Università degli Studi di Milano Bicocca
Dottorato di ricerca in Fisica e Astronomia XXIII CICLO

Data Analysis for Neutrinoless Double Beta Decay

Thesis submitted
by
Marco Andrea Carrettoni

A doctoral dissertation in partial fullfillment
of the requirements to obtain the Degree of
Doctor Philosophiae in Physics

March 2011

Supervisor:	**Prof. Maura Pavan**
School Coordinator:	**Prof. Claudio Destri**

It was while gliding through these latter waters
that one serene and moonlight night, when all
the waves rolled by like scrolls of silver; and by
their soft, suffusing seethings, made what seemed
a silvery silence, not a solitude: on such a silent
night a silvery jet was seen far in advance of the
white bubbles, at the bow.

- Hermann Melville, Moby Dick
or the Whale

A mio nonno Franco

Contents

Introduction

Double Beta Decay ($\beta\beta$) is a rare transition between two isobars, involving the change of the nuclear charge Z by two units. In Nature we have several even-even nuclei for which this is the only allowed decay mode. While the transition where 2 electrons and 2 neutrinos are emitted (2ν-DBD) does not imply "special" properties for the neutrinos (and has been observed for various isotopes), this is not the case for the neutrinoless channel (0ν-DBD). Indeed, despite being energetically possible, the neutrinoless transition violates the lepton number by 2 units and is possible only if the neutrino is a massive Majorana particle [1, 2, 3]. 0ν-DBD searches have been pursued since more than half a century and today they experience a renewed interest, thanks to the discovery of neutrino oscillations [4, 5, 6, 7]. However oscillations are not enough to investigate the very heart of neutrino physics i.e. what is their nature and why neutrinos are extraordinary light. Several theoretical speculations point toward a mass generation mechanism that imply a Majorana character of neutrinos and that indicates in the 0ν-DBD process the unique tool with a discovery potential. The 0ν-DBD transition can proceed through different mechanisms among which the simplest and favourit one is the "pure Majorana mass". In that case the 0ν-DBD observation would not only provide evidence of lepton violation and of the Majorana character of this particle, but would result in a measurement of the Majorana mass $m_{ee} = |\sum m_i U_{ei}^2|$ (where m$_i$ are the three mass eigenstates of neutrinos and U$_{ei}$ are the PNSM matrix elements) through the relation:

$$\frac{1}{\tau_{1/2}^{0\nu}} = m_{ee}^2 F_N = m_{ee}^2 G^{0\nu} |M^{0\nu}|^2 \tag{1}$$

Here $G^{0\nu}$ is the two-body phase-space factor and $M^{0\nu}$ is the 0ν-DBD Nuclear Matrix Element (NME) their product F_N being called "nuclear factor of merit". The name refers to the fact that, according to (1), F_N directely influences the experimental sensitivity to m$_{ee}$. The main uncertainty in deriving m$_{ee}$ (or an upper limit on it) from the experimental result on $\tau_{1/2}^{0\nu}$ comes from the NME element which is a theoretical evaluation still affected by a large spread among the adopted nuclear models and their implementation [8, 9, 10, 11]. As a practical

solution to this uncertainty source, several isotopes candidates to 0ν-DBD have been studied so far. Among them those that yielded the most stringent limits on m$_{ee}$ (within the NME spread) are ^{76}Ge [12, 13, 14], ^{100}Mo [15], ^{130}Te [16] and ^{136}Xe [17]. In all but one [18] case only upper bounds on the Majorana mass have been reported. In this paper we discuss the final 0ν-DBD result of the CUORICINO experiment, that yields the more stringent bound on m$_{ee}$ based on ^{130}Te studies. In this thesis I will describe my contributions to the data analysis of CUORICINO data, giving at the same time the theoretical and experimental context of this experiment.
This thesis is subdivided in two parts: the first three chapters should give an idea of the CUORICINO, from its theoretical context up to the evaluation of the 0ν-DBD limit, while the last chapters are a description of my contributions to the analysis of this experiments.
At first, I'll summarize the Neutrinoless Double Beta Decay theory, describing why its observation would be so important in the Standard Model. In the second chapter, I'll give a description of the experimental technique used in CUORICINO, the bolometric approach, which has been chosen to face the challenge of observing such a rare process. In the third chapter, a summary of the main steps required to infere a limit on 0ν-DBD , from the detector's DAQ to the fit of the final spectrum will be given.
In the fourth chapter, I'll focus on my contributions on Digital Signal Processing, one of the core part of my thesis work, which consisted on the analysis perfomed on bolometric signals used to gain in resolution and reduce the background. The fifth chapter gives a detailed description of the statistical studiesthat lead to the evaluation of the limit on the half-life of , while in the sixth chapter I'll describe a different way of proposing a scientific result when dealing with quantities close to the experiment's sensitivity as the CUORICINO limit on 0ν-DBD .

Chapter 1

Neutrinoless Double Beta Decay

Neutrino properties are well described by the standard electroweak theory that was finally formulated in the late 60's in the works of S. Glashow, A. Salam, and S. Weinberg. Together with quantum chromodynamics (QCD), this theory forms the so called Minimal Standard Model (MSM) of particle physics. All the existing experimental data are in good agreement with MSM, except for observed anomalies in neutrino processes. Small neutrino masses can not be explained by the Standard Higgs mechanism used to describe fermion masses. The most simple extention of the MSM that could explain the neutrino mass generation (the "see-saw" mechanism) is based on the assumption of the violation of the total lepton number and a Majorana nature of the neutrino, which means that it can be considered as its own anti-particle.
The succesfull results of the oscillation experiments did prove that at least one neutrino has a mass different to zero but they are not sufficient to solve the neutrino puzzle.
A probe which is very sensitive to the neutrino nature is the Neutrinoless Double Beta decay.
In this chapter the theoretical context of this rare process is summarized. Few details will be given concerning the Standard Model, the neutrino oscillations theory and the standard see-saw mechanism. This will give a feeling of how many informations can be extracted by the study of a single phenomenon and, finally, a description of this rare process will show why it represents such an impressive challenge.

1.1 The Standard Model

The Standard Model is the most complete description of weak, strong and electromagnetic interactions that we posses. Since its birth in the 70's it proved to be extremely resistent to every experimental test.

The founding principle of this model is based, as many of the most elegant axioms, on strong requirements, or desire, of symmetry: the differential equations that describe the evolution of fields that reprent particles are bound to be invariant under a particular kind of transformations called "local gauge group". The term *gauge* refers to a class of operations that act on internal degrees of freedom of a theory.

In Field Theory, the simplest gauge transformation takes the form:

$$\psi(x,t) \rightarrow \; e^{i\theta(x)}\psi(x,t) \tag{1.1}$$

where θ is a scalar (but can also be a vector or even a tensor), called the generator of a generic group of transformations (in this case, phase shifts). If θ is constant these trasformations are called "global" and are indeed important, since it is from them that, by means of Noether's theorem, we can derive important conservation laws (i.e. energy, momentum or charge conservation). If θ has instead an explicit dependence from a coordinate, equation (1.1) will describe a "local" gauge.

Imposing a simmetry under this kind of transformation is not trivial. The Lagrangian $\mathcal{L}$ that describes the evolution a particle field ψ is:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{1.2}$$

and it can't be invariant under a local gauge, since dependence of θ by the position will add a term, proportional to $\partial_{\mu}\,\theta$.

The covariance of this equation will be guaranteed if we choose an more general definition for the operator ∂_{μ} that can absorb the term $\partial_{\mu}\theta(x)$. The new form for this operator contains new fields that satisfy specific transformation laws.

A possible functional form of this operator is a generalization of the derivative, called "covariant derivative". To form this operator we must include a new vector field with proper trasformation properties under (1.1) so that the additional term in the lagrangian can be canceled. A possible choice is a vector field A_{μ} that transforms as:

$$A_{\mu} \rightarrow A_{\mu} - \partial_{\mu}\theta. \tag{1.3}$$

This field can be included in the new operator, which will be defined as:

$$D_{\mu} \equiv \partial_{\mu} - iA_{\mu} \tag{1.4}$$

The price we must pay, to guarantee the symmetry, is that the lagrangian of free particle becomes:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}D_{\mu}\psi - m\bar{\psi}\psi = \; = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + \bar{\psi}\gamma^{\mu}\psi A_{\mu} \tag{1.5}$$

The last term has exactly the form of an interaction term: the electromagnetic interaction is in fact expressed as the product of a term like $\bar{\psi}\gamma^{\mu}\psi$ (which is a density current) times a vector field [19].
In order to guarantee the symmetry, an interaction field must to appear.
In this case, $\theta(x)$ is a scalar field depending on the position and can be considered a potential. This symmetry belongs to the group of phase shifts, generally labelled as $U(1)$, and it is an Abelian group. Weak and Strong forces are derived by extending this procedure to non-Abelian groups. As an example, weak interactions can be considered as rotations (whose generator is the group $SU(2)$) in a bidimentional space on vectors whose components are fermions with the same flavour:

$$\text{leptons} \qquad \begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{1.6}$$

$$\text{quarks} \qquad \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} s \\ c \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.7}$$

Actually, in order to include the experimental evidence of parity violation in weak interaction fenomenology, the requirement of invariance under $SU(2)$ is imposed only between left-handed projections of these doublets (or right-handed if we are dealing with anti-particles). This is the main reason why, in the β^- decay, only right-handed anti-neutrino can be produced.
In a similar way, the strong interactions will emerge by imposing a symmetry of the Lagrangian under $SU(3)$ (the group of rotations in a three dimentional space) that will act on the three components of color space in which the quarks are grouped.
Gathering together all these symmetry information, the MSM can be simply defined by the combined symmetry group:

$$U(1) \times SU(2)_L \times SU(3) \tag{1.8}$$

The fields that emerges for the symmetry of this group are called "gauge fields", and each has a corresponding particle, called "boson". Example of these bosons are the photon for the electromagnetism, or the W^+,W^- and Z^0 for the weak interactions.

1.2 Particle masses

The symmetries of the Lagrangian summarized in (1.8) have an important requirement: the bosons describing the interactions have to be massless. While

this is true for the photon, this is not true for the bosons that mediates weak interactions.
The problem is that the most general term that can give a mass to a free particle has the form $m\bar{\psi}\psi$, which explicitly violates the required simmetry.
This problem is overcome by the mechanism of “spontaneus simmetry breaking” [19] that, in a similar way that was followed for the forces, introduces in the Lagrangian a new scalar field, the Higgs field, that interacts with the bosons giving the required mass term for the particles W^+, W^- and Z^0. The Higgs field gives mass also to the fermions by their coupling with this scalar field.
Anyhow, in this description of the Standard Model, there are still many open questions concerning neutrinos.
If we express the term in the Lagrangian that gives mass to the fermions on chirality eigenstates we obtain:

$$\mathcal{L}_m = m_\nu \overline{\nu}\nu = m_\nu(\overline{\nu}_L \nu_R + \overline{\nu}_R \nu_L) \tag{1.9}$$

but the Standard Model can’t include the existance of right-handed neutrinos, thus it is not possible to build a massive term like this: within its contest we are not able to explain simoultaneousely the fact that neutrinos have a defined elicity but do have a mass, as shown by the phenomenon of oscillation that will be described in the next section.

1.3 Neutrino Oscillation

The most astonishing evidence of the massive character of neutrinos has come from oscillation experiments.
In quantum mechanics, there is an oscillation phenomenon whenever the probability of measuring the flavour eigenstate of a particle has a periodic behaviour which is a function of the travelled distance. The base assumption needed to explain this quantum effect is linked to the possibility that the flavour eigenstates $|\nu_\alpha\rangle$ (or eigenstates of the interactions) are linear superposition of mass eigenstates $|\nu_i\rangle$:

$$|\nu_\alpha\rangle = \sum_{i=1}^{n} U^\star_{\alpha i} |\nu_i\rangle \tag{1.10}$$

here n is the number of eigenstates and $U^\star_{\alpha i}$ the generic matrix element (the “mixing” matrix) that gives the weights of the linear combination.
In case of two neutrinos generations, it can be shown [4] that the matrix U (in this case 2×2) can be parametrized only by an angle θ. By studying the projection of a flavour eigenstate on the other, expressed as mass eigenstate

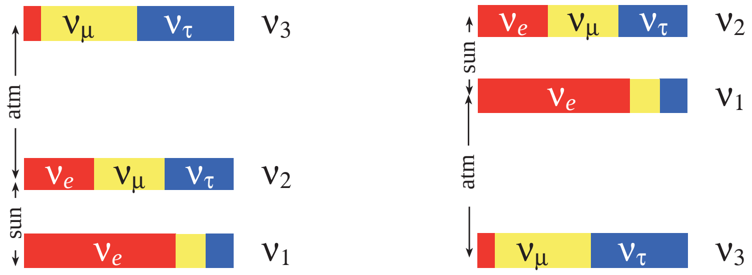


Figure 1.1: Possible neutrino spectra: inverted (left), normal (right)

		Best fit	99% CL range
Δm^2_{12}	$=$	$(7.58 \pm 0.21)\, 10^{-5}\text{eV}^2$	$(7.1 \div 8.1)\, 10^{-5}\text{eV}^2$
$\lvert\Delta m^2_{23}\rvert$	$=$	$(2.40 \pm 0.15)\, 10^{-3}\text{eV}^2$	$(2.1 \div 2.8)\, 10^{-3}\text{eV}^2$
$\tan^2\theta_{12}$	$=$	0.484 ± 0.048	$31^\circ < \theta_{12} < 39^\circ$
$\sin^2 2\theta_{23}$	$=$	1.02 ± 0.04	$37^\circ < \theta_{23} < 53^\circ$
$\sin^2 2\theta_{13}$	$=$	0.07 ± 0.04	$0^\circ < \theta_{13} < 13^\circ$

Table 1.1: Summary of present information on neutrino masses and mixings from oscillation data.

superpositions, it will be possible to obtain the transition probability by the formula:

$$P(\nu_\alpha \to \nu_\beta) \simeq \sin^2 2\theta \sin^2\left(1.27\frac{L[Km]}{E_\nu[GeV]}\Delta m^2_{12}\right) \tag{1.11}$$

where L ed E are, respectively, the distance from the origin and energy of the particle, while $\Delta m^2_{12} = |m_1^2 - m_2^2|$.
By measuring the disappearance probability, as the ratio between the observed and expected flux of a certain flavour, one can link this information to this squared mass differences. If an experiment finds that this probability is different from zero, it means that at least one neutrino has a mass eigenvalue different from zero.

The most important experiments deal with the disappearance of electronic neutrinos (solar) and muonic neutrinos (atmospheric). The standard notation is that the three neutrino masses m_i are ordered such that m_3 is the most splitted state and $m_2 > m_1$. With this choice Δm^2_{23} and θ_{23} are the “atmospheric parameters” and Δm^2_{12} and θ_{12} are the “solar parameters”. Whatever the spectrum of neutrino masses (“normal hierarchy” so that $\Delta m^2_{23} > 0$ or “inverted hierarchy” so that $\Delta m^2_{23} < 0$, see figure 1.1), the two neutrino mixing is a valid description of oscillation experiments. This is a consequence of the experimental evidence for a hierarchical ordering of the mass splittings:

$$\left|\Delta m^2_{12}\right| \ll \left|\Delta m^2_{13}\right| \simeq \left|\Delta m^2_{23}\right| \tag{1.12}$$

Figure 1.1 shows the possible scenarios for the neutrino mixings and mass spectrum, while table 1.1 summarizes the oscillation interpretation of the two established neutrino anomalies (a further and detailed discussion on this results can be found in [4].

As already stated, oscillation experiments gave many informations concerning many properties and aspects of neutrino physics but are blind to their Dirac or Majorana nature or to their absolute mass scale.

1.4 Dirac or Majorana

While an experimental set-up whose purpose is the discrimination between electrons and positrons would be trivial, thanks to their coupling with an electromagnetic field, an experiment designed to distinguish a neutral lepton by its own anti-particle would be a non obvious challenge.

Suppose that, in a generic frame of reference a left-handed neutrino ν_L is moving with negative elicity; since it is proved that neutrinos do have a mass, there will be an appropriate Lorentz boost that will identify another frame of reference in which the neutrino will have an opposite helicity: as it will be moving in the opposite direction,with a simple change of reference frame, we have transformed a left-handed neutrino ν_L into a right-handed ν_R. If we assume the validity of CPT theorem, there must exist the CPT-coniugate of this particle: an anti-neutrino with positive helicity $\bar{\nu}_R$.

At this point, we could ask ourselves if there is a quantum number that would allow us to distiguish between ν_R and $\bar{\nu}_R$. If this number exists, there would be four different eigenstates with the same mass: ν_L, ν_R, $\bar{\nu}_R$ and $\bar{\nu}_L$. Their physics would be described by a Dirac field, exactly as for the charged leptons. If no such quantum number exists, the neutrino eigenstates would be only two, distiguished by their helicity and it would become impossible to distinguish between $\bar{\nu_R}$ and ν_R. In this case, where no distinction bewtween particle and antiparticle is possible, the neutrino will be described by a Majorana field. These fields can be simply obtained by means of simple linear combinations of Dirac fields, for example:

$$\nu = \frac{1}{\sqrt{2}}(\psi_L + (\psi_L)^c), \text{ or } N = \frac{1}{\sqrt{2}}(\psi_R + (\psi_R)^c) \tag{1.13}$$

where ψ_R and ψ_L are the generic projection of a Dirac spinor on its two chirality components, defined as $\psi_{L,R} = [(1 \mp \gamma_5)/2]\psi$, while ψ^c is the charge-conjugate of ψ.
It is straighforward to see that these fields are invariant under charge conjugation and they are constitued only by left-handed and right-handed spinors respectively.
To summarize: in the Dirac situation neutrinos are described by a single field, used to describe its four degenerate mass eigenstates (left-handed, right-handed, particle, antiparticle); in the Majorana case, light neutrinos are defined by a single field, with only two components: left-handed or right-handed.
The Majorana picture would be ruled out if the quantum number to distinguish between ν_R and $\bar{\nu}_R$ exists, but in the Standard Model such a number is included: it's the lepton number. Nevertheless, there is nothing sacred about lepton number whose conservation is not a requirement arising from a higher principle, it is rather an experimental fact. If an experiment proves that its conservation can be violated, there is thus no reason why ν_L or $\bar{\nu}_R$ cannot be the same counterpart of the same object.

1.5 The see-saw mechanism

Another open puzzle of the Standard Model is the huge different mass scale between neutrinos and the corresponding charged leptons.
In order to build up a theoretical context that would include this feature we must extend the gauge mechanism to a more general symmetry group.
The main insight is the extention of the symmetry group of the MSM defined in equation (1.8) to a more general one is:

$$SU(3) \times SU(2)_L \times SU(2)_R \times U(1). \tag{1.14}$$

In this equation, $SU(2)_R$ could be seen as the symmetry group connected to a weak interaction similar to the standard one but mediated by a right-handed boson W_R. The presence of this field would link the right-handed components of fermion fields by means of new right-handed bosons. In order to explain an idea of this mechanism we will consider the simplest case of one type of neutrino. Let us assume the standard Higgs mechanism with one Higgs doublet, which is the mechanism of the generation of the masses of quarks and leptons. If we use the Majorana fields defined in the previous section, we can construct a Dirac mass term as follows:

$$\mathcal{L}^D = -M_D(\bar{\nu}N + \bar{N}\nu) \tag{1.15}$$

It is natural to expect that the mass M_D is of the same order of magnitude as masses of the corresponding lepton or quark. We know, however, from experimental data that neutrino masses are much smaller than the masses of leptons and quarks. In order to "suppress" neutrino mass let us assume that there exist also the lepton number violating Majorana mass terms:

$$-\mathcal{L}^M = M_L \bar{\nu}\nu + M_R \bar{N} N \qquad (1.16)$$

the former (M_L) is experimentally bound to be basically zero, the latter (M_R) has to be of energy scales of the order of the TeV. We are in fact assuming that the mechanism that gives rise to M_R is the result of a spontaneous symmetry breaking of the group $SU(2)_R$, which must have happened at energies greater than the scales of the Great Unification Theory. Combining equations (1.15) and (1.16) in a more compact form as:

$$-L_M = \begin{pmatrix} \overline{\nu} & \overline{N} \end{pmatrix} \mathcal{M} \begin{pmatrix} \nu \\ N \end{pmatrix} \quad \text{where} \quad \mathcal{M} \equiv \begin{pmatrix} M_L & M_D \\ M_D & M_R \end{pmatrix} \qquad (1.17)$$

it will be possible to diagonalized the mass matrix $\mathcal{M}$ obtaining two eigenvalues:

$$M_\nu \simeq \frac{M_D^2}{M_R}\, , M_N \simeq M_R \qquad (1.18)$$

where $\simeq$ is used by virtue of the experimental approximation of $M_L = 0$ e $M_R >> M_D$.
As mass eigenstates one of the order of meV, the other of the order of M_{GUT}. By including this more general symmetry group, an explanation of the mass difference between light neutrinos and charged fermions. The price for this is the appearance of a new particle, whose mass M_N is so heavy that only LHC could be a fair challenge for its discovery.

1.6 Neutrinoless Double Beta Decay

Double Beta Decay (DBD) is a rare spontaneous nuclear transition, whose existence was first proposed by Maria Goeppert–Mayer in 1935 [20]. In this transition, a nucleus (A,Z) changes the nuclear charge of two units maintaining the same mass number, becoming therefore a (A,Z±2) nucleus. Suppose that an (A,Z+2) nuclide appears in the final state. Normally DBD is not favored in respect to single β-decay: the transition is observable only when the nuclide has single β-decay energetically forbidden (which occurs when the intermediate (A,Z+1) nucleus has a binding energy greater than the (A,Z) and (A,Z+2)

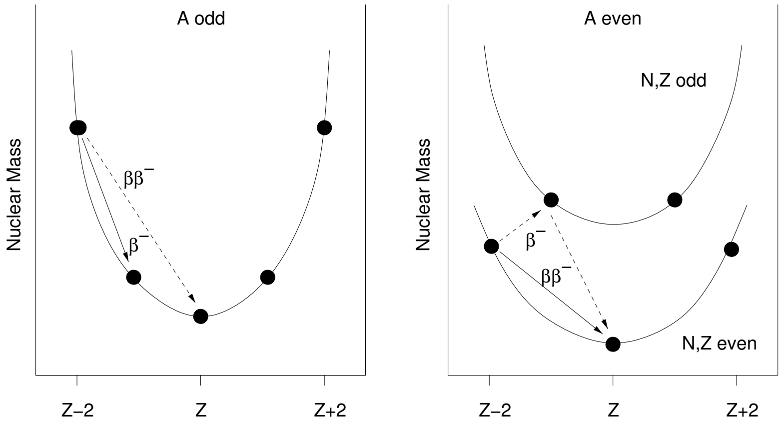


Figure 1.2: Nuclear mass as function of the proton number.

nuclei).

This situation can be understood by looking at the Weizsäker expression for the atomic mass as a function of the mass number A and the number of neutrons N and protons Z. In particular, the Weizsäker formula contains the "pairing" term that takes into account the increase in binding energy of the nucleus when protons or neutrons are coupled to give an angular momentum equal to zero:

$$\delta = \begin{cases} +12/A^{1/2} & \text{for even A and odd Z,N} \\ -12/A^{1/2} & \text{for even A and even Z,N} \\ 04 & \text{for odd A} \end{cases} \tag{1.19}$$

If isobaric nuclei are considered and their atomic masses are plotted as a function of Z it is easy to find that, for odd A, the nuclei are positioned as described on the left side of Fig. 1.2. If even A nuclei are considered, it is found that the nuclear masses are disposed as shown in the right part of Fig. 1.2.

Two different DBD modes are usually considered. First, the decay with two neutrinos (2ν-DBD). Lepton number is conserved by this decay, which is allowed by the SM:

$$2\nu\text{−DBD}: \quad (A, Z) \rightarrow (A, Z+2) + 2e^- + 2\bar{\nu}_e \tag{1.20}$$

Second, the decay without emission of neutrinos (0ν-DBD), given by

$$0\nu\text{−DBD}: \quad (A, Z) \rightarrow (A, Z+2) + 2e^- \tag{1.21}$$

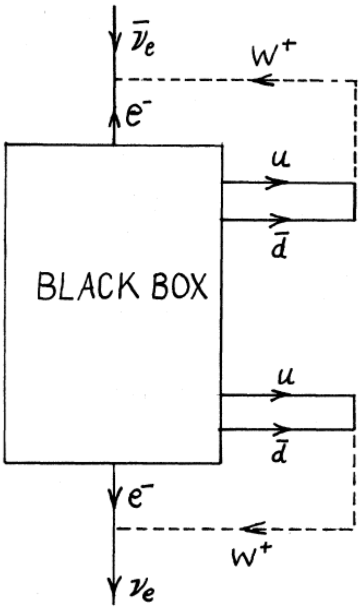


Figure 1.3: The "Black box" of 0ν-DBD: regardless of the physical mechanism that dominates the decay, this diagram shows that a modification in the neutrino propagator results in a Majorana mass term for neutrinos [21].

This decay mode is forbidden in the Standard Model, as it violates the lepton number by two units.

A brief look at the reaction above makes it evident that lepton number conservation is violated. Neutrinoless double beta decay can proceed through many different mechanisms: almost any physics that violates the total lepton number can generate it [22]. We will focus on the simplest channel, where only two electrons are produced in the final state.
An experimental confirmation of this decay mode will thus constitute an important step in the study of elementary particle physics beyond the SM. The Feynman diagrams for both decay modes are shown in Fig. 1.4. While the 2ν-DBD diagram contains only SM interactions, the 0ν-DBD requires the known $V - A$ interactions in addition to a massive Majorana neutrino. The virtual neutrino in the diagram can be thought of as produced as an anti-neutrino at one vertex and absorbed as a neutrino, which is equal to the anti-particle thanks to the Majorana nature, at the other vertex; moreover, a non-zero neutrino mass is required to flip the helicity, since the neutrino emitted with the electron at vertex 1 is right-handed and the one absorbed at the other vertex is left-handed. Should 0ν-DBD occur, its rate would be much lower than the one of 2ν-DBD because of the helicity flip and the smallness of the neutrino mass.

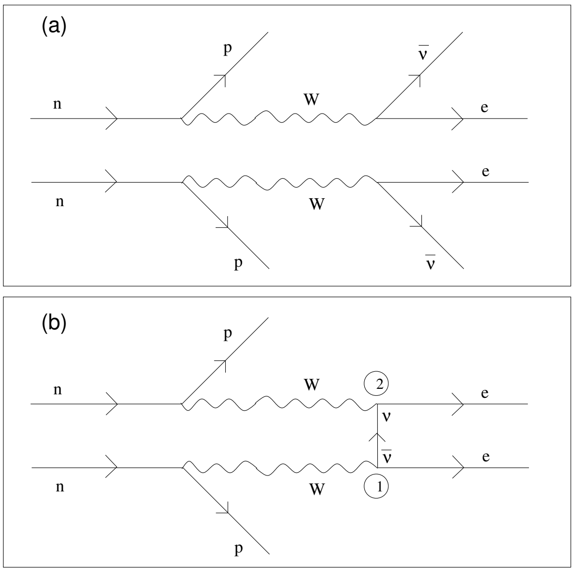


Figure 1.4: Feynmann diagrams for the two Double Beta Decay channels: (a) 2ν. (b) 0ν: an anti-neutrino is produced at vertex 1 and a neutrino is absorbed at vertex 2; this process is allowed only for Majorana neutrinos.

The discrimination between these decay modes is, in principle, very simple and is based on the shape of the spectrum obtained by summing on the energies of the two emitted electrons. In fact, this spectrum is determined by the phase space of the other emitted particles. As shown in Fig. 1.5, 2ν-DBD is a four body decay, with a continuous spectrum featuring a maximum value around one third of the Q-value. On the contrary, the two electrons retain all the available kinetic energy in 0ν-DBD (neglecting nuclear recoil). For this reason, the spectrum is just a spike at the transition energy.

In both cases, DBD is a semi-leptonic second-order weak interaction, and thus is characterized by a very long lifetime: for example, experimental half-lives for this decay mode are $T^{2\nu}_{1/2} \sim 10^{18} - 10^{22}$ years for the 2ν-channel. This remarks that the experimental observation of this decay turns out to be a great challenge: very rare events have to be detected and disentangled from the traces of other radioisotopes, which have similar transition energies but decay times even 10 orders of magnitude shorter. Presently, 2ν-DBD has been observed for ~ 10 nuclides, while there exists only one positive result for 0ν-DBD [18].
The probability for 0ν-DBD to occur is usually expressed using the general relation derived from Fermi's golden rule:

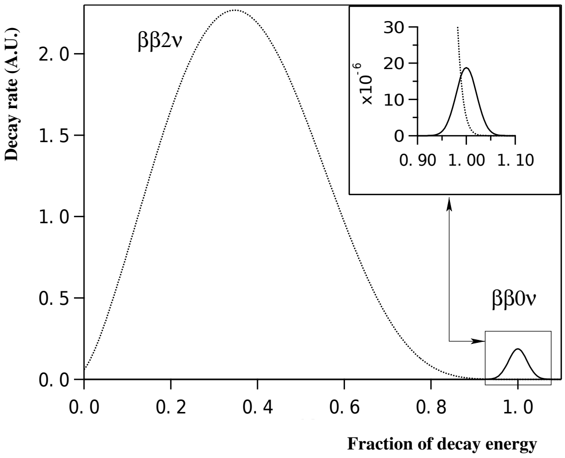


Figure 1.5: Schematic representation of the spectra obtained by summing on the two electrons energies for 2ν-DBD (dashed line) and 0ν-DBD. In the inset, the relative intensity of the 2ν-channel is overestimated in order to underline its contribution to the 0ν background. Spectra are obtained with the convolution of a 5% energy resolution, common to many experiments.

$$\left[T^{0\nu}_{1/2}\right]^{-1} = G^{0\nu}|M^{0\nu}|^2\langle {m_{\beta\beta}}^2\rangle \tag{1.22}$$

In the previous equation, $G_{0\nu}$ is the phase space integral, approximately proportional to $\left[Q^{0\nu}_{\beta\beta}\right]^5$ and determined exactly, $|M^{0\nu}|^2$ is the decay matrix element and $\langle m_{\beta\beta}\rangle$ the effective Majorana mass, defined by

$$\langle m_{\beta\beta}\rangle \equiv \sum_i \phi_i m_i |U_{ei}|^2 \qquad (i = 1, 2, 3) \tag{1.23}$$

The ϕ_i phases that appear in the last equation are the CP intrinsic neutrino parities, which are connected to the Majorana phases. Their presence implies that cancellations are possible, and if the neutrino is a Dirac particle the cancellation is total (it is equivalent to a couple of degenerate Majorana neutrinos with opposite phases), so that the decay doesn't take place. It is true, however, that mechanisms other than the exchange of a light Majorana neutrino cannot be excluded for 0ν-DBD. However, it has been shown [21] that if 0ν-DBD occurs, then neutrinos are Majorana particles even if the decay is dominated by a different process: the dominant mechanism that drives 0ν-DBD is inserted into the *black-box* shown in Fig. 1.3, and by connecting the external lines appropriately, a contribution to the neutrino propagator that turns ν into $\overline{\nu}$ induces always a Majorana mass term. The conclusion is that the experiments looking for 0ν-DBD provide a model-independent determination of the neutrino na-

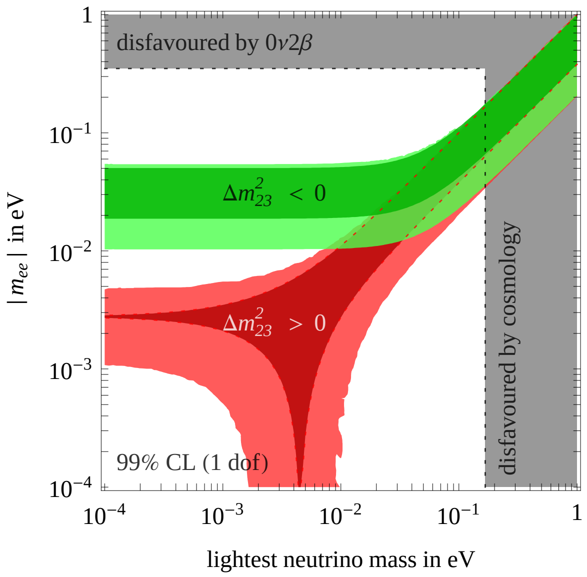


Figure 1.6: The Majorana effective mass in function of the lightest ν mass, in a parameter space constrained by previous experimental achievements. The green region corresponds to the inverted hierarchy and the red one to the normal hierarchy; the darker bands are the regions allowed if oscillation parameters were to be known with infinite precision [4].

ture, no matter which is the dominant mechanism driving the decay.

The importance of 0ν-DBD in the field on neutrino physics does not end here: Eq. ((1.23)) points out also its importance in mass hierarchy discovery. In fact, even if $\langle m_{\beta\beta}\rangle$ has a dependence on ϕ_i, its upper and lower limits depend on the absolute values of the mixing matrix elements. There is a relationship between the effective Majorana mass and the lightest neutrino mass: it depends on the valid hierarchy, according to which the lightest mass eigenstates is determined. The $\langle m_{\beta\beta}\rangle$ vs. m_{light} relationship can be determined easily for both the normal and the inverted hierarchy. The result in the case of normal hierarchy, where $m_{\text{light}} = m_1$, is reported as an example:

$$m_{\beta\beta} = |cos^2\theta_{12}cos^2\theta_{13}e^{i\alpha 1}m_1 \tag{1.24}$$

$$+sin^2\theta_{12}cos^2\theta_{13}e^{i\alpha 2}\sqrt{\Delta m^2_{12} + m^2_1} + sin^2\theta_{13}e^{-2i\delta}\sqrt{|\Delta m^2_{13}| + m^2_1}|$$

Incorporating the values of mixing angles and squared-mass differences measured by oscillation experiments gives m_ν in function of m_1 for the normal hierarchy and m_3 for the inverted hierarchy, for a given set of phases. Fig. 1.6

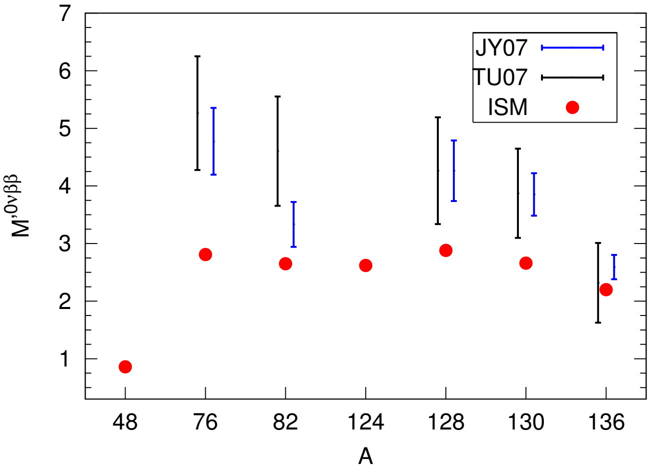


Figure 1.7: NME values for different nuclei. NME (red dots) and two different QRPA calculations are compared (black and blue bars). The spread is due indicates the limits of the values obtained by varying the input parameters [10]

shows the plot of m_ν in the two cases, allowing phases to vary from 0 to 2π. Thus, an observation of 0ν-DBD, together with data provided by oscillation experiments, would define a range for the absolute neutrino mass and disentangle the hierarchy scheme of the neutrino mass eigenvalues. Experiments on 0ν-DBD search give an upper limit to the effective Majorana mass and rule out regions of the plot 1.6. In particular, a sensitivity of $\sim 10 - 50$ meV on $\langle m_{\beta\beta} \rangle$ could definitely exclude the inverse and quasi–degenerate hierarchy.

1.6.1 Nuclear matrix elements

Eq. ((1.22)) shows clearly that the evaluation of $\langle m_{\beta\beta} \rangle$ from an experimental measurement of $T^{0\nu}_{1/2}$ requires the exact knowledge of the Nuclear Matrix Elements $\langle M^{0\nu} \rangle$. From a theoretical point of view, this is the main limit when describing 0ν-DBD. It has been discussed in the previous section that without making any assumption on the mechanism that generates neutrinoless double beta decay, the existence of this process implies that neutrinos are Majorana particles. However, in order to extract informations on the neutrino mass hierarchy and the absolute mass scale, some assumptions must be made on the mechanism that is responsible for this process. The simplest possibility is to assume that the 0ν-DBD occurs by the exchange of a light Majorana neutrino. It is clear that, even under this assumption, the phase space factor $G_{0\nu}$ and the nuclear matrix element (NME) $M_{0\nu}$ must be known in order to extract the value of $m_{\beta\beta}$. Moreover, a strong evidence of neutrinoless double beta decay can be obtained only if a positive signal is seen in several isotopes. This condition is necessary to rule out the possibility that the observed signal is produced by some other unknown rare process able to mimic the 0ν-DBD experimental signature. However, since different isotopes have different values for $G_{0\nu}$ and $M_{0\nu}$, these quantities must be known in order to compare experiments. While the phase space factor can be evaluated exactly, NME represents the biggest source of theoretical uncertainty in the evaluation of $m_{\beta\beta}$. Nuclear matrix elements depend on the structure of the parent and daughter nuclei, as well as the intermediate one. Since a many bodies problem must be solved, the calculation cannot be carried out analytically, but requires numerical computations in which several approximations are introduced. There are different approaches for the evaluation of nuclear matrix elements, two of the most important one are the quasi-particle random phase approximation (QRPA) [23] and the nuclear shell model (NSM) [24]: since their correcteness and accuracy are strictly related to the nuclear transition we are dealing with, there is still no simple verification, thus one possibility is to consider the spread of the theoretical values as a measure of their uncertainties: figure 1.7 gives an idea of such spreads.

Chapter 2

Bolometric Detectors for Rare Events Searches

2.1 Introduction

The passage of an elementary particle, an ion or a photon through matter gives rise to ionization and excitation of nuclei and atoms of the traversed material, resulting in a transfer of the particle kinetic energy to the traversed material.
Generally ionization represents a minority fraction of the total energy lost by the particle, the majority fraction is indeed converted into *phonons*, the quanta of elementary vibrational excitations of the atoms. In a material at thermal equilibrium phonons account for the thermal motion of the atoms, their mean energy being therefore proportional to the temperature.
Conventional techniques for the measurement of the energy deposition by an elementary particle (or an ion or a photon) are based on the direct or indirect detection of ionization [25]. More recent is the developement of devices based on phonon detection, called thermal detectors or bolometers [26]. In these devices the phonon signal is detected long after ion pair recombination, when - if radiative energy losses are negligible - also the ionization energy have been converted into phonons. It is quite common to read-out the phonon signal when the initial non equilibrium energy distribution of phonons has reached thermal equilibrium. In this approximation, where all the incident energy is thermalized, the detector provides a true calorimetric measurement of the energy deposited by the particle and the macroscopic variable measured by the phonon sensor is the temperature (this the reason for calling them thermal detectors).

2.2 Calorimeters

A calorimeter is a detector measuring, by means of a dedicated sensor, the temperature rise of the material in which a particle interacted, releasing a franction ΔE of its energy. If $C(T)$ is the thermal capacity of the material and T its initial temperature, the energy convertion ΔE produces a temperature rise ΔT. If $C(T) \simeq C(T + \Delta T)$ (which is a reasonable approximation for energy depositions of the order of 1 MeV are deposed in tens of milligrams of materials at low temperature) the the temperature variation can measure the energy released in the detector via:

$$\Delta T = \frac{\Delta E}{C(T)} \tag{2.1}$$

The intrinsic resolution of the measurement is limited only by the thermodynamic fluctuations of the internal energy U of the absorber. The squared mean value σ of these fluctuations is given by:

$$\sigma = \sqrt{k \cdot T \cdot C(T)^2} \tag{2.2}$$

where k is the Boltmann constant.
At room temperature these kind of detectors wouldn't have the sensitivity to be used as a particle detector. This is the reason why these instruments are generally used at temperatures of the order of mK where they can become even more competitive than the state of the art solid state detectors.

2.3 Thermal signal

A thermal detector consists of three main components:

- the particle absorber, which is the mass where the particles interact, depositing a fraction of their energy. Through the detection of its temperature rise it's possible to measure the amount of energy deposited. The absorber material can be chosen quite freely, the only requirements being, in fact, reasonable thermal and mechanical properties. The absorber mass can range from few micrograms to few kilograms;

- the temperature sensor which is used to measure the temperature rise induced by the particle interaction in the absorber. Different kind of thermometers exists, depending on the material and on the process through which the temperature signal is converted into an electric signal. The sensor is connected to the absorber by a thermal conductance G_{AS};

- the thermal link which provides the restoring of the equilibrium temperature of the absorber after the heat pulse. It is the thermal conductance G_{ab} between absorber and the heat sink. Together with the system heat capacity it influences the time response of the detector.

The heat sink is a low temperature refrigerator able to mantain the detectors at temperatures ranging from about 0.1 K to 10 mK. By working at this temperatures two important conditions are satisfied: the heat capacity of the absorber is very low and the statistical fluctuation of temperature are small. Figure 2.1 shows this simplified model.

The deposed energy is transformed in heat and it is transmitted thorugh G_{AS} to the sensor whose temperature rises by consequence. The conductance G_{AB} guarantees that the system returns to thermal equilibrium with the heat sink after the particle induced a temperature signal.

