Journal of Physics: Conference Series

Off-line data processing and analysis for the GERDA experiment

To cite this article: M Agostini *et al* 2012 *J. Phys.: Conf. Ser.* **368** 012047

Off-line data processing and analysis for the GERDA experiment

M Agostini[1], L Pandola[2] and P Zavarise[2,3]

[1] Physik Department E15, Technische Universität München, Garching, Germany
[2] INFN, Laboratori Nazionali del Gran Sasso, Assergi (AQ), Italy
[3] Dipartimento di Fisica, Università dell'Aquila, L'Aquila, Italy

E-mail: matteo.agostini@ph.tum.de, pandola@lngs.infn.it, zavarise@lngs.infn.it

Abstract. Gerda is an experiment designed to look for the neutrinoless double beta decay of ^{76}Ge. The experiment uses an array of high-purity germanium detectors (enriched in ^{76}Ge) directly immersed in liquid argon. Gerda is presently operating eight enriched coaxial detectors (approximately 15 kg of ^{76}Ge) and about 25 new custom-made enriched BEGe detectors will be deployed in the next phase (additional 20 kg of ^{76}Ge). The paper describes the Gerda off-line analysis of the high-purity germanium detector data. Firstly we present the signal processing flow, focusing on the digital filters and on the algorithms used. Secondly we discuss the rejection of non-physical events and the data quality monitoring. The analysis is performed completely with the Gerda software framework (Gelatio), designed to support a multi-channel processing and to perform a modular analysis of digital signals.

1. Introduction

Gerda [1, 2] is a low-background experiment designed to search for the neutrinoless double beta decay of ^{76}Ge using an array of bare high-purity germanium (HPGe) detectors isotopically enriched in ^{76}Ge. The detector array is operated directly in ultra radiochemically-pure liquid argon, allowing a substantial background reduction at the $Q_{\beta\beta}$-value of ^{76}Ge (2039 keV) with respect to the previous experiments [3–5]. The experiment pursues a staged implementation. In the present phase (Phase I) eight enriched coaxial detectors are being used (approximately 15 kg of ^{76}Ge). In Phase II, about 25 new custom-made enriched BEGe detectors [6] will be deployed providing additional ∼20 kg of ^{76}Ge. The experiment is located in the underground Laboratori Nazionali del Gran Sasso of the INFN (Italy).

If no background events are observed in the region of interest, the sensitivity of Gerda will scale with the product of the ^{76}Ge mass and the exposure time, $M \cdot t$. Otherwise, the sensitivity will scale with the square root of this product divided by the background index (BI) and the energy resolution (ΔE) in the region of interest, $\sqrt{(M \cdot t)/(BI \cdot \Delta E)}$. To enhance the experimental sensitivity, an advanced off-line analysis of the HPGe detector signals is performed. Digital filters are applied to reconstruct the event energy with high resolution. In addition the analysis of the signal time-structure is a powerful tool to identify background events [7–10].

To ease this kind of analysis a new software framework (Gelatio) has been recently developed [11]. The framework is presently used to handle the full data analysis flow of Gerda and the R&D activities related to the experiment. It is implemented in C++ and is based on the MGDO library [12]. The framework implements a modular analysis. Each module handles a precise and self-consistent task of the signal processing and is implemented as a dedicated C++ class. The output of the modules, which is either a scalar parameter (e.g. the amplitude of the signals) or a shaped trace, can be used as input for other modules and/or stored to disk. The users can easily create chains of modules and customize the parameters of each algorithm. The software supports also the processing of multi-channel data. While the analysis is performed event-per-event in order to keep the output synchronized, each channel can be processed along its own dedicated and customized chain of modules.

The framework was used for the reference analysis of the data acquired in the GERDA commissioning phase (from June 2010 to October 2011), when up to seven HPGe detectors have been operated simultaneously. The commissioning data were used as a benchmark to validate GELATIO against other independent analysis codes and to prove its suitability for the use in GERDA Phase I [11].

