Development of Pulse Shape Discrimination Methods as Tools for Background Suppression in High Purity Germanium Detectors used in the GERDA Experiment

Doctoral dissertation by

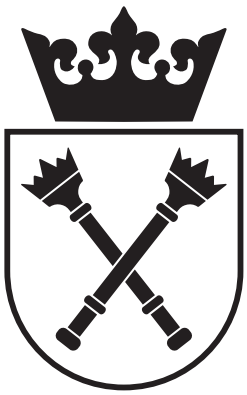
Krzysztof Panas

prepared under the supervision of

Prof. Marcin Wojcik

and co-supervision of

Dr Grzegorz Zuzel



M. Smoluchowski Institute of Physics
Jagiellonian University in Cracow
2018

Wydział Fizyki, Astronomii i Informatyki Stosowanej
Uniwersytet Jagielloński

Oświadczenie

Ja niżej podpisany Krzysztof Panas (nr indeksu: 1110101), doktorant Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego, oświadczam, że przedłożona przeze mnie rozprawa doktorska pt. „Development of Pulse Shape Discrimination Methods as Tools for Background Suppression in High Purity Germanium Detectors used in the GERDA Experiment"jest oryginalna i przedstawia wyniki badań wykonanych przeze mnie osobiście, pod kierunkiem prof. dr hab. Marcina Wójcika. Pracę napisałem samodzielnie.

Oświadczam, że moja rozprawa doktorska została opracowana zgodnie z Ustawą o prawie autorskim i prawach pokrewnych z dnia 4 lutego 1994 r. (Dziennik Ustaw 1994 nr 24 poz. 83 wraz z późniejszymi zmianami).

Jestem świadom, że niezgodność niniejszego oświadczenia z prawdą ujawniona w dowolnym czasie, niezależnie od skutków prawnych wynikających z ww. ustawy, może spowodować unieważnienie stopnia nabytego na podstawie tej rozprawy.

Kraków, dnia . . . . . .

. . . . . .
(podpis doktoranta)

Abstract

Neutrinoless double beta decay experiments are well recognized as the most powerful probes of the nature of the neutrinos (whether they are Dirac or Majorana particles), their absolute mass scale and the neutrino mass hierarchy. Observation of the decay would require physics beyond the Standard Model and it is the only process, which allows to experimentally the limit effective neutrino masses down to the meV level.

The GERDA – GERmanium Detector Array – experiment searches for the neutrinoless double beta decay decay of the ^{76}Ge isotope. It uses a matrix of High Purity Germanium detectors enriched in ^{76}Ge from the natural abundance of 7.8% to about 86%. In the searches for rare nuclear processes, like the above mentioned decay, one of the critical parameters of the detectors is their background, which needs to be reduced to an unprecedentedly low level, making the experiment practically "background-free". Pulse Shape Discrimination methods, developed in the frame of this study, are essential to achieve that goal.

The presented work covers development and application of the technique to the GERDA Phase II data, as well as to γ-ray spectrometers, for which the decrease of the Compton continuum background increases their sensitivity. The method was tested on the germanium detectors with different geometries, namely semi-coaxial and Broad Energy Germanium, yielding very good results in both cases.

Streszczenie

Detektory wykorzystywane do poszukiwania podwójnego bezneutrinowego rozpadu beta są uważane za najczulsze narzędzia do badania natury neutrina (cząstka Majorany czy Diraca) oraz bezwzględnej skali jego masy. Wyniki eksperymentów poszukujących tego procesu być może pozwolą na określenie hierarchi mas neutrin. Obserwacja tego procesu oznaczałaby, iż mamy do czynienia z fizyką spoza Modelu Standardowego, a pomiar jego czasu połowicznego rozpadu pozwoliłby na ograniczenie efektywnej masy neutrina na poziomie pojedynczych meV.

Eksperyment GERDA (GERmanium Detector Array) został zaprojektowany do poszukiwań podwójnego bezneutrinowego rozpadu beta izotopu ^{76}Ge. W tym celu użyta jest matryca detektorów germanowych, wzbogaconych w ^{76}Ge do ok. 86%. Przy poszukiwaniu tak rzadkich procesów jak wspominany rozpad, jednym z krytycznych parametrów detektora jest jego tło, które musi być zredukowane do niespotykanie niskiego poziomu, czyniąc eksperyment praktycznie "beztłowym".

Metody analizy kształtu impulsu, opracowane w ramach niniejszej pracy, są niezbędne do osiągnięcia tego celu. Rozprawa opisuje proces opracowania tych metod oraz ich zastosowanie do danych z drugiej fazy eksperymentu GERDA, oraz do spektrometrów γ, w których analiza kształtu impulsu pozwala na obniżenie tła komptonowskiego oraz zwiększenie ich czułości. Opracowane techniki zostały przetestowane zarówno na danych z detektorów semi-koasjalnych jak i typu Broad Energy Germanium, uzyskując bardzo dobre rezultaty w obu przypadkach.

Contents

Introduction

> *Skepticism is like a microscope whose magnification is constantly increased: the sharp image that one begins with finally dissolves, because it is not possible to see ultimate things: their existence is only to be inferred.*

Stanisław Lem, *His Master's Voice*, 1968
(translation by Michael Kandel)

The pursuit of the nature of neutrinos, often called the "elusive particles" due to their weak interaction with matter, can answer many questions of particle physics. First proposed by WOLFGANG PAULI as massless, they soon surprised the physicists with the manifestation of an oscillation mechanism, that could be explained only if they are in fact massive. The Pontecorvo–Maki–Nakagawa–Sakata matrix relates neutrino mass eigenstates (usually numbered 1, 2 and 3 [1, 2]) with flavor states (e, μ, τ). It allows for the calculation of a probability to find the flavor components of each mass eigenstate [1]. There are also some open questions regarding the nature of neutrinos, like whether they are Dirac or Majorana particles, or if the total lepton number is a conserved quantity. They could be answered by the observation of the neutrinoless double-beta ($0\nu\beta\beta$) decay, a process beyond the Standard Model. Furthermore, it can resolve whether the neutrino mass eigenstates follow a "normal" or "inverted" hierarchy, i.e. the sign of the difference of the squared mass eigenstates 2 and 3 $\Delta m^2_{23} = m^2_2 - m^2_1$ (it is already known that $\Delta m^2_{12} = m^2_1 - m^2_2 > 0$ [3]). The positive sign indicates the normal hierarchy, while negative the inverted one. On the other hand, the possible observation of lepton number violation in the $0\nu\beta\beta$ decay would support the theory that the leptons have their share in the creation of the matter-antimatter asymmetry observed in the Universe [1].

