



GRAN SASSO SCIENCE INSTITUTE

Search for double beta decay of ¹³⁰Te to the 0⁺ states of ¹³⁰Xe with the CUORE experiment

PhD thesis

submitted May 15, 2020 revised July 5, 2020

Advisors

Carlo Bucci INFN, Laboratori Nazionali del Gran Sasso

Claudia Tomei INFN, Sezione di Roma

PhD Candidate

Guido Fantini Gran Sasso Science Institute

Gran Sasso Science Institute

XXXII Cycle – A. Y. 2016-2019

Search for double beta decay of 130 Te to the 0⁺ states of 130 Xe with the CUORE experiment

G. Fantini

Abstract

The CUORE experiment is the largest bolometric array ever built with the main goal of searching for the lepton number violating neutrinoless double beta decay of ¹³⁰Te. The closely packed arrangement of the CUORE crystals allows us to look for coincident signals between detectors. The latest results on neutrino-less double beta decay search of ¹³⁰Te to the ground 0_1^+ state of ¹³⁰Xe and first 0_2^+ excited state will be presented and discussed. The presence of de-excitation gammas in the final state helps reducing the background by studying coincident events in two or more bolometers. The Standard Model $2\nu\beta\beta$ decay channel of the same isotope on the 0_2^+ excited state of ¹³⁰Xe will be investigated as well. No evidence for any of the mentioned decay modes was observed and a Bayesian upper bound at 90 % C.I. on the decay rate will be presented and discussed.

Contents

Introduction

1	Dou	ıble-be	eta decay	5
	1.1	Neutri	no physics	5
		1.1.1	Neutrinos in the Standard Model	7
		1.1.2	Majorana neutrinos	8
		1.1.3	Neutrino mixing	9
	1.2	Theor	etical overview of double-beta decay	11
		1.2.1	Nuclear Matrix Elements	13
		1.2.2	General remarks	15
		1.2.3	Baryon asymmetry in the universe (BAU)	15
	1.3	Exper	imental considerations	17
		1.3.1	Sensitivity	17
		1.3.2	Detector design	19
		1.3.3	Competing experiments	21
2	CU	ORE		25
2	CU 2.1	ORE Histor	y and predecessors	25 25
2	CU 2.1	ORE Histor 2.1.1	y and predecessors	25 25 26
2	CU 2.1	ORE Histor 2.1.1 2.1.2	y and predecessors	25 25 26 27
2	CU 2.1 2.2	ORE Histor 2.1.1 2.1.2 The b	y and predecessors	 25 25 26 27 27
2	CU(2.1)2.2	ORE Histor 2.1.1 2.1.2 The b 2.2.1	y and predecessors	 25 25 26 27 27 28
2	CU(2.1 2.2	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2	y and predecessors	 25 26 27 27 28 28
2	CU(2.1 2.2	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2 2.2.3	y and predecessors	 25 25 26 27 27 28 28 30
2	 CU(2.1) 2.2 2.3 	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2 2.2.3 The C	y and predecessors	 25 26 27 27 28 28 30 31
2	CU(2.1)2.22.3	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2 2.2.3 The C 2.3.1	y and predecessors	 25 25 26 27 27 28 28 30 31 32
2	CU(2.1)2.22.3	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2 2.2.3 The C 2.3.1 2.3.2	y and predecessors	 25 25 26 27 28 28 30 31 32 35
2	CU(2.1)2.22.3	ORE Histor 2.1.1 2.1.2 The b 2.2.1 2.2.2 2.2.3 The C 2.3.1 2.3.2 2.3.3	y and predecessors	 25 26 27 28 28 30 31 32 35 37

1

3.1Data acquisition43.1.1Trigger algorithms43.2Data quality43.3Data processing43.3.1Preprocessing43.3.2Amplitude evaluation43.3.3Thermal gain stabilization43.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.1.1Trigger algorithms43.2Data quality43.3Data processing43.3.1Preprocessing43.3.2Amplitude evaluation43.3.3Thermal gain stabilization43.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.2 Data quality 4 3.3 Data processing 4 3.3.1 Preprocessing 4 3.3.2 Amplitude evaluation 4 3.3.3 Thermal gain stabilization 4 3.3.4 Energy calibration 5 3.3.5 Coincidences 5 3.3.6 Pulse shape analysis 5
3.3 Data processing
3.3.1Preprocessing43.3.2Amplitude evaluation43.3.3Thermal gain stabilization43.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.3.2Amplitude evaluation43.3.3Thermal gain stabilization43.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.3.3Thermal gain stabilization43.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.3.4Energy calibration53.3.5Coincidences53.3.6Pulse shape analysis5
3.3.5Coincidences53.3.6Pulse shape analysis5
3.3.6 Pulse shape analysis
3.3.7 Blinding \ldots 5
4 Detector characterization 5
4 Detector characterization 4
4.1 Energy threshold
4.1.1 Higger threshold
4.1.2 Confidences threshold
4.1.5 Tuise Shape Analysis threshold
4.2 Enciency
$4.2.1$ Containment enciency $\dots \dots \dots$
4.2.2 Detection enciency
$4.2.9$ Accidential contribution endency $\dots \dots \dots$
4.3 Energy scale reconstruction
4.3 Line shape response function 66
4.3.2 Energy bias and resolution evaluation
5 Monte Carlo techniques and background control 6
5.1 Material selection and background control \ldots \ldots \ldots \ldots \ldots 6
5.2 Monte Carlo simulations
5.2.1 Particle generation and propagation
5.2.2 Detector response simulation $\ldots \ldots \ldots \ldots \ldots \ldots \ldots $
5.3 Background model
6 Neutrino-less Double Beta Decay search to the ground state of ¹³⁰ Xe 7
6.1 Event selection
6.2 Model of the reconstructed spectrum
6.3 Bayesian fit
6.4 Results
6.4.1 Exclusion sensitivity and blinded fit
6.4.2 Lower limit and systematics on $T_{1/2}^{0\nu}$

7	Doι	ible-Beta Decay search to the first 0^+ excited state of 130 Xe	93	
	7.1	Motivation and previous results	. 93	
	7.2	Physical process and experimental signature	. 94	
	7.3	Event selection and efficiency contributions	. 101	
	7.4	Model	. 103	
	7.5	Blinding and sensitivity computation	. 104	
	7.6	Systematic uncertainties and final results	. 110	
8	Cor	nclusion	121	
\mathbf{A}	Unł	oinned Extended Maximum Likelihood derivation	125	
в	Bayesian inference with Markov Chain Monte Carlo			
	B.1	Bayesian inference	. 127	
	B.2	Markov Chain Monte Carlo	. 128	
С	Bayesian Result Combination			
	C.1	Random variable transformations	. 133	
	C.2	Combining the results from excited and ground state modes $\ . \ . \ . \ .$. 134	
\mathbf{Li}	st of	Abbreviations	137	
Li	st of	Figures and Tables	139	
Ac	knov	wledgements	143	
Bi	bliog	graphy	147	

Introduction

The observation of neutrino mixing, and the consequent discovery that neutrinos are massive, has boosted the interest and importance of searching for neutrinoless double beta decay $(0\nu\beta\beta)$. This rare process provides a direct probe of Beyond Standard Model physics since it violates lepton number conservation by 2 units at the tree level. In addition it allows the investigation of the Majorana nature of neutrinos and provides insight on the neutrino absolute mass scale and ordering.

This work was carried out within the CUORE Collaboration. The Cryogenic Underground Observatory for Rare Events (CUORE) is an array of 988 TeO₂ closely packed bolometer crystals, operated underground in the Hall A of the INFN-LNGS¹ facility in Italy. Its main scientific goal is to search for neutrinoless double beta decay in ¹³⁰Te, but is sensitive to other processes as well. Among those, the double beta decay of ¹³⁰Te nuclei to the excited states of ¹³⁰Xe is investigated in this thesis. The development of the data analysis for both decay channels was a central part of the Author's work, as further detailed in the following.

In Ch. 1 a short, non-exhaustive review of some milestones of neutrino physics and double beta decay is given. The scope of the introductory Sec. 1.1 and 1.2 is far from being a comprehensive review of the subject. It is meant just as a logically consequent subset of the main theoretical and experimental achievements relevant - in the Author's opinion - to put the following results in context, and to provide a minimal set of references to further explore such topics. In Sec. 1.3 and 1.3.1 an introduction, again not meant to be complete but rather an overview of a selected subset of the relevant topics, is given from the experimental point of view. The main parameters and challenges related to the sensitivity to double beta decay are shown and discussed, with special focus to the ones that will have an impact or drive choices in the analyses outlined in the following chapters. A selection of experimental results from other experiments are included, together with a small description of the corresponding setup.

In Ch. 2 the basics to understand the CUORE experiment structure and operation

¹The Istituto Nazionale di Fisica Nucleare (INFN) is the italian national institute for nuclear physics research, the Laboratori Nazionali del Gran Sasso (LNGS) one of its laboratories, primarily devoted to searches for rare processes in a low radioactivity environment.

from the hardware point of view are mentioned. A short historical section 2.1 lists some of the predecessors and demonstrator experiments that were built prior to CUORE and allowed the successful operation of such a large and complex bolometer array. In Sec. 2.2 the principles of operation of bolometric detectors are explained, in Sec. 2.3 the techniques needed to cool down such a large detector mass to temperatures of $\sim 10 \text{ mK}$ are briefly reviewed. Finally in Sec. 2.4 the multiple CUORE calibration systems are sketched, together with some results of early operation of such devices. References are given throughout the chapter to allow the interested Reader a deeper investigation on any of the mentioned topics.

In Ch. 3 an overview of the hardware and software components related to data acquisition and processing is given. The steps that allow to readout the signal waveforms, store them on disk and trigger events are explained in Sec. 3.1. The data quality checks performed to ensure a clean sample for physics analyses are outlined in Sec. 3.2, together with the processing steps that allow to build the final spectra of energy release in the CUORE bolometers (Sec. 3.3). Some definitions and terminology needed to understand the data structure are included. A significant fraction of the available time during the Author's PhD was devoted to development, implementation, test and processing of data, especially with regard to the coincidence algorithms detailed in Sec. 3.3.5.

The characterization of the detector performances and behavior is a crucial prerequisite before a measurement of any physical observable is performed. In Ch. 4 an overview of the most important parameters that allow such characterization is given. In particular the definition and computation methods for the energy threshold (Sec. 4.1), the efficiency 4.2, the energy scale uncertainty and resolution (Sec. 4.3) are investigated. Some part of the Author's PhD work was devoted to the computation of the detector response function for each bolometer-dataset pair. Some time was devoted to the modeling of the energy scale reconstruction bias and resolution. Specifically, both the development of the needed software tools and their integration with previous (data processing) and subsequent (Bayesian fit) pieces of software of the data analysis chain was performed during the time of the PhD.

In Ch. 5, after an introduction (Sec. 5.1) about material selection and background control, the Monte Carlo techniques used in CUORE will be explored. A technical description of the software tools deployed to simulate and propagate particles originating from all known background sources are detailed in Sec. 5.2. The inference model and the results of the best fit to the CUORE data are included in Sec. 5.3, together with the latest determination of the half life of ¹³⁰Te, due to double beta decay to the ground state of ¹³⁰Xe in the $2\nu\beta\beta$ mode.

Chapters 6 and 7 contain the core of the work included in this thesis and the original contributions of the Author. In Ch. 6 the analysis techniques used to place a Bayesian lower bound to the neutrinoless double beta decay half life of 130 Te to the 0^+_1 ground state of 130 Xe are detailed. In Sec. 6.1 the criteria, nomenclature, quantitative results of the selection of candidate events are explained. Sec. 6.2 describes the model used to simultaneously describe the distribution of candidate events and infer a posterior probability distribution

for the rate of $0\nu\beta\beta$ decay events. A Bayesian fit implemented with a Markov Chain Monte Carlo method was used to eventually compute such posterior distribution, and it is detailed in Sec. 6.3. The most technical parts are detailed in the Appendix B. The results are shown and discussed in Sec. 6.4. The work described in this chapter contains significant contributions of the Author. It was carried out in a working group of 3 people, which contributed equally to all the detailed items, from the development of the Bayesian fit model, the implementation of the needed software tools, the definition, modelling and implementation of systematic uncertainties, to the Toy Monte Carlo data production and fit for the extraction of the expected limit setting sensitivity. In addition many other members of the CUORE Collaboration, including the members of the above mentioned working group, contributed to the data taking in terms of detector operation, data processing, data quality evaluation.

In Ch. 7 the analysis techniques and results to place the most stringent available lower limit on the double beta decay half life of 130 Te to the 0_2^+ excited state of 130 Xe are described. An overview of the available theoretical predictions and experimental results on this process is given in Sec. 7.1. A description of the relevant nuclear structure of the daughter 130 Xe nucleus together with the experimental signature definition constitutes Sec. 7.2. The selection of candidate events together with the corresponding efficiency is detailed in Sec. 7.3. The model used to perform inference on the investigated process is described in detail in Sec. 7.4. A blinding technique developed specifically for this analysis is described in Sec. 7.5, together with the techniques and results used to compute the limit setting sensitivity. Finally Sec. 7.6 includes a discussion on the systematic uncertainties, their model and effect on the final result, which is shown along with the spectra of unblinded data.

Finally in Ch. 8 a short conclusion and summary of the results is given. Future perspectives are sketched, pointing to the relevant items that need improvement in order to fully profit from the CUORE potential in the excited states double beta decay mode. In this last chapter the Author's contribution is explained in more detail, using the information and nomenclature the Reader will have extracted from the thesis.

l Chapter

Double-beta decay

Double-beta decay is a rare nuclear process where a nucleus transforms in its isobar emitting radiation,

$$(A, Z) \to (A, Z+2) + X.$$
 (1.1)

The simultaneous transmutation of a pair of neutrons into the same number of protons is predicted by the Standard Model (SM) and was detected in several nuclei together with the emission of two electrons and two anti-neutrinos. The neutrino-less mode of this decay is a posited Beyond Standard Model process that could shed light on many open aspects of modern particle physics and cosmology such as the neutrino mass ordering and scale, the existence of lepton flavor violation and elementary Majorana fermions, the baryon asymmetry in the universe. In the following a short review of the current knowledge and latest achievements in neutrino physics will be given, neutrino-less double beta decay will be outlined together with a summary of recent experimental results.

1.1 Neutrino physics

Less than a century ago nuclei with mass number A and atomic number Z were thought to be made out of A protons and A - Z electrons. Among the many inconsistencies with experimental data it is worth remembering the continuous spectrum of beta decay and the statistics of ⁶Li and ¹⁴N. Regardless of the fact that in beta decays the electron was not thought to be created but just emitted from the nucleus, energy conservation would anyway require a monochromatic final state electron. Considering the statistics, a nucleus made of an odd number of spin 1/2 fermions would need to obey the Fermi statistics, instead Raman spectroscopy measurements of rotational spectra of molecular lithium and nitrogen would contradict this prediction. In 1930 the *neutron* was introduced (this is how Pauli would call a neutrino) in a letter [1] sent by Pauli himself to the *radioactive ladies and gentlemen* attending a meeting in Tübingen, explaining he had come up with a *desperate remedy* to save both energy conservation and the exchange theorem of statistics in the interpretation of such experimental evidence. The remedy was the introduction of a neutral massive particle of spin 1/2 in the nucleus that would be emitted in beta decays together with the electron carrying away some energy and justifying at once both the inconsistencies. Such particles would need to have a mass of the order of the electron mass and, in any case, not larger than 0.01 proton masses¹. Soon later, after the discovery of the neutron by Chadwick in 1932, Fermi published [2] [3] [4] his theory of beta decay where he postulated a point interaction between the nucleon field, whose two states describe the proton and the neutron, and the electron and neutrino fields. Since the theory of nuclear electrons would require already from the uncertainty principle a momentum of the electron of the order of 100 MeV/c, way more than the available energy in beta decays, and the neutron and the posited particle from Pauli could not be the same particle due to the very different mass, it was interesting to unify all these theoretical developments in agreement with experimental results just allowing the possibility of particles to be created at the very moment of the decay rather than being pre-existent in the atomic nucleus.

The first direct experimental detection of neutrinos is due to Cowan and Reines in the Savannah-River nuclear reactor in 1952 [5]. There the inverse beta-decay $\bar{\nu}_e + p \rightarrow e^+ + n$ was detected in two tanks of 200 L of CdCl₂ water solution each, through the delayed (~ 10 μs) coincidence of the scintillation light due to the annihilation of the final state positron and the ~ 9 MeV γ emission from the Cd neutron capture. Both the annihilation and the de-excitation γ rays were detected by two out of the three scintillator detectors that surrounded the water tanks in a "club sandwich" fashion [6]. Reines was awarded the Nobel Prize for this discovery in 1995. Another milestone was reached in Brookhaven with the discovery of the muon neutrino due to Lederman, Schwartz, Steinberger and coworkers (1962 [7], Nobel Prize 1988). The picture was completed with the discovery of the tau neutrino in 2000 [8] and the observation of ν_{τ} appearence by the OPERA experiment from a muon neutrino beam produced at CERN [9].

The neutrino deficit first observed by the Homestake experiment (see [10] for a review) opened a long standing puzzle in neutrino physics known as the solar neutrino problem. It was eventually solved around the turn of the millennium, with the discovery of neutrino oscillations, which shows that neutrinos have mass. For this reason a Nobel Prize was awarded in 2015 to T. Kajita and A. B. McDonald. Kajita, together with the Super-Kamiokande collaboration, showed evidence of a ν_{μ} deficit in atmospheric neutrinos compatible with a two-flavor $\nu_{\mu} \leftrightarrow \nu_{\tau}$ oscillation model [11]. Meanwhile McDonald had been leading the SNO Collaboration, which provided a solution to the solar neutrino problem. The SNO experiment was sensitive not only to charged and neutral current neutrino interactions on deuterium, but also to elastic scattering. The different sensitivity of such detection channels as a function of the incoming neutrino flavor provided evidence for a non-active ν_e component in the ⁸B neutrino flux from the Sun [12] [13]. The SNO result showed a very good agreement with the standard solar model [14] [15]. A deep understanding of such experimental results requires some basics of neutrino physics such as neutrino mixing (see

¹Quoting from Pauli's letter "Die Masse der Neutronen müsste von derselben Grössenordnung wie die Elektronenmasse sein und jedenfalls nich grösser als 0.01 Protonenmasse."

Sec. 1.1.3), the propagation of neutrinos in vacuum and through matter [16] [17], see e.g. [18] and references therein for an introduction.

It is indeed the non vanishing neutrino mass, demonstrated by flavor oscillations and observed in many sources (the Sun, cosmic ray interaction with the atmosphere, beta decays in nuclear reactors, accelerators), that emphasizes the question whether neutrinos are Majorana fermions (see Sec. 1.1.2). The key point of the possibility for neutrinos to be Majorana particles resides in the definition of their only non-vanishing internal quantum number: lepton number. By definition an (electron) neutrino is the particle emitted in a β^+ decay and carries a lepton number L = +1 while an (electron) anti-neutrino is the one emitted in a β^- decay and carries a lepton number L = -1. Apart from this definition, there is no experimental evidence nor fundamental symmetry that forces neutrinos and anti-neutrinos to be different particles. They could as well be just two helicity states of the same massive Majorana fermion.

The last statement is of capital importance, because many decades of experimental efforts, including this work, stem from there. It is worth spending some more time deepening its foundations and implications. For a review of the main stages of the neutrino history starting from the Pauli proposal and finishing with the discovery and study of neutrino oscillations see [19] and references therein.

1.1.1 Neutrinos in the Standard Model

The Standard Model (SM) of electroweak interactions does not include massive neutrinos. The leptonic building blocks (quantized fields) of the theory are an isospin doublet for each flavor f and a charged lepton singlet,

$$L_{f,L} \doteq \begin{pmatrix} \nu_f \\ l_f \end{pmatrix}_L, \quad l_{f,R} \quad \text{where} \quad f = e, \mu, \tau$$
 (1.2)

and the L(R) subscript corresponds to the chiral left (right) projector $P_L = (1 \mp \gamma^5)/2$. The electroweak lagrangian is required to be gauge invariant under $SU(2)_L$ transformations. Since the left-handed components of the fermion fields transform differently from the right-handed ones, the presence of mass terms of the form $\bar{f}f = \bar{f}_L f_R + \bar{f}_R f_L$ in the lagrangian density is forbidden. In fact the generation of fermion masses in the SM happens through spontaneous symmetry breaking (SSB) with the Higgs mechanism. The Higgs-lepton Yukawa term in the SM lagrangian,

$$\mathcal{L}_{H-L} = -\sum_{f,f'} Y_{ff'}^{l} \bar{L}_{f,L} \Phi l_{f',R} + \text{h.c.}$$
(1.3)

where Φ is the Higgs doublet and Y^l is a complex 3×3 diagonal matrix of Yukawa couplings generates mass terms for the charged leptons of the form

$$m_f = \frac{Y_{f,f}^l v}{\sqrt{2}} \tag{1.4}$$

The experimental evidence of neutrino flavor oscillations has demonstrated that neutrinos instead do have mass. In order to embed this feature in the (minimally extended) SM one possibility is to add a right-handed component to the neutrino field in order to allow for a Dirac mass term in the lagrangian to be generated with SSB. The most relevant feature of such a component, rather than its right-handedness, is that the $\nu_{f,R}$ field would be invariant under all the SM symmetries being a singlet of $SU(3)_C \times SU(2)_L$ with hyper $charge^2 Y = 0$. For this reason the right-handed neutrino fields are ofter referred to as sterile, because they lack any gauge SM interaction. Proceeding with the addition of a new matrix of Yukawa couplings to the Higgs boson the corresponding neutrino masses would be, as for all the other leptons, proportional to the Higgs vacuum expectation value v. Since neutrino masses are known to be much smaller than the ones of the charged leptons and quarks, this theory lacks any explanation why the eigenvalues of the Yukawa coupling matrix for neutrinos are so small. In addition, even if the right-handed components of the massive Dirac neutrino fields are sterile, they can still mix with the ordinary degrees of freedom through the mass term. Since at present there is no indication of such additional mixing, other possibilities must be explored.

1.1.2 Majorana neutrinos

The Dirac equation for a massive free fermion field $\psi = \psi_L + \psi_R$ reads

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0 \tag{1.5}$$

and is equivalent to the system of equations

$$\begin{cases} i\gamma^{\mu}\partial_{\mu}\psi_{L} = m\psi_{R} \\ i\gamma^{\mu}\partial_{\mu}\psi_{R} = m\psi_{L} \end{cases}$$
(1.6)

where the space-time evolution of the ψ_L and ψ_R fields is coupled by the mass term. In the case of a massless fermion the two chiral projections of the fermion field obey independent Weyl equations (1.6) and are called Weyl spinors. After the discovery of parity violation, since there was no evidence of a non-vanishing neutrino mass and the neutrinos were known to weakly interact just with their left-handed component, they were incorporated in the SM with one Weyl spinor.

Since a Weyl spinor has just 2 independent components, and they are enough to describe

²The electric charge Q is related to the third component of the weak isospin and the hypercharge through $Q = I_3 + \frac{Y}{2}$

a massless fermion, it is natural to ask whether 4 independent components are necessary to describe a massive fermion. The answer is no, as discovered by Majorana in 1937 [20]. It is sufficient to assume that ψ_L and ψ_R are not independent, showing a possible dependence that can solve eq. 1.6 such as

$$\psi_R = \xi \, \mathcal{C} \, \bar{\psi_L}^T \tag{1.7}$$

where C is the charge conjugation operator and $|\xi|^2 = 1$ is an arbitrary phase factor. Such fermion fields are called Majorana fermions and obey the Majorana field equation

$$i\gamma^{\mu}\partial_{\mu}\psi_{L} = m\mathcal{C}\psi_{L}^{T} \tag{1.8}$$

It is very important to point out now that Dirac and Majorana neutrinos have different phenomenological implications if and only if $m_{\nu} \neq 0$. In a massless Dirac theory ψ_L and ψ_R are independent and separately obey the Weyl equations. In a massless Majorana theory the very same equations hold with the relation 1.7. Since only the left-handed component interacts, it obeys the same equations in both formalisms and the right-handed component is irrelevant for the dynamics, the resulting physics is the same and the two theories are indistinguishable.

The most general neutrino mass term does not conserve the total lepton number L and includes both the left-handed flavor fields and right-handed sterile fields. In the simplified framework of just one generation of neutrinos, the lagrangian density reads

$$\mathcal{L}^{D+M} = -\frac{1}{2}m_L\bar{\nu}_L\nu_L^c - m_D\bar{\nu}_L\nu_R - \frac{1}{2}m_R\overline{(\nu_R)^c}\nu_R + h.c.$$
(1.9)

where m_L, m_D, m_R are real parameters and CP invariance is assumed. To summarize, there exist three phenomenologically possible neutrino mass terms.

- the Dirac mass term, where flavor left-handed and sterile right-handed fields enter, and L is conserved.
- the Majorana mass term, where just left-handed fields enter and L is violated.
- the most general Dirac and Majorana mass term (Eq. 1.9) where both flavor lefthanded and sterile right-handed neutrino fields are needed. The total lepton number L is not conserved in this scenario.

Further discussion about the possible neutrino mass terms, their implications and the generation mechanisms is available in [21].

1.1.3 Neutrino mixing

It is well established that the neutrinos that take part in charged current interaction exist in three orthogonal flavors: ν_e, ν_μ, ν_τ . In the hypothesis of massless neutrinos the *appearence* probability of a neutrino of a given flavor f and energy E to convert to another flavor f' after having travelled a distance L, $P(\nu_f \rightarrow \nu_{f'}; E, L)$ is null, and so is the *disappearence* probability $1 - P(\nu_f \rightarrow \nu_f; E, L)$. A Nobel Prize was awarded in 2015 to Takaaki Kajita and Arhur McDonald for the discovery of neutrino oscillations, which show that neutrinos have mass.

Neutrino oscillations in vacuum can indeed be explained if the fields of flavor neutrinos are linear combinations of the fields of three massive neutrinos,

$$\nu_{\alpha} = U_{\alpha,i} \,\nu_i \tag{1.10}$$

where a summation is assumed over repeated indices and U is the 3×3 unitary neutrino mixing matrix, known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. The PMNS matrix can be parametrized with 3 angles and either 1 or 3 independent CP-violating phases depending on whether neutrinos are Dirac or Majorana particles respectively

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \cdot \operatorname{diag}(1, e^{i\eta_1}, e^{i\eta_2})$$

$$(1.11)$$

where $c_{ij} \doteq \cos \theta_{ij}$ and $s_{ij} \doteq \sin \theta_{ij}$, $\delta \in [0, 2\pi]$ is the Dirac phase, $\eta_i \in [0, 2\pi]$ are the 2 Majorana phases and $\theta_{ij} \in [0, \pi/2]$ are the mixing angles between the ν_i mass eigenstates. In this convention, there are two non-equivalent orderings (hierarchies) for the spectrum of neutrino masses (Fig. 1.1),

- Normal Ordering (NO) where $m_1 < m_2 < m_3$
- Inverted Ordering (IO) where $m_3 < m_1 < m_2$



FIG. 1.1: Illustration of the mass spectra compatible with the data from neutrino oscillations; left, normal hierarchy; right, inverted hierarchy. The length of the colored bars are proportional to the flavor content of each mass state, $|U_{li}|^2$. [18]

In vacuum oscillations the transition probability of a neutrino of flavor α and energy *E* to a flavor β after a flight of lenght *L* is given by

$$P_{\alpha\beta} = \delta_{\alpha\beta} - 4\sum_{i$$

where

$$X_{ij} = \frac{(m_i^2 - m_j^2)L}{4E} = 1.267 \frac{\Delta m_{ij}^2}{\text{eV}^2} \frac{L/E}{\text{m/MeV}}$$
(1.13)

and depends on the neutrino masses just through their squared differences Δm_{ij}^2 . Since $\sin^2(\theta_{13}) \sim 0.022$ is small, neutrino oscillations can be effectively described in many applications with a simpler 2-flavor mixing model. This means that the neutrino mass scale and ordering cannot be determined by (vacuum) oscillation experiments. Exceptions are experimental searches where the second sum in Eq. 1.12 is relevant [22]. In addition it can be shown that the oscillation observables cannot depend on the Majorana phases either. Nevertheless the coherent interaction with matter of neutrinos propagating in a dense medium is sensitive to mass ordering via the MSW effect, for this reason the sign of $\Delta m_{21}^2 > 0$ is known. A review of the most recent experimental results on the neutrino masses, mixing and oscillations can be found here [23].

1.2 Theoretical overview of double-beta decay

A SM double beta decay consists in the conversion, within an atomic nucleus, of two neutrons in protons with the simultaneous emission of 2 electrons and 2 anti-neutrinos. It was first pointed out by M. Goeppert-Mayer in 1935, together with the first half life calculations based on the Fermi theory [24]. From those it was already clear that the half life of $\beta\beta$ decaying nuclei must have been $T_{1/2} > 10^{17}$ y. Since the mass number A = N + Zis conserved, it is useful to distinguish between even and odd mass number nuclei. The available energy for the final state particles is

$$Q = (M_i - M_f - 2m_e)c^2 (1.14)$$

where M is the mass of the initial and final state nuclei, m_e is the electron mass and the neutrino mass is neglected. In addition, since the $\beta\beta$ decay is a second order weak decay, it can effectively happen just in nuclei where the single-beta decay is energetically forbidden. To this regard, since the nuclear binding energy has a pairing term that favors even configurations of nucleons, the even mass number isobars lie on two different curves as a function of the atomic number³ while the odd mass number ones lie on a single curve (Fig. 1.2). The existence of an energetically allowed $\beta\beta$ decay mode for an odd mass number nucleus, implies that the single β decay is allowed as well, and since the latter process has

 $^{^{3}}$ An even number can be obtained either from the summation of two even numbers, or from the summation of two odd ones, while odd numbers can just come from the summation of an even to an odd number.



FIG. 1.2: Qualitative representation of the nuclear mass dependence on atomic number for even (left) and odd (right) mass number isobars. For some even-even nuclei whose nuclear mass structure resembles the above sketch, the $\beta\beta$ decay mode is possible.

a much larger decay rate, the first one effectively never happens. Some even mass number nuclei, instead, can have a configuration where the single β decay is forbidden because of the pairing term, while the $\beta\beta$ decay is allowed. This is the case for the nuclei shown in Fig. 1.6.



FIG. 1.3: Diagrams of double beta decay processes in the electro-weak SM framework (left) and with a light Majorana neutrino exchange (right).

The rate Γ of double beta decay can be parametrized differently according to the decay mode,

$$\Gamma_{2\nu} = G_{2\nu} |M_{2\nu}|^2 \tag{1.15}$$

$$\Gamma_{0\nu} = G_{0\nu} |M_{0\nu}|^2 \frac{\langle m_{ee} \rangle^2}{m_e^2}$$
(1.16)

where G is the phase space factor and can be computed with very high accuracy [25]. M is the nuclear matrix element connecting the initial and final nuclear states, m_e is the electron mass and $\langle m_{ee} \rangle$ is the effective Majorana neutrino mass,

$$\langle m_{ee} \rangle = \left| \sum_{i} U_{ei}^2 \, m_i \right| \tag{1.17}$$

where m_i are the neutrino mass eigenvalues and U_{lj} the elements of the PMNS matrix defined in Eq. 1.11. It is important to stress that cancellations might occur in Eq. 1.17 hence the neutrino-less double beta decay rate may vanish even if neutrinos were Majorana particles. The same argument that motivates the search for an explanation for the *smallness* of the neutrino masses provides hope that, unless some underlying symmetry or principle comes into play, an exact cancellation of the effective Majorana neutrino mass is unlikely.

1.2.1 Nuclear Matrix Elements

The nuclear matrix elements $M_{0\nu}$ and $M_{2\nu}$ connect the same nuclear states and can be written in a similar form, i.e. as a bracket of the wavefunction describing the final state of the daughter even-even nucleus, the one of the parent nucleus and an operator in between. A very different role is played by the intermediate nuclear states in the two double-beta decay modes. In the $0\nu\beta\beta$ mode the momentum transfer $q \sim 100 \text{ MeV}^4$ is much higher than the excitation energy of the intermediate nuclear states, thus the energy of the intermediate states can be approximated with an average energy (closure approximation). In $2\nu\beta\beta$ decay instead, since the momentum transferred to the nucleons is of the same order of magnitude as the neutrino momentum ($\sim 1 \text{ MeV}$), just a subset of the intermediate nuclear states contribute to the matrix element. For this reason the comparison between the $M_{2\nu}$ and $M_{0\nu}$ and especially the experimental validation of the latter is not straightforward. The next generation $0\nu\beta\beta$ decay experiments will aim at fully covering the $m_{\beta\beta} > 10 \text{ meV}$ region, so that if neutrinos are Majorana particles a signal will be detected in the hypothesis of inverted hierarchy. A wide range of theoretical approaches to the calculation of NME for $\beta\beta$ decay exists, and the Reader is referred to [26] [27] for a review. In the following an overview of the available techniques will be given, together with relevant references to 130 Te $\beta\beta$ NME calculations. The uncertainty on NME propagates through Eq. 1.16 from a rate measurement to the determination of the corresponding effective Majorana neutrino mass. At the present date, the uncertainty on the NME evaluation ranges from a factor 2 to 3. In addition, the NME evaluation suffers from the so called g_A problem, as detailed in the following. The approaches to the NME calculation rely on one of the following techniques, as far as the neutrinoless mode is concerned

• ISM (Interacting Shell Model) [28] [29], is based on the idea that most of the nu-

⁴This result can be obtained from the uncertainty principle considering that a virtual neutrino must be exchanged between two nucleons that are ~ 1 fm apart.

cleons do not contribute to the interaction, except the ones close to the Fermi sea. Such *valence* nucleons are the degrees of freedom of ISM models, and the correlations among them are relevant. The dynamics of valence nucleons is restricted to a small number of single-particle levels close to the Fermi surface. An effective interaction is used, and many-particle states are obtained from linear combination of Slater determinants of single-particle nucleon states that diagonalize the effective hamiltonian operator. Since correlations (collective excitation of the available degrees of freedom) are important, the width of configuration space is limited by the corresponding model complexity.

- QRPA (Quasi Random Phase Approximation) [30] [31] [32], considers a much larger space of single-particle orbitals with respect to ISM, but with a restricted set of correlations. The interaction is a realistic nucleon-nucleon potential where the interaction strength of the proton-neutron pairing g_{pp} is usually modified to match some experimental input, e.g. match the $2\nu\beta\beta$ decay rate and then the same interaction strength is used to extract $M_{0\nu}$. The $\beta\beta$ NME are very sensitive to g_{pp} and investigations on this topic have produced attempts to cure the part of this behavior that is thought to be an artifact of the method (Renormalized QRPA). The results of QRPA NME calculations are systematically larger than the corresponding ISM results.
- EDF (Energy Density Functional theory) [33] [34] [35] are based on the minimization of an energy functional with respect to local density functions (e.g. nucleon number density). Given the functional, its minimization provides ground-state energy and density. The functional parameters are adjusted to match ground-state properties of a set of nuclei, and then used with no change for any nucleus. These methods can provide a good description of collective properties, but they do not work for all quantities in all nuclei. In general the calculations from EDF methods are larger than QRPA.
- IBM (Interacting Boson Model) [36], aims at sharing both the inclusion of all correlations around the Fermi surface and a careful treatment of collective motion at the cost of including more abstract degrees of freedom than just nucleons, and a phenomenological approach to nuclear structure. The IBM-2 version of this method, used to treat $\beta\beta$ decay, uses separate boson states to treat sets of many neutron and protons with different angular momenta. The results of $0\nu\beta\beta$ NME usually lie in between the EDF and ISM ones.

Some of these models can be used to compute $2\nu\beta\beta$ decay NME as well. Nonetheless a systematical overestimation of the $\beta\beta$ decay rates with respect to the observed ones is found, which motivates the introduction of a quenched axial coupling constant g_A . In this respect, the validation of theoretical calculations with experimental results is crucial. The source of the quenching factor can be twofold: omission of non-nucleonic degrees of freedom, limitation of the space in which the calculation is performed. The first one yields a g_A quenching that does not depend on the mass number, the other produces a model dependent quenching that grows with increasing mass number. The axial vector coupling constant appears to the second power in the NME, hence to the fourth power in the half life. If the $0\nu\beta\beta$ calculation needs the same quenching as $2\nu\beta\beta$, the expected half lives might be off by about one order or magnitude. See [37] for a review.

1.2.2 General remarks

The search for neutrino-less double beta decay provides insight on many open aspects of modern particle physics and cosmology. As already pointed out, because of their tiny mass neutrinos do not fit the SM description and this is why the possibility of them being the only known fundamental Majorana fermions is open. A $0\nu\beta\beta$ discovery would corroborate this hypothesis, might solve the neutrino mass hierarchy problem and constrain the absolute neutrino mass scale. The green and red areas in Fig. 1.4 correspond to the allowed values of effective Majorana neutrino mass as a function of the lightest neutrino mass for inverted and normal ordering respectively. The allowed points in this space do not fall on a line because any value of the Majorana phases in the $[0, 2\pi]$ interval is allowed. Stronger upper limits on the neutrino-less double beta decay rate are converted in upper bounds to the $\langle m_{ee} \rangle$ parameter. Other independent pieces of information about excluded areas in the plot of Fig. 1.4 come from the kinematics of single β decay. A precise measurement of β decay endpoints (e.g. the KATRIN experiment recently placed a $m_{\nu} < 1.1 \,\mathrm{eV}$ upper limit on the neutrino mass [38] from the endpoint of the emitted electron from tritium decay) provides a model independent limit on m_{lightest} . Other observables from cosmology such as the cosmic microwave background (CMB) anisotropy allow to constrain the sum of neutrino masses to much smaller values such as $\sum m_{\nu} < 120 \text{ meV}$ [39], correspondingly excluding lightest neutrino mass values larger than few tens of meV. Further considerations on the combination of cosmological data and $0\nu\beta\beta$ decay searches can be found here [40].

1.2.3 Baryon asymmetry in the universe (BAU)

Neutrino-less double beta decay is a lepton number violating process; 2 leptons are created out of an initial baryonic state. It is well established that the known universe favors matter over anti-matter as quantified by the asymmetry parameter

$$\eta = \frac{n_B - \bar{n}_B}{n_\gamma} \tag{1.18}$$

where $n_B(\bar{n}_B)$ is the baryonic (anti-baryonic) matter density and n_{γ} is the photon number density, allowing for the formation of complex structures such as heavy nuclei and, eventually, biological life. BAU can be measured independently of the abundances of light elements in the intergalactic medium and from the power spectrum of temperature fluctuations in the CMB. Both yield an asymmetry of the order of $\eta \sim 10^{-10}$. A review of this topic can be found in [44]. It can be shown that in order to attain baryogenesis, three



FIG. 1.4: Assuming $0\nu\beta\beta$ decay is mediated by light Majorana neutrinos, a lower limit on the $T_{1/2}^{0\nu}$ parameter corresponds to an upper limit on the effective Majorana mass $m_{\beta\beta}$ depending on the nuclear matrix elements used (yellow band). The width of the band corresponds to the spread of the NME used [41]. The expected CUORE sensitivity after 5 yr of live time is shown for comparison (blue band). The red and green bands correspond to the allowed values of the effective Majorana mass as a function of the lightest neutrino mass [42] in the normal and inverted ordering hypothesis respectively. The dark (light) areas show the allowed parameter space assuming the best fit (3 σ gaussian intervals) for the neutrino masses and mixing parameters reported in [43]. The Majorana phases range from 0 to 2π .

Sakharov [45] conditions must be met:

- baryon number violation,
- C and CP violation,
- interactions out of thermal equilibrium.

The Standard Model meets all these criteria, but insufficiently to explain the observed asymmetry. The expansion of the universe brings the primordial plasma out of thermal equilibrium, so that the last condition is met. CP violation has been observed in the quark sector, meeting the second condition. The baryon number in the SM is conserved at each perturbative order, together with three lepton numbers, but is violated by non-perturbative effects. The amount of CP violation in the quark sector can be parameterized with the lowest order invariant of the CKM matrix, i.e. the Jarlskog determinant [46] D, and an order of magnitude of the corresponding baryon asymmetry parameter can be given by

$$D/T_{sph}^{12} \sim 10^{-20} \ll \eta$$
 (1.19)

where T_{sph} is the temperature at which sphaleron processes⁵ can take place. This is not a direct proof of impossibility, but the large mismatch with the observed baryon asymmetry in the universe makes SM baryogenesis challenging.