This paper describes the basic off-line analysis of the GERDA data performed with the GELATIO framework. In section 2 we present the flow of the signal processing and analysis along the module chains and the shaping algorithms. Then, in section 3, we discuss the identification of non-physical events or of signals not properly processed along the analysis pipeline. Also, the data quality monitoring will be described. Finally, summary and conclusions are presented in section 4.

2. Signal processing flow

The charge signals of the HPGe detectors deployed in GERDA are composed by pulses with a $\sim 1\,\mu$s long rising part followed by an exponential decay tail ($\sim 100\,\mu$s decay time) folded by the front-end charge-sensitive preamplifier. The rising part (leading edge) is induced by the drift of electron-hole pairs – generated by γ-ray or charge particle interactions – toward the detector electrodes and it contains information concerning the event topology.

The signals are digitized at 100 MHz sampling frequency by 14-bit flash-ADCs (FADC) equipped with integrated anti-aliasing bandwidth filters [13–15]. For each event, the FADC computes in run-time and writes to disk two traces. The first trace is sampled at 100 MHz and is $4\,\mu$s long (high-frequency-short trace). It is used for identifying background events through the analysis of the signal time-structure. The second trace is sampled at 25 MHz and is $160\,\mu$s long (low-frequency-long trace). It is used for those operations, as energy reconstruction, which involve the integration of the pulse.

The two traces are processed along different chains of GELATIO modules, as shown in figure 1. The first module, GEMDTop, extracts from the input file the traces requested by the following modules. Before publishing the traces the module checks, and possibly changes, the pulse polarity in order to always have positive-polarity pulses. The read-out traces are the starting point for two chains: the low-frequency-long trace is processed along Chain1 while the high-frequency-short trace along Chain2.

Chain1 starts with GEMDBaseline. This module analyzes the baseline of the signal by computing the average value, the root-mean-square deviation (RMS) and the linear slope before the leading edge. In addition, the module performs a baseline restoration — a subtraction of the average baseline value to the trace — and provides the new signal to the other modules:

- GEMDTrigger. The module is used to identify the beginning of the pulse leading edge (trigger position). It implements a leading-edge discriminator with threshold defined dynamically as three times the RMS of the signal baseline. After the trigger, the signal has to remain above threshold for at least $40\,\mu$s, otherwise the trigger is rejected.
- GEMDFTTrigger. While the previous module is tuned to determine the trigger position with high precision and stability, this module is used to identify events with multiple physical pulses occurring within the same trace. The module applies to the input signal a $1.5\,\mu$s