Presently, the most stringent limit for the half-life of the $0\nu\beta\beta$ decay ($T^{0\nu\beta\beta}_{1/2}$) comes from the KamLAND-Zen experiment, which is looking for the decay of ^{136}Xe isotope ($T^{0\nu\beta\beta}_{1/2} > 1.07 \cdot 10^{26}$ yr [4]). In the case of ^{76}Ge, the best limit was achieved in the GERDA experiment (described in detail in Chap. 1). The chapter contains also a short history of the $0\nu\beta\beta$ decay searches. Since double-beta decays are second-order nuclear processes, they are characterized with the extremely long half-lifes $T^{2\nu\beta\beta}_{1/2}$. For ^{76}Ge it is $1.84 \cdot 10^{21}$ yr [5]. The $0\nu\beta\beta$ decay half-life $T^{0\nu\beta\beta}_{1/2}$ shall be orders of magnitude longer (with the present limit of $8 \cdot 10^{25}$ yr [6]) and therefore only a handful of $0\nu\beta\beta$ decay events are expected to be observed in the experiment. To put it into perspective, for a 5 yr long experiment, with 100 kg of ^{76}Ge, only ≈ 27 events are expected if $T^{0\nu\beta\beta}_{1/2}$ is equal to 10^{26} yr. It is therefore of highest importance to obtain the lowest possible radioactive background. Several techniques exist to achieve this goal, e.g.: construction materials prescreening (with $\alpha/\beta/\gamma$ spectrometers) and passive/active[1] shielding of the detectors.

[1] In GERDA liquid argon (LAr) is used as an active shield, since aside from passive γ-rays attenuation it also provides a veto signal from the LAr scintillation light. Since the $0\nu\beta\beta$ decay is takes place in the detector, there is no ionization of LAr (in 92% of all decays [7]). A LAr anti-coincidence cut is therefore used as a tool to reduce the background.

Another technique, that can be applied after data acquisition, is Pulse Shape Discrimination (PSD). The idea is to determine the event topology through an analysis of the detector's pulse shape, since the $0\nu\beta\beta$ decay deposits all of its energy in a small volume (single site) inside the detector, while high energy γ-rays usually do it in multiple sites. Chap. 2 begins with a derivation of simple analytical models of High Purity Germanium (HPGe) detectors for the true-coaxial and hemispherical geometries. Using the models it can be easily shown how the above mentioned differences in energy deposition topology result in different pulse shapes. The principles of the PSD technique, as well as previous methods developed in the field are also included in the chapter.

The main topic of this thesis is to estimate the efficiency of the PSD methods based on multivariate classifiers. In this context, the multivariate adjective refers to the usage of multiple input variables extracted from the digitized trace from a detector. This is in contrast to the single variable classifier, like e.g. A/E (also described in Chap. 2), which takes into account only one parameter derived from a waveform. The extraction is done for two groups of events, one representing the "background" and the other "signal" sample. Then, the variables are fed into the neural-network or Projective Likelihood based algorithms. The procedure is described in detail in Chap. 3, along with the measurements performed for the various types of the HPGe detectors. Aside from the applications in the $0\nu\beta\beta$ decay searches, it is also possible to increase the sensitivity of HPGe-based γ spectrometers via PSD [8]. This was done for two types of the detectors: an n-type semi-coaxial and a p-type with Broad Energy Germanium (BEGe) geometry. It should be noted that these two detector geometries are also used in the GERDA experiment. In fact, BEGe detectors were introduced in Phase I of the experiment due to their excellent PSD capabilities.

On December 20$^{\text{th}}$ 2015 GERDA experiment started its Phase II with 37 ^{76}Ge-enriched detectors, making up 35.6 kg of the total mass [9]. Chap. 4 includes the analysis of the Phase II data, acquired up to the 17$^{\text{th}}$ of June 2016. Several variants of the analysis were tested – in the case of the BEGe detectors dataset, a Multi-Layer Perceptron neural-network based method was used. Semi-coaxial detectors were analyzed with the Projective Likelihood classifier, the same method as the one developed for Phase I of the experiment by the Jagiellonian University group. Additionally, a novel approach was developed for the dimensionality reduction with Principal Component Analysis – it shows the increased efficiency in the discrimination of the γ-induced background. Lastly, the classifier performance for semi-coaxial detectors was tested using the Monte Carlo data, including the calculation of the efficiency for the $0\nu\beta\beta$ decay. Since the reliable testing data for $0\nu\beta\beta$ decay is not available (due to the different energy range or events topology), Monte Carlo simulations are currently the only way to obtain it. Discrepancies between Phase II data and simulation, as well as volumetric distribution of the events vetoed by PSD is also discussed. Tables with detailed results for the detectors in GERDA Phase II were collected in App. A. App. B contains the derivation of a mathematical model of the electronics response, which was applied to the Monte Carlo data to take into account the effect of a preamplifier.

Chapter 1

Search for the $0\nu\beta\beta$ decay with the GERDA experiment

The GERDA (GERmanium Detector Array) experiment was proposed in 2004 to probe the neutrinoless double-beta ($0\nu\beta\beta$) decay in ^{76}Ge. The detector is located in the Laboratori Nazionali del Gran Sasso (LNGS) of Istituto Nazionale di Fisica Nucleare, Italy. Since the laboratory is situated underground, a rock overburden (3500 m w.e.[1]) removes the hadronic components of cosmic ray showers and reduces the muon flux in the experiment by six orders of magnitude. The search for the decay is motivated by the non-zero neutrino mass, which in turn is needed to explain the neutrino oscillation phenomenon, firstly observed by the SuperKamiokande and SNO experiments. Neutrino oscillations provide evidence for the mass of at least some neutrino flavors, however, they do not provide information concerning the absolute neutrino mass, the mass hierarchy or whether the neutrinos are Dirac or Majorana particles [10]. The claim of the observation of the $0\nu\beta\beta$ decay, published by a part of the Heidelberg-Moscow collaboration [11] in 2001, created a rather urgent need for a new ^{76}Ge-based experiment to investigate it with a better precision and lower background. This was the main purpose of the first phase of the experiment and required the application of new low-background techniques, in order to lower the background by at least a factor of 10. The goal of the second phase is an improvement of the half-life sensitivity to $>10^{26}$ yr, by acquiring about 100 kg·yr of exposure and decrease the background by an order of magnitude, with respect to Phase I. In case of a negative results, the experiment was planned to transition to its third phase by forming a world-wide collaboration, capable of probing the effective neutrino mass on 10 meV scale, with O(0.5 t) of enriched germanium [12].