A large number of models that fulfill the Sakharov conditions have been developed, and they can be split into those that directly generate a BAU in the baryonic sector and the ones that initially generate a lepton asymmetry which is eventually transferred to baryons (leptogenesis) either via sphalerons or via other Beyond SM processes (see e.g. [47]).

1.3 Experimental considerations

The experimental search for $0\nu\beta\beta$ decay is based on the different spectral shape of the neutrino-less mode with respect to the $2\nu\beta\beta$ decay mode. As shown in Fig. 1.5 as a function of the sum energy of the emitted electrons, the first appears as a sharp peak at the Q value while the latter has a continuous spectrum whose endpoint is the Q value. Both of the spectra need to be convoluted with the detector response function, and this is why a good energy resolution is a fundamental parameter to build a sensitive experiment.

1.3.1 Sensitivity

In this section a sensitivity formula for the discovery sensitivity of a process whose signature is a monochromatic peak over an uniform background will be given. Two cases must be distinguished: background-free and finite-background search. In any experimental setup a resolution function must be taken into account.

⁵Non-perturbative B-violating SM processes



FIG. 1.5: Sum energy spectrum of electrons emitted in double beta decay in both standard $(2\nu\beta\beta)$ and neutrino-less $(0\nu\beta\beta)$ mode.

Let us assume an initial number N_0 of emitting sources ($\beta\beta$ emitters) are present in the detector. From the law of radioactive decay at time t the number of emitters will be

$$N(t) = N_0 e^{-t/\tau} (1.20)$$

where τ is the decay constant. Assuming that the measurement time $t \ll \tau$ is much shorter than the half life of the emitter, from Eq. 1.20 follows that in a time Δt the number of $\beta\beta$ disintegrations (signal events) is

$$\mathcal{N} = -\frac{dN}{dt} \,\Delta t = N_0 \frac{\ln 2}{T_{1/2}} \Delta t \tag{1.21}$$

where $T_{1/2} = \tau \ln 2$ is the half life of the emitter. Let us restrict to the case where the emitter is ¹³⁰Te, embedded in ^{nat}TeO₂ crystals. The number of emitters can then be expressed as a function of the total mass of the crystals $M(^{nat}\text{TeO}_2)$, the molecular mass of tellurium dioxide $m(^{nat}\text{TeO}_2)$ and the natural isotopic abundance η of the isotope of interest,

$$N_0 = \frac{M(^{nat}\text{TeO}_2) \,[\text{kg}]}{m(^{nat}\text{TeO}_2)[\text{g/mol}]} \,\eta \,N_A[\text{mol}^{-1}] \,10^3$$
(1.22)

where N_A is the Avogadro constant. In order to simplify even further, let us consider the smallest energy range ΔE that contains a sufficiently high integral of the resolution function of the detector, e.g. a $\pm 3\sigma$ range for a gaussian response function, in order to perform a counting experiment. The number of detected signal counts $\mathcal{N}_D = \epsilon \mathcal{N}$ can be derived from the number of signal events applying a detection efficiency correction. With the assumption of a constant background distribution as a function of the reconstructed energy, the expected number of background counts within the ΔE energy range is

$$B = b M \Delta t \,\Delta E \tag{1.23}$$

where M is the total active detector mass and b is a constant background index, which we will assume to know with very high accuracy. Let us also naively define the discovery sensitivity $\mathcal{N}^{5\sigma}$ as the number of detected signal counts whose p-value in the backgroundonly hypothesis respects the identity

$$p = \frac{1 - \operatorname{erf}(n/\sqrt{2})}{2} = \int_{-\infty}^{n\sigma} g(x|0,\sigma) dx \quad \text{where} \quad n = 5$$
(1.24)

and $g(x|\mu,\sigma)$ is a normalized gaussian distribution of mean μ and variance σ^2 .

Background free hypothesis Assuming the background expectation is much smaller than unity $B \ll 1$, the p-value can be computed as

$$p = \int_{\mathcal{N}^{5\sigma}}^{+\infty} e^{-\mu} d\mu \tag{1.25}$$

Enforcing the condition in Eq. 1.24 and substituting, the sensitivity in the background-free case can be expressed in terms of the inverse of the decay rate of the process of interest as

$$S_{\rm bkg-free}^{5\sigma} \simeq \frac{N_A \eta \, 10^3}{m(^{nat} {\rm TeO}_2)} \frac{\ln(2) \epsilon}{\left[-\ln(3 \, 10^{-7})\right]} M\Delta t$$
 (1.26)

A different choice of the p-value threshold would have implied just a modification in the argument of the logarithm in the denominator.

Finite background hypothesis This is also a limit case, where the number of expected background counts over the whole measurement is much greater than unity $B \gg 1$ so that the distribution of the measured background counts can be safely assumed gaussian with mean B and variance B as well. In this case Eq. 1.24 implies that

$$S_{\text{finite-bkg}}^{5\sigma} = \frac{N_A \eta \, 10^3}{m(^{nat}\text{TeO}_2)} \frac{\ln(2) \epsilon}{5} \sqrt{\frac{M \,\Delta t}{b \,\Delta E}}$$
(1.27)

Score function definition It might come handy to factorize some terms in Eq. 1.26, 1.27 defining a score function F such that

$$S^{5\sigma} = \frac{N_A \eta \, 10^3 \, \ln(2)}{m(^{nat} \text{TeO}_2)} \frac{M\Delta t}{5} \cdot F \tag{1.28}$$

$$F = \begin{cases} \frac{\epsilon}{\sqrt{B}} & \text{finite background} \\ \frac{5\epsilon}{-\ln\left(3\,10^{-7}\right)} & \text{background free} \end{cases}$$
(1.29)

1.3.2 Detector design

Let us stress in the very beginning that there is no optimal choice, as pointed out in [48], in terms of source. All double beta decay emitter isotopes are (within a factor 3 or

so) equivalent from the point of view of the relation between effective Majorana mass and neutrinoless double beta decay rate. This result was obtained from the anti-correlation between the phase space and the square of the nuclear matrix element for a wide range of isotopes. Nonetheless, as pointed out in Sec. 1.3.1, other parameters are involved in determining the sensitivity such as the ones that appear in Eq. 1.26 and Eq. 1.27. Those are the following,

- isotopic fraction, appears linearly in the sensitivity expression regardless of the expected background. The double beta decay emitter with the highest natural isotopic abundance is ^{130}Te and this allowed the construction of CUORE from natural Te avoiding the cost and complexity of isotopic enrichment
- background, together with the energy resolution can determine whether the sensitivity scales linearly or with the square root of live time. One of the most challenging features of experimental searches for $0\nu\beta\beta$ decay is indeed the background reduction, particularly from the point of view of the definition of specifications for material selection, detector production and quality control, transportation, assembly procedures (see e.g. [49] [50]). In addition not only the total activity of the detector and surrounding material matters, but also its distribution. Surface alpha contaminations, in the case of CUORE, dominate the background close to $Q_{\beta\beta}$
- resolution, in background dominated experiments the sensitivity is inversely proportional to the square root of the resolution. Otherwise a background free operation can be attained, if the expected number of counts within few energy resolutions is below unity
- efficiency, is a crucial parameter because it sets constraints to the detector geometry. The size of the CUORE bolometer crystal directly impacts the probability of a $\beta\beta$ decay event to fully release its energy in one detector
- detector mass, directly proportional to the number of emitter nuclei is also often used to normalize the background and quote it in terms of a background index (i.e. number of background counts per unit of detector mass and energy). In background free experiments the sensitivity grows linearly with detector mass. In a background dominated regime the half life sensitivity is proportional to its square root. In any case large detector masses help not only because the number of emitters is larger but because typically a relevant fraction of the background sources does not come from the source material itself, rather from the surrounding infrastructure. In CUORE the innermost copper vessel is a non-negligible source of background but the closely packed arrangement of the detector array helps reducing it via self-shielding. On the other hand scalability can be an issue for ton-scale detectors, and it is not always possible to grow indefinitely in mass without facing significant practical challenges. One such example comes, again, from the CUORE experiment. A custom made cryostat had to be designed, built and operated in order to house it (see Sec. 2.3).



FIG. 1.6: Double beta decay emitters Q value as a function of their natural isotopic abundance.

1.3.3 Competing experiments

A large international effort is under way to search for neutrinoless double- β decay in several emitter nuclei. In Table 1.7 a selection of the latest and current active experiments searching for neutrino-less double beta decay is reported. A short description follows, for each of the mentioned experiments.

CUPID-0 CUPID (CUORE Upgrade with Particle IDentification) is a next generation project for measuring $0\nu\beta\beta$ decay with scintillating bolometer crystals. The bolometric technique is widely discussed in Sec. 2.2. The main advancement with respect to CUORE is the capability to discriminate and reject the alpha background. CUPID-0 is the first

Experiment	Isotope	FWHM	${ m T_{1/2}^{90\%}}$	Sensitivity
		[keV]	$[10^{\overline{25}} \mathrm{y}]$	$[10^{25} \ { m y}]$
[41] CUORE	¹³⁰ Te	7.0	3.2	1.7
[51] CUPID-0	$^{82}\mathrm{Se}$	20	0.35	0.50
[52] EXO-200	$^{136}\mathrm{Xe}$	67	3.5	5.0
[53] GERDA	$^{76}\mathrm{Ge}$	3.3	9	11
[54] KamLAND-Zen	$^{136}\mathrm{Xe}$	270	10.7	5.6
[55] MAJORANA	$^{76}\mathrm{Ge}$	2.5	2.7	4.8

TABLE 1.7: A selected set of experimental searches for $0\nu\beta\beta$ decay together with the latest results in terms of 90% C.I. lower limits and expected sensitivity to the inverse of the neutrino-less double beta decay rate. The resolution at $Q_{\beta\beta}$ is reported as well.

pilot experiment of CUPID, is made of 2 natural ZnSe crystals and 24 enriched ones at 95 % in the ⁸²Se emitter. Each crystal is equipped with an Neutron Transmutation Doped (NTD) Ge thermistor, surrounded by 3M Vikuiti plastic reflective foils and sandwiched between two bolometer light detectors, equipped with NTD thermistors as well. They are arranged in 5 towers and placed in a dilution refrigerator. The energy resolution at the Q value of ⁸²Se (2997.9 ± 0.3 keV) is 20.05 ± 0.34 FWHM. The background index around the Q value is $3.5^{+1.0}_{-0.9} \times 10^{-3}$ counts/(keV kg yr), dominated by muon interactions. The 90 % C.I. median sensitivity is 5.0×10^{24} yr [51].

EXO-200 The EXO-200 detector is made of a cylindrical Time Projection Chamber (TPC) filled with liquid Xenon (LXe) enriched in the ¹³⁶Xe emitter isotope at 80.6%. A common cathode splits the TPC, enclosed in a radiopure copper vessel, into two identical drift regions. Interactions in LXe produce ionization charges that are read out by wire planes on the anodes placed at each end of the two volumes. The scintillation light, instead, is read out by arrays of large area avalanche photodiodes placed behind the anode planes. The sensitivity to different topologies of energy releases in the detector volume allow background rejection. A recent publication [52] reports the results of a $0\nu\beta\beta$ search with a total ¹³⁶Xe exposure of 234.1 kg yr acquired from September 2011 to December 2018 in two data taking campaigns. The resolution at $Q_{\beta\beta} = 2457.83 \pm 0.37$ keV is $\sigma/E = (1.15\pm0.02)\%$. The reconstructed backgrounds are $(1.7\pm0.2) \times 10^{-3} \text{ cts/(keV kg yr)}$ and $(1.9 \pm 0.2) \times 10^{-3} \text{ cts/(keV kg yr)}$ for the two data taking campaigns respectively, normalized to the total fiducial Xe mass, including all isotopes. A median 90 % C.L. sensitivity of 5.0×10^{25} yr is computed and a lower limit of $T_{1/2} > 3.5 \times 10^{25}$ yr at 90% C.L. is set for the $0\nu\beta\beta$ half life.

GERDA The GERmanium Detector Array explores double beta decays of ⁷⁶Ge through 37 enriched (> 85 %) Germanium diode detectors with different geometries operated in liquid Argon (lAr). The surrounding lAr cools the detector down to operating temperature. At the same time it acts as a scintillating high-purity shielding from background radiation, read out by wavelength shifting fibers coupled to photomultipliers. A water Cherenkov veto surrounds the experiment. The GERDA detectors features an excellent energy resolution at $Q_{\beta\beta} = 2039.061 \pm 0.007$ keV of either 3.6 ± 0.1 keV or 3.0 ± 0.1 keV, depending on the diode geometry. Moreover the GERDA search for $0\nu\beta\beta$ decay is backgroundfree, i.e. < 1 count is expected in the energy region of interest with an exposure of 82.4 kg yr, thanks to the radiopurity of the detector and surrounding material and the powerful topological background discrimination allowed by pulse shape discrimination in Germanium detectors. The reconstructed background index is $5.7^{+4.1}_{-2.6} \cdot 10^{-4} \text{ cts}/(\text{keV kg yr})$ or $5.6^{+3.4}_{-2.4} \cdot 10^{-4} \text{ cts}/(\text{keV kg yr})$ depending on the detector geometry. A 90 % C.L. lower limit for the $0\nu\beta\beta$ decay is reported as $T_{1/2} > 0.9 \cdot 10^{26}$ yr. The corresponding sensitivity for limit setting is $1.1 \cdot 10^{26}$ yr [53]. **KamLAND-Zen** Made of 13 tons of Xe-loaded liquid scintillator placed in a spherical nylon film balloon (Inner Balloon, IB), KamLAND-Zen searches for neutrinoless double beta decay of ¹³⁶Xe. The IB is contained in a 1 kton liquid scintillator outer balloon acting as veto shield. The scintillation light is read out by photomultiplier tubes. The Xe content of the scintillator in the IB is (2.91 ± 0.04) % by weight, and the isotopic fraction of ¹³⁶Xe emitter nuclei is (90.77 ± 0.08) %. The observed energy resolution is $\sigma/E \sim 7.3 \ \%/\sqrt{E(MeV)}$ which corresponds to 270 keV FWHM at $Q_{\beta\beta}$. The background cannot be described with a simple number due to its dependence on the fiducial volume and on time. A lower limit is placed on the neutrinoless double beta decay half life of $T_{1/2} > 10.7 \cdot 10^{25}$ yr at 90 % C.L. and an expected sensitivity of $5.6 \cdot 10^{25}$ yr [54].

MAJORANA The MAJORANA Demonstrator is an array of 58 high-purity Ge (HP-Ge) detectors aimed both at searching for $0\nu\beta\beta$ decay of ⁷⁶Ge and motivating the construction of a tonne-scale Ge experiment. The detector is enclosed in low-radioactivity copper shieldings surrounded by 45 cm of high purity lead. A passive neutron shielding is deployed together with active radon control and muon veto systems. With an exposure of 26.0 ± 0.5 kg yr, the MAJORANA Demonstrator reached an energy resolution at $Q_{\beta\beta}$ of 2.53 keV FWHM. A strong topological background rejection capability allows, thanks to the short range of electrons in the *Ge* detectors, to disentangle single-site signal-like events from backgrounds. Surface alpha particle can be tagged and rejected. A background free operation⁶ allowed to place a lower limit on the $0\nu\beta\beta$ half life of $T_{1/2} > 2.7 \cdot 10^{25}$ yr at 90% C.L. with a median expected sensitivity of $4.8 \cdot 10^{25}$ yr [55].

⁶The expected number of background events in the $0\nu\beta\beta$ region of interest is 0.65, one event is observed.

Chapter 2

CUORE

The CUORE experiment is a ton-scale cryogenic bolometer array of 988 detectors aimed at searching neutrinoless double-beta decay in 130 Te [56] [57] [58]. It stems from decades of development and the pioneering work of Ettore Fiorini and collaborators as further described in the following Sec. 2.1. The basic operating principles as well as a simplified model of a bolometer will be outlined in Sec. 2.2. More details on the large cryogenic infrastracture that hosts the detectors can be found in the following Sec. 2.3. Finally, in order to extract physics data from temperature variations, a calibration of the energy scale is needed. The calibration system of the CUORE detector is described in Sec. 2.4.

It is worth pointing out in the very beginning that a nomenclature difference exists between *bolometer* and *micro/macro-calorimeter* in the astroparticle physics community, the first being devoted to radiation detection and the latter being aimed at a more general single particle detection. In addition, the word *calorimeter* in (high energy) particle physics refers just to detectors able to stop the impinging particle possibly containing the whole shower that follows the interaction of the primary particle with the calorimeter, collect the emitted scintillation light, process and reconstruct from it the energy of the primary. Since in the double beta decay community only macro-calorimeters are used there is no ambiguity and it is possible to refer to these detectors as *bolometers*. This convention will be applied in all the following manuscript.

2.1 History and predecessors

Let us assume a closed system made of an impinging particle and a crystal working as a detector. A straightforward consequence of the 1st law of thermodynamics is that, if the detector absorbs the particle without getting lattice deformations (i.e. all the particle energy is converted in kinetic energy of the detector atoms), an amount of heat equal to the work necessary to stop the particle is transferred to the detector that, unavoidably, heats up. The basic idea of bolometric detectors is to measure the energy of the impinging particle via a measurement of the temperature variation induced in the detector. In the late 1970s it was already proposed to use a calorimetric approach in particle physics, in



FIG. 2.1: The running of the sensitivity to neutrino-less double beta decay with bolometric experiments is shown as a function of time, from the first pioneering works to date.

particular to detect a possible *gentle neutrino scattering* that could explain the missing solar neutrino flux in the Homestake experiment [59] [60].

Conventionally the birth of the low-temperature calorimetry technique dates back to 1984 when Fiorini and Niinikoski proposed [61] to apply it to the detection of neutrinoless double beta decay in an underground low-counting environment.

A key requirement for rare event detection is indeed a low background environment. The optimal place to perform such experiments is an underground laboratory such as the Laboratori Nazionali del Gran Sasso (LNGS) of the Istituto Nazionale di Fisica Nucleare (INFN), located in Assergi, Italy. The LNGS provide a 3600 m.w.e. rock overburden and in the experimental hall of the laboratory the total flux of neutrons Φ_n and muons Φ_{μ} is as low as

$$\Phi_n = 4 \cdot 10^{-6} \,\mathrm{cm}^{-2} \mathrm{s}^{-1} \tag{2.1}$$

$$\Phi_{\mu} = 3 \cdot 10^{-8} \,\mathrm{cm}^{-2} \mathrm{s}^{-1} \tag{2.2}$$

respectively [62] [63].

A dilution refrigerator was installed there in 1989 and the first tests of various crystals from 6 g up to 340 g were performed (see [64] for a complete review). After the successful operation of a single 340 g crystal, arrays of bolometers started to be tested [65]. In the following, for the sake of simplicity, just the latest steps before the realization of CUORE will be presented in greater detail.

2.1.1 CUORICINO

CUORICINO ran in Hall A of LNGS in 2003-2008. It was a single tower with active mass of 40.7kg. Made of both natural $5 \times 5 \times 5$ cm³ and enriched $6 \times 3 \times 3$ cm³ TeO₂ crystals,

it collected 19.75 kg · yr of ¹³⁰Te exposure over the whole data taking. Besides from being a proof of concept to demonstrate the feasibility of running a complex bolometer array for 5 years, it provided relevant physics results on its own. The best detector performances, both in terms of resolution and background, were achieved with the $5 \times 5 \times 5$ cm³ crystals. The resolution of the array was computed as harmonic mean FWHM¹ at the ²⁰⁸Tl line, weighted for the physics exposure (5.8 ± 2.1) keV. The background index in the region of interest (ROI) for the ^{nat}Te crystal was (0.153 ± 0.006) cts/(keV kg yr) [66]. The CUORICINO experiment was able to set a lower limit on the $0\nu\beta\beta$ decay half life of ¹³⁰Te of $T_{1/2}^{0\nu} >$ $2.8 \cdot 10^{24}$ yr at 90 % C.L. [67].

2.1.2 CUORE-0

The CUORE-0 detector was the last step before the CUORE construction, needed to validate new standardized and automated procedures for the detector assembly. They included the gluing and bonding of sensors and heaters, the tower wiring, the storage and transportation of materials. Up to CUORICINO all the above activities were handcrafted and this could not be scaled to the CUORE size. CUORE-0 was made of just one CUORElike tower. 52 cubic crystals of 5 cm side made of natural TeO₂ and weighting about 750 g each were arranged in 13 floors of 4 crystals each.

CUORE-0 collected data from 2013 to 2015 and showed a greater uniformity in terms of bolometric performance than CUORICINO. In addition the energy resolution improved and the exposure weighted FWHM at the $Q_{\beta\beta}$ was measured to be (4.9 ± 2.1) keV [66], compatible with the 5 keV design goal for CUORE.

Another important step that CUORE-0 was supposed to validate was the effectiveness of the material selection and cleaning procedures identified (see [64] for a review) to mitigate the background in the ROI. The result was successful, the measured background index $(0.058 \pm 0.004_{\text{stat}} \pm 0.002_{\text{syst}}) \text{ cts}/(\text{keV kg yr})$ [68] was a factor 3 lower than the predecessor allowing to predict that the design goal of CUORE of a background index in the ROI of ~ 0.01 cts/(keV kg yr) was within reach.

Finally, with an exposure of 9.8 kg · yr of ¹³⁰Te, a lower limit on the $0\nu\beta\beta$ half life was placed at $T_{1/2}^{0\nu} > 2.7 \cdot 10^{24}$ yr at 90 % C.I. with a Bayesian analysis [68]. Combining the result with the one from CUORICINO it was obtained $T_{1/2}^{0\nu} > 4.0 \cdot 10^{24}$ yr at 90 % C.I.

2.2 The bolometric technique

A bolometric detector is a device where the energy of incoming particles is converted to phonons and ultimately measured from the temperature variation of the absorber. Since the thermal variations are of the order of $10 - 100 \,\mu\text{K/MeV}$, in order to be able to measure the tiny temperature variations induced by natural radioactivity, the operating temperature needs to be kept at $\sim 10 \,\text{mK}$.

¹Full Width Half Maximum

2.2.1 Minimal thermal model

The necessary components needed to build a bolometric detector are an absorber with a low heat capacity at low temperature, a weak thermal coupling to a heat sink and a thermometer. Let us model the absorber as a temperature independent thermal capacity C linked through a thermal conductance G to a heat sink (i.e. an object with infinite thermal capacity) of temperature T_0 . If at the time t = 0 an amount of energy E is released in the absorber, its temperature will change by

$$\Delta T = \frac{E}{C} \; . \label{eq:deltaT}$$

The istantaneous change of temperature is an approximation, but it is valid as long as the thermalization time is much shorter than all the other time scales of the system. The energy released produces phonon excitations of the lattice that travel at the speed of sound c_s in the absorber, and ultimately are thermalized in a timescale of the order of

$$au_{phonon} \sim \frac{L}{c_s}$$

where L is the typical size of the absorber. For a TeO₂ cube of 5 cm side, the speed of sound is roughly $c_s \sim 4 \cdot 10^3$ m/s so the thermalization time is of the order of $\tau_{phonon} \sim 10 \,\mu$ s. At the moment, the only other timescale is set by the ratio of the crystal heat capacity and the thermal conductance of the thermal link with the heat bath. Neglecting the transient in which thermalization happens, the time evolution of the crystal temperature can be described as

$$T(t) - T_0 = \frac{E}{C} e^{-t/\tau} \quad \text{where} \quad \tau = \frac{C}{G} .$$
(2.3)

The thermal conductance needs to be matched to the heat capacity of the deployed absorber, which is required to be as low as possible in order to maximize the amplitude of the thermal pulses. For this reason, the best materials are dielectric and diamagnetic crystals as well as superconductors below the transition temperature. In these cases the dominant contribution to the heat capacity comes from the crystal lattice and the Debye law can be applied to describe it at low temperatures,

$$C \propto \left(\frac{T}{\Theta_D}\right)^3 \tag{2.4}$$

where Θ_D is the Debye temperature and is a property of the material. For TeO₂ the Debye temperature is $\Theta_D(\text{TeO}_2) = (232 \pm 7) \text{ K}$ [69].

2.2.2 The temperature sensor

The device described in the previous section cannot be used to detect particles simply because a temperature sensor is still missing. The addition of the temperature sensor and
its operation implies the insertion of new objects into the thermal model of the detector.

A good temperature sensor should be strongly coupled to the absorber in order to follow its temperature variations, should have a small heat capacity (possibly negligible, with respect to the absorber) and have a well known steep dependence of one of its physical parameters as a function of temperature in the range of operation of the absorber. Example of such detector are semiconductor thermistors, transition-edge sensors, metallic magnetic calorimeters. A complete review of detection techniques can be found at this reference [70].

In the following we will focus on semiconductor thermistors because this is the device used in CUORE. Generally speaking, a semiconductor thermistor is made of a small silicon



FIG. 2.2: Picture of a NTD thermistor (left) and diagram (right) of its geometry. The yellow parts represent gold pads, used for electrical contacts. Typical size: L = 3.0 mm, W = 2.9 mm, H = 0.9 mm, P = 0.2 mm [71]

or germanium crystal with a doped region. The Neutron Transmutation Doping (NTD²) technique [72] consists in irradiating the original semiconductor wafer with a controlled flux of neutrons which, through nuclear reactions, produces the needed dopants with a uniform distribution in the semiconductor volume. In Ge semiconductors the energy gap between the valence and conduction band ($E_{gap} \sim 0.7 \text{ eV}$) [73] is well above the typical thermal energy already at room temperature $k_BT \sim 25 \text{ meV}$, and it stays above it at lower temperature as well. The number of available conduction electrons is very low and the crystal behaves as an insulator. Adding impurities in the semiconductor lattice lowers the gap between the valence and conduction bands in specific points of the crystal. If the concentration of dopant is high enough, the wavefunction of the charge carrier associated with each impurity overlaps and the behavior of the material changes abruptly from an insulator to a conductor because an impurity band in between the valence and conduction band in specific points of the conduction band becomes available. This is called Metal-Insulator Transition (MIT).

NTD Ge thermistors are devices with a steep dependence of the resistance as a function of temperature. This behaviour can be achieved by doping a Ge wafer just below the concentration needed to achieve a MIT transition at the desired operating temperature. In this way a small temperature variation can significantly modify the number of charge carriers able to tunnel from one impurity to the other and have the macroscopic effect of

²In the following the NTD acronym will be used to refer both to the Neutron Trasmutation Doping technique and to any of the NTD Ge thermistors used in CUORE.

making the resistivity of the thermistor strongly temperature dependent, as desired. This mechanism is called "hopping", and since the tunneling probability depends exponentially on the distance between the neighboring dopant sites, the homogeneity of the dopant distribution is of fundamental importance. The analytical dependence of the NTD resistance as a function of temperature is described by the following expression

$$R = R_0 \cdot \exp\left(\frac{T_0}{T}\right)^{\gamma} \tag{2.5}$$

where R_0 and T_0 can be tuned adjusting the dopant concentration and the NTD geometry while $\gamma = 1/2$. Typical values for the CUORE NTDs are $R_0 \sim 1 \Omega$ and $T_0 = 6$ K.

In order to operate NTDs as temperature sensors it is necessary to measure their resistance. This can be done biasing them with a constant current I_b and measuring the voltage difference at their edges. In order to do so in CUORE, gold wires are linked to the gold pads of each thermistor and the signal is extracted from the 10 mK volume of the cryostat to room temperature where it is low-pass filtered, amplified, digitized and stored on disk. More details on the CUORE signal processing are available in Sec. 3.

2.2.3 Detector operation

As anticipated in the previous section, the operation of an NTD as a thermometer coupled to a cryogenic bolometer can not disregard the need to bias the NTD and to read out the voltage at its ends. In CUORE each NTD is connected with golden wires to a wiring tray. This is another path for the heat to flow from the crystal absorber to the heat sink. In addition each CUORE bolometer is equipped with a silicon heater, with a similar wiring, in order to allow the periodic injection of well known amounts of power in order to measure the thermal gain of the detector and correct for small temperature drifts of each channel individually. A small temperature change in a crystal, in fact, affects its heat capacity and the amplitude of the thermal pulse induced by a given amount of injected energy varies as a function of this parameter (see Eq. 2.8). The effective thermal conductance between the CUORE bolometers and the heat sink can be evaluated from the decay time of the temperature pulses. An approximate value for the effective heat conductance is $G \sim 2 \text{ nW/K}$.

Each thermistor is biased with the scheme reported in Fig. 2.3. It is connected in series with a DC power supply and a load resistor $R_L \gg R_{bol}$ so that the current I_{bol} that flows in the circuit is approximately steady. The voltage at the ends of each NTD is extracted, filtered, amplified and digitized as further described in Sec. 3.

The bias current in the NTD injects power in the system via Joule heating. The injected power is proportional to the NTD resistence and, in a stationary situation, flows completely through the conductance G to the heat sink. This implies that the equilibrium



FIG. 2.3: Bias circuit and readout scheme of the NTD thermistors in CUORE (left). Load curve of a semiconductor thermistor. The working point is defined by the intersection of the V-I characteristic of the NTD and the load line $V_{bol} = V_B - I_{bol} R_L$ (right).

temperature of the crystal-NTD system is different from the one of the heat bath,

$$T_{bol} = T_{base} + \frac{I_{bol}^2 R_{bol}(T_{bol})}{G}$$

$$\tag{2.6}$$

where T_{base} is the temperature of the heat bath. This phenomenon causes a non-ohmic behavior and is referred to as "electrothermal feedback". In order to model the response of such device it is convenient to define a static resistance $R_{bol}^{(s)}$ and a dynamic resistance $R_{bol}^{(d)}(V_{bol})$ where

$$R_{bol}^{(s)} \doteq \frac{V_{bol}}{I_{bol}} \quad \text{and} \quad R_{bol}^{(d)}(V_{bol}) \doteq \frac{\mathrm{d}V_{bol}}{\mathrm{d}I_{bol}} \tag{2.7}$$

The dynamic resistance is the reciprocal of the derivative of the I-V load curve shown in Fig. 2.3. From this definition it can be shown that, at a given base temperature and for a low bias voltage, the dynamic resistance equals the static resistance. Increasing the bias voltage V_B the working point moves towards a region where $R^{(d)}$ vanishes (the inversion point) and eventually turns negative. As a consequence of Eq. 2.5 the amplitude ΔV_{bol} of the voltage pulse caused by an energy release E in the crystal absorber has the following expression

$$\Delta V_{bol} = \gamma \left(\frac{T_0}{T_{bol}}\right)^{\gamma+1} \frac{E}{CT_0} \sqrt{PR_{bol}}$$
(2.8)

where P is the power injected by the Joule heating of the NTD. This expression vanishes if $P \to 0$ and $P \to \infty$, then a maximum must exist. The optimal working point should be chosen in such a way to maximize the signal amplitude keeping the detector in a region where the I-V curve is approximately linear in order to achieve a pulse shape independent on energy.

2.3 The CUORE cryostat

The operation of a bolometric detector needs a cryogenic infrastructure able to reach and maintain for long periods of time the working temperature of ~ 10 mK. A suitable device is a dilution refrigerator (DR), which exploits the unique properties of the ${}^{3}\text{He}/{}^{4}\text{He}$ mixture. In order to reach such a low temperature the cooling is performed in different stages: the first cools the system to ~ 4 K with Pulse Tube cryocoolers. The ${}^{3}\text{He}/{}^{4}\text{He}$ mixture is first thermalized on this stage and then inserted in the following part of the dilution unit to reach the base temperature.

The CUORE cryostat [74] is made of six nested copper vessels. The experimental volume of about 1 m^3 is located inside the innermost vessel. In order to reduce the radiative heat flow to the coldest stage of the cryostat the vessels are thermalized to the decreasing temperatures³ of 300 K, 40 K, 4 K, 600 mK (the still), 50 mK (the heat exchanger), and 10 mK (the mixing chamber).

2.3.1 Mechanical structure, materials and shielding

The CUORE cryostat is installed underground in the Hall A of LNGS. Cryogenic bolometers are very sensitive to vibrations and a good isolation is mandatory in order to reduce the induced noise and ultimately achieve a good energy resolution. In addition vibrations can dissipate power by means of micro-frictions on any stage of the cryostat contributing to the total heat load that the cryogenic system has to remove. Since the available cooling power at low temperature is limited, vibrations need to be minimized because they could prevent the system from even reaching the base operating temperature. Finally the LNGS is located in a seismic area and for this reason a structure able to respond to seismic events without suffering damages was designed (Fig. 2.5).

The CUORE cryostat mechanical support structure is a two-stage system placed at room temperature devoted to isolate the system from low frequency vibrations. On the floor a concrete structure stands on top of 4 elastomers. On top of the concrete basement, four columns rise and hold a squared steel support called the Main Support Plate (MSP). Both the columns and the MSP are shown in red in Fig. 2.5. The MSP in turn, is the support structure of the Y-beam, a Y-shaped support from which the 988 detectors are hanged. The Y-beam is mechanically linked to the MSP with 3 Minus-K mechanical insulators placed at each end of the Y-beam. The Minus-K insulators are a particular arrangement of springs able to carry a consistent load of some tons, behaving like a soft spring when subject to small displacements. This is the reason why the natural frequencies of the spring-mass system are very low and the net effect is a low-pass filtering of the transferred vibration spectrum.

The Pulse Tubes operation is an important source of noise. As further detailed in Sec. 2.3.3 they provide cooling power at the 40 K and 4 K stage. In order to reduce the noise transferred to the cryostat, flexible connections were installed at all stages.

The volume inside the 300 K shield of the cryostat is divided in two vacuum chambers: the Inner Vacuum Chamber (IVC) and the Outer Vacuum Chamber (OVC). The IVC is the volume within the 4 K shield, while the OVC is the one between the 4 K and the 300 K

³Nominal temperatures. The actual temperature of each stage might differ from the nominal value.



FIG. 2.4: Scheme of the the CUORE cryostat. The plates corresponding to the different thermal stages, the lead shielding and the vacuum chambers are highlighted.



FIG. 2.5: Scheme of the CUORE cryostat support structure. The external shield sits on a movable platform that can be lifted to surround the cryostat during the detector operation.

shield.

In order to host a rare event search experiment, the materials used to build the cryostat were chosen taking into account the content of long-lived isotopes. The largest masses, i.e. plates and vessels, were made of a high purity copper produced by Aurubis with the exception of the 300 K plate, made of stainless steel. Inner layers such as the 40 K, the 4 K, Still and HEX plates and vessels were built in the highest available purity Oxygen-Free Electrolytic (OFE) copper. A material assay was performed and the OFE copper provided an upper limit of $< 65 \,\mu\text{Bq kg}^{-1}$ (²³²Th) and $< 54 \,\mu\text{Bq kg}^{-1}$ (²³⁸U). The 40 K and 4 K shields were also equipped with a superinsulation wrapping of double layer high-reflectivity aluminized mylar low thermal conductivity polyester.

The inner part of the cryostat, the 10 mK shield and plate, were made of another copper alloy with minimal hydrogen content (Cu NOSV). A different alloy for the innermost shield was selected in order to minimize the thermal load due to the ortho-para hydrogen transition [75] [76] and because of its high conductivity at low temperature.

The radiation shielding of the CUORE detector is made of a room temperature external shielding and a low temperature one. The external shielding is placed on a movable platform, and (even if it weights ~ 80 tons) can be lifted to surround the CUORE cryostat with a set of screwjacks. It has an octagonal shape and it is arranged in layers (Fig. 2.5). In order to efficiently shield neutrons and environmental gamma radiation an outer 20 cm thick layer of borated polyethilene is present while in the inner part a modern lead shielding of 25 cm was installed. In order to prevent the accumulation of radon in the volume between the outer shield of the cryostat and the inner wall of the shielding, this is constantly flushed with N_2 gas. The internal shielding is very close to the detector and is meant to further protect it from the surviving environmental radioactivity but mostly from the one provided by the cryostat and support structure itself. Stainless steel used for the suspension system and the 300 K plate as well as silver powder used in the dilution unit are just some examples of the sources of radioactive background the internal shielding is designed to absorb. The internal shielding was installed inside the IVC and is made of a top, lateral and bottom shield. The top shielding is positioned below the 10 mK plate, thermalized at the previous stage (HEX) and is of cylindrical shape, 30 cm thick commercial low radioactivity lead, sandwiched between two copper plates. The lateral and bottom shield are instead 6 cm thick and are made of ultra radio-pure archaeological roman lead. The lateral shielding begins below the Still plate, to which it is mechanically coupled, but it is thermalized to the 4 K stage.

Finally, in the innermost part of the cryostat, below the lower plate enclosing the top lead shield, the Tower Support Plate (TSP) is placed. The detector is suspended from the TSP, which is not anchored to the MSP directly but is held instead by the Y-beam. In this way the detector is further vibrationally decoupled from the rest of the cryostat thanks to the Minus-K springs that link the Y-beam to the MSP.

2.3.2 Dilution refrigerators

In this section a general overview of the basic working principles of dilution refrigerators will be given. A more detailed and complete description can be found in dedicated references [77] [78]. The idea underneath this method was first suggested in 1951 by H. London [79]. In 1962, together with Clarke and Mendoza [80] he published a more detailed proposal of dilution refrigerator able to reach temperatures below 1 K. Four years later Das, De Bruyn Ouboter and Taconis reached a temperature of 220 mK with this technique, experimentally demonstrating for the first time the feasibility of the method [81].

The main feature of interest for building dilution refrigerators is the finite solubility of ³He into ⁴He at arbitrarily low temperature as shown in Fig. 2.6 and the negative heat of solution. Pure liquid ⁴He has a superfluid transition at ~ 2.18 K. Mixing ³He into liquid ⁴He has the effect of lowering the temperature of the superfluid phase transition of the mixture down to at most $T_{sep} = 870$ mK and ³He concentration $x_3 < 68\%$. Above that concentration, lowering the temperature of the mixture below T_{sep} triggers a separation in a ³He-rich phase floating on top of the heavier ⁴He-rich one effectively forbidding the realization of any state in the shaded area of Fig. 2.6. Nonetheless it is possible to have a diluted phase at arbitrarily low temperature if the ³He concentration is kept below $x_3 < 6.6\%$. This happens because ³He atoms obey the Fermi statistics and they are more strongly bound to ⁴He than among each other because of the lower zero-point energy of ⁴He due to its higher mass. Due to Fermi statistics, the kinetic energy of ³He atoms in the diluted phase grows with their number density and the effective binding energy vanishes



FIG. 2.6: Phase diagram of ${}^{3}\text{He}/{}^{4}\text{He}$ mixture at saturated vapour pressure. Figure adapted from [77].

at a concentration $x_3 \sim 6.6\%$. Another important property of ${}^{3}\text{He}/{}^{4}\text{He}$ mixtures is that at $\sim 800 \text{ mK}$ the vapor pressure of ${}^{3}\text{He}$ is much higher than the one of ${}^{4}\text{He}$. This implies that pumping the vapor of the helium mixture removes ${}^{3}\text{He}$ much more effectively than ${}^{4}\text{He}$.

The core of a dilution refrigerator is the mixing chamber, the coldest part of the ${}^{3}\text{He}/{}^{4}\text{He}$ circuit shown schematically at the bottom of Fig. 2.7. The whole apparatus made of the mixing chamber, the heat exchangers and the distiller (still) is located in a vacuum chamber which, in wet cryostats, is kept at a temperature of $\sim 1.5 K$ by a liquid ⁴He bath. This complex circuit is needed in order to reduce the heat load on the mixing chamber. Since at the operating temperature of DRs ³He has a finite solubility of $x_3 = 6.6$ % it is possible to achieve a phase separation in which the heavier saturated mixture of He lies on the bottom of the mixing chamber and a ³He rich phase floats on top of it. In order to force the ³He to dissolve from the pure to the dilute phase one needs to remove some ³He from the latter. This is done connecting just the dilute phase to the still and pumping from the surface of it. The temperature of the still is kept at 0.7 - 0.8 K in order to achieve a sufficient pumping efficiency. Due to its higher vapor pressure, even though the 3 He concentration in the still is < 1 %, it accounts for the majority of evaporated gas. The extracted ³He is then cleaned outside the cryostat and returned to the circuit. It is thermalized with a series of heat exchangers with the cold mixture flowing from the mixing chamber to the still and injected again in the ³He-rich phase.