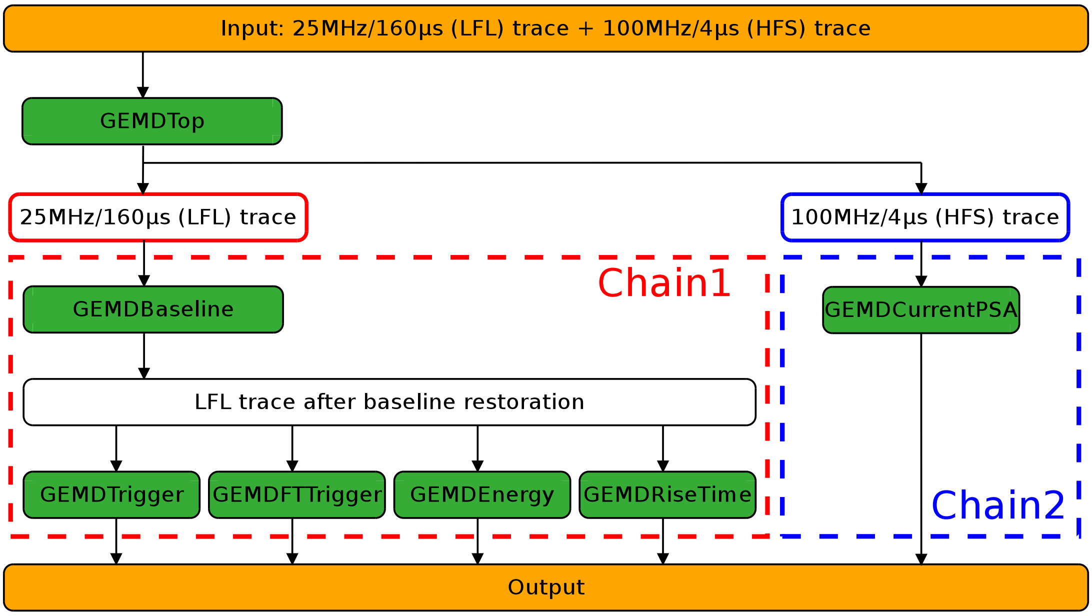

Figure 1. Flow chart of the signal processing. The two traces saved by the digitizer are processed along two different chains of GELATIO analysis modules. The low-frequency-long (LFL) trace is used for reconstructing the energy, the trigger and the rise time. The high-frequency-short (HFS) trace is used to analyse the time-structure of the signals.

moving differentiation filter and a 1 μs moving average filter for noise reduction (see figure 2). The resulting trace has a peak for each sharp variation of the signal (such as the leading edge of a pulse). The peak width is similar to the size of the moving differentiation and was chosen to maximize the pile-up identification efficiency and to avoid the misidentification of highly-multiple-site events. The number and the position of the peaks are estimated by applying a leading-edge discriminator, whose threshold is four times the RMS of the baseline. After this condition is met, the signal has to remain above the threshold for at least 1 μs.

- GEMDEnergyGauss. The module reconstructs the event energy using an approximate Gaussian filter [16, 17]. The pulse is differentiated by a moving differentiation filter and then integrated 15 times by a moving average filter to achieve an approximated Gaussian shape[4]. The energy information is eventually stored in the maximum amplitude of the quasi-Gaussian pulse. The width of the moving filters has been set to 10 μs in order to minimize losses due to ballistic effects. The intermediate steps of the shaping are shown in figure 3.
- GEMDRiseTime. The module computes the rise time between 10% and 90% of the maximum amplitude of the pulse. The maximum amplitude is computed as the difference between the maximum of the pulse and the average baseline value. Then, the first samples below the 10% and 90% of the maximum amplitude are found by moving backwards from the position of the maximum.

The second chain is used to evaluate parameters relevant for the identification of background events and it will be better defined during the future data taking. The chain presently includes

[4] Historically, the energy reconstruction filters for γ-ray spectroscopy also perform a deconvolution of the exponential function which is folded in the signal by the charge sensitive preamplifier. However, the approach suggested in this paper was found to provide better results on the GERDA data with respect to the usual filters.

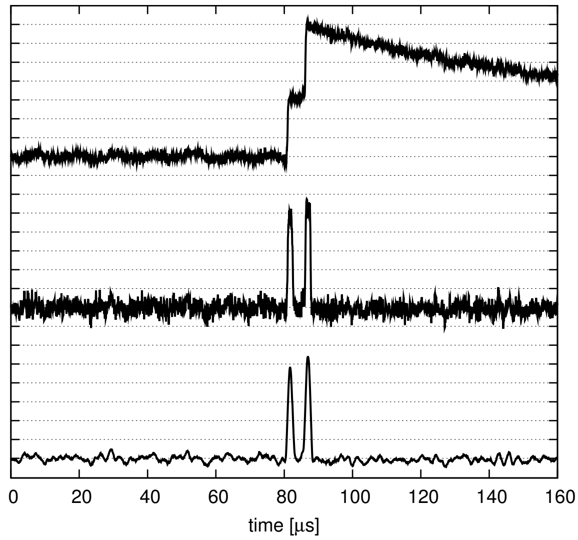

Figure 2. Digital signal processing performed by GEMDFTTrigger. The incoming trace (top trace) is differentiated (middle trace) and integrated (bottom). The output trace has a peak for each pulse in the incoming trace. The illustrative input trace contains two physical pulses with leading edge shifted by $\sim 5\,\mu$s.