Since the results of Phase II indicate that $T_{1/2}^{0\nu\beta\beta}$ is larger than $8 \cdot 10^{25}$ yr [13], the initially thought Phase III has been recently realized as the LEGEND project [14, 15]. The collaboration has been already formed and includes 219 members from 48 institutions in 16 countries [16]. The first phase of the project (LEGEND-200) will be accommodated in the current GERDA cryostat in LNGS, which can house up to 200 kg of detectors. The ultimate phase is LEGEND-1000, with the ambitious goal of acquiring 10 t·yr of exposure by operating O(1 t) of enriched detectors for about 10 yr.

[1] Muon shielding power of the rock overburden in underground laboratories is usually given in the units of meters of water equivalent (m w.e.). Because the stopping power of muons strongly depends on the rock composition the water equivalent is used to compare muon shielding properties of various sites. For example, Waste Isolation Pilot Plant (New Mexico, U.S.), a deep geological radioactive waste repository, is located 660 m underground (salt deposits), which corresponds to the 1585 m.w.e. On the other hand, Soudan Mine (Minnesota, U.S.) is slightly deeper at 713 m, but the water equivalent is 2100 m due to the iron deposits in the overburden rock.

1.1 History of $2\nu\beta\beta$ and $0\nu\beta\beta$ decays searches

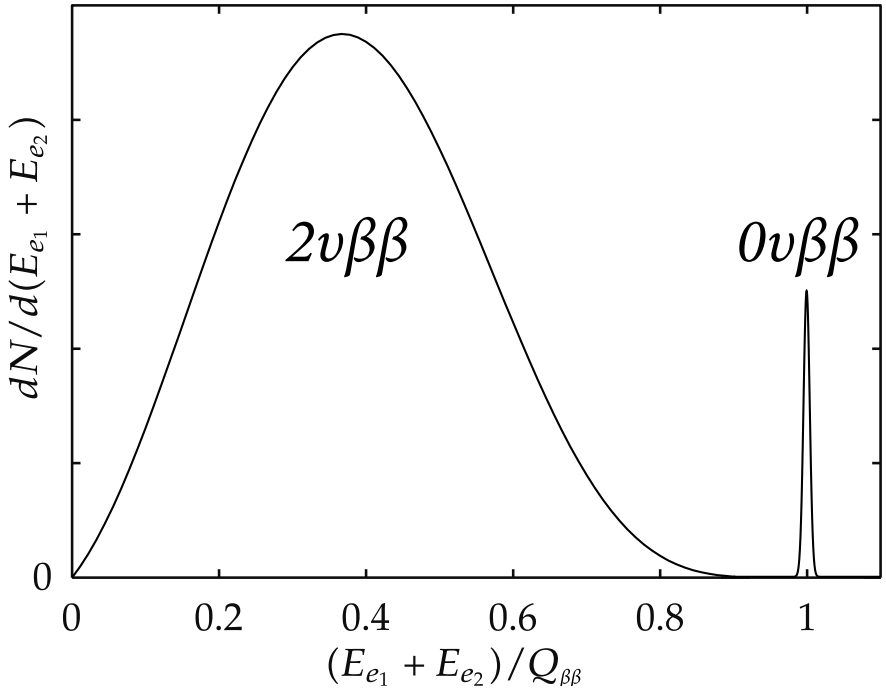


Fig. 1.1: Distribution of the sum of the electron energies ($E_{e1}+E_{e2}$) in the $2\nu\beta\beta$ and $0\nu\beta\beta$ decays [17]. The assumed energy resolution at $Q_{\beta\beta}$ (Q-value of the decay) is 2% (in terms of full width half maximum – in ^{76}Ge experiments it is actually better, at the level of 0.2%). A ratio between the rates of both decays was chosen arbitrarily.

The history of the neutrino accompanied double-beta ($2\nu\beta\beta$) decay dates back to the 1930s, when it was firstly suggested by MARIA GÖPPERT-MAYER in 1935 [18]. For a number of even-even nuclei the single beta decay is energetically forbidden, while the double-beta decay results in a lower energy state and thus is possible. Since the double beta decay is a second-order weak process [19], its half-life is extremely long, in the order of 10^{21} yr (with the exception of ^{128}Te with 10^{24} yr). From a theoretical point of view there are 35 isotopes for which the process can take place [20]. So far it was experimentally confirmed for 11 isotopes. For 9 of them (^{48}Ca, ^{76}Ge, ^{82}Se, ^{96}Zr, ^{100}Mo, ^{116}Cd, ^{130}Te, ^{136}Xe, ^{150}Nd) the measurement was possible by applying counting techniques, but the remaining two (^{128}Te and ^{238}U) were investigated using radiochemical methods. Generally, the $2\nu\beta\beta$ decay can be described with the following equation:

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z+2}X + 2e^{-} + 2\bar{\nu}_e \tag{1.1} $$

Since neutrinos can carry away some part of the relased energy ($Q_{\beta\beta}$ - Q-value of the decay), the distribution of the sum of the electron energies is continuous. This is illustrated in Fig. 1.1.