FIG. 2.7: Scheme of ${}^{3}\text{He}/{}^{4}\text{He}$ mixture flow in a dilution refrigerator

2.3.3 Pulse Tube cryocoolers

The CUORE cryostat is a multi-stage cryogen-free dilution refrigerator. This means that, instead of having a liquid He bath, the warmer stages, at 40 K and 4 K, are kept at their nominal temperatures by the cold heads of pulse tube cryocoolers (PT). The main difference between PT and other closed cycle refrigerators is the absence of moving parts at low temperature. On top of the advantage of dry refrigerators that intrinsically do not need cryogen refills, this feature is especially suitable for the CUORE infrastructure because it prolongs the lifetime of the cryocooler, reducing the need for maintainance. This is a key requirement since CUORE is supposed to operate for 5 years. In addition, the vibrations on the low temperature stages are reduced and this contributes to a better resolution of the bolometer detectors.

Pulse tubes in general are composed of a compressor, providing high and low pressure ends, connected to a rotary valve that allows the generation of an oscillating pressure (pulse), a regenerator material that acts as heat reservoir and a thin-walled *tube* with heat exchangers at both ends. The CUORE cryostat is equipped with 5 custom-adapted PT415-RM PTs by Cryomech (see Fig. 2.8), able to provide a cooling power of 1.2 W at 4.2 K and 32 W at 45 K each.

Despite the careful design and the intrinsic low vibration trasfer of PTs, they are one of the most relevant sources of noise in the CUORE detectors. To address and reduce the impact of this problem, a control system for driving the PT rotary valves was developed. The system allows to maintain the relative phases of the generated pressure waves constant



FIG. 2.8: Scheme of a CUORE pulse tube.

and to scan different phase configurations in order to select the one in which the coherent superposition of the mechanical vibrations transmitted to the detectors yields the minimum amount of noise [82].

2.4 Calibration system

In Sec. 2.2 the basic principles that allow the conversion from an energy release in a bolometric detector to a voltage pulse were outlined. Before the final energy spectrum can be reconstructed other steps are needed and described in detail in the subsequent chapters, nevertheless just one of those is intrinsically linked to the hardware component of CUORE that was not yet discussed in this chapter, the calibration system.

In order to map the amplitude of the reconstructed pulses to energy, known peaks from radioactive gamma emissions are used. The CUORE detectors, as already described in Sec. 2.3.1, are shielded from the environmental radiation with both an internal extremely low radioactivity Roman lead shielding [83] and an external modern lead one. So in order to calibrate the energy scale, calibration sources can be either placed at room temperature and be designed with a higher activity in order to produce a high enough rate in the bolometers, or at low temperature with looser requirements on their activity and composition but with much more challenging cryogenic and mechanical constraints. Both the systems were developed, built and successfully operated. The low temperature calibration is referred to as Detector Calibration System (DCS) and the other one as External DCS (EDCS).

The DCS was designed to deploy 12 strings (thoriated tungsten commercial welding rods) through thin tubes from room temperature to the inner part of the cryostat, inside the lead shielding, within the detector volume. This methods allows to have uniform illumination of all the channels, unachievable with the EDCS. This calibration system was designed to allow the calibration string deployment without the need of warming up the detector, since its response is not necessarily unchanged after a warm-up. As a drawback, the maintenance of the DCS heavily affects the data taking, forcing to long interruptions in the event of a technical failure. Moreover, in absence of a backup solution, since the stability of the detector response is monitored with monthly calibrations, a technical failure that completely prevented the operation of the DCS not only would affect future data taking because of the needed maintenance, but would also imply a loss of all the data acquired between the last successful calibration and the technical failure.

For redundancy and as a backup solution, the EDCS was designed and installed at the beginning of 2018. It consists of 8 PVC pipes running around the 300 K shield of the CUORE cryostat. The string composition for the EDCS is different from the DCS one. In this case ²³²Th sources are alternated with ⁶⁰Co sources and the visible peaks that can be used to calibrate the detector are just four: 511 keV from positron annihilation, 1173keV and 1332keV from ⁶⁰Co and 2615keV from ²⁰⁸Tl. The activity of the sources was defined with dedicated Monte Carlo simulations simultaneously optimizing the efficiency due to pile-up⁴ rejection in the external most illuminated channels, and the event rate in the innermost channels, suffering from the self-shielding of the other crystals. The need for additional ⁶⁰Co in the EDCS comes from the effectiveness of the internal lead shielding in stopping the external gamma radiation at lower energy and the need to have a sufficiently high number of different calibration peaks in order to constrain a 2nd order polynomial calibration function.

The DCS was used to calibrate all the datasets acquired since the beginning of the CUORE data taking till August 2018. From the experience acquired during the first year of data taking, the operation of the DCS was found to be time consuming and significantly reducing the duty cycle of the experiment, very likely to disturb the cryogenic system stability and failure-prone, hence too risky. The deployment of EDCS strings takes about 10 minutes, while the operation of the DCS could require from few hours to weeks and could not anyway always ensure the uniform illumination it was designed for on all the channels due to partial failures of the system. For this reason from August till December 2018 test EDCS calibration sources with no 60 Co were used instead of the DCS. The final setup, currently in use, includes 60 Co and has been successfully and safely operated during all 2019.

 $^{^{4}}$ A pile-up event is defined as two distinguishable energy releases happening in the same bolometer less than 7 s apart, where 7 s is the time between the trigger and the end of the analysis window.

Chapter 3

Data acquisition and processing

In the previous chapter some details were given about the CUORE detector, the underlying technology, cryogenics and each of its components separately. A very rough summary might state that CUORE is a machine able to convert energy releases in its detectors into voltage pulses. The goal of this chapter is to provide an overview of the steps that allow to readout such voltage pulses, store them on disk, identify (trigger) energy releases (Sec. 3.1) and make the data available for further analysis (Sec. 3.3) while ensuring that its quality is not spoiled by unwanted behaviors of the experimental apparatus nor unwanted environmental conditions (Sec. 3.2).

3.1 Data acquisition

The analog voltage signal at the ends of each NTD resistor is brought out from the cryostat, processed by the readout chain and eventually stored in files on commercial hard drives before being transferred to 2 different computing clusters where it becomes available for processing. The signal readout chain is composed of three steps:

- amplification, done with a programmable gain differential amplifier in order to reduce the common mode noise
- filtering, performed with a 6-poles anti-aliasing Bessel filter
- digitization, through commercial boards with a sampling frequency of 1 kHz and a 18 bit resolution.

In order to monitor and correct for thermal gain drifts of each channel individually, the bolometers are equipped with a $100 \,\mathrm{k}\Omega$ Si resistor (heater) used to periodically inject known amounts of energy thanks to a ultra-stable pulse generator. In order not to rely just on the software identification of the heater pulses and avoid any possible mis-classification of a heater pulse as a particle pulse, the Si heaters of each column¹ are connected in parallel so

¹The CUORE detector array is arranged in 19 towers where 13 floors of 4 detectors each are piled one on top of the other. A tower is then made of 52 detectors arranged in 4 columns.

Parameter	Value
Number of channels	988
Crystal mass	750 [m g]
Responsivity	$\sim 100 [\mu V/MeV]$
NTD resistance	$\sim 100 [\text{G}\Omega]$
Pulse rise time $(10\% - 90\%)$	$100 \ [ms]$
Pulse decay time $(10\% - 90\%)$	$400 \ [ms]$
Amplifier gain	$20 - 10^4 [V/V]$
Sampling frequency	$1 [\mathrm{kHz}]$
Event window	10 [s]
Pre-trigger	3 [s]

TABLE 3.1: Summary of some parameters of the CUORE bolometers, sensors and DAQ system settings. Some of the reported values should be interpreted as estimates of the order of magnitude of the corresponding parameter in optimal / average working conditions.

that a voltage spike in any heater channel produces a simultaneous pulse in 13 detectors, easing the identification of such events. This wiring is particularly convenient also from the cryogenic point of view because it reduces the heat load on the mixing chamber with respect to a design with independent connections for each channel.

The amplification stage is embedded in the Front End (FE) electronics that also provide the bias to the NTDs. The signal coming from the FE boards is fed to custom made antialiasing low-pass filters. The Bessel filter boards provide an attenuation of 120dB/decade and have a programmable cutoff frequency that can be set to 15, 35, 100, 120 Hz. The digitization of the output of the Bessel boards is performed with the Apollo software² and commercial NI-PXI-6284 boards from National Instruments. The main specifications of the digitizer boards are the maximum sampling frequency of 500 kS/s, the resolution of 18 bit and the maximum input voltage range of ± 10.5 V. The digitization resolution is well below the RMS of the baseline fluctuation of the CUORE bolometers, and this can be obtained with the simple following considerations. The typical gain of the amplification step is 5000 V/V, while the response of the bolometer-NTD system is $\sim 100 \ \mu V/MeV$, so the voltage amplitude of the pulses at the digitization stage is approximately 500 μ V/keV while the ADC resolution corresponds to $\sim 160 \,\mathrm{eV}$, which is much smaller than the typical detector resolution of few keV. The digitizer boards are hosted in six NI-PXI chassis, each of which has a controller module PCI-PXI-8336 that is connected to the following hardware components of the data acquisition chain via an optical link. This guarantees the electrical decoupling of the readout system from the computers power supplies. To enforce syncronization of the data digitization among different digitizer boards, a single clock source is connected in daisy-chain to all the chassis and then delivered from each chassis to all its boards with a PXI bus.

The acquisition and storage on disk of the digitized data is performed by the Apollo software. It is beyond the scope of this manuscript to describe in detail the hardware and

 $^{^{2}}$ A custom <code>ROOT</code> based software

software components and challenges that allow to acquire the CUORE data [84], but an overview and the main features will be outlined. The data transfer from the digitizer boards to ROOT files [85] is performed on a chassis-by-chassis basis by a dedicated computer. The samples are transferred to shared memories and untriggered *continuous files* are created, meanwhile up to 4 software trigger algorithms are run in parallel on each channel and the corresponding flags are attached to the samples each trigger fired upon. An estimate of the amount of data processed online by CUORE is given by

$$N_d \cdot B \cdot f_s = 1024 \cdot 18 \,[\text{b}] \cdot 10^3 \,1/\text{s} \sim 18 \,\text{Mb/s}$$
 (3.1)

where N_d is the number of available digital channels, B is the bit depth of the digitizers, f_s is the sampling frequency. This data transfer rate is even below what modern 4G smartphones can achieve (roughly ~ 60 Mb/s) when used as streaming media players. Given the low data throughput it is possible to afford both a software trigger and the storage of all the digitized data on disk. This allows to re-run, in case it is needed, any triggering algorithm. Furthermore the online triggering allows the real-time monitoring of some quantities such as the trigger rate or the noise RMS as well as other indicators of the data quality.

A separate process running on a dedicated machine accesses all the shared memories associated to each chassis and builds another set of ROOT files where an instance of complex objects storing quantities such as the trigger type, timestamp and involved channel is stored for each fired trigger flag. Such complex objects mirror what is called *event*. For each event, in the following steps of the analysis, a waveform can be retrieved at run time from the continuous data files. Any event-related observable can be stored, during the following data processing steps, as an additional object associated with each event whereas all the other global pieces of information are stored in a database.

3.1.1 Trigger algorithms

Different trigger algorithms are run in parallel in order to distinguish different event categories such as signal events, noise events and pulser (heater) events. Each trigger corresponds, in the analysis, to a waveform of 10 s extracted from 3 s before the trigger (pre-trigger) to 7 s after (post-trigger). The noise trigger is just a random flag placed with a given probability per unit time. Noise events are used to define waveforms free from pulses, used to monitor the detector conditions and to build average noise power spectra on a channel basis employed at later stages of the data processing. The pulser trigger identifies the time when the Si heater is used to inject pulses in order to correct for small thermal drifts of the detectors. At the moment two signal trigger algorithms are available, with the aim of detecting particle signals.

Derivative Trigger (DT) It fires when the derivative of the acquired voltage stays above threshold for a given amount of time and is artificially inhibited after having fired,



FIG. 3.2: Overlayed distribution of trigger thresholds at 90 % efficiency for DT (purple) and OT (red) algorithms as a function of the reconstructed energy variable [88]. Analysis thresholds applied on the same data for the neutrino-less double beta decay search [89] [41] are shown as vertical lines.

for a fixed amount of dead time. It is completely defined by the following parameters:

- derivative threshold
- time interval (i.e. number of samples) used for derivative computation
- time-over-threshold
- dead time

DT provides a fast, easy to configure and effective (thresholds of the order of few tens of keV can be achieved) trigger. It is used for online monitoring of the detector and for the online data processing (see Sec. 3.2 and 3.3).

Optimum Trigger (OT) This is a more advanced algorithm, described in [86] in greater detail, based on a matched filtering technique. A sub-sampled version of the continuous data is processed with an optimum filter (OF) [87] and then a simple threshold trigger is applied. Running the OT is demanding in terms of configuration because it requires knowledge of the (average) detector response and noise in the frequency domain. It also has higher demands in terms of computing power. Since the CUORE DAQ system allows for offline retriggering of the data, the OT is not enabled as online triggering algorithm. On the other hand, since it allows to significantly lower the energy threshold (Fig. 3.2) reaching values comparable with the detector energy resolution, it was adopted as the default triggering algorithm for the last step of the (offline) data processing and physics analyses.



FIG. 3.3: Effect of a small earthquake in a town close to the detector location on the detector baseline (left) with time interval, highlighted in red, flagged as *bad* (i.e. to be rejected in early data processing steps). Some details (right) about the seismic event made publicly available on the Istituto Nazionale di Geofisica e Vulcanologia (INGV) website.

3.2 Data quality

In order to ensure a good quality of the acquired data, thanks to the online triggering algorithm, a set of parameters can be constantly monitored. The voltage at the ends of each NTD (baseline), the RMS of the baseline, the trigger rate are examples of such channel-based parameters. Some environmental events that interfere with a stable detector operation have a clear pattern on such parameters. An earthquake that causes seismic waves to propagate up to the detector location, may slightly increase the temperature of the whole apparatus with the energy transmitted vibrationally. Such an event can be detected as a sudden and time-coincident (within few seconds) rise of the temperature of many crystals. Such temperature rise corresponds to a rise in the baseline and is eventually dumped with a timescale much longer than the one of particle pulses. This is because the amount of heat that needs to be removed is large and is distributed in many components rather than concentrated just in a single crystal. A cross-check of the earthquake hypothesis is the simultaneous detection of a seismic event by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) as in Fig. 3.3. In these cases a bad interval is set in the affected time period so that all the events contained in the bad interval where the detector conditions are not optimal can be identified and rejected, properly accounting for the introduced down time. Other sources of detector instability are electrical problems such as jumps in the grounding or any source of (non-stationary) vibration close to the detector (construction works, loud speakers, ...).

Other parameters are constantly monitored in order to ensure a safe, stable and reproducible operation of the detector such as pressures in the helium mixture circulation system, the temperature of different plates of the cryostat, the phase of the rotary valves of the active pulse tubes, the temperature of the water cooling system are, again, just a subset. The trend, over time, of such parameters is also used to schedule ordinary maintenance of the system and, extraordinarily, to decide data taking interruption. In case



FIG. 3.4: Accumulated exposure

any of the monitored parameters exceeds a specific threshold, action is taken to restore normal operating conditions and, depending on the impact the event might have on the reconstructed data, a decision is taken whether to include the corresponding time period in the data sample used for further physics analyses or not.

3.3 Data processing

The atomic unit of the data processing is the run. A CUORE run typically lasts 24 hours and can be classified as

- *background* if it is acquired with the purpose of collecting data to perform any physics analysis, trying to ensure detector stability, the lowest possible radioactivity and noise
- calibration if it is recorded with stable detector conditions and low noise but with the purpose of evaluating, in the shortest possible amount of time, the response function of the detector to convert amplitude of digitized voltage pulses to energy (see Sec. 2.4)
- *test* if it is recorded during any other unstable condition such as during maintenance of the cryogenic infrastructure

During stable data taking, background runs are acquired continuously, one per day, for at least 1 month. Before starting background data taking, a calibration is performed (see Sec. 2.4). Calibration sources are deployed and calibration runs are acquired, again one per day, up to the moment when a sufficient amount of channels has collected enough statistics to compute the calibration coefficients as detailed in the following. A typical calibration lasts

4-7 days. Another calibration campaign is performed after one month of background data taking is completed. The initial and final calibration runs as well as all the intermediate background ones define a *dataset*. The final calibration is shared between two contiguous datasets, acting as final calibration for the previous one and as initial calibration for the latter.

The acquired data undergo two subsequent data processings with the DIANA software, a custom C++ and ROOT-based [85] framework. The first one happens within ~ 1 day from the moment when the data acquisition of the run has ended, uses the events identified by the online derivative trigger and is meant to provide a fast feedback in order to spot possible problems that were not detected with the real-time data quality checks described in Sec. 3.2. This first processing (online processing) is done run-by-run and does not include some steps such as the computation of coincident events and the pulse shape discrimination filters. After the online processing of a complete dataset, the data is retriggered offline with the optimum trigger algorithm and a reprocessing is performed. Since the processing steps included in the reprocessing include - but are not limited to - those of the online processing chain, in the following just the reprocessing steps will be described.

Bearing in mind the main goal of reconstructing a combined energy spectrum of the entire detector, the acquired digital waveform in the time domain for each bolometer channel ch is modelled as the superposition of a known detector response function $s_{ch}(t)$ and an unknown noise term $n_{ch}(t)$ as

$$v_{ch}(t) = A_{ch}(E)G_{ch}(T)s_{ch}(t) + n_{ch}(t)$$
(3.2)

where $A_{ch}(E)$ depends only on the energy released in the corresponding detector, and $G_{ch}(T)$ is the bolometer gain and depends on the operating temperature of the detector. Following [90], the steps that the data undergo are aimed at

- computing some basic parameters of each event, used in the following sequences (preprocessing, 3.3.1);
- measuring the amplitude $B_{ch} \doteq A_{ch}G_{ch}$ with the maximum possible signal-to-noise ratio in order to achieve the best energy resolution (amplitude evaluation, 3.3.2);
- building a temperature independent estimator of the response function A of the detector to any given energy release (thermal gain stabilization, 3.3.3);
- inverting the response function A_{ch} in order to extract an energy estimator (energy calibration, 3.3.4);
- identifying coincident energy releases (coincidences, 3.3.5);
- identifying and rejecting unphysical events based on the pulse shape (pulse shape analysis, 3.3.6);

• hiding features of the reconstructed energy spectrum around the Q-value of ¹³⁰Te $\beta\beta$ decay in order to avoid possible bias in the event selection procedure (blinding, 3.3.7).

3.3.1 Preprocessing

In this preliminary part of the processing basic estimators of the noise level, signal amplitude and pile-up identification parameters are computed. A pile-up event happens when 2 energy releases are recorded in the same bolometric detector *close* in time. This definition needs further specification because the time distance threshold is not well defined. The purpose of classifying signal events either as pile-ups or not is to make sure the energy estimator that will be computed in the following processing steps is free from a bias coming from the fact that the considered pulse builds up on the tail of the previous one. All the following parameters are computed from the digitized voltage samples within a waveform, a time window of 10 s built around a trigger flag from 3 s before (pre-trigger) to 7 s after (post-trigger). The first 3 parameters are computed from the first 2.25 s of the pre-trigger.

• Baseline

average of the digitized voltage samples

- BaselineSlope slope coefficient extracted from a least squares linear fit of the digitized waveform
- BaselineRMS

standard deviation of the voltage samples from the linear fit function

• SingleTrigger

boolean flag, true when exactly 1 signal trigger flag is present in the waveform, false otherwise

• NumberOfPulses

number of detected peaks in the waveform.

3.3.2 Amplitude evaluation

In order to evaluate the amplitude $B_{ch}^{(i)}$ of each triggered pulse the optimum filtering (OF) technique [87] is used. It acts on the frequency domain and relies on the assumption that the noise is stationary and that the functional form of the detector response s_{ch} does not depend on energy, i.e. there is no sizeable shift in the working point of the detector across a whole dataset nor any other change in the working conditions that could induce a modification in the pulse shape. This is part of the data quality checks and is monitored with weekly NTD resistance measurements. Leveraging the entire waveform, rather than just few data samples around the peak of the pulse, each filtered waveform can be expressed

in terms of the original waveform as

$$V_j^{OF}(\omega) \propto e^{i\omega t_{max}} \frac{S_{ch}^*(\omega)}{|N_{ch}(\omega)|^2} V_j(\omega)$$
(3.3)

where V_j^{OF} is the output of the OF for the *j*-th triggered waveform in channel ch, $V_j(\omega)$ is the Fourier transform of the input waveform, $S_{ch}(\omega)$ and $|N_{ch}(\omega)|^2$ are the Fourier transform of the response function for the considered channel and the noise power spectral density respectively, ω is the angular frequency and t_{max} is the time of the pulse maximum.

Each channel has a different response function, for this reason it is computed, for each channel individually, from calibration data. Calibration runs are very effective for this task because the average calibration trigger rate is ~ 20 - 100 mHz while during background data taking it hardly reaches 6 mHz. A clean sample of pulses is selected requiring that just one trigger flag is present in the selected window (SingleTrigger), the BaselineSlope in the pre-trigger is compatible with 0 and the difference between the maximum and the minimum of the triggered waveform is greater than a fixed multiple of the BaselineRMS. Averaging Eq. 3.2 over a sufficiently large amount of waveforms cancels the noise term because its mean value is null and yields $s_{ch}(t)$ up to a multiplicative scaling factor. This object, after a proper normalization, will be called *average pulse* (AP).

Noise samples are instead acquired with a dedicated trigger algorithm and are selected according to the **BaselineSlope** in order not to include in the noise sample any component coming from the tail of signal pulses. Each noise waveform is Fourier transformed, the square modulus is computed for each frequency and an average over all noise events is performed. The result defines the *average noise power spectrum* (ANPS).

3.3.3 Thermal gain stabilization

Small variations of the base temperature induce variations in the bolometer gain which, if they were not taken into account and corrected, would cause a significant worsening of the energy resolution. Since the baseline dependence on temperature can be assumed linear for small temperature variations, the correction of this effect is just a correction of the dependence of the amplitude of the pulse reconstructed by the OF, namely OFAmplitude, as a function of the Baseline parameter,

$$\mathsf{OFAmplitude} = A(E) \left[p_0 + p_1 \cdot \mathsf{Baseline} \right] \tag{3.4}$$

The correction is implemented in two steps. First a linear regression of the amplitude dependence on the baseline value is performed on a sample of triggered heater events with the same injected energy (stab-pulser events) in order to compute the p_0 and p_1 parameters for each channel. Then a new stabilized amplitude variable is defined and attached to all events according to the following definition,

StabAmplitude
$$\doteq A_0 \cdot \frac{\text{OFAmplitude}}{\left[p_0 + p_1 \cdot \text{Baseline}\right]}$$
 (3.5)

Energy [keV]	Source
511	e^+ annihilation
583	208 Tl
911	^{228}Ac
969	^{228}Ac
1173	$^{60}\mathrm{Co}$
1332	$^{60}\mathrm{Co}$
2615	208 Tl

TABLE 3.5: Summary table of the most prominent available calibration peaks. The 583 keV, 911 keV and 969 keV peaks are visible with sufficient strength just with internal DCS strings while 60 Co ones are available just with the external calibration system because its strings are the only ones containing the source.

where A_0 is an appropriate normalization constant that sets the average reconstructed stabilized amplitude for stab-pulser events to 5000. This stabilization procedure is called *heater-TGS*.

The above mentioned procedure cannot be performed on channels with malfunctioning heater. There exists a different procedure to perform the reconstruction and correction of the **OFAmplitude** dependence on **Baseline** across a whole dataset, based on 208 Tl decays recorded during calibration runs. One additional challenge of this procedure are thermal fluctuations across both calibrations, which are larger than the ones that can happen on the timescale of a single run. For this reason the dependence on the baseline (and, ultimately on temperature) is modeled with a second order polynomial. This latter stabilization procedure will be referred to as *calibration-TGS* and is not performed during the online processing to allow for a faster feedback. The channels that fail either of the procedures are taken track of in the database and consistently undergo different following processing steps.

3.3.4 Energy calibration

The reconstruction of the energy release corresponding to each triggered event requires modeling and inversion of the A(E) relation. This is achieved assuming that the StabAmplitude response of the detector to a monochromatic energy release (see Tab. 3.5) is well described by a CrystalBall probability density function³ multiplied by an appropriate normalization factor. A simpler Gaussian model was tested as well, but it was found not adequate to model the distribution of counts around the considered calibration lines, whereas a CrystalBall function provides a better description. Assuming the following dependence of the original energy released on the mean value of the reconstructed StabAmplitude distribution,

$$Energy = a_1 \cdot StabAmplitude + a_2 \cdot (StabAmplitude)^2$$
(3.6)

³The CrystalBall function is defined by a gaussian core smoothly connected to a power-law left tail.

a CrystalBall fit is performed on each channel-peak pair on all signal events from the initial calibration runs that pass some basic pile-up rejection cuts. Then a least squares quadratic regression allows the extraction of the calibration coefficients a_i for each channel. Such coefficients are used to assign an **Energy** value to each event according to Eq. 3.6. During the online processing the final calibration is first processed separately, with the calibration coefficients obtained from the initial calibration only. Then the reconstructed spectra of both calibration campaigns are superimposed in order to check for consistency. The channels that are found compatible with each other in terms of the reconstructed energy of the most prominent background peaks for both calibration coefficients are extracted again, from both the calibration campaigns processed together, while the others are assigned a flag in the database and eventually excluded. The online processing ends with just one more step (blinding) after energy calibration, described in Sec. 3.3.7.

3.3.5 Coincidences

Up to this step every channel is processed in parallel and independently of all the others. Coincidence analysis, instead, allows a very powerful background rejection as well as a deeper understanding of the background contaminations, combining and relating events that belong to different channels. In order to proceed further it is important to provide some definitions:

- a collection of coincident events defines a *multiplet*;
- Multiplicity, is the number of events within a multiplet;
- the *main event* of a multiplet is one particular event within the multiplet;
- all the other events are called *coincident events* with respect to the main;
- the TotalEnergy of a multiplet is the sum of the Energy releases of the main event and all coincident events.

A good coincidence algorithm must be symmetric. In other words it must ensure that, if A and B are events, the proposition A is coincident with B is necessary and sufficient for B to be coincident with A. In addition, the coincidence relation must be transitive⁴. Many coincidence algorithms were developed within the CUORE analysis framework. Each algorithm defines the conditions that allow to identify the coincidence relation between events.

Unbiased event time estimator The first relation that coincident events must satisfy is time coincidence, namely being reconstructed within a given amount of time. An estimator of the event time can be obtained, channel by channel, as the time corresponding to the

 $^{^{4}}A$ coincident with B and B coincident with C implies A coincident with C.

maximum of the OF filtered pulse: OFTime. Due to the different intrinsic pulse shapes of the channels, simultaneous events in different channels can exhibit a time delay up to several tens of ms. We will refer to such delay as *jitter*. Under the approximation that the pulse shape does not depend on energy, the jitter depends just on the difference between the response of different channels. Being so, it can be computed for each dataset and subtracted so that the distribution of the time difference between simultaneous energy releases in different channels a and b, computed as

$$\Delta t_{a,b} = (\mathsf{OFTime})_b + \mathcal{J}_b - (\mathsf{OFTime})_a - \mathcal{J}_a \tag{3.7}$$

has 0 mean, where \mathcal{J}_{ch} is the jitter of channel ch with respect to one reference channel, common for all the detector array.

One method to compute the *absolute* jitter of a channel \mathcal{J}_{ch} is to start from the relative jitter among channel pairs that share a sample of (physically) coincident events⁵. From the relative jitter it is possible to reconstruct the absolute one by summation of the contributions along the path that connects the channel to the reference one minimizing the uncertainty. The sample of physically coincident events used to compute relative jitters is extracted from coincidence multiplets built with a large time window in calibration data, selecting Multiplicity = 2 events that sum at the 208 Tl line at 2615 keV. In order to have enough events to perform the jitter computation the coincidence algorithm is run tower-by-tower, otherwise accidental coincidences, given the high rate of calibration and the large W = 100 ms time window, would impede the measurement. A reference channel is chosen for each tower, and the syncronization among different towers is done based on the average time difference between the trigger flag and the maximum of the filtered amplitude for pulser events in the reference channels. The effectiveness of this method is shown in Fig. 3.6, where the width of the time difference distribution after jitter syncronization is $\sim 10 ms$. For this reason the smallest time window chosen to perform coincidence analyses is ± 5 ms. Since some analyses might benefit from a wider time coincidence window, the data processing chain includes coincidence multiplets computed with both ± 5 ms and $\pm 30 \text{ ms.}$

Time coincidence window In order to define a symmetric coincidence relation between events it is not enough to set

$$(|\Delta t_{a,b}| \le W/2) \Leftrightarrow (a \nleftrightarrow b) \tag{3.8}$$

where the $\leftrightarrow stands$ for a symmetric coincidence relation between event a and event band W is a suitable coincidence time window. Eq. 3.8 holds as long as Multiplicity < 3. Let us assume e.g. events A, B, C with $t_A < t_B < t_C$, $\Delta t_{A,B} = \Delta t_{B,C} < W$ and

 $^{{}^{5}}$ By *physically coincident events* it is meant energy releases originating from a process that guarantees simultaneous (negligible time delay) energy releases in different bolometers, e.g. a photon interacting subsequently with more than one crystal.



FIG. 3.6: Time difference distribution for Multiplicity = 2 events with TotalEnergy compatible with the 2615 keV ²⁰⁸Tl line before jitter syncronization (coincidence window W = 100 ms, black) and after (W = 10 ms, red).

 $\Delta t_{A,C} > W$. Building coincidences according to the definition in Eq. 3.8 would associate a Multiplicity = 2 multiplet to event A and C, and a Multiplicity = 3 multiplet to event B. This breaks the transitive property of the coincidence relation, so (3.8) it is not a good definition. Without entering the implementation details, a good definition of time coincidence is

$$(|\Delta t_{a,b}| \le W/2) \Rightarrow (a \nleftrightarrow b) \tag{3.9}$$

so that, referring to the example above, three Multiplicity = 3 multiplets would be associated to events A, B, C as

$$A \nleftrightarrow \{B, C\}, B \nleftrightarrow \{A, C\}, C \nleftrightarrow \{A, B\}$$

$$(3.10)$$

where A, B, C are the main events and the ones in curly brackets represent the associated coincident events for each multiplet. This means that the coincidence algorithm can be thought as a recursive one, where a W wide time window is symmetrically opened around the main event. All events that fall in the coincidence window are included as coincident events. Then the coincidence window is enlarged accordingly, so that new coincident events are included in the multiplet \mathcal{M} if their event time t_x satisfies

$$\min_{i \in \mathcal{M}} \left[(\mathsf{OFTime})_i + \mathcal{J}_i \right] - W/2 \le t_x \le \max_{i \in \mathcal{M}} \left[(\mathsf{OFTime})_i + \mathcal{J}_i \right] + W/2 \tag{3.11}$$

and then this procedure is iterated until no more events can be added to the multiplet.

The problem of finding the optimal coincidence window W is not related just to the

probability of correctly including genuine coincident events in the same multiplet, otherwise the larger the coincidence window the better. The rate of accidental coincidences is also an important parameter, because it directly influences the overall detection efficiency of any process. Since the CUORE detector array is made of 988 channels, the probability of mis-reconstructing the multiplicity of a coincident energy release can be approximated as

$$\mathcal{P}(n=0|\mu=R\cdot W)^{988-1} = e^{-987\cdot R\cdot W}$$
(3.12)

where $P(n|\mu)$ is the Poissonian probability of observing *n* events expecting μ , *R* is the average trigger rate and *W* is the time coincidence window. With an average rate of 6 mHz the efficiency term in Eq. 3.12 varies from 94% to 41% for coincidence windows of ± 5 ms and ± 75 ms respectively.

Radial cut A simple way to reduce the reconstruction inefficiency due to accidental coincidences is to leverage the granularity of the CUORE detector. Being a closely packed array of high density crystals, it is very unlikely that two simultaneous events that happen very far apart are causally related. For this reason an additional requirement in the coincidence algorithm can be the spatial distance between involved detectors,

$$\left(\Delta t_{a,b} \le W\right) \& \left(R_{a,b} \le R_{max}\right) \Rightarrow \left(a \nleftrightarrow b\right) \tag{3.13}$$

where $R_{a,b}$ is the distance between the centers of the corresponding crystals. The distance between any pair of CUORE crystal centers ranges from 80 - 1200 mm. A typical radial cut is 150 mm.

3.3.6 Pulse shape analysis

The purpose of this processing step is to build a filter that discriminates single particle pulses from other triggered events such as pile-up events, spikes or fast baseline fluctuations. It does so combining the following six pulse shape parameters:

- RiseTime (and DecayTime) time difference between the samples where 10% and 90% of the maximum of the pulse amplitude are reached, computed on the rising (falling) edge of the pulse,
- BaselineSlope slope of a first order polynomial fit done on the first 3/4 of the pre-trigger samples,
- OFDelay time difference between the beginning of the window and the maximum of the filtered pulse,
- TVL (and TVR) sum of the squares of the differences between the filtered pulse and filtered pulse template, properly normalized, before (after) the maximum. The higher these values, the further the shape of the pulse with respect to the average response of the detector.

A small energy dependence of the pulse shape parameters as a function of energy is observed as well as a widening of their distribution at low energy. From an appropriate clean sample of M2 (short for Multiplicity = 2) events, the energy dependence of the median and median absolute deviation (MAD) of each parameter as a function of energy is extracted with phenomenological functions and normalized pulse shape parameters are defined as

$$P_i^{norm} \doteq \frac{p_i - f_{median}^{(i,ch,ds)}(E)}{f_{MAD}^{(i,ch,ds)}(E)}$$
(3.14)

where f_{median} and f_{MAD} are the phenomenological best fit functions for each parameter i, channel ch and dataset ds. It might be useful to remind the reader that the CUORE bolometers have the same response to α and β/γ particles, because nearly⁶ all the released energy in the crystal is dissipated in the phonon channel. For this reason it is possible to define just a single band for each pulse shape parameter, regardless of the source. Finding appropriate fit functions on a wide energy range for all parameters is challenging and cannot be completely automated.

The final step takes into account the correlations among the parameters and computing the Mahalanobis distance [91] M_{dist} of each event from the average value of the six normalized pulse shape parameters computed for a reference data sample. A cut on the Mahalanobis distance is the final step of the pulse shape analysis, that eventually defines, on an event basis, the ones that are allowed to enter the final data sample that can be used for physics analyses. For neutino-less double beta decay search in the gs-gs transition⁷ the probability of the energy of the two electrons to be fully contained in a single crystal is ~ 88%. For this reason the analysis can be safely performed just on the M1 event sample using coincidences just as an anti-coincidence veto system to reduce the background from e.g. surface α contaminations. Nonetheless it is important to produce a set of reliable multiplets for any multiplicity. For this reason a consistency check is performed after the pulse shape analysis is done, flagging as invalid any multiplet where either the main event or any of the coincident events was rejected by the cut on M_{dist} .

3.3.7 Blinding

The final processing step, right before data reduction⁸, is blinding. Different blinding algorithms are available depending on the analysis, but since the $0\nu\beta\beta$ decay search is the main scientific goal of CUORE, the one described in the following is the only applied at the processing level. It consists in a data salting procedure that shifts a random fraction (around 40%) of the events that reconstruct in a ±10 keV energy window around the 2615keV peak from ²⁰⁸Tl decay down by 87keV so that they mimic a peak at the $Q_{\beta\beta}$ value

⁶A small amount of energy can propagate as Cerenkov radiation and escape the crystal, but it is a negligible fraction.

⁷Transition from the ground state of ¹³⁰Te to ground state of ¹³⁰Xe.

⁸Since most of the parameters computed along the processing are not needed by higher level analyses, the data is copied in light **ROOT** files containing ntuples with the smallest possible number of variables.

of the ¹³⁰Te double beta decay. The same fraction of events is moved in the other direction, from around $Q_{\beta\beta}$ to the thallium peak position. The original value of the reconstructed energy is encrypted and stored. Until all the details of the $0\nu\beta\beta$ decay search analysis are finalized the data is kept *blinded*, so that no bias can be introduced optimizing selections or any other fitting procedure on the data around $Q_{\beta\beta}$.

Chapter 4

Detector characterization

In this chapter an overview on some of the parameters that allow to model the detector response will be given: energy threshold (Sec. 4.1), efficiency (Sec. 4.2), energy scale uncertainty and resolution (Sec. 4.3). Some details on the definition of each of the above mentioned effects, the underlying assumptions of their model and the methods employed to derive such descriptions are shown.

4.1 Energy threshold

The meaning of *energy threshold* strongly depends on the context. In general it means the minimum energy above which, with more than a fixed probability (efficiency), events are detected and get properly reconstructed. Depending on the selected data sample the energy threshold can change because different selection cuts might be applied. In the following section all the selections and processes in general that result in an energy threshold will be explored. In order of execution in the data acquisition and processing they are

- trigger algorithm
- coincidences computation
- pulse shape analysis

4.1.1 Trigger threshold

The trigger threshold is defined as the minimum energy E_{th}^{trigger} above which signal events (i.e. energy releases in a bolometer) get properly flagged as signal events by the trigger algorithm with an efficiency $\epsilon_{th}^{\text{trigger}} \geq 90\%$. In order to evaluate the trigger thresholds, a dedicated test run (*NPulser* run) is acquired for each dataset, where the pulser is fired with N different amplitudes that are expected to reconstruct at different energies depending on the channel, spanning the whole dynamic range up to ~ 3 MeV. For each amplitude the pulser is fired N_{fired} number of times. Then the standard online processing



FIG. 4.1: Derivative trigger threshold evaluation for two different channels. The intersection of the trigger efficiency curve (red line) with the target efficiency of 90% (grey dashed line) defines the trigger threshold (green solid line) for each channel.

is performed in order to assign an energy to each reconstructed pulser event and efficiency points are extracted as a function of energy as

$$\epsilon^{trigger}(E_e) = \frac{N_{sel}^{(e)}}{N_{fired}} \tag{4.1}$$

where $N_{sel}^{(e)}$ and N_{fired} are the triggered and fired number of pulses for each channel, with injected amplitude indexed by the e = 1...N superscript. E_e is the mean of the reconstructed energy distribution for the *e*-th pulser amplitude. Then, for each efficiency point, an uncertainty is assigned with the following rule:

$$\sigma(\epsilon) = \sqrt{\frac{\epsilon \cdot (1-\epsilon)}{N_{fired}}} \quad \text{if} \quad \epsilon \neq 0, 1 \tag{4.2}$$

$$\sigma(\epsilon) = \sqrt{\frac{1/N_{fired} \cdot (1 - 1/N_{fired})}{N_{fired}}} \quad \text{if} \quad \epsilon = 0, 1 \tag{4.3}$$

in order to be able to assign a non vanishing uncertainty to the efficiency term for each $0 \le \epsilon \le 1$ [92]. Finally the dependence of the trigger efficiency on energy is derived with a least squares fit to the expression

$$\epsilon^{trigger}(E) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{E - a_0 \cdot a_1}{a_1}\right)$$
(4.4)

hence the trigger energy threshold, as defined above, is

$$E_{th}^{trigger} \sim a_0 \cdot (0.9 + a_1) \tag{4.5}$$

4.1.2 Coincidences threshold

Depending on the purpose of the analysis it might be more important to have a uniform response of the detector or to exploit as much as possible the knowledge coming from low-threshold¹ channels. It is possible to set a software threshold, at processing level, on the candidate events used to build multiplets, E_{th}^{coin} . Events with energy above threshold do enter coincidence computation, the others do not. Two extreme choices are possible, for the sake of:

- uniformity of the detector response, to choose a coincidence threshold such that most (if not all) channels have a trigger threshold below the coincidence threshold $E_{th}^{coin} > E_{th}^{trigger}$
- detecting as much low energy coincident processes (such as nuclear recoils) as possible, to choose a coincidence threshold lower than any trigger thresholds $E_{th}^{coin} < E_{th}^{trigger}$

Choosing the lowest possible coincidence threshold has two important consequences:

- a channel dependent efficiency must be implemented to correctly interpret the data;
- not only the accidental coincidence rate is higher (the trigger rate increases at low energy, see Eq. 3.12) but it is sufficient that the noise level of one channel changes over time enough to allow some noise events (fast baseline fluctuations) to be triggered, to make the accidental coincidence probability time dependent in a way that is difficult to detect and properly model.