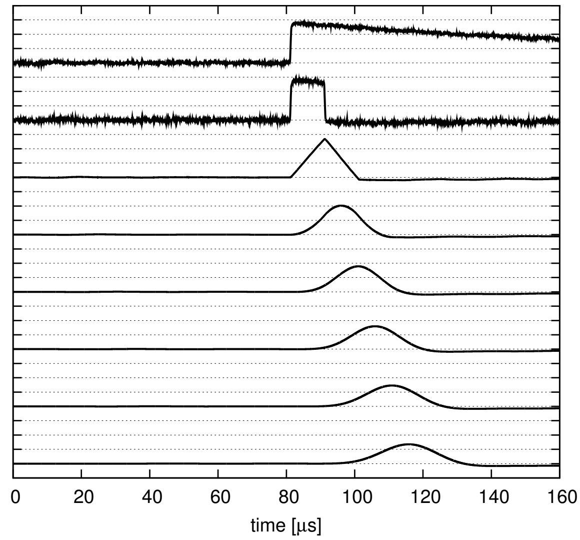

Figure 3. Digital signal processing performed by GEMDEnergyGauss. The incoming signal (top trace) is differentiated (second trace) and then integrated several times (following traces) by a moving average filter. The output signal has a Gaussian shape and its maximum is proportional to the event energy.

only one module, GEMDCurrentPSA, which computes the current signal as the derivative of the charge signal and then extracts the basic features of the current pulse, like rise time, width and area.

3. Data selection and monitoring

In the Gerda data sets there are two main classes of signals that have to be identified and tagged: 1) signals corrupted or produced by non-physical events, i.e. discharges, cross-talk, pick-up noise; 2) signals which are not properly processed along the analysis pipeline, as pile-ups and accidental coincidences.

The first class includes signals with anomalous shape, wrong polarity, extremely short/long rise time or exceeding the dynamic range of the FADC (see figure 4). To identify these events a sequence of cuts based on four parameters is applied. The first parameters are the trigger position computed by GEMDTrigger and the time position of the maximum amplitude of the Gaussian pulse (maxAmpTime) computed by GEMDEnergyGauss. If the signal has a leading edge at the proper position, followed by an exponential decay tail, then the trigger has to be reconstructed roughly in the center of the trace and maxAmpTime has to be in a well-defined range. The third parameter is the 10-90% rise time which can be used to identify signals that are extremely fast or slow, and hence inconsistent with well-behaved physical events. Finally, signals that saturate the dynamic range of the FADC are identified by scanning the individual samples of the traces.

The second class includes signals generated by the superposition of multiple physical pulses, or having the leading edge not aligned with the center of the trace (see figure 5). These signals can be identified using the baseline slope computed by GEMDBaseline, the number of pulses provided by GEMDFTTrigger and the trigger position yielded by GEMDTrigger. Their amount is proportional to the event rate and can reach up to 15% in the calibration data sets, while it

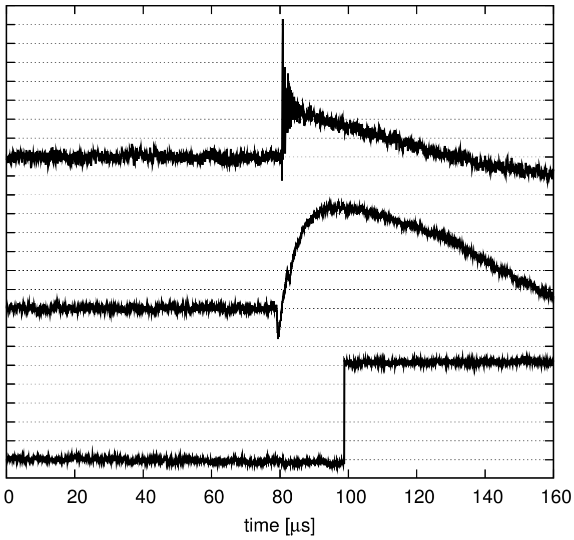

Figure 4. Illustrative traces generated by non-physical events. Note that these pulses do not have the typical exponential decay tail after the leading edge.