In 1939 FURRY performed the first theoretical calculation of approximate rates of the $0\nu\beta\beta$ decay [21]. Since no neutrinos are emitted in this process, if both electrons are absorbed in the detector's active volume, the experimental signature is a sharp peak at $Q_{\beta\beta}$ in the energy spectrum (Fig. 1.1). The decay can be described schematically with:

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z+2}X + 2e^{-} \tag{1.2} $$

Observation of such a process can be only possible if the neutrino is a massive Majorana particle and the lepton number is not conserved (no emission of the antineutrinos). The requirement regarding the mass can be considered as already fulfilled, the neutrino oscillation mechanism (proposed

by BRUNO POTECORVO in 1957) can only take place if the neutrino is a massive particle. Since the lepton number conservation violation was yet not observed, the experimental proof of the $0\nu\beta\beta$ decay would be a beginning of a new physics beyond the Standard Model. However, it needs to be mentioned that there is no deep justification for the conservation of lepton and baryon numbers – it is possible that it is just a circumstantial observation for the energy ranges in the current experiments [1].

Geochemical and counting methods. Two distinct approaches have been taken to observe the $2\nu\beta\beta$ and $0\nu\beta\beta$ decays: a geochemical and a direct counting one. The former brought the first observation of the $2\nu\beta\beta$ decay in ^{130}Te as early as in 1950, its half-life was determined to $T_{1/2}^{2\nu\beta\beta} = 1.4 \cdot 10^{21}$ yr [22]. It was followed by the observation of the $2\nu\beta\beta$ decay in ^{82}Se in 1967 by KIRSTEN with $T_{1/2}^{2\nu\beta\beta} = 6 \cdot 10^{19}$ yr [23]. The $2\nu\beta\beta$ and $0\nu\beta\beta$ decays cannot be distinguished by the geochemical techniques, since the final product is the same in both cases. The success of the geochemical methods at that time can be explained by their much better sensitivity when compared with the counter experiments [24]. The $2\nu\beta\beta$ decay of ^{82}Se was first directly observed as late as in 1987, using a time projection chamber [25]. So far no convincing evidence for the $0\nu\beta\beta$ decay was found, but it is worth noting that numerous "false starts" were also reported in the literature. For example, in 1949 FIREMAN performed an experiment with a 25 g sample of tin, isotopically enriched in ^{124}Sn (to 54%), and observed the half-life for the $0\nu\beta\beta$ decay $T_{1/2}^{0\nu\beta\beta} = (4-6) \cdot 10^{19}$ yr, but his result was not later confirmed by any other experiment [24].

^{76}Ge-based $0\nu\beta\beta$ decay experiments . The first double beta experiment based on ^{76}Ge was performed by FIORINI ET AL. in 1967 using a lithium drifted germanium detector (Ge(Li)) [26]. The detector contained ^{76}Ge at the natural abundance level of 7.67%, since isotopic enrichment was seen as a cost prohibitive procedure at that time. A rather small (17 cm^3, 90.5 g) germanium diode was placed inside a multi layered shield and muon veto detectors. The obtained limit on the $0\nu\beta\beta$ decay half-life was $T_{1/2}^{0\nu\beta\beta} > 3.1 \cdot 10^{20}$ yr after 712 h of measurement at sea level. The search was later repeated with a larger Ge diode (68.5 cm^3) in Mount Blanc underground laboratory (4200 m w. e.) and the result was improved to $T_{1/2}^{0\nu\beta\beta} > 5 \cdot 10^{21}$ yr at the 68% confidence level (C.L.).

The non-enriched detector approach was continued until late 1980s. In 1990 D.O. CALDWELL published the best result obtained with a natural germanium detector ($T_{1/2}^{0\nu\beta\beta} > 1.2 \cdot 10^{24}$ yr) [27] from the UCSB/LBL experiment (a cooperation between University of California Santa Barbara and Lawrence Berkeley Laboratory), taking place underground (Oroville Dam, 600 m w. e.). It was then pointed out that by simply continuing the measurement, the limit would be hardly improved – at the dawn of the experiment, the background index (BI) was already vastly lower, compared to the earlier approach[2] (0.3 vs 1.7 cts/(keV·kg·yr)[3]). Since the limit is proportional to a square root of an exposure[4], the elongation of measurement time was not an effective way for its improvement. Larger detector mass, with such low cosmic rays induced background, could not be easily increased due to a rather mundane reason – the only active germanium mine in the U.S. at the time went out of business and only low-grade zinc ore was available. Lower grade meant longer processing time and in turn higher background from cosmic-ray induced isotopes. Clearly the most reasonable way to improve the limit for ^{76}Ge-based $0\nu\beta\beta$ decay experiment was to increase the enrichment factor of the applied detector.

Meanwhile, the first experiment using detectors enriched in ^{76}Ge was under preparation in the USSR. ITEP–ErPhI was a joint collaboration of Institute for Theoretical and Experimental Physics

[2] Lower background was achieved by shortening processing time of germanium at sea level.

[3] The abbrevation "cts" stands for counts.

[4] An exposure in a $0\nu\beta\beta$ decay experiment is defined as a product of active mass of the detector and measurement time

from Moscow and Yerewan Physical Institute. Data taking with 2 Ge(Li) detectors enriched to 85% in ^{76}Ge (the abundance was increased more than tenfold, with the total weight of 1.1 kg) started in 1987. Simultaneous measurements with enriched and natural diodes allowed for the direct background subtraction in the $2\nu\beta\beta$ decay energy region using spectra from both types of the detectors. The calculated $2\nu\beta\beta$ decay half-life was $T_{1/2}^{2\nu\beta\beta} = (9 \pm 1) \cdot 10^{20}$ yr. The limit for the $0\nu\beta\beta$ decay half-life was slightly increased to $T_{1/2}^{0\nu\beta\beta} > 1.3 \cdot 10^{24}$ yr (68% C.L.), with respect to $T_{1/2}^{0\nu\beta\beta} > 1.2 \cdot 10^{24}$ yr from UCSB/LBL experiment [27]. The new result was not very impressive, but one should take into account that it was obtained with the even slightly larger BI than in UCSB/LBL (2.5 vs 1.2 cts/(keV·kg·yr)) and, what's more important, almost 14 times smaller exposure (1.6 vs 22.6 kg·yr). The proof of concept of using enriched germanium detectors in the $0\nu\beta\beta$ decay experiments was demonstrated and the limit was improved with just two diodes (with the mass of ca. 0.5 kg each) and less than 1.5 yr of measurement time. It was the beginning of a new era of experiments using detectors enriched in ^{76}Ge, which were soon able to reach limits on $T_{1/2}^{0\nu\beta\beta}$ larger than 10^{25} yr. The detector production was also simplified by using the newly developed High Purity Germanium (HPGe) crystals, which could be reliably produced as early as in 1980s [28].