In order to collect all the information, the transmission time t_{AB} of the heat to the heat sink must be longer than t_{AS}, the time need to reach equilibrium between absorber and sensor.

Using some math, if we keep both the heat capacities of absorber and sensor constant and being $P(t)$ the thermal power dissipated on the sensor, the differential equation that describes that heat flow from the absorber to the sensor are:

$$P(t) = G_{AB}(T_A - T_0) + G_{AS}(T_A - T_S) + C_A \frac{\mathrm{d}T_A}{\mathrm{d}t} \tag{2.3}$$

$$G_{AS}(T_S - T_S) = C_s \frac{\mathrm{d}T_S}{dt} \tag{2.4}$$

where C_A and C_S are the thermal capacities of absorber and sensor respectively, while T_A,T_S and T_0 are the temperatures of the absorber of the sensor and of the heat bath.

These equations are easily solvable by means of Fourier Transform, which leads to the solution for the time developement of the temerature $T(t)$:

$$T(t) = T_0 + \frac{\Delta E}{(\alpha - \beta)} \left(\mathrm{e}^{-\frac{t}{\alpha}} - \mathrm{e}^{-\frac{t}{\beta}} \right) \tag{2.5}$$

where α and β are liked by the relation:

$$\begin{cases} \alpha + \beta & = \tau_{AS} + \tau_{AB} + \frac{C_{AS}}{G_{AB}} \\ \alpha \times \beta & = \tau_{AS} \times \tau_{AB} \end{cases} \tag{2.6}$$

where $\tau_{AS} = \frac{C_{AS}}{G_{AS}}$ and $\tau_{AB} = \frac{C_{AB}}{G_{AB}}$: the expected temperature signal rises in a finite time and returns to T_0 with an exponential behaviour.

This thermal model can, in some cases, be even more simplified if that absorber

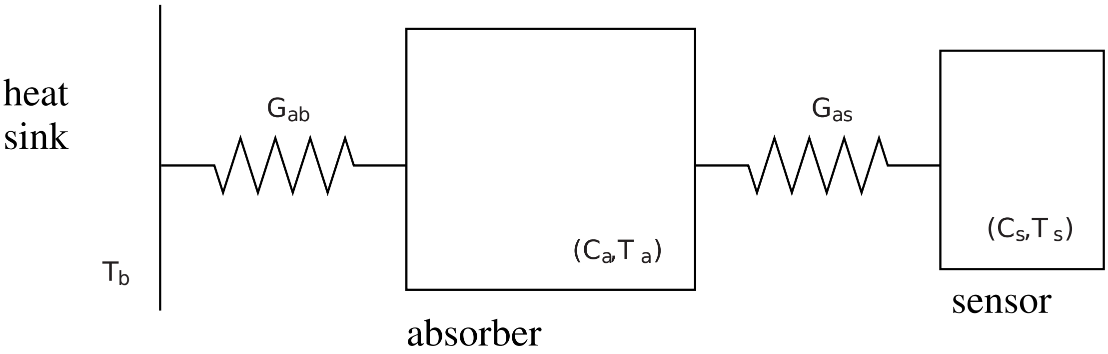


Figure 2.1: Schematics of a simple thermal detector.

and sensor are considered as the same object (we will refer to this approximation as *monolithic bolometer*). In this case there's only one thermal conductance and capacity (G and C) and the signal has an instant rise time and an exponential decay:

$$\Delta T(t) = \frac{\Delta E}{C(T)} e^{-\frac{t}{\tau}} \tag{2.7}$$

where $\tau = \frac{C}{G}$.

2.4 Phonon physics

In order to understand the model described in the previous section, it can be useful to discuss the process that leads to developement of a thermal signal after a particle interaction.

A ionizing particle loses energy primarly by means of scattering with the electrons of the material it's crossing. The energy of this interaction is converted in the production of electron-hole couples and in phonons. As we stated ad the beginning of the chapter, the phonon is the quantum of energy associated to the different vibrational frequency mode of the crystal.

There are two kind of phonons that can be produced: *optical* and *acoustical*, depending whether their vibrational pattern is longitudinal or transversal.

After the electron-hole recombination, mainly optical phonons are produced. In times of the order of tens of ns their energy is converted into acoustic phonons. At this point, phonons can follow different kind of propagation modes in the crystal: when the optical phonons decay, phonons are still confined in the region of the primary interaction; this region expands due to the phonon propagation up to dimentions of their propagation length; this propagation is called "balistic", since the phonons basically evolve like free particles; by means of scattering with the impurities in the crystal the propagation becomes diffusive until the thermal equilibrium is reached. At this point, they are distributed in frequency with a Plank spectrum whose temperature is shifted with respect the equilibrium one (proportionally to the energy of the incident particle).

2.4.1 Phonon sensors

In principle two kinds of phonon can be detected: *thermal* or *non-thermal* [27].
Non-thermal phonons propagates with non-equilibrium distribution in the crystal; being subjected by few scatterings, they mantaining a "memory" of the interaction site: using several fast sensors it is possible in principle to determine the particle interaction point using the phonon signal relative time. This phonons are difficult to be detected.
Thermal phonons instead are uniformely distributed in the crystal and their signal can be measured as a thermal pulse described in section 2.3.
In order to achieve good signal-to-noise ratio, a phonon detector neeeds a high sensitivity phonon sensor. The phonon sensor is a device that collects the phonons produced in the absorber and generates an electrical signal, usually proportional to the energy contained in the collected phonons. A simple realization of this device can be accomplished through the use of a thermistor, whose resistance, as a function of the temperature, has a steep slope. In practical devices, there are two main classes of thermistors which give the best results: semiconductor thermistors (STs) and transition edge sensors (TESs).
Thermistors are usually characterized by their "logarithmic sensitivity" A, defined as:

$$A = \left| \frac{d \log R(T)}{d \log T} \right| \tag{2.8}$$

The value of the sensitivity is usually in the range 1-10 for STs and in the range 102-103 for TESs. In the rest of this work, we will focus on semiconductor thermistors.

2.4.2 Semiconductor thermistors

STs consist normally of Ge or Si small crystals with a doped region. The doping process must be chosen to provide a uniform concentration of dopants in the thermistor volume. Neutron Transmutation Doping (NTD) is usually the best choice for Ge. Standard planar Si technology (ion implantation) works also well.
The electrical conductance of semiconductors is due to different phaenomena,

depending on the working temperature and on the relative concentrations of donors and acceptors of the material.
At high temperatures the conductance is $intrinsic$: it takes place through thermal eccitation of the electrons from the valence band to the conduction band. The intrinsic conductance decreases with the temperature: below a specific thermal bound, the phonon energy is not sufficient for that kind of eccitation.
In doped semiconductors a second mechanism guarantees electrical conductance at low temperatures: this is due to the thermal activations of the charge carriers of the impurities. The activation energy of impurities is inferior to the energy gap and has a weak dependence from the temperature: for doped Si or Ge, the activation energy varies from a few meV to 100 meV.
Doped semiconductors offer a high $extrinsic$ conduction above nearly 1K. At lower temperatures the activation energy of charge carriers is higher than the thermal energy thus they become insulators.
Nevertheless, it is possible to realize semiconductors with a sufficiently high conductance at low temperatures, by playing with the doping concentrations.
For example, in an n-doped semiconductor, by increasing the donors (decreasing the average distance between impurities) the electronic levels of donors start to interfere due to the superposition of their wave functions. For sufficiently high concentration, this gives rise to a dual-band structure inside the forbidden gap: the Hubbard band [28].
For donor concentrations below this transition, the electrons can tunnel from an impurity site to another: this effect is known as "hopping" conduction. Figure 2.2 shows a nice view on this process.

Electrons with energies below the Fermi level can be thermally excited (by the absorption of a thermal phonon) above the coulombian gap, moving from a donor site to another. In case $k_B T$ is of the same order of the gap, the tunneling happens between lower energy impurity sites, independently from their spacial distance: this is called Variable Range Hopponing (VRH). In this conduction regime the resistivity of the sensor has a steep dependence on the temperature:

$$\rho(T) = \rho_0 \exp\left(\frac{T}{T_0}\right)^{\gamma} \tag{2.9}$$

where γ, ρ_0 and T_0 depend on the doping level.
A flux of thermal phonons produces then a change in the conductance and this variation can be transmuted into an electrical signal if the thermistor is polarized.
The processes described previously do not account entirely for the behaviour of a real sensor. In particular, the introduction of a phenomenological model

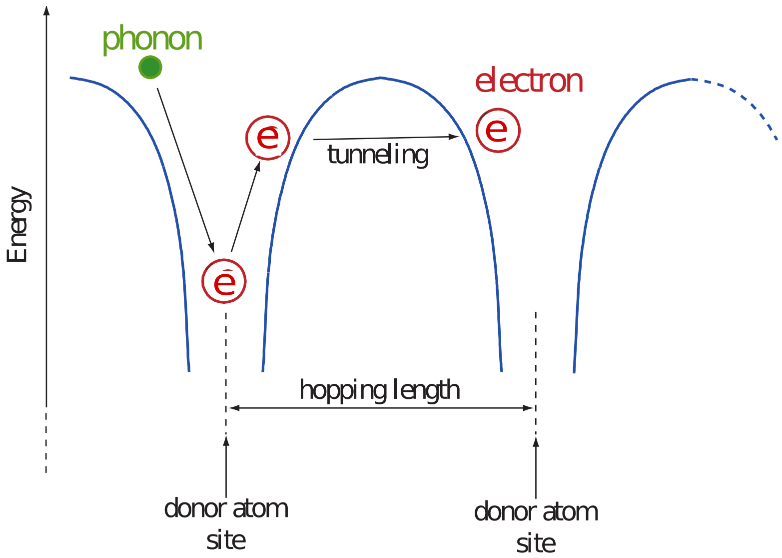


Figure 2.2: Graphical representation of the "hopping" conduction mechanism.

known as Hot Electron Model (HEM) [29] helps to understand the non-linearities of VRH. The basic idea of HEM is that, at low temperatures, the thermal coupling between electrons and the lattice in the thermistor is weakened. This implies the introduction of an electronic temperature T_e and of a phononic temperature T_{ph} : when a power P is injected into the system, T_e results larger than T_{ph}. The HEM answers the question about which temperature should be considered in equation (2.9). For example, the temperature of metals is defined through their conduction electrons; in semiconductor thermistors, where conduction electrons and phonons of the lattice form two separate systems connected by a finite thermal conductance, th resistance will depend on T_e. In this case the resistance of the thermistor is determined by equation (2.9) and the thermistor can be used as thermometer.

2.5 Bolometers

The general theory describing bolometers has been developed by Mather in 1982 [30] for what concerns the monolithic case, in which there is no distinction between sensor and absorber.

Bolometers use the variation of a resistance to measure temperature variation of its absorber is a *bolometer*, so the main difference from the generic thermal detector is that the bolometer temperature T_B, in absence of a signal, is different

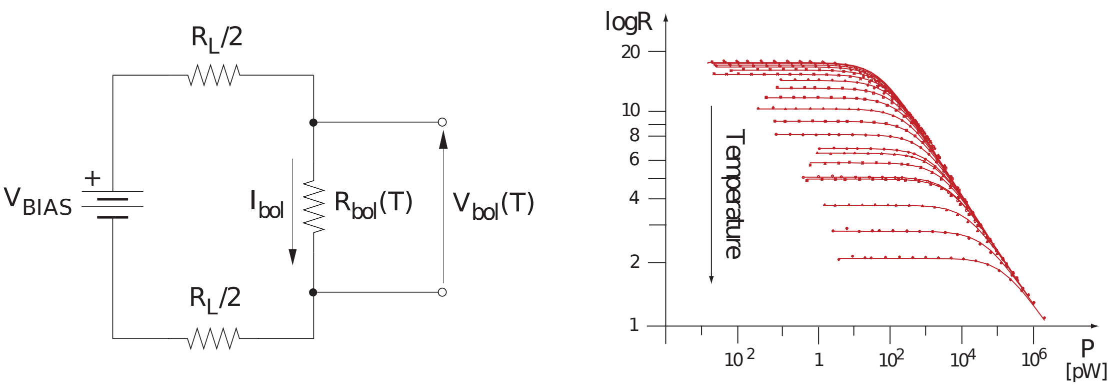


Figure 2.3: The left panel shows the electric scheme of a simple bolometer circuit. The right panel shows the dependence of the resistance on the power dissipation for various values of the base temperature.

from the heat sink T_0: this is due to two sources of power dissipated on the bolometer itself: the electric and background power.
While the background power P_b is the superposition of all other dissipative components, the electric power P_e is simply the heat flux due to Joule effect since the bolometer is polarized by a small current i_{bias} and is given by:

$$P_e = V_{\text{bias}} \times i_{\text{bias}} \tag{2.10}$$

where V_{bias} is the voltage applied to the bolometer.
At thermal equilibrium we have:

$$P_e + P_B = G_{eff}(T_1 - T_0) \tag{2.11}$$

where G_{eff} is the effective conductance of the sensor to the heat sink. A schematic representation of a bolometric detector is shown on the left panel of figure 2.3, while on the right panel the dependence of the resistance of the bolometer on the power dissipation for diffeent temperatures.
In the model description given in the previous paragraph we neglected this dynamic component. However, for bolometers this is an non-negligible source of non linearities of the detector's response. This non-linearity is called "electrothermal feedback". and can be appreciated by looking at the $I-V$ dependence (the *load curve*) for a TS shown on the left panel of figure 2.4. For small values of the bias current the temperature rise produced by power dissipation can be neglected and the IV curve is almost linear. For bigger values of I_B the slope of the $I-V$ curve starts to increase, until an inversion point is reached, where a further increase of the bias current causes a decrease of the sensor voltage.

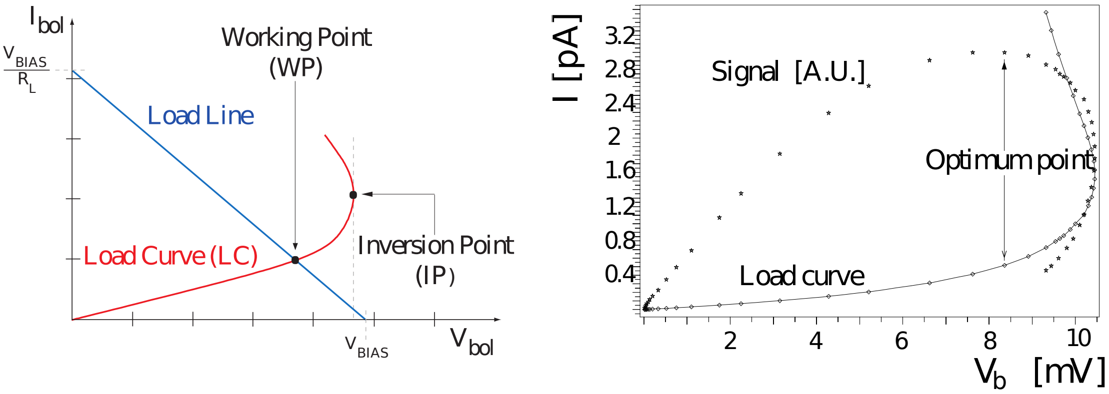


Figure 2.4: Load curves for semiconductor thermistors. On the left picture the working point is determined by intersection of the sensor characteristic curve with the bias circuit load line. On the right the load curve is shown together with the corresponding signal amplitude.

The right panel shows instead the relation between the resistance of the thermistor and the dissipated power for different temperatures.
In static conditions, the thermistors electric and thermal parameters are described by a point on the load curve. The intersection of the straight line of equation $V = V_b I \cdot R_L$ and the load curve $I = I(V)$ determines the working point of the sensor. Usually an optimal working point is chosen so that the signal amplitude or the signal-to-noise ratio is maximum.

In this simple modelization in fact, it is straightforward to obtain the relationship between the maximum voltage signal (ΔV), the thermistor parameters and the deposited energy E:

$$\Delta V = \frac{R_L}{R_L + R_{bol}} \cdot V_{bol} \cdot A \cdot \frac{\Delta T_{bol}}{T_{bol}} \simeq \frac{E}{C T_{bol}} \cdot \left\{ A \cdot \sqrt{P \cdot R_{bol}} \right\} \tag{2.12}$$

As expected, the highest the thermistor sensitivity, the highest the signal amplitude. Moreover, since ΔV goes to zero if $P \rightarrow 0$ but also if $P \rightarrow \infty$ there must be a point where the signal reaches a maximum. This point, called "optimum", must be found experimentally. The easiest way to accomplish this procedure is by means of a fixed resistance (heater) directly glued onto the crystal through which it is possible to dissipate a known joule power into the crystal in a manner equivalent to a particle energy deposit. On the right panel of figure 2.4 the trend of the heater energy as a function of the bias voltage is shown (superimposed to the load curve): the optimum point is the value of V_b that maximasises the heater signal.

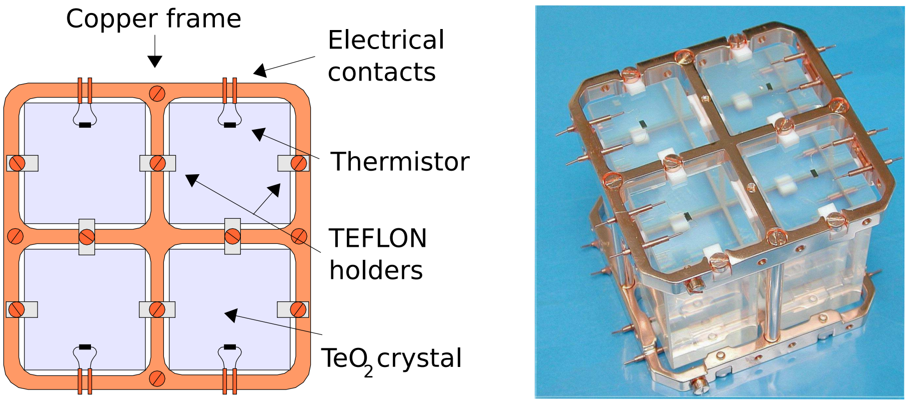


Figure 2.5: Left panel: schematics of a TeO_2 single module detector. Right panel: picture of a single module used for the CUORICINO experiment.

2.6 TeO_2 single module

In chapter 3 we will describe the CUORICINO experiment, which operates with bolometers having TeO_2 crystals as absorber and a Ge thermistor as thermal sensor.
In this section we describe the basic structure of CUORICINO, or *single module*. This is a crystal, $3 \times 3 \times 6$ or $5 \times 5 \times 5$ cm^3 in size, provided with an NTD Ge thermistor acting as phonon sensor and operated as a completely independent detector. The germanium thermistors are glued to the TeO_2 crystal by 9 spots of Araldit rapid epoxy. Their bias circuit scheme is the same as the one shown in figure 2.3. Silicon chips with a heavily doped meander structure and a resistance between 50 to 100 kΩ are also glued on to each crystal. They are used as Joule heaters to inject a uniform energy in the crystal in order to monitor the thermal gain and correct its variations off-line to stabilize the response of the bolometers. This topic will be discussed in more details later on. The signal is read by means of two 50 mm diameter gold wires, ball bonded to metalized surfaces on the thermistor and thermally coupled to a copper frame. This copper structure constitutes the heat sink, being in thermal contact with the coldest point of the dilution refrigerator used to reach the operating temperature. PTFE (Teflon) is used between the copper frame and the crystals: it has low heat conductance and low heat leak preventing any signal loss as described in section 2.3. It compensates also for the differences between coefficients of thermal expansion of copper and of TeO_2. Figure 2.5 shows a schematic of a TeO_2 single module, holding four TeO_2 crystals.

Now that the basic principles of the energy absorber and of the temperature

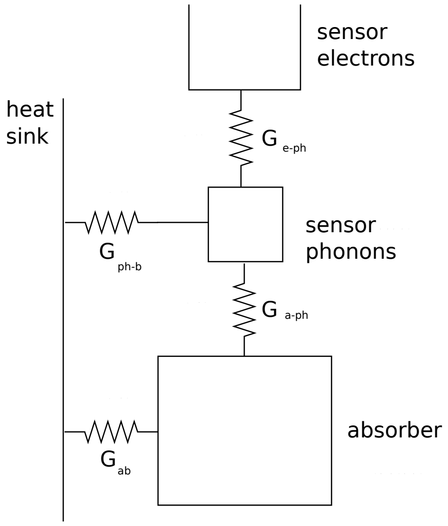


Figure 2.6: Left panel: schematics of a TeO$_2$ single module detector. Right panel: picture of a single module used for the CUORICINO experiment.

sensor are known, a more refined bolometric model can be presented. The single module can be modelized as in figure 2.6. It is composed by the following parts:

- the heat bath, which is the thermal reference point;
- the absorber, connected by means of distinct conductances to the heat bath (G_{ab}) and to the thermistor's phonon system (G_{a-ph});
- the thermistor's phonon system, connected by means of distinct conductances to the heat bath (G_{ph-b}) and to the thermistor's electronic system, which is uncoupled due to the HEM model [29];
- the thermistor's electron system, connected to the thermistor's phonon system by a conductance (G_{ph-e})

In this case the differential equation governing the dynamics of the thermal signal has an exponential solution characterized by three different time constants depending on the main conductances of the model. If we take into account also the thermo-electric feedback, these time constants are affected also by the parameters of the sensor's circuit.
In first approximation - if the thermal capacity of the sensor is inferior to that of the absorber and if the thermal conductance between sensor and absorber is better that the ones that connect them to the heat bath, the expression for the

Model parameter	**Value**
C_{abs}	2.3×$10^{-3} \cdot T^3$ [J/K]
C_{lat}	2.7×$10^{-8} \cdot T^3$ [J/K]
G_{ph_b}	4.8×$10^{-5} \cdot T^{2.4}$ [W/K]
$G_{abs-bath}$	4.0×$10^{-5} \cdot T^2$ [W/K]
G_{ph-el}	7.0×$10^{-1} \cdot T^{4.37}$ [W/K]
G_{th-abs}	1.6×$10^{-3} \cdot T^3$ [W/K]

Table 2.1: Values of the parameters in the composite model applied to the Cuoricino basic detector, with 5×5×5 cm^3 absorber and 3×3×1 mm^3 thermistor.

signal amplitude in equation (2.12) is still a good approximation (if the capacity in the expression becomes that of the absorber). Table 2.1 reports the results of different characterization measurements ([31, 32, 33]) and shows that this approximation to be correct.

2.7 Noise sources

The main noise sources limiting the resolution for thermal detectors can be classified in two main categories:

- **Intrinsic noise**: this is an unavoidable source of noise, strictly dependent on the physical characteristics of the absorber and of the sensor. It gives the reachable theoretical limit for the energy resolution.
- **Extrinsic noise**: it accounts for all the noise sources due to the experimental set-up, as the cryogenic system and the electronic read-out system. In this category can be also included other sources of noise, as electromagnetic interferences and mechanical microphonic noise. This noise actually dominates the energy resolution of low temperature experiments.

In this chapter only a description of the intrinsic noise will be given, leaving the analysis of the extrinsic sources for the chapter devoted to the detector details.

Johnson noise.
From the electric point of view and using a naive model, the thermistor can be considered a resistor. It constitutes therefore a source of Johnson noise. The power spectrum noise for a typical resistance $R \simeq 100M\Omega$ working at a temperature $T \simeq mK$ is:

$$\sqrt{e_n^2(\omega)} = \sqrt{k_B T R_b} \simeq 10^{-8} \frac{V}{\sqrt{Hz}} \tag{2.13}$$

thus meaning an rms value of about of 10 nV. Since the typical bolometric signal is on average 150 μV/MeV/kg, this is absolutely negligible. In a more realistic model, that takes into account the elettrothermic feedback effect, the thermistor should be considered a complex impedance. This implies that its noise power spectrum depends on the frequency. In any case the value given in equation (2.13) gives a good estimation for an upper limit of the actual value.

Thermodynamic fluctuations.
As previously said the number of phonons exchanged between the thermistor and the heat sink is subject to statistical fluctuation. This fact limits the intrinsic energy resolution, according to eq. (2.2) and, for working temperatures of tens of mK, it corresponds to and rms value of 20 100 eV for our bolometers.

2.8 The sensitivity of a calorimetric experiment

A useful way to introduce the main goals that a calorimetric experiment should achieve in order to observe a rare nuclear decay in its absorber (the so-called "source=detector" approach) is defyning its *sensitivity*. The sensitivity $S^{0\nu}$ to 0ν-DBD is defined as the half-life corresponding to the number of decays $N_{\beta\beta}$ that can still be considered as background fluctuations at a given confidence level. First of all, the radioactive decay law for the process under investigation can be written as

$$N(t) \sim N_0 \left(1 - \ln 2 \cdot \frac{t}{T^{0\nu}_{1/2}}\right) , \tag{2.14}$$

where N_0 is the number of nuclei under control at t=0; the equation is obtained in an approximated form by considering the inequality $T^{0\nu}_{1/2} \ll t$, which is allowed by the current limits on 0ν-DBD half-lives of order 10^{24} or greater. As a consequence, the number of 0ν-DBD decays $N_{\beta\beta}$ expected during the measurement time t is

$$N_{\beta\beta} = \frac{M \cdot N_A \cdot \eta \cdot a}{W} \cdot \ln 2 \cdot \frac{t}{T^{0\nu}_{1/2}} \cdot \epsilon \tag{2.15}$$

In the previous formula, M is the total absorber mass, N_A is the Avogadro number, η is the number of nuclides candidate to 0ν-DBD for absorber molecule, W is the molecular weight of the absorber, a is the isotopic abundance and ϵ is the detector's efficiency. When no decays are observed in the measurement time, $N_{\beta\beta}$ represents the background fluctuation in the energy region under investigation, and the $T^{0\nu}_{1/2}$ written in the previous equation becomes the limit on the half-life of the decay. In order to detect $N_{\beta\beta}$ in an energy window ΔE, which corresponds to the energy resolution of the detector, the presence of radioactive background counts is a severe limitation. The calculation of sensitivity can be performed in the case both of a zero-background experiment, and of background fluctuations that can cover the expected peak: only the second option will be considered, as it is suitable to the Cuoricino experiment and its next-generation evolutions (actually, someone may argue that we can never be sure that there is no background). Let's assume that the background level is b, measured in c/(kg·keV·y): for a detector with mass M, the number of background events B in a time t and in an interval equal to the energy resolution FWHM ΔE, centered around the Q-value of the decay, is given by

$$B = b \cdot t \cdot M \cdot \Delta E \tag{2.16}$$

Two assumptions are made: the number of background events scales with the absorber mass of the detector and b is measured independently and precisely, as it commonly happens in bolometric experiments; the background events in the region of interest will follow a Poisson distribution, with standard deviation equal to the square root of the number of events (if the number of counts allow us to approximate the Poisson distribution with a gaussian). The minimum number of events that are needed to reach a certain significance of the peak, for example at a ξ% Confidence Level, is given by the probability that the fluctuation of background events is equal to the peak. In other words,

$$N_{\beta\beta} = n_\sigma \cdot \sqrt{B}\,, \tag{2.17}$$

with n_σ equal to the number of deviations corresponding to a probability of (1-ξ)%. The combination of Eq. (2.15) and Eq. (2.17) defines the sensitivity of a calorimetric experiment for 0ν-DBD search with non-zero background rate:

$$S^{0\nu} = 4.174 \times 10^{26} \left(\frac{a}{A}\right) \sqrt{\frac{M \cdot t}{b \cdot \Delta E}} \cdot \epsilon_{\Delta E} \tag{2.18}$$

The previous expression is valid for 1σ. It is evident by looking at the formula that action on the parameters M, t, b, ΔE (and possibly a) are the crucial parameter required to increase the sensitivity of an experiment.

Chapter 3

The CUORICINO experiment

In this chapter, the final 0-DBD result of the CUORICINO experiment is presented. The lower limit on the half-life of Neutrinoless Double Beta Decay computed on CUORICINO data yielded one of the most stringent bound on me e based on ^{130}Te studies, and one of the best in general. CUORICINO data acquisition started in April 2003 and ended in June 2008. The data are separated into two runs (RUN I and RUN II), due to a major maintenance interruption. The chapter is organized as follows: after a short description of the experimental set-up in section 3.1, the details of RUN II data analysisisare presented: signal processing in section 3.2, data reduction in section 3.3 and efficiency evaluation in section 3.4. In section 3.5, the Bayesian approach used for the 0-DBD half-life limit evaluation is described. The limit is evaluated on the entire data set. This is done treating RUN I and RUN II as two independent experiments whose likelihoods are combined. This choice was motivated by the difference in detector configuration between the two runs (increased number of active detectors, improved performance) and a presumable difference in background composition (due to detector exposure to air).

3.1 CUORICINO

CUORICINO was an array of 62 TeO_2 bolometers arranged in a tower of 13 floors. Each floor consists of a single module detector as described in section 2.6. Eleven floors are made of four $5 \times 5 \times 5$ cm^3 crystals each, while the two remaining floors are composed by nine $3 \times 3 \times 6$ cm^3 crystals each.

The array was mounted in the dilution refrigerator installed in Hall A of the underground Laboratori Nazionali del Gran Sasso (L'Aquila, Italy), a location that guarantees a high degree of suppression of cosmic ray flux thanks to an average coverage of 1400 m of rock (nearly 3500 m.w.e depth).

Crystal Type	Mass [g]	^{130}Te Mass [g]	Exposure Run II [kg(^{130}Te)y]	Exposure Run I [kg(^{130}Te)y]
big	790	217	15.7	0.94
small	330	91	1.98	0.094
enriched	330	199	0.75	0.145

Table 3.1: CUORICINO crystals informations and statistics. Crystal mass is the average measured mass for CUORICINO detectors. For the ^{130}Te mass, we used the measured isotopic abundances.

The array, closed in a copper structure and hang in vacuum inside the Inner Vaccum Chamber of the refrigerator (the detectors are at a temperature of about 10 mK) is surrounded by a 1 cm thick roman lead cylindrical shield closed with a bottom and a top lead discs of thicknesses of 7.5 cm and 10 cm respectively. The refrigerator itself is shielded with a 20 cm thickness of low activity lead and a 10 cm thickness of borated PET. Nitrogen is fluxed between the external lead shield and the cryostat to avoid any Rn contribution to the detectors background.
Figure 3.1 shows a picture and the schematics of the detector's set-up.

All crystals have natural abundance of 130 Te (34%), apart for four $3 \times 3 \times 6$ cm^3 crystals: two of them are enriched in 130 Te (82%) and the other two are enriched in 128 Te (75%). This kind of detectors search for 0ν-DBD processes with the so called "source=detector" technique, where the candidate is contained inside the active mass of the detector itself. Te$\beta\beta$ candidate with a rather favorable factor of merit[1].
The signature of the decay would consist in a monochromatic peak appearing in the spectra of the bolometers at an energy equal to the Q-value of the decay: 2527.5±0.013 keV (for the experimental measurement of this quantity we refer to [34]). Indeed, in about 85% of cases the two electrons emitted by ^{130}Te 0ν-DBD decay are fully contained within one crystal.
The difficulty of the experiment consists in the control and reduction of all the background events that can mimic such a signal. These can be non-particle signals, due to electronic or thermal noise, and particle signals, due to radioactivity and cosmic rays. The former are rejected on the basis of a pulse shape discrimination technique (see section 3.3 and appendix C for further details).

---

[1] A second $\beta\beta$ candidate with high natural abundance is the isotope ^{128}Te (i.a. 31.7%). This is not as much interesting as ^{130}Te because of its lower transition energy, that reduces the nuclear factor of merit and shifts the signal in a higher background region (lowering therefore also the achievable experimental sensitivity).

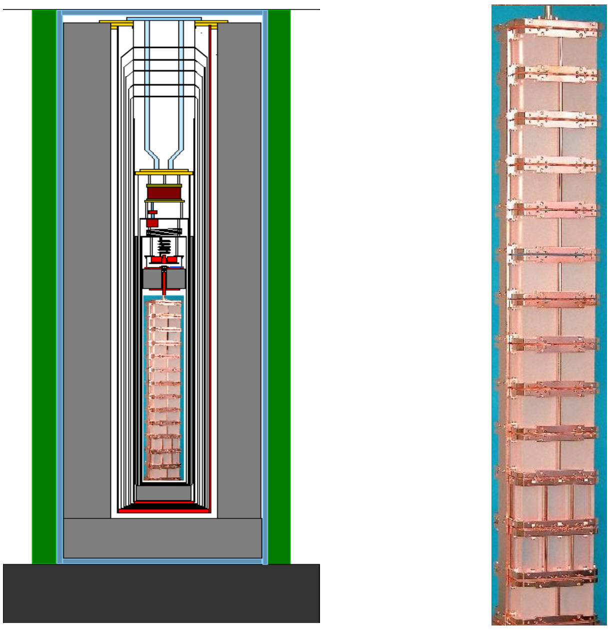


Figure 3.1: On the left, a picture of the CUORICINO tower. On the right panel, the schematics of the shieldings surrounding the detector's array.

The latter are controlled during the experiment design and construction - by a proper material selection and shielding [16, 35, 36, 37], and - at the stage of data-analysis - by coincidence cuts (see section 3.3).
A detailed description of the array, of the cryogenic set-up, of the shields and of the front-end electronics and DAQ can be find in [16] and references therein. Here we recall the main step of data acquisition and data handling which are relevant for the discussion.

3.1.1 Data and acquisition

Once the optimal working point of the detectors is chosen (2.5) cooled after a proper study of their their operating point down at a working temperature of nearly 8 mK. The electrical connections of the bolometric signals coming out from each single module runs along the CUORICINO tower over a twisted pair of low radioactivity Constantan wires. Passing through several thermalization stages these wires reach the front-end electronics boards. The electrical read-out configuration is composed of different stages: a differential preamplifier (with a gain G=128 V/V), a bessel filter to reduce high frequency noise (the cut-off frequency is set above 12 Hz on average) and a successive gain stage.

At this point, the output voltage of each detector is monitored by a constant fraction trigger. When the output voltage exceeds the trigger threshold, the acquisition system records 512 samples (″event″ in the following), corresponding to a 4 s window sampled at 125 Hz. The acquired time window fully contains the pulse development providing an accurate description of its waveform so that the energy or different informations can be computed off-line.
The existence of a pre-trigger interval just prior to the production of the pulse (″baseline″) is used to measure the DC level of the detector (which can be directly correlated to the detector temperature).

3.1.2 Extrinsic Noise

In the previous chapter we defined the intrinsic noise sources characterizing TeO_2 bolometers and we roughly estimated that they give contained contributions to the observed FWHM of these detectors. The extrinsic noise sources instead are not generated inside the detector, such as cryogenic apparatus, electronics and read-out set-ups.
The electronic read-out introduces an RMS of nearly 130 eV [33]: since the observed average FWHM of these detectors is nearly 6 keV, these sources are not dominant. Vibrations of the wires introduce the so called *microphonic noise*: since a parasitic capacitance between wires is unavoidable, these oscillations propagate to the voltage signal.
Finally, the most dangerous factor that spoils our resolution is the presence of thermodinamic fluctuations of the system and heat power arising from mechanical vibrations. Being these low-frequency vibrations, they occur in a similar frequency band as the one of the bolometric signals thus it is difficult to reduce their contribution with an off-line filtering procedure.

3.1.3 CUORICINO spectra

Each single CUORICINO measurement lasted about 22 hours on average, with the time between measurements (about 2 hours) dedicated to cryogenic system maintenance. A routine calibration with an external ^{232}Th source was performed approximately once per month, lasting for about 3 days. The accumulated data between two calibrations is referred to as a “data-set”. The spectrum obtained by summing all the CUORICINO collected data (i.e. summing over detectors and data-sets) is shown in figure 3.1.3. The background recorded by the detectors is clearly dominated, in this region, by gamma emissions due to radioactive contaminations of the detector and of the surrounding apparatus.

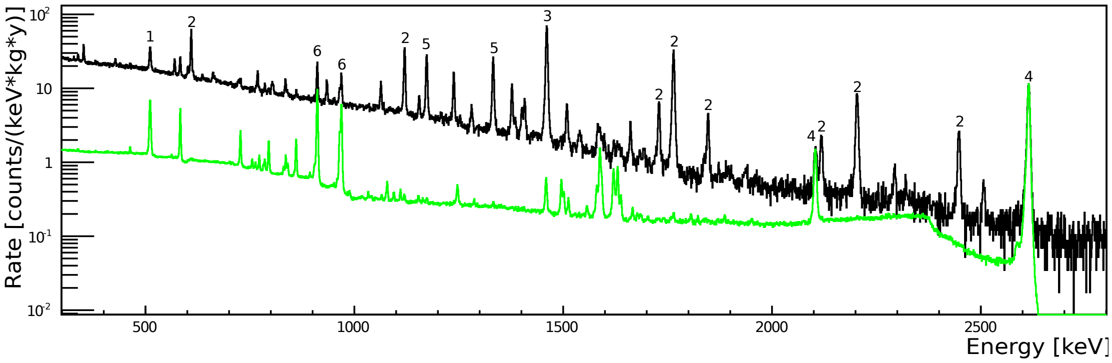


Figure 3.2: Total energy spectrum of all CUORICINO detectors (black). The most prominent peaks are labeled and come from known radioactive sources such as: e^+e^- annihilation (1), ^{214}Bi (2), ^{40}K (3), ^{208}Tl (4), ^{60}Co (5) and ^{228}Ac (6). The total energy spectrum of all CUORICINO detectors during calibration measurements is also shown (red). For convenience it is normalized to have the same intensity of the 2.6 MeV line of ^{208}Tl as measured in the background spectrum.

The most intense gamma lines are listed in reference [16]. In figure 3.1.3, the spectrum corresponding to the sum of all calibration data is also shown. For convenience the calibration spectrum is normalized to have the same intensity of the 2615 keV line of ^{208}Tl as measured in the background spectrum.

3.2 Data processing

The analysis of CUORICINO data starts with the collection of all the triggered events. For clarity, we will model the single waveform $V(t)$ induced by a particle interaction in the crystal as:

$$V(t) = V_0\, s(t) + n(t) \tag{3.1}$$

$$V_0 = G(T) \cdot A(E) \tag{3.2}$$

In the first equation, V_0 is the maximum value of the raw signal acquired at time t_0, $s(t)$ describes the shape of the particle signal and $n(t)$ is an additive noise source. The second equation describes the dependence of the signal amplitude on the detector working temperature. Here we assume that the dependence of the gain on temperature, $G(T)$, and that of the amplitude on energy, $A(E)$, can be factorized. This is not true in general, however it is a good approximation when dealing — as in our case — with small temperature drifts.

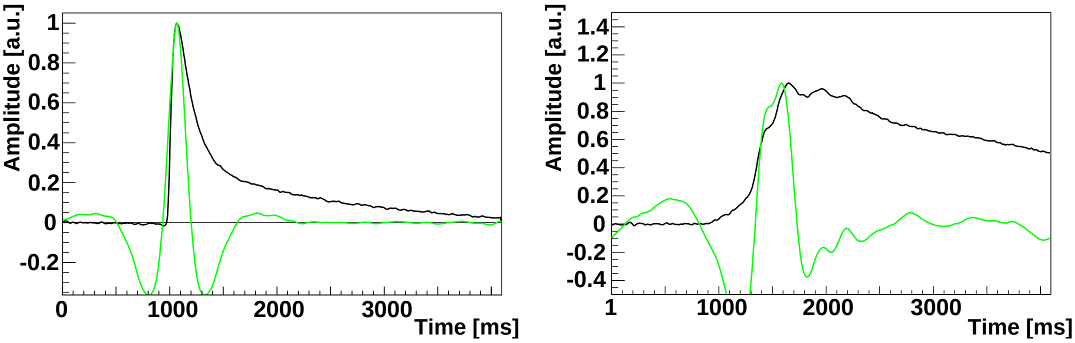


Figure 3.3: A bolometric particle event (left) and a spurious non-physical signal (right), superimposed with their optimum filter output (in color) in the time domain.

Although it describes a naive model, this formula highlights the key points of the analysis. In order to estimate E, we need:

1. a technique able to measure V_0, reducing the effect of $n(t)$ as much as possible, in order to improve our resolution (amplitude evaluation);
2. an algorithm to control for the variation of $G(T)$ produced by detector temperature drifts (gain instability correction);
3. a technique to measure the form of $A(E)$ (energy calibration).

Amplitude evaluation. This is done by maximizing the signal-to-noise ratio by means of optimum filtering [38]: each waveform is convolved with a transfer function $h(t)$ whose Fourier transform is defined as:

$$H(\omega) = \mathrm{e}^{i\omega t_{max}} \frac{S^*(\omega)}{N(\omega)} \tag{3.3}$$

where $S(\omega)$ is the Fourier transform of the average detector response function $s(t)$, $N(\omega)$ is the spectral power density of the noise characterizing the detector, and t_{max} is the time at which the pulse reaches its maximum. The functions $s(t)$ and $N(\omega)$ are computed by an averaging procedure of the bolometric pulses and of the Fourier transformed baselines. Since the developement of the Optimum Filter for Cuoricino was one of the main topics of my ph.D work, a detailed description of this technique can be found in chapter 4.

Figure 3.3 shows an example of an event due to a particle interaction and and example of a non-particle event, most likely due to an abrupt temperature increase produced by an electric disturbance or by vibrations. Each waveform is superimposed with its optimum-filtered counterpart.