In addition, the time coincidence relation of two events relies on the (jitter-corrected) OFTime difference according to Eq. 3.7. Let us define OFDelay as the difference between the reconstructed event time and the beginning of the corresponding waveform. The distribution of OFDelay shows a small energy dependence (low-energy events tend to have lower OFDelay) and its dispersion grows at low energy. Hence, in order to correctly identify coincident events where a high energy release (e.g. alpha decay) happens in coincidence with a low energy one (e.g. nuclear recoil), it is necessary not only to lower the coincidence energy threshold but also to enlarge the time coincidence window.

4.1.3 Pulse Shape Analysis threshold

As already pointed out in Sec. 3.3 the six parameters used to characterize the pulse shape of events have an energy dependence that must be corrected. The description of such dependence is poor at low energy and this is the main reason why the PSA efficiency reaches its plateau at a few hundred keV. Details about the methods that allow to model such efficiency are outlined in the following Sec. 4.2. For the purpose of defining the energy threshold due to PSA it is enough to mention that the efficiency as a function of energy is

¹Trigger threshold

well described by the curve in Fig. 4.2,

$$\epsilon_{PSA}(E) = \theta(E - p_0) \cdot p_0 \left[1 - e^{-(E - p_1)/p_2} \right]$$
(4.6)

where $\theta(x)$ is the Heaviside step function, and p_i are parameters estimated from the data. With this definition is possible to define the PSA energy threshold as the energy at which the efficiency reaches q = 90% of its maximum value

$$E_{th}^{PSA} = p_1 - p_2 \ln \left(1 - q\right) \tag{4.7}$$

4.2 Efficiency

The efficiency evaluation is a complex task because it not only depends on choices done at the processing level, such as the coincidence window size or the threshold set on the M_{dist} variable for the pulse shape analysis, but primarily depends on the process of interest. The aim of this section is to disentangle the different reasons why an efficiency term might be introduced and provide the methods to evaluate its contribution.

4.2.1 Containment efficiency

In order to select candidate events of some process (e.g. neutrino-less double beta decay) a region of interest (ROI) must be identified. This means selecting (at least) one multiplicity, setting energy selection cuts and fixing some requirements on the candidate signal events (e.g. data quality, pile-up rejection). The containment efficiency is defined as the probability of the process of interest to release energy in the selected ROI, regardless of how such events perform with respect to the selection cuts on variables other than **Energy** and **Multiplicity**. This is the only efficiency contribution that must be evaluated with Monte Carlo simulations. The process of interest (signal) is simulated N_{MC} times independently and the number of events that reconstruct in the ROI (N_{ROI}) are evaluated. From the ratio

$$\epsilon_{cont} \doteq \frac{N_{\rm ROI}}{N_{MC}} \tag{4.8}$$

the containment efficiency can be evaluated. In the case of neutrinoless double beta decay search on the ground state the containment efficiency is $\epsilon_{cont}^{0\nu} \sim 88\%$, while for the search of double beta decay with neutrino emission on the first 0⁺ excited state of ¹³⁰Xe in 2-site coincidences (M2 events) the containment efficiency can be as low as $\epsilon_{cont}^{2\nu(0_1^+)} \sim 25\%$.

4.2.2 Detection efficiency

Under this category falls the probability that a signal energy release with energy E is reconstructed (triggered) without being discarded by the pile-up rejection cuts². In

 $^{^{2}}$ The PSA cut, which also contributes to pile-up rejection, is not included in the definition of the detection efficiency.

the following such a selection will be referred to as *base cuts*. Besides the trigger efficiency contribution, already described in Sec. 4.1.1, a data driven detection efficiency computation can be performed. Based on pulser events it is possible to know a priori the number of injected signal events. Given that, within a certain dataset, the heater fired N_0 times, the detection efficiency can be expressed as

$$\epsilon_{base} \doteq \left[\frac{N_S}{N_0}\right] \left[\frac{N_{3\sigma}}{0.997 \cdot N_{10\sigma}}\right] \left[\frac{N'_S}{N_S} \cdot \left(1 - f_{\text{pulser}} T_w\right)\right]$$
(4.9)

where N_S is the number of triggered events among the N_0 original ones, and measure the trigger efficiency³ at high energy (well above threshold). $N_{n\sigma}$ is the number of pulser events within n standard deviations from the mean of the triggered pulser events energy distribution. At the denominator a correction factor equal to the integral of a gaussian distribution from -3σ to $+3\sigma$ is introduced. Finally N'_S is the number of pulser events that are not rejected because of pile-up (i.e. pass the base cuts), f_{pulser} is the pulser firing frequency and $T_w = 10$ s is the size of the time window used for the data processing. The last term accounts for the probability of a pile-up on a signal event. A correction dependent on the frequency of pulser events is necessary since this efficiency contribution is computed from pulser events themselves⁴. Overall, the detection efficiency is then parameterized as the product of the three factors

- trigger efficiency
- energy reconstruction efficiency
- pile-up rejection efficiency

of Eq. 4.9, respectively.

4.2.3 Accidental coincidence efficiency

Every time a coincident energy release in different bolometers happens, there is a small probability of an uncorrelated trigger in some other channel. The probability that such coincident energy releases get properly reconstructed and flagged with the correct multiplicity is called accidental coincidence efficiency. It can be computed from data for each dataset, combining the information from the entire detector. ⁴⁰K is very well suited for this study because it has an electron capture (EC) decay channel to ⁴⁰Ar with a neutrino and a 1460 keV γ emission. The photons due to ⁴⁰K decay approximate very well single-site events and for this reason they are the optimal benchmark to study the anti-coincidence efficiency. A binned maximum likelihood fit with polynomial background subtraction of

³The statistics of pulser events is much higher in this case with respect to the sample used to evaluate the trigger threshold. This is the reason why in this case the trigger efficiency at high energy is not assumed equal to unity.

⁴The pulser frequency is designed in such a way to prevent pile-ups of pulser events on other pulser events. Rather, pulser events can pile-up to signal events, and this is the reason for the correction.



FIG. 4.2: Pulse shape analysis efficiency curve (red, Eq. 4.6) overlaid on a binned histogram (blue) of Multiplicity = 2 signal events. Point evaluations of the same quantity on Multiplicity = 1 γ lines (green) are also shown.

the energy released at the $^{40}{\rm K}$ γ line in different event samples allows to determine the efficiency as

$$\epsilon_{ac} = \frac{N_{\rm M1}^{1460}}{N^{1460}} \tag{4.10}$$

where $N_{\rm M1}^{1460}$ is the background subtracted estimate of signal events in Multiplicity = 1 events while N^{1460} is the same quantity computed for any value of Multiplicity. Both samples require the events to be triggered and to pass the base cuts. Since accidental coincidences are uncorrelated with the energy release they happen in coincidence with, this efficiency contribution does not depend on energy.

4.2.4 PSA efficiency

The pulse shape analysis efficiency is the probability of correctly classifying a signal event that has already passed the pile-up rejection cuts with a Mahalanobis distance below threshold

$$\epsilon_{\text{PSA}}(E) \doteq \mathcal{P}(E, M_{dist} < M_{dist}^{cut} | S) \tag{4.11}$$

where *signal event* in this context means any pulse induced by a particle releasing energy in a bolometer.

A very simple, yet powerful, way of getting from data the complete energy dependence of the PSA efficiency is to leverage Multiplicity = 2 events whose sum energy corresponds to the Q value of known α or γ decays. Since in the standard data processing the coincident multiplets are built before the PSA M_{dist} variable is even computed, it is possible to select a sample of Multiplicity = 2 events that pass the base cuts, where the main event is required to fulfill the $M_{dist} < M_{dist}^{cut}$ requirement while its coincident is not. Building such a sample before and after the PSA cut is applied, we can consider the histogram of ratios of the bin contents

$$\frac{N(M2)^{PSA+base}}{N(M2)^{base}}$$
(4.12)

as a function of the energy of the main event. A fit to the histogram of count ratios as defined in Eq. 4.12 with the expression given in Eq. 4.6 allows the extraction of the parameters that describe the PSA efficiency curve.

4.3 Energy scale reconstruction

Every detector designed to measure some observable x has a response function f(x) which ideally would be a Dirac delta function $f = \delta(x)$ but in practical situations is not so sharp, rather in first approximation it can be written as a gaussian distribution $f = e^{-(x^2/2\sigma^2)}/\sqrt{2\pi\sigma}$. Assuming the observable is randomly extracted from a spectrum g(x), the outcome of the measure will be a convolution of the original *true* spectrum with the detector response function

$$\tilde{g}(\tilde{x}) = \int g(x')f(x' - \tilde{x})dx'$$
(4.13)

so that the measured spectrum $\tilde{g}(\tilde{x})$ is obtained smearing the original one with the response function. In general the detector response can introduce a bias in the observable reconstruction, i.e. a monochromatic input at $x = x^*$ not necessarily will be reconstructed at $\tilde{x} = x^*$, not even on average. The difference between the expected response and the input

$$\Delta(x^*) \doteq \langle \tilde{g}(x) \rangle - x^* \tag{4.14}$$

defines the experimental reconstruction bias on the observable x. It is worth mentioning that the bias can be a function of the observable rather than just a constant. In addition, the shape of the response function might depend on the observable as well, rather than staying constant.

In the following sections the CUORE detector response function is defined, the techniques to extract from data its shape parameters and their dependence on energy will be detailed.

4.3.1 Line shape response function

The functional form of the response function of the CUORE bolometers to monochromatic energy releases is defined phenomenologically for each channel-dataset pair. It consists of the superposition of three gaussian distributions that share the same width but



FIG. 4.3: The best fit to the 2615 keV line from ²⁰⁸Tl in calibration data (black dots), summed over all channels (red solid), is shown. Other components such as the compton shoulder, X-ray escape peak and coincidence peak are also shown (blue solid).

have different amplitude and mean value slightly shifted one from the other as

$$f(E|\mu,\sigma,A_L,A_R,a_L,a_R) \doteq \frac{\mathcal{G}(E|\mu,\sigma) + A_L \,\mathcal{G}(E|a_L \cdot \mu,\sigma) + A_R \,\mathcal{G}(E|a_R \cdot \mu,\sigma)}{(1 + A_L + A_R)} \quad (4.15)$$

where $\mathcal{G}(E|\mu,\sigma)$ is a normalized Gaussian distribution with mean μ and variance σ^2 , A_L , $A_R < 1$ are amplitude scaling parameters that control the relative amplitude of the two gaussians with respect to the main peak, $a_L < 1$ and $a_R > 1$ instead tune the position of the mean of the sub-peaks. The shape of the distribution is defined, for each channel-dataset pair, by the six parameters $\mu, \sigma, A_{L,R}, a_{L,R}$.

To infer the values of such parameters a fit to the ²⁰⁸Tl line in calibration data is performed (Fig. 4.3). The data sample includes triggered events, no matter which multiplicity they are assigned, in order to benefit as much as possible from the high event rate in calibration data. Additional components are required to properly model the spectrum around the 2615 keV peak. Those are

• a Compton shoulder, modeled as a smeared step function, to describe scattering γs

$$f_{step} = \frac{1}{2} \operatorname{erfc}\left(\frac{E-\mu}{\sqrt{2}\sigma}\right)$$

an X-ray escape peak, due to the absorption and subsequent emission (and escape) of de-excitation X-rays from ¹³⁰Te. The strongest lines have energies ranging from 27 to 31 keV, resulting in a peak 30 keV below the ²⁰⁸Tl line. This component is modeled as the superposition of 6 gaussian peaks, one for each of the strongest known


FIG. 4.4: An example of the distribution of the FWHM resolutions attained by the CUORE bolometers in one dataset is shown. The numerical FWHM values are extracted from the simultaneous fit of the lineshape response function to the 2615 keV line in calibration data. The exposure weighted harmonic mean of the individual channels' resolutions is shown (red dashed line). The fraction of active channels for the considered dataset is reported in the legend.

emission lines of ¹³⁰Te, properly weighted to account for the different intensity;

- a coincidence peak, centered at 2687 keV, resulting from the simultaneous absorption in one crystal of the 2615 keV and 583 keV γ rays from ²⁰⁸Tl, followed by pair production, positron annihilation and the escape of a 511 keV γ ;
- a polynomial background.

The normalization of all the additional components is let free to vary in the fit. Since the statistics is not enough to constrain both the line shape and the background parameters for each channel-dataset pair individually, the normalization of the background components is shared, within each dataset, among all the channels of the same tower. A simultaneous fit on each tower is then performed, letting the normalization of the individual background components float as a function of the tower, in order to extract the free parameters of Eq. 4.15 for each channel-dataset pair.

4.3.2 Energy bias and resolution evaluation

The available statistics in background data is much lower than the one available in calibration data. This is the reason why no channel dependent quantity can be evaluated, for what concerns the detector response function. It is nevertheless mandatory to cross check the validity of the line shape model extracted from calibration. For reasons still under investigation, the detector resolution is found to be energy dependent. In addition it



FIG. 4.5: Energy scale reconstruction bias $\Delta(Q)$ (left) and resolution scale factor $\Sigma(Q)$ (right) fit examples. A second order polynomial dependence is assumed for both. A significant improvement in the resolution at low energy is observed, while the bias level stays below 500 eV over the whole explored range.

is possible that the calibration procedure does not properly map the reconstructed energy variable, introducing a small bias. Both those contributions are modeled from known peaks in background data as dataset dependent functions (Fig. 4.5). The line shape response function in Eq. 4.15 is modified as follows

$$\mu_{ch,ds} \longrightarrow Q \cdot \left(\frac{\hat{Q}_{ch,ds}}{Q_{\text{Tl}}}\right) + \Delta(Q)_{ds}$$

$$\sigma_{ch,ds} \longrightarrow \Sigma(Q)_{ds} \,\hat{\sigma}_{ch,ds}$$

$$(4.16)$$

where $\hat{Q}_{ch,ds}$ and $\hat{\sigma}_{ch,ds}$ are the best fit line shape parameters from calibration data, whereas

$$Q_{\rm Tl} = 2614.51 \, \rm keV$$

is the literature value [93] of the fitted calibration peak from ²⁰⁸Tl, so that the fraction in round brackets is ~ 1 and any deviation from it describes the individual channel-dataset mismatch between the expected and reconstructed Thallium peak position in calibration data. The $\Delta(Q)_{ds}$ parameter describes the bias in the energy reconstruction, namely the difference between the reconstructed position and the expected one for an energy release Q, while $\Sigma(Q)_{ds}$ is a scaling factor to parameterize the dependence of the resolution on energy. Ideally $\Delta(Q) = 0$ and $\Sigma(Q) = 1$.

The result of the fit for the bias term typically satisfies

$$\Delta(Q)_{ds} \ll \sigma_{ch,ds} \tag{4.17}$$

at least for the majority of channel-dataset pairs, whereas the resolution is not constant as a function of energy (Fig. 4.5). The expectation of the resolution to be energy independent comes from the following considerations:

• the intrinsic energy resolution, due to the Poisson fluctuation of the number of carriers

(phonons), does have a clear energy dependence, but is negligible;

• a stationary noise superimposed to an energy independent signal template produces a constant noise term.

A mismatch (excess noise) is observed between the expected noise computed from the integral of the average noise power spectrum and the observed resolution. Even if this *feature* could be explained just with a small time dependence of the noise level that is not taken into account, there is also a mismatch between the pulser-induced event resolution and the particle-induced event resolution. This points to a noise source that depends on the geometrical details of the energy release and the thermalization process. It is beyond the scope of this manuscript to further investigate this issue but it is necessary to mention that work is still in progress to proceed from a phenomenological characterization of the detector response to a predictive physics-based description.

Chapter 5

Monte Carlo techniques and background control

In this chapter the Monte Carlo techniques used in CUORE to describe radioactive contaminations and extract the half life of ¹³⁰Te from its SM $\beta\beta$ decays will be explored. First a short description of the background control techniques that allowed the CUORE construction is given in Sec. 5.1. In Sec. 5.2 the simulation techniques of the known background sources are detailed together with the software tools. The included background components and the best fit to the CUORE data are considered in Sec. 5.3. The best fit result for the half life of ¹³⁰Te due to double beta decay with the emission of 2 neutrinos is reported. For further details about Bayesian analyses and related numerical algorithms a basic introduction is given in Sec. B, along with a description of the software tools used to perform it.

5.1 Material selection and background control

A key requirement for a rare decay search is the background control, as already discussed in Sec. 1.3.2. The target background for CUORE was a rate not higher than 10^{-2} cts/(keV kg yr) around the Q value of $\beta\beta$ decay. Detailed Monte Carlo simulations, based on the results from a material screening campaign, radioassays and bolometric measurements, were developed to evaluate the expected background of the CUORE experiment prior to its construction (see [56] [66] [94] and references therein). These were used in the past years to guide the construction strategies and to produce a projected background model. For each component of the detector, specific selection processes, purification techniques, assays were used to ensure cleanliness and radio-purity, some of which will be mentioned in the following. The Reader is referred to the bibliography for further details.

A number of sources, in addition to $0\nu\beta\beta$ decay, can produce events in an energy region of 100 keV around the double beta decay transition energy $Q_{\beta\beta}$ (Region Of Interest, ROI). They include natural and artificial environmental radioactivity, cosmogenically activated isotopes, cosmic rays. In most cases the contribution to the spectrum in the ROI is a continuum. Exceptions are the 2448 keV line from ²¹⁴Bi, the 2587 keV Te escape peak due to the 2615 keV ²⁰⁸Tl line, which can be excluded from the ROI, and the ⁶⁰Co sum peak at 2505 keV due to the simultaneous detection in a single crystal of both the ⁶⁰Co γ rays. Past experience from the MiDBD and the CUORICINO [95] [96] experiments allowed to identify the main contributors to the ROI event rate:

- multi-site Compton events from the 2615 keV γ line of ²⁰⁸Tl originating from ²³²Th contaminations of the cryostat
- surface 238 U and 232 Th contaminations of TeO₂ crystals, including the implantation of 210 Pb from environmental 222 Rn
- surface contaminations of inert materials close to the detectors, such as the copper holder structure, by ²³⁸U, ²³²Th and their decay products

A pre-selection of materials with very low radioactive contamination was performed, for both passive and sensitive parts of the detector. All materials were investigated through a very wide range of radio-assay techniques. Inductively Coupled Plasma Mass Spectroscopy (ICPMS), α and γ spectroscopy were performed [97] [98] [99]. An analysis on neutron activation was carried out in collaboration with the Laboratory of Applied Nuclear Energy (LENA) in Pavia, Italy [100] [101]. Then in order to reduce backgrounds from the detector materials the design was aimed at minimizing the copper mass close to the crystals and at developing procedures to avoid the introduction of contaminants during construction and storage of the detector components. An example is passing from the CUORICINO structure, where each floor was independently assembled as a sandwitch of crystals between two copper frames, to the CUORE-0 one where each tower was assembled as a single unit where each copper frame was used as top and bottom holder of neighboring floors, hence reducing the amount of copper between detector planes by a factor ~ 2 in mass and surface area. Techniques for crystal polishing and cleaning procedures for the copper structure were developed and described in detail in [66]. Custom designed glove boxes flushed with nitrogen were developed in order to prevent recontamination of materials after polishing and cleaning [102]. Nitrogen atmosphere assembly prevents ²²²Rn-induced surface contaminations, which can induce background [103]. Special packaging techniques were developed in order to prevent recontamination of detector parts during long-term storage. Dedicated production lines for raw material synthesis, crystal growth, and surface processing were built compliant with specific radio-purity constraints. High sensitivity measurements of radio-isotope concentrations in raw materials, reactants, consumables, ancillaries, and intermediary products used for TeO₂ crystals production are reported in [50]. Validation runs to test the performances and the radiopurity, measure bulk and surface contaminations of the CUORE crystals produced at SICCAS (Shanghai Institute of Ceramics, Chinese Academy of Sciences) for the CUORE experiment were carried out since the end of 2008 at the Gran Sasso National Laboratories [49].



FIG. 5.1: Model of the CUORE detector array, internal shielding and part of the surrounding support and cryogenic infrastructure as implemented in the Monte Carlo simulation software [94]

5.2 Monte Carlo simulations

A reliable simulation of the effects that radioactive contaminants have on the reconstructed CUORE spectrum is crucial not only to disentangle different contributions and measure their activity but to extract unknown information, such as a precise measurement of the $2\nu\beta\beta$ decay rate of ¹³⁰Te, to make predictions and design the next generation experiments knowing which technological improvements to aim at. The tools used in CUORE to produce Monte Carlo (MC) simulations are based on the **Geant4** toolkit [104] [105] [106], and are called **qshields** and **g4cuore**. The first one takes care of simulating particle interactions across the volumes implemented to model the detector array and the materials nearby, the other is a reprocessing tool that implements the detector response and behaviour in order to produce an output as close as possible to the one at the end of the data processing.

5.2.1 Particle generation and propagation

The objects (materials and geometries) included in the detector model for MC simulations include the 988 TeO₂ tellurite crystals, their copper holder structure, PTFE supports, wire trays, NTDs, internal DCS guide tubes, the thermal shields and cryostat flanges, the internal and external lead shield, the neutron borated polyethilene shield as in Fig. 5.1. All simulated primary and secondary radiation is propagated through the implemented volumes in discrete steps, and energy releases in the active detector components are stored. An optimization between accuracy and required computing time is required and was performed tuning differently the range of production cuts in different materials. Some eletromagnetic processes (e.g. Bremsstrahlung or δ -ray production) involve infrared divergences. Production cuts limit the generation of particles below threshold (i.e. particles that have a range below a user-defined and material dependent range cut), treating the remaining divergent part as a continuous effect. The particle generation and propagation in the CUORE setup is performed by the **qshields** software. Possible simulated background sources differ by the generated particle types, rates, and their geometry. Particles that can be simulated are

- single particles, either monochromatic or with a given energy spectrum (e.g. deexcitation gammas);
- singular nuclear decays;
- decay chains, possibly accounting for breaks of secular equilibrium;
- double beta decays, where different calculations in literature [107] [108] [25] [36] are implemented to generate the shape of the sum energy spectrum of the emitted electrons in the $2\nu\beta\beta$ decay mode.

Other than different sources of radiation and their expected energy spectrum, their geometrical distribution can be set. Contaminations in the detector materials can be either uniformly distributed in any of the available volumes (bulk contaminations) or distributed with an exponentially decaying density profile on the surface of materials. This is the case of surface contaminations of α emitters both in the TeO₂ crystals and in the inner copper shields of the cryostat. Building reliable Monte Carlo simulations crucially depends on the physics list (included processes) and on the production cuts in each material. Further discussion on the volume by volume optimization of such simulation parameters is important but it lies beyond the scope of this manuscript.

5.2.2 Detector response simulation

Once the output of a **qshields** simulation is available, a simulation of the detector response and the data processing chain (see Sec. 3.3) with the **g4cuore** software is needed. This step implements:

- trigger thresholds, allowing the user to input different parameters for each channel;
- excluded channels, importing a dataset dependent list of channels that are excluded from the data production because of some failure in the data processing (e.g. miscalibration);
- hit reconstruction and pile-up discrimination, the trigger dead time is implemented as well as an integration time. All energy releases happening in the same crystal within one integration time are summed together because they happen too close in time to be detectable as pile-ups;
- detector resolution, it is implemented as a gaussian smearing with channel-dependent resolution;
- base cut and PSA cut efficiency, the efficiency curve in Eq. 4.6 is implemented, with channel dependent parameters. The plateau efficiency is adjusted to include the base cut contribution;

- coincident multiplets reconstruction, according to the settings provided in the data processing concerning time-based and space-based windows;
- particle-dependent quenching factor, since the reconstructed energy of α particles was observed to be reconstructed at slightly higher values than expected [109], this feature is implemented to shift the reconstructed energy of α particles with a multiplicative correction (quenching) factor.

Splitting the production of simulated data in two steps has the main advantage of decoupling the most (CPU) time-consuming part related to particle generation and propagation across the detector geometry, which stays always the same, from the simulation of the detector response and processing chain. The last part requires much less computational resources but is much more likely to change over time. For example a retriggering of the data with improved trigger thresholds or with modified settings for the coincidence algorithm changes the detector response without changing its geometry. This architecture allows to timely update the simulations just running again the last step with different settings, without having to simulate again the particle generation and propagation.

5.3 Background model

The most common background sources for low-counting underground experiments are long lived radioactive contaminants of the detector and surrounding materials, radioactive isotopes coming from cosmogenic activation of materials, environmental radiation and cosmic muons. With the exception of the radiation entering the experimental setup from outside the external shielding, all the other components are modeled specifying the following parameters: contaminated physical volume, contaminant, geometrical distribution (bulk or surface). Bulk contaminations are distributed uniformely within the physical volume, surface contaminations have an exponentially decaying density profile with depth of $0.1 \,\mu\text{m}$, $1 \,\mu\text{m}$ or $10 \,\mu\text{m}$. This is mostly important for contaminations of α emitters, due to their short range. For this reason all contaminations located outside the 10 mK vessel are simulated as bulk only. ⁴⁰K, ²³⁸U and ²³²Th are present in all volumes, ⁶⁰Co is present both in TeO₂ and Cu materials due to cosmogenic activation. Concerning cosmic muons and environmental γ and neutron fluxes they are modeled according to known experimental distributions [110] [111].

Once the background sources are identified, the problem of building a background model can be expressed as the one of finding the activity of each included source able to describe the observed counts in the CUORE detector. This task is performed taking advantage from the granularity of the detector to disentagle and constrain different contaminations. Two *layers* are defined, namely inner (L0) and outer layer (L1), as depicted in Fig. 5.2. A simultaneous fit on single-site (Multiplicity = 1, M1) and double-site (Multiplicity = 2, M2) events is performed in order to extract the activity of each contamination. More specifically, the following binned spectra are produced



FIG. 5.2: Side (left) and top (right) view of the CUORE detector array. In green (cyan) is highlighted the inner (outer) layer.

- M1L0, energy of the single site events in the inner layer;
- M1L1, energy of the single site events in the outer layer;
- M2, energy of the double site events in any layer (both coincident channels included);
- M2sum, sum energy of the double site events in any layer.

The bin size depends on energy in order to include peaks from known α or γ emissions in one single bin. This choice is especially appropriate to describe α decays, where nongaussian tails are observed and it is difficult to disentangle the origin of the reconstructed shape between detector response and density profile of the contaminants (e.g. surface contaminations). In this way systematic uncertainties due to a mismatch between the resolution function applied to smear the simulated data and the detector response are minimized. Then a set of simulated data is produced, one for each included background source, with a known (large) number of generated primaries. The fit is performed scaling the histograms of the above mentioned binned spectra from each simulated source s with a scaling factor f_s and summing together all the contributions. The fit consists in a Bayesian estimation of the normalization factors, with the following likelihood function

$$\mathcal{L}(\boldsymbol{n}|\boldsymbol{f}) = \prod_{h,b} \mathcal{P}(n_{h,b}|N_{h,b}(\boldsymbol{f}))$$
(5.1)

where $\mathcal{P}(n|N) = N^n e^{-N}/n!$ is the Poisson probability of observing n counts expecting N, n is the matrix of observed counts $n_{b,h}$ in bin b for spectrum $h \in (M1L0, M1L1, M2, M2sum)$, fis a vector of normalization factors f_s of each source s and $N_{h,b}(f)$ is the number of expected counts in bin b of spectrum h given the normalization factors f. The normalization factors are independent, as a consequence the prior probability distribution is the product of the priors for each source. When no information is available the prior probability distribution function (pdf) is chosen uniform in the range $f_s \in [0, f_s^{max}]$ where f_s^{max} is the normalization factor corresponding to the maximum activity of the corresponding source compatible with



FIG. 5.3: The best fit of the background model (red) to the binned data (blue) is shown (top) together with their ratio (bottom). [112]

the CUORE data. Otherwise, if a measurement of one of the activities exists, the prior pdf is Gaussian distributed according to the measurement result.

The Bayesian fit is performed with the JAGS software [113], which implements a Markov Chain Monte Carlo (MCMC) via the Gibbs sampling algorithm. Details about both Bayesian analyses and MCMC techniques are available in Sec. B. The best fit histograms and residuals for the M1L1 spectrum (inner layer) are shown in Fig. 5.3. The marginalized posterior pdf is used to compute estimators for each source activity, including ¹³⁰Te in the $2\nu\beta\beta$ mode. This process accounts for most of the observed counts in the 1-2 MeV region (Fig. 5.4). In order to extract the activity of each source it is enough to invert the relation

$$f_s N_{MC}^{(s)} = A_s \,\Delta t \tag{5.2}$$

where $N_{MC}^{(s)}$ is the number of generated chains¹ and A_s the activity of source s. In order to extract the half life of ¹³⁰Te its dependence from the activity must be considered

$$T_{1/2}[s] = \frac{\ln(2)}{A[\text{Bq}]} \frac{M(^{nat}\text{TeO}_2) \,[\text{kg}]}{m(^{nat}\text{TeO}_2)[\text{g/mol}]} \,\eta \,N_A[\text{mol}^{-1}] \,10^3$$
(5.3)

¹By *chain* we mean a nuclear decay or a primary particle that is simulated and propagated together with all related secondaries.



FIG. 5.4: The binned energy spectrum of single site events in the inner layer is shown (blue) together with the background best fit for the $2\nu\beta\beta$ component (cyan) and all the others (yellow). [112]

where $M(^{nat}\text{TeO}_2)$ is the total mass of the crystals, $m(^{nat}\text{TeO}_2)$ is the molecular mass of tellurium dioxide, η is the natural isotopic abundance of the isotope of interest and N_A is the Avogadro constant. Plugging Eq. 5.2 into Eq. 5.3 we can extract a measurement [114] of the ¹³⁰Te half life from the best fit normalization factor

$$T_{1/2}^{2\nu}(^{130}\text{Te}) = (8.2 \pm 0.1_{\text{stat}}) \times 10^{20} \text{ y}$$
 (5.4)

Chapter 6

Neutrino-less Double Beta Decay search to the ground state of 130 Xe

In the following chapter the latest results on the neutrino-less double beta decay of 130 Te to the ground state of 130 Xe [41] are detailed. The work described in this chapter contains significant original contributions of the Author. It was carried out in a working group of 3 people, which contributed equally to all the detailed items, from the development of the Bayesian fit model, the implementation of the needed software tools, the definition, modelling and implementation of systematic uncertainties, to the Toy Monte Carlo data production and fit for the extraction of the expected limit setting sensitivity. In addition many other members of the CUORE Collaboration, including the members of the above mentioned working group, contributed to the data taking in terms of detector operation, data processing, data quality evaluation. The work [41] includes an introduction about the physics case (Ch. 1), the CUORE detector (Ch. 2) and data processing (Ch. 3) that are not discussed here.

Differently from $2\nu\beta\beta$ where neutrinos carry some amount of *invisible* (i.e. not detected) energy, causing the $\beta\beta$ sum energy spectrum to be continuous from 0 to $Q_{\beta\beta} = (2527.518 \pm 0.013) \text{ keV} [115] [116] [117]$, the neutrino-less double beta decay of ¹³⁰Te to the ground state of ¹³⁰Xe corresponds to a monochromatic energy release at $Q_{\beta\beta}$. Its spectral signature is then a small peak at the endpoint of the $\beta\beta$ spectrum. The analysis consists in selecting good (see Sec. 6.1) single-site events that reconstruct energetically close to $Q_{\beta\beta}$. Then the reconstructed spectrum is modeled with a flat background, a peak due to ⁶⁰Co decay and the posited peak at $Q_{\beta\beta}$ (Sec. 6.2). The $0\nu\beta\beta$ rate is evaluated with a Bayesian fit using a Markov Chain Monte Carlo (Sec. 6.3 and B), with the results shown in Sec. 6.4.

dataset	start_date	stop_date	active channels	TeO_2 exposure $(kg \cdot yr)$	exposure $(kg \cdot yr)$
3519	2017-05	2017-06	905	38.57	10.73
3522	2017-08	2017-09	954	49.68	13.81
3552	2018-05	2018-07	911	70.63	19.64
3555	2018-07	2018-08	900	65.93	18.33
3561	2018-11	2018-12	931	18.97	5.27
3564	2019-03	2019-05	945	57.58	16.01
3567	2019-05	2019-07	947	71.12	19.78

TABLE 6.1: Summary of the processed data at the time of writing with breakdown of the exposure for each dataset. The number of active channels refers to the number of channels that successfully underwent all the processing steps.

6.1 Event selection

The analysis presented in the following is based on the data collected from May 2017 to July 2019 (Fig. 3.4). The total analyzed TeO₂ exposure is 372.5 kg · yr, which corresponds to 103.6 kg · yr of ¹³⁰Te exposure, divided in 7 datasets as detailed in Tab. 6.1. With a probability $\epsilon_{cont} = 88 \%$ [118] the energy released by the two emitted electrons is fully contained in one crystal (containment efficiency). For this reason the search for this process is performed selecting a Region Of Interest (ROI) in the M1 spectrum of single-site events around $Q_{\beta\beta}$ as

$$E \in [E_{\text{ROI}}^{min}, E_{\text{ROI}}^{max}] \tag{6.1}$$

where E is the reconstructed energy. For this analysis the coincidences' threshold is set at 40 keV. Thanks to the optimum trigger algorithm > 97 % of the channels have > 90 % trigger efficiency at the coincidences' threshold. At the same time this choice ensures that 2615 keV γ rays from ²⁰⁸Tl decay that simultaneously release energy in two crystals are either flagged as M2 or reconstructed in the M1 spectrum above 2575 keV. Choosing the upper edge of the ROI $E_{\rm ROI}^{max} \leq 2575$ keV prevents such events to enter the region of interest. In addition, for the sake of sample purity, no radial cut is applied to coincidences. The effect of this choice is that the whole CUORE experiment acts as an anti-coincidence veto. The coincidence time window is set to ±5 ms.

Other selections used to define candidate events for $0\nu\beta\beta$ search in this channel involve data quality, rejection of pulser-induced events, pile-up events and baseline fluctuations incompatible with the average response of the detector to particle-induced pulses. As detailed in Sec. 3.2 for each channel it is possible to define periods of time where the detector behavior is not optimal, for any reason. Such *bad intervals* are excluded from the analysis and the live time of the involved channel is corrected accordingly. Then candidate events are required to be triggered (Sec. 3.1.1) in an asymmetric 10 s window¹ free from pulser events. In addition all candidate events must belong to a channel-dataset pair that underwent successfully all steps of the data processing. Channel-dataset pairs that fail

¹Each trigger flag defines a time window from 3 s before (pre-trigger) to 7 s after (post-trigger) the trigger fired.



FIG. 6.2: The CUORE spectrum after the base cut selection (grey), anti-coincidence veto (red) and PSA cut (blue) are applied. Reprinted from [41]

some step of it are referred to as *bad channels*, or *inactive channels*. Eventually, pile-up events are loosely rejected relying just on the waveform derivative to count the number of detected pulses. Together with the rejection of bad intervals, this selection defines the *base cuts*. Finally pulses that differ more than a given threshold in terms of the Mahalanobis distance from a reference sample are rejected (PSA cut)². The PSA cut has a much better pile-up rejection capability than the base cuts. The energy spectra of reconstructed events after base, anti-coincidence and PSA cuts are shown in Fig. 6.2. The effect of such selection cuts is modeled with dataset dependent efficiency terms (details in Sec. 4.2 and summarized in Tab. 6.4), namely

- containment efficiency, probability of a $0\nu\beta\beta$ decay to be fully contained in a single crystal, evaluated with Monte Carlo simulations, not included in the *total analysis* efficiency because it depends only on the geometry of the detector rather than on the data analysis;
- base cut efficiency, includes detection (trigger), energy reconstruction and pile-up rejection efficiencies, is evaluated with pulser events and takes into account the probability of triggering, correctly reconstructing the energy and rejecting pile-up;
- anti-coincidence efficiency, describing the signal selection efficiency of the anti-coincidence veto, computed as the survival probability of the 40 K γ line at 1460 keV. The γ emis-



FIG. 6.3: Left: spectrum in the [2490, 2575] keV energy range (ROI) with best fit curve (solid red) and the model corresponding to the marginalized 90 % C.I. upper limit on the $0\nu\beta\beta$ rate (dashed blue). Right: marginalized posterior distribution as a function of the decay rate $\Gamma_{0\nu}$ with all systematics included for the fit in both the physical range ($\Gamma_{0\nu} > 0$) and the full range. The 90 % C.I. is shown in blue. Reprinted from [41].

TABLE 6.4: Efficiency contributions involved in the $0\nu\beta\beta$ decay search in the gs-gs transition.

Efficiency [%]	Value
Detection	99.8440(9)
Energy reconstruction	99.1010(21)
Pile-up rejection	96.821(3)
Reconstruction (total)	95.802(3)
Anti-coincidence selection	98.70(13)
Pulse shape selection	92.60(15)
Total analysis	87.50(20)
Containment	88.35(9)

TABLE 6.5: Analysis cuts and effective parameters for $0\nu\beta\beta$ decay search in the gs-gs transition. When a single parameter is not representative or the spread among datasets is relevant (active channels, resolution scaling, PSA cut) a range is reported. Otherwise the sum (exposure), an average (background index) or an exposure-weighted value (resolution) is shown. The resolution is expressed in terms of the Full Width Half Maximum (FWHM). In a gaussian distribution the FWHM corresponds to ~ 2.4 σ .

Parameter	Unit	Effective value
Active channels	-	$900 \leftrightarrow 954$
TeO_2 exposure	$[m kg \cdot yr]$	372.46
¹³⁰ Te exposure	$[m kg \cdot yr]$	103.58
Resolution at 2615 keV in calibration data	FWHM [keV]	7.72 ± 0.02
Resolution scaling at $Q_{\beta\beta}$	-	$0.7(2) \leftrightarrow 1.0(9)$
Resolution at $Q_{\beta\beta}$ in physics data	FWHM [keV]	7.06 ± 0.02
Bias at $Q_{\beta\beta}$	$[\mathrm{keV}]$	< 0.3
Systematic effect on PSA	-	0.0056 ± 0.0013
Mahalanobis distance M_{dist} cut (PSA)	-	$3.5 \leftrightarrow 5$
Background index	$[\mathrm{counts}/(\mathrm{keV}\cdot\mathrm{kg}\cdot\mathrm{yr})]$	0.0138 ± 0.0007

sion follows an electron capture with a K-shell energy of $\sim 3 \text{ keV}$ (below threshold), so it is a sample of pure M1 events;

• pulse shape selection efficiency, extracted from a sample of double-crystal events summing at the energy of known prominent γ lines and cross checked with single-site events with full absorption of ⁶⁰Co and ⁴⁰K γ lines. The efficiency value is computed as the average of the efficiency estimates obtained with the above methods while the difference (±0.7%) is accounted for as a systematic effect.

6.2 Model of the reconstructed spectrum

The spectrum of reconstructed events as a function of energy is modeled with three components:

- flat background, this is motivated by Monte Carlo simulations and validated by previous experience [89]. The main background contribution (~ 90%) comes from *degraded* α particles [41], i.e. contaminants distributed over surfaces of both the detectors and the surrounding materials that undergo α decays where the emitted α particle escapes the crystal and releases part of its energy in some non-active part of the apparatus, circumventing the anti-coincidence veto. The rest originates from ²⁰⁸Tl 2615 keV γ rays undergoing Compton scattering in multiple crystals;
- sum peak at ~ 2505 keV due to ⁶⁰Co decays where both the 1173 keV and 1332 keV emitted de-excitation γ rays are fully absorbed in the same crystal;
- posited $0\nu\beta\beta$ decay peak at $Q_{\beta\beta}$.

All the included monochromatic energy releases are modeled with the detector response function in Eq. 4.15 corrected for the bias and scaling (dependence) of the resolution as a function of energy (Eq. 4.16).

The choice of the ROI ends is dictated by the need to include the widest possible region around $Q_{\beta\beta}$ in order to best constrain the background without including non-needed spectral features (other background peaks or structures) that, instead, introduce systematic uncertainties and unnecessary model complexity. The fit is performed on events reconstructed in the [2490, 2575] keV energy range but a wider range is visible in Fig. 6.3. A small ~ 2σ excess around 2480 keV is visible, at the moment not explained by any known contaminant, but more data is needed to assess whether it has a different origin than just plain statistical fluctuations. Meanwhile the choice of excluding such region from the ROI is taken. As previously mentioned, the upper end of the fit range is instead dictated by the threshold applied to the coincidence algorithm.