1.2 Experimental aspects of the $0\nu\beta\beta$ decay searches

Before proceeding to the description of the GERDA experiment, it is worth noting that from the experimental point of view, various parameters have to be taken into account while designing a $0\nu\beta\beta$ decay experiment, just to name a few: detection efficiency, background and the energy resolution. They will be briefly covered in this section – a large number of different approaches in this field are mostly due to the different physical properties of the $0\nu\beta\beta$ decay candidate isotopes (called in short "$0\nu\beta\beta$ material" or "$0\nu\beta\beta$ isotope" – the possibility of decaying through the $0\nu\beta\beta$ process is of course still hypothetical).

Detection efficiency. Enrichment of the $0\nu\beta\beta$ decay material is a costly and complicated process, therefore it is a logical requirement to obtain as high detection efficiency as possible. The best case scenario is the calometric (internal source) configuration, where the detector is made from the $0\nu\beta\beta$ material [29] and the efficiency is close to 100% (source = detector). This approach was implemented for the first time in 1966 by MATEOSIAN and GOLDHABER for the $0\nu\beta\beta$ decay search in ^{48}CaF crystals [24]. Apart from using the $0\nu\beta\beta$ isotope as the detector material (HPGe detectors with ^{76}Ge, scintillation crystals with ^{48}Ca, ^{116}Cd semiconductor detectors), it can be also mixed with a liquid-based scintillator detector (^{136}Xe). External source experiments also have some advantages like e.g. the possibility for an event topology reconstruction in gaseous time projection chambers (like in the NEMO3 experiment), which is also one of the background reduction techniques. However, the background from the $2\nu\beta\beta$ decays with the energy of electrons close to $Q_{\beta\beta}$ is not reduced, because of the same event topology. Additionally, gasous/liquid radiation detectors have worse energy resolution. The detection efficiency is at the approximate level of 30% [1] and to avoid the self-absorption the source should be produced as a thin foil. This geometry constraint can possibly be an important limitation in achieving large isotope masses.

Q-value of the decay. Another important aspect is $Q_{\beta\beta}$ – it is desirable for it to be larger that the energy deposited by the most energetic background component. In such situation the peak at $Q_{\beta\beta}$ would appear in the low-background region and high signal to background ratio would be obtained[5]. The gamma energy spectrum of the natural radioisotopes ends with the ^{208}Tl line at

[5] Assuming a full energy deposition in the detector.

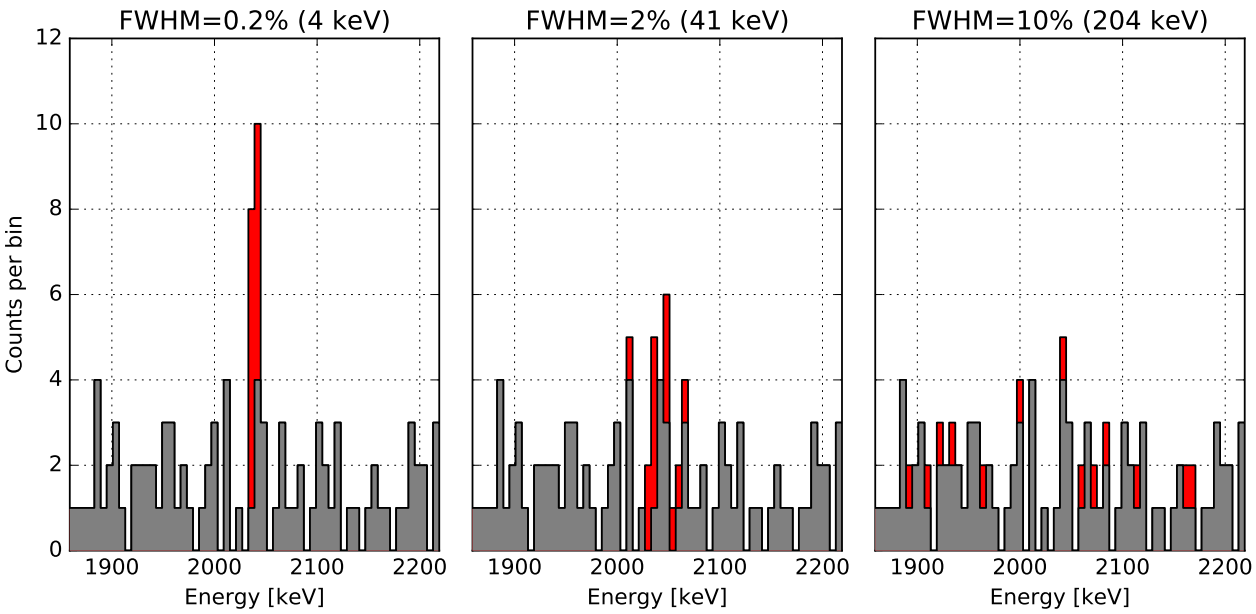


Fig. 1.2: Effect of detector's energy resolution on energy spectra for a 5 yr long, 50 kg active mass ^{76}Ge-based ($Q_{\beta\beta} = 2039.06$ keV) $0\nu\beta\beta$ decay experiment with the assumed background index of $1 \cdot 10^{-3}$ cts/(keV·kg·yr) and $T_{1/2}^{0\nu\beta\beta} = 1.0 \cdot 10^{26}$ yr. The background events (gray) were generated from a flat probability distribution. The signal (red – 13 counts) was simulated using a gaussian distribution with the given FWHM. The spectra show summed signal and background counts (i.e. red color shows the excess of events above the background, due to the $0\nu\beta\beta$ decay). For the very good energy resolution (left panel: 0.2% FWHM, ≈ 4 keV, typical value for germanium detectors) the signal peak clearly protrudes over the background level. In a case of the poor energy resolution (right panel, FWHM of 10%, typical value for a liquid scintillator detector e.g. KamLAND Zen [33]) the signal structure is almost indistinguishable from the background and only the increase in total number of events can be observed. The central panel shows an intermediate situation (2% FWHM, a factor 2 better than e.g. SuperNEMO detector [33]) – the peak structure starts forming in the region of interest. Idea for the plot adapted from [34].

