrino rate in Borexino and MSW oscillation parameters,” *Phys. Lett.*, vol. B707, pp. 22–26, 2012. [Cited on page [28](#).]
- [87] G. Bellini *et al.*, “First evidence of *pep* solar neutrinos by direct detection in Borexino,” *Phys. Rev. Lett.*, vol. 108, p. 051302, 2012. [Cited on pages [28](#) and [43](#).]
- [88] G. Bellini *et al.*, “Measurement of the solar  ${}^8\text{B}$  neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector,” *Phys. Rev.*, vol. D82, p. 033006, 2010. [Cited on page [28](#).]
- [89] G. Bellini *et al.*, “Neutrinos from the primary proton–proton fusion process in the Sun,” *Nature*, vol. 512, no. 7515, pp. 383–386, 2014. [Cited on pages [28](#) and [63](#).]
- [90] S. Kumaran, “Updated Geoneutrino Measurement with the Borexino Detector,” Master’s thesis, RWTH Aachen, 2018. [Cited on page [28](#).]
- [91] M. Agostini *et al.*, “Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data,” *Phys. Rev.*, vol. D96, no. 9, p. 091103, 2017. [Cited on page [28](#).]

- [92] M. Agostini *et al.*, “Modulations of the Cosmic Muon Signal in Ten Years of Borexino Data,” *JCAP*, vol. 1902, p. 046, 2019. [Cited on pages 28 and 61.]
- [93] S. K. Agarwalla *et al.*, “Constraints on flavor-diagonal non-standard neutrino interactions from Borexino Phase-II,” *JHEP*, vol. 02, p. 038, 2020. [Cited on pages 28 and 121.]
- [94] G. Alimonti *et al.*, “The Borexino detector at the Laboratori Nazionali del Gran Sasso,” *Nucl. Instrum. Meth.*, vol. A600, pp. 568–593, 2009. [Cited on pages 29, 33, and 36.]
- [95] G. Bellini *et al.*, “Final results of Borexino Phase-I on low energy solar neutrino spectroscopy,” *Phys. Rev.*, vol. D89, no. 11, p. 112007, 2014. [Cited on pages 29, 34, 43, 50, 52, 55, 58, 61, and 62.]
- [96] G. Consolati, D. Franco, S. Hans, C. Jollet, A. Mereaglia, S. Perasso, A. Tonazzo, and M. Yeh, “Characterization of positronium properties in doped liquid scintillators,” *Phys. Rev.*, vol. C88, p. 065502, 2013. [Cited on page 31.]
- [97] G. Bellini *et al.*, “Muon and Cosmogenic Neutron Detection in Borexino,” *JINST*, vol. 6, p. P05005, 2011. [Cited on pages 31, 59, and 60.]
- [98] “The Borexino Experiment.” <http://borex.lngs.infn.it>. [Cited on page 32.]
- [99] H. Back *et al.*, “Borexino calibrations: Hardware, Methods, and Results,” *JINST*, vol. 7, p. P10018, 2012. [Cited on page 33.]
- [100] B. Caccianiga and A. C. Re, “The calibration system for the Borexino experiment,” *Int. J. Mod. Phys.*, vol. A29, p. 1442010, 2014. [Cited on page 33.]
- [101] D. Basilico, *First indication of solar neutrinos from the CNO cycle reactions with the Borexino experiment*. PhD thesis, University of Milan, 2020-02-07. [Cited on pages 34, 82, 83, 99, 100, 101, 126, and 127.]
- [102] D. D’Angelo, *Towards the detection of low energy solar neutrinos in BOREXino: data readout, data reconstruction and background identification*. PhD thesis, Munich, Tech. U., 2006. [Cited on pages 34 and 37.]
- [103] C. Arpesella *et al.*, “Measurements of extremely low radioactivity levels in BOREXINO,” *Astropart. Phys.*, vol. 18, pp. 1–25, 2002. [Cited on pages 33, 52, and 56.]
- [104] F. Gatti, V. Lagomarsino, P. Musico, M. Pallavicini, A. Razeto, G. Testera, and S. Vitale, “The Borexino read out electronics and trigger system,” *Nucl. Instrum. Meth.*, vol. A461, pp. 474–477, 2001. [Cited on page 35.]
- [105] G. Alimonti *et al.*, “Science and technology of BOREXINO: A Real time detector for low-energy solar neutrinos,” *Astropart. Phys.*, vol. 16, pp. 205–234, 2002. [Cited on page 35.]
- [106] C. Galbiati and K. McCarty, “Time and space reconstruction in optical, non-imaging, scintillator-based particle detectors,” *Nucl. Instrum. Meth.*, vol. A568, pp. 700–709, 2006. [Cited on page 41.]