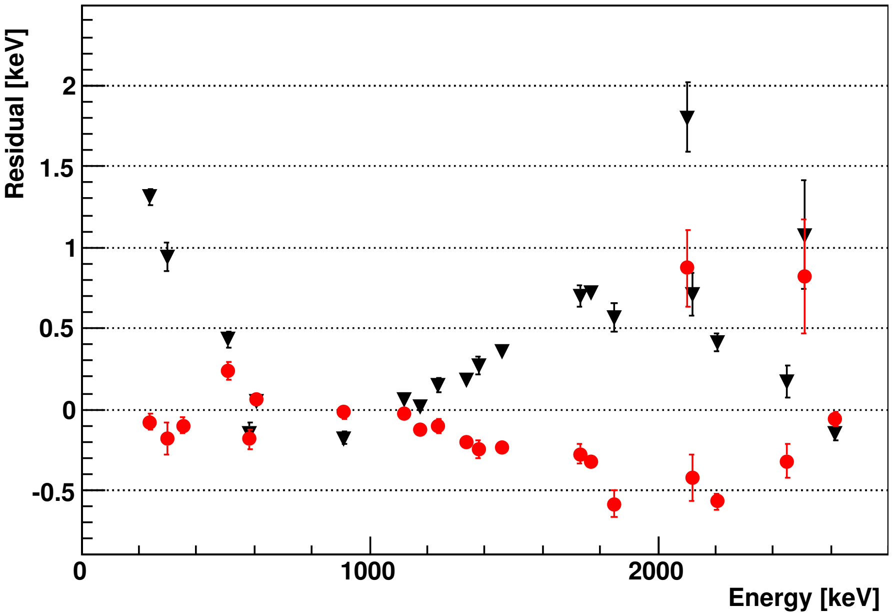


Figure 3.4: Residuals (nominal energy – calibrated energy) vs. nominal energy evaluated on the main gamma lines identified in the CUORICINO background spectrum. Circles (in color) refer to a calibrated energy obtained with the third order polynomial, triangles (in black) with the log-polynomial. Figure from M.Martinez.

Once the optimum filter is applied, the amplitude of the signal is inferred from the maximum of the filtered waveform in time domain.

Gain instability correction. This correction is achieved by measuring the voltage amplitude, V_{ref}, of a monochromatic reference pulse. This pulse is produced by depositing a fixed amount of energy into the crystal by the Joule dissipation from a heavily doped silicon resistor glued to the crystal (see section 2.5. Because the energy deposited is fixed, any variation of V_{ref} would be due to a variation in $G(T)$, which can be measured and used to correct the amplitudes of all the triggered events. For a more detailed discussion of this method we refer to [39, 40].

Energy calibration. The voltage-energy relationship is reconstructed by means of routine source calibrations: two wires of thoriated tungsten are periodically inserted between the cryostat and its external lead shield. The voltage amplitude of the pulses corresponding to the main gamma lines of ^{232}Th are used for the determination of the parameters describing the $A(E)$ relationship. This function is characterized by different non-linearity sources [16], the dominant one being the dependence of the thermistor resistance on the tem-

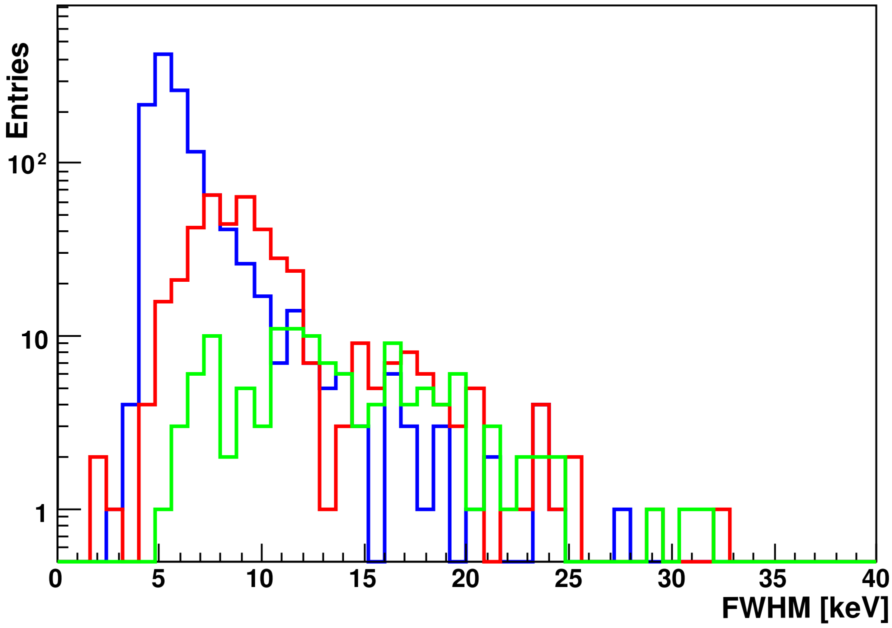


Figure 3.5: Distribution of the energy resolutions (FWHM) measured in calibration for the three groups of crystals during the 33 data-sets belonging to RUN II.They are divided according to the crystal type: big crystals (blue), small natural crystals (red) and ^{130}Te enriched crystals (green).

perature [41]. In this work, $A(E)$ is parametrized with a third-order polynomial, which can be considered as the truncated Taylor's expansion of the real unknown calibration function. In the previous CUORICINO analysis [16], a different calibration function was used. This was a second order polynomial in log(V) and log(E), based on a thermal model describing our bolometric detectors. While this function performs better at extrapolation (i.e. above the highest calibration line at 2615 keV), the third order polynomial performs better in the interpolation region (i.e. between threshold and 2.6 MeV). The difference between the two parameterizations was studied using the total background spectrum recorded by CUORICINO (figure 3.1.3); this spectrum contains several gamma peaks whose origins, and therefore nominal energies, are clearly identified. The difference between the nominal energy of each peak and its measured position (the residual) is plotted against the nominal energy in figure 3.4, showing the slightly better performance of the 3rd order polynomial. These residuals also provide important information concerning the precision of our calibration: their spread, nearly 0.4 keV, can be used as an estimator of the uncertainty in the energy position of a peak, including that produced by the 0ν-DBD signal.

Source calibration measurements are repeated for each data-set and are also used to check the detector performances over time. Figure 3.5 shows the distribution of all the resolutions measured in calibration for the three crystal groups (big, small, and enriched crystals).

3.3 Data reduction

The final CUORICINO spectrum is composed of events which survived two different types of data selection: global and event-based cuts.

Global cuts: these are applied following quality criteria decided a priori (e.g. an excessive noise level or an incompatibility between the two calibration measurements at the beginning and the end of a data-set). They identify bad time intervals to be discarded. This kind of cut introduces a dead time that is accounted for by properly reducing the live time of the detector of interest. The cuts are generally based on off-line checks that monitor the detector performances and flag excessive deviations from global control quantities (average resolution, average rate, etc.). The total dead time introduced by these global cuts is $\sim$5%. A further dead time is introduced by the rejection of a short time window centered around each reference pulse (the frequency with which the reference pulses are generated is about 3 mHz). This cut ensures the rejection of possible pile-up of a particle signal with the reference pulse (the impact of this cut is reported as an efficiency in table 3.4).

Event-based cuts: these are the pulse-shape and the anti-coincidence cuts. The former is used to reject non-physical and pile-up events (the presence of a pile-up prevents the optimum filter algorithm from providing a correct evaluation of the pulse amplitude). The latter allows for the reduction of the background counting rate in the region of interest (ROI). The 0ν-DBD signature we look for consists of a “single-hit” event (only one detector at a time involved), while many of the background counts in the ROI are due to “multiple-hit” events. These include events due to alpha decays on the crystal surfaces that deposit energy in two neighboring crystals and events due to gammas that Compton scatter in one crystal before interacting in another one.

The pulse shape parameters used in this analysis are the rise time and decay time of the raw waveform and a parameter that measures the consistency of one of the basic statements of optimum filter theory. This “Optimum Filter Test” parameter (OFT) is the difference (expressed in percentage of the total amplitude) between the evaluation of the pulse height in the time domain (as the maximum value of the filtered pulse) and that in the frequency domain (as the

Source	Signal efficiency (%)
Energy escape	87.4±1.1 (big crystals) 84.2±1.4 (small crystals)
Pulse-shape cuts	98.5±0.3
Anti-coincidence cut	99.3±0.1
Noise	99.1±0.1
Pile-up with reference pulses	97.7
Total	82.8±1.1 (big crystals) 79.7±1.4 (small crystals)

Table 3.2: Contributions to the CUORICINO 0ν-DBD signal efficiency.

integral of the filtered-pulse power spectrum). Indeed, if the shape of an event is identical to the average detector response, the two methods yield the same result. However, if the shape of the signal is different from the expected one (such as in the case of a non-physical or a pile-up event), they differ. Figure 3.6 shows the scatter plot of OFT as a function of energy for a CUORICINO detector (here only one data-set is reported). The main trend reflects the change of the signal shape with energy, and has a minimum in the region where the average response was measured (1-2 MeV). This variation in the pulse shape is caused by non-linearities introduced by the thermistor. Outliers on this plot correspond to misshapen events which will be discarded; the colored vertical bars identify confidence regions evaluated automatically on this distribution in order to obtain cuts which are independent of the signal amplitude. A full description of all the shape parameters used for Cuoricino can be found in 4, while the shape linearization procedure is described in C. The anticoincidence cuts require that only one detector fires within a time window of 100 ms.

3.4 Signal efficiency

The signal efficiency is the probability that a 0ν-DBD event is detected, its energy is reconstructed accurately, and that it passes the data selection cuts. This parameter must be accurately determined, since it is used to obtain the number of ^{130}Te 0ν-DBD events. The overall signal efficiency is: (82.8 ± 1.1)% for the CUORICINO big crystals and (79.7 ± 1.4)% for the small and the ^{130}Te-enriched ones. These efficiencies were computed as discussed below.

There are two main sources of inefficiencies, one "physical" that can be com-

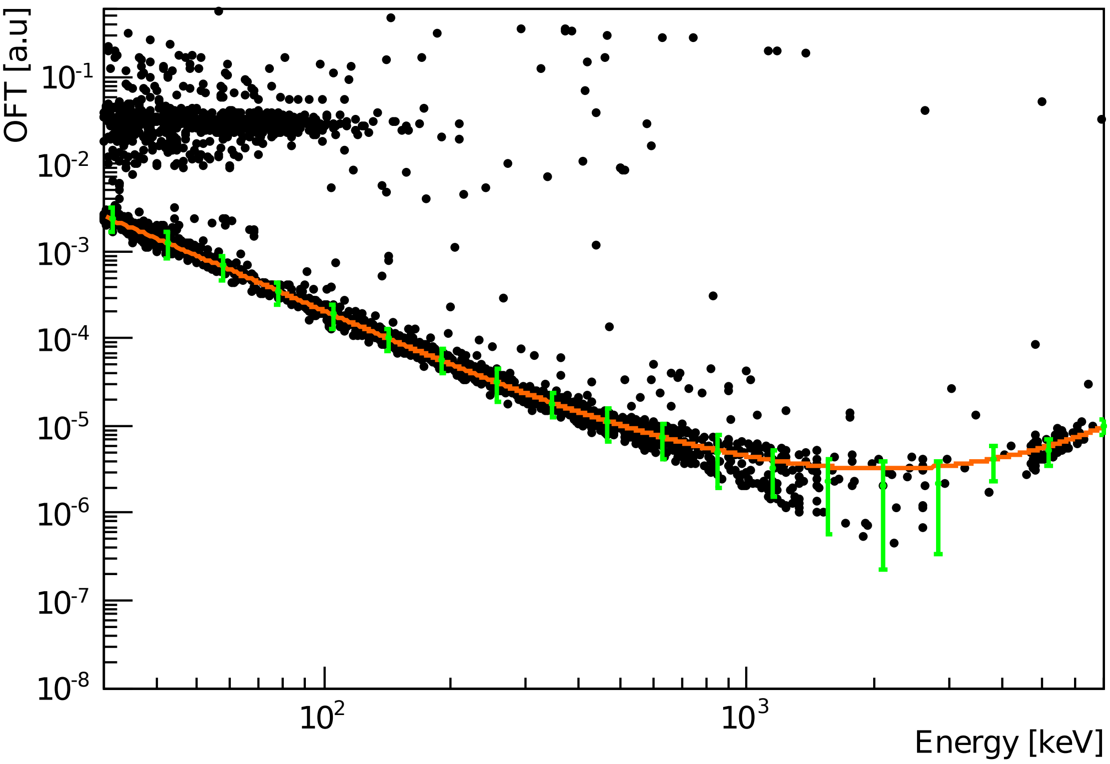


Figure 3.6: Typical scatter plot of the OFT (deviation of filtered raw signal from the average detector response) as a function of the signal's evaluated energy. The main trend identifies "good" events. Pile-up and non-physical pulses are outside the confidence regions (identified by the colored vertical bars). At low energy, there is a high density distribution of events for OFT values between 10^{-2} and 10^{-1}; these events are due to electric disturbances.

puted by simulations, and the other "instrumental" that must be measured from the data. The mechanism of "physical" efficiency loss is the escape of a fraction of the 0ν-DBD energy from the source crystal. Mechanisms for the "instrumental" efficiency loss are: the pulse-shape cut, the anti-coincidence cut and an incorrect assignment of the energy of the signal (mainly due to noise and pile-up). Their contributions to the total signal efficiency are summarized in table 3.4.

Physical efficiency: the 0ν-DBD signature is a sharp peak centered at the transition energy (Q-value) of the decay. The peak is produced by 0ν-DBD decays fully contained within the source crystal. The containment probability was evaluated using a Geant4-based Monte Carlo simulation that takes into account all the possible energy escape mechanisms (i.e. electrons, X-rays or bremsstrahlung photons escaping from the source crystals). Since the escape probability depends on the crystal geometry, the efficiency is slightly different for the big and the small crystals (see table 3.4). **Instrumental efficiency:** this is the product of the pulse-shape cut, anti-coincidence and excess noise efficiencies. To evaluate the efficiency of the pulse shape cut, the background photopeak at 2615 keV due to ^{208}Tl was used as a proxy for the 0ν-DBD peak. The 2615 keV peak was chosen because of its proximity to the 0ν-DBD energy and its relatively high intensity. In principle, an ideal pulse-shape cut should leave the main peak untouched and should only reduce the flat background. The area of the peak can then be computed in terms of the the total number of signal events (N_{sig}), the signal efficiency (ϵ_{PS}), the total number of background events (N_{bkg}) and the background efficiency (ϵ_{bkg}). A simultaneous fit was done on both the spectra of accepted and rejected events. The area of the peak in the accepted events spectrum is given by $\epsilon_{PS} \times N_{sig}$, while for the rejected events it is $(1 - \epsilon_{PS}) \times N_{sig}$. Similarly, the background yield for the accepted events is $\epsilon_{bkg} \times N_{bkg}$, and the background yield for the rejected events is $(1 - \epsilon_{bkg}) \times N_{bkg}$. By including ϵ_{PS} directly in the parametrization of the fit, correlations among the fit parameters are automatically taken into account when the error on ϵ_{PS} is calculated. The result is $\epsilon_{PS} = (98.5 \pm 0.3)\%$, and $\epsilon_{bkg} = (64 \pm 2)\%$. The pulse shape cut is clearly very powerful, rejecting approximately 36% of the events in the continuum background while retaining 98.5% of the signal events in the peak. The events discarded in this region are mainly pile-up with real or spurious signals. To estimate the efficiency of the anti-coincidence cut, we used the same procedure but considered the only available high intensity peak that is produced by a nuclear decay with no detectable coincidence radiation. This is the 1460 keV gamma line emitted in ^{40}K electron capture (the only coincident radiation, a 3 keV X-ray, is far below our threshold). Since events in this photopeak are single hits, their reduction after an anti-coincidence cut can be ascribed only to random coincidences. The last source of inefficiency is the loss of 0ν-DBD events due to excess noise which can distort the pulse shape and introduce an error in the reconstructed energy. If such an error is greater than the resolution, the event can be considered as lost in the continuum background. In order to estimate this efficiency we compare the number of reference pulses generated during the measurements (the signals used for the gain instability correction, see section 3.2) with the number actually measured in the correct energy range.

3.5 0ν-DBD analysis

The definition of the energy window used to fit the 0ν-DBD spectrum, the hypothesis assumed for the background shape and the number of free parameters

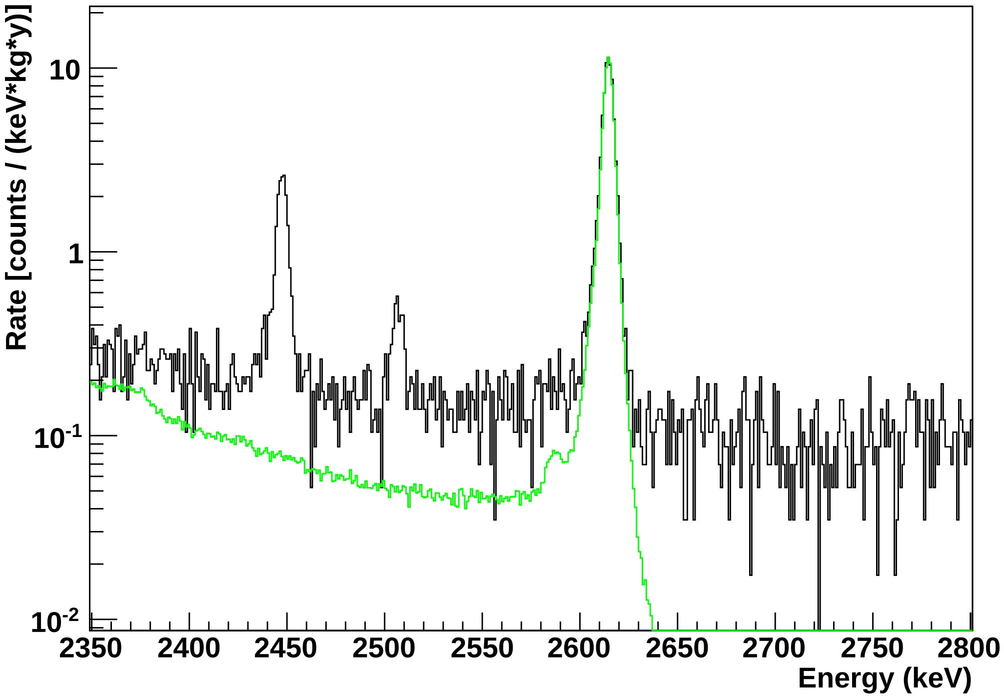


Figure 3.7: A closer view of the CUORICINO anticoincidence spectrum (presented in figure 3.1.3) near the 0ν-DBD ROI. Note that the efficiencies listed in table 3.4 are not yet included.

used to describe the background itself are extremely important for the choice of the analysis procedure and for the determination of its systematics. The choice of the energy window is somewhat arbitrary, but it influences the background representation. If the energy window is too wide (compared to the signal FWHM) a very precise knowledge of the background shape is necessary. Obviously there is also a minimum width necessary to be able to evaluate the background level beyond the 0ν-DBD peak. In our case, there is a background line near the 0ν-DBD energy, at $\sim$2505 keV, due to ^{60}Co (sum of the two photons emitted in cascade by ^{60}Co decay), which should also be included in the window. Given these considerations, our final choice for the fit window is 2474–2580 keV. This is the widest window centered on the $\beta\beta$ Q-value that allows the following two background peaks to be excluded from the fit: the 2448 keV line of ^{214}Bi and the 2587 keV Te X-ray escape peak of the ^{208}Tl line. The latter peak is clearly visible in the CUORICINO calibration spectrum shown in figure 3.7, although — due to the lower statistics — it is not visible in the background spectrum shown in the same figure.

Within this picture, we choose the simplest possible model for the shape $f_{i,j}(E)$ of the spectrum (normalized to mass, live time, efficiency and isotopic abundance) of each single detector (index i) in each data-set (index j) as:

$$f_{i,j}(E) = B_i + \Gamma^{60Co}_{i,j} \; g_{i,j} \left(E - E^{Co}\right) + \Gamma^{0\nu} \; g_{i,j} \left(E - E^{0\nu}\right) \tag{3.4}$$

Here $g_{i,j}(E)$ is the function describing the shape of monochromatic energy lines in the i^{th} detector, during the j^{th} data-set, i.e. the response function that is represented by a gaussian with a width determined from calibration data[2]. $E^{0\nu}$ is the ^{130}Te $\beta\beta$ Q-value, fixed at its measured value (2527.5 keV). E^{Co} is the sum energy of the two ^{60}Co gamma lines (2505.7 keV). B_i is the flat background component for the i^{th} detector (here we assume a time independent background). Finally $\Gamma^{60Co}_{i,j}$ and $\Gamma^{0\nu\beta\beta}$ are, respectively, the ^{60}Co activity for the i^{th} detector during the j^{th} data-set (^{60}Co has a half-life of 5.27 years), and the absolute activity for 0ν-DBD , both expressed in counts/kg/y.

Free parameters are B_i, $\Gamma^{60Co}_{i,j}$, $\Gamma^{0\nu}$ and E^{Co}, the parameters of the response function being fixed at the values measured during calibrations. Note that the dependence of $\Gamma^{60Co}_{i,j}$ on the index j is determined by the ^{60}Co half-life; therefore, the total number of free parameters is determined only by the number of detectors (i.e. by the index i).

3.5.1 Statistical approaches

The CUORICINO spectrum shows no evidence of a 0ν-DBD signal, thus we will provide a limit for the half life of ^{130}Te by means of a Bayesian approach. Unlike other low statistics methods such as that of Feldman and Cousins [42], this technique does not require an exact evaluation of the expected number of background events (which is unknown). In our case, all the uncertainties are marginalized in the process of the limit computation, and our prior knowledge for the rate $\Gamma^{0\nu}$ will be represented by a flat distribution, excluding the non-physical region.

Once the statistical method is chosen, we need to decide how to model the experiment: every CUORICINO detector can in fact be imagined as an independent search for 0ν-DBD, characterized by its own background and resolution. There are three natural approaches which can be chosen to search for a 0ν-DBD signal, to which we will refer as *ideal*, *multi-spectrum* and *average-spectrum*, respectively:

---

[2] To evaluate the energy resolution in the 0ν-DBD region we use the 2615 keV peak since this is the nearest peak to the Q-value clearly visible in our calibration spectra.

ideal: treat the different detectors separately;

multi-spectrum: treat the different detectors separately, assuming an identical background for all detectors within each group (big, small, ^{130}Te-enriched);

average-spectrum: sum the spectra of all detectors belonging to the same group.

In the *ideal* approach , each detector and each data-set is fit with its own function $f_{i,j}(E)$. In principle this is the best approach since it uses all the information available; however, the number of free parameters is huge (about a hundred).

The *multi-spectrum* approach lowers the number of free parameters by forcing the background and the ^{60}Co rates to be identical on detectors of the same group. In this approach each detector and data-set is still described individually by its own function $f_{i,j}(E)$ but the total number of free parameters is reduced since B$_i$ can assume only 3 values (for big, small and ^{130}Te-enriched crystals) and the same is true for $\Gamma_{i,j}^{60Co}$. This could be considered a strong assumption, but it is motivated by the fact that the low statistics prevent us from being sensitive to background variations among crystals of the same group in the 0ν-DBD region. This method also offers the advantage of being less sensitive to fluctuations in the counting rate of a single detector over time, and takes into account the decay rate of ^{60}Co.

The *average-spectrum* method removes the background assumption of the previous model, at the price of a certain degree of information loss. The counting rate is simply averaged over all data-sets and detectors for the three mentioned groups. A variation of the background and of the ^{60}Co rate over time is then irrelevant, provided that the response function does not change with time. The average is done simply by summing over all the data collected with detectors belonging to the same group, thereby obtaining three spectra that can be represented by the function:

$$f_k(E) = B_k \,+\, \Gamma_k^{60Co} \; g_k\left(E - E^{Co}\right) + \Gamma^{0\nu} \; g_k\left(E - E^{0\nu}\right) \tag{3.5}$$

Here, the index k has three allowed values for big, small and ^{130}Te-enriched crystals while the response function $g_k(E)$ is defined as:

$$g_k(E) = \frac{1}{\sum\limits_{i,j} A_{i,j}} \sum_{i,j} \frac{A_{i,j}}{\sqrt{2\pi}\sigma_{i,j}} \exp\left(-\frac{(E - E_0)^2}{2\sigma_{i,j}^2}\right) \tag{3.6}$$

where the sum over i extends on all the detectors belonging to the k^{th} group and j runs over all the data-sets. $A_{i,j}$ and $\sigma_{i,j}$ are the corresponding background

Fit Parameter	big crystals	small crystals	^{130}Te-en. crystals
RUN I: flat bkg rate	0.20±0.02	0.20±0.02	0.8±0.4
RUN II: flat bkg rate	0.153±0.006	0.17±0.02	0.35±0.05
RUN I: ^{60}Co rate	4.6 ± 1.5	9 ± 6	0 ± 14
RUN II: ^{60}Co rate	2.5 ± 0.3	1.7 ± 0.8	0 ± 3.5

Table 3.3: Background and ^{60}Co rates expressed in counts/keV/kg/y obtained by the combined fit of RUN I and RUN II data. The absolute rates are about 20% higher because the efficiencies have not been yet included. The high background rate obtained for ^{130}Te-enriched detectors is mainly due to a higher intrinsic contamination of these crystals.

exposure and energy resolution measured during calibration. Note that — as is true also for the other two approaches — the response function is built using measured quantities, i.e. it does not contain any free parameters.

We discarded the first approach due to the excessive number of free parameters, and we performed two parallel (and independent) analyses on real and Monte Carlo-simulated data for the limit computation, following an unbinned likelihood technique [43] for the second approach and the standard CUORICINO Likelihood-Chi-Square technique [16, 44] for the *average-spectrum* one. The goal was to choose the most reliable procedure, checking for possible biases and comparing performances. This Monte Carlo analysis, described in details in chapter 5, proved an impressive correspondence between the two approaches. Since the choice of the one or the other is somewhat arbitrary. We opted for the third method because it is consistent with the previous analysis of CUORICINO data [16, 44] and because of its intrinsic simplicity.

3.6 The CUORICINO final result

As a next step, we added the contributions from big, small and enriched crystals from RUN I, combining their likelihoods with the RUN II data and using a similar reconstruction for the response function as described in section 3.5.1. Finally, an evaluation of the overall systematic error on $\Gamma^{0\nu}$ has been carried out. This study will be documented in chapter 5.
The background rates are shown in table 5.5 while in figure 3.8 (left panel) we show the best fit and the corresponding 68% and 90% C.L. limits. Figure 3.8 (right panel) shows the logarithm of the combined likelihoods of RUN I and RUN II before and after the systematic uncertainties are included.

Isotope	$\tau^{0\nu}_{1/2}$ [y]
^{130}Te (CUORICINO, this work)	$> 2.8\ 10^{24}$
^{76}Ge (Heidelberg-Moscow collaboration [12])	$> 1.910^{25}$
^{100}Mo (NEMO collaboration [15])	$> 5.8\ 10^{23}$
^{136}Xe (Dama/LXe [17])	$> 1.2\ 10^{24}$
^{76}Ge (Heidelberg-Moscow experiment [18])	$= 2.23^{+0.88}_{-0.62} 10^{25}$

Table 3.4: We compare the most stringent 90% C.L. half-life lower limits present in literature (first 4 rows). In the last row is reported the claim for a 0ν-DBD signal of ^{76}Ge.

The resulting best fit for the 0ν-DBD rate of ^{130}Te is:

$$\Gamma^{0\nu}_{\rm best} = (-0.25 \pm 1.44({\rm stat}) \pm 0.3({\rm syst})) \times 10^{-25}{\rm y}^{-1}$$

This result is compatible with zero, and the corresponding 90% C.L. half life lower bound is:

$$\tau^{0\nu}_{1/2} \geq 2.8 \times 10^{24}{\rm y}$$

This limit is almost identical to the one we published in [16], despite the increase of the total exposure by a factor ~1.6. In general it is expected that the limit scales with the square root of the exposure, i.e. we would expect an improvement of about a factor 1.3 which is by far smaller than the spread in the 90% C.L. limits that different experiments (with the same exposure and sensitivity) can yield (figure 5.14). This is the reason we prefer to quote the sensitivity of the experiment together with the limit. It should to be mentioned that, in reference [16], we used an older value of the ^{130}Te transition energy energy which had a much larger error and a slightly higher central value than the recently measured one [34, 45]. With the same data used in [16], using the new result for the transition energy (with its smaller error) pushes the limit toward a lower half-life.

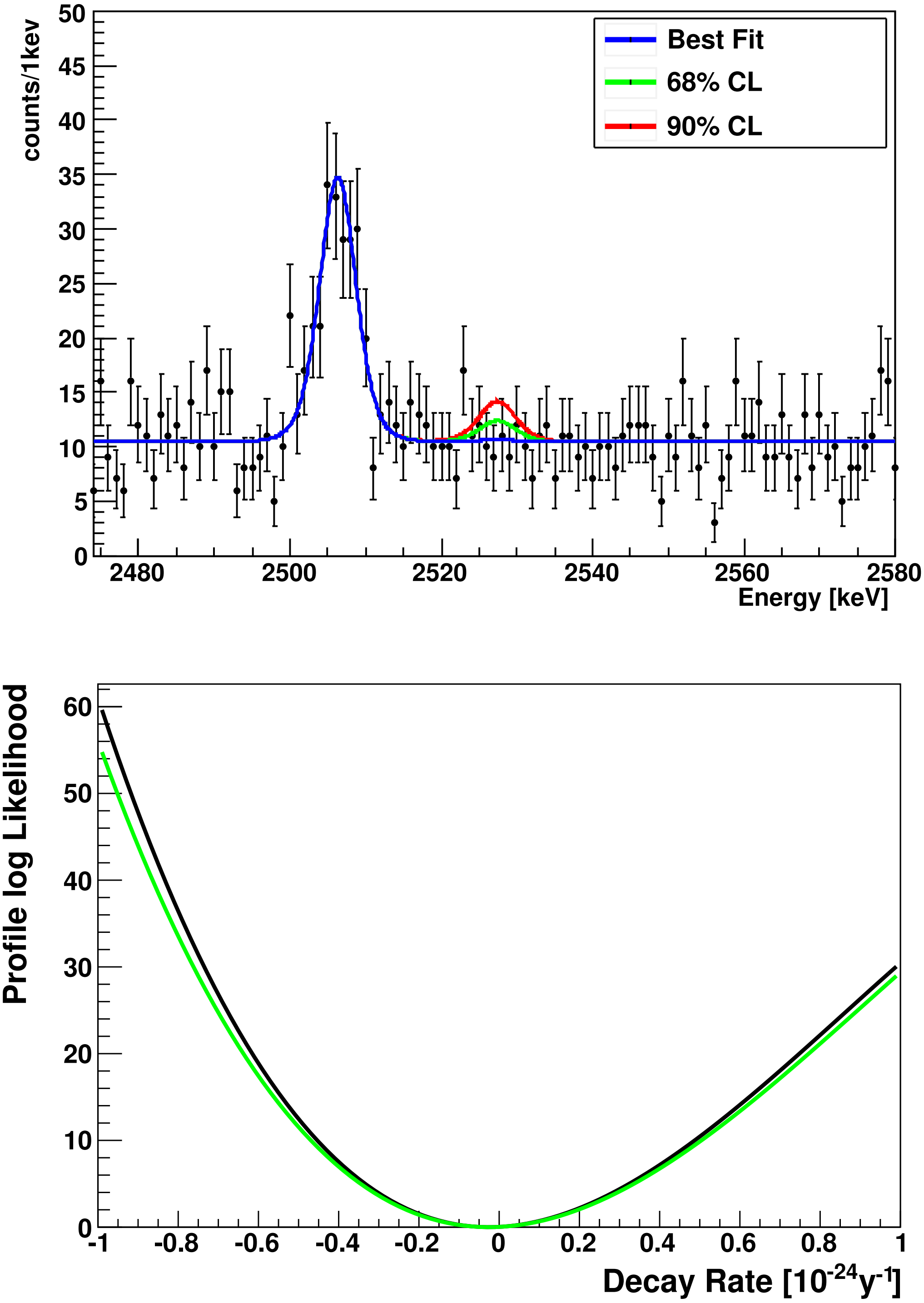


Figure 3.8: Top panel: best fit, 68% and 90% confidence intervals for the total statistics (RUN I+RUN II) superimposed on the CUORICINO sum spectrum of the three groups of crystals (each scaled by efficiency and exposure) in the 0ν-DBD region. (The purpose of the plot is to give a pictorial view of the result; in fact the fit was performed separately on 6 spectra whose likelihood are combined, as described in the text). Bottom panel: negative profile of the combined log likelihoods of RUN I and RUN II before (black) and after (color) the systematic uncertainty is included.

Chapter 4

Signal processing

Digital Signal Processing (DSP) refers to various techniques for improving the accuracy and reliability of informations computed on a digital signals.
In CUORICINO, when the trigger fires, a sample of N converted data, corresponding to a time window T, is saved to disk. The time window is optimized in order to reduce as far as possible signal truncation: in CUORICINO the sampling frequency is 125 Hz, corresponding to 512 points acquired in 4 seconds.
The purpose of signal processing is to extract as many informations as possible from each single event, such as the energy of a recorded particle interaction. The main obstacles that need to be treated as much carefully as possible are:

- the bolometric noise superimposed to the signal that sploils the resolution and must be filtered;
- the identification of spurious signals: non-physical events that triggered the acquisition;
- the pile up with other signals that leads to a wrong amplitude evaluation;

The core part of my thesis work has been focused on the implementation of different DSP algorithms to study bolometric signals.
These are: the Optimum Filter (OF), a digital filter that mitigates the influence of noise for the first item; the calculation and study of pulse shape parameters for the second and third items; the Wiener Filter (WF) technique, a digital filter to gain time resolution resolution.
For each shape parameter, an example on a bolometric run characterized by a high pile-up rate will be given in order to give an idea of their general distributions.
After a detailed description of these algorithms, a validation of their performances on Monte Carlo simulations is also presented.

The difficulty of a modeling of the complex noise patterns characterizing TeO_2 made this kind of test a real difficult task. In order to overcome the lack of a theoretical modelization of CUORICINO noise behaviour, I developed a Monte Carlo method to produce random noise time series having, on average, the same measured noise spectra of the bolometers. The general idea of the algorithm is summarized in Appendix B. This is a starting point of a full simimulation of the detector, which, in the future, will be combined by the modeling of the bolometers response function recently proposed [41].

4.1 Notation

In this context, a signal, or an *event*, is considered as a collection of couples $\{R_i, t_i\}$, where R_i is a recording of the detector's bias voltage, aquired at an instant $t_i = i \times \tau$, where i is an integer index running from 0 to the number of acquired points and $1/\tau$ is the sampling frequency. To simplify the notation, I will identify a generic vector of samples simply as $\{R_i\}$, omitting the occuring time t_i to reduce redundancy.
$\mathcal{R}_i$ will instead identify the complex number associated to the i^{th} frequency of the Discrete Fourier Transform of $\{R_i\}$.

4.2 Optimum filter

In our bolometric detectors, a signal, amplified by the front-end electronics, is fed into a high resolution ADC. The Bessel filter at the end of the analog chain is used as anti-aliasing that could spoil the signal information. It's cut-off frequency is - by design - opportunely chosen to be higher than the band width of the signal but much lower than the ADC sampling frequency, avoiging thus any aliasing that can reduce the signal.
The acquired voltage signal - as a function of time - can be modelled as:

$$V(t) = b + a \cdot s(t - t_0) + n(t), \tag{4.1}$$

where b is the baseline DC level of the detector, a is the pulse amplitude, t_0 is the start time of the pulse (tipically 1 s after the the time start of the window), $s(t)$ is the detector response function (the pulse shape without noise, normalized to unit height), and $n(t)$ is the stochastic noise. The baseline, b, is the detector's bias voltage, and is generally estimated by averaging over the $\sim$1 s pre-pulse interval.
The simplest estimate of the pulse amplitude is the pulse maximum minus the

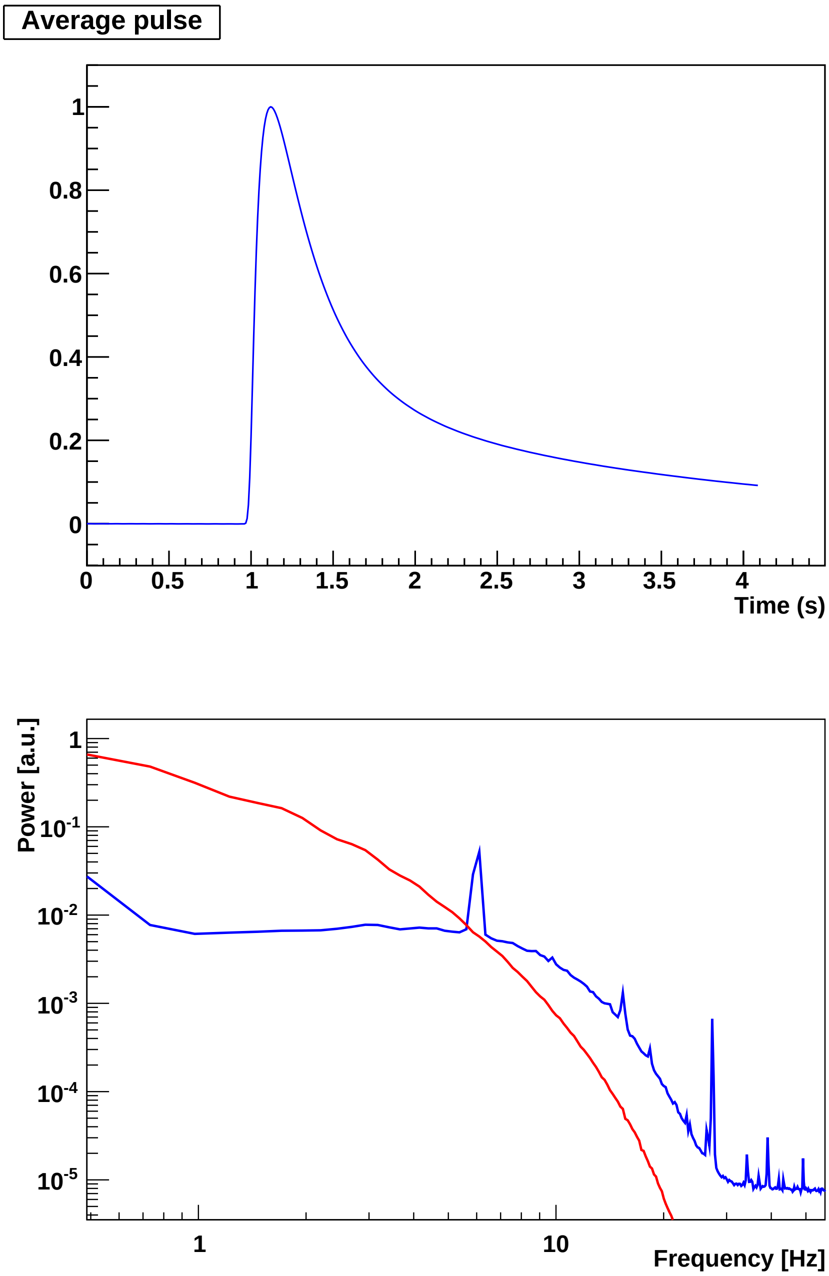


Figure 4.1: Top: The average pulse for CUORICINO channel 1 (Figure from A. Bryant). Bottom: The average noise power spectrum (red) and the modulus of the discrete Fourier transform of the average pulse (blue) for CUORICINO channel 1 (scaled to be equal at the lowest frequency).

baseline value: this measure is highly effected by $n(t)$. It can be shown, following [38], that if $n(t)$ has a stationary behaviour and if the shape $s(t)$ doesn't change, we can obtain the best estimate for the signal amplitude by filtering it with a transfer function whose components in the frequency domain are:

$$H(\omega) = \eta \frac{\mathcal{S}(\omega)^*}{|N(\omega)|^2} e^{i\omega t_M} \tag{4.2}$$

where $S(\omega)^*$ is the complex conjugate of the Fourier transform of the ideal response $s(t)$, $|N(\omega)|^2$ is the noise power spectrum of the main disturbs affecting the signal, t_M is the instant when the maximum of $s(t)$ is reached and η is a proper normalization constant. The OF technique weights different Fourier components of the pulse differently depending on the expected signal-to-noise ratio (SNR) of each frequency component: frequencies that are highly influenced by noise are down weighted, thus mitigating their influence on the pulse amplitude.

At this point, as it will be shown in the next session, the estimate for the amplitude a will be the maximum value of the OF's output in time domain. Since we are dealing with discrete signals, it can be useful to consider the DFT counterpart of equation (4.2):

$$\mathcal{H}_i^{\mathrm{OF}} = \frac{\mathcal{S}_i^*}{\mathcal{N}_i} \tag{4.3}$$

where $\mathcal{S}_i$ and $\mathcal{N}_i$ are the discrete components of the Fast Fourier Transform of the average pulse and of the noise power spectrum. In this equation no phase shift is applied to the transfer function: this simply implies that the output of the filter in the time domain $\{O_i\}$will be shifted so that the time axis will represent the delay between the raw pulse and the average pulse, without modifying its amplitude.