For any given channel-dataset pair, the probability density function used to model the data can be written as follows

$$f_{ch,ds}(E|\Gamma_{0\nu},\boldsymbol{\theta}) = \frac{\lambda_S f_S(E|\Gamma_{0\nu},\boldsymbol{\theta}) + \lambda_{\rm Co} f_{\rm Co}(E|\boldsymbol{\theta}) + \lambda_B f_B(E)}{\lambda_S + \lambda_B + \lambda_{\rm Co}}$$
(6.2)

where each of the f functions in the right hand side is a normalized p.d.f. as a function of energy in the ROI

$$\int_{E_{ROI}^{min}}^{E_{ROI}^{max}} f_i(E|\boldsymbol{\theta}) dE = 1 \quad \text{where} \quad i = S, \text{Co}, B$$
(6.3)

and models the signal, ⁶⁰Co and background components respectively. The λ_i components represent instead the number of counts assigned to each of the components, they all depend on $\boldsymbol{\theta}$, λ_S depends on the ¹³⁰Te decay rate $\Gamma_{0\nu}$ in the $0\nu\beta\beta$ mode as well. The $\boldsymbol{\theta}$ symbol represents, collectively, all nuisance parameters and will be thoroughly investigated in the following.

Cobalt component It is well known that ⁶⁰Co is produced by cosmogenic activation [119], which mainly happens above ground. In the underground LNGS environment the production rate of ⁶⁰Co is negligible. This implies that, in the absence of a source of ⁶⁰Co nuclei, their number N(t) at time t obeys

$$N(t) = N e^{-t/\tau} \tag{6.4}$$

where N is the initial number of emitter nuclei. When a measurement is done on a nonnegligible timespan with respect to the decay constant τ , it is not possible to neglect the dependence of the activity on time. Since the lifetime of ⁶⁰Co is $\tau_{60} \sim 7.2 y$, which is not much larger with respect to the data taking timespan of $\sim 2 y$, this effect must be taken into account. Nonetheless the duration of a dataset $\sim 0.1 y \ll \tau_{60}$ is negligible with respect to the decay constant, for this reason a dataset dependent correction factor is introduced rather than an event-dependent one. The number of counts attributed to 60 Co decay is parameterized as

$$\lambda_{Co} = \Gamma_{60} \epsilon \, e^{-t_{ds}/\tau_{60}} \tag{6.5}$$

where $\epsilon = \epsilon_{reco} \cdot \epsilon_{PSA}$ is the product of the reconstruction and PSA efficiencies and t_{ds} is the number of days between the beginning of the physics data taking in dataset ds and the beginning of the first included dataset. The efficiencies, the τ_{60} and Γ_{60} are nuisance parameters, the others are constants.

Signal component The signal component normalization factor depends linearly on the decay rate as

$$\lambda_S = \Gamma_{0\nu} N_0 \Delta t \epsilon \tag{6.6}$$

where N_0 is the number of emitter nuclei in the specified crystal (channel) and Δt its live time in the considered dataset and ϵ is an efficiency term. More explicitly

$$\lambda_S = \Gamma_{0\nu} \frac{M(^{nat}\text{TeO}_2)\Delta t}{m(^{nat}\text{TeO}_2)} \eta N_A \cdot \epsilon_{cont} \cdot \epsilon_0 \cdot \epsilon_{PSA}$$
(6.7)

where $M(^{nat}\text{TeO}_2)$ is the mass of the crystal, $m(^{nat}\text{TeO}_2)$ the molecular mass, η the isotopic fraction of ^{130}Te and N_A the Avogadro constant and ϵ_i are efficiency terms as detailed in Tab. 6.4.

Both the signal and the Cobalt components are modeled with the lineshape function in Eq. 4.15:

$$f_i(E|\Gamma_{0\nu}, \boldsymbol{\theta}) = f(E|\mu, \sigma, A_L, A_R, a_L, a_R) \quad \text{where} \quad i = S, \text{Co}$$
(6.8)

centered at the expected energy Q of the corresponding contribution and correspondingly scaled in width to match the expected detector resolution at that energy, as detailed in Eq. 4.16.

$$\mu_{ch,ds} = Q \cdot \left(\frac{\hat{Q}_{ch,ds}}{Q_{\text{Tl}}}\right) + \Delta(Q)_{ds}$$
$$\sigma_{ch,ds} = \Sigma(Q)_{ds} \,\hat{\sigma}_{ch,ds}$$

Both the Q parameters are included as nuisance parameters, incorporating the actual knowledge on the endpoint of the double beta decay spectrum of ¹³⁰Te and the sum energy of the γ s that follow ⁶⁰Co decay. The $\Delta(Q)$ and $\Sigma(Q)$ functions are parameterized as second order polynomials. The six coefficients needed to describe them are nuisance parameters as well. **Background component** The shape of the background component is simply a normalized uniform distribution in the ROI,

$$f_B(E) = \frac{1}{E_{\text{ROI}}^{max} - E_{\text{ROI}}^{min}}$$
(6.9)

while the number of background events is channel-dataset dependent as

$$\lambda_B = \mathrm{BI}_{ds} \cdot (M\Delta t)_{ch,ds} \cdot \left(E_{\mathrm{ROI}}^{max} - E_{\mathrm{ROI}}^{min} \right) \tag{6.10}$$

where the $M\Delta t$ term is the TeO₂ exposure and is a constant as well as the ends of the ROI while BI_{ds} are nuisance parameters, one per included dataset.

6.3 Bayesian fit

The inference on the parameter of interest $\Gamma_{0\nu}$ is performed with a Bayesian unbinned extended maximum a posteriori probability fit performed simultaneously on all the involved channel-dataset pairs. It can be shown (Sec. A) that the unbinned extended likelihood can be obtained from a binned likelihood in the limit of vanishing bin size. Since, by definition, the likelihood does not have to be normalized but just proportional to the probability of the data given a model (i.e. the functional form of the model and the values of the parameters that define it), it is possible to write it as

$$\mathcal{L}_{ch,ds} = e^{-\lambda} \prod_{i \in \text{ROI}} \left[\lambda_S f_S(E_i | \boldsymbol{\theta}) + \lambda_{\text{Co}} f_{\text{Co}}(E_i | \boldsymbol{\theta}) + \lambda_B f_B(E_i) \right]$$
(6.11)

starting from Eq. A.4, where $\mu = \lambda$ and $\lambda = \lambda_S + \lambda_{Co} + \lambda_B$ is the total number of expected counts. Realizing that the *n*! term is irrelevant because it does not depend on $\Gamma_{0\nu}$ nor on θ and that the normalization term at the denominator of Eq. 6.2 cancels with the λ^n term³, Eq. 6.11 is obtained. Then, in order to write the overall likelihood term, we must compute the product of the individual channel-dataset contributions as

$$\mathcal{L}(\mathcal{D}|\mathbf{M}) = \prod_{ch,ds} \mathcal{L}_{ch,ds}$$
(6.12)

where \mathcal{D} denotes the data and M the model detailed in the previous section.

The fit parameters are summarized and listed in Tab. 6.6 together with the implemented prior. A discussion is mandatory to motivate the prior choice. In Bayesian reasoning there is no such thing as (literally) *non-informative* prior probability distribution. The purpose of a Bayesian analysis is to consistently update the knowledge about the model's parameters. Omitting to plug knowledge in the priors is not possible. The principle of indifference motivates⁴ the uniformity of the prior for $\Gamma_{0\nu} > 0$ up to a value much larger

 $^{^{3}\}lambda$ does not depend on the event and can be factored out

⁴The writer is well aware that the choice of a flat prior as a function of one parameter, say the decay

Parameter	Symbol [units]	Prior
$0\nu\beta\beta$ decay rate	$\Gamma_{0\nu} \left[yr^{-1} \right]$	uniform
Background index (per dataset)	$BI_{ds} [counts/(keV kg yr)]$	uniform
⁶⁰ Co sum peak amplitude	$\Gamma_{60} \left[\text{counts/kg/yr} \right]$	uniform
60 Co sum peak position	$Q_{60} [\mathrm{keV}]$	uniform
Analysis efficiency (except PSA)	ϵ_0 [1]	Gaussian
Analysis efficiency (PSA)	ϵ_{PSA} [1]	uniform
Containment efficiency	$\epsilon_{cont} [1]$	Gaussian
Energy resolution scaling	$\Sigma(Q)$ [1]	multivariate Gaussian
Energy bias	$\Delta(Q) [1]$	multivariate Gaussian
$\beta\beta$ decay Q-value of ¹³⁰ Te	$Q_{etaeta} \; [{ m keV}]$	Gaussian
Isotopic fraction	$\eta \left[1 ight]$	Gaussian

TABLE 6.6: Parameters used in the $0\nu\beta\beta$ decay search in the gs-gs transition together with their symbol, units and related prior probability distribution. The parameters are divided as parameter of interest (top), nuisance parameters (middle), nuisance parameters included for systematic uncertainty evaluation (bottom).

than the current limit because the result does not depend on the upper end of the prior's domain. Rather, for non-physical negative decay rates, the prior is null. For debug purposes the fit was also performed extending the $\Gamma_{0\nu}$ parameter's domain to negative values, and this is how the distribution in the shaded gray area of Fig. 6.3 was obtained.

The background model does provide knowledge about the background, both regarding its value and its shape. The choice of not including such knowledge is motivated by two main issues:

- 1. the background model comes from a fit to the very same data that are included in this analysis. Plugging an information about the background index coming from the same source would bias the result. Moreover, such information would be extracted from a model that does not include $0\nu\beta\beta$;
- 2. the assumptions that underlie the background model construction are source of systematic uncertainty on the background index estimation. This choice avoids such uncertainties.

For the same reason the rate of the ⁶⁰Co sum peak has an associated uniform prior. Specifically this is linked to the lack of knowledge about the distribution of the ⁶⁰Co contaminant within the experimental setup. Regarding its position, it was observed [89] that the reconstructed peak is shifted by ~ 1 keV with respect to the expected position. The source of this shift is currently unknown. Nevertheless it is important to mention that the prior choice on the Cobalt sum peak has a negligible effect on the fit result. Even with a different prior (e.g. a gaussian prior centered on the literature value for Q_{60}) the marginalized posterior pdf for the parameter of interest is effectively unchanged.

rate, because no information is available that allows to favor some rates rather than others (principle of indifference), automatically implies that re-parametrizing the problem as a function of $T_{1/2} \doteq \log(2)/\Gamma$ forces a non-flat prior on the new parameter due to the Jacobian of the non-linear coordinate transformation function. There is plenty of literature on this topic and we wish not to discuss it here.



FIG. 6.7: Left: knowledge update plot for the Q_{60} parameter where the prior pdf (grey) is superimposed to the marginalized posterior pdf (black). Right: knowledge update plot for the Q_{60} (ordinate) and the Γ_{60} parameters (abscissa). The marginalized mode is shown (black box) within the 68 % smallest interval (black contour).

Nuisance parameters for systematics Before entering the following discussion, it is useful to anticipate that the effect of systematics is largely sub-dominant. Introducing additional nuisance parameters to account for known sources of systematic uncertainty, the global mode of the marginalized posterior is shifted by at most 0.04%. That said, since by definition the systematic uncertainties do not get better accumulating further statistics if no additional external information is plugged in, a very useful cross-check that the fit procedure is implemented correctly is the comparison of the prior and posterior distributions (knowledge update) for systematics nuisance parameters. The two distributions should be compatible because no information is available in the data regarding such parameters (Fig. 6.8). Moreover the prior knowledge should be adequate to describe the data and no modification should be needed unless the model itself is inaccurate. The other nuisance parameters, instead, embed a piece of indispensable - yet possibly not well known - information to model the data (Fig. 6.7) and the knowledge update shows how much the data is informative with respect to prior knowledge. This possibility is unique of Bayesian reasoning (frequentists do not update their degree of belief since they do not have any).

As detailed in Sec. 4.3.2 and Fig. 4.5 the energy scale reconstruction bias and resolution scale factor are evaluated from the data themselves. They are extracted with separate fits and parameterized as second order polynomials. For this reason the associated covariance matrix C (a second order approximation of the likelihood used for the fit, around its maximum) is block diagonal, i.e. the 3 parameters that describe the energy scale reconstruction (C_{Δ} block) are uncorrelated to the ones describing the resolution scaling with energy (C_{Σ} block). For this reason the prior pdf for these parameters is written in the following factorized form

$$\pi(\boldsymbol{\Delta}, \boldsymbol{\Sigma}) = \frac{\exp\left[\frac{1}{2}\boldsymbol{a}_{\Delta} - \hat{\boldsymbol{a}}_{\Delta}\right)^{\mathrm{T}}\boldsymbol{C}_{\Delta}^{-1}(\boldsymbol{a}_{\Delta} - \hat{\boldsymbol{a}}_{\Delta})\right]}{\sqrt{(2\pi)^{3}}|\boldsymbol{C}_{\Delta}|} \cdot \frac{\exp\left[\frac{1}{2}\boldsymbol{a}_{\Sigma} - \hat{\boldsymbol{a}}_{\Sigma}\right)^{\mathrm{T}}\boldsymbol{C}_{\Sigma}^{-1}(\boldsymbol{a}_{\Sigma} - \hat{\boldsymbol{a}}_{\Sigma})\right]}{\sqrt{(2\pi)^{3}}|\boldsymbol{C}_{\Sigma}|} \quad (6.13)$$

where a_i are column vectors including the polynomial coefficients of the energy bias function $(i = \Delta)$ or the resolution scaling function $(i = \Sigma)$ and the \hat{a}_i symbol indicates the



FIG. 6.8: Knowledge update plot for one coefficient of the $\Delta(Q)$ bias function for one dataset. The comparison between the prior (grey) and posterior (black) distributions shows agreement since no information about the energy reconstruction bias can be extracted just from the data in the ROI.

best fit values in either case.

6.4 Results

No evidence of $0\nu\beta\beta$ decay is found, and a Bayesian lower limit is placed (Sec. 6.4.2)

$$T_{1/2}^{0\nu} > 3.2 \cdot 10^{25} \, [yr] \quad \text{at} \quad 90\% \text{C.I.}$$
 (6.14)

corresponding to the blue dashed curve on Fig. 6.3. The inclusion of systematics yields a 0.4% weaker limit. The probability of obtaining a stronger limit is 3% (Sec. 6.4.1). This is either due to an unlikely local statistical fluctuation or to additional spectral features of the background that were not taken into account. When more exposure will be accumulated it will be possible to better model the background and to discriminate between these hypotheses. The reconstructed average background index is

$$\overline{\mathrm{BI}} = (1.38 \pm 0.07) \cdot 10^{-2} \, \mathrm{cts}/(\mathrm{keV \ kg \ yr}) \tag{6.15}$$

and is in agreement with the expectations derived from a preliminary background model (see Ch. 5 for further details). Nevertheless the exposure used to fit the latest version of the CUORE background model is smaller than the one used to extract the result in Eq. 6.15. Work is in progress to refine the modeling of the available exposure.



FIG. 6.9: Distribution of the 90 % C.I. lower limits obtained from the S+B minimal model fit to 10^4 independent ToyMC ensembles of 7 datasets each (blue solid). The median expected sensitivity of $1.7 \ 10^{25} \ [y]$ is superimposed (red dashed).

6.4.1 Exclusion sensitivity and blinded fit

In the early stages of the analysis some parameters are subject to change such as e.g. the lower (upper) end of the ROI, the detector response functions, the PSA Mahalanobis distance cut. Until all the selections, prior probability distribution functions for the nuisance parameters and any other parameter that could impact on the fit result are not fixed and agreed on by the Collaboration, only blinded data (see Sec. 3.3.7) are available. After all such parameters are defined, a fit to the blinded spectrum is performed. The results of that fit are used to compute a preliminary median expected limit setting sensitivity at 90%C.I. with the technique described in the following. Then the data unblinding is performed and a new fit is done with the signal plus background (S+B) minimal model in the ROI. By minimal model we mean the fit with the minimal number of nuisance parameters needed to fully describe the observed spectral shape. All the systematics' nuisance parameters are excluded (i.e. set to constant). The model that instead includes such parameters will be called *full model*. If the minimal S+B model yields, as in this case (Fig. 6.3), a rate of $0\nu\beta\beta$ compatible with 0, the fit is repeated with the background only (B model). The background indices extracted with this procedure are used to generate a set of toy monte *carlo* simulations (ToyMC) from which the final limit setting sensitivity is extracted.

Each ToyMC simulation consists in a set of simulated datasets, each according to the actual exposure of the acquired datasets included in the analysis, with events drawn from the best fit minimal background-only model. A total of 10^4 independent Toy Monte Carlo simulations are produced and then fit with the minimal S+B model to extract the distribution of 90 % credibility interval (C.I.) limits. The median of such distribution is quoted as the *median expected sensitivity* of CUORE with the considered datasets. The comparison of the lower limit obtained from a fit of the same S+B minimal model to the unblinded data with the distribution of ToyMC limits gives a measure of the probability of the observed statistical fluctuation at $Q_{\beta\beta}$. It is possible to define a threshold (e.g. the Gaussian 3-sigma quantile) above which an underfluctuation becomes *suspicious*. This process can be interpreted as testing the hypothesis whether the unblinded data could have been extracted from the minimal B-only model or not. Fixing such threshold defines the maximum acceptable probability of type I errors (i.e. probability of false positives, the probability of mistakenly rejecting the hypothesis just because of an unlikely statistical fluctuation). The obtained median expected limit setting sensitivity (see Fig. 6.9 for the distribution of ToyMC limits) is

$$S_{1/2}^{0\nu} = 1.7 \cdot 10^{25} \,\mathrm{yr} \tag{6.16}$$

As anticipated at the beginning of the section, the probability of setting a more stringent limit, with the same exposure, rather than the one obtained with the analyzed data, is 3%.

6.4.2 Lower limit and systematics on $T_{1/2}^{0\nu}$

The result of the ROI fit is a set of quasi-independent samples $(\Gamma_{0\nu}, \theta)^{\rm T}$ (Markov Chain steps) drawn from the joint posterior distribution. The best fit of the minimal S+B model is shown in Fig. 6.3. The red solid curve corresponds to the model of Eq. 6.2, summed over all active channel dataset pairs, and computed at the global mode

$$\hat{\boldsymbol{\theta}}' \doteq (\Gamma_{0\nu}, \boldsymbol{\theta})^{\mathrm{T}}$$
 where $P(\hat{\boldsymbol{\theta}}' | \mathcal{D}, \mathrm{S} + \mathrm{B}) = \max_{\Theta'} P(\boldsymbol{\theta}' | \mathcal{D}, \mathrm{S} + \mathrm{B})$ (6.17)

where $\Theta' \doteq [0, \Gamma_{0\nu}^{max}] \times \Theta$ is the joint parameter space for the parameter of interest and the nuisance parameters. The 90 % credibility interval C.I. limit on the neutrinoless double beta decay raye $\Gamma_{0\nu}^{lim}$ is computed integrating the marginalized posterior distribution as

$$0.9 = \int_0^{\Gamma_{0\nu}^{lim}} d\Gamma_{0\nu} \int_{\Theta} P(\theta') d\theta$$
(6.18)

The integral of the marginalized posterior in the $[0, \Gamma_{0\nu}^{max}]$ is highlighted as the blue area of Fig. 6.3. An under-fluctuation is observed at $Q_{\beta\beta}$ and its significance is quantified as the integral of the distribution $l(\Gamma)$ of expected 90 % C.I. limits obtained from minimal S+B model fit to ToyMC data (Fig. 6.9),

$$\int_{\Gamma_{0\nu}^{lim}}^{\infty} l(\Gamma) d\Gamma \sim 0.03 \tag{6.19}$$

The marginalized mode $\hat{\Gamma}_{0\nu}$ coincides with the end of the allowed range for the rate parameter. In order to evaluate the effect of systematics on the joint posterior distribution the fit was also performed artificially extending the prior on $\Gamma_{0\nu}$ to negative values. This is how the curve in the shaded grey area of Fig. 6.3 is obtained. In this way it was possible to find a global mode that does not lie on the endpoint of the $\Gamma_{0\nu}$ domain, even if it lacks



FIG. 6.10: Assuming $0\nu\beta\beta$ decay is mediated by light Majorana neutrinos, a lower limit on the $T_{1/2}^{0\nu}$ parameter corresponds to an upper limit on the effective Majorana mass $m_{\beta\beta}$ depending on the nuclear matrix elements used (yellow band). The width of the band corresponds to the spread of the NME used [41]. The expected CUORE sensitivity after 5 yr of live time is shown for comparison (blue band). The red and green bands correspond to the allowed values of the effective Majorana mass as a function of the lightest neutrino mass [42] in the normal and inverted ordering hypothesis respectively. The dark (light) areas show the allowed parameter space assuming the best fit (3 σ gaussian intervals) for the neutrino masses and mixing parameters reported in [43]. The Majorana phases range from 0 to 2π .

physical meaning,

$$\hat{\Gamma}_{0\nu} = (-3.5^{+2.2}_{-1.1}) \times 10^{-26} \, [1/\mathrm{yr}] \tag{6.20}$$

corresponding to a Gaussian $\sim 1.6\sigma$ background underfluctuation, compatible with the result obtained from the limit setting sensitivity mentioned above.

This extended-range minimal S+B fit was used as a starting point to evaluate the effect of the systematics, one nuisance parameter at a time, on the posterior pdf. As a proxy for the position of the posterior distribution the global mode in the extended-range fit can be used, while the 90 % C.I. limit reflects its width. To this respect, it is found that the maximum shift of the global mode due to the insertion of a nuisance parameter (bottom box of Tab. 6.6) is 0.04% (uniformly distributed analysis efficiency) while the overall effect on the limit adding all the systematics' nuisance parameters is a 0.4% shift towards lower values (weaker limit).

Under the assumption that light neutrino exchange is the only mediator of $0\nu\beta\beta$ decay

and using the phase space factors from reference [120], the result in Eq. 6.14 corresponds to an upper limit on the effective Majorana mass ranging from 75 meV to 350 meV due to the spread of the results of the NME calculations available in literature [27–29, 31–36, 121, 122] as shown in Fig. 6.10.

l Chapter

Double-Beta Decay search to the first 0^+ excited state of ^{130}Xe

7.1 Motivation and previous results

The interpretation of an experimental result on the neutrinoless double beta decay rate in different isotopes as a limit (or a measurement) of the $m_{\beta\beta}$ parameter, in the hypothesis of light Majorana neutrino exchange, requires knowledge of both phase space factors and nuclear matrix elements (NME) connecting the final and initial state nuclei. While the phase space factors can be computed with good accuracy [25], this is not the case for NME as discussed in 1.2. Besides being predicted but not yet discovered, a measurement of the rate of a known standard model process such as the double beta decay to the excited states is interesting because, together with the ground state transition, provides an important cross check for nuclear model calculations [123].

At the moment $2\nu\beta\beta$ decay to the 0_2^+ excited state was observed in just 2 isotopes. First in ¹⁰⁰Mo (1995, [124]), then in ¹⁵⁰Nd (2004, [125]) with half lives¹ of the order of few 10^{20} yr. Investigation of the same process in ⁹⁶Zr, ⁸²Se, ⁴⁸Ca, ¹¹⁶Cd, ⁷⁶Ge, ¹³⁶Xe, ¹³⁰Te, ¹²⁴Sn yielded upper limits at 90% C.L. from $3.1 \cdot 10^{20}$ y to $8.3 \cdot 10^{23}$ y for the T_{1/2} parameter (see [126] for a review). Previous attempts of measuring the $\beta\beta$ decay of ¹³⁰Te to the first 0⁺ excited state of ¹³⁰Xe ($2\nu\beta\beta$ gs-es transition in the following) were made by both the CUORICINO [127] and CUORE-0 [128] collaborations. The latter (not included in [126]) was published recently and includes the combination of the predecessor's results.

$$(T_{1/2})_{0^+_2}^{0\nu} > 1.4 \cdot 10^{24} \text{ yr}, 90\% \text{ C.L.}$$
 (7.1)

$$(T_{1/2})_{0_2^+}^{2\nu} > 2.5 \cdot 10^{23} \text{ yr}, 90\% \text{ C.L.}$$
 (7.2)

¹Here the *half life* term is used inappropriately to indicate, up to a factor of $\log 2$, the reciprocal of the decay rate. It would be appropriate to call such number so, if the father nucleus had only one decay channel.

Being a standard model process, it is possible to compute the $T_{1/2}^{0_2^+}$ observable in the $2\nu\beta\beta$ decay channel. The most recent predictions [129] are based on the QRPA approach and yield a range of predictions of

$$^{th}(T_{1/2})_{0_2^+}^{2\nu} = (7.2 - 16) \cdot 10^{24}$$
 y (7.3)

where the lower (upper) bound corresponds to the implementation of a linear model of the effective axial coupling constant as a function of the mass number A and the assumption $g_A = 0.6$ respectively. The first assumption is derived from the analysis of 24 isobaric triplets in the A = 100 - 136 range. The other instead is derived from data of both β^- and $2\nu\beta\beta$ for A = 100, 116, 128 nuclei [130] [131] with the ansatz that the effective g_A parameter to be used in calculations follows a different dependence from the mass number A in β^- and $\beta\beta$ decays. It should be noted that the same calculations predict the half life of ¹³⁰Te due to the dominant decay channel on the 0_1^+ ground state as

$${}^{th}(T_{1/2})^{2\nu}_{0^+_1} = (1.6 - 3.4) \cdot 10^{20}$$
 y (7.4)

whereas the latest experimental value (Eq. 5.4) is a factor 2-5 higher

$$^{exp}(T_{1/2})^{2\nu}_{0_1^+} = (8.2 \pm 0.1_{\text{stat}}) \times 10^{20} \text{ y}$$

indicating a possible systematic uncertainty in the theoretical evaluation of the half life. To this respect, the range of predictions about the decay rate to the first 0^+_2 excited state must be taken with a grain of salt, as the authors of [129] themselves acknowledge, stating that their predictions are accurate up to one order of magnitude. Nonetheless, since the experimental sensitivity for this process is starting to explore the above mentioned range of half lives, both a measurement or a more stringent limit with respect to [128] are informative from the point of view of refining and validating the theoretical computations.

Regarding the neutrino-less double beta decay channel instead, its observation on the excited state can help disentangle the underlying mechanism [129]. Both 0ν and 2ν decay channels allow a great background reduction with respect to the corresponding transitions on the ground state, due to the emission of coincident de-excitation γ rays.

7.2 Physical process and experimental signature

In order to define a selection criterion for double beta decay to the excited states some time must be spent investigating the nuclear structure of the daughter ¹³⁰Xe nucleus. Since $\beta\beta$ decay of ¹³⁰Te stems from the 0⁺ ground state of ¹³⁰Te, the 0⁺ final states are allowed while 2⁺ states are suppressed. The decay to the 2⁺ states, even if they sit at a lower energy and provide for this reason a higher phase space, is strongly suppressed and practically inaccessible to detection at present time. For this reason these nuclear final states will be neglected. The nuclear states of ¹³⁰Xe are depicted in Fig. 7.1, where the



FIG. 7.1: The decay scheme of ¹³⁰Te is shown with details about the involved excited states of ¹³⁰Xe up to its first 0⁺ excited state. An energy scale is shown (right) where the ¹³⁰Xe ground state is taken as reference. Gamma de-excitation channels are shown for each state (red arrows) with associated probabilities (%). Global γ de-excitation patterns from the 0⁺₂ state to the ground state of ¹³⁰Xe are shown as orange (channel A), green (channel B) and magenta (channel C) arrows, in decreasing order of probability. [132]

nomenclature $0_1^+, ..., 0_n^+$ indicates states with the same angular momentum in increasing order of excitation energy. The lowest excited state with a given angular momentum is called *yrast* and has subscript 1, the others *yrare* and have subscript n > 1. The deexcitation of the ¹³⁰Xe nucleus matches three possible patterns i.e. paths through states of decreasing energy from the 0_2^+ to the 0_1^+ ground state as shown in Fig. 7.1. Details about the probability of each de-excitation channel and the energy of the emitted γ rays are reported in Table 7.2. It is important to remember that there are 3 possibilities of de-excitation which will be referred in the following as pattern A, B and C in decreasing order of probability.

Full containment requirement Since neither $0\nu\beta\beta$ decay nor $2\nu\beta\beta$ decay were ever observed in ¹³⁰Te on any excited state, this experimental search has the goal of discovering either (or both) of the processes. For this reason, in order to provide convincing evidence for a possible claim of discovery, the requirement of *full containment* of the final state was set. More specifically this means that, except for neutrinos:

- 1. each final state particle must be detected, for this reason experimental signatures where one or more of the final state γ s escapes the source crystal and is absorbed by some non-active part of the experimental apparatus will not be included;
- 2. each final state particle must fully release its energy in no more than one crystal. This requirement is of paramount importance because it rejects all the situations where a Compton scattering allows the energy of a single de-excitation gamma to be splitted among two or more detectors.

Another consequence of the full containment requirement is that it sets a limit on the

maximum number of simultaneously involved crystals (Multiplicity). In fact the maximum allowed Multiplicity for a candidate multiplet is set by the maximum number of secondaries. In this case the limit is Multiplicity = 4 set by de-excitation pattern B where the $\beta\beta$ and each of the de-excitation γ s release energy in different crystals.

Experimental signatures definition and labelling The problem of counting the possible candidate signatures is purely combinatorial, and the answer is known since 1730 [133]. Rephrased in mathematical words it consists in counting the number of ways to partition a set of n labelled objects (the secondaries of each de-excitation pattern) into $k \leq n$ non-empty unordered subsets. It is referred to as Stirling number of the second kind and holds

$$S(n,k) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^{i} \binom{k}{i} (k-i)^{n}$$
(7.5)

In this case, the number of candidate *partitions* S_p for each de-excitation pattern p is given by

$$S_p = \sum_{i=2}^{k_p} S(k_p, i) \text{ where } p \in \{A, B, C\}$$
 (7.6)

where k_p is the number of emitted secondaries for pattern p. The sum runs from i = 2 because for Multiplicity = 1 the neutrino-less channel would be indistinguishable from the same decay on the ground state of ¹³⁰Xe. The $2\nu\beta\beta$ channel, instead, would appear as a kink in the

$$\left[Q^{0^+_1}_{\beta\beta}-Q^{0^+_2}_{\beta\beta},Q^{0^+_1}_{\beta\beta}\right]$$

portion of the energy spectrum of single-site events, where $Q_{\beta\beta}^{0_1^+}$ and $Q_{\beta\beta}^{0_2^+}$ indicate the double beta decay Q-value of ¹³⁰Te on the ground and 0_2^+ excited state respectively, suffering from high background from the decay to the ground state. For this reason Multiplicity = 1 signatures are not taken into account.

The number of partitions is 4 for pattern A and C, 14 for pattern B. Each of the partitions is labelled with strings of 3 characters each with the following convention

[Multiplicity] [Pattern] [Index]

where Multiplicity = 2,3,4 indicates the number of simultaneously involved crystals, Pattern = A,B,C stores the originating de-excitation pattern, and Index is a unique integer counter to distinguish partitions that share multiplicity and pattern. Table 7.3 collects all the possible partitions of the secondaries in the full containment hypothesis. Some rows contain more than one partition because the fundamental quantity is the expected energy release of each signature. Signatures sharing the same expected energy release are equivalent and are merged. For this reason instead of handling a total number of 22 signatures we are left with 15.

Pattern	Branching ratio	Energy γ_1	Energy γ_2	Energy γ_3
A	86%	$1257 \mathrm{keV}$	$536 \mathrm{keV}$	-
В	12%	$671 \mathrm{keV}$	$586 \mathrm{keV}$	$536~{ m keV}$
С	2%	$1122 \mathrm{keV}$	$671 { m keV}$	-

TABLE 7.2: The de-excitation γ rays emitted by 130 Xe^{*} in the transition from the 0_2^+ to the ground state. Each row corresponds to a different path through intermediate states. The energies of the emitted γ s are listed, in order of energy, along with the branching ratio of each pattern [132].

Ranking the experimental signatures Handling simultaneously 15 independent signatures to extract a measurement of the double beta decay rate is possible but requires effort. It is possible to perform a down-selection thanks to Monte Carlo simulations of both signal and background. The rationale behind such down-selection is that not all signatures contribute equally to the sensitivity. It could be possible to set a threshold on the sensitivity contribution to the $\beta\beta$ decay rate of down-selected signatures, that reflects a reasonable trade-off between model complexity, time required and expected sensitivity. It should be remembered that increasing the model complexity corresponds to the introduction of systematic uncertainties and increases the required computation time. In order to perform the ranking, the expected discovery sensitivity of each signature must be computed. From an approximate expression for the discovery sensitivity (see Appendix 1.3.1 for details on the underlying hypotheses and derivation) it is possible to define a score function factoring the common terms that appear in a background-limited peak search and the corresponding background-free case

$$F(\epsilon_s, B_s) \doteq \theta(B_s - 1) \frac{\epsilon}{\sqrt{B}} + \theta(1 - B_s) \frac{5\epsilon}{-\ln(3 \cdot 10^{-7})}$$
(7.7)

where θ is the Heaviside function, ϵ_s and B_s are the signal efficiency and expected number of background counts for signature s. The first term in Eq. 7.7 is valid for high-background signatures, the other for background free ones. The transition is set sharply at $B_s = 1$. The relative score of each scenario is computed simply as

$$S_s \doteq \frac{F(\epsilon_s, B_s)}{\sum_{s'} F(\epsilon_{s'}, B_{s'})}$$
(7.8)

where s is an index running on all the considered experimental signatures, i.e. the entries in the first column of Table 7.3. Some arguments about this choice can be found in the Appendix C. A threshold of $S_s > 5\%$ was set on the normalized score parameter S_s and three signatures were selected both for the $0\nu\beta\beta$ and $2\nu\beta\beta$ search, namely the 2A0-2B1, 2A2-2B3 and 3A0 listed in Table 7.3. The selected experimental signatures account for a 83.5% (86.9%) total score in the $0\nu\beta\beta$ ($2\nu\beta\beta$) search respectively.

The computation of the expected efficiency and number of background events is based on very strict selection cuts around the expected energy release in each scenario. For $0\nu\beta\beta$ decay scenarios the signal is monochromatic in all the involved crystals, so the signal region is expected to lie around a specific point in the M-dimensional space of coincident

$\begin{tabular}{ c c c c }\hline & {\bf S}_{2\nu}({\bf S}_{0\nu}) \\ & {\bf Signature} \\ \hline \end{array}$	Crystal 1 [keV]	Crystal 2 [keV]	Crystal 3 [keV]	Crystal 4 [keV]
39.8% (38.5%) 2A0-2B1	$egin{array}{c} etaeta + \gamma_{(A2)} \ etaeta + \gamma_{(B3)} \ 536 \ - \ 1270 \end{array}$	$ \begin{array}{c} \gamma_{(A1)} \\ \gamma_{(B1)} + \gamma_{(B2)} \\ 1257 \end{array} $		
2.3% (2.3%) 2A1-2B2-2C1	$\gamma_{(A1)} + \gamma_{(A2)} \ \gamma_{(B1)} + \gamma_{(B2)} + \gamma_{(B3)} \ \gamma_{(C1)} + \gamma_{(C2)} \ 1793$	$\begin{array}{c} \beta\beta\\ \beta\beta\\ \beta\beta\\ \beta\beta\\ 0 - 734 \end{array}$		
21.6% (24.7%) 2A2-2B3	$egin{array}{l} etaeta+\gamma_{(A1)}\ etaeta+\gamma_{(B1)}+\gamma_{(B2)}\ 1257$ - 1991	$\gamma_{(A2)}$ $\gamma_{(B3)}$ 536		
25.5% (20.3%) 3A0	$\begin{array}{c} \gamma_{(A1)} \\ 1257 \end{array}$	$\beta\beta$ 0 - 734	$\gamma_{(A2)} \ 536$	
2.4% (2.8%) 2B0-2C2	$egin{array}{l} etaeta+\gamma_{(B2)}+\gamma_{(B3)}\ etaeta+\gamma_{(C1)}\ 1122$ - 1856 \end{array}	$\begin{array}{c} \gamma_{(B1)} \\ \gamma_{(C2)} \\ 671 \end{array}$		
$\begin{array}{c c} 0.2\% \ (0.1\%) \\ 2B4 \end{array}$	$egin{array}{l} etaeta+\gamma_{(B2)}\ 586$ - 1320 \end{array}	$\gamma_{(B1)}+\gamma_{(B3)} onumber \ 1207$		
$ \begin{array}{c c} 1.4\% (2.0\%) \\ 2B5 \end{array} $	$egin{array}{l} etaeta+\gamma_{(B1)}+\gamma_{(B3)}\ 1207$ - 1941	$\begin{array}{c} \gamma_{(B2)} \\ 536 \end{array}$		
1.0% (0.9%) 2B6-2C0	$egin{array}{l} etaeta+\gamma_{(B1)}\ etaeta+\gamma_{(C2)}\ eta71$ - 1405	$egin{array}{l} \gamma_{(B2)} + \gamma_{(B3)} \ \gamma_{(C1)} \ 1122 \end{array}$		
1.1% (1.1%) 3B0	$egin{array}{l} etaeta+\gamma_{(B3)}\ 536$ - $1270 \end{array}$	$\begin{array}{c} \gamma_{(B1)} \\ 671 \end{array}$	$\frac{\gamma_{(B2)}}{586}$	
0.8% (0.5%) 3B1-3C0	$\gamma_{(B2)}+\gamma_{(B3)} \ \gamma_{(C1)} \ 1122$	$\begin{array}{c} \beta\beta\\ \beta\beta\\ 0 - 734 \end{array}$	$\begin{array}{c} \gamma_{(B1)} \\ \gamma_{(C2)} \\ 671 \end{array}$	
$\begin{array}{c c} 0.1\% (0.1\%) \\ 3B2 \end{array}$	$rac{\gamma_{(B1)}+\gamma_{(B2)}}{1257}$	$\beta\beta$ 0 - 734	$\gamma_{(B3)}$ 536	
1.3% (1.0%) 3B3	$egin{array}{l} etaeta+\gamma_{(B2)}\ 586$ - 1320	$\begin{array}{c} \gamma_{(B1)} \\ 671 \end{array}$	$\begin{array}{c} \gamma_{(B3)} \\ 536 \end{array}$	
$ \begin{array}{c c} 0.1\% & (0.1\%) \\ 3B4 \end{array} $	$\gamma_{(B1)}+\gamma_{(B3)} onumber 1207$	$\begin{array}{c} \beta\beta \\ 0 - 734 \end{array}$	$\begin{array}{c}\gamma_{(B2)}\\536\end{array}$	
1.5% (1.0%) 3B5	$egin{array}{c} etaeta+\gamma_{(B1)}\ eta 71$ - 1405	$\frac{\gamma_{(B2)}}{586}$	$\begin{array}{c}\gamma_{(B3)}\\536\end{array}$	
0.8% (0.6%) 4B0	$\frac{\beta\beta}{0-734}$	$\begin{array}{c} \gamma_{(B1)} \\ 671 \end{array}$	$\begin{array}{c}\gamma_{(B2)}\\586\end{array}$	$\begin{array}{c}\gamma_{(B3)}\\536\end{array}$

TABLE 7.3: In the above table all the independent combinations of fully contained energy release patterns following ¹³⁰Te decay to the first 0_2^+ excited nuclear final state are listed, one per row. Signatures originating from different de-excitation patterns that produce identical experimental signatures are merged together. For each signature, monochromatic ordered expected energy releases are specified, one per column, where appropriate. The final state particle combinations expected to generate each signature are listed as well. The $\beta\beta$ energy release is specified with an energy range corresponding to the endpoints of the $\beta\beta$ spectrum. For the $0\nu\beta\beta$ decay mode just the upper endpoint should be considered. energy releases. Resolution effects smear the signal, nevertheless most of it lies within one FWHM² of the energy resolution distribution from its expected position. The CUORE detectors have a resolution that depends on energy with a FWHM at $Q_{\beta\beta} = 2527$ keV of $\sim 8 \text{ keV}$ (Fig. 4.4), which improves at low energy. From Fig. 4.5 the improvement is a factor ~ 2 for energies below 1.2 MeV. It is possible, in principle, to smear simulations with any resolution function and to introduce whatever channel dependence. This choice would imply that the same model be used to fit simulated data, in order not to bias the result (e.g. efficiency evaluation). For the sake of simplicity simulations are processed with an energy independent Gaussian resolution smearing of 5 keV FWHM. A selection is enforced, on simulations, with a *box cut*, i.e. a selection interval for energy releases in each crystal, defined as

$$|E_i - Q_i| < 5 \text{ keV} \quad \text{where} \quad i = 1...M \tag{7.9}$$

where E_i is the reconstructed energy release in the ordered energy space³ and Q_i is the corresponding expected energy release. In $2\nu\beta\beta$ signatures the same selections apply except the one crystal where the energy release from the $\beta\beta$ is expected. Since the emitted neutrinos carry away a variable (on an event basis) amount of invisible⁴ energy, the expected energy release is not monochromatic. It is instead expected to vary from Q_j^{min} to Q_j^{max} where *j* indicates the channel where the $\beta\beta$ release their energy. For that channel, in each multiplet the following selection is applied

$$Q_j^{min} - 5 \operatorname{keV} < E_j < Q_j^{max} + 5 \operatorname{keV}$$

$$(7.10)$$

The efficiency ϵ_s for scenario s is simply defined as the ratio between number of selected simulated signal events and total simulated signal events corrected for the branching ratio of the de-excitation channel(s) populating s. The background B_s is computed in a similar way, counting the number of simulated background events for each source included in the background model used, weighting it with the corresponding normalization factor f(see Sec. 5.3) and summing the contributions from all sources. In order to minimize the inclusion of quantities related to data processing in the score function computation, the efficiency term ϵ_s does not include any of the efficiency contributions mentioned in Sec. 4.2 except the containment efficiency term better detailed in the following Sec. 7.3. This is acceptable for the computation of an approximate analytical score function, since the containment term by far dominates the overall efficiency (Tab. 7.13, 7.14). Once the downselection of signatures is performed, all efficiency contributions are included in the model (see Sec. 7.4) to extract an unbiased measurement of the double beta decay rate.