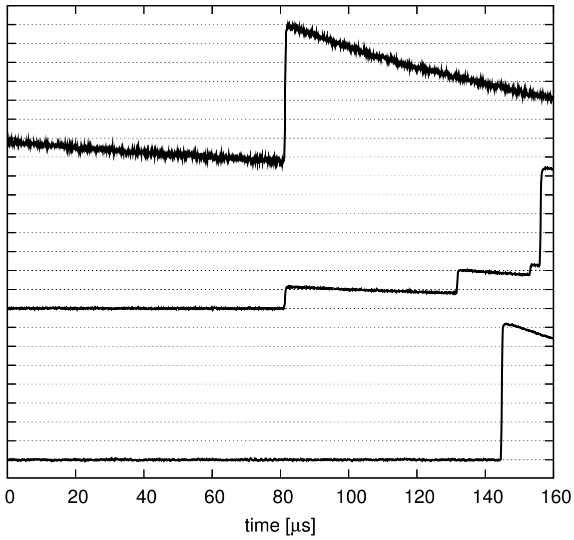

Figure 5. Example of traces generated by pile-up events (top-middle trace) and accidental coincidences (bottom).

is usually negligible in the physics data sets.

The identification of this second class of events is of primary importance for GERDA since the background rejection techniques based on the analysis of the signal time-structure are developed and tuned by using calibration data. Therefore the calibration data have to be filtered to extract samples of events as similar as possible to the physics run data and a small inefficiency could bias the following analysis. The result of the data selection applied to a calibration run with a ^{228}Th source is shown in figure 6. The cuts remove efficiently bad signals and pile-up events, improving the shape of the γ-ray peaks and the agreement with the standard analytical functions used to model γ-line peaks.

Besides cuts for removing undesirable classes of signals, there are also parameters which can be used to monitor the quality of the data taking and the stability of the set-up, the most important being the average value and the RMS of the baseline. These parameters are sensitive to noise changes and to gain variations in the read-out chain. Figure 7 shows these parameters as a function of time for a 10-day commissioning run. The parameters are stable over the whole data taking, except for a few hours during day 7. These instabilities can be correlated with hardware operations in the set-up and the corresponding data can be removed by applying a cut on the two parameters.

4. Conclusions

The GERDA experiment is currently starting the data taking of Phase I. The off-line analysis of the HPGe detector signals will be performed with the GELATIO software framework, a tool specifically developed for GERDA which supports a multi-channel analysis and implements signal processing chains based on a modular approach. A reference analysis pipeline has been defined and optimized. The signal processing is performed along chains of modules and includes the estimate of the trigger position, the amplitude and several basic pulse shape analysis parameters. The digital filters have been improved and optimized during the GERDA commissioning phase and the shaping parameters have been tuned. Also, a set of cuts has been defined to identify