2614.5 keV[6] (99.75% intensity), isotopes with $Q_{\beta\beta}$ smaller than this value will have a spectral signature in the region with background induced by the scattered ^{208}Tl gammas. Therefore, an isotope with a high value of $Q_{\beta\beta}$ should be selected, keeping in mind the maximal background reduction. Although the $Q_{\beta\beta}$ of ^{76}Ge isotope does not fulfill this requirement ($Q_{\beta\beta} = 2039.06$ keV [30]), the excellent energy resolution of germanium detectors allows for the powerful separation of the peak from a flat background [1]. Furthermore, semiconductor detectors are intrinsically very radiopure – the ^{232}Th and ^{238}U radioimpurities are well below 3 µBq/kg and 12 µBq/kg, respectively [31]. Radiopurity of the GERDA Phase I detectors was confirmed in [32], setting the limits at the level of few nBq/kg for the ^{226}Ra, ^{227}Ac and ^{228}Th isotopes.

Another background sources are the α and β decays on the detector's surface. Even if the $Q_{\beta\beta}$ is larger than the 2.6 MeV energy of the ^{208}Tl peak, there is still a possible background contribution from the low energy tail of the alpha emitters (mostly radon daughters). The highest $Q_{\beta\beta}$ of $0\nu\beta\beta$ isotopes (^{48}Ca – 4276 keV) is still smaller than the energy of an α particle from e.g. ^{210}Po decay (5304 keV) and therefore α-related background should be reduced using other countermeasures. The same conclusion is true also for some β emitters like ^{42}K or ^{214}Bi, with the β endpoint energies of 3.52 MeV and 3.27 MeV, respectively.

[6] γ-rays with higher energies can be emitted by the ^{214}Bi isotope from the ^{238}U decay chain, but they have very low intensities when compared with ^{208}Tl (highest ones are: 2694.7 keV at 0.030%, 2769.9 keV at 0.025% and 3053.9 keV at 0.021%).

Energy resolution. The effect of a detector's energy resolution is visualized in Fig. 1.2, where several energy spectra are shown for the different values of the energy resolution (given in the relative full width at half maximum (FWHM) value) in an example ^{76}Ge-based $0\nu\beta\beta$ decay experiment. Exposure and background levels are the same for all plots, the expected number of $0\nu\beta\beta$ decay events was calculated from the following formula:

$$N = \log 2 \cdot \varepsilon \cdot N_{Av} \cdot \frac{m}{M} \cdot \frac{t}{T_{1/2}^{0\nu\beta\beta}} \tag{1.3}$$

where:
ε – detection efficiency,
N_{Av} – Avogadro number,
m – active mass of the $0\nu\beta\beta$ isotope,
M – molar mass of the $0\nu\beta\beta$ isotope,
t – measurement time.

It is clear that the sharp, peak-like signal structure can be observed only if the good energy resolution is available. Otherwise, the signal can be deducted only by the excess of counts over the expected background in the $Q_{\beta\beta}$ region of interest (ROI).

Lastly, the excellent energy resolution is the only countermeasure against the intrinsic background from the $2\nu\beta\beta$ decay. A ratio $R_{0\nu/2\nu}$ of counts originating from the $0\nu\beta\beta$ and $2\nu\beta\beta$ decays, respectively, can be approximated as [35]:

$$R_{0\nu/2\nu} \sim \left(\frac{Q_{\beta\beta}}{\Delta}\right)^6 \cdot \frac{T_{1/2}^{2\nu\beta\beta}}{T_{1/2}^{0\nu\beta\beta}} \qquad \Delta \text{ – energy resolution of the detector.} \tag{1.4}$$

and therefore the good energy resolution is a critical parameter in this regard. Thus, solid-state detectors with good energy resolution (low Δ) will always have the highest discrimination power against the $2\nu\beta\beta$ decay background. However, a minimum requirement on the energy resolution depends also on the chosen isotope (the $T_{1/2}^{2\nu\beta\beta}$ term) [1] – the best candidates are the isotopes with the longest $2\nu\beta\beta$ half-lifes, namely ^{136}Xe and ^{76}Ge: $(2.17 \pm 0.06) \cdot 10^{21}$ yr and $(1.93 \pm 0.09) \cdot 10^{21}$ yr, respectively [5, 29].

1.3 The GERDA experiment

The main design feature of GERDA is to use cryogenic liquid argon (LAr) as a shield against the gamma radiation [31], the dominant background in earlier experiments [37]. Since HPGe detectors are immersed directly in the cryogenic liquid, which also serves as a cooling medium, the amount of possibly radioactive materials around the diodes can be significantly reduced. The cryostat with LAr is installed in a tank containing ultra-pure water, acting as an additional gamma and neutron shield (as shown in Fig. 1.3). The water buffer also serves as a muon veto – the Cherenkov light generated by muons is detected with photomultipliers.

Detectors array. A class 10 000 cleanroom is located above the water tank. Any operations on the detector strings are performed in a glove-box connected to the cryostat entrance. The glove-box is an important element from both radiopurity and detector operation points of view – it prevents the contamination from Rn daughters present in the air and provides nitrogen atmosphere, absent of water vapor. The latter is needed to avoid moisture condensation on the detector surface while immersing the detectors in LAr. The germanium detector array is arranged in vertical strings, every one containing 4–6 diodes hooked one under another using special low-mass holders. The holders used in the first and the second phase of the GERDA experiment are shown in Fig. 1.4, on the left and right panels, respectively.