Signature	Cuts	Source
2A0 - 2B1	$1247 \mathrm{keV} < E^*_{min} < 1280 \mathrm{keV}$	$E_{min}: \ \gamma_{(A1)} \ / \ \gamma_{(B1)} + \gamma_{(B2)} \ (1257 { m keV})$
	$1247 \mathrm{keV} < E_{max} < 1280 \mathrm{keV}$	$E_{max}:~etaeta~(734~{ m keV})+\gamma_{(A1)}~/~\gamma_{(B3)}~(536~{ m keV})$
2A2 - 2B3	$523 \mathrm{keV} < E^*_{min} < 573 \mathrm{keV}$	$E_{min}: \ \gamma_{(A2)} \ / \ \gamma_{(B3)} \ (536 \ { m keV})$
	$1981 \text{ keV} < E_{max} < 2001 \text{ keV}$	$E_{max}: \beta\beta (734 \text{ keV})$ (1257 heV)
		$\gamma(A1) / \gamma(B1) + \gamma(B2) (1257 \text{ keV})$
3A0	$526 \text{ keV} < E_{min} < 546 \text{ keV}$	$\mid E_{min}: \ \gamma_{(A2)} \ / \ \gamma_{(B3)} \ (536 \ { m keV})$
	$700 \mathrm{keV} < E^*_{med} < 760 \mathrm{keV}$	$E_{med}: \beta\beta (734 \mathrm{keV})$
	$1247 \mathrm{keV} < E_{max} < 1267 \mathrm{keV}$	$E_{max}: \gamma_{(A1)} \; / \; \gamma_{(B1)} + \gamma_{(B2)} \; (1257 { m keV})$

TABLE 7.4: Selected experimental signatures for double beta decay search on the 0_2^+ excited state of ¹³⁰Xe in the $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ (bottom) channel are listed. For each signature the corresponding Regions Of Interest (ROI, i.e. the applied selection cuts) are listed in terms of the ordered energy releases $E_{min} \leq E_{med} \leq E_{max}$. The component that will be used for the fit is highlighted with a * superscript (see Sec. 7.4 for further details). Such variable is called *projected energy* because the candidate events are projected on the corresponding component where a 1-dimensional unbinned fit is performed. For each signature the partition of the secondaries expected to contribute are listed.

Signature	Cuts	Source
2A0 - 2B1	$620 \text{ keV} < F \rightarrow 1150 \text{ keV}$	E_{min} : $\beta\beta (0-734 \mathrm{keV})$
	$020 \text{ KeV} \leq L_{min} \leq 1150 \text{ KeV}$	$\gamma_{(A1)} \;/\; \gamma_{(B3)} \; (536 \ { m keV})$
	$1220 \text{ keV} < E^*_{max} < 1300 \text{ keV}$	$E_{max}: \ \gamma_{(A1)} \ / \ \gamma_{(B1)} + \gamma_{(B2)} \ (1257 \ { m keV})$
2A2 - 2B3	$523 \mathrm{keV} < E^*_{min} < 573 \mathrm{keV}$	$E_{min}: \gamma_{(A1)} / \gamma_{(B3)} \ (536 \ { m keV})$
	$1360 \ \mathrm{keV} < E_{max} < 1990 \ \mathrm{keV}$	$E_{max}: \beta\beta (0-734 \text{ keV})$
		$\gamma_{(A1)} \;/\; \gamma_{(B1)} + \gamma_{(B2)} \;(1257 { m keV})$
3A0	$400 \mathrm{keV} < E_{min} < 523 \mathrm{keV}$	$E_{min}: \beta\beta \ (0-734 \text{ keV})$
	$523 { m keV} < E^*_{med} < 573 { m keV}$	$E_{med}: \ \gamma_{(A2)} \ / \ \gamma_{(B3)} \ (536 \ { m keV})$
	$1779 \mathrm{keV} < E_{med} + E_{max} < 1807 \mathrm{keV}$	$E_{max}: \ \gamma_{(A1)} \ / \ \gamma_{(B1)} + \gamma_{(B2)} \ (1257 \ { m keV})$
7.3 Event selection and efficiency contributions

The most relevant experimental signatures from the point of view of the sensitivity are included in the analysis and are listed in Table 7.4 for the $0\nu\beta\beta$ and $2\nu\beta\beta$ decay search respectively. The selection cuts are tuned, for each signature, on Monte Carlo simulations taking simultaneously into account the location and spectral features of the specific background components that populate each of the selections as well as the spectral shape of the signal (e.g. the $\beta\beta$ energy spectrum is not uniform in the 0-734 keV energy range, for this reason the impact of a modification of the selection cuts on the selection efficiency is non-linear) and all relevant spectral features related to data processing such as the PSA threshold (see Sec. 3.3.6).

Each of the experimental signatures listed in Table 7.4 is populated by signal counts originating from one or more partitions (i.e. combination of secondaries from a specific deexcitation pattern that simultaneously hit a given number of detectors). In the following a detailed description of all the signal efficiency contributions will be given, along with the techniques used to evaluate each component.

Containment efficiency The containment efficiency is defined for each signature, it represents the overall probability that a nuclear decay of ¹³⁰Te to the 0_2^+ state of ¹³⁰Xe is reconstructed within the selection cuts of the signature, neglecting all other possible sources of inefficiency. This component can be decomposed in as many contributions as there are de-excitation patterns p populating the same scenario s

$$\epsilon_s^{(cont)} = \sum_p BR_p \cdot \frac{\left[N_{MC}^{(sel)}\right]_p^{(s)}}{\left[N_{MC}^{(tot)}\right]_p}$$
(7.11)

where BR_p is the branching ratio of pattern p, $[N_{MC}^{(sel)}]_p^{(s)}$ and $[N_{MC}^{(tot)}]_p$ are respectively the selected and total number of simulated decays in the de-excitation channel of interest. The number of selected MC counts is evaluated with a binned maximum likelihood fit to simulated events from each involved de-excitation pattern. The energy releases belonging to each event are ordered decreasingly, selected according to Table 7.4, projected on one of the energy axes in order to obtain a 1-dimensional histogram and then fit with a simplified model for signal and background. The variable selected for projection is called *projected energy*, and is chosen taking into account that a better sensitivity can be achieved at lower energy, where the resolution function is narrower. An example is shown in Figure 7.5. In this efficiency evaluation, the signal is modeled as a single Gaussian distribution, while the background is the superposition of a linear component on a Gaussian error function centered at the expected signal position and with the same width, modelling possible

 $^{^{2}}$ Full Width Half Maximum, a measure of the width of a distribution. In this case is used to quantify the energy resolution effect.

³The energy releases of each M-channels multiplet are ordered in descending order so that $E_i > E_{i+1}$. ⁴Undetected.



FIG. 7.5: Distribution of simulated $2\nu\beta\beta$ decay events from pattern 2A0 for containment efficiency computation. The binned maximum likelihood fit is performed with a gaussian (signal) over a linear background plus a Compton shoulder on the left of the peak, described by a Gaussian error function.

Compton scatterings of the signal photons. Modelling the signal with a simple Gaussian rather than the complex line-shape of the CUORE detectors does not bias the evaluation of the signal efficiency because the signal simulations are processed with a single Gaussian smearing as well.

Other efficiency contributions Additional efficiency contributions come from the probability of correctly detecting, in each of the involved channels, a signal energy release and from the probability of correctly assigning the multiplicity to the event due to accidental simultaneous energy releases. The first efficiency term will be called ϵ_{cut} , where *cut* is referred to the data processing cuts needed to select triggered events that pass the *base* and *PSA* cuts (see Sec. 3.3), the other ϵ_{acc} . The ϵ_{cut} term must be raised to the Mth power because it models channel-related efficiencies and a multiplet is selected if and only if all of the involved channels pass the selection cuts. Summarizing, the signal efficiency for signature *s* is

$$\epsilon_s = \left[\sum_p BR_p \cdot \frac{\left[N_{MC}^{(sel)}\right]_p^{(s)}}{\left[N_{MC}^{(tot)}\right]_p}\right] \epsilon_{cut}^M \epsilon_{acc}$$
(7.12)

Since the cut efficiency and accidental term are evaluated separately for each dataset, this efficiency term must be thought of as the signal efficiency for signature s for a specified dataset.

7.4 Model

In this section details about the method designed to perform Bayesian inference on the rate of both neutrino-less and $2\nu\beta\beta$ decay of ¹³⁰Te to the 0_2^+ excited state of ¹³⁰Xe will be explained. The mathematical and computational basics are detailed in Appendix B. The inference (i.e. the Bayesian fit) is performed separately for either process, simultaneously for all involved signatures in the same process. For the sake of readability the nomenclature used in this section will be briefly outlined here. For further reference the reader can refer to previous sections. The word *process* will be used to refer to $0\nu\beta\beta$ or $2\nu\beta\beta$ gs-es transition, unless otherwise stated. Each process can be followed, with a different branching ratio, by a different *de-excitation pattern*, i.e. a different set of secondary γ emissions (see Fig. 7.1). For a given process and de-excitation pattern many *partitions* derive, i.e. the energy released by final state particles can be combined differently and involve simultaneously a different number M of crystals. For each process a number of independent experimental signatures are populated. Each signature is, in turn, populated by one or more partition of the secondaries from one or more de-excitation patterns, as detailed in Table 7.3 and 7.4. Every multiplet of multiplicity M can be represented as a point \vec{E}_{ev} in a M-dimensional space of energy releases. The energy releases are ordered so that $E_i > E_{i+1} \forall i = 1, .., M-1$. One of the components of the \vec{E}_{ev} vector is selected to perform the fit, and it is referred to as projected energy E_{ev} as detailed, for each signature, in Table 7.4. In the following equations, the subscripts s, ch, ds will be used to refer to a specific signature, channel, or dataset respectively.

An unbinned extended maximum a posteriori probability fit is implemented, where the likelihood can be decomposed, for each scenario and dataset, as follows

$$\log \mathcal{L}_{s,ds} = -(\lambda_{S\,s,ds} + \lambda_{B\,s,ds}) + \sum_{ev \in (s,ds)} \log \left[\lambda_{S\,s,ds} \frac{M\Delta t_{ch}}{M\Delta t_{ds}} f_S(E_{ev}) + \lambda_{B\,s,ds} \frac{M\Delta t_{ch}}{M\Delta t_{ds}} f_B(E_{ev}) \right]$$
(7.13)

where the λ_S and λ_B terms are the expected number of signal and background events respectively, $M\Delta t$ is the exposure, f_S and f_B are the normalized signal and background probability density functions. They depend just on the projected energy variable E_{ev} . The number of signal and background counts can be written as

$$\lambda_{S\,s,ds} = \Gamma^{(p)}_{\beta\beta} [\mathrm{yr}^{-1}] \left[\frac{N_A 10^3 \eta (^{130} \mathrm{Te})}{m (\mathrm{TeO}_2) [g/mol]} \right] \epsilon^{(cont)}_s (M\Delta t)_{ds} [\mathrm{kg} \cdot \mathrm{yr}] (\epsilon_{cut})^M_{ds} (\epsilon_{acc})_{ds}$$
(7.14)

$$\lambda_{B\,s,ds} = \mathrm{BI}_s \, (M\Delta t)_{ds} \, (E_{max}^{proj} - E_{min}^{proj})_s \tag{7.15}$$

where $\Gamma_{\beta\beta}^{(p)}$ is the decay rate of process p, N_A the Avogadro constant, η the isotopic abundance of ¹³⁰Te in natural tellurium, $m(\text{TeO}_2)$ the molecular mass of a tellurium dioxide

molecule. The combined log-likelihood then reads

$$\log \mathcal{L}(\mathcal{D}|\mathcal{H}_{S+B}) = \sum_{s,ds} \log \mathcal{L}_{s,ds}$$
(7.16)

where H_{S+B} indicates that the likelihood is written in the signal-plus-background model hypothesis, i.e. that the existence of the process of interest is assumed. The signal term $f_S(E_{ev})$ models the expected shape of a monochromatic peak of energy Q_s in the channeldataset pair the energy E_{ev} was released in, according to Eq. 4.15. A correction for the bias in the energy scale reconstruction is implemented together with the resolution dependence on energy, as detailed in Sec. 4.3.2. The background term $f_B(E_{ev})$ is parameterized as

$$f_B(E_{ev}) = \frac{1}{\Delta E} \left[1 + m_s \left(E_{ev} - E_0^{(s)} \right) \right]$$
(7.17)

where $\Delta E = E_{max}^{proj} - E_{mins}^{proj}$ is the width of the region of interest, $E_0^{(s)}$ is the center of it and m_s describes the slope of the background for signature s.

The prior probability distribution functions for all the fit parameters will be discussed in detail together with systematics in Sec. 7.6. It is worth to mention here that the Γ_p parameter describing the rate of the process under investigation is never allowed to have negative values for two reasons

- negative rates are unphysical and do not describe any situation included in H_{S+B}, for this reason a Bayesian analysis must exclude such values from the domain of its parameters
- mathematically a probability distribution function can never assume negative values, because probability is positive definite. Since a background free signature is present in both $0\nu\beta\beta$ and $2\nu\beta\beta$ decay searches, it is not possible to allow the rate parameter to go negative because the corresponding pdf would soon enter a meaningless regime. This is not the case for peak searches where a high background is expected and negative values of the decay rate still produce valid probability distribution functions for the spectrum of experimental counts.

7.5 Blinding and sensitivity computation

After the selection cuts are defined, all the efficiency contributions are evaluated either from data or simulations and the functional form of the H_{S+B} model is fixed, the final result could be extracted directly from the data, even disregarding the background prediction from simulations. Nonetheless it is important to evaluate the sensitivity of the analysis technique to the process of interest, reducing as much as possible the experimenter's bias and possibly refining the selection or the model itself before extracting the final result. In order to minimize the experimenter's bias, to validate the model and compute the analysis sensitivity, a blinding method is implemented (see [134] for a review). The chosen approach is the injection of a large enough - yet unknown to the analyst - number of simulated events in the data in order to mimic the expected signal. In this way the spectral features of the data in the signal region are effectively hidden while the background region remains unchanged. This allows to extract a measurement of the amount of expected background in the signal region and to compute a reliable sensitivity estimate. In the following, details about the implementation of the blinding technique will be given, the results of a blinded fit in both $0\nu\beta\beta$ and $2\nu\beta\beta$ decay channels will be presented and discussed. Finally the methods and results of the computation of a 90 % C.I. bayesian limit setting sensitivity will be outlined.

Blinding technique In order to inject simulated signal events for a given process, a range of allowed simulated rates is defined according to the current available upper limit $\Gamma_p^{90\%}$ on the rate of the process

$$\Gamma_p^{inj} \in \left[\Gamma_p^{min}, \Gamma_p^{max}\right] \tag{7.18}$$

so that $\Gamma_p^{min} > \Gamma_p^{90\%}$. This choice is appropriate because the expected sensitivity [135] is greater than the one of the CUORE-0 and CUORICINO experiments, which placed the $\Gamma_n^{90\%}$ limit. For this reason such a blinding peak is always visible. Moreover it lies in a region of the parameter space that was already explored, for this reason the number of signal events that could be present in the data would be much smaller than the ones coming from the blinding peak, and would be hidden beneath it. Each process generates multiple experimental signatures and each candidate event belonging to a signature is identified by the vector of ordered simultaneous energy releases \vec{E}_{ev} . Since every such vector is projected on one of its components, depending on the signature, and the response of different channels is uncorrelated, the event injection can be performed taking care just of correctly reproducing the expected detector response on the projected component. The other components can be distributed in any way, since the information on those components will be projected away, provided that the generated \vec{E} falls within the selection cuts of the signature it was generated from. Blinded data are then produced extracting randomly Γ_p^{inj} from a uniform distribution in the domain specified by Eq. 7.18. Then the expected number of counts in each of the derived signatures is computed, according to known efficiencies and exposures, for each dataset. The actual number of generated signal events is extracted from a Poisson distibution centered on the expected number of counts. Each generated signal event is then randomly assigned a channel, according to its exposure within the considered dataset. Finally the projected energy of the signal event is generated according to the detector response function $f_S(E_{ev}|Q_s)$ centered at the expected position Q_s of the monochromatic energy release in the projected energy space (see Eq. 4.15).

For both $0\nu\beta\beta$ and $2\nu\beta\beta$ decay, blinded data were produced with a rate comprised between

and then fit to extract data driven estimates of the background indices. The result of the blinded fits is shown in Fig. 7.7 and 7.8 for the $0\nu\beta\beta$ and $2\nu\beta\beta$ decay process respectively, together with some marginalized posterior distributions. For neutrino-less double beta decay signatures the blind analysis shows compatibility with zero background for the 3A0 signature, indication for a non-null background is found in the other signatures. In the $2\nu\beta\beta$ data instead the background is non-null for both the 2A0 – 2B1 and 2A2 – 2B3 signatures, while no background is expected for the 3A0 one.

Limit setting sensitivity Such distributions are used as input for the generation of background-only toy Monte Carlo simulations (ToyMC) and, eventually, to extract the median expected 90% C.I. upper limit on the decay rates reported in Tab. 7.6. A background-only ToyMC simulation is an ensemble of simulated datasets, according to the following procedure which is iterated N_{toy} times, to produce the same number of ToyMC ensembles

- 1. a set of signatures is defined, together with the multiplicity and cuts in the ordered energy variables that identify candidate events;
- 2. for each signature a functional form for the background pdf is defined (either flat or linear);
- for each signature and independent variable describing the background (background index BI and/or background slope m) a value is randomly extracted from the posterior pdf of the corresponding blinded fit;
- 4. the number of expected background events for each signature and dataset is computed according to Eq. 7.15;
- 5. the actual number of background events is randomly extracted from a Poisson distribution with expectation value equal to the number of expected counts;
- 6. each simulated ToyMC event is stored as a vector of ordered energy releases \vec{E}_{ev} and related channels \vec{ch}_{ev} , where the channels are randomly extracted from the active channels of each dataset according to their exposure in the real data, while the energies are generated according to the selected shape of the background pdf computed with the parameters (e.g. background index) generated according to the posterior pdfs obtained with the blinded fit to the data.

Each ToyMC is then fit with the signal-plus-background model H_{S+B} and an upper limit for the decay rate parameter is extracted as the 90% quantile of the posterior pdf. The distribution of such limits is shown in Fig. 7.9 for both the $0\nu\beta\beta$ and $2\nu\beta\beta$ decay process. The blind analysis was repeated for two different processing paths on the same data that differ just on the settings used to define coincident multiplets called *narrow* and *wide* respectively. The difference will be detailed in the following paragraph. The expected limit setting sensitivity distribution correctly weights the background estimation, the different signal efficiency and the resolution effects in both cases. It was found that the *narrow* and

Observable	$\mathbf{Mode} \pm \mathbf{Small.Int.}$	Units
	$({f Mean}\pm{f Stdev})$	
Γ_{etaeta}	$2.847\substack{+0.087 \\ -0.094}$	$10^{-23} [\mathrm{vr}^{-1}]$
	(2.845 ± 0.090)	10 [91]
$\mathrm{BI}_{\mathrm{2A0-2B1}}$	$3.987^{+4.052}_{-2.422}$	10 ⁻⁴ [cts/keV/kg/vr]
	(6.105 ± 3.603)	10 [000/110 //118/31]
$\mathrm{BI}_{\mathrm{2A2-2B3}}$	$1.829^{+1.658}_{-1.045}$	10 ⁻⁴ [cts/keV/kg/vr]
	(2.704 ± 1.540)	[, /0, / -]
BI _{3A0}	$0.291\substack{+9.880 \\ -0.291}$	10 ⁻⁵ [cts/keV/kg/vr]
	(8.682 ± 8.346)	
$S_{1/2}$	4.63 ± 1.25	$10^{24} [m yr]$

TABLE 7.6: Results of the blinded fit to $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ (bottom) candidate events in the signatures of Table 7.4. For each parameter the mode and smallest interval with 68% probability content of the corresponding marginalized posterior distribution are reported together with the mean and standard deviation. The $S_{1/2}$ parameter indicates the median expected sensitivity for limit setting at 90% C.I. on the $T_{1/2}$ parameter together with the MAD of its distribution.

Observable	$\mathbf{Mode} \pm \mathbf{Small.Int.}$	\mathbf{Units}
	$({f Mean} \pm {f Stdev})$	
$\Gamma_{\beta\beta}$	$5.097\substack{+0.146\\-0.149}$	$10^{-23} [vr^{-1}]$
	(5.099 ± 0.147)	
BI_0	$3.232\substack{+0.418\\-0.365}$	10^{-3} [cts/keV/kg/vr]
	$\bigl(3.288\pm0.393\bigr)$	[,,
BI_1	$4.009_{-0.619}^{+0.537}$	10 ⁻³ [cts/keV/kg/yr]
	(4.020 ± 0.580)	
BI_2	$0.129\substack{+7.806 \\ -0.129}$	10 ⁻⁵ [cts/keV/kg/yr]
	(6.864 ± 6.829)	
m ₀	$-5.641^{+4.268}_{-4.402}$	$10^{-3} [1/\text{keV}]$
	(-5.467 ± 4.258)	
$S_{1/2}$	4.63 ± 1.25	$10^{24} [yr]$

wide processings are equivalent for the $0\nu\beta\beta$ decay search, while for the $2\nu\beta\beta$ the wide settings yield a median expected limit setting sensitivity ~ 20 % higher than the *narrow* ones. For this reason the data processed with the *wide* settings were unblinded and used to produce the final result.



FIG. 7.7: Top: simultaneous blinded fit result (blue solid) of $0\nu\beta\beta$ candidate events in the signatures of Table 7.4. The background is modeled as a uniform distribution, the background index is compatible with 0 in all signatures. An indication of non-vanishing background was found in the 2A2-2B3 signature. Bottom: posterior pdfs. The subscripts 0, 1, 2 refer to the 2A0-2B1, 2A2-2B3 and 3A0 signature respectively.



FIG. 7.8: Top: simultaneous blinded fit result (blue solid) of $2\nu\beta\beta$ candidate events in the signatures of Table 7.4. The background is modeled as a uniform distribution in all signatures except the 2A0-2B1, where a linear component was introduced. The background index is compatible with 0 just in signature 3A0. An indication of non-vanishing background slope was found in the 2A0-2B1 signature. Bottom: posterior pdfs. The subscripts 0, 1, 2 refer to the 2A0-2B1, 2A2-2B3 and 3A0 signature respectively.

Narrow and *wide* coincidences Two different settings in the data processing were used with regard to coincidences. In the text they are referred to as wide and narrow, where the first one accepts coincidences regardless of the distance between crystals, provided that the events are reconstructed within a time window of $\pm 5 \,\mathrm{ms}$. The other implements a looser cut in terms of time window $(\pm 30 \text{ ms})$ while compensating with a tight⁵ requirement on the distance (150 mm). Looser requirements in terms of coincidence time window are needed to enhance the reconstruction efficiency for surface α decays. Since this analysis handles just β/γ energy releases it has no effect on the signal cut efficiency. Nonetheless the probability of accidental coincidences is affected. For this reason a looser cut on the time coincidence window is compensated with a tighter one on distance. Evaluating the effect on the background selection efficiency is not trivial because it strongly depends on the distribution of background contaminants in the crystal and surrounding inert materials. Even simulations are not completely reliable on this quantity because of the assumptions made on such distributions of radioactive contaminants. For this reason a blinded fit is crucial as input for the sensitivity evaluation, because it allows an unbiased data driven evaluation of the expected background in the signal region.

7.6 Systematic uncertainties and final results

The evaluation of results on ToyMC is affected by the large number of fits that are needed in order to sample e.g. the expected limit setting sensitivity. For this reason it is not computationally affordable to perform such fits nor with a very high statistics in terms of MCMC steps nor with a large number of parameters. The former impacts linearly on computation time and affects the resolution on the posterior pdfs. The latter scales differently in terms of computation time, because the problem's complexity does not depend linearly on its dimensionality (i.e. number of parameters). Increasing the dimensionality allows to plug in prior information about nuisance parameters. In this way it is possible to handle systematic uncertainties. For these reasons the fits to ToyMC were performed fixing all nuisance parameters to their best available estimate. This is what will be called *stat-only* fit, because it correctly takes into account uncertainties coming from the limited statistics in the data, but it does not consider other contributions. To evaluate blinded fit results, instead, it is possible to afford a more complex problem since just one fit is needed instead of $N_{toy} \sim 10^4$. Nonetheless, since the unblinded fit result (shown in Fig. 7.11) is compared to the expected distribution of stat-only results obtained from ToyMC fits, it was chosen to increase as much as possible the statistics (number of samples) of the MCMC, without including systematic uncertainties, in order to compare homogeneous quantities. In the stat-only fit to unblinded data for both $0\nu\beta\beta$ and $2\nu\beta\beta$ decay to the first 0_2^+ excited state of ¹³⁰Xe no evidence for signal was found. The stat-only 90% C.I. marginalized lower limit on the $T_{1/2}$ parameter lies within the $\pm \; 2 \; \sigma$ (gaussian) central interval from the median expected limit setting sensitivity. This is an indication that the

⁵The clearance between CUORE crystal does not allow neighbor crystals to be closer than 86 mm.



Expected limit setting sensitivity - Median: 4.63e+24 [yr] MAD: 1.25e+24 [yr]

FIG. 7.9: Distribution of 90% C.I. marginalized upper limits on $T_{1/2} = \log 2/\Gamma$ for $0\nu\beta\beta$ decay (top) and $2\nu\beta\beta$ decay (bottom) obtained from Toy MC simulations. The median of each distribution is highlighted (black dashed line), the upper limit obtained with a fit to the unblinded data is shown (red solid line).

Observable	$Mode \pm Small.Int.$	Units
	$({f Mean} \pm {f Stdev})$	
$\Gamma_{\beta\beta}$	$2.745^{+5.037}_{-2.707}$	10^{-26} [vr ⁻¹]
	(6.370 ± 4.846)	IU [yI]
$\mathrm{BI}_{\mathrm{2A0-2B1}}$	$1.336 \ ^{+1.474}_{-0.900}$	10^{-4} [cts/keV/kg/yr]
	(2.149 ± 1.378)	
$\mathrm{BI}_{\mathrm{2A2-2B3}}$	$2.136\ ^{+1.199}_{-1.031}$	10^{-4} [cts/keV/kg/yr]
	(2.563 ± 1.198)	
BI _{3A0}	$0.080 {}^{+6.644}_{-0.080}$	10^{-5} [cts/keV/kg/yr]
	(5.673 ± 5.422)	
$T_{1/2}$	> 5.373	$10^{24} [yr]$

model used to describe the data is appropriate. The results of the fit are summarized in Table 7.10 for both the neutrino-less and $2\nu\beta\beta$ decay processes.

TABLE 7.10: Results of the unblinded stat-only fit to $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ (bottom) candidate events in the signatures of Table 7.4. For each parameter the mode and smallest interval with 68% probability content of the corresponding marginalized posterior distribution are reported together with the mean and standard deviation. The result on the last line indicates the lower limit set at 90% C.I. on the marginalized posterior pdf for the T_{1/2} parameter.

Observable	$\mathbf{Mode} \pm \mathbf{Small.Int.}$	Units
	$({f Mean} \pm {f Stdev})$	
$\Gamma_{\beta\beta}$	$2.174 {}^{+2.305}_{-1.895}$	$10^{-25} [vr^{-1}]$
	(3.328 ± 2.176)	10 [51]
$\mathrm{BI}_{\mathrm{2A0-2B1}}$	$2.989 {}^{+0.359}_{-0.309}$	10 ⁻³ [cts/keV/kg/vr]
	$\bigl(3.012\pm0.332\bigr)$	
$\mathrm{BI}_{\mathrm{2A2-2B3}}$	$4.193 {}^{+0.537}_{-0.455}$	10^{-3} [cts/keV/kg/vr]
	(4.234 ± 0.495)	
BI _{3A0}	$0.104 {}^{+6.191}_{-0.104}$	10 ⁻⁵ [cts/keV/kg/vr]
	(5.366 ± 5.358)	[,, /0, J -]
$m_{2A0-2B1}$	$-5.447_{-4.194}^{+4.416}$	$10^{-3} [1/\text{keV}]$
	(-5.168 ± 4.229)	
$T_{1/2}$	> 1.097	$10^{24} [yr]$

In order to evaluate the effect of systematic uncertainties on the fit result and, ultimately, on the lower limit on the $T_{1/2}$ parameter, the most appropriate method in Bayesian analyses is the insertion of a prior probability distribution function on all the parameters



FIG. 7.11: Best fit (blue solid line) corresponding to the global mode of the posterior pdf for unblinded data in the $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ decay channel. The signal component is shown (blue dashed line). The model corresponding to the marginalized 90% C.I. upper limit on the decay rate is overlayed (red solid line). The data are binned for plotting purposes, and Poisson uncertainties are shown as error bars. Systematic uncertainties are not included.

that are expected to contribute, i.e. those among all included parameters needed to build the H_{S+B} model that are known with the highest uncertainty. A summary of such parameters is given in Table 7.12. The detector response function (Sec. 4.3) is parameterized for each channel and dataset as the superposition of three Gaussian components with the same width. Two of them model non-gaussianity of the left and rigth tail of the energy response function and feature a constant relative amplitude with respect to the main peak. For each dataset an average bias correction is implemented (i.e. the position of the response function is slightly shifted with respect to the nominal energy) as well as a dependence of the variance with energy. Each of these contributions is evaluated from data and parameterized with a second order polynomial. The correlation among the polynomial coefficients are taken into account with the same prior implemented in the search for $0\nu\beta\beta$ decay in the gs-gs transition (Eq. 6.13).

Each of the efficiency contributions included in the model is known with some uncertainty. In particular the uncertainty on cut efficiency comes from the limited statistics used to compute it and is modeled with an gaussian prior for each dataset with the domain restricted to the [0, 1] range. An additional uncertainty is due to the evaluation of the PSA cut efficiency, which is one of the contributions to the cut efficiency. The PSA efficiency can, in fact, be evaluated with two independent methods (Sec 4.2.4) that rely on single site $\epsilon_{\text{PSA}}^{\text{M1}}$ and double site $\epsilon_{\text{PSA}}^{\text{M2}}$ events. For reasons yet under investigation a tension is observed between the two methods, hence an additional systematic is included. Assuming the discrepancy between the two methods can be described with a uniformly distributed shift of the reconstructed PSA efficiency with respect to the expected value of its distribution, the *nominal* PSA efficiency is obtained as the average

$$\hat{\epsilon}_{PSA} = \frac{\epsilon_{PSA}^{M1} + \epsilon_{PSA}^{M2}}{2} \tag{7.20}$$

while the difference of the two measurements is used to sample the variance of the uniform distribution

$$\hat{\sigma}(\epsilon_{PSA}) = \frac{|\epsilon_{PSA}^{M1} - \epsilon_{PSA}^{M2}|}{2} \tag{7.21}$$

and from that value to reconstruct the range of the uniformly distributed parameter used to describe the systematic uncertainty as

$$\Delta \epsilon_{PSA} = \sqrt{12} \cdot \hat{\sigma}(\epsilon_{PSA}) \tag{7.22}$$

For this reason the cut efficiency, a constant in the stat-only fit, is substituted by

$$\epsilon_{cut} \longrightarrow \epsilon_{cut} + \delta_{PSA} \tag{7.23}$$

where δ_{PSA} is uniformly distributed in a $\Delta \epsilon_{PSA}$ wide interval with null expectation value. Additionally, the containment efficiency term (i.e. the probability of selecting signal events applying the cuts on the Multiplicity and Energy variables on each experimental signature) is affected by two independent sources of uncertainty. The first one comes from the limited available statistics in signal Monte Carlo simulations used to evaluate it. This is reflected by the uncertainty on the amplitude of the peak in e.g. Fig. 7.5, and is taken into account with a Gaussian prior. Moreover a term due to the accuracy of the Geant4 description of Compton scattering and, in general, of the electromagnetic processes beneath the simulated the energy release for each generated decay is added in quadrature. Its effect is < 1 % [136] for energies above 100 keV, < 5 % for energies 10 keV < E < 100 keV. The statistical one is of the same order of magnitude, ~ 1 %. For this reason a relative uncertainty of 5% is assumed to model both the effects on the containment efficiency. Eventually the uncertainty on the isotopic abundance of ¹³⁰Te in natural Tellurium is considered according to the results reported in [137] and modeled with another Gaussian prior on the corresponding parameter.

The impact of systematic uncertainties on the fit results is summarized in Tab. 7.12. The fit to the data was repeatedly performed inserting each time, in addition to the parameters of the stat-only fit, a different sub-set of nuisance parameters. Each fit of this kind defines a new model for the data, and is represented by a row in Table 7.12. Eventually the fit is performed adding all the nuisance parameters at the same time (combined), hence allowing the Markov Chain Monte Carlo to explore the correlations between different sub-sets. To quantify the effect of the introduction of a model (i.e. a set of nuisance parameters and the corresponding prior pdfs) for each contribution to systematic uncertainties, the following statistical indicators are shown

- $\hat{\hat{\Gamma}}$, projection of the global mode of the joint posterior pdf $P(\Gamma, \vec{\theta} | D)$ on the rate axis
- $\hat{\Gamma}$, mode of the marginalized posterior pdf for the rate parameter $P(\Gamma | \mathcal{D}) = \int P(\Gamma, \vec{\theta} | \mathcal{D}) d\vec{\theta}$
- Δ_{50} , width ⁶ of the smallest interval with $\geq 50\%$ probability content in the marginalized pdf P($\Gamma | \mathcal{D}$)
- $T_{1/2}^{90}$ lower limit at 90% Bayesian Credibility Interval (C.I.) on the marginalized distribution of the $T_{1/2} \doteq \log 2/\Gamma$ parameter

Since the domain of the rate parameter forces $\Gamma > 0$ and there is no evidence for signal, the marginalized and global mode for Γ are always included in the 50% smallest interval, very close to 0. For this reason the relative variation reported in Table 7.12 as

$$\frac{\hat{\hat{\Gamma}}_{\text{syst}} - \hat{\hat{\Gamma}}_{\text{stat-only}}}{\hat{\hat{\Gamma}}_{\text{stat-only}}}$$
(7.24)

can fluctuate significantly, and large positive fluctuations are much more likely than negative ones. The width of the smallest interval with 50% probability content reflects instead the spread of samples within the distribution (analogue of the second central moment, the

⁶In mathematical terms $\sup_{I_{50}} |x_i - x_j| \quad \forall x_i, x_j \in I_{50}$ where I_{50} is the smallest interval with $\geq 50\%$ probability content of the marginalized pdf.

variance). Smaller widths correspond to *smaller uncertainty* though, in this case, they do not guarantee a tighter upper limit on the rate parameter because that depends on the position of the mode as well. Due to the definition

$$T_{1/2} \doteq \frac{\log(2)}{\Gamma} \tag{7.25}$$

a smaller (more stringent) upper limit on Γ corresponds to a larger (more stringent) lower limit on T_{1/2}. All relative variations of parameters in Table 7.12 are reported up to the $\%_0$ level because numerical uncertainties do not allow a more precise determination.

The unusual situation in which the insertion and modeling of systematic uncertainties yields a more stringent upper limit needs further explanation. Fixing the shape and position of the detector response function is prone to mis-interpretation of background fluctuations at the Q value as signal counts. On the other hand the insertion of a model that allows slight modifications of the detector response function position and resolution reduces the significance of such fluctuations after marginalization is performed. In addition it is possible to categorize nuisance parameters as *informative* and *non-informative*. The first ones (e.g. background indices) are parameters the data are able to provide information about, for the others (e.g. efficiency) all the available information is stored in the prior and no knowledge update is expected from the selected data. A cross check for unexpected fit behaviors can be performed inspecting the difference between the prior and marginalized posterior of all non-informative nuisance parameters. If a significant change were observed between the input prior distribution and the marginalized posterior, it would point to a mistake either in the definition or implementation of the model used to perform the inference. No such change was observed for any non-informative parameter. This provides confidence that the obtained result is correct. Nonetheless further investigation is in progress to disentangle the contribution of the energy bias correction from the one coming from the uncertainty of the resolution interpolation as a function of energy. In addition a different approach to the systematic uncertainty due to the PSA efficiency is under investigation, following the sceptical combination of experimental results mentioned in [138].

Including contributions from all sources of systematic uncertainty listed in Table 7.12 no signal is observed in either decay mode and the following Bayesian lower bounds on the corresponding half life parameters are set

$$^{0^+}_2(T^{0\nu}_{1/2}) > 5.9 \cdot 10^{24} \, [yr] \quad 90\% C.I.$$
 (7.26)

$${}^{0^+_2}(T^{2\nu}_{1/2}) > 1.3 \cdot 10^{24} \,[yr] \quad 90\% C.I.$$
 (7.27)

Source distribution	$\hat{\hat{\Gamma}}_{0 u}$	$\hat{\Gamma}_{0 u}$	Δ_{50}	$T_{1/2}^{90}$
Stat-only fit	4.278	2.587	5.264	5.373
uniform	$10^{-26} [1/yr]$	$10^{-26} [1/yr]$	$10^{-25} [1/yr]$	$10^{24} [yr]$
Detector response	+66.2%	-3.5%	+35.1%	+10.1%
multivariate				
Cut efficiency (stat)	+20.4%	+3.5%	+23.7%	-0.1%
gaussian				
PSA efficiency (syst)	+5.2%	+5.8%	+17.8%	-0.7%
uniform				
Accidental coincidences	-12.8%	+6.0%	+26.7%	-0.4%
gaussian				
Containment efficiency	-2.5%	+8.8%	+18.4%	-0.6%
gaussian				
Isotopic abundance	-3.0%	+7.3%	+34.9%	-0.1%
gaussian				
Combined	+43.3%	-11.5%	+5.7%	+10.1%
multivariate				

TABLE 7.12: Summary table of the effect of the introduction of each systematic uncertainty as the corresponding set of nuisance parameters for $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ decay search (bottom). The effect of the introduction of each contribution to systematic uncertainty and the stat-only fit is shown as a % difference with respect to the stat-only fit. The relevant parameters are the position of the the global mode $\hat{\Gamma}$ of the rate parameter, the marginalized mode $\hat{\Gamma}$, the width of the 50% smallest interval of the marginalized posterior pdf Δ_{50} , and the 90 % C.I. lower limit on $T_{1/2} \doteq \log 2/\Gamma$.