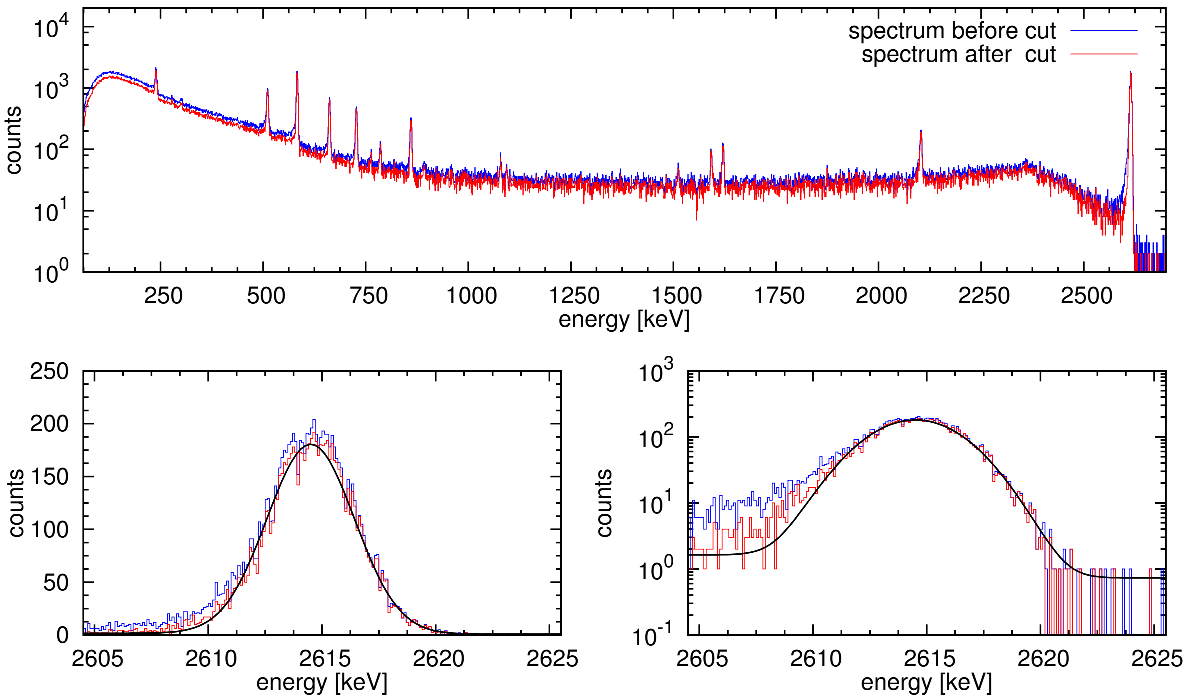

Figure 6. Energy spectrum reconstructed by an enriched detector deployed in the GERDA set-up during a calibration run with a ^{228}Th source. The cuts remove approximately 15% of the total events (∼10% in the γ lines). The bottom panels show the peak of the 2.614 MeV γ-ray line in linear and logarithmic scale. The reduction of the tail due to pile-up events improves significantly the fit with the analytical model (the χ^2 per degree of freedom is significantly reduced, from 1.77 to 0.99).

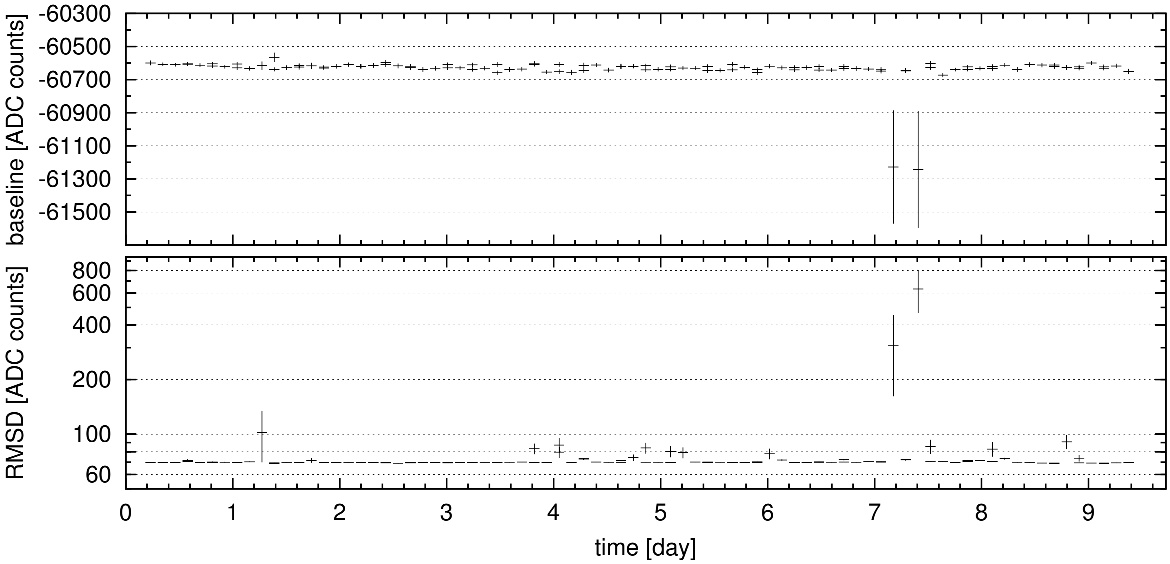

Figure 7. Average value (top panel) and RMS (bottom panel) of the signal baseline vs. time for a detector operated in GERDA during a 10-day run. The bin content represents the mean value of the parameter and the error bars are related to the width of the distribution. The bins are 2.4 hours wide.

signals induced by non-physical events or signals which are not properly processed. In addition, a set of parameters was identified to monitor the data quality and possibly to discriminate corrupted data.