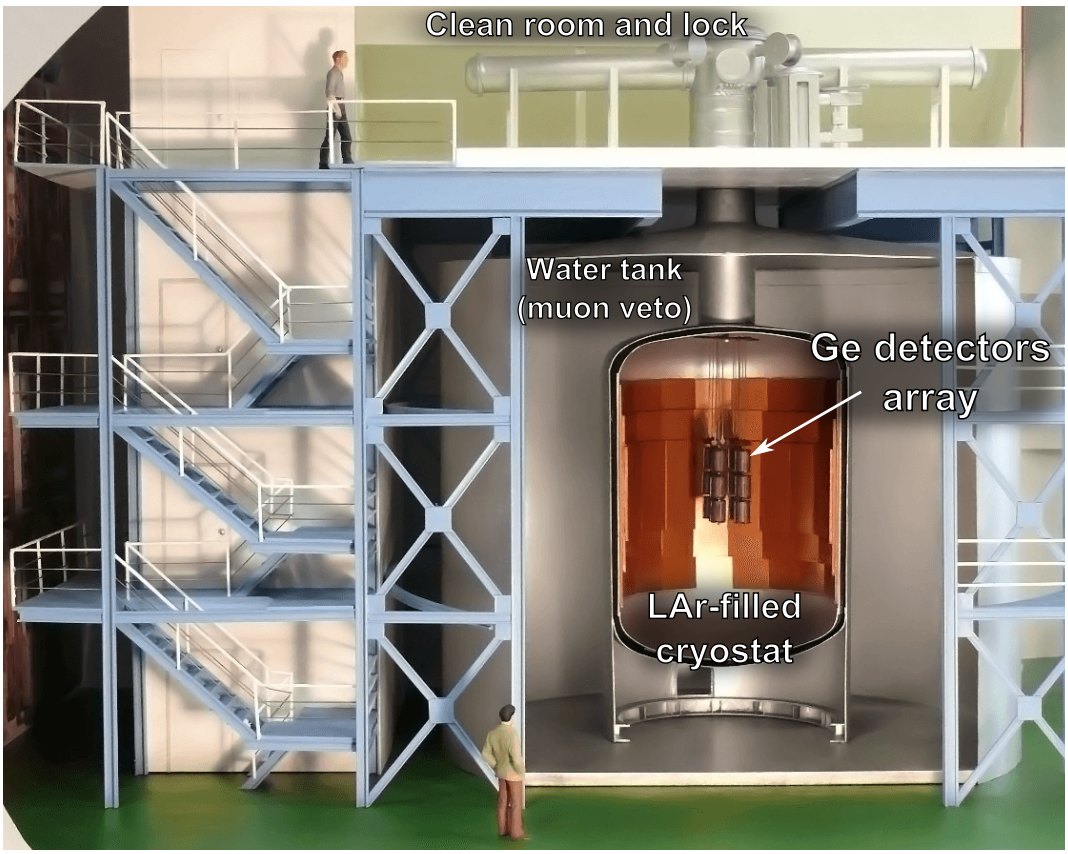


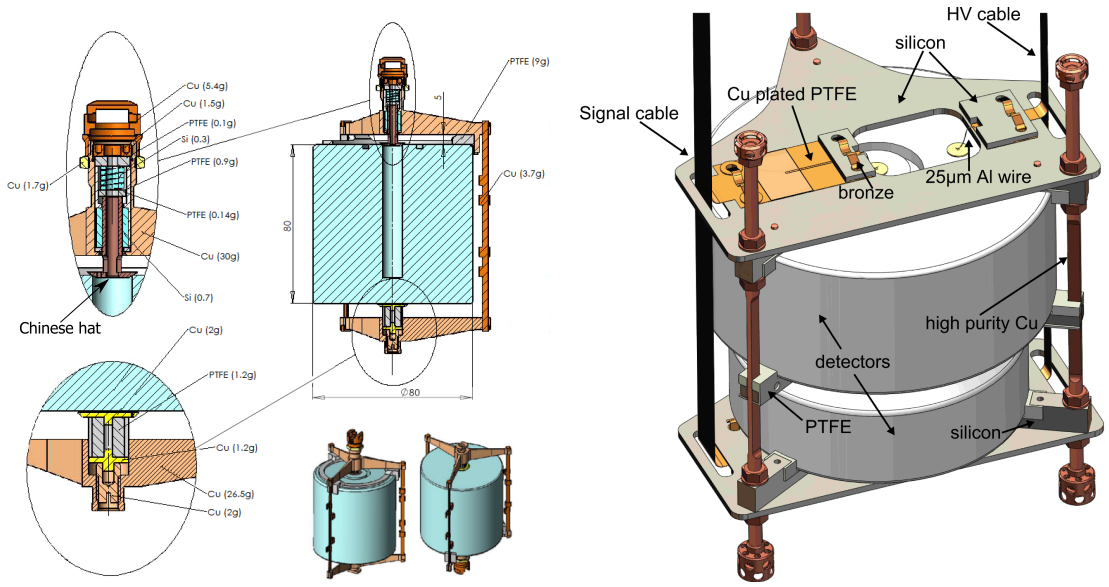
Fig. 1.3: Artistic view of the GERDA experiment setup [36]. Most important components are: the liquid argon cryostat, the detector array, the muon veto tank and the clear room above the lock.

Data acquisition. Front-end electronics consist of a resistive feedback charge sensitive preamplifier, housed on the Cuflon[7] printed circuit board (PCB) [39]. The Very Front End part, consisting of the input JFET with the RC feedback is placed on the silicon plate, close to the detector (Fig. 1.4b). The second stage of the preamplifier is based on AD8651 operational amplifier, mounted on a separate PCB 30 cm away from the detector. This is not an optimal configuration from the electronics performance point of view, but additional distance reduces the background originating from the preamplifier components. Preamplifier signals, after a linear (non-shaping) amplification, are digitized with a 100 MHz/14 bit Flash ADC (model SIS3301, manufactured by Struck Innovative Systeme). A precise energy reconstruction and a pulse shape analysis are performed off-line on the digitized waveforms.

Energy calibration. Data taking is divided into two modes: physics and calibration. Physics mode lasts 1–2 weeks and after that is interrupted by short calibration runs with ^{228}Th sources, introduced in the vicinity of the detectors by a specially designed vacuum-tight mechanical system, from the parking location in the cryostat, Three sources are used in total and the system is controlled by the rotary feedthrough from the outside of the cryostat. During physics data taking, the sources are shielded by tantalum absorbers.

^{228}Th isotope is the third daughter in the ^{232}Th decay chain, which also includes ^{208}Tl. ^{208}Tl emits one of the highest energy γ-rays that can be observed in the nature– the 2614 keV line. In principle, the ^{232}Th source could also be used, but it decays to ^{228}Ac, which would give rise to

[7] Copper traces on the PTFE substrate.