- [107] A. Minotti, *Exploitation of pulse shape analysis for correlated background rejection and ortho-positronium identification in the Double Chooz experiment*. PhD thesis, U. Strasbourg, 2016-11-17. [Cited on page 43.]
- [108] E. G. et al., “Particle identification by pulse shape discrimination,” *Energia Nucleare*, vol. 17, p. 34, 1970. [Cited on page 44.]
- [109] G. Bellini *et al.*, “Final results of Borexino Phase-I on low energy solar neutrino spectroscopy,” *Phys. Rev.*, vol. D89, no. 11, p. 112007, 2014. [Cited on pages 45 and 119.]
- [110] S.-I. Kroon, “Towards CNO measurement with Borexino:  $^{210}\text{Bi}$  homogeneity in the liquid scintillator,” Master’s thesis, RWTH Aachen, 2019. [Cited on page 45.]
- [111] M. Agostini *et al.*, “The Monte Carlo simulation of the Borexino detector,” *Astropart. Phys.*, vol. 97, pp. 136–159, 2018. [Cited on pages 46 and 47.]
- [112] N. Rossi, M. Agostini, P. Cavalcante, X. Ding, D. Guffanti, A. Ianni, M. Misiaszek, and F. Calaprice, “Feasibility study of the link between  $^{210}\text{Bi}$  and  $^{210}\text{Po}$  in the Borexino Phase-II.” Internal document, Nov 2017. [Cited on pages 55 and 126.]
- [113] X. Ding, *High Precision Solar Neutrino Spectroscopy with Borexino and JUNO*. PhD thesis, SISSA, Trieste, 2019. [Cited on pages 55, 78, 126, 127, and 131.]
- [114] L. Cadonati, F. Calaprice, C. Galbiati, A. Pocar, and T. Shutt, “The nylon scintillator containment vessels for the Borexino solar neutrino experiment,” *Int. J. Mod. Phys.*, vol. A29, p. 1442004, 2014. [Cited on page 56.]
- [115] M. Agostini *et al.*, “Simultaneous precision spectroscopy of  $pp$ ,  $^7\text{Be}$ , and  $pep$  solar neutrinos with borexino phase-ii,” *Phys. Rev. D*, vol. 100, p. 082004, Oct 2019. [Cited on pages 58, 106, 109, 113, 118, 128, 129, 130, and 183.]
- [116] G. Bellini *et al.*, “Cosmogenic Backgrounds in Borexino at 3800 m water-equivalent depth,” *JCAP*, vol. 1308, p. 049, 2013. [Cited on pages 59 and 61.]
- [117] X. Ding, “Proposal of a new effective Monte Carlo Quantum Efficiency.” Internal document, Oct 2018. [Cited on page 78.]
- [118] C. Arpesella *et al.*, “Direct measurement of the  $^7\text{Be}$  solar neutrino flux with 192 days of borexino data,” *Phys. Rev. Lett.*, vol. 101, p. 091302, Aug 2008. [Cited on page 105.]
- [119] M. Agostini *et al.*, “Improved measurement of  $^8\text{B}$  solar neutrinos with 1.5 kt·y of Borexino exposure,” *Phys. Rev.*, vol. D101, no. 6, p. 062001, 2020. [Cited on pages 115, 116, 117, and 118.]
- [120] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, and T. Schwetz, “Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity,” *JHEP*, vol. 01, p. 087, 2017. [Cited on pages 119 and 120.]
- [121] X. Ding, F. Calaprice, A. Ianni, and N. Rossi, “Search for CNO solar neutrinos in Borexino.” Internal document, Apr 2019. [Cited on pages 126 and 127.]