An estimation for $\{S_i\}$ and $\{\mathcal{N}_i\}$, is produced by averaging on acquired signals. Since this is done from scratch, it is necessary to follow a recursive process: the reason is that it is necessary to separate particle signals from pure noise events (baselines) or to reject spurious signals such as pile-ups that would spoil the result.

This is accomplished following this procedure:

- The first step is a rough hardware classification: particle signals are flagged by the ADC under the veto of a constant fraction discriminator, while baseline events are random caption of the detector voltage bias acquired in anticoincidence with the particle trigger.

- The second step consists on an evaluation of the amplitude and other shape indicators to select signals as "good" as possible for a first averaging. For noise events, the signals are firstly windowed with a Kaiser

Notation	Meaning
$\{R_i\}$	Raw pulse
$\{S_i\}$	Average pulse
$\{O_i\}$	OF time-domain output of $\{R_i\}$
$\{A_i\}$	OF time-domain output of $\{S_i\}$
$\{\mathcal{N}_i\}$	Noise Power Spectrum
$\{\mathcal{H}_i\}$	OF transfer function
$\mathcal{R}_i, \mathcal{S}_i, \mathcal{O}_i, \mathcal{A}_i$	DFT components

Table 4.1: Notation Summary

function and then their average is computed by summing over their Fast Fourier transform. Particle signals are also shifted so that they are synchronized with an arbitrary reference point (the position of a pulse can in fact change due to the jitter of the trigger. In this way we have a first raw evaluation of $\{\hat{\mathcal{H}_i}\}$ of $\{\mathcal{H}_i\}$.

- Finally, using $\{\hat{\mathcal{H}_i}\}$ the amplitude and the shape indicators are computed using the OF output so that the noise fluctuation are as much contained as possible to produce a refined evaluation of $\{\mathcal{H}_i\}$.

As an example, figure 4.2 displays the average pulse and average noise power spectrum for CUORICINO channel 1.
There is still one more point: leakage. Leakage is an error introduced by DFT when the a sampled function is not completely contained in the acquired window. As can by see on the top panel of figure , the signal is trucated on its right tail. The simple solution that was chosen is that signals to be filtered and the OF's transfer function are differentiated. Since bolometric signals have very flat tails, their average derivative at the beginning and at the end of the acquired will be zero, so that all the information is fully contained. It can be easily shown that this operation doesn't modify the output of the filter (in Fourier space derivation becomes multiplication), if the differentiated counterparts of the average pulse and spectral density are used.
For clarity, a summary on the notation that will be used in this chapter is given on table 4.1.

4.3 Amplitude evaluation

In Fig. 4.2, a typical bolometric signal and the output of OF are shown. The simplest amplitude estimator for the signal energy is the sample $A_{i_{max}}$, where:

$$i_{max} = \arg\max_i \{O_i\} \tag{4.4}$$

Since the OF can be seen as an autocorrelation function, it is simply the point where the auto-correlation is maximal.
However, there is an intrinsic limit on the resolution of the pulse height measurement imposed by the sampling frequency and by the Bessel filter used in the acquisition chain, which limit the spectral information (the real instant at which the pulse reaches its maximum value can infact fall between two acquired samples).
A technique to improve the uncertainty on the pulse height due to discrete sampling is to consider the parabola that intersects $O_{i_{max}}$, $O_{i_{max}-1}$ and $O_{i_{max}+1}$. By considering the analytical maximum of this parabola as the height of the pulse, we should gain in resolution. In figure Fig. 4.3 three superimposed peaks are shown, corresponding to three different techniques of amplitude estimation (in mV). These are events caused by the gamma emission at 2615 keV of ^{208}Tl deposing their whole energy in the bolometer. In blue it shown the height of the raw pulse E_{raw}, defined as:

$$E_{raw} = R_{i_{max}} - \hat{b} \tag{4.5}$$

where $\hat{b}$ is an average of the first samples to extimate the baseline level.
In red, the simple height of the OF output and in black, the height of the interpolated parabola. Even though at these energies the SNR is quite high, the improvement in the resolution is evident. The amplitude can also be evaluated in the frequency domain by integrating over frequency components after applying Eq. (4.3); this alternative will be discussed in Sect. 4.5.3.

The OF provides also a valuable information on the pulse timing: in data analysis, for example to identify coincidence patterns between detectors, it is important to measure the position of a signal with sufficient accuracy. but when secondary pulses are present, identification of the primary pulse (the one that triggered the acquisition) is important: the position of the primary pulse is identified as the analytical value of the position of the maximum of the interpolated parabola.

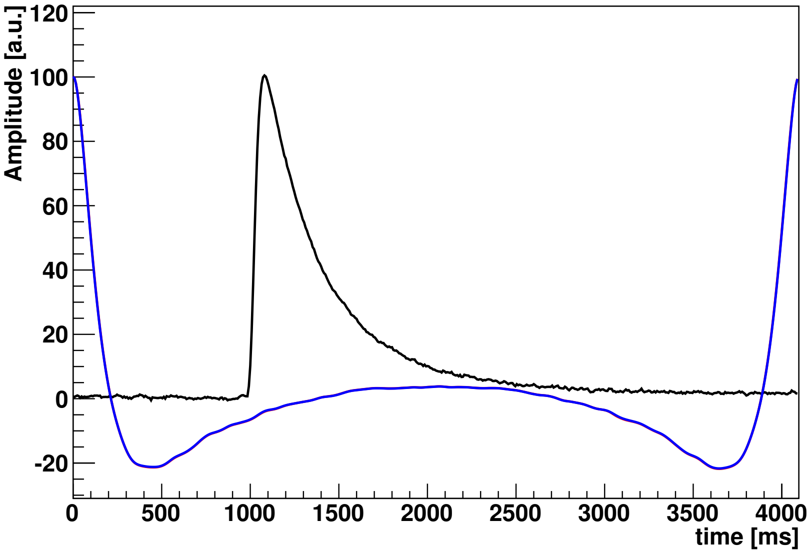


Figure 4.2: A raw bolometric pulse (black) and its time domain OF output (blue).

4.4 Main pulse identification and Wiener filter

When a secondary pulses are present, the identification of the primary pulse with the OF alone can be difficult and - sometimes - even impossible. An example of this issue can be seen in figure 4.4 where a close pile up is not resolved in OF output.

The technique implemented for the identification of the main pulse is the "Wiener Filter": a deconvolution of the signal from the detector response in order to obtain a narrow time resolution. The Wiener filter transfer function is the following:

$$H_i^{\text{Wiener}} = \frac{\mathcal{S}_i^*}{\mathcal{N}_i + \alpha|\mathcal{S}_i|^2} \tag{4.6}$$

where $\mathcal{S}_i$ and $\mathcal{N}_i$ are the, respectively, the i^{th} DFT component of the average pulse and the i^{th} component of the noise average power spectrum, while α is a parameter that depends on the energy of the signal to be filtered. Basically, α scales the average pulse power so that it is correctly compared to the noise average contribution.

Given R_i the i-th DFT component of the OF output, α is computed as:

$$\alpha = \sum_{i=1}^{N} \sqrt{|\mathcal{O}_i|^2} \Bigg/ \sum_{i=1}^{N} \sqrt{|\mathcal{A}_i|^2} \tag{4.7}$$

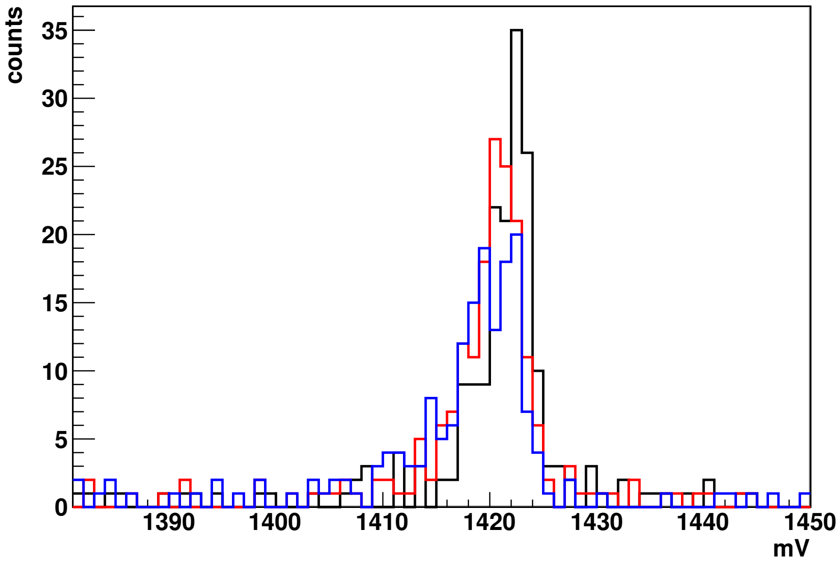


Figure 4.3: Resolution gain by applying the different amplitude estimation techniques described in the text: raw pulse height (blue), OF height (red), and height of the interpolated parabola (black).

which corresponds to the ratio of the amplitudes of the filtered signal and of the filtered average pulse, computed in the frequency domain (see section 4.5.3).
If the noise component at the denomitator is zero, the WF becomes the exact formula for the deconvolution of a signal, whose output should be a delta function centered at the signal starting point. The noise component plays the same role as in the OF's transfer function: it downweights the frequencies where the noise power is high.
As an example, figure 4.5 shows the WF output (red curve) applied to the same raw pulse shown in figure 4.4: the two peaks are reasonably resolved.
The next step is to use a proper algorithm for maximum identifications: this is due to the fact that the noise contribution cannot be suppressed as much as with the OF technique, so it can happen that some random noise fluctuations could be identified as signals.

In order to perform a robust signal identification, the following steps are iterated:

1. the RMS of the filtered signal is computed
2. given a factor λ (> 1), the position t_{max} of the first maximum above λ·RMS is flagged as signal.

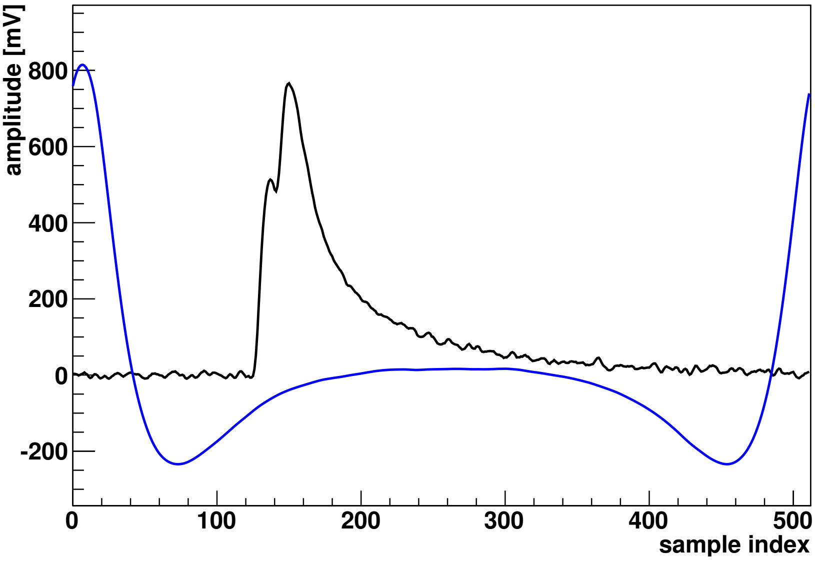


Figure 4.4: A close pile up in the raw signal (black curve) is not resolved by the OF's output (blue curve)

3. all the samples within a time window between $t_{max} - \Delta t_l$ to $t_{max} + \Delta t_r$ are set equal to zero, where Δt_l and Δt_r are the time distance in which the filtered signal goes from t_{max} to its first minimum on the left and on the right, respectively.

4. the procedure is then iterated until there are no maxima above the threshold.

In this way it is possible to identify which peak is the particle signal that triggered the acquisition just by considering the one that is nearest to the trigger sample which is fixed by the DAQ setup. This technique can be also applied in pile-up rejection but presents many difficulties, since it is not simple to find an automatic tuning procedure of the parameter λ for large amounts of data so it has been used in CUORICINO mainly for the primary pulse identification.

4.5 Pulse Shape Parameters

An important feature of bolometric signals is that their time constants depend mainly on intrinsic characteristics of the detector (in particular, the heat capacities and the thermal conductances of the components of the assembly). It is

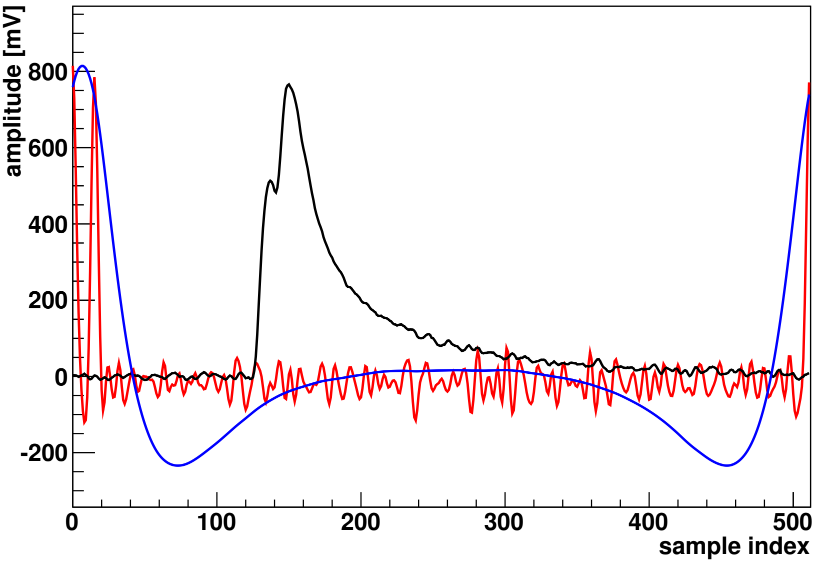


Figure 4.5: A raw bolometric pulse with a pile-up (black), compared to its Wiener (red) and Optimum (blue) filter outputs.

then reasonable to extract shape parameters from the raw pulse or from the OF output to measure the deviation from the average response function of the detector in order to discard spurious events. The main shape parameters used in our pulse shape analysis are: the rise time and decay time, evaluated on the raw pulses, and three different measures of the discrepancy between the filtered pulse and the average detector's response (TVL, TVR and OFT).
In the following section I will analyze in detail the ideas behind this shape parameters. In each section I will show a scatter plot for each single parameter as a function of the energy, evaluated on a bolometric run with a 5x5x5 cm^3 crystal of TeO$_2$, characterized by a high rate of pile-up. This is done to give a picture of the distribution of these parameters in the worst possible situations. In these measurements, in fact, the main source of pile-up was the ^{210}Po alpha decay at 5.4 MeV: on each scatter plot we separated events identified as pile-ups by the WF (red dots), in order to have a better feeling of the discrimination power of each parameter.

4.5.1 Rise and Decay Time

These are the most "natural" shape parameters that can be used to characterize a bolometric signals.

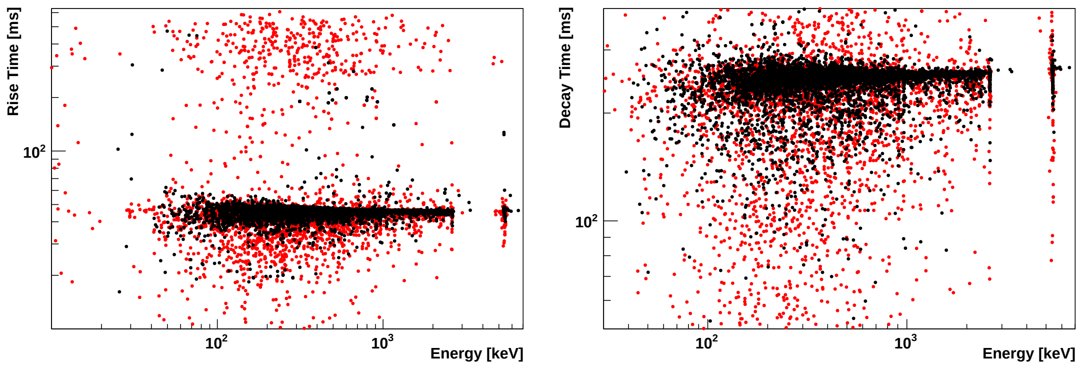


Figure 4.6: Scatter plots of the logarithm of rise time (left) and decay time (right). Red dots are pile-ups identified by the Wiener filter.

We define the rise time of a signal as the time interval between 10% and 90% of the primary pulse height. Since the sampling is discrete, a linear interpolation is applied in order to approximate better the instant where the two fractions are reached. The same procedure is performed to calculate the decay time of a signal, with the only difference that is the time interval intercurring from 90% of the amplitude to its 30%.

Since these parameters are computed directly on the raw signal, a sort of control of noise fluctuations is also applied: the raw pulse is smoothed with a triangular function: a weighted sum of each bin with its neighbours where the weights decays linearly with the distance. The i$^{\mathrm{th}}$ sampled value R_i becomes:

$$\hat{R}_i = \frac{1}{N^2}\sum_{j=1}^{N}(N \cdot R_i + (N-j) \cdot R_{i-j} + (N-j) \cdot R_{i+j}) - \hat{b} \tag{4.8}$$

where N is number of samples used for averaging and $\hat{b}$ the average baseline value. In our analysis N=3.

Then, the OF and WF technique are used to find the position and amplitude of the primary pulse. Starting from the maximum position we find the time instants $t_{90\%}$, $t_{10\%}$ and $t_{30\%}$ when the desired percentage of the maximum are reached. Rise and Decay time (RT and DT) will then be:

$$RT = |t_{90\%} - t_{10\%}| \tag{4.9}$$

$$DT = |t_{90\%} - t_{30\%}| \tag{4.10}$$

In Fig. 4.6 we show two typical scatter plots of these two quantities as function of the time domain amplitude: the black dots are those flagged as single events

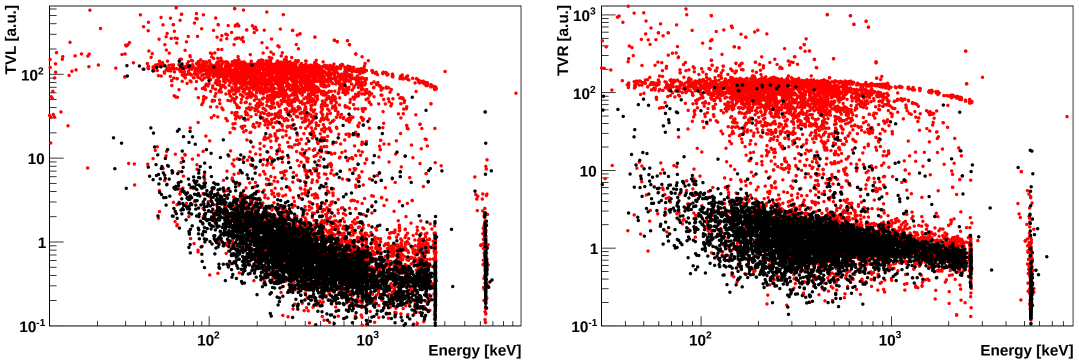


Figure 4.7: TVL (left) and TVR (right) as functions of the energy. Red dots are pile-ups identified by the Wiener filter.

by the Wiener filter, the red ones are pile-ups. The fact that a pile-up event can have a "good" rise (respectively, decay) time is due to a secondary event occurring after (respectively, before) the maximum of the main pulse.
In this measure, the distribution of rise and decay are nearly two horizontal line as functions of the energy: this means that the shape of the pulses can be considered almost constant with the energy.

4.5.2 Test Values

Under the assumption that the average response of the detector is independent from the energy, it should be possible to produce an estimate, a test value (TV) of the deviation of each acquired signal $R(t)$ from the properly normalized and synchronized average pulse $A(t)$ as:

$$TV = \frac{1}{N}\sum_{i}^{N}(A_i/A_{max} - O_i/O_{max})^2 \tag{4.11}$$

where A_{max} and O_{max} are maxima of the filtered average pulse.
The synchronization between the two waveform is carried on as follows.
The aim is to shift each A_i so that $A_{i_{max}}$ and $O_{i_{max}}$ are as close as possible. The index i_{max} identifying the maximum of the pulse O_i is the nearest integer to t_{max}/τ, where t_{max} is the position of the maximum identified by the parabolic interpolation as described in 4.3. A shift of i_{max} is easy with sampled signals, since it is sufficient to substitute each sampled A_i with $A_{i-i_{max}}$ (the OF output is perfectly periodic by construction), but it will be necessary to compensate also for the fractional part of the delay $i_{fraq} = t_{max}/\tau - i_{max}$, which, due to the

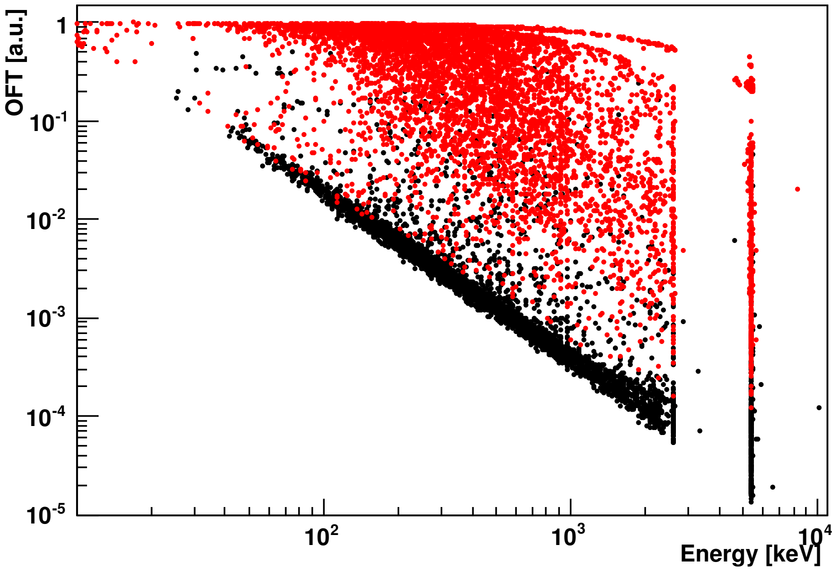


Figure 4.8: Scatter plot of the deviation from unity of the ratio of amplitudes in the time and frequency domains (log scale) as a function of the OF maximum. Red dots are pile-ups identified by the Wiener filter.

fluctuation of the triggering jitter, will be generally different from zero. This is done with the following steps:

- a linear interpolation between A_i and A_{i+1} is performed, obtaining an equation $f(i)$;
- then A_i becomes simply $f(i + i_{fraq})$

Once the average pulse and the signal are synchronized, it still remains to be decided the number of samples to be used in equation (4.11). Averaging on all the time samples will result in an information loss: we decided to construct two different etimator, for the left and the right pulse of the pulse near its maximum. The number of point is twice the number need to reach half of the signal maximum (in CUORICINO, on average there are 20 points). As can be appreciated in figure 4.7, the pile-up distribution is more resolved is more separated from the single events: a simple cut like TVL¡10 and TVR¡10 would discard a consistent fraction of pile-ups.

4.5.3 An Optimum Filter Test (OFT)

As described in Sect. 4.3, the amplitude of a single particle event can be measured as the height of the OF output in time domain, but the theory of this digital filter offers a different approach, which deals with the problem in the frequency domain. At first, we define the average of the square root of the OF output power spectrum as:

$$E_f = \frac{1}{N} \sum_{i=1}^{N} \sqrt{|O_i|^2} \tag{4.12}$$

where O_i is again the i^{th} DFT component of the filtered signal. It can be shown that, as long as the SNR is sufficiently high and the main hypotheses of the OF theory hold (basically if the noise is stationary and the waveform to be filtered is as similar as possible to the response of the detector to a delta stimulation - [38]), this is another estimator for the signal energy.

Intuively, each DFT compontent of the filtered signal is an independent estimation of the signal energy: thus, equation (4.12) can be seen as an average over these different estimates.

The main difference between these two approaches is that, by working in the frequency domain, we are dealing with the information contained in all the acquired window while by working in time domain, we consider only the region near the pulse maximum. It follows that, when the signal shape is different from the standard detector response function or when there is a secondary pulse in the acquired window, these two numbers will differ.

The shape parameter that has to be extracted following this assumption can be considered as an Optimum Filter Test (OFT) and is defined as:

$$OFT = 1 - \frac{E_{td}}{E_{fd}} \tag{4.13}$$

where E_{td} and E_{fd} are, respectively, the amplitudes evaluated in frequency and time domain.

In Fig. 4.8 we show the scatter plot of OFT as a function of the time domain amplitude E_{dt}. Given the stability of rise and decay time, the steep decent of this parameter could not be caused simply by a shape variation. It must be addressed also to the way this parameter is expressed: for example, since the frequency-domain is an integral of all the information of the recorded waveform, it could affected by a costant noise component which becomes less and less important as the pulse height increases.

In this case, the good pulses form a high density band, and as can be seen in the figure and almost no events identified as pile-up fall in the high density

region, proving a good sensitivity of this parameter to the presence of a secondary pulse.
A technical discussion on how to use these quantities - as they are applied to CUORICINO data - as pulse shape indicators in Appendix C.

4.6 Amplitude reconstruction in pile-up events

When dealing with pile-up events, combining Wiener and Optimum filtering, we can try to extract as much information as possible from each event (i.e. from each recorded pulse in a single window): we can use the Wiener filter to guess the presence of secondary pulses and use the OF to measure their relative amplitudes. To perform this task we follow a simple least square technique.
Under the assumption that signals superpose in a linear way, we model $\{O_i\}$ with a function $\{A'_i\}$, which is a superposition of the N_p signals identified by the WF:

$$A'_i(\vec{a}) = \sum_{p=1}^{N_p} a_p \cdot A_i^p \tag{4.14}$$

where the index i identifies the i^{th} recorded time sample, while the index p refers to the p^{th} pile-up. The vector $\vec{a}$ is the vector of the unknown amplitudes $\{a_p\}$ that we want to infere. A_i^p is the output of the OF applied to the average pulse, shifted in order to be synchronized to the p^{th} pile-up, whose occurring time has been measured with the WF.
With this modelization, we can build up an error function for the actual output of the Optimum filter $O(t_i)$

$$\text{ERR}(\vec{a}) = \sum_{i=1}^{N} (O_i - A_i(\vec{a}))^2 \tag{4.15}$$

The idea is to find the vector $\vec{a}$ of amplitudes that minimizes this function: by solving the set of linear equations:

$$\left\{ \frac{\partial}{\partial a_p} \text{ERR}(\vec{a}) = 0 \right\} \tag{4.16}$$

Since we have N_p equation with N_p unknown quantities, an algebric solution always exists.

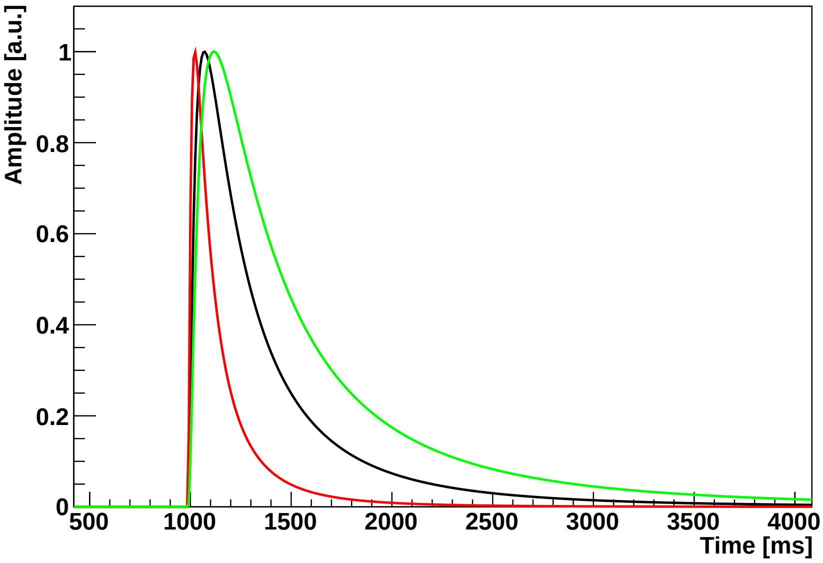


Figure 4.9: Three examples of lognormal function for different values of σ in equation (4.17). To be compared with a real CUORICINO pulse, shown in figure 4.2.

4.7 A study on simulated raw pulses

To check the many assumption and performances of the algorithms described in the previous sections, we performed a check on simulated raw pulses. This study has been carried out as a validation of the software I implemented: it shouldn't be considered as an efficiency test. The simulation is kept as much simple as possible: I simulated particle events superposing to a function similar to the detector response, a simulated noise time serie generated with the algorithm described in Appendix B.

We used, for the reference pulse an analytic function, the "lognormal", defined as:

$$f(t) = \frac{A}{\sigma t}\exp\left(-\frac{(\log(t) - c)^2}{2\sigma^2}\right) \tag{4.17}$$

which shows a shape quite similar to bolometric signals, as can be seen in figure 4.9, where different shapes are superimposed.

Combining the noise simulation with this "analytical" pulse we can reproduce a good variety of signal shapes at different SNRs. As an example, figure 4.10 shows a simulated electronic spike (fast decay time) and an instability of the criostat (slow decay time) both superimposed to a simulated noise time serie.

The idea is to check the performances of all the algorithms described in this

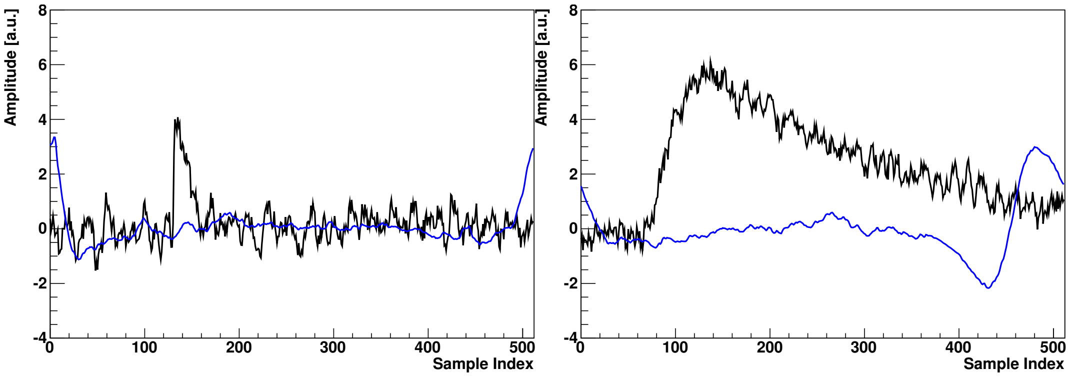


Figure 4.10: Simulated electronic spike (left) and a cryostat instability (right). The black curve is the raw signal, the blue one the OF Filter output.

chapter varying:

- the SNR;
- the shape of the acquired signals (by means of parameter σ in equation (4.17));
- relative amplitudes and positions in case of pile up;

Signal Amplitude. The first question that can arise from section 4.3 is: "is the OF amplitude evaluation affected by differences between the signal to be filtered and the reference pulse?"

In Cuoricino, particle signals show slow variations of their decay time (and rise time) with the energy (they are due to non-linearities of the detector response described in 2): tipical variations like these are of the order of nearly 1-2% within an energy range from 0 to 2.6 MeV. This condition violates one the hypothesis for the OF appliability (fixed response function) and it might have two important consequences that I'm going to discuss below: a variation of the amplitude evaluated for the primary pulse (i.e. the amplitude computed by the OF is different from the real one) and a variation of energy resolution (since the OF is not working in ideal conditions). The effects on amplitude are less relevant - in this case - since they can be accounted for by the amplitude to energy conversion, while quite relevant is a possible degradation of the energy resolution.

I firstly checked the OF output behaviour when the shape of the signals to be filtered are different from the average response function used in equation (4.3). Figure 4.11 shows the height of a noiseless pulse evaluated with the OF, in which the rise time of $\{S_i\}$ is mantained fixed, while the decay time of $\{R_i\}$

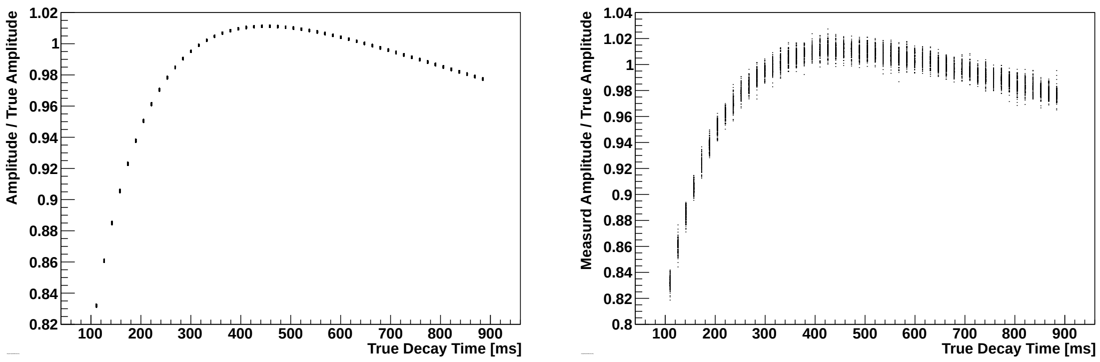


Figure 4.11: Ratio between the measured OF amplitude and the true amplitude on noiseless simulated raw pulses (left) and on events generated with a SNR equal to 10 (right): the amplitude varies as the decay time is increased.

is progressively increased. As a reference pulse, we used a signal with a decay time of nearly 320 ms. From this we can draw few conclusions:

- the amplitude is systematically underestimated on pulses with decay time smaller than the correct one;
- with decay times greater than the reference one, it is slowly overestimated until it reaches a maximum and then decreases slowly
- the variation in the amplitude is faster for decay times smaller than the reference one;
- a variation of a factor 9 in the decay time, is reflected in a variation of nearly 20% in the amplitude proving a good stability of the output of the OF with the signal shape: significant variations in the signal shape do not have a strong impact in the signal amplitude.

In order to check if this shape variation could be reflected in the resolution, we performed the same analysis after superimposing a noise component to the simulated events. The simulations has been carried out by generating 500 events for each different decay time. The right panel of figure 4.11 shows the result of this test.
The vertical spread associated to the different decay times can be considered as an indicator of the resolution for that specific decay time. The most relevant consideration we can extract from this plot is that the trend doesn't change and that the resolution is similar at each different decay time.

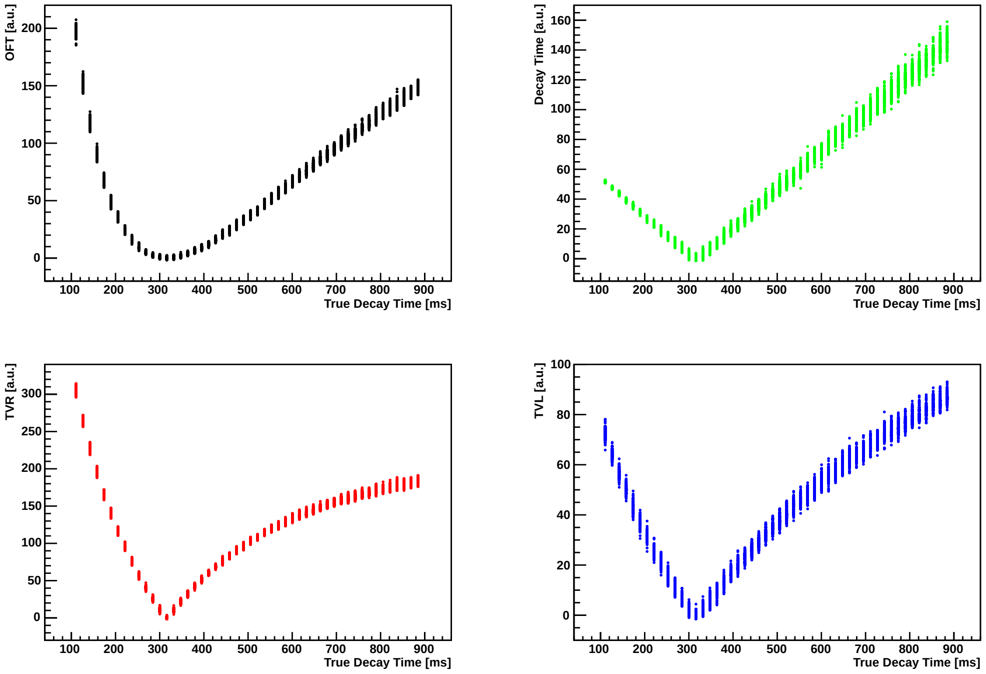


Figure 4.12: Performances of the main pulse shape parameter on the discrimination of different decay time. Each simulation is carried out with a SNR of 100. Each parameter is normalized so that its mean and variance at the correct decay time are, respectively 0 and 1. Top left: OFT; Top Right: Decay Time; bottom left: TVR; bottom right: TVL

We can conclude that, if we use a response function that is different from the average behaviour of the signals (in the ranges tested in the simulations), only a bias is introduced in the amplitude evaluation . This bias could be absorbed in the calibration procedure for the same run. What would really matter would be a change of the detector response during the same measurement, which couldn't be reabsorbed in the same way.
Compared to the variation we have tested with the simulation, we can state that, within the variation observed on CUORICINO data, the OF theory is perfectly appliable.

Shape Discrimination. I used the simulations also to validate the algorithm I implemented for the shape indicators. The Monte Carlo pulses can be used also to check parameters' performances on the recognition of different decay times, when the SNR changes. As a measure of good performance, I will use the minumum decay time difference that can be resolved (at least at one sigma) using the corresponding shape parameter distributions.

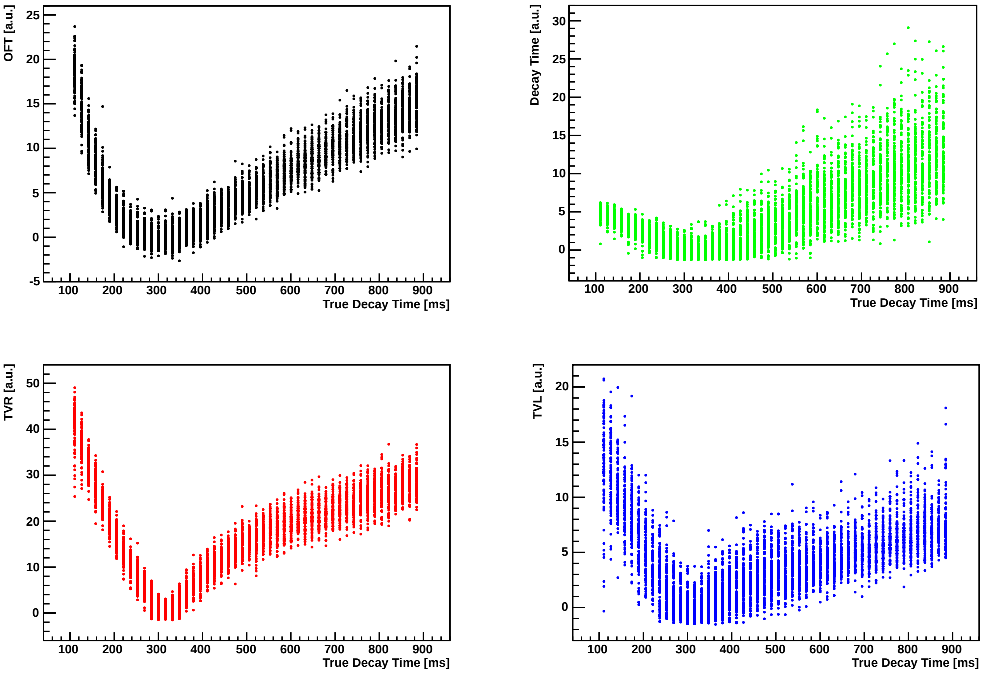


Figure 4.13: Performances of the main pulse shape parameter on the discrimination of different decay time. Each simulation is carried out with a SNR of 10. Each parameter is normalized so that its mean and variance at the correct decay time are, respectively 0 and 1. Top left: OFT; Top Right: Decay Time; bottom left: TVR; bottom right: TVL

Figure 4.12 shows the behaviour of TVR, TVL, OFT and the Decay Time (expressed as absolute difference from the true value) on a simulations generated with a SNR of 100 (since the average RMS of Cuoricino crystal peaks is nearly 3 keV, this should correspond, basically, to a 600 keV signal). Here all the parameters have a good resolving power, and TVR seems to be the most sensitive parameter even to slight variation of decay time (the curvature is high near the minimum).

Near the lowest Cuoricino energy threshold (the lowest one is 30 keV) the SNR would be nearly 10. Figure 4.13 shows the perfomances of the pulse shape parameter in this conditions. Also here, the one with the lowest RMS and best resolving power remains TVR, followed by OFT.

This software validation proved that the combined use of these shape parameters seems to be a powerful tool to recognize even small variations of decay time near the energy thereshold. However, the relative decay-time distance of the spurious signals from the reference pulse is never known a priori: this is the reason why we can't put an arbitrary energy threshold. The fine tun-

Table 4.2: Fractions of false pile-up identification on 1000 simulations for different SNRs and WF configurations.

λ	SNR=100	SNR = 50	SNR = 20	SNR = 10
2	0%	0%	5%	10%
3	0%	0%	0%	2%
4	0%	0%	0%	0%

ing required to find the minimum SNR needed for the discrimination is again mandatory. For example, in a measurement with similar noise and average pulse to the ones used in these simulations, if the electronic spikes have a decay time of nearly 200 ms, a SNR of 10 would not be sufficient to distinguish them by particle signals, thus the only way to discriminate them would be to increase the energy threshold.