Source	Ê.	Êa	Δεο	$T_{1/2}^{90}$
distribution	$1 2\nu$	± 2ν		-1/2
Stat-only fit	2.107	2.448	2.325	1.097
uniform	$10^{-25}[1/yr]$	$10^{-25}[1/yr]$	$10^{-24}[1/yr]$	$10^{24}[yr]$
Detector response	-15.4%	-16.0%	-23.9%	+17.6%
multivariate				
Cut efficiency (stat)	-6.3%	+0.4%	-7.5%	< 0.1%
gaussian				
PSA efficiency (syst)	+2.4%	-5.3%	-9.6%	-0.2%
uniform				
Accidental coincidences	+8.3%	-10.7%	-6.2%	+0.1%
gaussian				
Containment efficiency	+32.7%	-9.1%	-4.5%	-0.4%
gaussian				
Isotopic abundance	-16.3%	-2.9%	+4.0%	-0.2%
gaussian				
Combined	-5.7%	-23.9%	-26.0%	+17.2%
multivariate				

Observable	${\bf Mean} \pm {\bf Stdev}$	\mathbf{Units}
$\Gamma_{\beta\beta}$	2.875 ± 1.841	$10^{-25} [\mathrm{yr}^{-1}]$
$\mathrm{BI}_{\mathrm{2A0-2B1}}$	3.026 ± 0.330	$10^{-3} [\mathrm{cts/keV/kg/yr}]$
$\mathrm{BI}_{\mathrm{2A2-2B3}}$	4.273 ± 0.489	$10^{-3} [\mathrm{cts/keV/kg/yr}]$
$\mathrm{BI}_{\mathrm{3A0}}$	5.394 ± 5.424	$10^{-5} [\mathrm{cts/keV/kg/yr}]$
$m_{2A0-2B1}$	-5.178 ± 4.222	$10^{-3} [1/\text{keV}]$
$\epsilon_{ m cut}^{ m ds3519}$	8.579 ± 0.037	10^{-1} []
$\epsilon_{ m cut}^{ m ds3522}$	9.209 ± 0.025	10^{-1} []
$\epsilon_{ m cut}^{ m ds3552}$	9.073 ± 0.027	10^{-1} []
$\epsilon_{ m cut}^{ m ds3555}$	9.301 ± 0.023	10^{-1} []
$\epsilon_{ m cut}^{ m ds3561}$	9.142 ± 0.060	10^{-1} []
$\epsilon_{ m cut}^{ m ds3564}$	9.125 ± 0.037	10^{-1} []
$\epsilon_{ m cut}^{ m ds3567}$	7.911 ± 0.050	10^{-1} []
$\epsilon_{ m scenario}^{ m 2A0-2B1}$	4.145 ± 0.208	10^{-2} []
$\epsilon_{ m scenario}^{ m 2A2-2B3}$	2.430 ± 0.121	10^{-2} []
$\epsilon_{ m scenario}^{ m 3A0}$	1.859 ± 0.093	10^{-3} []
$\delta^{\mathrm{ds}3519}_{\mathrm{PSA}}$	0.021 ± 5.499	10^{-3} []
$\delta^{\mathrm{ds}3522}_{\mathrm{PSA}}$	0.004 ± 1.001	10^{-3} []
$\delta^{\mathrm{ds}3552}_{\mathrm{PSA}}$	-0.080 ± 7.785	10^{-3} []
$\delta^{\mathrm{ds}3555}_{\mathrm{PSA}}$	-0.011 ± 1.496	10^{-3} []
$\delta^{\mathrm{ds}3561}_{\mathrm{PSA}}$	-0.107 ± 8.202	10^{-3} []
$\delta^{\mathrm{ds}3564}_{\mathrm{PSA}}$	0.011 ± 1.169	10^{-2} []
$\delta^{\mathrm{ds}3567}_{\mathrm{PSA}}$	-0.092 ± 2.322	10^{-2} []
$\epsilon_{ m AC}^{ m ds3519}$	9.881 ± 0.041	10^{-1} []
$\epsilon_{ m AC}^{ m ds3522}$	9.935 ± 0.030	10^{-1} []
$\epsilon_{ m AC}^{ m ds3552}$	9.883 ± 0.025	10^{-1} []
$\epsilon_{ m AC}^{ m ds3555}$	9.839 ± 0.027	10^{-1} []
$\epsilon_{ m AC}^{ m ds3561}$	9.814 ± 0.063	10^{-1} []
$\epsilon_{ m AC}^{ m ds3564}$	9.849 ± 0.030	10^{-1} []
$\epsilon_{ m AC}^{ m ds3567}$	9.860 ± 0.025	10^{-1} []
$f(^{130}Te)$	3.417 ± 0.000	10^{-1} []
$T_{1/2} (90\% \text{ C.I. lower limit})$	$> 5.909 \cdot 10^{24}$	[yr]

TABLE 7.13: Unblinded fit results $0\nu\beta\beta$ with all systematics. Scaling lineshape parameters are omitted for clarity. The results are displayed as mean and standard deviation of the marginalized posterior pdf for each fit parameter.

Observable	$\mathbf{Mean} \pm \mathbf{Stdev}$	\mathbf{Units}
Γ_{etaeta}	5.742 ± 4.452	$10^{-26} [yr^{-1}]$
$\mathrm{BI}_{\mathrm{2A0-2B1}}$	2.126 ± 1.371	$10^{-4} [\mathrm{cts/keV/kg/yr}]$
$\mathrm{BI}_{\mathrm{2A2-2B3}}$	2.685 ± 1.200	$10^{-4} [\mathrm{cts/keV/kg/yr}]$
BI _{3A0}	5.788 ± 5.441	$10^{-5} [\mathrm{cts/keV/kg/yr}]$
$\epsilon_{ m cut}^{ m ds3519}$	8.579 ± 0.037	10^{-1} []
$\epsilon_{ m cut}^{ m ds3522}$	9.209 ± 0.025	10^{-1} []
$\epsilon_{ m cut}^{ m ds3552}$	9.073 ± 0.027	10^{-1} []
$\epsilon_{ m cut}^{ m ds3555}$	9.301 ± 0.023	10^{-1} []
$\epsilon_{ m cut}^{ m ds3561}$	9.143 ± 0.060	10^{-1} []
$\epsilon_{ m cut}^{ m ds3564}$	9.125 ± 0.037	10^{-1} []
$\epsilon_{ m cut}^{ m ds3567}$	7.912 ± 0.050	10^{-1} []
$\epsilon_{ m scenario}^{ m 2A0-2B1}$	4.553 ± 0.228	10^{-2} []
$\epsilon_{ m scenario}^{ m 2A2-2B3}$	2.935 ± 0.147	10^{-2} []
$\epsilon_{ m scenario}^{ m 3A0}$	2.491 ± 0.125	10^{-2} []
$\delta^{\mathrm{ds3519}}_{\mathrm{PSA}}$	-0.043 ± 5.494	10^{-3} []
$\delta^{\mathrm{ds3522}}_{\mathrm{PSA}}$	0.003 ± 9.976	10^{-4} []
$\delta^{\mathrm{ds3552}}_{\mathrm{PSA}}$	-0.025 ± 7.800	10^{-3} []
$\delta^{\mathrm{ds}3555}_{\mathrm{PSA}}$	-0.003 ± 1.499	10^{-3} []
$\delta^{\mathrm{ds3561}}_{\mathrm{PSA}}$	0.001 ± 8.202	10^{-3} []
$\delta^{\mathrm{ds3564}}_{\mathrm{PSA}}$	0.001 ± 1.167	10^{-2} []
$\delta^{\mathrm{ds3567}}_{\mathrm{PSA}}$	-0.050 ± 2.323	10^{-2} []
$\epsilon_{ m AC}^{ m ds3519}$	9.881 ± 0.041	10^{-1} []
$\epsilon_{\rm AC}^{\rm ds3522}$	9.935 ± 0.030	10^{-1} []
$\epsilon_{ m AC}^{ m ds3552}$	9.883 ± 0.025	10^{-1} []
$\epsilon_{ m AC}^{ m ds3555}$	9.839 ± 0.027	10^{-1} []
$\epsilon_{\rm AC}^{\rm ds3561}$	9.813 ± 0.063	10^{-1} []
$\epsilon_{ m AC}^{ m ds3564}$	9.849 ± 0.030	10^{-1} []
$\epsilon_{ m AC}^{ m ds3567}$	9.860 ± 0.025	10^{-1} []
$f(^{130}Te)$	3.417 ± 0.000	10^{-1} []
$T_{1/2}$ (90% C.I. lower limit)	$> 1.295 \cdot 10^{24}$	[yr]

TABLE 7.14: Unblinded fit results $2\nu\beta\beta$ with all systematics. Scaling lineshape parameters are omitted for clarity. The results are displayed as mean and standard deviation of the marginalized posterior pdf for each fit parameter.

Chapter 8

Conclusion

My PhD thesis work has been primarily carried out at the Laboratori Nazionali del Gran Sasso (LNGS) of the Istituto Nazionale di Fisica Nucleare (INFN), within the CUORE collaboration. I actively contributed to the data taking at different levels that can be summarized in the following macro-areas:

- support to detector operation and data processing
- development of software tools for data processing and analysis
- data analysis and extraction of physics results

Detector operation and data processing The CUORE data taking needs, among other expertise, data analysts able to process quickly but reliably the data in order to extract the basic pieces of information needed by the Collaboration to take decisions on how to schedule maintenance and operation in the safest and most effective possible way. I have took part in both the online data processing (Sec. 3.3), and online analysis of data specifically acquired for detector noise optimization in terms of the relative phase of the rotary valve of the Pulse Tube cryocooles 2.3.3. Such optimization campaigns have taken, up to now, $\sim 5\%$ of the live time of the detector. Keeping the amount of live time dedicated to optimization low while ensuring optimal noise configuration is the goal of the pulse tube scan analysis. Finally I took part in the offline computation of the detector response function. As detailed in Sec. 4.3.1 in order to characterize the response function of each detector in each dataset the statistics provided by physics data is not enough, for this reason calibration data are used. I had the task of evaluating the detector response for the first time with the external detector calibration system was deployed.

Software tools The data processing relies on a modular software framework called **DIANA**. CUORE, for the first time after the long story of its predecessors, offers a high enough granularity that allows a design of the coincidence algorithms based on the distance among crystals. I was responsible for the development of such coincidence algorithm, its

implementation and its validation. The search for double beta decay on the ground state is effectively performed looking just at single-site events. For this reason it is not mandatory to have optimized algorithms to select signal events where Multiplicity > 1. I have developed a dedicated software module, within the same DIANA framework, to handle this. The process was successful for both tasks and the corresponding modules were included in the latest official data reprocessing campaing used to extract the results on double beta decay.

Any analysis would benefit from a narrower detector resolution, and the resolution is strictly correlated with the integral of the noise in the signal bandwidth. It was found that CUORE is affected by very low frequency noise f < 1 Hz that the current analysis chain cannot suppress effectively because its frequency resolution is limited by the 10 s width of the time window used to process digitized voltage pulses. It is foreseen an improvement up to $\sim 10\%$ in the noise RMS if such components were effectively suppressed. I have been working to the development of a software notch filter to address such low frequency noise components. The development of such tool is still in progress.

The evaluation of the detector response function dependence on energy is crucial because it directly impacts any analysis that involves the **Energy** observable. I have been in charge of coding a software for extracting the energy bias and resolution dependence on energy, starting from an existing **RooFit** infrastructure. The outcome of this work allowed me to go further and develop the corresponding Bayesian description of the same parameters as Eq. 6.13, which was implemented successfully both in the double decay search in the ground-state to ground-state (gs-gs) transition as well as in the ground-state to excited-state (gs-es) one.

Finally I was involved, together with other collaborators, in the software development of a Bayesian fitter for $0\nu\beta\beta$ decays reconstructed as single-site events. The expertise acquired allowed me to develop an analogue tool for searching $\beta\beta$ decay events to the first 0^+ excited state of ¹³⁰Xe.

Data analysis Using the above described software tools I was able to contribute to the data analysis for $0\nu\beta\beta$ decay search (gs-gs transition) and to search for double beta decay events in the gs-es transition. Some of the results obtained were presented on 9-13 September 2019 at the TAUP conference, Toyama (Japan) [135] and were published (gs-gs transition) as a collaboration paper on Physical Review Letters [41]. The lower limits for the double beta decay half life parameter in the investigated decay channels follow

$$^{0^+}_1(T^{0\nu}_{1/2}) > 3.2 \cdot 10^{25} \,[\text{yr}] \quad 90\% \text{C.I.}$$
(8.1)

$$_{2}^{0^{+}}(T_{1/2}^{0\nu}) > 5.9 \cdot 10^{24} \, [yr] \quad 90\% C.I.$$
 (8.2)

$$^{0^+}_2(\mathbf{T}^{2\nu}_{1/2}) > 1.3 \cdot 10^{24} \, [yr] \quad 90\% \text{C.I.}$$
 (8.3)

It is useful to remember, with regard to double beta decay search to the excited states

of 130 Xe, that the currently available limits yield (Eq. 7.1 and 7.2)

$$\begin{split} (T_{1/2})_{0_2^+}^{0\nu} &> 1.4 \cdot 10^{24} \text{ yr}, 90\% \text{ C.L.} \\ (T_{1/2})_{0_2^+}^{2\nu} &> 2.5 \cdot 10^{23} \text{ yr}, 90\% \text{ C.L.} \end{split}$$

and that this analysis improves by a factor 4-5 the result in both decay channels. This result benefits from both a ~ 4 -fold increase in exposure and lower background.

The CUORE data taking is proceeding smoothly and continuously towards the goal of accumulating 5 yr of live time, corresponding to a ^{nat}TeO₂ exposure of ~ 3.7 ton · yr. The first milestone of 1 ton · yr of exposure is expected by the end of this year. Assuming the median limit setting sensitivity scales as the square root of the exposure, an additional factor 3 – 4 improvement is expected by the end of the data taking, bringing the median expected limit setting sensitivity very close to the lowest bound of the half life theoretical predictions. It is useful to remind here the range of theoretical estimates for the $2\nu\beta\beta$ decay on the 0^+_2 excited state (Eq. 7.3)

$${}^{th}(T_{1/2})^{2\nu}_{0^+_2} = (7.2 - 16) \cdot 10^{24} \text{ y}$$

Small additional improvements in the analysis technique might allow, by the end of the CUORE data taking, to begin exploring such range. At the time of writing, the results in Eq. 8.2 and 8.3 constitute the world most sensitive search for double beta decay of 130 Te to the excited states of 130 Xe.

Appendix A

Unbinned Extended Maximum Likelihood derivation

In this section it will be shown that constructing an unbinned extended likelihood as the product of a Poissonian term and the model probability density function evaluated in each data point is equivalent to constructing the binned likelihood of the same data in the limit of small bin size.

Let us assume we have a random variable x distributed with a normalized pdf f(x)in some domain. We sample from that distribution expecting μ counts, instead we get nvalues x_i i = 1...n. Let us bin the domain of f and compute the binned likelihood of our dataset, given the model $f(x), \mu$:

$$\log \mathcal{L}(\vec{x}|f,\mu) = \log \prod_{b} \frac{\mu_b^{n_b} e^{-\mu_b}}{n_b!}$$
(A.1)

where b is the index of the bin, $\mu_b = \mu \int_{b^{th} bin} f(x) dx$ is the expected number of counts in bin b. Explicitly,

$$\log \mathcal{L} = \sum_{b} n_b \log \mu_b - \sum_{b} \mu_b - \sum_{b} \log n_b!$$
(A.2)

where n_b is the number of counts in bin b. If now we let the bin width go to 0 so that n_b is either 1 or 0 we have

$$\log \mathcal{L} = \sum_{i} \log \mu_{i} - \mu \tag{A.3}$$

where the summation is done on the index i, which runs just on non-empty bins (i.e. collected samples), and the last term vanishes because $\log 1! = \log 0! = \log 1 = 0$.

Let us now define the Unbinned Extended Likelihood as the product of the poissonian probability of getting n counts expecting μ , and the product of the probability density function evaluated at each collected sample x_i

$$\log \mathcal{L}_U \doteq \log \frac{e^{-\mu} \mu^n}{n!} + \log \prod_i f(x_i)$$
(A.4)

$$\log \mathcal{L}_U = -\mu + n \log \mu - \log n! + \sum_i \log f(x_i)$$
(A.5)

If we now want to insert a tiny binning (so that the number of counts in each bin can either be 0 or 1) we can consider that the expected number of counts in each bin is μ_b , and goes to 0 as the binning becomes thinner. Since $\mu_b/\mu = f(x_b) \cdot \epsilon$ where x_b is any point in bin band ϵ is the bin width, we can easily substitute and get

$$\log \mathcal{L}_U = -\mu + n \log \mu - \log n! + \sum_i \log \frac{1}{\epsilon} \frac{\mu_i}{\mu}$$
(A.6)

$$\log \mathcal{L}_U = -\mu + n \log \mu - \log n! + \sum_i \log \mu_i - n \log \mu + const$$
(A.7)

where we dropped the term with ϵ since it is a constant and the sum is on non-empty bins. We can now see that the $n \log \mu$ terms cancel, and we can include also $\log n!$ into the irrelevant constant, because n is constant at fixed data sample. From the equivalence of Eq. A.3 and A.7 up to an additive constant we can deduce that building an unbinned extended likelihood as in Eq. A.4 is equivalent to the binned product of Poisson terms in a binned likelihood in the limit of small bin size, because the likelihood function is defined up to an arbitrary multiplicative normalization constant.

Appendix B

Bayesian inference with Markov Chain Monte Carlo

The scientific community is generally split in two sub-communities as far as the approach to data analysis is concerned: Bayesians and Frequentists. The first ones use Bayes' theorem to define a degree of belief in a model, or in general make probabilistic statements about unknown physical quantities, plugging into their analyses prior information coming from a source potentially different from the analyzed data itself (e.g. previous measurements of the same observable). The others believe that physical quantities do have a *true value* that the observer does not know (and cannot know with arbitrary precision), hence making probabilistic statements about nature's parameters does not make sense. In this respect, probabilistic statements can be rightfully made solely about possible outcomes of repeated experiments given an assumed true value of the parameter of interest. In the following a bayesian approach will be adopted.

B.1 Bayesian inference

Bayes' theorem in general states that

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$
(B.1)

where A and B are logical propositions, P(A) denotes both the degree of belief in proposition A and the probability of A, P(A|B) is the conditional probability of A given B. It is possible to rewrite Eq. B.1 as

$$P(\boldsymbol{\theta}|\mathcal{D}, \mathbf{H}) = \frac{\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \mathbf{H})\pi(\boldsymbol{\theta}|\mathbf{H})}{\int_{\boldsymbol{\Theta}} \mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \mathbf{H})\pi(\boldsymbol{\theta}|\mathbf{H})d\boldsymbol{\theta}}$$
(B.2)

where $\boldsymbol{\theta}$ is a column vector of parameters θ_i that define a model H of the data \mathcal{D} . $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, H)$ is the likelihood, i.e. the conditional probability, up to a normalization factor, of the data

given the model H defined by the $\boldsymbol{\theta}$ parameters value, $\pi(\boldsymbol{\theta}|\mathbf{H})$ is the joint prior probability density of the $\boldsymbol{\theta}$ parameters whose domain is Θ . $P(\boldsymbol{\theta}|\mathcal{D},\mathbf{H})$ is the posterior probability density of the $\boldsymbol{\theta}$ parameters given the data \mathcal{D} , the prior knowledge $\pi(\boldsymbol{\theta}|\mathbf{H})$ and the model H. Prior and posterior probability densities, in Bayesian inference, represent the state of knowledge about the $\boldsymbol{\theta}$ parameters before and after seeing the data, and have no frequentist equivalent.

The outcome of a statistical inference about one parameter of interest (POI) can be expressed as the value of a point estimator (a single valued result) and/or an interval estimator (a range of values). Let us restrict to one single parameter of interest Γ , being $\hat{\Gamma}$ its point estimator. An example of point estimator is the maximum a posteriori probability (MAP) estimator

$$\hat{\Gamma} \doteq \max_{\Gamma} P(\Gamma | \mathcal{D}) \tag{B.3}$$

defined as the global mode of e.g. the marginalized posterior probability distribution. The θ parameters, being not-POI, are called *nuisance* parameters and the distribution on the right hand side of Eq. B.3 can be obtained integrating the joint posterior over the nuisance parameters. A consistent example of interval estimator is the shortest interval with a given probability content α which, for a unimodal distribution, is defined as

$$\alpha = \int_{\hat{\Gamma}_{-}^{\alpha}}^{\hat{\Gamma}_{+}^{\alpha}} P(\Gamma|\mathcal{D}) d\Gamma \quad P(\hat{\Gamma}_{-}^{\alpha}|\mathcal{D}) = P(\hat{\Gamma}_{+}^{\alpha}|\mathcal{D}) \quad \hat{\Gamma}_{-}^{\alpha} < \hat{\Gamma} < \hat{\Gamma}_{+}^{\alpha}$$
(B.4)

B.2 Markov Chain Monte Carlo

A Markov Chain is a stochastic process in which any future state is independent on past states, rather it depends just on the present state. Such a random walk for the purpose of bayesian analyses takes place in the Θ space and is made of time-indexed steps $\boldsymbol{\theta}^{(t)} \in \Theta$ that satisfy the Markov property

$$P(\theta^{(t+1)}|\theta^{(t)}, \theta^{(t-1)}, ..., \theta^{(1)}) = P(\theta^{(t+1)}|\theta^{(t)})$$
(B.5)

It is completely defined by the initialization $\theta^{(0)}$ and the transition kernel $f(\theta^{(t+1)}|\theta^{(t)})$. It can be shown [139] that a Markov Chain

- without any non-trivial time periodicity (aperiodic)
- allowing to explore the whole domain regardless of the starting point, in any number of steps (irreducible)
- returning to any given state with probability 1 in a finite amount of time (positive recurrent)

is ergodic, a unique stationary distribution $\Pi(\theta)$ exists and is eventually reached. This means that the probability to be in a given state is stationary. Moreover, given any

function $g(\boldsymbol{\theta})$ with expected value $E[\boldsymbol{\theta}] < \infty$, holds

$$\frac{1}{M} \sum_{i=1}^{M} g(\boldsymbol{\theta}^{(i)}) \xrightarrow{M \to \infty} \int_{\Theta} g(\boldsymbol{\theta}) \Pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$$
(B.6)

The strong law of large numbers states that given an infinite sequence of identically distributed independent random variables, their sample average converges to the expected value. The result in Eq. B.6 is fundamental, because even without being able to sample independent identically distributed (according to the posterior) random variables, being able to devise a Markov Chain whose stationary distribution is the posterior allows to use slightly correlated random variables (i.e. the Markov Chain steps) to compute any integral of this form. In the following, two algorithms to generate Markov Chains with limiting distributions $\Pi(\boldsymbol{\theta}) = P(\boldsymbol{\theta}|\mathcal{D})$ will be outlined.

Gibbs sampling A recipe for sampling steps of a Markov Chain is Gibbs sampling. It requires knowledge of the full conditional distributions

$$P(\theta_i | \boldsymbol{\theta}_{-i}) \quad \forall i = 1.. \dim \boldsymbol{\theta}$$
(B.7)

where θ_{-i} represents all the θ components except the *i*-th one, which can be derived from the joint probability distribution $P(\theta)$. It proceeds as follows,

- 1. pick any starting value $\boldsymbol{\theta}^{(0)} \in \Theta$ for t = 0
- 2. generate $\theta_1^{(t+1)}$ from $P(\theta_1^{(t)}|\boldsymbol{\theta}_{-1}^{(t)})$
- 3. generate $\theta_2^{(t+1)}$ from $P(\theta_2^{(t)}|\theta_1^{(t+1)}, \theta_3^{(t)}, ..., \theta_d^{(t)}$ where $d = \dim \theta$
- 4. proceed in the same way recursively refreshing the conditioning variables until all directions are scanned
- 5. output $\theta^{(t+1)}$ and restart from step 2 to generate the next step of the Markov Chain setting t = t + 1

Metropolis-Hastings algorithm Another procedure to sample steps of a Markov Chain is the Metropolis-Hastings algorithm. The main advantage with respect to Gibbs sampling is that the knowledge of all the conditional distributions is not needed. On the other hand a suitable proposal function (kernel) must be defined. A Markov Chain is produced as follows:

- 1. pick any starting value $\boldsymbol{\theta}^{(0)} \in \Theta$ for t = 0
- 2. sample $\boldsymbol{\theta}^*$ from a proposal function $f(\boldsymbol{\theta}^*|\boldsymbol{\theta}^t)$
- 3. set $\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^*$ with probability $\min(1, r)$ where $r = \frac{P(\boldsymbol{\theta}^* | \mathcal{D})}{P(\boldsymbol{\theta}^t | \mathcal{D})}$, which corresponds to deterministic acceptance when the proposed point $\boldsymbol{\theta}^*$ has higher posterior with respect to the previous step, otherwise the acceptance happens randomly with probability r

4. update t if necessary and go back to step 2

The proposal function f must not depend on the step. An optimal choice of the proposal function (kernel) of the Metropolis-Hastings algorithm guarantees efficient sampling (i.e. a high step-acceptance rate), and a lower number of steps needed before a stationary state is reached.

A Markov Chain Monte Carlo (MCMC) is then a procedure in which, leveraging the generation of pseudo-random numbers, it is possible to simulate time-ordered draws $\boldsymbol{\theta}^{(t)}$ that are slightly correlated and are approximately distributed as the posterior distribution. Then it is possible to use such draws to compute quantities of interest involving the posterior distribution.

Practical usage of MCMC methods forces to consider one additional conceptual step: the burn in. If, from a mathematical point of view, a Markov Chain is guaranteed to converge (i.e. the distribution of the $\theta^{(t)}$ approximates the limiting distribution) for some time step $t > t_{\text{th}}$, it is important to be able to assess when convergence is reached, to understand which parameters can influence the time of convergence and, possibly, tune them in order to minimize it. Every bayesian analysis based on MCMC has a burn-in, namely a part in which samples are drawn waiting for convergence. In the following this topic will be explored in the context of the BAT (Bayesian Analysis Toolkit) implementation of the Metropolis-Hastings algorithm [140].

3sectionMCMC prerun and convergence in the BAT framework

BAT is a toolkit for data analysis based on Bayes' Theorem and Markov Chain Monte Carlo techniques. One crucial ingredient of a MCMC is the definition of a convergence criterion. BAT runs multiple independent chains in parallel and declares convergence if the chains mix, i.e. if they end up exploring the same region of the parameter space. Non-convergence is an issue that can have many causes: the most common are bugs in the implementation of the posterior. Additionally, if some parameters are strongly correlated or if the posterior has multiple well separated local modes, the MCMC will very likely have difficulties in properly exploring the full parameter space.

An additional key element of Metropolis-Hastings is the algorithm for the choice of a new point given the current one, i.e. the choice of the proposal function $q(\tilde{\theta}|\theta_t,\xi)$ with some adjustable parameter ξ . In BAT, the kernel is symmetric:

$$q(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}_t,\xi) = q(\boldsymbol{\theta}_t|\tilde{\boldsymbol{\theta}},\xi)$$
(B.8)

A BAT fit is divided in two stages: the *prerun*, devoted to the optimization of the proposal function internal parameter ξ until the convergence is reached and that correponds to the burn-in, and the *main run* for the mapping of the posterior with a proper MCMC with a fixed value of ξ . In the following, we illustrate how convergence is defined during the prerun phase and briefly describe the available proposal functions.

Prerun and proposal functions During the prerun the parameter ξ of the proposal function is gradually updated based on past iteration. Three criteria are relevant:

- 1. the MCMC acceptance, or *efficiency*, i.e. the ratio of accepted over generated samples. A small efficiency means the chain rarely moves, but may do large moves. A large efficiency means the chain explores well the parameters space locally, but might take long in probing the whole domain and reaching the region around the global mode. Optimal efficiency values exist only for special cases. The BAT default range for the MCMC efficiency is [0.15, 0.35].
- 2. the R value, which compares the mean and variance of the expectation value of each parameter for a single chain with that of all chains together. If the chains mix despite starting from different initial points, they are likely independent of the initial value and the Markov Chain requirements are satisfied. During the prerun, BAT periodically monitors R and declares convergence once R is below a given threshold for all chains. An important remark is that a steadily increasing R for any parameter is a symptom of non-convergence usually due to a buggy posterior implementation, multi-modal distributions that trap the chains, or strongly correlated parameters.
- 3. the minimum and maximum allowed number of prerun steps. BAT declares convergence if the efficiency is within a predefined range for all chains, R below threshold for all parameters, and terminates the prerun after reaching convergence and after having drawn at least the minimum number of samples. If the convergence is not reached within the maximum number of steps, BAT will anyway switch to the main run, and warn the user about the prerun fail. This can happen either for models with a large number of parameters or, more likely, if the implementation of the posterior has some bug. The user can in any case modify the minimum and maximum prerun length, but caution is highly recommended in doing so.

BAT offers two categories of proposal functions: *factorized* and *multivariate*, which is the default one. In the factorized case, the joint distribution is a product of 1-dim distributions, whereas in the multivariate case a covariance matrix is used to treat correlations. A detailed description of the two cases is out of scope but is available in the BAT documentation. A proposal function is efficient if it allows to sample from the entire target support in a finite time, it resolves small and large scale features of the target, and it quickly reaches the asymptotic regime, leading to a Markov Chain. With a factorized proposal, BAT sequentially varies one parameter at a time and completes one iteration of the chain once a move has been accepted in every direction. The independent move in every direction typically leads to a higher efficiency because the new generated point differs from the previous one only for one parameter, but also involves a larger number of tested samples, thus is not advisable if the computation of the posterior is CPU expensive. On the other hand, in the multivariate case BAT moves in all directions at once: if the proposed move

parameters space, otherwise no progress is made. The efficiency is lower, but the number of tested points decreases. In general, a multivariate proposal function is advisable in case of highly correlated priors, or for highly CPU expensive posteriors. For the analysis described in Chap. 6 and Chap. 7 a multivariate proposal function is always used.

Posterior sampling vs mode finding We want to stress here that a MCMC aims at mapping the posterior distribution, not ad finding its global mode. For every step of the chain, BAT stores the corresponding coordinates in the parameter space, and its likelihood, prior and posterior probabilities. Nothing guarantees that the point with the highest posterior probability found by the MCMC corresponds to the global mode. BAT allows the user to find the global mode calling BCIntegrate::FindMode. The user can specify the starting point (usually the maximum found by the MCMC) and the optimization method, with Minuit as default. Minuit is a gradient follower that moves deterministically in the direction that increases the posterior, until it reaches the maximum. As such, Minuit performs well at finding the exact location of a maximum, but it does not guarantee that this maximum is the global mode.

Appendix

Bayesian Result Combination

One of the most useful features of Bayesian inference is the straightforward possibility of consistently update knowledge with appropriate choice of priors. For example let us assume, in the hypothesis that $0\nu\beta\beta$ decay is mediated by light Majorana neutrino exchange, that both the gs-gs¹ and the gs-es² transitions are explored. In the absence of specific prior knowledge in either case a uniform prior is defined in the $[0, \Gamma^{\max}]$ domain for the $\Gamma_{0\nu}^{gs}$ and $\Gamma_{0\nu}^{es}$ parameters is chosen, where the parameters represent the rate of $0\nu\beta\beta$ decay in the gs-gs and gs-es channel respectively. The result of the analysis is a marginalized posterior probability distribution for each of the Γ parameters. Since they both depend on the square of the effective Majorana mass parameter $m_{\beta\beta}$ as

$$\Gamma_{0\nu}^{ch} = \ln\left(2\right) G_{0\nu}^{ch} g_A^4 |M_{0\nu}^{ch}|^2 \left(\frac{m_{\beta\beta}}{m_e}\right)^2 \tag{C.1}$$

where ch denotes the decay channel, it is interesting to understand how the results can be combined to possibly achieve a more stringent limit on the parameter of physical interest, m_{β} .

C.1 Random variable transformations

It might be useful to recall some basics about how the probability distribution functions transform under an invertible random variable transformation. For this reason let us consider a variable Γ , the rate of some unknown process, and its probability density function

$$f_{\Gamma}(\Gamma) : \mathbb{R} \longrightarrow \mathbb{R}^+ \tag{C.2}$$

Please note that the domain is not restricted to positive rates because this physical restriction can be implemented in the prior pdf setting $f_{\Gamma} \propto \theta(\Gamma)$ where θ is the Heaviside function. Since a redefinition of the random variable according to $\phi = \Phi(\Gamma)$ cannot imply

 $^{^1\}mathrm{Ground}$ state to ground state

²Ground state to excited state, in this case implicitly we refer to the 0_2^+ state of ¹³⁰Xe

probability non-conservation, the following holds

$$f_{\Gamma}\left[\Phi^{-1}(\phi)\right] \left| \frac{d\Gamma}{d\phi} \right| d\phi = f_{\phi}(\phi) d\phi \tag{C.3}$$

where the left hand side is just the probability content of an infinitesimal portion $d\Gamma$ of the Γ domain centered at $\Gamma = \Phi^{-1}(\phi)$, expressed in terms of the transformed variable ϕ . This, in turn, implies that the probability density function in the transformed space reads

$$f_{\phi}(\phi) = \left| \frac{d\Phi}{d\Gamma} \right|^{-1} f_{\Gamma} \left[\Phi^{-1}(\phi) \right]$$
(C.4)

and means that provided that the Φ transformation is invertible, it is possible to compute (either analytically or numerically) the probability distribution function in the transformed space according to Eq. C.4.

C.2 Combining the results from excited and ground state modes

Let us assume, to ease computation, that no signal is observed nor in the gs-gs nor in the gs-es mode of neutrino-less double beta decay of 130 Te and that the posterior pdf is of the form

$$P(\Gamma_{0\nu}^{ch}|\mathcal{D}_{ch}) \propto e^{-\Gamma_{0\nu}^{ch}/\Lambda_{ch}}$$
(C.5)

where Λ_{ch} is the scale of the limit that the search for $0\nu\beta\beta$ was able to place in channel ch. Assuming no nuisance parameter is present in the derivation of the result and a uniform prior on the decay rate parameter, the likelihood function is proportional to the prior. From Eq. C.1 substituting the term in round brackets follows

$$\Gamma_{0\nu}^{es} = \left[\frac{G_{0\nu}^{es}}{G_{0\nu}^{gs}} \frac{|M_{0\nu}^{es}|^2}{|M_{0\nu}^{es}|^2}\right] \Gamma_{0\nu}^{gs} \tag{C.6}$$

under the hypothesis that $0\nu\beta\beta$ is mediated by light Majorana neutrino exchange.

Assuming the result of a $0\nu\beta\beta$ decay search in the gs-gs mode is available at the time the gs-es mode is explored, it would be possible to plug the knowledge about $\Gamma_{0\nu}^{gs}$ as a prior for the gs-es analysis. Doing so and plugging the result from Eq. C.4 into Bayes Theorem, follows

$$P(\Gamma_{0\nu}^{es}|\mathcal{D}) \propto e^{-\Gamma_{0\nu}^{es}/\Lambda_{es}} \cdot e^{-\Gamma_{0\nu}^{es}/\epsilon\Lambda_{gs}}$$
(C.7)

where $\epsilon \sim 3 \cdot 10^{-2}$ is the proportionality factor in square brackets in Eq. C.6. The ratio of the squares of the nuclear matrix elements is of order unity, while the phase space ratio is much smaller, since the phase space of double beta decay is proportional to $Q_{\beta\beta}^5$. It is found that the scale of the limit on the half life parameter $T_{ch} \doteq \Lambda_{ch}^{-1}$ is

$$T_{es}' = \frac{T_{gs}}{\epsilon} \left[1 + \left(\frac{\epsilon T_{es}}{T_{gs}} \right) \right]$$
(C.8)

For this reason, unless the sensitivity to neutrino-less double beta decay to the excited state surpasses the one to the ground state by at least one order of magnitude, the combined result is by far dominated by the limit on the ground state. It can be shown that assuming a different shape of the posterior pdf (e.g. a gaussian centered at null rate) the combined limit still depends on the term in round brackets as

$$T_{es}' = \frac{T_{gs}}{\epsilon} \sqrt{1 + \left(\frac{\epsilon T_{es}}{T_{gs}}\right)^2}$$
(C.9)

leaving the conclusions unchanged.