The software and the digital filters have been validated during the GERDA commissioning and in several R&D activities related to the experiment. The new pipeline has been tested and it proved to be stable and ready to be used for the reference analysis of GERDA Phase I data.

Acknowledgments

We would like to acknowledge our colleagues of the GERDA Collaboration, especially D. Budjáš and B. Schwingenheuer, for many invaluable advices concerning signal analysis and data quality control. We want also to thank D. Bazzacco, C. A. Ur and T. Kihm for all the stimulating discussions and suggestions concerning digital signal processing and the algorithms used in γ-ray spectroscopy.

This work was supported in part by the Transregio Sonderforschungsbereich SFB/TR27 "Neutrinos and Beyond" by the Deutsche Forschungsgemeinschaft and by the Munich Cluster of Excellence "Origin and Structure of the Universe".

References

[1] Abt I *et al.* (GERDA Collaboration) 2004 GERDA: The GERmanium Detector Array for the Search of Neutrinoless $\beta\beta$ Decay of ^{76}Ge at LNGS *Laboratori Nazionali del Gran Sasso - Proposal*

[2] Schönert S *et al.* (GERDA Collaboration) 2005 The GERmanium Detector Array (GERDA) for the search of neutrinoless beta beta decays of Ge-76 at LNGS *Nucl. Phys. Proc. Suppl.* B **145** 242-245

[3] Gunther M *et al.* 1997 Heidelberg - Moscow beta-beta experiment with Ge-76: Full setup with five detectors *Phys. Rev.* D **55** 54-67

[4] Klapdor-Kleingrothaus H V, Krivosheina I V, Dietz A and Chkvorets O 2004 Search for neutrinoless double beta decay with enriched Ge-76 in Gran Sasso 1990-2003 *Phys. Lett.* B **586** 198-212

[5] Aalseth C E *et al.* (IGEX Collaboration) 2002 The IGEX Ge-76 neutrinoless double beta decay experiment: Prospects for next generation experiments *Phys. Rev.* D **65** 092007

[6] CANBERRA Broad Energy Ge (BEGe) Detector, URL http://www.canberra.com/products/485.asp

[7] Budjáš D, Barnabé Heider M, Chkvorets O, Khanbekov N and Schönert S 2009 Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double beta decay experiment *JINST* **4** P10007 (*Preprint* arXiv:0909.4044)

[8] Agostini M *et al.* 2011 Signal modeling of HPGe detectors with a small read-out electrode and application to neutrinoless double beta decay search in Ge-76 *JINST* **6** P03005 (*Preprint* arXiv:1012.4300)

[9] Gonzalez D *et al.* 2003 Pulse shape discrimination in the IGEX experiment *Nucl. Instrum. Meth.* A **515** 634-643

[10] Hellmig J and Klapdor-Kleingrothaus H V 2000 Identification of single-site events in germanium detectors by digital pulse shape analysis *Nucl. Instrum. Meth.* A **455** 638-644

[11] Agostini M, Pandola L, Zavarise P and Volynets O 2011 GELATIO: a general framework for modular digital analysis of high-purity Ge detector signals *JINST* **6** P08013 (*Preprint* arXiv:1106.1780)

[12] Agostini M *et al.* 2011 The MGDO software libraries for Germanium neutrinoless double beta decay experiments to appear in *J. Phys.: Conf. Series* XII International Conference on Topics in Astroparticle and Underground Physics (TAUP 2011), September 5-9 2011, Munich, Germany

[13] MIZZI Computer Software GmbH, URL www.mizzi-computer.de

[14] Isocrate R *et al.* 2004 MD2S - Digital sampling electronics for the MARS detector, A data acquistion system for the MD2S digital sampling electronics *Laboratori Nazionali di Legnaro - Annual Report* 226-227

[15] Ur C A *et al.* 2004 A data acquistion system for the MD2S digital sampling electronics *Laboratori Nazionali di Legnaro - Annual Report* 228-229

[16] Smith S W 1999 *The Scientist and Engineer's Guide to Digital Signal Processing* (California: California Technical Publishing San Diego)

[17] Van Trees H L 1968 *Detection, Estimation, and Modulation Theory* (New York: John Wiley and Sons)