Pile-up. To get the situation more complicated, I added the contribution of a pile-up. This time two parameters are changed: keeping their shape fixed, I summed two lognormal curves varying their relative distance and amplitude. The main problem introduced by the pile-up, is that, if we deal with filtered (and thus periodic) signals, a pile-up can affect the main pulse even if it happens $after$ it: obviously it is a computational effect introduced by the passage to the frequency domain when the OF is applied. The amplitude of the secondary pulse instead is $physically$ altered by the main pulse because the secondary pulse occurs during an instability of the detector's temperature caused by the primary event.

First of all, we perfomed a validation of the algorithm defined in section 4.4 for the identification of secondary pulses. The first check has been a measure of the probability of rejecting a good pulse, at different SNR. Starting with a signal-to-noise ratio of 100 and λ=2 (defined in section 4.4), the signal amplitude is lowered until the signal-to-noise ratio of Cuoricino threshold is reached. 1000 time series has been generated for different sets of SNR values: the results are shown in table 4.2. As can be seen, near the threshold we have a 10% of "probability" of rejecting a true signal as a pile-up. Increasing λ we can obtain better results but, as we addressed before, this algorithm is too sensitive to small fluctuations and its fine-tuning can become difficult if the noise changes with time. Anyhow, its resolving power can be seen as we switch on pile-ups: at a fixed SNR we superimposed to the main pulse a secondary event whose magnitude and distance to the main pulse are progressively increased. In this way a simple test of the efficiency of the algorithm can be appreciated if we

look at the reconstructed energies evaluated with the procedure described in 4.6. Figure 4.14 shows the amplitude of the primary pulse (expressed as SNR), as a funtion of the relative distance between the two pulses. When they are perfectly synchronized (zero delay) the pulse will be recognized as a sigle event. Without the WF, the amplitude is generally incorrect untill the relative distance is not sufficiently high. As the pile-up begins to reach the end of the acquired window the amplitude suffers an underestimation due to the leakage because the shape of the pile-up is truncated.
If we switch on the Wiener Filter, a reasonable range of delays (in this case, for example, between nearly 400 and 2000 ms) could still be accepted for a measure of the main pulse energy. The bottom panel of figure 4.14 shows in fact the reconstructed primary-pulse amplitude as a function of the delay of the secondary pulse . Within the mentioned delay range, the amplitude is correctly reconstructed but the resolution is worsened of less than a factor of 1.3-1.5.

At the time the CUORICINO result was finalized, the rule was to discard all the pile-ups (it was the safer approach, since this validation had never been carried out): this is a small fraction of the events since the rate on a single crystal is so small that the probability of a pile-up in a 4s time window is not so relevant for the analysis. Anyhow, in the future, in situation of a bolometric analysis with higher rate, this algorithm can become a powerful tool to gain live time of the detector by keeping as much information as possible also from the events that would be rejected: even if their amplitude is not perfectly reconstructed, they could be, for example, be used for coincidences.
To check if the secondary pulse is actually the one that has been superimposed (and not just a random noise fluctuation), I checked two quantities: the deviations of the measured amplitudes and positions of the pile-up from their true values (which are recorded during the simulation). Figure 4.15 shows these quantities as functions of the pile-up delay from the reference pulse: above delays comparable to the rise time the secondary pulse is always recognized.

Pile-ups can be recognized also by means of shape indicators.
Figure 4.16 shows the behaviour of all the shape indicators in case of a pile-up between two pulses with amplitude 100 and 10 SNR, expressed as a function of the relative delay of the two events. Note that events with delay greater that 3 seconds do not contain pile-ups and are reported to compare the average behaviour of the pulse shape parameter in absence of a secondary pulse. As expected, TVL and TVR are indeed sensitive untill the two pulses are close, while losing the information for distant pile-ups. The most sensitive indicator is OFT, showing a high sensitivity to the presence of pile-up even at higher delays.

4.8 Pulse shape analysis for CUORICINO

As already stated in chapter 3, the rejection of spurious events is a critical issue of CUORICINO analysis, in which we classify events through confidence intervals determined for each recorded pulse by the distributions of the shape parameters such as the ones described in this chapter.
Signals falling within these intervals are defined as ″true″ (or physical) pulses, while signals having one or more of their parameters outside the relevant intervals are rejected as noise or pileups.
The main problem is that the evaluation of this confidence intervals in the context of an experiment like CUORICINO had been indeed difficult.
If the detector′s response function wouldn′t be energy-dependent, the trend of any pulse shape parameter as a function of the energy should be an horizontal line.
Figure 4.17, where the distribution of TVR and OFT for a CUORICINO run are shown, gives an idea of how reality is far from being linear. In this situations, it is evident that it is not possible to apply simple box-shaped cuts in the parameters′ space, otherwise we would obtain an energy dependent efficiency: it becomes necessary to control the trend of any quantity used in pulse shape analysis before applying any criteria of rejection. One of the main goal of my thesis work has been focused on the linearization procedure briefly described in chapter 3 (and in Appendix C for a more description).
This non-linear trends are surely due to two main factors: a computational one and physical one. First of all, I have shown that basically every parameters evaluation algorithm contains an estimate of the amplitue itself, so every shape indicator and energy are correlated by construction. On the other hand, it has been shown by a recent work on the thermal model of bolometric pulses that the time constants of bolometric pulses contain an energy dependent term determined by the thermistor′s resistance [41]. Since every detector had its own working point, its thermal conductances, the main problem with CUORICINO data is the strong differences in the shape parameters distributions among the detectors. CUORICINO is composed of 52 working channels, whose measurements are subdivided in 32 data-sets: this means that a control on over 1760 different scatter-plots needs to be carried out as much automately as possible.
This has lead to the choice of a compromise between the sensitivity of a parameter and its stabilitity during CUORICINO data-taking.
In the previous section we have shown that the most sensitive parameter to shape variation is TVR. In the CUORICINO case this translates in parameter whose distributions are difficult to be linearized since it is indeed to much sensitive even to small fluctuations of the signal′s shape (figure 4.17 is an example of the differences between TVR and OFT, for the same channel).

OFT has shown, on average, a smoother behaviour, and seemed to offer this kind of compromize. Thus, it has been chosen as the main parameter for pulse shape analysis of CUORICINO, using also the rise and decay time (whose scatter plots can be easiliy interpolated with a polynomial function) for a better resolving power. The efficiency of this technique, shown in table 3.4 proved this to be a reasonable choice: the probability of discarding a good event was nearly 2%, while the rejecting nearly 36% of the background level in the region of interest for the 0ν-DBD .

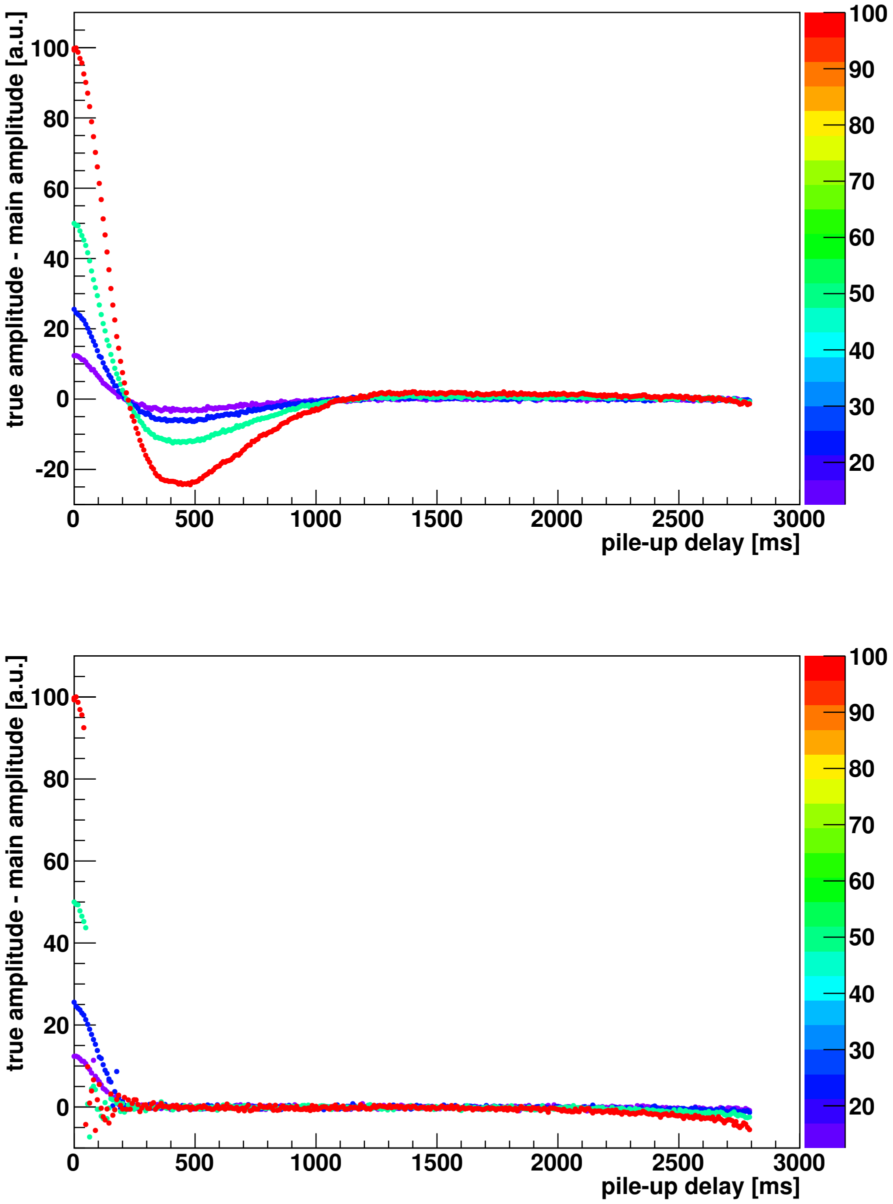


Figure 4.14: Top panel: amplitude of the primary pulse without the amplitude correction technique as a function of the distance to the pile-up maximum. Bottom panel: the same plot is shown for the reconstructed amplitude. Different colors represents the different amplitudes of the secondary pulse - expressed as signal-to-noise ratio

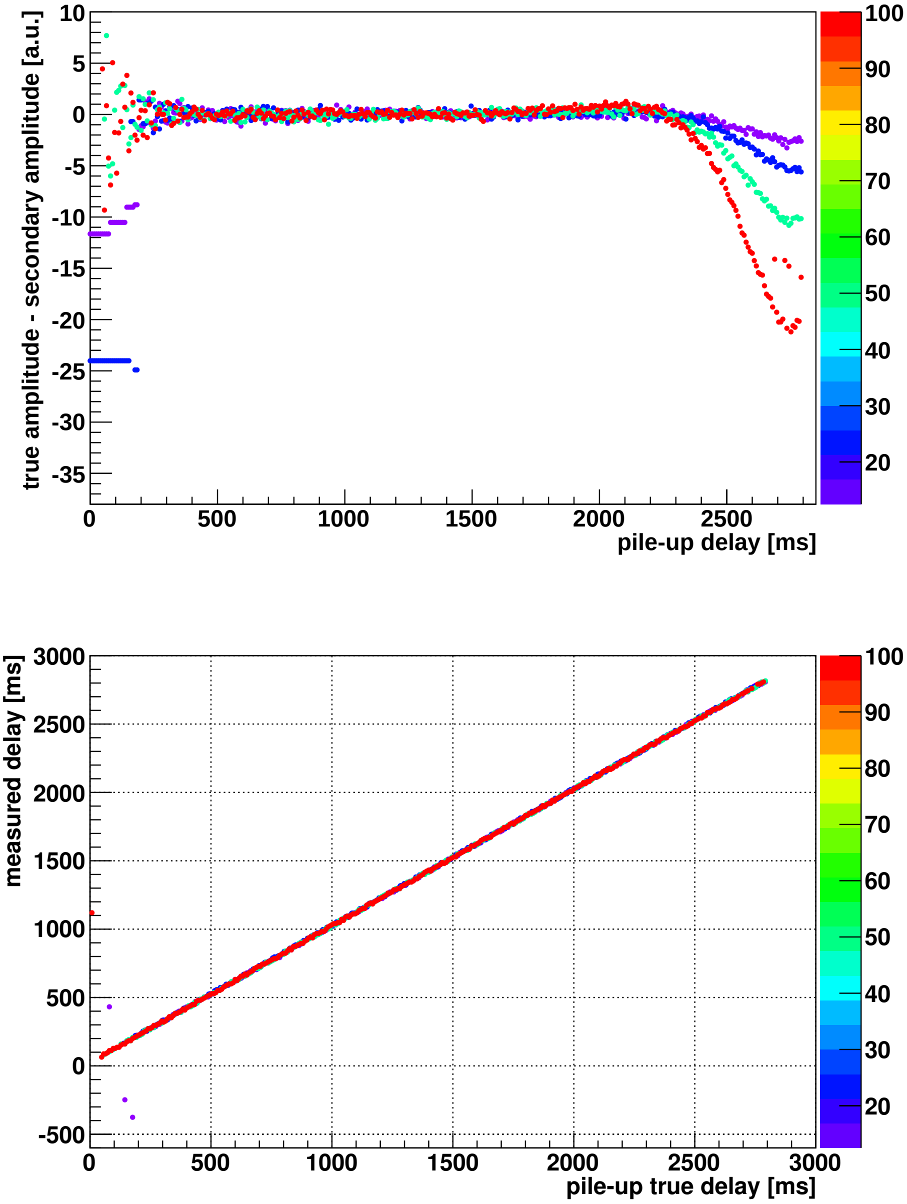


Figure 4.15: Top panel: reconstructed amplitude of the secondary pulse as a function of its distance to the main pulse. Bottom panel: measured delay of the secondary pulse as a function of its the true delay known from the Monte Carlo truth. Different colors represents the different amplitudes of the secondary pulse - expressed as signal-to-noise ratio

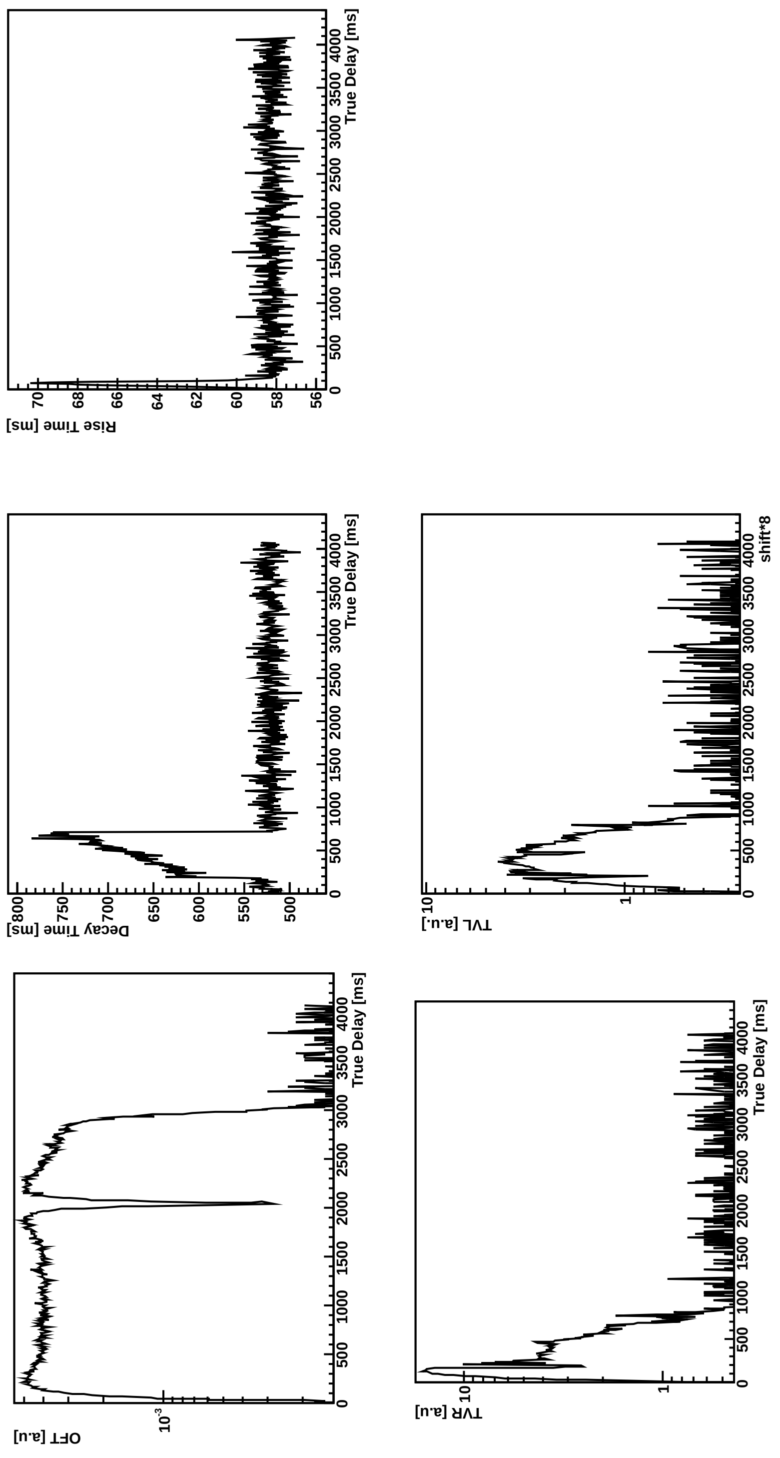


Figure 4.16: Schematic representation of the resolving power of the different pulse shape estimators for pile-up rejection. Each parameter value is plotted as a function of the relative distance between the main and the secondary pulse. Events with delay greater that 3 s contain no pile-up and are reported for reference of the average behaviour of the shape parameter in absence of pile-up.

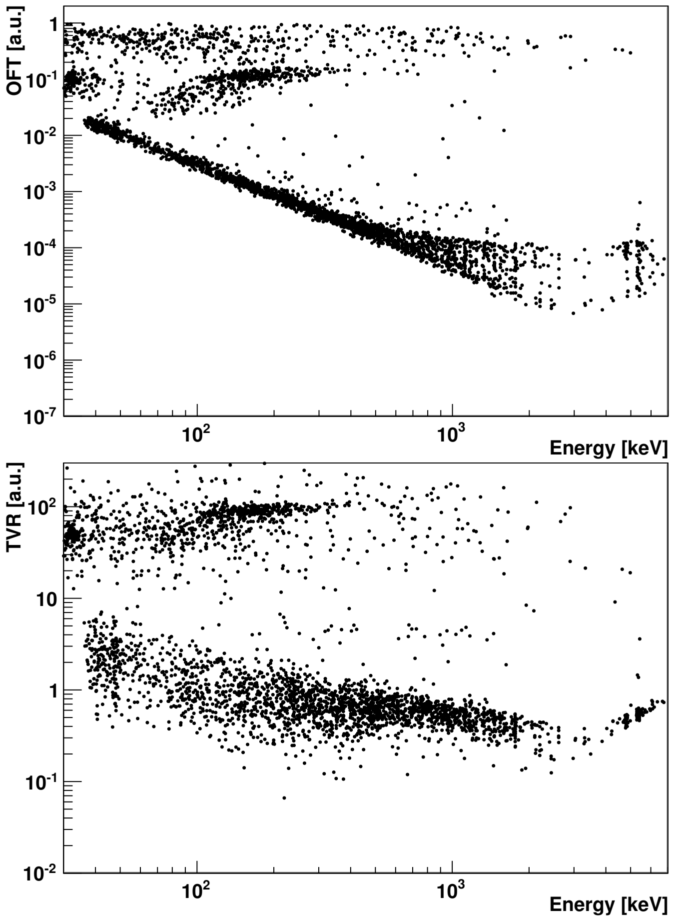


Figure 4.17: Examples of two CUORICINO scatter-plots of TVR (bottom) and OFT (top). The main trends identify particle signals. The dense distribution on top are electronic spikes. Even if the spikes' distribution is more resolved by TVR, OFT shows a trend which is more easy to be linearized.

Chapter 5

Cuorino Analysis For Neutrinoless Double Beta Decay

This chapter presents the main steps followed for the analysis of CUORICINO data for the evaluation of Neutrinoless Dobule Beta Decay. This is a documentation of the stuties perfomed on RUN II which was analyzed from scratch with the new software that will be used for the CUORE, whose prototype is CUORICINO itself. As described in chapter 3, CUORICINO limit has been evaluated by means of a bayesian approach used to study the spectrum in the region of interest: in the first part of this chapter instead, there will be a description of the studies performed for the optimization the data, the model and algorithms used for such analysis. Finally, the comparison between the Monte Carlo approaches is discussed in detail.

5.1 Analysis framework and data

The entire RUN II data of Cuoricino has been used to test the software frame work developed by the CUORE collaboration: DIANA. This software is based partially on algorithms already used for Cuoricino that were transported from the old environment (FORTRAN based) to the new one (DIANA) with improve ments and revisions, and partially on new algorithms introduced mainly to be able to authomatize most of the data processing steps. DIANA, written entirely in C^{++}, is designed to deal with a huge amount of data as much automately as possible. Each operation of the data chain that leads to the evaluation of the energy of a single event belongs to different "modules", C^{++}class structures inheriting from a common base class. Each module is configured at the start-up of the analysis so that the data flow is mainly automated. The collection of the DSP algorithms described in chapter 4 is an example of such modules, as the methods for the energy calibration and gain instability corrections already mentioned in chapter 3. Many of these techniques have been translated from the old FORTRAN code with improvements and automation tools. The reprocessing of CUORICINO RUN II data (besides the change in the code) consisted in a more uniform treatment of data (such as an identical denition of the Optimum Filter parameters, the use of an automatic or semiautomatic procedures for calibrations and stabilization) and in the application of an improved algorithm for the Pulse Shape Analysis and stabilization. For this reasons Cuoricino final result have been evaluated using - for RUN II - the new reprocessed data.

5.2 Fit method

The numerical method used for the limit evaluation is the same applied for the computation of the results published in [46]. The same code (TIMEZZO), based on MINUIT , is here used. For a more detailed description of TIMEZZO , we refer to Appendix A.
The limit for the 0ν-DBD half life of ^{130}Te is evaluated by means of a Bayesian approach.
The three anticoincidence sum spectra of Big, Small and Enriched crystals (identified with the index k) are modeled with the following set of equations:

$$f_k(E) = B_k + \Gamma_k^{^{60}Co} g_k\left(E - E_{Co}\right) + \Gamma^{0\nu} g_k\left(E - E_{0\nu}\right) \tag{5.1}$$

where E_i is the energy associated to the i^{th} bin, $g_k(E)$ is the function describing the shape of the peak and $E_{0\nu}$ is the Q-value of the ^{130}Te DBD transition. The free parameters of this model are:

- B_k [c/kg/y], the three flat background rates;
- $\Gamma_k^{^{60}Co}$ [c/kg/y], the three ^{60}Co activities;
- $\Gamma^{0\nu}$ [c/kg/y], the absolute 0νDBD rate;
- E_{Co} [keV], the Q-value for the ^{60}Co line.

Assuming Poisson statistics for the binned data, and scaling each f_k by its exposure and efficiency in order to obtain $y_{i,k}$ (i.e. the number of counts predicted by the model for the i^{th} bin in the k^{th} spectrum), the overall CUORICINO likelihood will be:

$$\mathcal{L}(\overline{n}|\overline{y}) = \prod_{i,k} \exp\left(-y_{i,k}\right) \cdot \frac{y_{i,k}^{n_{i,k}}}{n_{i,k}!} \tag{5.2}$$

where $n_{i,k}$ is the number of counts observed in the i^{th} bin in the k^{th} spectrum (in this notation, $\overline{n}$ and $\overline{y}$ indentify the vectors of the measured and expected number of counts, respectively). Instead of maximizing $\mathcal{L}$, we follow the prescription of a likelihood chi-square analysis proposed by Baker and Cousins [47], which gives the possibility of constructing a general χ^2 statistics directly from the likelihood function. A chi-square statis tics is any function of the data and of the parameters which is asymptotically distributed in the classical chi-square distribution. The desired function comes from a likelihood ratio :

$$\lambda = \frac{\mathcal{L}(\overline{n}|\overline{y})}{\mathcal{L}(\overline{n}|\overline{n})} \tag{5.3}$$

which is the ratio between the likelihood of obtaining the collection of all the counts $\overline{n}$, given a general vector of expected counts y, and the model-independent likelihood for which the best estimate of each $y_{i,k}$ is simply $n_{i,k}$ (its a generalization of the one-bin case where, if n counts are observed in a time T, then the maximum likelihood estimator for the expected rate y is just n/T).

At this point we can compute our χ^2 as:

$$\chi^2_L = -2 \cdot \log \lambda = 2 \cdot \sum_i \sum_j (y_{i,k} - n_{i,k} + n_{i,k} \log(n_{i,k}/y_{i,k})) \tag{5.4}$$

This function has the advantage that it can be used both for goodness of fit tests as for computation of confidence intervals.

In order to obtain the 90% C.L. limit value on $\Gamma^{0\nu}$, we express the likelihood as a function of the free parameter of the model of equation (5.1), we exclude the unphysical region from the likelihood function with a flat prior (basically we won't allow negative rates to be measured) and then compute the confidence interval with TIMEZZO as described in Appendix A.

5.3 Hypotheses behind the fit

5.3.1 Background modelization

The definition of the energy window used to fit the DBD spectrum, the hypothesis assumed for the background shape and the number of free parameters used to describe the background itself are important for the choice of the analysis procedure and for the determination of its systematic.

The choice of the energy window is somehow arbitrary, but it influences the background representation. If the energy window is too wide (compared to the

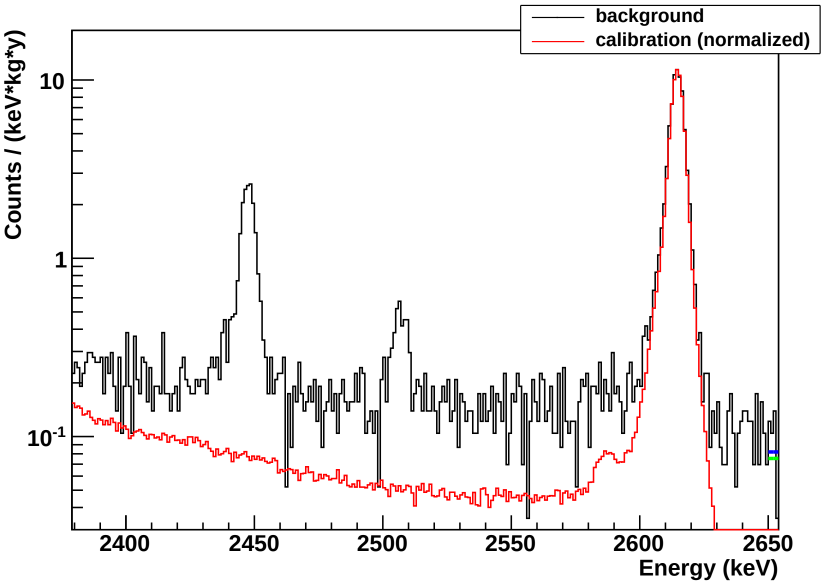


Figure 5.1: Background summed spectrum superimposed to the normalized calibration sum spectrum. The same pulse shape cuts are applied on both cases.

FWHM) a physical model able to describe the background shape is mandatory (but today we don't have a sufficiently reliable model). Obviously there is also a minimum size of the energy window necessary to be able to evaluate the background level beyond the 0ν-DBD peak. Moreover, in our case, there is a peak very close to the $\beta\beta$ Q-value, at 2505 keV due to ^{60}Co, which has therefore to be included in the window. The final choice for the fit window is 2474-2580 keV, that is the wider window, centered on the $\beta\beta$ Q-value, that allows to exclude the following background peaks: the 2448 keV line of ^{214}Bi and the 2587 Te X-ray escape peak of the ^{208}Tl line
The latter peak, not clearly visible in the background spectrum, can be appreciated in figure 5.1 that shows the CUORICINO spectra for the Big crystals compared to a calibration spectra whose rate is normalized on the ^{208}Tl line.

As already stated in chapter 3, once defined the energy window, it is nec essary to fix a technique for background evaluation. There are three possible general approaches:

- *ideal*: treat separately the different detectors;
- *multi-spectrum*: assume the background identical for groups of detectors;
- *average-spectrum*: average the background over the detectors belonging to

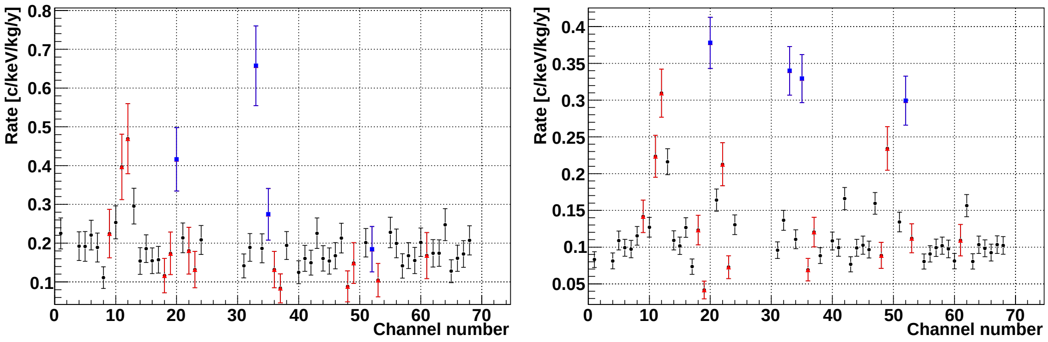


Figure 5.2: Left panel: counting rate [c/keV/kg/y] in the energy region between 2490 keV and 2590 keV (including therefore the 2505 keV peak due to ^{60}Co) vs. channel number. Right panel: counting rate [c/keV/kg/y] in the energy region between 2700 keV and 3200 keV vs. channel number. Red points refer to 330 g natural crystals, blue points to enriched ones. Rates are computed on the same data used for the DBD analysis (PSA and anticoincidence cuts included). Colors refer to the crystal type: Big (black), Small (red) and Enriched (blue).

the same group;

where the three groups of detectors are, as already stated, those having the same efficiency (i.e. mass) and isotopic abundance: Big, Small and Enriched crystals.

The *ideal* approach would introduce a high number of free parameters since in our analysis we have 55 detectors. Moreover it would require to have a good description, detector by detector, of the energy resolution. This is possible if we assume that the energy resolution is well represented by the estimation obtained on the calibration runs. With this approach, each channel and each data-set is treated separately, fitting each channel spectrum with its background model and using the energy resolution measured during the calibration for each data-set. In principle this is the best approach since it uses all the information available and does not imply any hypothesis on the uniformity of the background rate among detectors. Two weak points can be devised in this approach: the number of free parameters (55 background coefficients and 55 ^{60}Co rates), and the assumption that the energy resolution measured during a background data-set (∼1 month) is equal to the energy resolution measured in a 6 days calibration (∼3 days before the start of the background run and ∼3 days after).

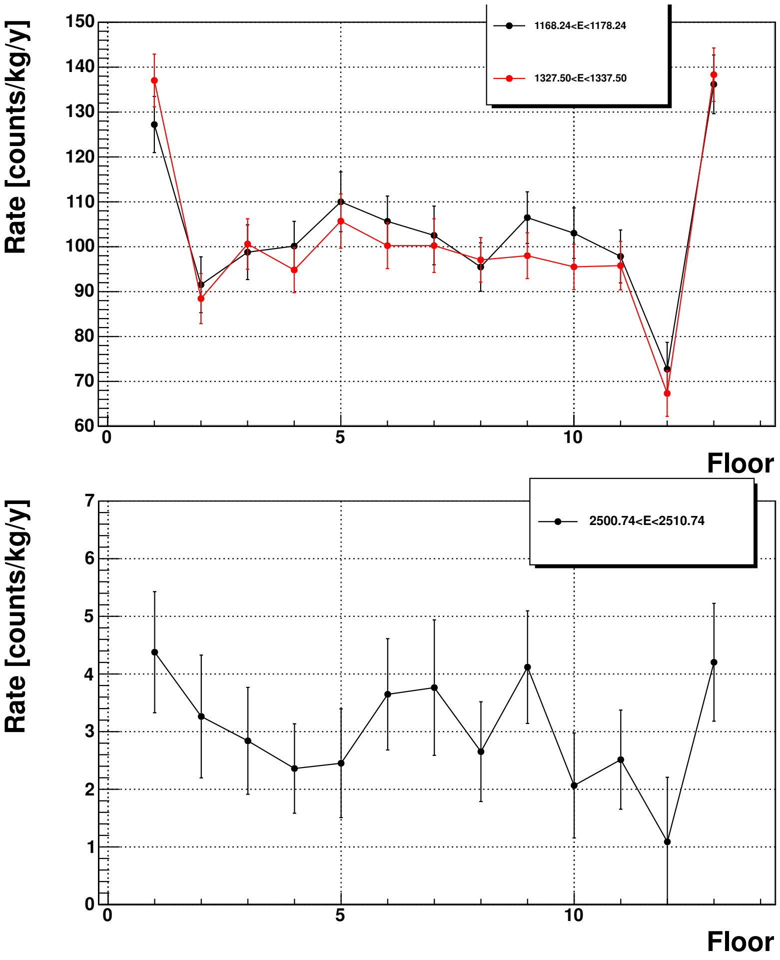


Figure 5.3: Counting rate [c/kg/y] vs. floor on the three ^{60}Co lines: 1173 and 1332 keV in the upper panel and the sum line at 2505 keV in the lower panel.

The *multi-spectrum* procedure avoids the high number of free parameters in the fit by forcing the background to be identical on the detectors of the same group: this assumption is motivated by the fact that the low statistics prevents us from being sensible to background variations among crystals of the same group in the DBD region (see left panel of fig. 5.2). Here too the energy resolution is evaluated from calibration data.

The *average-spectrum* approach avoids any strong assumption, at the price of a certain degree of information loss. Indeed, the energy resolution here can be determined on the background spectrum itself (see section 5.3.2) and no hypothesis on the background differences among detectors is used. The back-

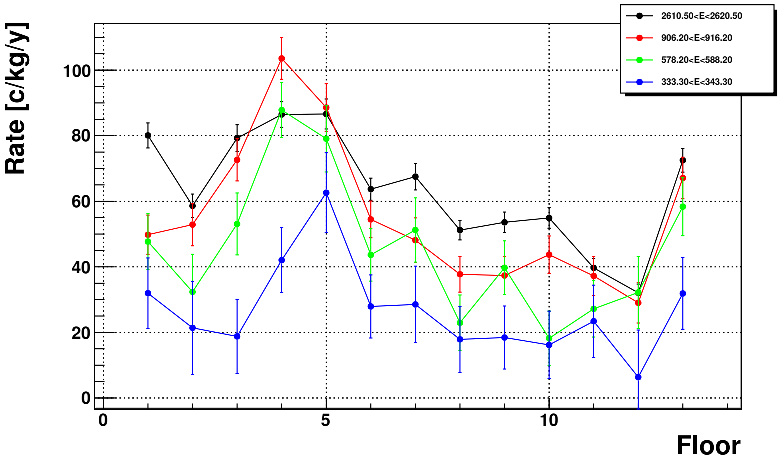


Figure 5.4: Counting rate [c/kg/y] vs. floor on the gamma lines of ^{232}Th: 338, 583, 911 and 2615 keV.

ground is simply averaged on all the statistics and all the similar crystals (those of the three mentioned groups). Also a variation of the background rate during time is here not relevant, provided that the response function does not change. This last approach is the one used for the study documented in this thesis, as a continuation of the standard CUORICINO analysis.

Anyhow, it should be stated that the hypothesis of a uniform counting rate on detectors of the same group is not supported by our present knowledge on DBD background composition but we expect that the extremely low statistics on the single channel spectra will prevent us from appreciating such differences in the ROI. For example, the 2505 keV peak has a too low intensity to be able to compare its rate among different detectors. However this peak is produced by the contemporary detection of the two gamma photons emitted in ^{60}Co decay. These lines (at 1173 and 1332 keV) have a much higher intensity and - as shown in figure 5.3 - their rate is dependent on the crystal position in the tower. ^{60}Co is a short-living isotope produced by neutron activation in Copper. The 1173 and 1332 keV lines reduce their intensity with the ^{60}Co half-life, showing that activation is not due to underground neutrons and that the contamination must have been introduced with the new Copper used for Cuoricino. We can therefore identify the source of ^{60}Co in the Copper of the mounting structure of the Cuoricino tower. This was constructed with two different kind of Copper and each piece of the tower (top, bottom, frames and the lateral coverages) was exposed to sea level neutrons for a different time. This difference in the exposure, and the variation of geometrical efficiency can account for the behavior observed in figure 5.3. Excluding the ^{60}Co, we have not identified any other contribution to the DBD background that could be time dependent. Indeed, ^{241}Bi lines (which have had very important variations during the experiment) have a negligible intensity above the 2448 keV peak. The few cosmogenic isotopes identified in the spectrum have Q-values far below the $\beta\beta$ transition energy and the ^{210}Po peak - which is observed to decay in time - is ascribed to an internal contamination of the crystals and therefore should not give any contribution in the region of interest. The approximately flat background under the DBD peak is most likely due to three different sources. In the LNGS Annual Report of year 2006 [48] is included a detailed description of the sources. They are:

- the ^{208}Tl that contributes through the Compton of the 2615 keV line (see figure 5.1 and 5.4);
- the surface alpha contaminations of the crystals;
- the surface alpha contamination of the crystal mounting structure (mainly Copper);

The first two sources have a clear signature and are quite dis-uniform in their intensity. The third is not clearly identified, but should account - together with the second source - for the continuum background measured in the 3-4 MeV region (see figure 5.2).

5.3.2 Response functions

The simplest model that can be assumed to represent the shape $g_k(E)$ of a peak centered at E_0 in the sum spectrum of the k^{th} crystal group is:

$$g_k(E) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(E-E_0)^2}{2\sigma^2}\right) \tag{5.5}$$

We will refer to this equation as *single-gaussian* response function.

The hypothesis subtended by this equation is that, although collecting events from different channels and data sets, the response function is still a simple gaussian. The resolution associated with this gaussian can be measured on k^{th} sum background spectrum itself, using a background line close enough to the 0ν-DBD Q-value (in our case the 2615 keV line of ^{208}Tl). The assumption that a single gaussian can represent correctly the response function of a collection of detectors could be considered an oversimplification of the problem, and - except for a simple χ^2 - there's no actual method to confirm the reliability of this approach.

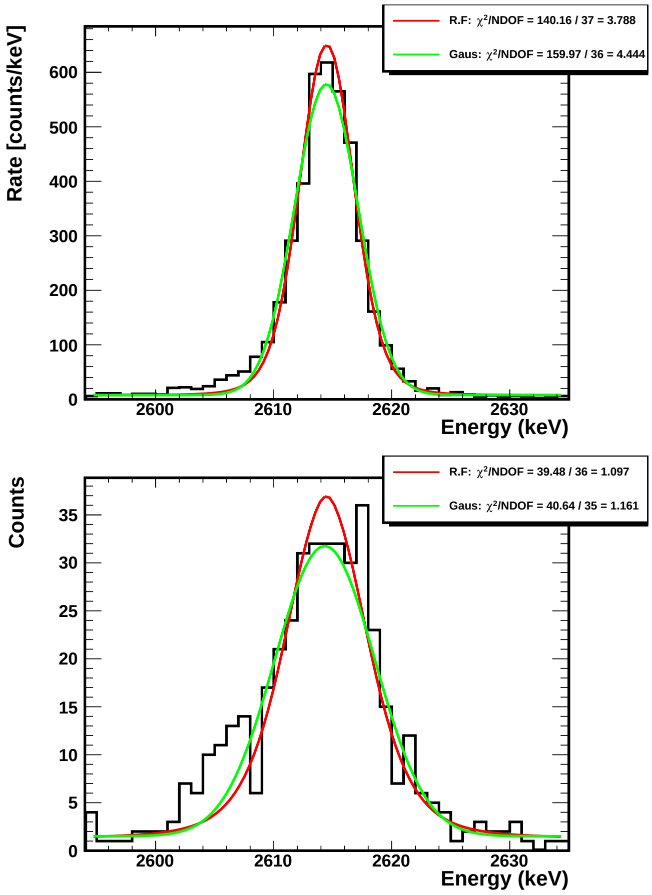


Figure 5.5: Fit of the 2615 keV line in the anticoincidence sum background spectrum (Big crystals on the top panel and Small crystals on the bottom one). The two response functions are compared: single gaussian (green) and sum of gaussians (red).