Combining independent results on the same mode Based on the same arguments and assumptions, should we attempt to plug the knowledge from previous searches about the same process (e.g. limits on the rate of $0\nu\beta\beta$ decay on the first 0^+ excited state of the same nucleus, coming from independent experiments or signatures), the same conclusions hold provided that $\epsilon = 1$ since Eq. C.6 would just be an identity between the decay rates observed from different data samples. It might be interesting to note that an exponentially decaying posterior can be obtained in a background-free scenario whereas a gaussian posterior can be expected in background-dominated searches.
List of Abbreviations

0 uetaeta	Neutrinoless double beta decay	MiDBD	Milan Double Beta Decay
2 uetaeta	Two-neutrino double beta decay		(experiment)
BAU	Baryon Asymmetry in the	MIT	Metal-Insulator Transition
	Universe	MSP	Main Support Plate
CMB	Cosmic Microwave Background	NH	Normal Hierarchy
CUORE	Cryogenic Underground	NME	Nuclear Matrix Element
	Observatory for Rare Events	NTD	Neutron Transmutation
CUPID	CUORE Upgrade with Particle		Doping/ed (technique/sensor)
	IDentification	OF	Optimum Filter
DAQ	Data AcQuisition	OFE	Oxygen-Free Electrolytic
DCS	Detector Calibration System		(copper) (OFHC C10100)
EDCS	External Detector Calibration	OFHC	Oxygen-Free High thermal
	System		Conductivity (copper)
EDF	Energy Density Functional	OVC	Outer Vacuum Chamber
DR	Dilution Refrigerator	pdf	Probability Distribution Function
DU	Dilution Unit	РТ	Pulse Tube (refrigerator)
FWHM	Full Width Half Maximum	PTFE	PolyTetraFluoroEthene
	(energy resolution)		
HEX	Heat EXchanger	ORPA	Quasiparticle Bandom Phase
IBM	Interacting Boson Model	QIUIII	Approximation
ISM	Interacting Shell Model	BOI	Region Of Interest
IH	Inverted Hierarchy	BMS	Boot Mean Square
INFN	Istituto Nazionale di Fisica	SM	Standard Model
	Nucleare	TSP	Tower Support Plate
ISM	Interacting Shell Model	101	
NSM	Nuclear Shell Model		
IVC	Inner Vacuum Chamber	TVL	Test Variable Left
LNGS	Laboratori Nazionali del	TVR UEML	Test Variable Right
	Gran Sasso		Unbinned Extended Maximum
MC	Mixing Chamber		Likelihood (fit)

List of Figures and Tables

1.1	Neutrino mass spectra and flavor content	10
1.2	Qualitative representation of the nuclear mass dependence on atomic number	
	for even (left) and odd (right) mass number isobars. For some even-even	
	nuclei whose nuclear mass structure resembles the above sketch, the $\beta\beta$	
	decay mode is possible	12
1.3	Diagrams of double beta decay processes in the electro-weak SM framework	
	(left) and with a light Majorana neutrino exchange (right)	12
1.4	Limit on $m_{\beta\beta}$	16
1.5	Sum energy spectrum of electrons emitted in double beta decay in both	
	standard $(2\nu\beta\beta)$ and neutrino-less $(0\nu\beta\beta)$ mode	18
1.6	Double beta decay emitters Q value as a function of their natural isotopic	
	abundance.	21
1.7	A selected set of experimental searches for $0\nu\beta\beta$ decay together with the	
	latest results in terms of 90% C.I. lower limits and expected sensitivity to	
	the inverse of the neutrino-less double beta decay rate. The resolution at	
	$Q_{\beta\beta}$ is reported as well.	21
2.1	Bolometer sensitivity improvement over time	26
2.2	NTD used in CUORE-0 - picture and schematic	29
2.3	Bias and I-V characteristic of the CUORE NTDs	31
2.4	CUORE cryostat - vessels, lead shields and volumes	33
2.5	Support structure of the CUORE cryostat	34
2.6	Phase diagram of ${}^{3}\text{He}/{}^{4}\text{He}$ mixture	36
2.7	Schematic of a dilution refrigerator	37
2.8	CUORE pulse tube	38
3.1	Summary of some parameters of the CUORE bolometers, sensors and DAQ	
	system settings. Some of the reported values should be interpreted as esti-	
	mates of the order of magnitude of the corresponding parameter in optimal	
	/ average working conditions	42

3.5 Summary table of the most prominent available calibration peaks. The 583 keV, 911 keV and 969 keV peaks are visible with sufficient strength just with internal DCS strings while ⁶⁹ Co ones are available just with the external calibration system because its strings are the only ones containing the source. 56 Jitter correction effect on time difference distribution	3.2	Comparison of derivative and optimum trigger thresholds	44
3.6 Jitter correction effect on time difference distribution	3.5	Summary table of the most prominent available calibration peaks. The 583 keV, 911 keV and 969 keV peaks are visible with sufficient strength just with internal DCS strings while ⁶⁰ Co ones are available just with the external calibration system because its strings are the only ones containing the source.	50
4.1 Derivative trigger threshold	3.6	Jitter correction effect on time difference distribution	53
4.1 Derivative trigger threshold 58 4.2 PSA efficiency curve 66 4.3 Line shape fit to calibration data 66 4.4 Distribution of resolutions FWHM in calibration data 66 4.5 Energy bias and resolution energy dependence 66 5.1 CUORE model implemented in Monte Carlo simulations 77 5.2 Layer definition for background modelling 74 5.3 Background model 76 5.4 2νββ spectrum reconstruction 76 6.1 Summary of the processed data at the time of writing with breakdown of the exposure for each dataset. The number of active channels refers to the number of channels that successfully underwent all the processing steps. 78 6.2 The CUORE spectrum after the base cut selection (grey), anti-coincidence veto (red) and PSA cut (blue) are applied. Reprinted from [41] 76 6.3 Left: spectrum in the [2490, 2575] keV energy range (ROI) with best fit curve (solid red) and the model corresponding to the marginalized 90% C.I. upper limit on the 0νββ rate (dashed blue). Right: marginalized posterior distribution as a function of the decay rate $\Gamma_{0\nu}$ with all systematics included for the fit in both the physical range ($\Gamma_{0\nu} > 0$) and the full range. The 90% C.I. is shown in blue. Reprinted from [41]. 80 6.4 Efficiency contributions involved			
4.2 PSA efficiency curve	4.1	Derivative trigger threshold	58
4.3 Line shape fit to calibration data	4.2	PSA efficiency curve	62
 4.4 Distribution of resolutions FWHM in calibration data	4.3	Line shape fit to calibration data	64
 4.5 Energy bias and resolution energy dependence	4.4	Distribution of resolutions FWHM in calibration data	65
 5.1 CUORE model implemented in Monte Carlo simulations	4.5	Energy bias and resolution energy dependence	66
 5.2 Layer definition for background modelling	5.1	CUORE model implemented in Monte Carlo simulations	71
 5.3 Background model	5.2	Layer definition for background modelling	74
 5.4 2νββ spectrum reconstruction	5.3	Background model	75
 6.1 Summary of the processed data at the time of writing with breakdown of the exposure for each dataset. The number of active channels refers to the number of channels that successfully underwent all the processing steps 78 6.2 The CUORE spectrum after the base cut selection (grey), anti-coincidence veto (red) and PSA cut (blue) are applied. Reprinted from [41]	5.4	$2\nu\beta\beta$ spectrum reconstruction	76
 6.2 The CUORE spectrum after the base cut selection (grey), anti-coincidence veto (red) and PSA cut (blue) are applied. Reprinted from [41]	6.1	Summary of the processed data at the time of writing with breakdown of the exposure for each dataset. The number of active channels refers to the number of channels that successfully underwent all the processing steps	78
 6.3 Left: spectrum in the [2490, 2575] keV energy range (ROI) with best fit curve (solid red) and the model corresponding to the marginalized 90% C.I. upper limit on the 0νββ rate (dashed blue). Right: marginalized posterior distribution as a function of the decay rate Γ_{0ν} with all systematics included for the fit in both the physical range (Γ_{0ν} > 0) and the full range. The 90% C.I. is shown in blue. Reprinted from [41]	6.2	The CUORE spectrum after the base cut selection (grey), anti-coincidence veto (red) and PSA cut (blue) are applied. Reprinted from [41]	79
 6.4 Efficiency contributions involved in the 0νββ decay search in the gs-gs transition. 6.5 Analysis cuts and effective parameters for 0νββ decay search in the gs-gs transition. When a single parameter is not representative or the spread among datasets is relevant (active channels, resolution scaling, PSA cut) a range is reported. Otherwise the sum (exposure), an average (background index) or an exposure-weighted value (resolution) is shown. The resolution is expressed in terms of the Full Width Half Maximum (FWHM). In a gaussian distribution the FWHM correspondent to at 2.4σ. 	6.3	Left: spectrum in the [2490, 2575] keV energy range (ROI) with best fit curve (solid red) and the model corresponding to the marginalized 90% C.I. upper limit on the $0\nu\beta\beta$ rate (dashed blue). Right: marginalized posterior distribution as a function of the decay rate $\Gamma_{0\nu}$ with all systematics included for the fit in both the physical range ($\Gamma_{0\nu} > 0$) and the full range. The 90% C.I. is shown in blue. Reprinted from [41]	80
6.5 Analysis cuts and effective parameters for $0\nu\beta\beta$ decay search in the gs-gs transition. When a single parameter is not representative or the spread among datasets is relevant (active channels, resolution scaling, PSA cut) a range is reported. Otherwise the sum (exposure), an average (background index) or an exposure-weighted value (resolution) is shown. The resolution is expressed in terms of the Full Width Half Maximum (FWHM). In a gaussian distribution the FWHM correspondent to $\alpha + 2.4\sigma$	6.4	Efficiency contributions involved in the $0\nu\beta\beta$ decay search in the gs-gs transition	80
~ 2.40 . ~ 2.40	6.5	Analysis cuts and effective parameters for $0\nu\beta\beta$ decay search in the gs-gs transition. When a single parameter is not representative or the spread among datasets is relevant (active channels, resolution scaling, PSA cut) a range is reported. Otherwise the sum (exposure), an average (background index) or an exposure-weighted value (resolution) is shown. The resolution is expressed in terms of the Full Width Half Maximum (FWHM). In a gaussian distribution the FWHM corresponds to ~ 2.4σ .	81

6.6	Parameters used in the $0\nu\beta\beta$ decay search in the gs-gs transition together with their symbol, units and related prior probability distribution. The pa-	
	rameters are divided as parameter of interest (top), nuisance parameters	
	(middle), nuisance parameters included for systematic uncertainty evalua-	
	tion (bottom)	85
6.7	Left: knowledge update plot for the Q_{60} parameter where the prior pdf (grey) is superimposed to the marginalized posterior pdf (black). Right: knowledge update plot for the Q_{60} (ordinate) and the Γ_{60} parameters (abscissa). The marginalized mode is shown (black box) within the 68% smallest	
	interval (black contour).	86
6.8	Knowledge update plot for one coefficient of the $\Delta(Q)$ bias function for one dataset. The comparison between the prior (grey) and posterior (black) distributions shows agreement since no information about the energy recon-	
	struction bias can be extracted just from the data in the ROI	87
6.9	Distribution of the 90 % C.I. lower limits obtained from the S+B minimal model fit to 10^4 independent ToyMC ensembles of 7 datasets each (blue solid). The median expected sensitivity of $1.7 \ 10^{25} \ [y]$ is superimposed (red	
	dashed)	88
6.10	Limit on $m_{\beta\beta}$	90
7.1	$^{130}{\rm Te}\ \beta\beta$ decay scheme to excited states $\ \ldots \ $	95
7.2	De-excitation patterns of the 0^+_2 state of 130 Xe	97
7.3	Double beta decay to excited states fully contained signatures	98
7.4	Selected signatures for gs-es $\beta\beta$ decay search and ROI $\ldots \ldots \ldots \ldots$	100
7.5	Distribution of simulated $2\nu\beta\beta$ decay events from pattern 2A0 for contain- ment efficiency computation. The binned maximum likelihood fit is per- formed with a gaussian (signal) over a linear background plus a Compton shoulder on the left of the peak, described by a Gaussian error function	102
7.6	Results of the blinded fit to $0\nu\beta\beta$ (top) and $2\nu\beta\beta$ (bottom) candidate events in the signatures of Table 7.4. For each parameter the mode and smallest interval with 68% probability content of the corresponding marginalized posterior distribution are reported together with the mean and standard deviation. The $S_{1/2}$ parameter indicates the median expected sensitivity for limit setting at 90% C.I. on the $T_{1/2}$ parameter together with the MAD of	
	its distribution.	107
7.7	Top: simultaneous blinded fit result (blue solid) of $0\nu\beta\beta$ candidate events in the signatures of Table 7.4. The background is modeled as a uniform distribution, the background index is compatible with 0 in all signatures. An indication of non-vanishing background was found in the 2A2-2B3 signature. Bottom: posterior pdfs. The subscripts 0, 1, 2 refer to the 2A0-2B1, 2A2-2B3	
	and 3A0 signature respectively	108

78	Top: simultaneous blinded fit result (blue solid) of $2\mu\beta\beta$ candidate events
1.0	in the signatures of Table 7.4. The background is modeled as a uniform
	listerilistics in all simulations second the 2A0 2D1 submer a linear second second
	distribution in all signatures except the 2A0-2B1, where a linear component
	was introduced. The background index is compatible with 0 just in signature
	3A0. An indication of non-vanishing background slope was found in the 2A0-
	2B1 signature. Bottom: posterior pdfs. The subscripts $0, 1, 2$ refer to the
	2A0-2B1, 2A2-2B3 and 3A0 signature respectively
7.9	Distribution of 90% C.I. marginalized upper limits on $T_{1/2} = \log 2/\Gamma$ for
	$0\nu\beta\beta$ decay (top) and $2\nu\beta\beta$ decay (bottom) obtained from Toy MC simu-
	lations. The median of each distribution is highlighted (black dashed line),
	the upper limit obtained with a fit to the unblinded data is shown (red solid
	line)
7.10	Unblinded stat-only $\beta\beta$ gs-es fit result
7.11	Unblinded fit $\beta\beta$ decay to excited 0^+_2 state
7.12	Summary table of the effect of the introduction of each systematic uncer-
	tainty as the corresponding set of nuisance parameters for $0\nu\beta\beta$ (top) and
	$2\nu\beta\beta$ decay search (bottom). The effect of the introduction of each con-
	tribution to systematic uncertainty and the stat-only fit is shown as a $\%$
	difference with respect to the stat-only fit. The relevant parameters are the
	position of the the global mode $\hat{\hat{\Gamma}}$ of the rate parameter, the marginalized
	mode $\hat{\Gamma}$, the width of the 50% smallest interval of the marginalized posterior
	pdf Δ_{50} , and the 90 % C.I. lower limit on $T_{1/2} \doteq \log 2/\Gamma$
7.13	Unblinded fit results $0\nu\beta\beta$ with all systematics. Scaling lineshape parame-
	ters are omitted for clarity. The results are displayed as mean and standard
	deviation of the marginalized posterior pdf for each fit parameter 118
7.14	Unblinded fit results $2\nu\beta\beta$ with all systematics. Scaling lineshape parame-
	ters are omitted for clarity. The results are displayed as mean and standard
	deviation of the marginalized posterior pdf for each fit parameter. $\dots \dots \dots 119$

Acknowledgements / Ringraziamenti

Welcome to the most interesting part of this thesis, maybe the only one meant for the general public, perhaps the first one (and only) that people will read. In the end who cares about double beta decay, right? I have a long list of people to thank this time, and I have to admit that - even if I normally try avoiding pleasantries - this time I feel the urge of writing them down, so that hopefully the people they're addressed at will read them, and I won't have to face the embarassing part where I stand close to them and thank every one in person, maybe even with a handshake or a hug. Before I start I want to reassure you that it's gonna be fully in english, as a form of respect for all non-italian colleagues that will ever read it. I had the chance of improving my english during my PhD and in CUORE, and I feel this is an appropriate way to say *thank you*.

First of all I have to mention Francesco Vissani, who not only contributed to my admission to GSSI and all that followed, but helped a lot in welcoming the all of us new students of Astroparticle Physics 32^{nd} cycle in the *foreign* city of L'Aquila. In the toughest times during my PhD I always could count on his encouragement, in easier times we could laugh together as we did in Rende. Thanks.

A special mention goes to Claudia Tomei and Carlo Bucci, who always managed to find some time to listen to my ruminations, to answer my questions, to review my work, to suggest which path would be the most effective to get to the final result. And I believe that since I ended up writing this section they were damn right. I believe the support from PhD advisors can be different from the expectations, but is always crucial. I've heard about a theory of neuron development in grad students from some Paolo I've met somewhere? Without that support it's way harder. Thanks for never leaving me without your help.

I need especially to thank Francesca, who found unexpected strength and dragged me out of despair when I was about to quit the whole PhD program. I am sure this work would not exist if it weren't for her.

Now let us move to some of the most notable colleagues and mates in the CUORE crew. Just as much as neutrinos are always left, Alice is always right, for example. Thanks for having done such a great job in the crazy days of the TAUP19-* git branches, for having questioned every single assumption or line of code and for our almost daily skype calls. Lots of the details of this work were clarified during those calls. Then I need to thank

Valentina and her alpaca. They always made good targets for my frustration, both literally and metaphorically. Thanks for understanding my mood and the frustration that comes together with software and data analysis. Thanks for the countless discussions about this or that detail of either software configuration, computing cluster slowness, data analysis, for having forced me to long breakfast and lunch breaks and for all the CAVALLO. Moreover, a long list of unpublished music hits must be documented:

- la valigia
- parapappappà
- miss-K la $\pi\pi$
- accende la tivi
- sempre dritto per un po'
- non riuscivo a chiudere la bocca
- la catapulta
- hit dei record
- basta Valentina, basta! (le cose senza senso)

At this point I realize I could make a whole chapter of acknowledgements rather than just few lines! Maybe you're tired of reading, but I need to go on. I need to mention Irene, for the support in my early days in CUORE. Thanks for getting me involved in the analysis of the temperature scan, for your help in the data processing, and for the patience when I was squashed on the sofa because I had enough of QThings. Not to mention the lampredotto evenings and the food truck dream. I need to thank Laura for the same reason, I felt I could help while learning something. And that there's nothing wrong in being a newbie as long as you want to work hard and learn, because there's people around who's willing to teach. Thanks for having taken me underground and explained what every single power cord is there for. Thanks to all the *calcetto* folks, even there I found out that you don't need to be a champion in order to have fun and that in the end, if you play often enough, there's free beer! I cannot mention everyone, because I am sure that I would forget someone and he might be sad about it. So I want just to say that, overall, was it for the clean air of Assergi, the people from there, the ones that I met from all corners of the world, but I've had such a great time overall that I wanted to cry when I had to move back to Roma. Should I mention all the shepards that allowed my body fat to hit levels I'd never thought possible? Let's do it! Should I also mention all what has happened when the Collaboration gathered in Sestri, in Berkeley or in Toyama? Maybe not. But I need to thank Stefano in particular and Jon. You are great singers, besides all the support I got from you about excited states. And I was about to forget Alex and Odysse, which made very good flatmates. I have always been happy to come back and share a meal with



FIG. C.1: Rey, a Jedi cat

you guys, and I need to thank you because you helped a lot in making the GSSI house my home. And for the record, Odysse's amatriciana is worth a try. I am pretty sure I wouldn't be able to make a better one. Special thanks especially for having taken care of Rey (see Fig. C.1), together with Mimmo and Marianna. She grew tall, furry and happy also because of you when I was away. Thanks, on her behalf as well.

Thanks to all the GSSI friends, not only the great 10 AP-XXXII folks, but everyone who was there to play table tennis or drink at Garibaldi's or eat at Puledro's.

In the end, to conclude, I want to mention one thing that I learned. Maybe the most important one. When it's a matter of integers 1 + 1 makes 2, no way. When it comes to people, though,

$$\sum_{i=1}^n 1 \gg n$$

where 1 stands for "a person" and the summation is meant as the combination of forces in a team, in a group of friends, colleagues or a family. n would be a sort of *measure* of the corresponding result that could be obtained if everyone were on his own. Cooperation is crucial, in all aspects of life. And the more cooperative, communicative and proactive you are, the better it is for you and the ones around you. Maybe this can sound a bit weird as an acknowledgement. But it's the main life lesson I was taught during the past 4 years. And I'm grateful for that.

Bibliography

- [1] W. Pauli. Letter to the Tübingen conference, dec 1930.
- [2] E. Fermi. Tentativo di una teoria dell'emissione dei raggi beta. Ric. Sci., 4:491–495, 1933.
- [3] E. Fermi. Tentativo di una Teoria Dei Raggi β. Nuovo Cim., 11: 1, 1934.
- [4] E. Fermi. Versuch einer Theorie der β-Strahlen. I. Z. Phys., 88:161–177, 1934.
- [5] F. Reines and C. L. Cowan. Detection of the free neutrino. Phys. Rev., 92:830–831, 1953.
- [6] Frederick Reines and Clyde L. Cowan. The neutrino. Nature, 178:446–449, 1956.
- [7] G. Danby, J.M. Gaillard, Konstantin A. Goulianos, L.M. Lederman, Nari B. Mistry, M. Schwartz, and J. Steinberger. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.*, 9:36–44, 1962.
- [8] K. Kodama et al. Observation of tau neutrino interactions. Phys. Lett. B, 504:218–224, 2001.
- [9] N. Agafonova et al. Final results of the opera experiment on ν_{τ} appearance in the cngs neutrino beam. *Phys. Rev. Lett.*, 120:211801, May 2018.
- [10] R. Davis. A review of the Homestake solar neutrino experiment. Prog. Part. Nucl. Phys., 32:13–32, 1994.
- Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562– 1567, 1998.
- [12] Q.R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ⁸B solar neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, 2001.
- [13] B. Aharmim et al. Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory. *Phys. Rev. C*, 88:025501, 2013.
- [14] John N. Bahcall, M.H. Pinsonneault, and Sarbani Basu. Solar models: Current epoch and time dependences, neutrinos, and helioseismological properties. Astrophys. J., 555:990–1012, 2001.
- [15] A.S. Brun, S. Turck-Chieze, and J.P. Zahn. Standard solar models in the light of new helioseismic constraints. 2. mixing below the convective zone. Astrophys. J., 525:1032–1041, 1999.

- [16] L. Wolfenstein. Neutrino Oscillations in Matter. Phys. Rev. D, 17:2369–2374, 1978.
- [17] S.P. Mikheyev and A.Yu. Smirnov. Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos. Sov. J. Nucl. Phys., 42:913–917, 1985.
- [18] Guido Fantini, Andrea Gallo Rosso, Francesco Vissani, and Vanessa Zema. Introduction to the Formalism of Neutrino Oscillations. Adv. Ser. Direct. High Energy Phys., 28:37–119, 2018.
- [19] S.M. Bilenky. Neutrino. History of a unique particle. Eur. Phys. J. H, 38:345–404, 2013.
- [20] Ettore Majorana. Teoria simmetrica dell'elettrone e del positrone. Nuovo Cim., 14:171–184, 1937.
- [21] Samoil Bilenky. Introduction to the physics of massive and mixed neutrinos, volume 817. 2010.
- [22] K. Abe et al. Constraint on the matter-antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, 2020.
- [23] P. A. Zyla and others (Particle Data Group). to be published in Prog. Theor. Exp. Phys. 2020, 083C01, 2020.
- [24] M. Goeppert-Mayer. Double beta-disintegration. Phys. Rev., 48:512–516, 1935.
- [25] J. Kotila and F. Iachello. Phase space factors for double-β decay. Phys. Rev., C85:034316, 2012.
- [26] F. T. Avignone, S. R. Elliott, and J. Engel. Double Beta Decay, Majorana Neutrinos, and Neutrino Mass. *Rev. Mod. Phys.*, 80:481–516, 2008.
- [27] J. Engel and J. Menéndez. Status and future of nuclear matrix elements for neutrinoless double-beta decay: a review. *Rep. Prog. Phys.*, 80(4):046301, 2017.
- [28] J. Menendez, A. Poves, E. Caurier, and F. Nowacki. Disassembling the Nuclear Matrix Elements of the Neutrinoless beta beta Decay. *Nucl. Phys. A*, 818:139–151, 2009.
- [29] Andrei Neacsu and Mihai Horoi. Shell model studies of the ¹³⁰Te neutrinoless double-beta decay. Phys. Rev. C, 91:024309, 2015.
- [30] Fedor Šimkovic, Vadim Rodin, Amand Faessler, and Petr Vogel. 0νββ and 2νββ nuclear matrix elements, quasiparticle random-phase approximation, and isospin symmetry restoration. Phys. Rev. C, 87(4):045501, 2013.
- [31] Juhani Hyvärinen and Jouni Suhonen. Nuclear matrix elements for $0\nu\beta\beta$ decays with light or heavy Majorana-neutrino exchange. *Phys. Rev. C*, 91(2):024613, 2015.
- [32] M.T. Mustonen and J. Engel. Large-scale calculations of the double-β decay of ⁷⁶Ge,¹³⁰ Te,¹³⁶ Xe, and ¹⁵⁰Nd in the deformed self-consistent Skyrme quasiparticle randomphase approximation. *Phys. Rev. C*, 87(6):064302, 2013.
- [33] Tomas R. Rodriguez and G. Martinez-Pinedo. Energy density functional study of nuclear matrix elements for neutrinoless $\beta\beta$ decay. *Phys. Rev. Lett.*, 105:252503, 2010.

- [34] Nuria López Vaquero, Tomás R. Rodríguez, and J. Luis Egido. Shape and pairing fluctuations effects on neutrinoless double beta decay nuclear matrix elements. *Phys. Rev. Lett.*, 111(14):142501, 2013.
- [35] J.M. Yao, L.S. Song, K. Hagino, P. Ring, and J. Meng. Systematic study of nuclear matrix elements in neutrinoless double-β decay with a beyond-mean-field covariant density functional theory. *Phys. Rev. C*, 91(2):024316, 2015.
- [36] J. Barea, J. Kotila, and F. Iachello. $0\nu\beta\beta$ and $2\nu\beta\beta$ nuclear matrix elements in the interacting boson model with isospin restoration. *Phys. Rev. C*, 91(3):034304, 2015.
- [37] Francesco Iachello, Jenni Kotila, and Jose Barea. Quenching of g_A and its impact in double beta decay. *PoS*, NEUTEL2015:047, 2015.
- [38] M. Aker et al. Improved upper limit on the neutrino mass from a direct kinematic method by katrin. *Phys. Rev. Lett.*, 123:221802, Nov 2019.
- [39] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. 2018. arXiv: 1807.06209.
- [40] G. Benato. Effective Majorana mass and neutrinoless double beta decay. The European Physical Journal C, 75:563, 2015.
- [41] D. Q. Adams et al. Improved Limit on Neutrinoless Double-Beta Decay in ¹³⁰Te with CUORE. Phys. Rev. Lett., 124:122501, Mar 2020.
- [42] Francesco Vissani. Signal of neutrinoless double beta decay, neutrino spectrum and oscillation scenarios. JHEP, 06:022, 1999.
- [43] M. Tanabashi et al. Review of Particle Physics. Phys. Rev. D, 98(3):030001, 2018.
- [44] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. Matter and Antimatter in the Universe. New J. Phys., 14:095012, 2012.
- [45] A D Sakharov. Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe. JETP Lett. (USSR)(Engl. Transl.), 5: 24-7(Jan. 1, 1967).
- [46] C. Jarlskog. Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation. *Phys. Rev. Lett.*, 55:1039, 1985.
- [47] Apostolos Pilaftsis and Thomas E.J. Underwood. Electroweak-scale resonant leptogenesis. *Phys. Rev. D*, 72:113001, 2005.
- [48] R. G. H. Robertson. Empirical Survey of Neutrinoless Double Beta Decay Matrix Elements. Mod. Phys. Lett. A, 28:1350021, 2013.
- [49] F. Alessandria et al. CUORE crystal validation runs: results on radioactive contamination and extrapolation to CUORE background. Astropart. Phys., 35:839–849, 2012.
- [50] C. Arnaboldi et al. Production of high purity TeO₂ single crystals for the study of neutrinoless double beta decay. J. Cryst. Growth, 312(20):2999–3008, 2010.
- [51] O. Azzolini et al. Final Result of CUPID-0 Phase-I in the Search for the ⁸²Se Neutrinoless Double-β Decay. *Phys. Rev. Lett.*, 123:032501, Jul 2019.

- [52] G. Anton et al. Search for Neutrinoless Double-β Decay with the Complete EXO-200 Dataset. Phys. Rev. Lett., 123:161802, Oct 2019.
- [53] M. Agostini et al. Probing Majorana neutrinos with double-beta decay. Science, 365(6460):1445–1448, 2019.
- [54] A. Gando et al. Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen. Phys. Rev. Lett., 117:082503, Aug 2016.
- [55] S. I. Alvis et al. Search for neutrinoless double- β decay in ⁷⁶Ge with 26 kg yr of exposure from the Majorana Demonstrator. *Phys. Rev. C*, 100:025501, Aug 2019.
- [56] R. Ardito et al. CUORE: A Cryogenic underground observatory for rare events. 1 2005. arXiv: hep-ex/0501010.
- [57] C. Arnaboldi et al. CUORE: A Cryogenic underground observatory for rare events. Nucl. Instrum. Meth. A, 518:775–798, 2004.
- [58] Irene Nutini et al. The CUORE Detector and Results. J. Low Temp. Phys., 199(1-2):519–528, 2020.
- [59] E. Lubkin. Gentle neutrino scattering and the davis experiment. *Physics Letters A*, 46(7):431 432, 1974.
- [60] T O Niinikoski and Fred Udo. Cryogenic detection of neutrinos? (CERN-NP-Internal-Report-74-6), 1974.
- [61] E. Fiorini and T. O. Niinikoski. Low Temperature Calorimetry for Rare Decays. Nucl. Instrum. Meth. A, 224:83, 1984.
- [62] M. Ambrosio et al. Vertical muon intensity measured with MACRO at the Gran Sasso Laboratory. Phys. Rev. D, 52:3793–3802, 1995.
- [63] H. Wulandari, J. Jochum, W. Rau, and F. von Feilitzsch. Neutron flux at the Gran Sasso underground laboratory revisited. Astropart. Phys., 22:313–322, 2004.
- [64] Chiara Brofferio and Stefano Dell'Oro. Contributed Review: The saga of neutrinoless double beta decay search with TeO₂ thermal detectors. Rev. Sci. Instrum., 89(12):121502, 2018.
- [65] C. Arnaboldi et al. A Calorimetric search on double beta decay of ¹³⁰Te. Phys. Lett. B, 557:167–175, 2003.
- [66] C. Alduino et al. CUORE-0 detector: design, construction and operation. J. Instrum., 11(07):P07009, 2016.
- [67] E. Andreotti et al. ¹³⁰Te Neutrinoless Double-Beta Decay with CUORICINO. Astropart. Phys., 34:822–831, 2011.
- [68] K. Alfonso et al. Search for Neutrinoless Double-Beta Decay of ¹³⁰Te with CUORE-0. Phys. Rev. Lett., 115(10):102502, 2015.
- [69] M. Barucci, C. Brofferio, A. Giuliani, E. Gottardi, I. Peroni, and G. Ventura. Measurement of Low Temperature Specific Heat of Crystalline TeO₂ for the Optimization of Bolometric Detectors. J. Low Temp. Phys., 123(5-6):303–314, 2001.

- [70] H. C. Stahl. Cryogenic Particle Detection. Springer, 2005.
- [71] C. Alduino et al. CUORE-0 detector: design, construction and operation. Journal of Instrumentation, 11(07):P07009–P07009, jul 2016.
- [72] E. E. Haller, N. P. Palaio, M. Rodder, W. L. Hansen, and E. Kreysa. Ntd germanium: A novel material for low temperature bolometers. In R. D. Larrabee, editor, *Neutron Transmutation Doping of Semiconductor Materials*, pages 21–36. Springer US, 1984.
- [73] Y.P. Varshni. Temperature dependence of the energy gap in semiconductors. *Physica*, 34(1):149 – 154, 1967.
- [74] C. Alduino et al. The CUORE cryostat: An infrastructure for rare event searches at millikelvin temperatures. Cryogenics, 102:9 – 21, 2019.
- [75] M. Schwark, F. Pobell, W. P. Halperin, C. Buchal, J. Hanssen, M. Kubota, and R. M. Mueller. Ortho-para conversion of hydrogen in copper as origin of time-dependent heat leaks. J. Low Temp. Phys., 53:685–694, 1983.
- [76] M. Koláč, B. S. Neganov, and S. Sahling. Low Temperature Heat Release from Copper: Ortho-Para Conversion of Hydrogen. J. Low Temp. Phys., 59:547–559, 1985.
- [77] F. Pobell. Matter and Methods at Low Temperatures. Springer, 2007.
- [78] C. Enss and H. Siegfried. Low-Temperature Physics. Springer, 2005.
- [79] H. London. Proc. 2nd Int. Conf. on Low Temp. Phys., page 157, 1951.
- [80] H. London, G. R. Clarke, and E. Mendoza. Osmotic Pressure of He³ in Liquid He⁴, with Proposals for a Refrigerator to Work below 1°K. *Phys. Rev.*, 128:1992–2005, 1962.
- [81] P. Das, R. De Bruyn Ouboter, and K. W. Taconis. Proc. 9th Int. Conf. Low Temp. Phys., page 1253, 1965.
- [82] A. D'Addabbo, C. Bucci, L. Canonica, S. Di Domizio, P. Gorla, L. Marini, A. Nucciotti, I. Nutini, C. Rusconi, and B. Welliver. An active noise cancellation technique for the CUORE Pulse Tube Cryocoolers. *Cryogenics*, 93:56–65, 2018.
- [83] A. Alessandrello et al. Measurements of internal radioactive contamination in samples of Roman lead to be used in experiments on rare events. *Nucl. Instrum. Meth. B*, 142:163–172, 1998.
- [84] S. Di Domizio, A. Branca, A. Caminata, L. Canonica, S. Copello, A. Giachero, E. Guardincerri, L. Marini, M. Pallavicini, and M. Vignati. A data acquisition and control system for large mass bolometer arrays. *Journal of Instrumentation*, 13(12):P12003–P12003, dec 2018.
- [85] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. Nucl. Instrum. Meth., A389:81–86, 1997.
- [86] S. Di Domizio, S. Orio, and M. Vignati. Lowering the energy threshold of large-mass bolometric detectors. *Journal of Instrumentation*, 6(02):P02007–P02007, feb 2011.

- [87] E. Gatti and P. F. Manfredi. Processing the Signals From Solid State Detectors in Elementary Particle Physics. *Riv. Nuovo Cim.*, 9:1–146, 1986.
- [88] A. Campani et al. Lowering the energy threshold of the cuore experiment: Benefits in the surface alpha events reconstruction. *Journal of Low Temperature Physics*, 6 2020.
- [89] C. Alduino et al. First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ¹³⁰Te. *Phys. Rev. Lett.*, 120:132501, Mar 2018.
- [90] C. Alduino et al. Analysis techniques for the evaluation of the neutrinoless double- β decay lifetime in ¹³⁰Te with the CUORE-0 detector. *Phys. Rev. C*, 93:045503, 2016.
- [91] P. C. Mahalanobis. On the generalised distance in statistics. Proc. Natl. Inst. Sci. India, 49, 1936.
- [92] L. Marini. The CUORE experiment: from the commissioning to the first $0\nu\beta\beta$ limit. *PhD* thesis, 2018.
- [93] S. Rahaman et al. Double-beta decay Q values of 116 Cd and 130 Te. *Physics Letters B*, 703(4):412 416, 2011.
- [94] C. Alduino et al. The projected background for the CUORE experiment. Eur. Phys. J. C, 77(8):543, 2017.
- [95] C. Bucci et al. Background study and Monte Carlo simulations for large-mass bolometers. Eur. Phys. J. A, 41(2):155–168, 2009.
- [96] C. Arnaboldi et al. Results from a search for the $0\nu\beta\beta$ -decay of ¹³⁰Te. *Phys. Rev. C*, 78:035502, 2008.
- [97] M Laubenstein, M Hult, J Gasparro, D Arnold, S Neumaier, G Heusser, M Köhler, P Povinec, J.-L Reyss, M Schwaiger, and P Theodórsson. Underground measurements of radioactivity. *Applied Radiation and Isotopes*, 61(2):167 – 172, 2004. Low Level Radionuclide Measurement Techniques - ICRM.
- [98] E. Sala, M. Clemenza, and E. Previtali. High-purity material selection techniques for Rare-Events Physics experiments. *Nuovo Cim. C*, 036(01):54–56, 2013.
- [99] C. Brofferio, A. Cesana, A. Fascilla, L. Garlati, A. Giuliani, M. Pedretti, G.L. Raselli, and M. Terrani. Characterization of an underground site in northern italy in view of low radioactivity measurements. *Journal of Environmental Radioactivity*, 71(2):159 – 173, 2004.
- [100] Andrea Borio, Antonio Cammi, Davide Chiesa, Massimiliano Clemenza, S. Manera, Marisa Nastasi, Luca Pattavina, R. Ponciroli, Stefano Pozzi, Michele Prata, E. Previtali, Andrea Salvini, and M. Sisti. Triga reactor absolute neutron flux measurement using activated isotopes. *Progress in Nuclear Energy*, 70:249–255, 01 2014.
- [101] Andrea Borio, Antonio Cammi, Massimiliano Clemenza, V. Memoli, Luca Pattavina, and E. Previtali. Benchmark evaluation of reactor critical parameters and neutron fluxes distributions at zero power for the triga mark ii reactor of the university of pavia using the monte carlo code mcnp. *Progress in Nuclear Energy - PROG NUCL ENERGY*, 52:494–502, 07 2010.

- [102] E. Buccheri, M. Capodiferro, S. Morganti, F. Orio, A. Pelosi, and V. Pettinacci. An assembly line for the construction of ultra-radio-pure detectors. *Nucl. Instrum. Meth. A*, 768:130–140, 2014.
- [103] M. Clemenza, C. Maiano, L. Pattavina, and E. Previtali. Radon-induced surface contaminations in low background experiments. *Eur. Phys. J. C*, 71:1805, 2011.
- [104] J. Allison et al. Recent developments in geant4. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 835:186 – 225, 2016.
- [105] S. Agostinelli et al. Geant4 a simulation toolkit. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3):250 - 303, 2003.
- [106] J. Allison et al. Geant4 developments and applications. IEEE Transactions on Nuclear Science, 53(1):270–278, Feb 2006.
- [107] W. C. Haxton and G. J. Stephenson. Double beta Decay. Prog. Part. Nucl. Phys., 12:409– 479, 1984.
- [108] Masaru Doi, Tsuneyuki Kotani, Hiroyuki Nishiura, and Eiichi Takasugi. Double Beta Decay. Progress of Theoretical Physics, 69(2):602–635, 02 1983.
- [109] S. Pozzi. Search for double-beta decay of ¹³⁰Te to the excited states of ¹³⁰Xe in cuore-0. PhD thesis, 2017.
- [110] M. et al. Aglietta. Muon "depth-intensity" relation measured by the lvd underground experiment and cosmic-ray muon spectrum at sea level. *Phys. Rev. D*, 58:092005, Oct 1998.
- [111] S. et al. (MACRO Collaboration) Ahlen. Muon astronomy with the macro detector. The Astrophysical Journal, 412:301–311, 06 1993.
- [112] CUORE Collaboration. The cuore background model. paper in preparation.
- [113] Martyn Plummer et al. Jags: A program for analysis of bayesian graphical models using gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing, volume 124, pages 1–10. Vienna, Austria, 2003.
- [114] CUORE Collaboration. Measurement of the 2 neutrino double beta decay half life in cuore. presented at TAUP 2019, paper in preparation.
- [115] Matthew Redshaw, Brianna J. Mount, Edmund G. Myers, and Frank T. Avignone. Masses of ¹³⁰Te and ¹³⁰Xe and double- β -decay q value of ¹³⁰Te. *Phys. Rev. Lett.*, 102:212502, May 2009.
- [116] N. D. Scielzo, S. Caldwell, G. Savard, J. A. Clark, C. M. Deibel, J. Fallis, S. Gulick, D. Lascar, A. F. Levand, G. Li, J. Mintz, E. B. Norman, K. S. Sharma, M. Sternberg, T. Sun, and J. Van Schelt. Double-β-decay q values of ¹³⁰Te, ¹²⁸Te, and ¹²⁰Te. *Phys. Rev. C*, 80:025501, Aug 2009.
- [117] S. Rahaman et al. Double-beta decay Q values of 116 Cd and 130 Te. *Physics Letters B*, 703(4):412 416, 2011.

- [118] C. et al. Alduino. Analysis techniques for the evaluation of the neutrinoless double- β decay lifetime in ¹³⁰Te with the cuore-0 detector. *Phys. Rev. C*, 93:045503, Apr 2016.
- [119] J. et al. Amaré. Cosmogenic activation of materials. AIP Conference Proceedings, 785(1):267– 271, 2005.
- [120] J. Kotila and F. Iachello. Phase space factors for double- β decay. *Phys. Rev. C*, 85:034316, 2012.
- [121] F. Šimkovic, V. Rodin, A. Faessler, and P. Vogel. $0\nu\beta\beta$ and $2\nu\beta\beta$ nuclear matrix elements, quasiparticle random-phase approximation, and isospin symmetry restoration. *Phys. Rev. C*, 87(4):045501, 2013.
- [122] A. Meroni, S. T. Petcov, and F. Simkovic. Multiple CP non-conserving mechanisms of $(\beta\beta)_{0\nu}$ -decay and nuclei with largely different nuclear matrix elements. *JHEP*, 02:025, 2013.
- [123] Björn Lehnert. Excited State Transitions in Double Beta Decay: A brief Review. EPJ Web Conf., 93:01025, 2015.
- [124] A. S. Barabash et al. Two neutrino double beta decay of Mo-100 to the first excited 0⁺ state in Ru-100. *Phys. Lett.*, B345:408–413, 1995.
- [125] A. S. Barabash, F. Hubert, P. Hubert, and V. I. Umatov. Double beta decay of Nd-150 to the first 0⁺ excited state of Sm-150. *JETP Lett.*, 79:10–12, 2004. [Pisma Zh. Eksp. Teor. Fiz.79,12(2004)].
- [126] A. S. Barabash. Double beta decay to the excited states: Review. AIP Conf. Proc., 1894(1):020002, 2017.
- [127] E. Andreotti et al. Search for double- β decay of ¹³⁰Te to the first 0⁺ excited state of ¹³⁰Xe with CUORICINO. *Phys. Rev.*, C85:045503, 2012.
- [128] C. Alduino et al. Double-beta decay of ¹³⁰Te to the first 0⁺ excited state of ¹³⁰Xe with CUORE-0. Eur. Phys. J., C79(9):795, 2019.
- [129] Pekka Pirinen and Jouni Suhonen. Systematic approach to β and $2\nu\beta\beta$ decays of mass a = 100 136 nuclei. *Phys. Rev. C*, 91:054309, May 2015.
- [130] Jouni Suhonen and Osvaldo Civitarese. Probing the quenching of g_A by single and double beta decays. *Physics Letters B*, 725(1):153 157, 2013.
- [131] J. Suhonen and O. Civitarese. Single and double beta decays in the a=100, a=116 and a=128 triplets of isobars. Nuclear Physics A, 924:1 – 23, 2014.
- [132] Balraj Singh. Nuclear data sheets for a = 130. Nuclear Data Sheets, 93(1):33-242, 2001.
- [133] Khristo N. Boyadzhiev. Close encounters with the stirling numbers of the second kind. Mathematics Magazine, 85(4):252–266, 2012.
- [134] J.R. Klein and A. Roodman. Blind analysis in nuclear and particle physics. Ann. Rev. Nucl. Part. Sci., 55:141–163, 2005.

- [135] G. Fantini on behalf of the CUORE collaboration. Sensitivity to double beta decay of ¹³⁰Te to the first 0⁺ excited state of ¹³⁰Xe in cuore. Journal of Physics: Conference Series, 1468, 02 2020.
- [136] J. Allison, K. Amako, John Apostolakis, Pedro Arce, Makoto Asai, T. Aso, Enrico Bagli, A. Bagulya, S. Banerjee, Guy Barrand, B.R. Beck, A.G. Bogdanov, D. Brandt, Jeremy Brown, H. Burkhardt, Ph Canal, Daniel Ott, Stephane Chauvie, K. Cho, and H. Yoshida. Recent developments in geant4. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 835, 07 2016.
- [137] Manuela A Fehr, Mark RehkÄmper, and Alex N Halliday. Application of mc-icpms to the precise determination of tellurium isotope compositions in chondrites, iron meteorites and sulfides. *International Journal of Mass Spectrometry*, 232(1):83 – 94, 2004.
- [138] Giulio D'Agostini. Bayesian Reasoning in Data Analysis. WORLD SCIENTIFIC, 2003.
- [139] James E Gentle, Wolfgang Karl Härdle, and Yuichi Mori. *Handbook of computational statistics: concepts and methods.* Springer Science & Business Media, 2012.
- [140] Allen Caldwell, Daniel Kollar, and Kevin Kroninger. Bat the bayesian analysis toolkit. Computer Physics Communications, 180(11):2197 – 2209, 2009.