- [122] D. Basilico, X. Ding, A. Göttel, and S. Kumaran, “Extraction of  $^{210}\text{Bi}$  rate via  $^{210}\text{Po}$ .” Internal document, May 2020. [Cited on pages 126 and 127.]
- [123] A. Göttel and S. Kumaran, “A Bayesian framework for the  $^{210}\text{Po}$  clean region analysis in the search for CNO neutrinos.” Internal document, May 2020. [Cited on page 127.]
- [124] A. Kurylov, M. Ramsey-Musolf, and P. Vogel, “Radiative corrections to low-energy neutrino reactions,” *Phys. Rev. C*, vol. 67, p. 035502, 2003. [Cited on page 128.]
- [125] M. Agostini *et al.*, “Sensitivity to neutrinos from the solar cno cycle in borexino,” 2020. [Cited on pages 128, 129, 130, and 132.]
- [126] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J. C*, vol. 71, p. 1554, 2011. [Erratum: *Eur.Phys.J.C* 73, 2501 (2013)]. [Cited on pages 130 and 139.]
- [127] X. Ding, “GooStats: A GPU-based framework for multi-variate analysis in particle physics,” *JINST*, vol. 13, no. 12, p. P12018, 2018. [Cited on page 133.]
- [128] D. Basilico, F. Calaprice, X. Ding, and A. Singhal, “Systematics for CNO discovery (Hypothesis test + uncertainty).” Internal document, May 2020. [Cited on pages 138 and 139.]
- [129] F. Calaprice, X. Ding, A. Ianni, and N. Rossi, “Systematics to CNO neutrino evidence from the  $^{210}\text{Bi}$  shape.” Internal document, May 2020. [Cited on page 138.]
- [130] H. Behrens and L. Szybisz, “On the  $\beta$ -decay  $^{210}\text{Bi}(\text{RaE}) \rightarrow ^{210}\text{Po}(\text{RaF})$ ,” *Nuclear Physics A*, vol. 223, no. 2, pp. 268 – 284, 1974. [Not cited.]
- [131] A. Carles and A. Malonda, “Precision measurement of the rae shape factor,” *Nuclear Physics A*, vol. 596, no. 1, pp. 83 – 90, 1996. [Not cited.]
- [132] H. DANIEL, “Shapes of beta-ray spectra,” *Rev. Mod. Phys.*, vol. 40, pp. 659–672, Jul 1968. [Not cited.]
- [133] F. D., W. Wiesner, R. Löhken, and H. Rebel, “ $\beta$ -Spektroskopie mit Halbleiterdetektoren beim Zerfall von  $^{32}\text{P}$ ,  $^{49}\text{Sc}$ ,  $^{204}\text{Tl}$  und  $^{210}\text{Bi}$ ,” *Zeitschrift für Physik A Hadrons and nuclei volume*, vol. 225, pp. 164–194, 1969. [Cited on page 138.]
- [134] L. Ludhova, “CNO detection summary.” Internal document, May 2020. [Cited on page 139.]





# Eidesstattliche Erklärung

Ich, Mariia Redchuk, erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt.

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt.
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt.
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt.
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate.
5. Alle wesentlichen Quellen von Unterstützung wurden benannt.
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde.
7. Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:
  - M. Agostini et al. (Borexino Collaboration), “Sensitivity to neutrinos from the solar CNO cycle in Borexino”, arXiv:2005.12829;
  - M. Agostini et al. (Borexino Collaboration), “Improved measurement of  $^8\text{B}$  solar neutrinos with 1.5 kt·y of Borexino exposure”, *Physical Review D* 101 (2020) 062001;
  - S.K. Agarwalla et al. (Borexino Collaboration), “Constraints on flavor-diagonal non-standard neutrino interactions from Borexino Phase-II”, *Journal of High Energy Physics* 02 (2020) 038;
  - M. Agostini et al. (Borexino Collaboration), “Comprehensive geoneutrino analysis with Borexino”, *Physical Review D* 101 (2020) 012009;
  - M. Agostini et al. (Borexino Collaboration), “Simultaneous precision spectroscopy of  $pp$ ,  $^7\text{Be}$  and  $pep$  solar neutrinos with Borexino Phase-II”, *Physical Review D* 100 (2019) 082004;

- M. Agostini et al. (Borexino Collaboration), “Modulations of the cosmic muon signal in ten years of Borexino data”, *Journal of Cosmology and Astroparticle Physics* 1902 (2019) 046;
- M. Agostini et al. (Borexino Collaboration), “Comprehensive measurement of  $pp$ -chain solar neutrinos”, *Nature* 562 (2018) 505;
- M.Redchuk (on behalf of the Borexino Collaboration), “Solar neutrino analysis with the Borexino detector”, *Journal of Physics: Conference Series* 1056 (2018) 012050;
- M. Agostini et al. (Borexino Collaboration), “Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data”, *Physical Review D* 96 (2017) 091103.

---

Datum

---

Unterschrift