On the contrary, if we follow the idea that the resolution measured in a calibration is a good estimator for the resolution of the corresponding background data-set, then it is reasonable to believe that in the sum spectrum the response function can be modeled as a superposition of all the gaussians whose resolutions were measured in calibrations. This for all the channels/data-sets summed in the spectrum. With this assumption, the response function has the form that we already defined in chapter 3:

$$g_k(E) = \frac{1}{\sum\limits_{i,j} A_{i,j}} \sum_{i,j} \frac{A_{i,j}}{\sqrt{2\pi}\sigma_{i,j}} \exp\left(-\frac{(E-E_0)^2}{2\sigma_{i,j}^2}\right) \tag{5.6}$$

where the sum over i extends on all the detectors belonging to the k^{th} group and

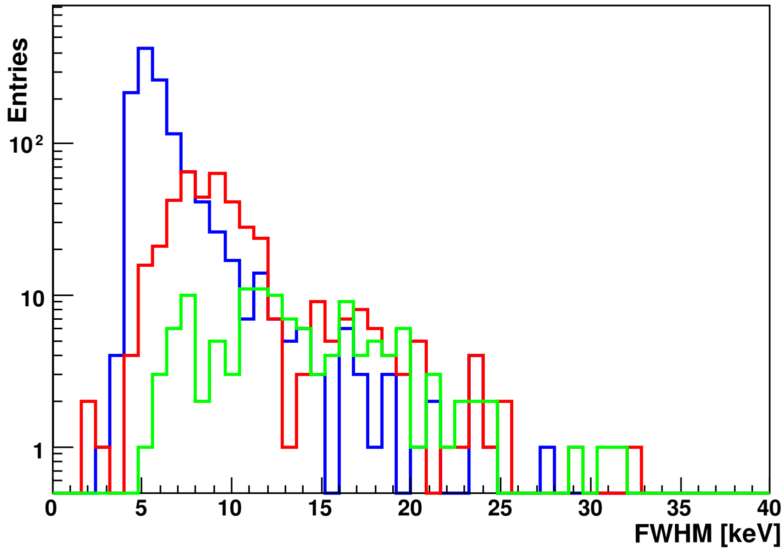


Figure 5.6: Distributions of the resolution measured in calibration for the three groups of crystals.

j runs over all the data-sets. $A_{i,j}$ and $\sigma_{i,j}$ are the corresponding background exposure and energy resolution measured during calibration.
Even in this case there is a simplification since calibrations run last for few days while background data-sets lasts for about 1 month. Equation (5.6) in principle provides a better description of all the possible differences among channel resolutions and their changes through time. While the first leads not only to averaging such differences but also measures the energy resolution exactly on the background spectrum that is being used for 0ν-DBD analysis. In section 5.4 we will show the impact of these two functons on the best fit values and on the 90% CL limits.
In order understand which of the two equation is more appropriate, we studied the performances of the two functions (on RUN II data) comparing the χ^2/ndf obtained when fitting the 2615 keV line in the Big and Small crystal background spectra (see figure 5.5). Practically, the two response functions have been constructed as described below.
The *single gaussian* response function described by equation (5.5) was obtained measuring the resolution at the 2615 keV line of ^{208}Tl for each of the three sum spectra on the right side of the peak: this is due to the presence of a tail on the left part which presently is not understood. The regression was performed with a gaussian peak superimposed to a linear background. These are the resulting FWHM's:

- Big crystals: 6.3 ± 0.1 keV

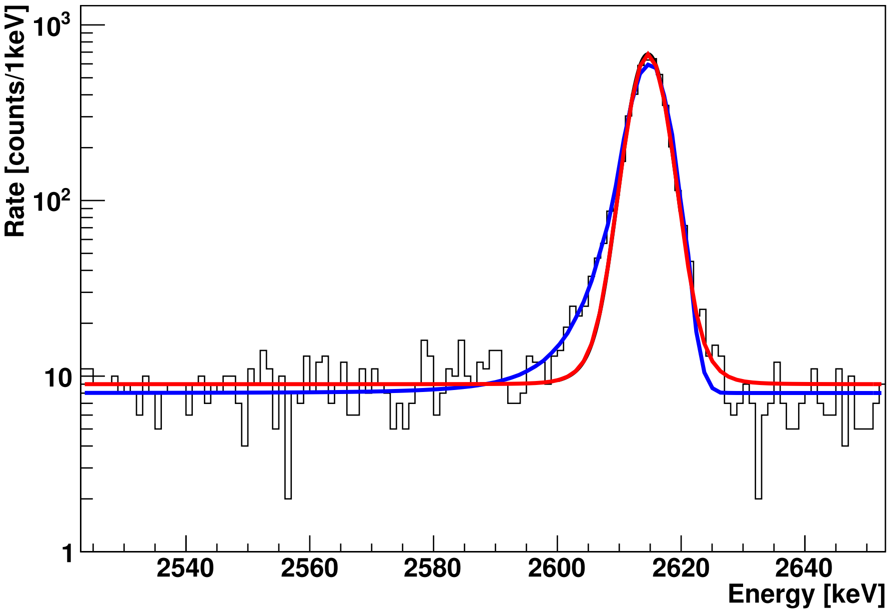


Figure 5.7: Comparison between the fits of the *multi-gaussian* response function (red) and the *crystal ball* function (blue).

- Small crystals: 8.9 ± 0.5 keV
- Enriched crystals: 10 ± 2.2 keV

The *sum of gaussians* response function described by equation (5.6) was obtained after a *binning* of the measured resolutions. Since the gaussians in equation (5.6) are more than a thousands for the Big crystals, almost four hundreds for the Small crystals and about 50 for the two Enriched crystals, we used the following averaging procedure for the construction of the response function:

- for each group of detectors we collected all the resolutions measured in calibrations into an histogram (see figure 5.6) with binning width $\Delta\sigma$;
- we assigned to the j^{th} bin a weight T_j which is the sum of the live-times of all the detectors whose resolutions are within $\sigma_j \pm \Delta\sigma/2$.
- the response function at energy E_0, for a group of crystal is then be:

$$g(E) = \frac{1}{\sum_j T_j} \sum_j \frac{T_j}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(E-E_0)^2}{2\sigma_j^2}\right) \tag{5.7}$$

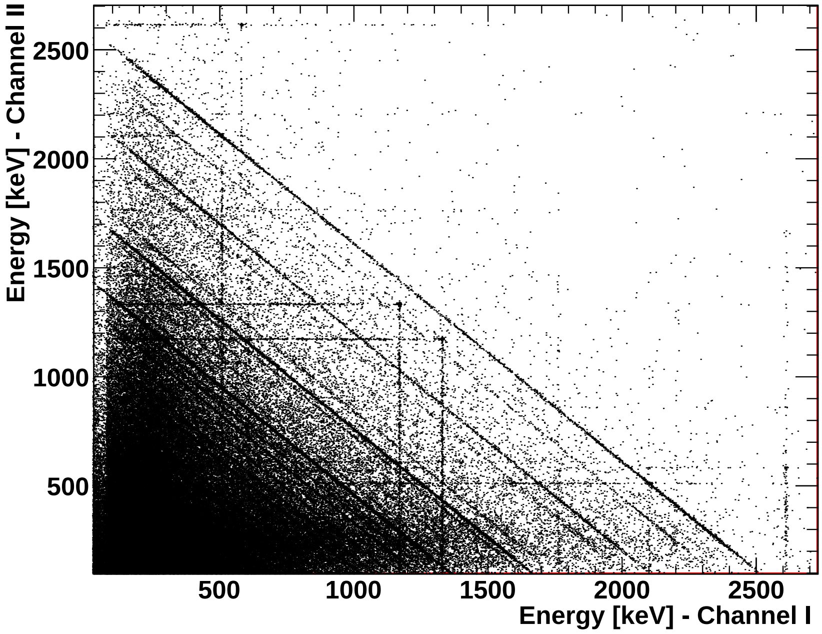


Figure 5.8: Cuoricino calibrations double-hit scatter plot.

Apparently there is no significant evidence for a better modeling of one of the two functions since the χ^2/ndf are similar. Anyhow, both the modelizations seem to describe poorly this peak: a low energy tail can be seen on the left side of the peak (possibily due to some sort of energy loss mechanism). The question that can arise is then whether this could be a characteristic of every monochromatic line in the Cuoricino spectrum. In fact, this energy tail is not present in the CUORICINO Monte Carlo simulation, thus it could be a side effect of some unknown energy degradation characterizing the bolometers (i.e. a quenching) and not accounted by the processes implemented in the simulations. A possible alternative, frequently used in particle physics, that can describe this peak is the so-called *crystal ball* function:

$$f(x,\alpha,n,\bar{x},\sigma)=\begin{cases}\exp\left[-\dfrac{(x-\bar{x})^2}{2\sigma^2}\right], \text{for } \dfrac{(x-\bar{x})}{\sigma}>\alpha\\ A\cdot\left(B-\dfrac{x-\bar{x}}{\sigma}\right)^{-n}, \text{for } \dfrac{(x-\bar{x})}{\sigma}\leq\alpha\end{cases} \tag{5.8}$$

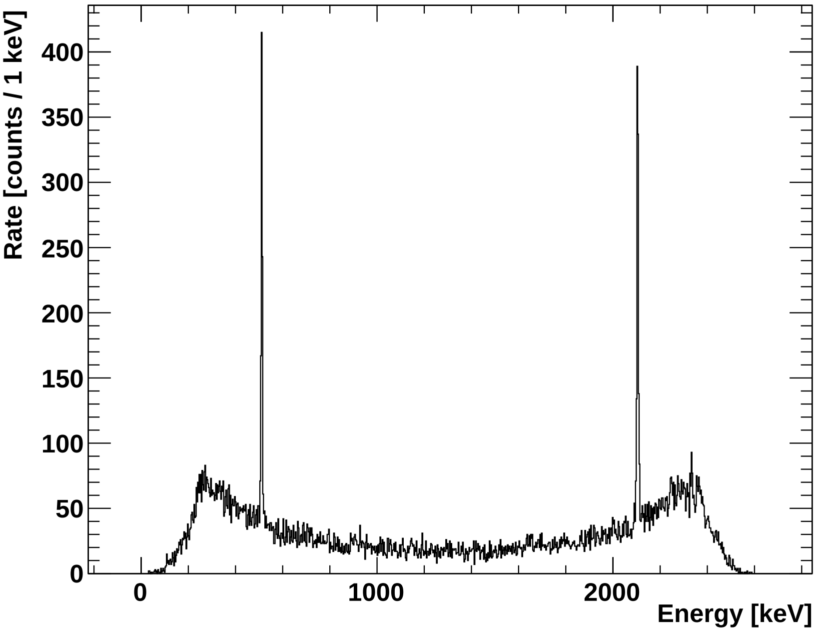


Figure 5.9: Cuoricino calibrations double-hit scatter plot.

where

$$
\begin{aligned}
A &= \left(\frac{n}{|\alpha|}\right)^n \cdot \left(-\frac{|\alpha|^2}{2}\right), \\
B &= \frac{n}{|\alpha|} - |\alpha|.
\end{aligned}
\tag{5.9}
$$

Although this function better describes the shape of the Tallium line, as can be seen in figure 5.7, in the CUORICINO spectrum we couldnt find any other peak with a comparable signal-to-background ratio to test this function. The question is: is this asymettry a characteristic of every monochromatic line?
We decided to use the coicidence information in order to find other peak with this characteristics. Figure 5.8 shows a scatter plot of the energy of all the double-hit events recognized in Cuoricino calibrations. Diagonal iso-energy lines identify events in which the whole energy of a gamma decay is divided in two crystals via Compton scattering. We identified the double-hit events whose energies summed to 2615±10 keV and plotted their comulative spectrum, that can be appreciated in figure 5.9. As expected the main trend is the Compton shape plus the contribution of two peaks due to the single escape of a 511 keV gamma after pair production (one crystal will record the energy of the two electron plus one of the annhilation photons - 2104 keV -, while the other

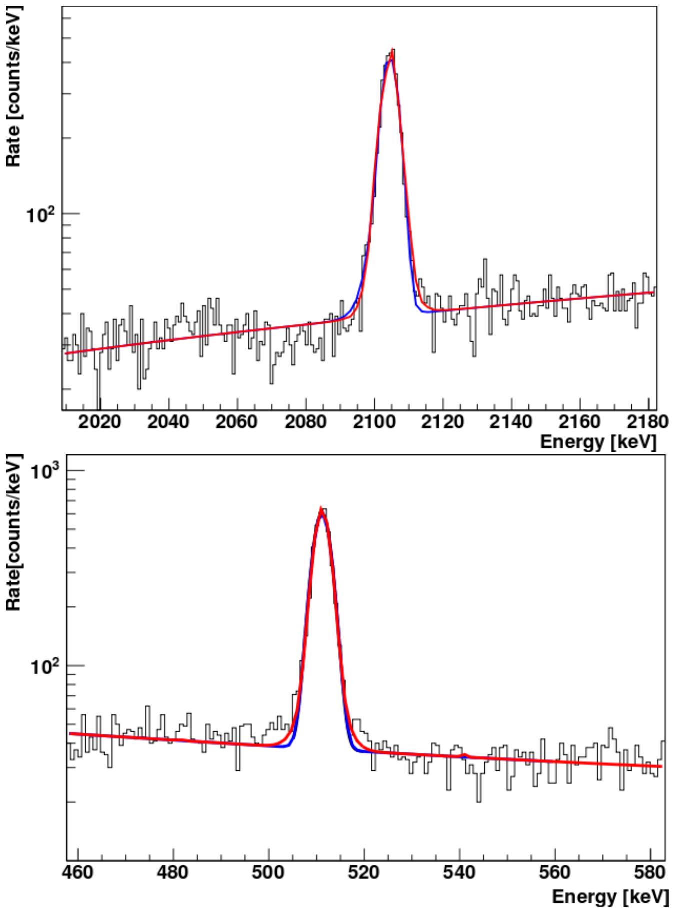


Figure 5.10: Different response function evaluated on the 511 keV (top) and 2104 keV (bottom) energy lines.

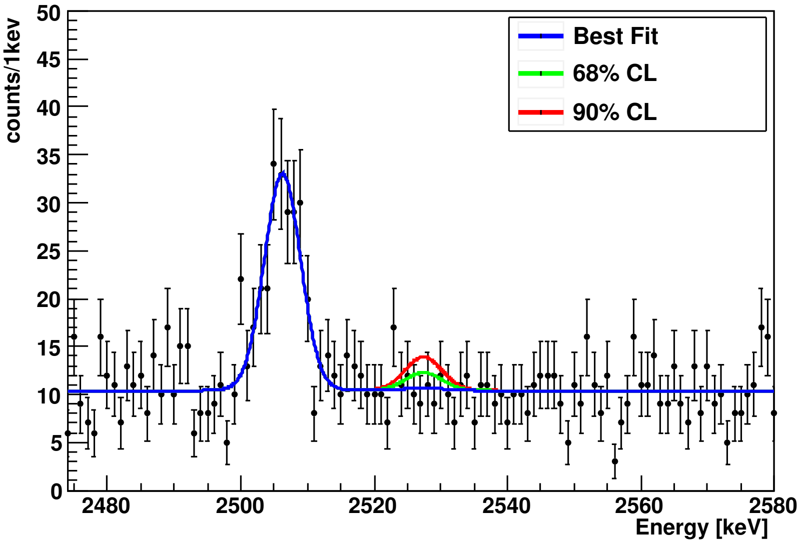
crystal stops the escaped gamma). As can be seen, this two peaks have high statistics and a great signal-to-background ratio (nearly 1000) and can thus be used to check the performances of different response function. The different height of the two line is due to the resolution improvement at lower energy. Firstly we interpolated the background with a second order polynomial on an energy region far from the peak. Then, once the background shape is fixed, we fitted the peak with the three different functions and checked the result with a χ^2 test and a likelihood ratio test. Figure 5.10 shows no significant difference between the two: both chi-square and likelihood ratio are nearly one, thus no evidence for an asymmetry was found.

The asymmetric part seems to be most likely due to an inefficiency in the energy collection when compton effect is envolved; anyhow, any physical analysis of these processes will require dedicated measurements in orther to gain a better understanding. Until further evidence, we will describe the shape of the peak caused by the containement of the 0ν-DBD event with a gaussian function.

5.4 Best Fit Results

We have studied the differences in the 0ν-DBD best fits obtained with the two models for the detector's response function. This study has been performed on RUN II data and can be considered as a measure of the impact of changing the hypothesis that the resolution of a detector could be measured during calibrations.
The parameters used in the fit procedure are summarized in table 5.1 while tables 5.2 and 5.3 show a summary of the results.

Figure figure 5.11 shows the results obtained by fitting DBD peak with the *single-gaussian* response function (equation (5.5)) and the *sum-of-gaussian* one (equation (5.6)).
The results are computed following the *average-spectrum* approach. The best fit, the background function, the 68% and 90% confidence limits are superimposed on a fictitious spectrum constructed as the sum of the three spectra (those that are used in the fit) after a proper normalization that takes into account the different live times, efficiencies, mass and isotopic abun dances of the three groups of crystals.
The difference between the results obtained with the two functions seems to be not only in the 0ν-DBD rate best fit value, but also in its error. Consequently the limit changes. As it can be seen from the figures, the *single-gaussian* distribution is wider and this translates in a higher number of counts, in larger error and, therefore, in a weaker limit. Anyhow, the hypothesis that the resolution measured in calibration can be a considered as a measure also of the resolution of background runs is not ruled by this analysis, so we decided to choose the sum of gaussian as a description of the detector response and we used, in order to be conservative, the difference between the best fits obtained with the two functions (0.08×10^{-25} y $^{-1}$) as a measure for the a systematic error on $\Gamma^{0\nu}$.
As a next step, we added the contribution from Big, Small and Enriched crystals from RUN I, combining their likelihoods in equation with the RUN II data and using a similar reconstruction for the response function. The results are shown in table 5.5 while in figure 5.12 we show the best fit and corresponding 68% and 90% limits.



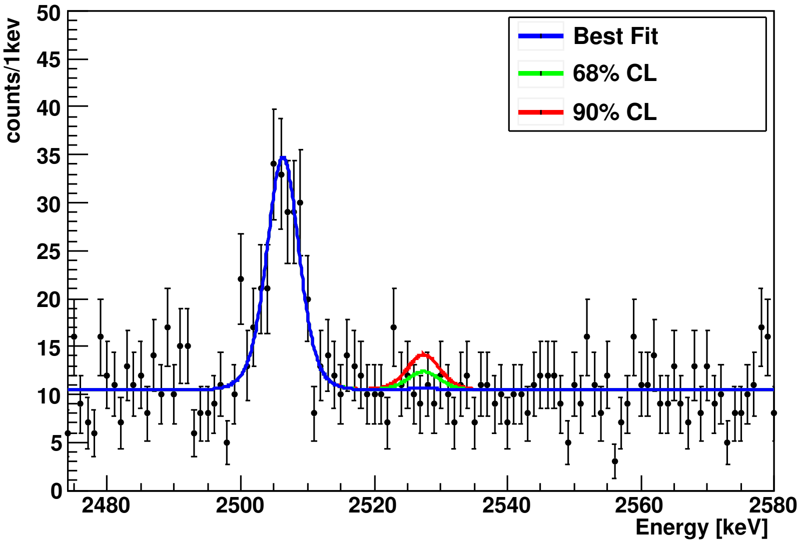


Figure 5.11: Fit result for the *fictitious* spectrum for the *single-gaussian* (top) and *multi-gaussian* response function.

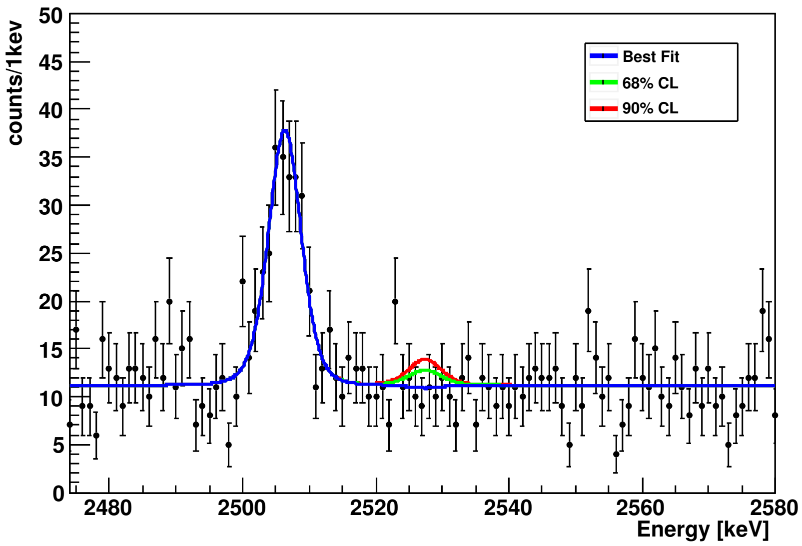


Figure 5.12: Fit result with the inclusion of RUN I. The response function here is the sum of gaussian and the background is flat.

5.5 Systematic uncertainties

In our analysis we have identified the following sources of systematic uncertainties:

- the calibration uncertainty;
- uncertainty in the signal efficiency;
- the background shape;
- the response function;
- the energy window of the fit.

In the case of the calibration uncertainty we have a direct estimate of its magnitude coming from a dedicated analysis of the residuals (see section 3.2): we reconstruct the position of a peak in the 0ν-DBD region with a precision $\Delta E = \pm 0.4$ keV. Being the position of the ^{60}Co fit more far than this value, we decided to use a more conservative value of 0.8 keV. This systematic error has been included directly in the fit as a gaussian fluctuation in the energy position. Since this uncertainty is significantly larger than the error (0.013 keV) on the Q-value [34], the systematic error discussed here also automatically includes the experimental uncertainty on the measured ^{130}Te transition energy.

To evaluate the uncertainty from the choice of energy window and shape of the background spectrum, we varied the model for the background (flat, linear and parabolic) at four different energy intervals centered at the Q-value of the 0ν-DBD decay (we increased the lower and the upper bound of an energy window starting from 2527±30 keV at increasing steps of 5 keV). The results of this analysis have shown an average variation in $\Gamma^{0\nu}$ of 3×10^{-26}y^{-1}.

The uncertainty on the signal efficiency is reported in table 3.4 to be 1.1% for the big crystals and 1.4% for the small crystals, both of which are negligible compared to the contributions from the energy scale and background parametrization.
Being the systematic error on the response functions also small (0.8×10^{-26} y^{-1}) we can state that we are dominated by the background uncertainty.

5.6 Confronting statistical methods

On CUORICINO RUN II data, the results of the *multi-spectrum* and *average-spectrum* approaches described in section 5.3.1 were indeed very similar. Table 5.4 shows the best fits and limits for the two approaches: no significant difference is observed.
In order evaluate the performances of the two approaches, we compared the distribution of the relevant statistical estimators for a thousand toy Monte Carlo simulated spectra for each CUORICINO detector, generated with rates corresponding to those directly measured in RUN II (see table 5.5) and no 0ν-DBD signal. Several pieces of information have been obtained from this comparison:

- Both methods lead to compatible results. Figure 5.13 demonstrates a strong correlation between them.
- Both are unbiased. Figure 5.14 (left panel) shows that for both methods, the distribution of the best fits divided by their statistical errors is compatible with a gaussian centered at zero with a variance equal to one. This is an important result which is not always guaranteed by maximum likelihood methods applied to low statistics.
- Both have similar sensitivities. The distributions of the 90% confidence intervals (figure 5.13 top panel) show a sensitivity of nearly 2.5x10^{24} y, evaluated as the median of the distribution, following the Feldmann and Cousins prescription [42].

The last point offers a nice synthesis for a better understanding of the numbers we are dealing with. A wide range of limits can be reached in experiments

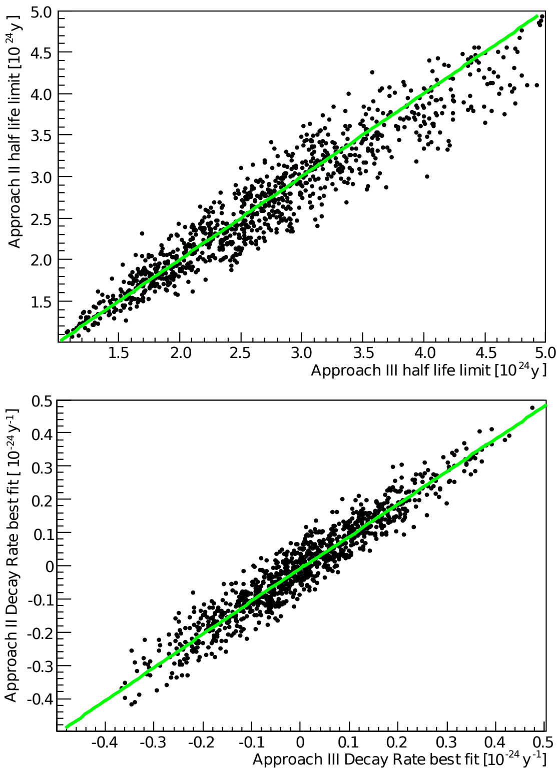


Figure 5.13: Results of a toy Monte Carlo simulation with no 0ν-DBD signal (1000 simulated CUORICINO-RUN II experiments). Scatter plot of the 90% C.L. limits (top panel) and of the best fits (bottom panel) obtained with the two different approaches. The colored line has slope=1 and shows the strong correlation between the two techniques.

with the same true background level (figure 5.14, right panel) and therefore with the same sensitivity. In this respect, quoting only the limit reached by an experiment can be misleading if the sensitivity is not also mentioned.

5.6.1 An interpretation of the MC results

We tried to find a possible - analytical - explanation for the result obtained with the MC simulations: the almost perfect compatibility between the *multi-spectrum* and the *average-spectrum* methods. To do this we model the problem making some assumptions:

- we consider Cuoricino data to be composed only by big crystals;
- we assume that the experimental data are a collection of binned spectra; one for each channel c and data-set d;

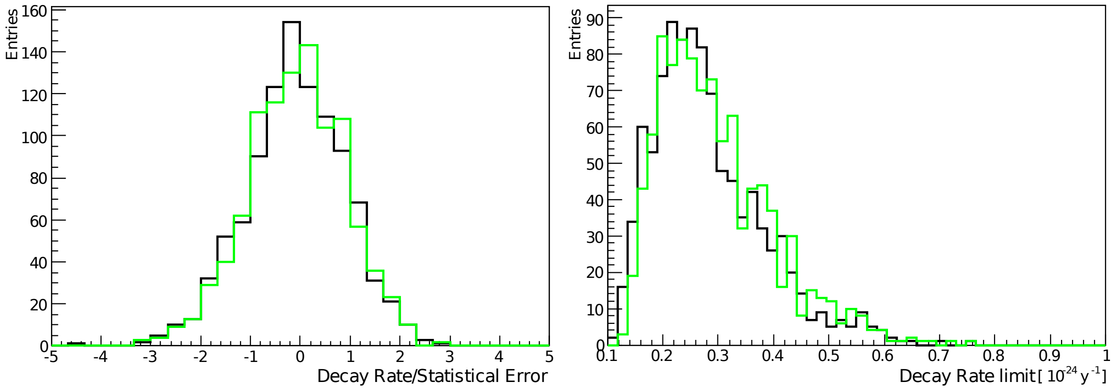


Figure 5.14: Results of a toy Monte Carlo simulation with no 0ν-DBD signal (*average-spectrum* in black and *multi-spectrum* in color). Left panel: pull distributions of the obtained best fits divided by their statistical error. Right panel: distribution of the 90% confidence level limit on the decay rate.

- $n_{c,d}(E_i)$ will be the number of counts in the i$^{\text{th}}$ bin with energy E_i for channel c and data-set d;
- we assume that the expected number of counts in our spectra $y_{c,d}(E_i)$ is the superposition of a gaussian peak (the 0ν-DBD signal) and only a flat background:

$$y_{c,d}(E_i) = m_{c,d} \cdot \left(\Gamma \cdot \frac{1}{\sqrt{(2\pi)}\Delta E_{c,d}} \exp\left[-\frac{(E_i - E_0)^2}{2\Delta E_{c,d}^2} \right] + B_{c,d} \right) \tag{5.10}$$

where $m_{c,d}$ is the corresponding exposure, while Γ and $B_{c,d}$ are the expected rate of signal and background in one energy bin.E_0 is the Q-value of the searched transition and $\Delta E_{c,d}$ the resolution measured for dataset d on channel c.

Finally we assume that:

- we know with infinite precision the value of the average expected background $B_{c,d}$, so that we can avoid any marginalization;
- we can work in a gaussian approximation of Poisson statistics;
- the number of counts $m_{c,d}\Gamma$ is negligible with respect to the background (this seems reasonable, since we actually set it to zero in the simulations).

Within this picture, the expected number of counts for one of the spectra will be:

$$N_{c,d} = \sum_i y_{cd}(E_i) \tag{5.11}$$

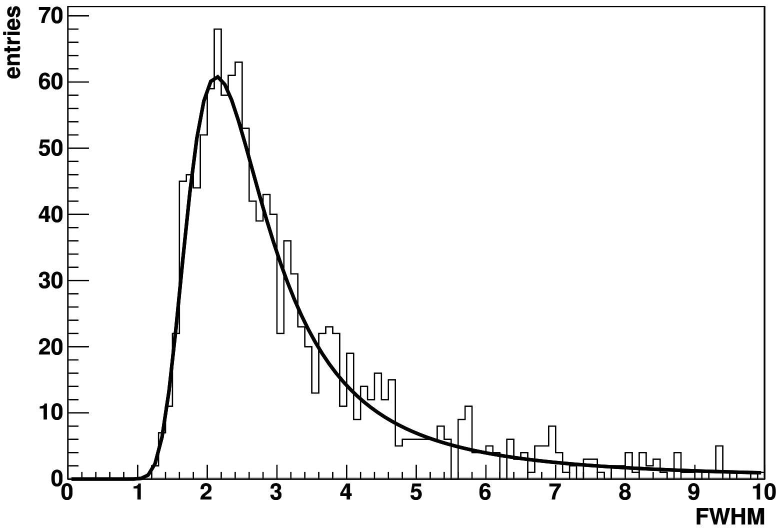


Figure 5.15: 90% upper limit on Γ as a function of the best fit. 1000 simulations are compared to the expected approximated behaviour.

Considering our assumptions, the likelihood of obtaining $n_{c,d}(E_i)$ will be a gaussian with mean value $y_{c,d}(E_i)$ and width $\sigma_{c,d}$, thus the logarithm of the likelihood for this spectrum will reduce to:

$$\log\left[\mathcal{L}(data|\Gamma)\right] = \text{const.} - \chi^2 \tag{5.12}$$

where χ^2 is defined as:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left[y_{c,d}(E_i) - n_{c,d}(E_i)\right]^2}{\sigma_{c,d}(E_i)} \tag{5.13}$$

Since we chose a flat prior for Γ, the upper limit for this quantity will be proportional to the width of the likelihood with respect to Γ near its maximum (Γ_{best}). Such width, σ_Γ, can be defined as:

$$\frac{1}{\sigma_\Gamma^2} \propto \left.\frac{\partial^2\chi^2}{\partial\Gamma^2}\right|_{\Gamma=\Gamma_{\text{best}}} \tag{5.14}$$

With the proper substitution of equation (5.11) and deriving twice (5.20), we obtain:

$$\frac{1}{\sigma_\Gamma^2} \propto \sum_{i}^{N} \frac{2m_{c,d}^2}{\sigma_{c,d}(E_i)} \frac{1}{2\pi\Delta E_{c,d}^2} \exp\left[-\frac{(E_i - E_0)^2}{\Delta E_{c,d}^2}\right] \tag{5.15}$$

Using again the gaussian approximation of Poisson statistics ($\sigma_{c,d}^2(E_i) = y_{cd}(E_i)$), the fact that we expect that Γ_{best} to be negligible with respect to B

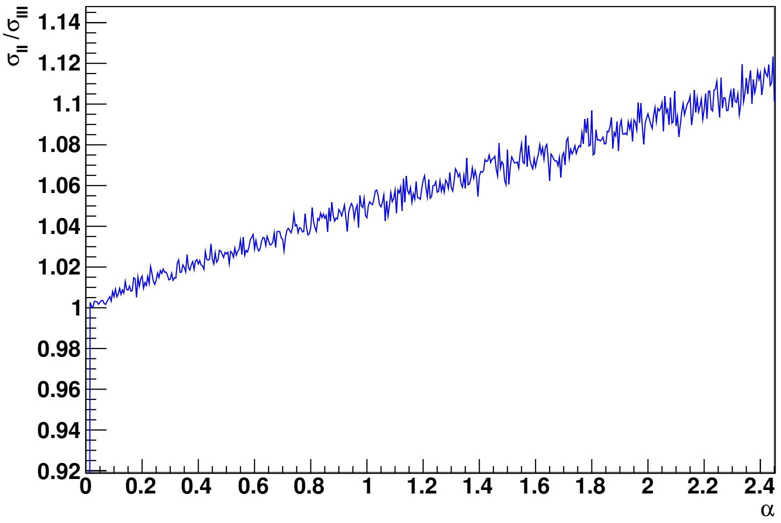


Figure 5.16: Ratio between the errors of approach II and approach III increasing the spread of the detectors' resolutions ($\alpha = 1$ correspond to the CUORICINO spread shown in figure 5.15).

and after some algebra to deal with the summation of the exponentials (for the details, we refer to [49]) we obtain a familiar expression:

$$\frac{1}{\sigma_\Gamma^2} \propto \frac{m_{c,d}}{B_{c,d}\Delta E_{cd}} \tag{5.16}$$

In which the error on Γ will be proportional to the expected background and resolution, and inversely proportional to the exposure: it's the definition of sensitivity. If we want to express the overall error, combining all channels and data-sets, we obtain the likelihood spread for the *ideal* approach [1]:

$$\frac{1}{\sigma_I^2} \propto \sum_{c,d} \frac{m_{c,d}}{B_{c,d}\Delta E_{c,d}} \tag{5.17}$$

At this point, in order to reduce the number of free parameters, a possible interpretation of the *multi-spectrum* approach [2] could be the application of the following substitution:

$$B_{c,d} \simeq \overline{B} = \frac{1}{K}\sum_{c,d} B_{c,d} \tag{5.18}$$

---

[1] This is the approach treating each detector as a different experiment without any assumption on the background

[2] Treating CUORICINO as a collection of different experiments with the same background

where K is the total number of spectra. Rather than saying that we force the detector to have the same background, we could say that we are simply evaluating the average value of the background for each channel.
As long as the backround are similar, this assumption will be correct, and the error will be proportional to:

$$\frac{1}{\sigma_{\text{II}}^2} \propto \sum_{c,d} \frac{m_{c,d}}{\overline{B} \Delta E_{c,d}}. \tag{5.19}$$

The *average-spectrum* approach [3] instead, computes also a different χ^2:

$$\chi^2 = \sum_{i=1}^{N} \frac{[y(E_i) - n(E_i)]^2}{\sigma(E_i)} \tag{5.20}$$

where $y(E_i) = \sum_{c,d} y_{c,d}(E_i)$ and $n(E_i) = \sum_{c,d} n_{c,d}(E_i)$.
Now, following the same steps that we used to reach (5.17), the rough estimate for the likelihood spread would be:

$$\frac{1}{\sigma_{\text{III}}^2} \propto \sum_{i} \frac{\left(\sum_{c,d} m_{c,d}\right) g^2(E_i)}{\sum B_{c,d}} \tag{5.21}$$

where $g(E_i)$ is the response function defined in equation (5.6). If we approximate also the response function as a gaussian whose sigma ΔE_{eff} is:

$$\Delta E_{eff} = \frac{1}{\sum_{c,d} m_{c,d}} \sum_{c,d} m_{c,d} \Delta E_{c,d} \tag{5.22}$$

we obtain:

$$\frac{1}{\sigma_{\text{III}}^2} \propto \frac{1}{\sum_{c,d} B_{c,d} \Delta E_{eff}} \sum_{c,d} m_{c,d} \tag{5.23}$$

The main difference from equation (5.19) and (5.23) is that in the latter we are averaging before summing all the terms. Clearly, the sum of a product is greater that the product of the sums, so the sensitivity of the second approach (in which we treat separately each detector fixing the background) is in principle greater.
What is the order of magnitude of this difference for CUORICINO data?
To get an idea of the ranges of appliability of the approximation of equation

---

[3] It averages all the detectors in the same spectrum, under the assumption that they have similar backgrounds

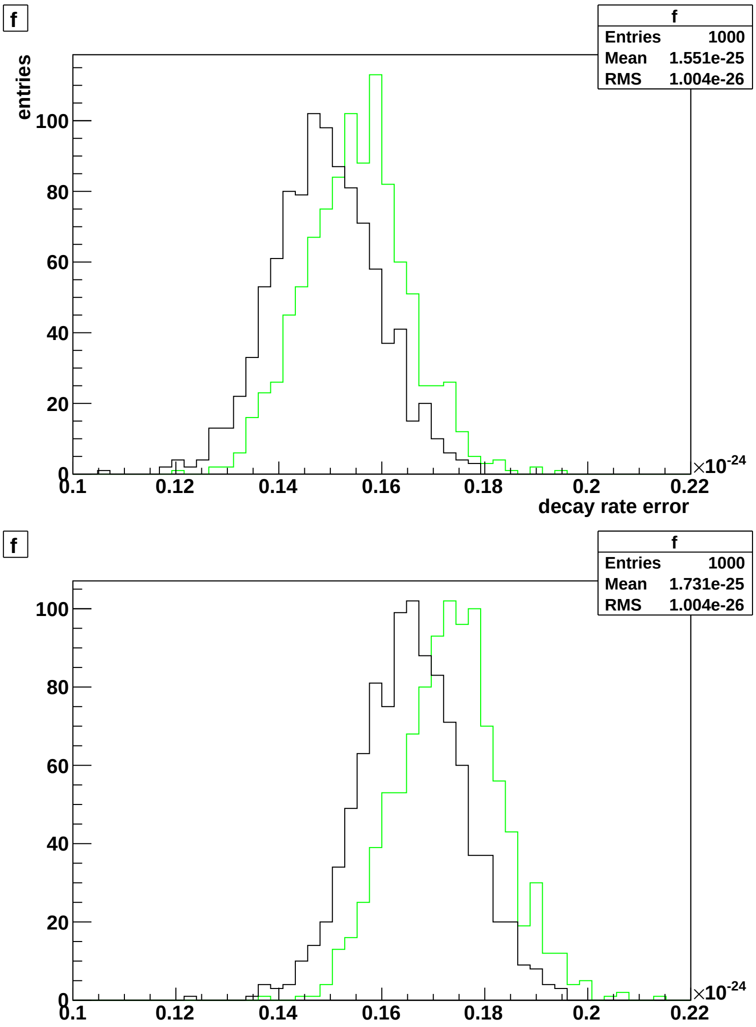


Figure 5.17: Errors on the decay rate computed on 1000 Cuoricino-RUNII-like experiments.Top panel: the true background rate is the same on each detecter. Bottom panel: the true background is different between detectors (with the same spread as the experimental one)

(5.23) we modelled the distribution of CUORICINO data with a Landau probability distribution. Figure 5.15 shows that this distribution fairly describes CUORICINO detectors. Defining σ_q as the width of this Landau function, we simulated different outcomes for CUORICINO resolutions sampled from a Landau distribution whose width is $\alpha \cdot \sigma_Q$ and we computed the ratio between (5.19) and (5.23) as a function of α.
Figure 5.16 shows the result of this evaluations.

In CUORCINO case (where $\alpha = 1$), the third approach has an error of the order of 3-5% greater than the second one. In this section surely we made many approximations, but if we confront the error of the two approaches on the simulations (see figure 5.17), we obtained that the second approach is better, with a 4% smaller error (1.55×10^{-25} y^{-1} vs. 1.44×10^{-24} y^{-1}) which is of the expected order of magnitude.
More over, we genereted a set of simulations in which the true background of each crystal can be different, taking values sampled from a flat distribution with the same width of the experimental errors. Looking at the bottom panel we can see that the two distribution are shifted, mantaining the same proportion (1.73×10^{-25} y^{-1} vs 1.66×10^{-25} y^{-1}) that they had when the background was the same: the two assumptions are in fact making the same approximations for the background introduced by equation (5.18) and this propagates in the same way for both approaches. We can then conclude that the two methods are actually making almost the same approximations concerning the background and that they gave comparable results basically thanks to the similar resultions of detectors belonging to the same group.

5.6.2 Sensitivity

An important conclusion extracted from equation (5.19) and (5.23) is that, if we have a strong knowledge of the background, the spread of our likelihood is not linked to the peculiar outcome of an experiment (the maximum Γ_{best} of its likelihood). This can be appreciated also by looking at the distribution of the experimental errors evaluated on the simulations in figure 5.17 and compare it to the spread of the limits distributions in figure 5.14. The statistical error seems to be an instrinsic characteristic of different experiments with the same background.
What really varies between experiment is the maximum of the likelihood, Γ_{best}, which surely depends on the recorded counts in the region of interest. This affects directly the limit: if we assume the likelihood to be almost gaussian with a constant error, the equation for the limit Γ_{lim} takes the form:

$$\frac{\int_0^{\Gamma_{\mathrm{lim}}} \mathrm{P}(\Gamma) \frac{1}{\sqrt{2\pi}\sigma_\Gamma} \exp\left[-\frac{(\Gamma-\Gamma_{\mathrm{best}})^2}{2\sigma_\Gamma^2}\right] \mathrm{d}\Gamma}{\int_{-\infty}^{+\infty} \mathrm{P}(\Gamma) \frac{1}{\sqrt{2\pi}\sigma_\Gamma} \exp\left[-\frac{(\Gamma-\Gamma_{\mathrm{best}})^2}{2\sigma_\Gamma^2}\right] \mathrm{d}\Gamma} = \mathrm{C.L.} \tag{5.24}$$

where $\mathrm{P}(\Gamma)$ is the prior we have chosen.

In our simulation, Γ_{best} will depend only on how much the background will be above or below its average value (this can be seen by deriving equation

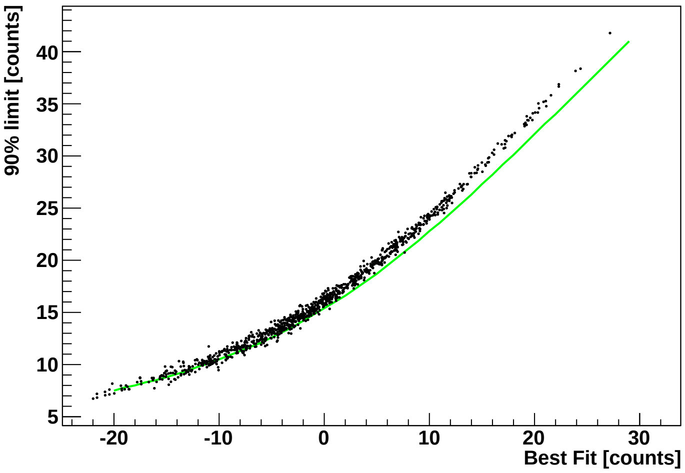


Figure 5.18: 90% upper limit on Γ as a function of the best fit. 1000 simulations are compared to the expected approximated behaviour.

(5.20) once with respect to Γ). Thus, instead of simulating N experiment then, we can gain a full picture of the behaviour of the limit as a function of Γ_{best} by computing the limit with equation (5.24) for every reasonable outcome that we expect (namely, for every resonable background fluctuations below or above its average).

Figure 5.18 shows a comparison between the resulting parameters (limits and best fits) evaluated on 1000 Cuoricino-like experiment and this approximation, in which we computed equation (5.24) as a function of Γ_{best}.

As we can se there's a good agreement between the two as Γ_{best} is close to zero. When we are looking for a signal whose contribution is compatible with the background fluctuations, the limit is actually a less stable estimate for an information on Γ than the likelihood width or - since basically they are the same - the sensitivity. This can be addressed to the fact that in the fit window we have more information concerning the background, rather than the searched signal.

Table 5.1: Fit ingredients

0νDBD Q-value [keV]	2527.5±0.013
Atomic molecular mass [g]	159.6
Average Efficiency RUN II (Big crystals)	0.827
Average Efficiency RUN II (Small crystals)	0.798
Average Efficiency RUN II (Enriched crystals)	0.798
Average Efficiency RUN I (Big crystals)	0.856
Average Efficiency RUN I (Small crystals)	0.825
Average Efficiency RUN I (Enriched crystals)	0.825
Crystal mass (Big crystals) [kg]	0.79
Crystal mass (Small crystals) [kg]	0.33
Crystal mass (Enriched crystals) [kg]	0.33
Isotopic abundances (Big crystals)	0.338
Isotopic abundances (Small crystals)	0.338
Isotopic abundances (Enriched crystals)	0.75

Table 5.2: RUN II fit results for the *single gaussian* response function. A calibration uncertainty of 0.8 keV is included in the fit.

Energy Region [keV]	[2474-2580]
Best Fit (Γ) [y^{-1}]	(0.34± 1.54)x10^{-25}
Limit ($\tau_{1/2}$) [y]	1.99x10^{24}
Flat Bkg (Big Crystals) [c/keV/kg/y]	0.154±0.006
Flat Bkg (Small Crystals) [c/keV/kg/y]	0.17±0.02
Flat Bkg (Enriched Crystals) [c/keV/kg/y]	0.34±0.05
^{60}Co Rate (Big Crystals) [c/kg/y]	2.5 ± 0.3
^{60}Co Rate (Small Crystals) [c/kg/y]	1.7 ± 0.8
^{60}Co Rate (Enriched Crystals) [c/kg/y]	0 ± 4
^{60}Co Peak Energy [keV]	2506.9±0.2

Table 5.3: RUN II fit results for the *sum of gaussians* response function. A calibration uncertainty of 0.8 keV is included in the fit.

Energy Region [keV]	[2474-2580]
Best Fit (Γ) [y^{-1}]	(0.25 ± 1.51)x10^{-25}
Half life limit [y]	1.99x10^{24}
Flat Bkg (Big Crystals) [c/keV/kg/y]	0.154±0.006
Flat Bkg (Small Crystals) [c/keV/kg/y]	0.17±0.02
Flat Bkg (Enriched Crystals) [c/keV/kg/y]	0.34±0.05
^{60}Co Rate (Big Crystals)[c/kg/y]	2.5 ± 0.3
^{60}Co Rate (Small Crystals) [c/kg/y]	1.7 ± 0.8
^{60}Co Rate (Enriched Crystals) [c/kg/y]	0 ± 3.5
^{60}Co Peak Energy [keV]	2506.5±0.3

Table 5.4: Comparison between the fit results for the two studied approaches on RUN II data.

	Multi-spectrum	Average-spectrum
Best Fit [y^{-1}]	(0.2 ± 1.5)×10^{-25}	(0.3 ± 1.5)×10^{-25}
Half-life limit [y]	2.5 × 10^{24}	2.4 × 10^{24}

Table 5.5: Fit Results combining run I and II using the sum of gaussian response functions. A calibration uncertainty of 0.8 keV is included in the fit.

Energy region [keV]	[2474-2580]
Best fit (Γ) $[y^{-1}]$	(-0.25 ± 1.44)x10^{-25}
Limit (τ)(without systematics) [y]	2.92x10^{24}
Limit (τ)(with systematics) [y]	2.81x10^{24}
Flat Bkg (Big Crystals) RUN II [c/keV/kg/y]	0.153±0.006
Flat Bkg (Small Crystals) RUN II [c/keV/kg/y]	0.17±0.02
Flat Bkg (Enriched Crystals) RUN II [c/keV/kg/y]	0.35±0.05
^{60}Co Rate (Big Crystals) RUN II [c/kg/y]	2.5 ± 0.3
^{60}Co Rate (Small Crystals) RUN II [c/kg/y]	1.7 ± 0.8
^{60}Co Rate (Enriched Crystals) RUN II [c/kg/y]	0 ± 3.5
Flat Bkg (Big Crystals) RUN I [c/keV/kg/y]	0.20±0.02
Flat Bkg (Small Crystals) RUN I [c/keV/kg/y]	0.17±0.07
Flat Bkg (Enriched Crystals) RUN I [c/keV/kg/y]	0.8±0.4
^{60}Co Rate (Big Crystals) RUN I [c/kg/y]	4.6 ± 1.5
^{60}Co Rate (Small Crystals) RUN I [c/kg/y]	9 ± 6
^{60}Co Rate (Enriched Crystals) RUN I [c/kg/y]	0 ± 14
^{60}Co Peak Energy [keV]	2506.5±0.3

Chapter 6

A prior-free analysis of Cuoricino data

> *"I trust you about as far as I can throw Manhattan"*
> *– Guybrush Threepwood*

6.1 Coverage and Priors

The confidence interval we computed for Cuoricino is obtained by means of a bayesian approach. This means that the confidence level we quote refers to the so-called "degree of belief" concerning the limit we obtained. A generic statement E - in our case: "if Neutrinoless Double Beta Decay exists, then $\tau^{0\nu}_{1/2} > 2.8 \times 10^{24}$ yrs" - is considered 90% probable. This *belief* can be expressed in many different ways, but the one of the most coherent with the Bayesian approach could be: "Given a box containing 90 white and 10 black balls, I am as confident that E will happen, as that the colour of the ball will be white. I am as confident about $\bar{E}$ (the complement of E) as of extracting a black ball".
If I have computed my limit correctly, I wouldn't see any difference between betting on E or extracting a white ball.
This probability is different from frequentist one: in that case, the limit would have been a property of an ensamble of similar experiments and the confidence level would be the fraction of these experiments for which E would be true. If a frequentist interval satisfy this requirment, we say that it has a correct "coverage".
The bayesian limit instead is related to a single experiment, the one on which E was computed. Anyhow checking the coverage can give precious hints on the

Table 6.1: Coverage of the Cuoricino limit evaluation as the sensitivity is reached

<n>	25	20	15	10	5
coverage	91%	96%	99%	99%	99%

limit evaluation behaviour, expecially on low counting experiment like Cuoricino.

For this study I performed a similar Monte Carlo study as the one presented in the previous chapter, with the addiction of a simulated signal of 0ν-DBD . Just to have a more practical idea of the number of counts we are dealing with, the average background rate used in this simulations corresponds to basically 9 counts/keV in the region of interest. The average RMS fluctuation of this background under a FWHM of th signal peak would be of nearly 7 counts. To this background we superimposed a simulated signal progressively lowered to reach this value: we started with the 0ν-DBD rate that would have caused 25 counts in the CUORICINO spectrum for the Big crystals (since the number of ^{130}Te per year is 7.3x10^{24}, this would correspond to $\tau^{0\nu}$ = 2x10^{24}yrs). Table 6.1 shows the frequentist coverage of the limit evaluation method as long as we reach the sensitivity.

As we reach the sensitivity region, whatever is the expected 0ν-DBD rate, we never fail: we always quote a correct higher limit. This is exactly what is shown in figure 5.18: there is a wide range of possible limits that are unaccesible, unless really deep background downward fluctuations, thus it is obvious that when the true signal is lower than this "bound" we will never be wrong on quoting our upper limit.

How is it, that the Bayesian prescription for a desired confidence level (i.e. 90%), ends up in a limit with different properties? The answer which is generally given in this case [50] is that the prior doesn't respect properly the researcher's real ideas about the quantity he is measuring. The flat prior we used seemed the most umbiased assumption but it hides some subtleties. One for all: with a flat prior we are stating that, before looking at Cuoricino data, we believe that:

$$P(0 < \Gamma^{0\nu} < 1) = \frac{1}{10}P(1 < \Gamma^{0\nu} < 10) = \frac{1}{100}P(10 < \Gamma^{0\nu} < 100), ... \qquad (6.1)$$

which may sound strange. If this prior is really the most unbiased, how could I say that Γ is more likely to be between 0 and 1, than between 10 and 100?

A common way to avoid this situation is to use another prior, frequently use in counting experiments: the so-called "Jeffrey's prior" [51], a distribution

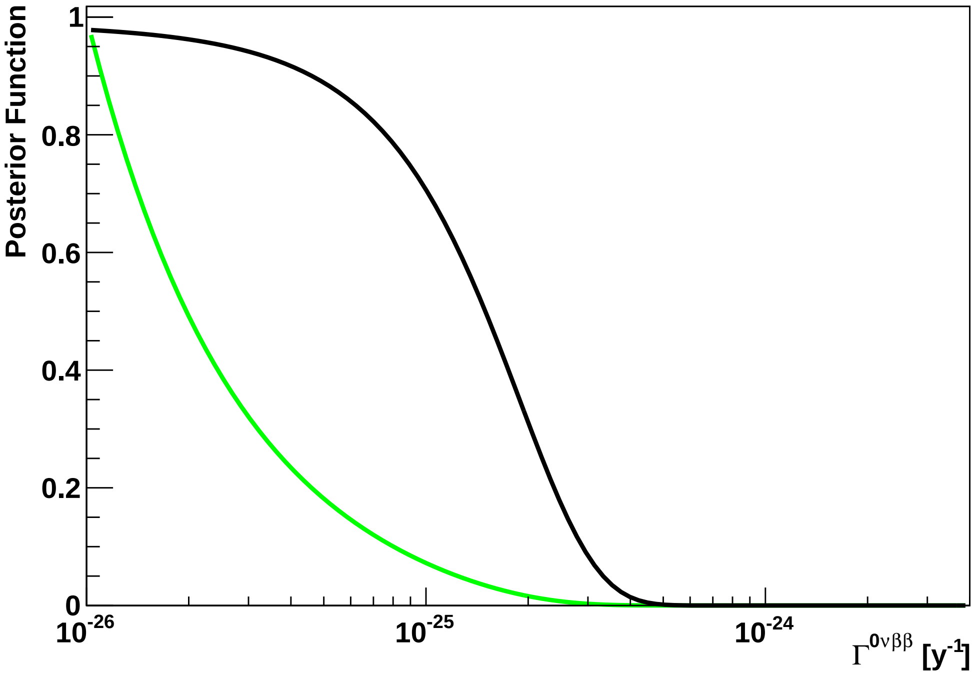


Figure 6.1: Comparison between different Cuoricino posterior functions: the black curve is obtained with a flat prior for $\Gamma^{0\nu\beta\beta}$, the green curve with a Jeffrey's prior. The distributions are not normalized.

based on the observation that what often seems uniform is not the probability per unit of $\Gamma^{0\nu}$, but rather its probability per decade (i.e., we are equally uncertain about its orders of magnitudes). Mathematically this translates in:

$$P(0 < \Gamma^{0\nu} < 1) = P(1 < \Gamma^{0\nu} < 10) = P(10 < \Gamma^{0\nu} < 100), ... \tag{6.2}$$

which implies that $\mathrm{d}P/\mathrm{d}\ \ln\Gamma^{0\nu} = k$, or $\mathrm{d}P/\mathrm{d}\Gamma^{0\nu} \propto 1/\Gamma^{0\nu}$. Since for $\Gamma^{0\nu} = 0$ there is a divergence, generally a cut-off Γ_{min} to approximate this function.

Let us check what changes between the two, if applied to Cuoricino final likelihood.

Figure 6.1 shows the comparison with the two posterior functions: Jeffrey's prior (for which we used a cut-off at $\Gamma_{min} = 10^{-28}\mathrm{y}^{-1}$) is evidently more peaked to zero. This results in a much stronger limit : $\tau^{0\nu} > 8.5 \times 10^{24}$ yrs. This is what we expected: as can be seen, this prior favors lower rates.

Anyhow, this strong dependence of the final distributions on the priors shown in this example should not be considered a bad feature, as were an artifact of Bayesian inference. It should not be a surprise: this is just a hint of the fact that

the experimental data are not so strong as to lead every scientist to the same conclusion.
As we reach the sensitivity the experiment becomes inable to change our prejudice, thus there can be subsets of the prior function which are left untouched by the likelihood of our experiment: in this case, any probabilistic statement about $\Gamma^{0\nu}$ becomes too belief-dependent.
Since a discussion on the correct prior could be endless, it would be useful to present Cuoricino data in a prior-free way, in order to gain as much informations as possible of how this experiment can effectively change our idea on Neutrinoless Double Beta Decay.

6.2 Prior-free results

Cuoricino can be considered as *frontier* experiment: it was designed to look for a tiny signal which is expected to be lower than the background which is measured during the long run of the data taking itself. It is interesting to note that, even when the signal is evident at an impressive number of sigma (one for all in this field is the Klapdor's claim [18]), the scientific community can still be doubtful about its physical interpretations. Since in Cuoricino there is no evidence of a signal, it's not then suprising the limit we quote is indeed sensitive to the researcher's prior assumption. Nevertheless, it is important to stress out that it is possible to give all the informations about Neutrinoless Double Beta Decay of ^{130}Te which are as much unbiased as possible.
The idea, proposed in [50], is very simple: presenting the results so that they should be ready (and easy) to be turned into probabilistic statements, which are needed to form one's opinion about the quantity of interest or to take decisions. It should be the scientific community to interpret and combine different experimental results and *then* translate them into statement about Nature. This can be done by noticing that the strongest tool that we have on taking decisions is the likelihood-ratio approach.
If we rewrite Bayes theorem in terms of the quantities we are measuring we obtain:

$$P(\Gamma^{0\nu}|n_{cd}) \propto L(\{n_{cd}\}\,|\Gamma^{0\nu}) \times P(\Gamma^{0\nu}) \tag{6.3}$$

we can note that we can use it to confront two different hypothesises on $\Gamma^{0\nu}$ (i.e. Γ_1 and Γ_2) as follows:

$$\frac{P(\Gamma_1^{0\nu}|\,\{n_{cd}\})}{P(\Gamma_2^{0\nu}|\,\{n_{cd}\})} = \frac{L(\{n_{cd}\}\,|\Gamma_1^{0\nu})}{L(\{n_{cd}\}\,|\Gamma_2)} \times \frac{P(\Gamma_1^{0\nu})}{P(\Gamma_2^{0\nu})} \tag{6.4}$$

In this way we keep the information coming from the data separated from the priors. The ratio of likelihoods on the right side of equation (6.4) known as

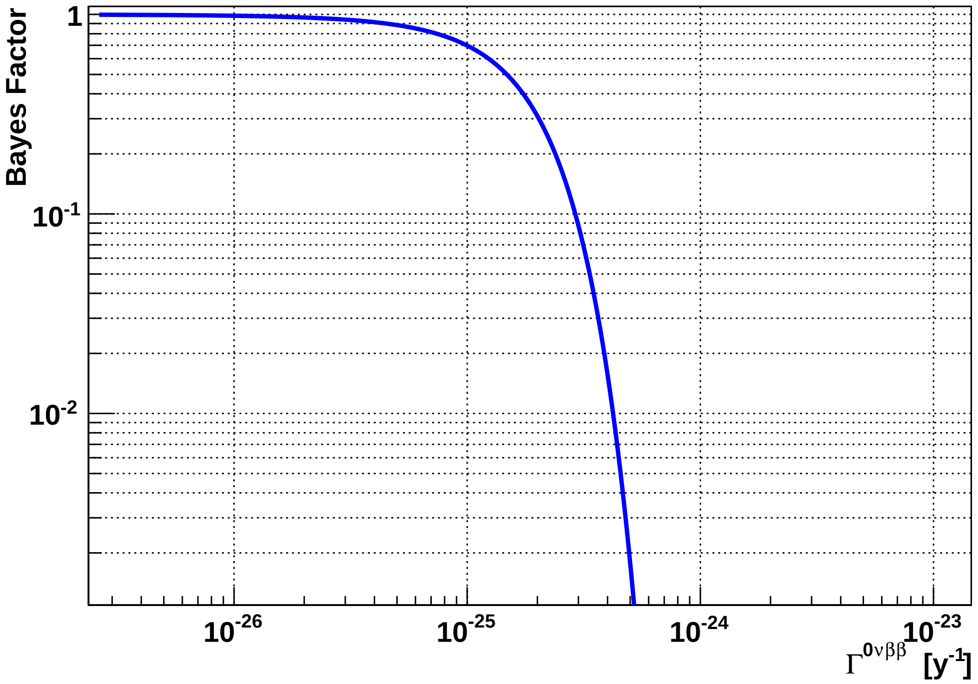


Figure 6.2: Bayes factor for Cuoricino data.

"the Bayes factor" and it quantifies the ratio of evidence provided by the data in favor of either hypothesis. The Bayes factor is considered to be practically objective because likelihoods (i.e. probabilistic description of the detector response) are usually much less critical than priors about the physics quantity of interest. The Bayes factor can be extended to a continuous set of hypotheses, by allowing Γ_1 to assume any possible value, and fixing Γ_2 to an arbitrary vale, for example $\Gamma_2^{0\nu} = 0$.

If we plot this function for Cuoricino data, we obtain the curve shown in figure 6.2. The logarithmic scale on x-axis is important: the likelihood ratio has a plateau which covers an infinite number of order of magnitudes for which it would be irrelevant if Γ is different or equal to zero (for example between 0 and 10^{-25} y^{-1}): this is the effect of the sensitivity, the data are blind to all those order of magnitudes. This function can be considered in fact as a sort of a transfer function of our prior knowledge about $\Gamma^{0\nu}$. If we believed that $\Gamma^{0\nu} = 0$, for example, then Cuoricino data are useless: our idea wouldn't have been changed by the data, since the likelihood ratio in that region is almost constant and equal to 1. Above $\Gamma^{0\nu} = 10^{-25}$ yrs^{-1} the function drops to zero, ruling out any other value.

This function gives a clear way on how someone needs to change its mind about $\Gamma^{0\nu}$ after looking at Cuoricino data and its more than a cosmetic rewriting of the likelihood ratio, since its gets rid of any propabilistic interpretation.
Looking at this function a different kind of bound can be set on $0\nu\beta\beta$ rate, for example as the conventional value of Γ where the Bayes factor equals 50%, 5% or 1% of the insensitivity plateau (becomes just a matter of taste). What is important is not to call this value a bound at a given probability level, but a sort of sensitivity bound that separates the ragion where we cannot say nothing about neutrinoless double beta decay (the sensitivity region) and a region where we can clearly rule out different order of magnitudes.
This approach combines the idea of limit and sensitivity, which are generally kept separated, at once. It also offer an easy way of interpret the results of different experiments: since the Bayes factor is a filter of our prior ideas, the combined effect of several experiments on our prejudices will be simply the product of the different Bayes factors.

Chapter 7

Conclusion

In this thesis the final result on 0ν-DBD in ^{130}Te is shown, obtained with an exposure of 19.75 kg·y of , including a detailed study of systematic errors for the first time. A half life limit of 2.8×10^{24} y at 90% C.L. is obtained (2.9×10^{24} y if systematic errors are not included), to be compared (as discussed in section 5.6) with an experimental sensitivity[1] of ∼2.6×10^{24} y. This limit can be used to extract an upper limit on m_{ee} using the theoretical NME evaluation for ^{130}Te nucleus. We report here results obtained using the most recent nuclear calculations found in literature:

- 300–570 meV using the Quasiparticle Random Phase Approximation (generally known as QRPA) evaluations of reference [8]
- 360–580 meV using the QRPA evaluations of reference [9]
- 570–710 meV using the Shell Model (SM) evaluations of reference [10]
- 350–370 meV using the Interacting Boson Model (IBM) evaluations of reference [11]

Note that, for each reference, a range (and not a single value) for m_{ee} is presented, reflecting the different results for the NME obtained by the authors when varying model parameters, such as the treatment of the short range correlations or the value of g$_A$ (the axial-vector coupling). Then, the interval 300–710 meV can be taken as the final range for the 90% C.L. upper bound on m_{ee} (at 95% C.L. this becomes 340–780 meV).

In table 7.1 we compare this result with the most stringent 90% C.L. half-life lower limits present in literature. For each experimental result, in table

---

[1] This is the sensitivity evaluated for the total (RUN I + RUN II) statistics.

Isotope	$\tau^{0\nu}_{1/2}$ [y]
^{130}Te (CUORICINO, this work)	$> 2.8\ 10^{24}$
^{76}Ge (Heidelberg-Moscow collaboration [12])	$> 1.910^{25}$
^{100}Mo (NEMO collaboration [15])	$> 5.8\ 10^{23}$
^{136}Xe (Dama/LXe [17])	$> 1.2\ 10^{24}$
^{76}Ge (Heidelberg-Moscow experiment [18])	$= 2.23^{+0.88}_{-0.62} 10^{25}$

Table 7.1: We compare the most stringent 90% C.L. half-life lower limits present in literature (first 4 rows). In the last row is reported the claim for a 0ν-DBD signal of ^{76}Ge.

7.2 we report the m_{ee} range obtained with the NME evaluations here considered. Despite the differences between the NME evaluations, it is evident that CUORICINO is one of the most sensitive experiments performed to date.

Finally, comparing the top and the bottom row in table 7.2, the 95% C.L. half-life limit on m_{ee} obtained in this work with the 2 sigma range corresponding to the positive signal quoted by [18] and obtained with a re-analysis of the Heidelberg-Moscow data. The two results are clearly compatible.
My contribution to this analysis has been mainly focused on the digital processing of raw signals, on pulse shape analysis and on the validations of the statistical tools developed to compute the final result: basically, the first and last step of the whole analysis chain. This forced me to face some of the main loop holes of the information flow from the detector DAQ up to the final spectrum. This is the reason why, in this conclusive section, Id like to focus also on the benchmark role of CUORICINO for the future DBD experiment that will be CUORE [52].
In my opinion, the main difficulty of CUORICINO analysis has been caused by the different behaviours of the detectors: this lead to a problematic automatization of the analysis process - such as pulse shape studies, calibration or stabilization - one of the main demands for CUORE. This was difficulty of many steps of the analysis, from pulse shape studies to stabilization or calibration, for example.
With this perspective, a lot of work for CUORE has been done on research and developement projects. Under the hardware point of view the collaboration is studying a more uniform assembly of the detectors. For what concerns the data analysis, we now possess a more detailed understanding of the bolometers thermal model.
Anyhow, as an example, the parametric approach to pulse shape analysis, as

Isotope	QRPA [8] m_{ee} [meV]	QRPA [9] m_{ee} [meV]	SM [10] m_{ee} [meV]	IBM [11] m_{ee} [meV]
^{130}Te (this work)	$<$ 300–570	$<$ 360–580	$<$ 570–710	$<$ 350–370
^{76}Ge ([12])	$<$ 230–400	$<$ 280–460	$<$ 530–640	$<$ 270
^{100}Mo ([15])	$<$ 610–1270	$<$ 810–1430	-	$<$ 830–850
^{136}Xe ([17])$<$ 700–1640	$<$ 800–1230	$<$ 1020–1270	$<$ 640–670	
^{76}Ge ([18])	= 180–430	= 220–500	= 410–700	= 210–290

Table 7.2: We report the different m_{ee} range obtained from the 90% C.L. of the most recent experimental result (rows 1-4) and for each of the considered NME evaluations. This identifies the upper bound on the neutrino Majorana mass according to the different results reported by the same author (when varying some of the parameters in the used nuclear model). This should be compared with the positive signal quoted by [18] (row 6). For this last case the m_{ee} range corresponds to the 2 sigma range in the measured half-life.

it has been applied to CUORICINO, will be impossible for an experiment like CUORE. In fact, the cross-checking of the performances of the implemented algorithms was one of the main difficulties: the only way was to manually identify any issue (i.e. a sudden variation of a detectors rate from its average behaviour, due to the appearance of an unexpeted pupulation of spurious signals) and then correct it.
The impressive and cautious work done by the CUORICINO analysis group on the fine-tuning of the software is something that we cannot afford in CUORE. We dont know yet the impact on CUORE data of all the refinements implemented for CUORICINO: a detailed study will be mandatory between obtaining a better resolution or a lower background and the effective work necessary to guarantee stable performances of these techniques.

For what concerns the statistical approaches to the final fit of CUORICINO spectrum for the 0-DBD limit evaluation, I think we gained a lot of informations by comparing the differences between two different approaches described in section 5.3.1. The main lesson I learned is that, no matter which statistical refinements we can use (at least the one we figured out), the background fluctuations in the DBD region prevents us from any significant gain in sensitivity. Also the valutation of the systematics, proved that the main source of uncertainty is the lack of a background model. Since the background is our main issue, and since it is the main source of errors on any evaluation of this elusive decay, I believe that the best approach for the future would be keeping its evaluation separated from the signal. The main systematic effects we got came from the background model uncertainty (fit window and background function), so I believe that this error should be computed directly on the background and then propagated on the DBD rate. As in blind analysis, if the fit is carried out excluding the energy region where the peak should appear (for example one FWHM under the DBD Q-value), we should be able to study and predict the number of counts under the peak with a statistical and systematic error.

For what I have been able to understand during this thesis work, it is my opinion that only the Bayesian approach (rather thant the Feldmann&Cousins one [42]) offers an easy and simple way to propagate this information (the expected number of background counts under the DBD peak) to produce an unbiased estimate on the neutrinoless double beta decay. In the last chapter I tried to summarize what I would choose about a way of presenting a scientific result near the sensitivity of an experiment, focusing on all the advantages that the Bayesian approach can offer as a clear way to present usable results. The main goal, rather than quoting and refining the limit evaluation of a few percent, should be to produce a result which is fairly understandable and comparable with other results. Even if we see nothing, we still can be as much clear and precise as possible about our uncertainty.

Appendix A

TIMEZZO

The operating principle of the program TIMEZZO is that the likelihood function of the data we are dealing with is asymptotically gaussian.
To describe TIMEZZO operating principles, we start from the monodimentional case, where only a parameter Γ has to be studied.
Under very general conditions, given a collection of data $\underline{x}$, when the number of observation becomes "infinitely" large, the likelihood function $L(\underline{x}|\Gamma)$ takes the form of a gaussian in Γ with mean value $\underline{\Gamma}$ and variance σ^2. This is a consequence of the Minimum Variance Bound theorem [53] and allows us to express $\mathcal{L}$ as:

$$\mathcal{L}(\underline{x}|\Gamma) = \mathcal{L}^{\max} \cdot \exp\left(-\frac{1}{2}\chi^2\right) \tag{A.1}$$

or

$$\log \mathcal{L}(\underline{x}|\Gamma) = \log\left(\mathcal{L}^{\max}\right) - \frac{1}{2}\chi^2 \tag{A.2}$$

where χ^2 is given by the formula:

$$\chi^2 = \left(\frac{\Gamma - \underline{\Gamma}}{\sigma^2}\right)^2. \tag{A.3}$$

The principle used by TIMEZZO for the evaluation of confidence intervals is the so called "intersection procedure": the interval is custructed by intersecting the gaussian $L(\Gamma)$ by the straight line $\mathcal{L}_{max} \cdot e^{-\alpha}$: if $\alpha = 0.5$ we are computing a 68% interval, $\alpha = 2$ identifies the 95% and so on.

The advantage we get from this approach is that this can be done also for non-symmetric gaussians: if there exists a generic transformation $g = g(\Gamma)$ which transforms the function $\mathcal{L}(\Gamma)$ into a gaussian of unit variance and mean

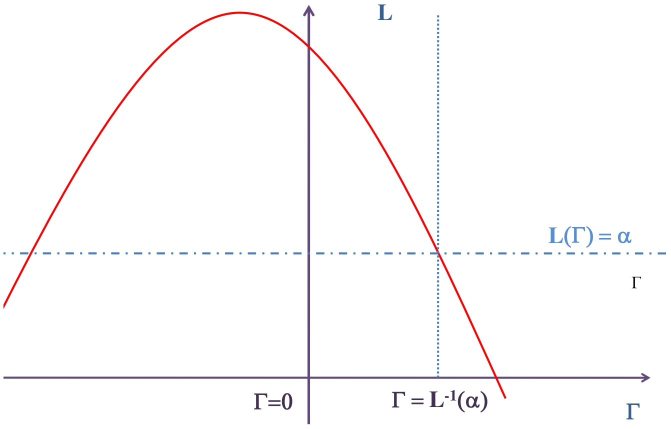


Figure A.1: Graphical representation of the intersection method in case an unphysical region is excluded (i.e. $\Gamma < 0$).

value $\bar{g}$. With this new parametrization in the new variable g, $\mathcal{L}$ takes the form:

$$\mathcal{L}_g(\underline{x}, g) \propto \exp\left[-\frac{1}{2}\left(g - \bar{g}\right)^2\right] \tag{A.4}$$

In principle, one could extract the parameters with the intersection procedure on $\mathcal{L}_g$, finding the confidence interval $[g_1, g_2]$ and then inverting the relation $g(\Gamma)$ in order to get the proper interval $[\Gamma_1, \Gamma_2]$ but this is actually non necessary. One of the most important properties of the likelihood function is that it is invariant under the parametrization (since it gives the joint probability of obtaining $\underline{x}$, this probability cannot change if we express it as $\mathcal{L}_\Gamma$ or $\mathcal{L}_g$. For all the possible Γ we should have:

$$\mathcal{L}_\Gamma(\underline{x}|\Gamma) = \mathcal{L}_g(\underline{x}|g) \tag{A.5}$$

Streactly speaking, this approximation can only be approximately correct. The existance of a transformation function $g(\Gamma)$ is not granted at will. As long as $\log \mathcal{L}$ has a single maximum and does not deviate too much from a parabola, the intersection procedure would work properly. Of course this implies that one has to check a posterior if the approximation is reasonable, given the actual shape of $\mathcal{L}_\Gamma$.

If a flat prior function is included, TIMEZZO numerically search for a proper value of the intersection α that would give the desired confidence interval in case part of the area. As an example, figure A.1 shows a situation in which the

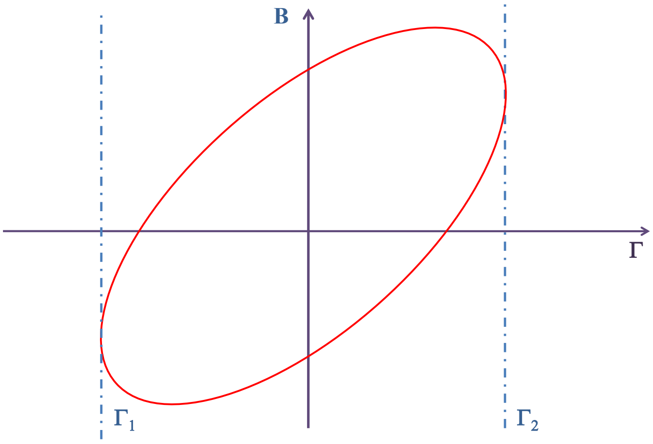


Figure A.2: Graphical representation of TIMEZZO operating principle for confidence interval evaluation.

maximum likelihood value is in the negative region: TIMEZZO then finds the value $alpha$ for which the interval $[\Gamma = 0, \Gamma = \mathcal{L}^{-1}(\alpha)]$ identifies the desired probability content of the physical region $\Gamma \geq 0$.
This idea can be generalized to the N-dimensional case, if we extend the hypothesis of a normal distribution also to the other parameters in our model. In this case the generalized likelihood becomes:

$$\mathcal{L}(\underline{x}|\underline{\theta}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\mathcal{V}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\left(\underline{x}-\underline{\theta}\right)^T \mathcal{V}^{-1}\left(\underline{x}-\underline{\theta}\right)\right] \tag{A.6}$$

where $\underline{\theta}$ is the N-dimensional vector of the parameters and $\mathcal{V}$ is the covariance matrix. In this case the χ^2 will be:

$$\chi^2(\underline{x}, \underline{\theta}) = \left(\underline{x}-\underline{\theta}\right)^T \mathcal{V}^{-1}\left(\underline{x}-\underline{\theta}\right), \tag{A.7}$$

and the region of space defined by:

$$\chi^2(\underline{x}, \underline{\theta}) = \alpha \tag{A.8}$$

will be a hypersurface of constant probability density, which will depend on the value of α. Figure A.2 shows the bidimensional case in which we have two parameters, Γ and B, in which the ellipse $\chi^2 = \alpha$ gives the region of space having a probability content of 68.3%.
In this context, the region $[\Gamma_1, \Gamma_2]$ which has the 68.3% probability content for Γ,

idependently by the value of B, is the intersection of the Γ axis with the vertical tangents of the ellipse [53].
In the bidimentional case TIMEZZO would then proceed as follows:

- using the gaussian approximantion, TIMEZZO finds the value α that would identify a region subtended by a normal distribution with the desired probability content (taking into account the proper normalization due to the inclusion of priors);
- it calls MINUIT to identify the isocountur corresponding to the intersection $\chi^2 = \alpha$.
- finally, it identifies the interval $[\Gamma_1, \Gamma_2]$ by intersecting the Γ-axis with the vertical tangents of the ellipsoid.

The N-dimentional case, is a generalization of this procedure.

Appendix B

A simulation of bolometric noise

A standard approach in noise time series simulation is based on Carson's theorem [54]. In fact, the superposition of randomly delayed pulses of a definite shape $f(t)$ with arbitrary coefficients a$_k$ gives rise to a pulse train

$$n(t) = \sum_k a_k f(t - t_k) \tag{B.1}$$

whose power spectrum $N(\omega)$ can be expressed in terms of the $f(t)$ power spectrum $P_f(\omega)$ according to

$$N(\omega) = \alpha P_f(\omega) = \alpha |\widehat{F}(\omega)|^2 \tag{B.2}$$

where α is a proper normalizing constant and $\widehat{F}(\omega)$ is the Fourier transform of f(t). The delays t_k's are distributed according to Poisson statistics and $f(t)$ is usually required to have a zero mean value so that $n(t)$ averages to zero. It can be shown that this result is independent of the density function used for the random amplitudes a_k, if the variance of such distribution is fixed by the choice of the normalizing constant α.

The basic idea of the method consists then in searching a proper expression for $f(t)$ so that the generated time series $n(t)$ have the desired power spectrum $N(\omega)$. In other terms one is faced with the problem of inverting equation (B.2) to obtain the expression of $f(t)$ given $P_f(\omega)$. Then, $n(t)$ can be built iteratively according to equation (B.1) with a proper choice of the random parameters a_k and t_k.

This approach can be quite difficult whenever the desired spectrum is different from a simple power law and analytical approaches to the determination of $f(t)$ are usually not straight-forward. In fact, the desired noise spectrum often shows complex behaviours in the measured frequency range. An example of such a common situation is shown in fig. B.1 (referring to the output of a few hundred microgram bolometric detector) where we have a complex forest of microphonic lines overimposed to a smoothly varying distribution. The anti-aliasing filter cut frequency is also apparent by the quick drop of the power spectrum in the 100-200 Hz region.

The method proposed here consists in the determination of an $f(t)$ which exactly matches a given noise power spectrum. This is essentially based on the choice of a particular (out of an infinite number of possibilities) function $f(t)$ which satisfies equation (B.2) and can be applied both to the case in which $N(\omega)$ is known analytically or experimentally measured.

Let us begin with the discrete case of a sampled sequence $f[k]$. If we consider a finite number L of samples and assume ergodicity, then the power spectrum can be approximated by an ensemble average over a sufficiently large number of finite sequences $f_L[k]$ of fixed length L according to

$$P_f[k] \simeq < |\mathcal{F}(f_L[k])|^2 > \tag{B.3}$$

where $\mathcal{F}$ stands for the Discrete Fourier Transform (DFT) operator. Since the sequence elements $f_L[k]$ are supposed to be real, the corresponding power spectrum is a real even function and only half of its values are independent. In fact all the informations concerning the phases are lost in the quadrature and this is the reason why the inversion of equation (B.3) does not admit a unique solution $f_L[k]$.

The simplest solution to equation (B.3) is obtained by deleting the ensemble average and inverting the resulting expression. Phases can be then randomly added assuming a flat distribution. The result represents the core of this work and can be summarized as follows:

$$F[k] \equiv \sqrt{N[k]} \cdot e^{i\theta_k} \tag{B.4}$$

$$f[k] = \mathcal{F}^{-1}(F[k]) \tag{B.5}$$

where $N[k]$ is the power spectrum according to which we want to generate our time series (known a priori) and θ_k are random numbers uniformely distributed between 0 and 2π. In order to guarantee the condition of reality of $f[k]$ upon inverse DFT the sYmmetry constraint $F[k] = F^*[-k]$ must be imposed.

Noise time series can now be built following equation (B.1), simply generating the delays t_k according a Poisson distribution:

$$t_k = t_{k-1} - \ln(1 - R)/\lambda \tag{B.6}$$

where R is a random number with a uniform distribution and λ, which represents the overlap rate of the signals $f[k]$, is the only free parameter of the method and is usually adjusted to improve the quality of the simulated time series.

As stated above, the choice of the normalization term α partially fixes the arbitraryness in the choice of the amplitudes density function by setting its variance. This can be better understood by following some of the steps in the derivation of the Carson's theorem. Actually, by substituting equation (B.1) into the the power spectrum definition we obtain:

$$\begin{aligned} N(\omega) &= \left\langle \left| \mathcal{F}\left[\sum_k a_k f(t - t_k) \right] \right|^2 \right\rangle = \\ &= \left\langle |F(\omega)|^2 \cdot \sum_{ij} a_i a_j e^{i\omega(t_j - t_i)} \right\rangle \end{aligned} \tag{B.7}$$

The last sum can then be separated into two different terms, one for $i = j$, the other for $i \neq j$. It can be shown that the latter is proportional to $F(0)$ and therefore vanishes since we required $f(t)$ to average to zero. We obtain therefore:

$$N(\omega) = \left\langle |F(\omega)|^2 \right\rangle \left\langle \sum_{k=1}^{L} |a_k|^2 \right\rangle \tag{B.8}$$

The averaged sum on the right hand part can be thought as the average of the squared amplitude of each single pulse times their average number M of occurencies. As a result we can express the average power spectrum as:

$$N(\omega) = \left\langle |F(\omega)|^2 \right\rangle M < |a|^2 > \tag{B.9}$$

By comparing equation (B.9) and equation (B.2) we can finally obtain the relation between the average rate λ, the normalization constant α and the amplitudes variance $|a|^2$:

$$|a|^2 = \frac{\alpha}{M} = \frac{\alpha}{\lambda T} \tag{B.10}$$

where T is the finite length of the time series. Any arbitrary distribution with a variance given by equation (B.10) can now be used to generate the random amplitudes a_k and the method is fully defined.

We can therefore summarize the most relevant steps for the method implementation as follows:

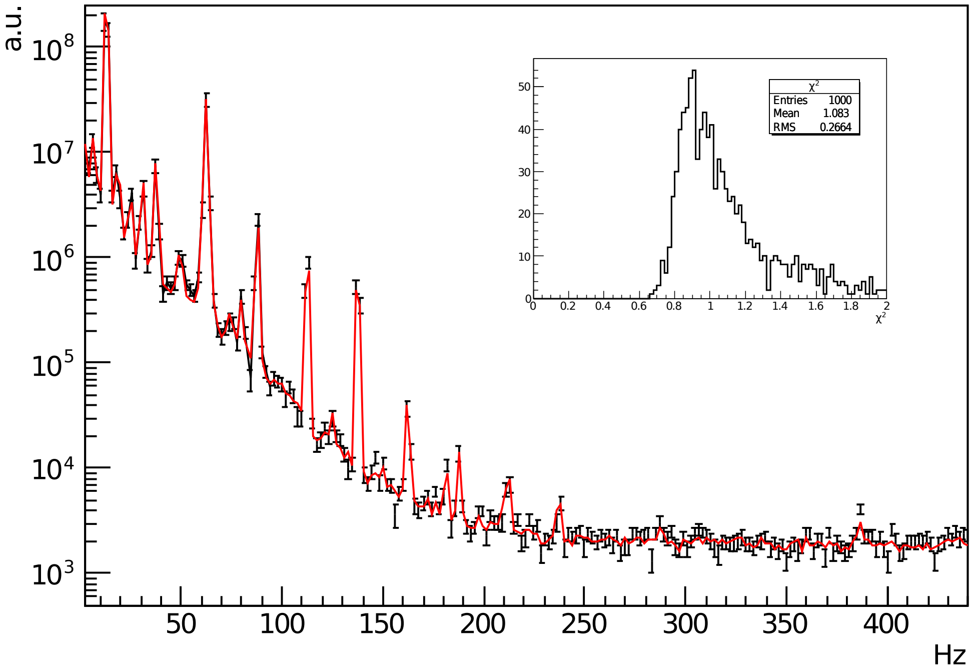


Figure B.1: Superposition between a simulated (black) and measured (red) bolometric noise power spectrum obtained with a bolometric detector of O(mg) mass. The simulated spectrum was obtained averaging 50 finite time series. Error bars are the average standard deviations of the simulated spectra. The distribution of the χ^2/d.o.f. for a set of simulated spectra is shown in the inset.

- Given the desired noise power spectrum $N(\omega)$ select the basis function f(t) according to equation (B.4) and (B.5). Then equation (B.2) fixes the constant α

- Generate a set of increasing delays t_k according to a Poisson distribution with average rate λ.

- Generate the random amplitudes a_k according to an arbitrary distribution with variance fixed by equation (B.10).

- Shift $f(t)$ by t_k imposing a periodicity constraint $f(t) = f(t+T)$ and multiply the result by a_k.

- Sum iteratively over t_k while $\sum t_k < T$

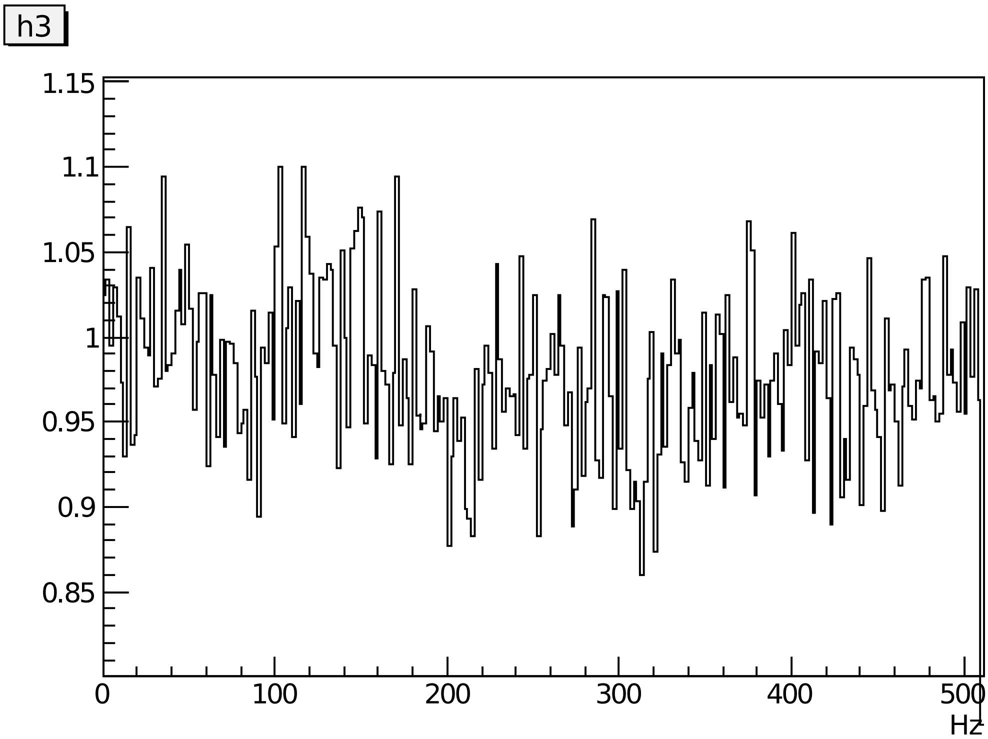


Figure B.2: Ratio between the simulated and measured bolometric noise power spectra shown in Fig.B.1.

The method can be used also when a theoretical (analog) estimate of $F(\omega)$ is known. This is a common situation when the performance of an electronic circuit or of a detector are studied. In this cases the discrete power spectrum F[k] can be in fact obtained by means of a proper frequency warping [55] of the analog theoretical estimate of the power spectrum.

B.1 Validation

Following the recipe described here, from the power spectra of a small mass bolometer, a set of 50 simulated time series was generated. The respective power spectra were obtained according to equation (B.3) by averaging their DFT's. The result is directly compared in fig. B.1 with the original noise power spectra obtained experimentally. The ratio between the experimental and simulated power spectra is also reported in fig. B.2.

In order to allow a more quantitative estimate of the level of agreement be-

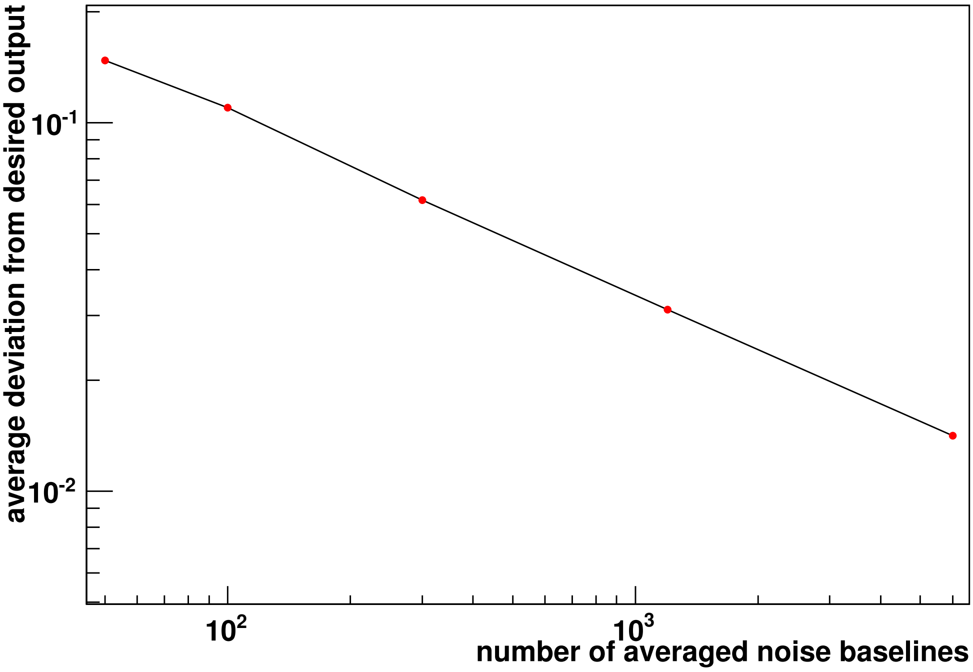


Figure B.3: Average normalized standard deviation (dots) of the simulated power spectra as a function of the number of simulated time series used for the average process. The black line joins simulated points while the red one is an inverse square root law shown for reference.

tween the simulated and the original power spectra, a full set of simulated power spectra was generated in order to evaluate the distribution of the simulated values at each frequency. Each power spectrum was obtained averaging 50 simulated time series. The obtained standard deviations are shown in fig. B.1. The distributions of the χ^2/d.o.f. are also shown to demonstrate the correct statistical behaviour. The fluctuations of the simulated power spectra values depend of course also on the number of time series used for the average process (equation (B.3)). Such a dependence was explicitely studied by generating a different set of power spectra and evaluating the corresponding average standard deviation from the desired otput, as a function of the number of averaged time series. The result is shown in fig. B.3 and found in excellent agreement with an inverse square root law.

Appendix C

Pulse shape analysis for the CUORICINO experiment

Pulse shape analysis is used to reject spurious signals.
We classify events through confidence intervals determined for each recorded pulse by the distributions of the shape parameters described in section 4.5. Signals falling within these intervals are defined as ″true″ (or physical) pulses, while signals having one or more of their parameters outside the relevant intervals are rejected as noise or pileups.
The use of more than one pulse shape parameter results in better reliability of the rejection technique but is important to stress out that the definition of these confidence levels is arbitrary: what really matters is the possibility to measure the efficiency of this procedure in order to control the dead time induced by such cuts.

C.1 Spurious signals

The definition of spurious events is kept as much general as possible: a triggered event containing any waveform which is not caused by a single particle event. Also true particle signals could be discarded if they are pile-up events or if the pulse shape is ″deformed″ by a fake event (see figure C.2). In both cases the choice of rejecting a particle event is motivated by the possible failure of the pulse amplitude evaluation algorithms due to the presence of a secondary waveform in the analyzed sample of data.
As clearly visible in figure C.1 below a certain energy a pulse shape parameter as the rise time can become unefficient and it is hard to separate good and bad events: particle signals are the higher horizontal line, the lower one is the dis-

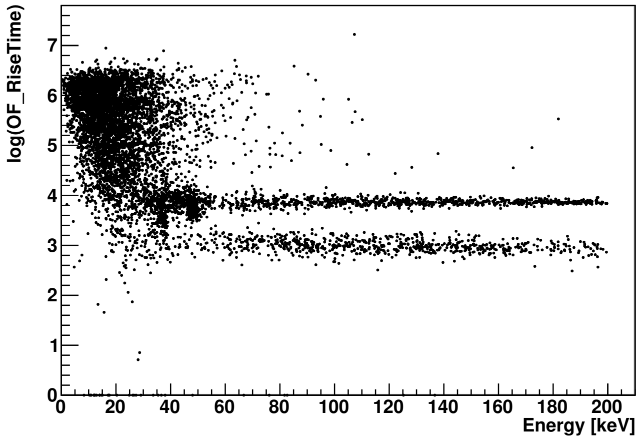


Figure C.1: Scatter plot of the logarithm of the rise time versus energy.

tribution of "spikes", impulsive electronics noise with small time constants. In the lower energy region, the distributions of good events and noise waveforms merge. To avoid any noise contamination in the data sample used for physics a software energy threshold (see section C.4) is usually set, selecting only those signals for which the pulse shape parameters are enough efficient.
Typical examples of rejected events are shown in figure C.2.

C.2 Shape parameters linearization

Pulse shape analysis has to make decisions based on the distributions of a set of shape indicators in order to identify spurious events. For example, an efficient pulse shape parameter r in general should satisfy this condition:

$$P(\text{particle}|r = \overline{r}) \gg P(\text{spurious}|r = \overline{r}) : \tag{C.1}$$

the signals for which r is close to its average value $\underline{r}$ are more likely caused by a particle interaction, rather than by a spurious event. In this situation, a reliable pulse shape cut could be:

$$|r - \overline{r}| < n\sigma_r \tag{C.2}$$

where σ_r is the width of the distribution of r and n identifies the desired number-of-sigma confidence level. This is because we are implicitly making the assumption that:

$$P(\text{particle}|r) \simeq P(r) \tag{C.3}$$

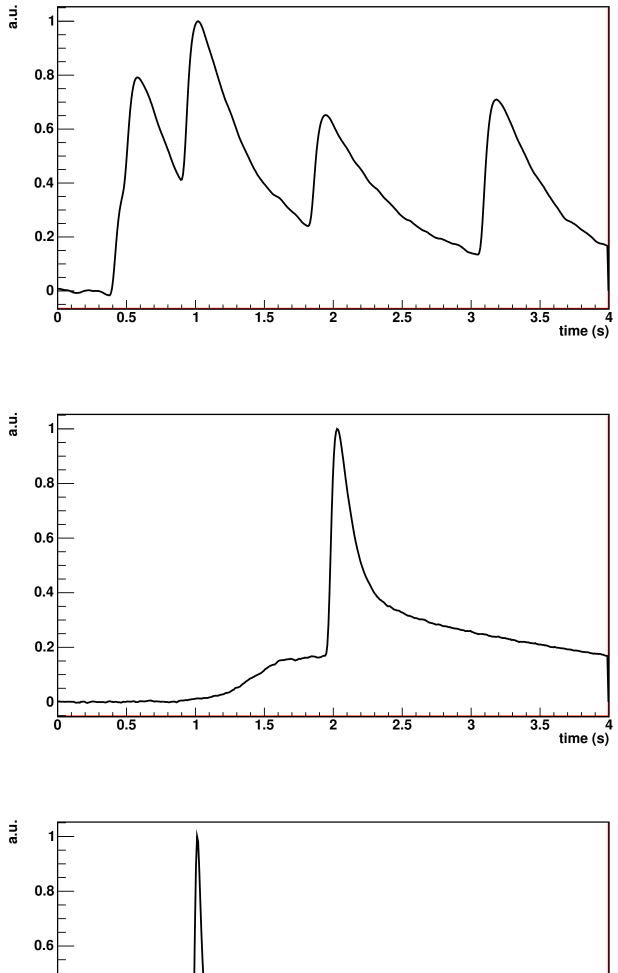


Figure C.2: Examples of spurious signals: a pile-up (top), a pileup with a non particle event (middle) and a spike (bottom).

Anyhow, it can happen that r depends on the energy E of the signal which it was computed on. In this case, for two different energies E_1 and E_2 , in general we could have have that:

$$P(\text{particle}|r(E_1)) \neq P(\text{particle}|r(E_2)) \tag{C.4}$$

which makes unreliable any simple cut like the one in equation (C.2).
The main idea behind linearization of a parameter r, is to find a transformation $r \to \widehat{r}$, such that:

$$P(\text{particle}|\widehat{r}(E_1)) \simeq P(\text{particle}|\widehat{r}(E_2)) \tag{C.5}$$

In order to obtain $\widehat{r}$ from r, we perform a fit on a scatter plot of this quantity versus the energy, interpolating with a function $f_r(E, \vec{a})$, where $\vec{a}$ is the vector of free parameters to be found. The parameter r of an event with measured energy E , will then become:

$$\widehat{r}(E) = \frac{r}{f_r(E, \vec{a}_{\text{best}})} \tag{C.6}$$

where $\vec{a}_{\text{best}}$ is the vector obtained by the regression.
It is important to stress out that this method is indeed rough, since we basically straighten the distributions without any significant knowledge of the parameters governing such non linear behaviour: this knowledge can come from a thermal model of the thermistor response in conjunction with a measurement of the main quantities governing the detector's response (resisistances, capacitances, and so on...). The details on this procedure are described in the next two sessions.

C.2.1 Interpolating functions

In order to measure the trend of shape parameters a simple least square fit is not efficient: the presence of outliers will cause a deformation of the interpolating function beacause each point has the same weight of the others; it is then necessary to give stress to regions where the probability of a "good" signal to occur is high.
The method used in order to control the presence of outliers is called "robust regression". The basic idea behind such technique is to change the error function to be minimized during the regression so that it is less sensitive to low density regions: for example, given N measured values E_i, r_i and a model $f(E_i, \vec{a})$, a robust regression will minimize, with respect to $\vec{a}$:

$$\text{Err}(\vec{a}) = \sum_i \log\left[\left(\frac{r_i - f(E_i, \vec{a})}{s}\right)^2\right] \tag{C.7}$$

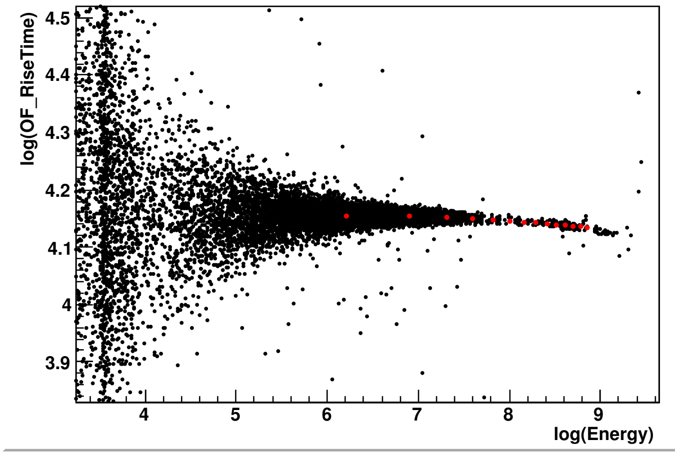


Figure C.3: Scatter plot of the logarithm of the rise time versus the logarithm of the energy: the non-linearity of the detector's response gets higher at greater energies. The energy range is 0-10 MeV: red dots identify intervals of 500 keV.

where Err($\vec{a}$) is the error function and s a scaling quantity which is generally iteratively measured.

In this way, bigger errors induced by outliers are reduced by the use of the logarithm and by the presence of s.

The dependence of rise and decay times (to which we will refer as RT and DT) from the energy had been modelled as a second order polinomial, while for the OFT parameter, which has shown different trends even in the same channel during CUORICINO, three are the main tendencies that have been used to force it to be linear. Examples of these trends are shown in figure C.4. The trend on the top panel is modelled with fourth order polynomial. The second trend has been derived by noticing that its derivative should behave like a sigmoid function, since it is characterized by a negative slope at low energy and a positive slope at higher energies. The interpolating function for the trend in the middle panel is then the integral of a sigmoid and will have five free parameters: the minimum value, the width of this minimum, the two slopes and the additive constant of the integral.

The bottom panel is a rare observed condition for some channels in which the shape parameter presents a minimum with an adrupt discontinuity. It has then been modelled as:

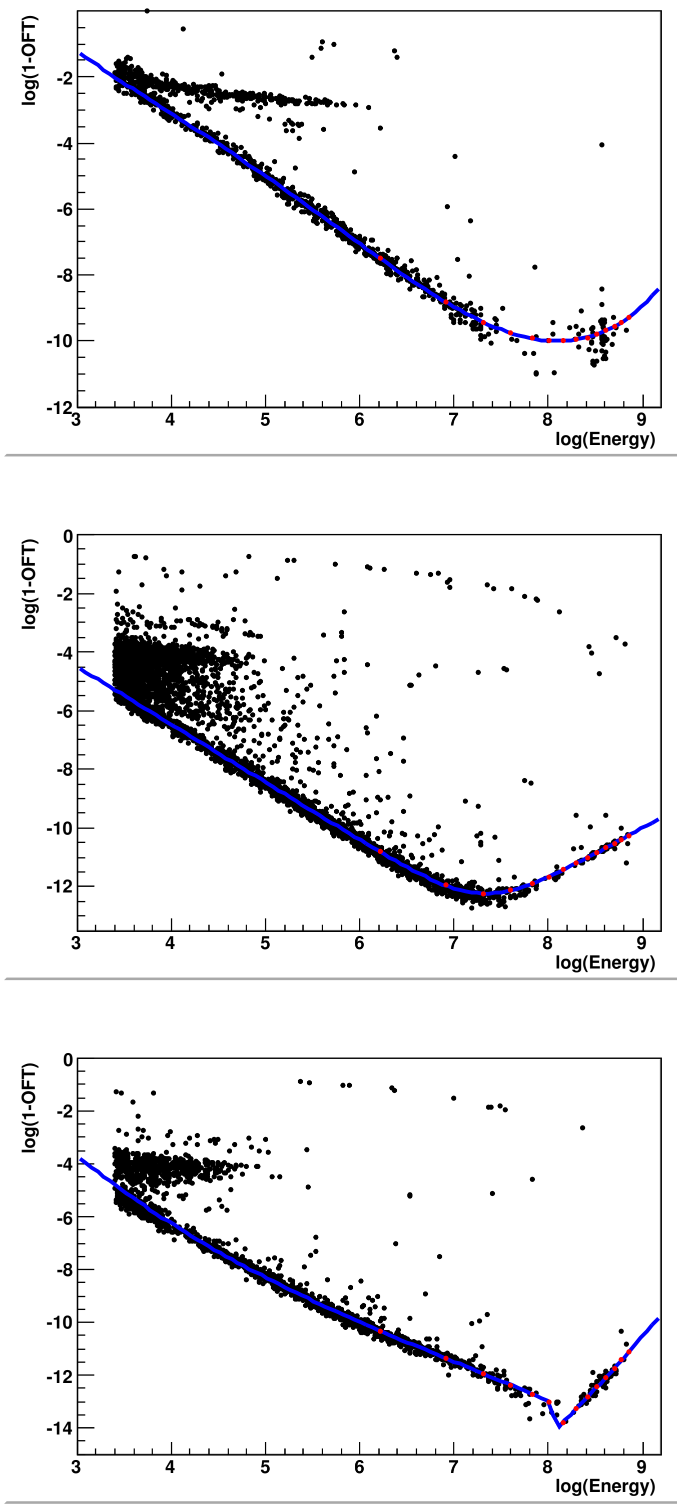


Figure C.4: Examples of the three interpolating functions used in the OFT parameter linearization. Outliers are most likely to be pileups and "triangoloni", while the high density regions far from the blue trends are spikes or low energy noise events. The energy range is 30-7000 keV: red dots identify intervals of 500 keV.

$$\begin{cases} \sum_{i=0}^{2} a_i \cdot E^i, \text{if } E < E_{min} \\ \sum_{i=0}^{2} b_i \cdot E^i, \text{if } E \geq E_{min} \end{cases} \tag{C.8}$$

where a_i ,b_i and E_{min} are the free parameters.

For each channel and data set we must then decide which is the best model for the OFT: since we have many outliers, the χ^2 test is not sensitive to the goodness of fit (the average error could be high, even with a good linearization, because of the presence of spurious events). Thus, the decision pointed to the one whose effect makes OFT more linear: we perform a regression with a simple first order polinomial on each of the three $\widehat{\mathrm{OFT}}$ obtained with the different functions and chose the one whose slope il closer to zero.

C.2.2 Variance normalization

In the first section, we described a general transformation of a shape parameter $r \to \overline{r}$ such that a cut of the form:

$$|\widehat{r} - \underline{\widehat{r}}| \leq n\sigma_r \tag{C.9}$$

can be considered a pulse shape cut. Actually, as can be appreciated from the top panel of figure C.4, also σ_r can vary with the energy. By variance normalization we mean then the control of the parameter's variance $\sigma_r(E)$, expressed as further tranformation for r, which could make a cut as much as possible independent of the energy:

$$|\widehat{r} - \underline{\widehat{r}}| \leq n\sigma_r(E) \to r' \equiv \frac{|\widehat{r} - \underline{\widehat{r}}|}{\sigma_r(E)} \leq n \tag{C.10}$$

This offers the possibility to perform a "number of sigma" cut which has the same confidence level at any energy. It also offers the opportunity of treating each channel in the same way, since by construction the parameter should average to the same mean and variance.
This procedure will lead to the normalized quantity ($\mathrm{PSA}_{\mathrm{nsigma}}$) applied in the main pulse shape cut:

- Produce a scatter plot of $\widehat{OFT}$ as a function of the logarithm of the energy.

- Divide the whole scatter plot in vertical energy slices. Here becomes important the use of $\log(E)$, instead of E: since the higher is the energy, the lower is the number of events, by dividing the histogram of $\log(E)$ in equally spaced slices we obtain wider regions as we increase the energy.

- For each of these slices produce an histogram of $\widehat{OFT}$.

- Since we expect $\widehat{OFT}$ to average to one by construction, we fit each of these histograms with a gaussian in a reasonable interval (if the fit doesn't converge due to low statistics effects, we take the mean and RMS from each events in the slice).

- Get the mean value m_j and variance σ_j from the interpolation on the j^{th} histogram.

- Apply the following transformation to any $\widehat{OFT}$ falling in the j^{th} histogram:
$$\widehat{OFT} \to \mathrm{PSA}_{\mathrm{nsigma}} \equiv \frac{\widehat{OFT} - m_j}{\sigma_j} \tag{C.11}$$

The effect of this procedure, applied to the OFT parameter, is shown in figure C.5.

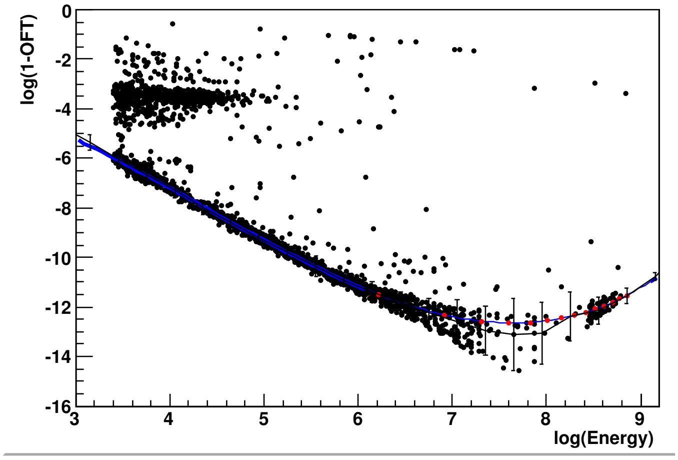
C.3 Pulse Shape Cuts

In order to obtain a procedure through which we can obtain linearized quantities as much automatically as possible, we followed this procedure:

- We select all the events from a single channel within a single data set.

- We perfom a rough cut (using the Wiener filter and mean values and RMS of the shape parameters) in order to produce a scatter plot of rise and decay time (RT and DT) versus energy with a controlled number of spurious events.

- We apply the linearization procedure, intepolating their distributions with a second order polinomial in order to obtain $\widehat{DT}$ and $\widehat{RT}$ from equation (C.6) (in the standard DIANA ntple, these quantities are labelled "linRiseTime" and "linDecayTime").



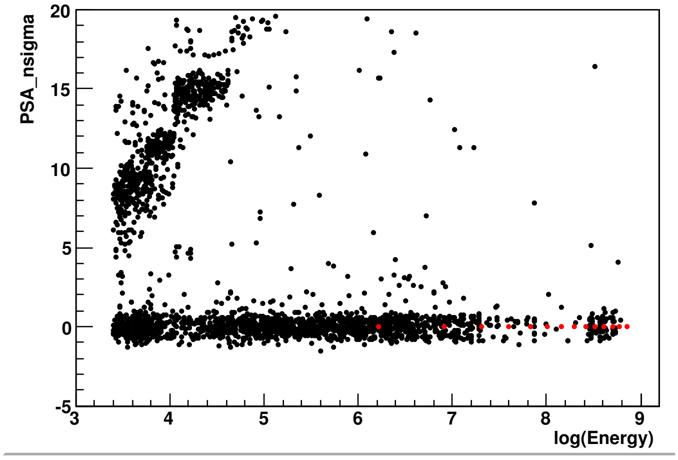


Figure C.5: Example of the process of linearization and normalization of the OFT parameter. The scatter plot is analyzed at different energy slices, where the mean values and variance are computed (top). After this process, the parameter becomes almost energy independent (bottom).The energy range is 30-7000 keV: red dots identify intervals of 500 keV.

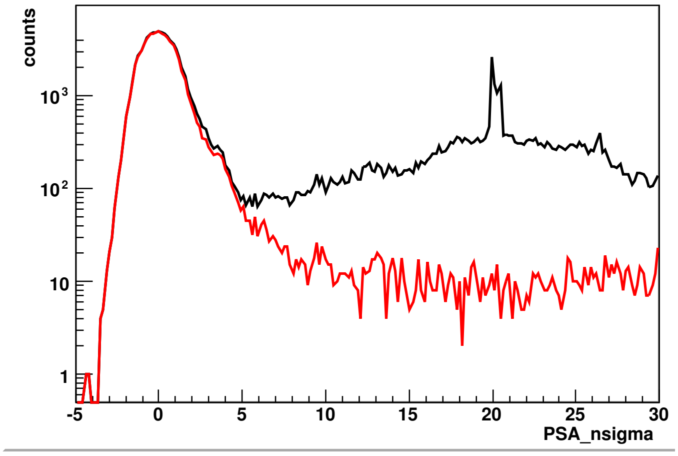


Figure C.6: Distribution of the variable $\mathrm{PSA}_{\mathrm{nsigma}}$ for channel 10 for the whole CUORICINO data (black curve). The red curve shows the effect of the cut on linearized rise and decay time, which leaves the left peak almost untouched.

- We use linearized rise time and decay time to identify spikes and "triangoloni" (which were difficult to identify with a rough cut) by selecting a belt in parameter shape in which we require these two parameters to have variations of 20% at most, compared to their average value.

- On these events we perfom the linearization and variance normalization in order to obtain $\mathrm{PSA}_{\mathrm{nsigma}}$.

Figure C.6 shows the distribution the variable $\mathrm{PSA}_{\mathrm{nsigma}}$ for all the events from channel 10 from all the CUORICINO statistics. The left peak corresponds to particle events, as shown by the red curve, obtained after applying the cut also on the linearized rise and decay times. In order to work with a distribution as much symmetrical as possible, we decided to apply the final "number of sigma" cut:

$$\mathrm{PSA}_{\mathrm{nsigma}} \leq 2$$

on all the CUORICINO channels at once. Figure C.7 shows the effect of this cut on the variable "linRiseTime" for channel 31 on data set 15.
In order to deal with the effect of these combined cuts we will need to deal with an efficiency measure.

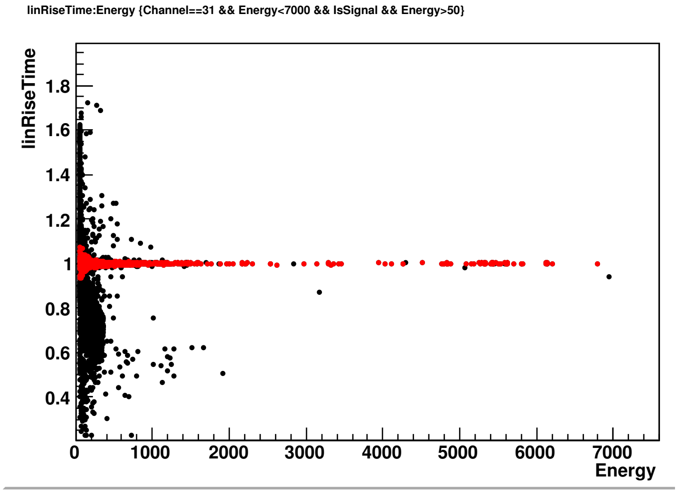


Figure C.7: Effect of the cut $\mathrm{PSA}_{\mathrm{nsigma}} \leq 2$ alone on channel 31 from data set 15.

C.4 Thresholds

The definition of threshold is useful to reduce spurious coincidences between true signals and triggered noise and also to reduce causal coincidenses between true pulses and cross-talks. At low energies good pulses and noise merge in a indistinguishable way so that even the study of the variation of pulse shape parameter distribution resulted unefficient for the definition of thresholds (at leas it resulted not simple to develop an authomatic algorithm to fix threshold). There are also occurrences where the trigger efficiency is reduced by unknown reasons (i.e. the trigger threshold unchanged but the detector triggering much less that in other runs, and - to a first check - the problem doesn't look as a noise rate variation).

Consequently definitions based on the variation of the $\mathrm{PSA}_{\mathrm{nsigma}}$ distribution with energy does not work, nor consistency with the average rate works because rates are not statistically compatible. What makes difficult the definition of a recipe for setting analysis threhsolds is not only the spread in the counting rate among crystals but also the variation of the counting rate with time due to Rn.

The only solution that we devised so far for the definition of thresholds was to compare the rate among crystals in different data sets. We assumed that the rate at low energy is dominated by contaminations that are similar for all the crystals. Two time variation of the rate are reasonably possible: a time decreasing rate (due to long-living activated isotopes not identificed but giving contribution at low energies) and Rn. The former seems not present, the latter has a clear signature: a correlation with Rn lines. We adopted therefore the following criteria for threshold definitions:

- define an acceptance interval for the rate of each channel/data set normalized to the average rate for that data set. This requires that rate time variations have the same behaviour for all channels since they are due to a ″global″ source: Rn. The acceptance interval was defined looking at a ″safe″ region i.e. above 600 keV.
- rise the thresholds of detectors that lie outside this acceptance interval
- check the result in terms of correlation with Rn

This technique was applied considering the following intervals 600-2000 keV, 200-80 keV and 80-30 keV. Whenever a channel was not satisfing the above mentioned criteria its threshold was rised to the lowest energy above which the criteria is satisfied. This definition of threshold is in someway arbitrary, however it does not bias the analysis provided that efficiencies are correctely evaluated.

Acknowledgements

This thesis would have never seen any light without the precious contributions or help I received in these three years. They have been tough years, shaken by so many deep changes in my life. It's indeed a precious gift, being able to look back and see all the support I've been given and all the changes I've been through thanks to the support of so many people.

I want to thank Prof. Maura Pavan, my advisor, at first. Considering my tendency on opening an infinite number of doors, she had the tough task of teaching me the importance of chosing a path while focusing on my purposes. It surely was difficult (for me as for her, I guess) but it is something I will bring with myself for the rest of my life.

I thank prof. Oliviero Cremonesi who has always been a model for me, not only as a physicist but, more importantly, as a human being. And I will never forget all the jazz lessons on the road to Gran Sasso too!

I thank the Milano group as a whole, for the continous support I recieved during these years. I'm especially grateful to Chiara Brofferio who has always shown such a deep care of us, Ezio Previtali for his curious and critical spirit and Luigi Zanotti for his inspiring power.

How can I forget about Maurizio Perego, for all the compelling discussions about philosophy, science, probabilitity, for all the coffees he offered me and for his WARP trigger.

I thank my collegues, Cecilia, Elena, Luca G., Luca P. and Matteo. We shared everything of this strange and peculiar work. Thanks for all the support but, especially, for all the laughters we shared.

I thank Maria Martinez, whose simple presence can turn a bad day into a shining one.

I thank every member of the CUORE/CUORICINO collaboration who I've been in touch with: Andrea G., for his aesthetic nerdiness, Claudia R., for sharing with me too much sun in Athens, Marco V., for the precious discussions and advices, and everyone else I'm forgetting (freudian alert!). A specific thank goes to Adam Bryant and Yury Kolomensky, for they have shown me that a collaboration can indeed be healthy.

I ringraziamenti più personali, li lascio alla mia lingua natia, vista la difficoltà nel trovare le giuste parole.
Ringrazio mia madre, che mi ha insegnato a vivere con gioia la mia vita e a scovare (anche ossessivamente, perchè no...) il bello in tutto ciò che mi circonda: questa qualità mi ha aiutato sempre ad affrontare con serenità le scelte che mi hanno permesso di arrivare sin qui.
Ringrazio i miei amici più cari: Simone, Marco e Alessia per tutto l'affetto che non smettono mai di dimostrarmi; Fabio e Benedetto per simboleggiare così bene il senso del mio passato, del mio presente e del mio futuro; Dario e Tommaso, perchè, nonostante tutte le maschere che mi metta, mi ricordano che ci sarà sempre qualcuno che mi conosce e che mi vuole bene per quel che sono.
Ringrazio mio nonno Franco, che con il suo ingegno è riuscito a farmi arrivare, proprio ora, un segno della sua presenza.
A Chiara va il ringraziamento più profondo: la felicità e la serenità che è riuscita e che continua a donarmi hanno sciolto così tanti nodi che per la prima volta in vita mia ho scoperto cosa vuol dire respirare.

Bibliography

[1] V. I. Tretyak and Y. G. Zdesenko, Atomic Data and Nuclear Data Tables **80**, 83 (2002). 2

[2] S. Elliott and P. Vogel, Annual Review of Nuclear and Particle Science **52**, 115 (2002). 2

[3] A. Morales and J. Morales, Nucl. Phys. B (Proc. Suppl.) **114**, 141 (2003). 2

[4] A. Strumia and F. Vissani, Neutrino masses and mixings, 2010. 2, 7, 9, 16

[5] G. L. Fogli *et al.*, Phys. Rev. D **70**, 113003 (2004). 2

[6] G. L. Fogli *et al.*, Phys. Rev. D **75**, 053001 (2007). 2

[7] S. Pascoli and S. Petcov, Phys. Rev. D **77**, 113003 (2008). 2

[8] F. Simkovic, A. Faessler, and V. Rodin, Phys. Rev. C **77**, 045503 (2008). 2, 116, 118

[9] O. Civitarese and J. Suhonen, JoP Conference Series **173**, 012012 (2009). 2, 116, 118

[10] J. Menendez, A. Poves, E. Caurier, and F. Nowacki, Nucl. Phys. A **818**, 139 (2009). 2, 17, 116, 118

[11] J. Barea and F. Iachello, Phys. Rev. C **79**, 044301 (2009). 2, 116, 118

[12] H. V. Klapdor-Kleingrothaus *et al.*, European Physics Journal **A12**, 147 (2001). 3, 50, 117, 118

[13] C. E. Aalseth *et al.*, Phys. Rev. D **65**, 092007 (2002). 3

[14] C. E. Aalseth *et al.*, Phys. Rev. D **70**, 078302 (2004). 3

[15] A. S. Barabash, Presentation at the XXXIII international conference on high energy physics, Moscow, July 26 - August 02, 2006. 3, 50, 117, 118

[16] C. Arnaboldi *et al.*, Phys. Rev. C **78**, 35502 (2008). 3, 36, 38, 40, 41, 49, 50

[17] R. Bernabei *et al.*, Physics Letters B **546**, 23 (2002). 3, 50, 117, 118

[18] H. V. Klapdor-Kleingrothaus, I. V. Krivosheina, A. Dietz, and O. Chkvorets, Mod. Phys. Lett. **A16**, 2409 (2001). 3, 14, 50, 113, 117, 118

[19] F. Halzen and A. D. Martin, *Quarks and Leptons* (John Wiley & Sons, 1984). 6, 7

[20] M. Goeppert-Mayer, Phys. Rev. **48**, 512 (1935). 11

[21] J. Schechter and J. W. F. Valle, Physics Review **D25**, 2951 (1982). 13, 15

[22] R. N. Mohapatra and P. B. Pal, *Massive Neutrinos in Physics and Astrophysics -3rd edition* (World Scientific, 2004). 13

[23] V. R. arXiv:0910.5866, 2009. 18

[24] E. Caurier, G. Martinez-Pinedo, F. Nowacki, A. Poves, and A. P. Zucker, 2005. 18

[25] G. F. Knoll, *Radiation Detection and Measurement* (John Wiley and Sons, 2000). 20

[26] E. Booth, B. Cabrera, and E. Fiorini, Annu. Rev. Nucl. Part. Sci. **46**, 471 (1996). 20

[27] N. Ashcroft and N. Mermin, *Solid State Physics* (Brooks Cole, 1976). 24

[28] J. Hubbard, Proc. Roy. Soc. A **276**, 238 (1963). 25

[29] J. Zhang *et al.*, Phys. Rev. B **57**, 4472. 26, 30

[30] J. Mather, Appl. Opt 21 **6**, 1125 (1982). 26

[31] M. Barucci *et al.*, Journal of Low Temperature Physics **123**, 303 (2001). 31

[32] A. A. et al., Czechoslovak Journal of Physics **46-S5**, 2893 (1996). 31

[33] P.Gorla, Optimization of the cuore detector performances - PhD thesis, 2005. 31, 37

[34] M. Redshaw, B. J. Mount, E. G. Myers, and F. T. A. III, Phys. Rev. Lett. **102**, 212502 (2009). 35, 50, 96

[35] M. Pavan *et al.*, European Physics Journal A **36**, 159 (2008). 36

[36] C. Bucci *et al.*, European Physics Journal A **41**, 155 (2009). 36

[37] C. Arnaboldi *et al.*, Phys. Lett. B **557**, 167 (2003). 36

[38] E. Gatti and P. F. Manfredi, Rivista del Nuovo Cimento **9**, 1 (1986). 39, 55, 65

[39] C. Arnaboldi *et al.*, IEEE Trans. Nucl. Sci. **49**, 2440 (2002). 40

[40] A. Alessandrello *et al.*, Nucl. Instr. Meth. A **412**, 451 (1998). 40

[41] M. Vignati, Journal of Applied Physics **108**, 084903 (2010). 41, 53, 74

[42] G. J. Feldman and R. D. Cousins, Physical Review D **57**, 3873 (1998). 47, 97, 119

[43] G. Cowan, *Statistical Data Analysis* (Oxford University Press, 1998). 49

[44] C. Arnaboldi *et al.*, Phys. Rev. Lett. **95**, 14501 (2005). 49

[45] N. D. Scielzo *et al.*, Physical Review C **80**, 025501 (2009), arXiv:0902.2376v1 [nucl-ex]. 50

[46] C. Arnaboldi *et al.*, Phys. Rev. C **78**, 156 (2008). 81

[47] S. Baker and P. D. Cousins, Nucl. Instrum. Methods Phys. Res. A **221**, 457 (1984). 82

[48] C. Collaboration, Lngs annual report, 2006. 87

[49] D. S. Sivia, *Data Analysis - a Bayesian tutorial* (Oxford Science Publications, 1996). 101

[50] G. D'Agostini, *Probabilistic reasoning in HEP - principles and applications* (Report CERN 99-03, 1999). 111, 113

[51] H. Jeffreys, *Theory of probability* (Oxford University Press, 1961). 111

[52] C. Arnaboldi *et al.*, Nucl. Instr. Meth. A **518**, 775 (2004). 117

[53] A. Frodesen, *Probability and Statistics in Particle Physics* (Universitetsforlaget, 1979). 120, 123

[54] J. Carson, Bell Syst. Techn. J. **10**, 374 (1931). 124

[55] A. Oppenheim and R.W.Schafer, *Digital Signal Processing* (Prentice Hall, 1975). 128