(a) **(b)**

Fig. 1.4: *Left panel a)* – detector holder from GERDA Phase I [36]. Mechanical elements were fabricated from ultrapure copper, detector's HV contact is separated from metal elements with PTFE spacers. Signal readout is realized with the conical copper piece mounted in the well ("Chinese hat").
Right panel b) – reduced mass holder from Phase II [38]. Structural copper was replaced with crystalline silicon plates. HV and signal contacts are made by ultrasonic bonding, using only minute amounts of material [38].

numerous peaks in the spectrum and complicate the calibration procedure, as well as the Pulse Shape Discrimination (PSD) efficiency estimation[8].

1.4 GERDA Phase I

The first phase of the experiment begun in November 2011 and lasted untill March 2013. It was initially started with refurbished semi-coaxial[9] detectors from the previous ^{76}Ge-based $0\nu\beta\beta$ decay experiments: Heidelberg-Moscow (designated ANG1, ..., ANG5) and International Germanium EXperiment (IGEX) (RG1, ..., RG3[10]). The detectors were mounted on a 3-string arm. Additionally, 3 natural (non-enriched) detectors from Genius Test Facility (GTF) [41] were present in the tank – namely GTF32, GTF45 and GTF112. GTF112 was installed in the 3-string arm, together with the enriched detectors. Additional, second arm contained the remaining two GTFs. The detailed arrangement was as follows (see Fig. 1.6):

- String 1: ANG1, ANG2, GTF112
- String 2: ANG4, RG1, RG2
- String 3: ANG3, ANG5, RG3
- String 4: GTF45, GTF32

The details regarding the detectors parameters like mass, dimensions and ^{76}Ge abundance are collected in Tab. 1.1.

[8] The efficiency estimation is based on the peak area calculation in the spectra before and after application of the PSD cut, therefore sparse peaks offer better background estimation possibilities – the procedure is described in more detail in chapter 4.

[9] The adjective "semi-coaxial" refers to the detector's geometry – for details see Sec. 2.1 and Fig. 2.1.

[10] ANG is an abbreviation of "angereichert" (ger. enriched), RG stands for "Rico Grande" – named after the experiment investigating the electron stability [40].

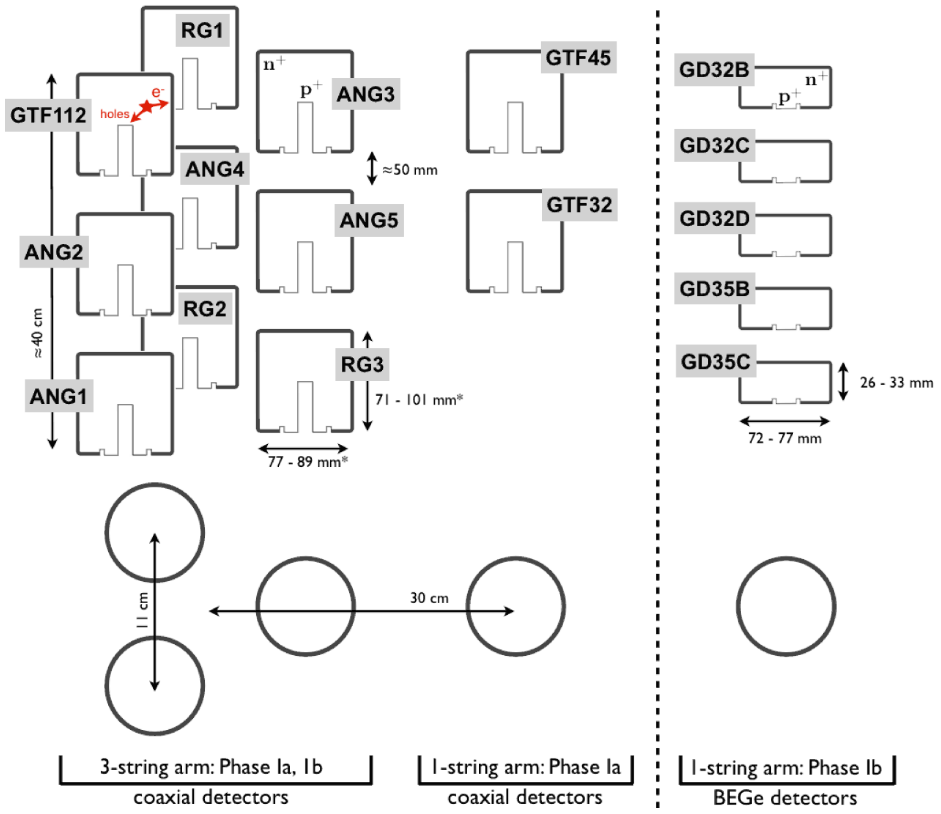


Fig. 1.6: Arrangement of the germanium detectors in GERDA Phase I. The detector fabricated from ^{enr}Ge were put in 3 strings (seen on the left), held on a 3-string arm. Two additional ^{nat}Ge detectors (GTF45 and GTF32) were mounted separately on the second arm. The two ^{nat}Ge detectors were later replaced with a string containing newly produced BEGe detectors. Illustration from [42].

Table 1.1: Parameters of the detectors used in GERDA Phase I [36]. Last 5 BEGe-type detectors (names starting with GD) were added after almost a year of data taking with only semi-coaxial detectors.

Name	Diameter [mm]	Length [mm]	Mass [g]	^{76}Ge abundance [%]	
ANG1	58.5	68	958	85.9	In Phase I from the beginning
ANG2	80	107	2833	86.6	
ANG3	78	93	2391	88.3	
ANG4	75	100	2372	86.3	
ANG5	78.5	105	2746	85.6	
RG1	77.5	84	2110	85.5	
RG2	77.5	84	2166	85.5	
RG3	79	81	2087	85.5	
GTF32	89	71	2321	7.8	
GTF42	85	82.5	2467	7.8	
GTF44	84	84	2465	7.8	
GTF45	87	75	2312	7.8	
GTF110	84	105	3046	7.8	
GTF112	85	100	2965	7.8	
GD32B	71.8	32.2	717	87.7	Added to Phase I in 07.2012
GD32C	72	33.2	743	87.7	
GD32D	72.2	32	723	87.7	
GD35B	76.6	32	812	87.7	
GD35C	74.8	26.4	635	87.7	

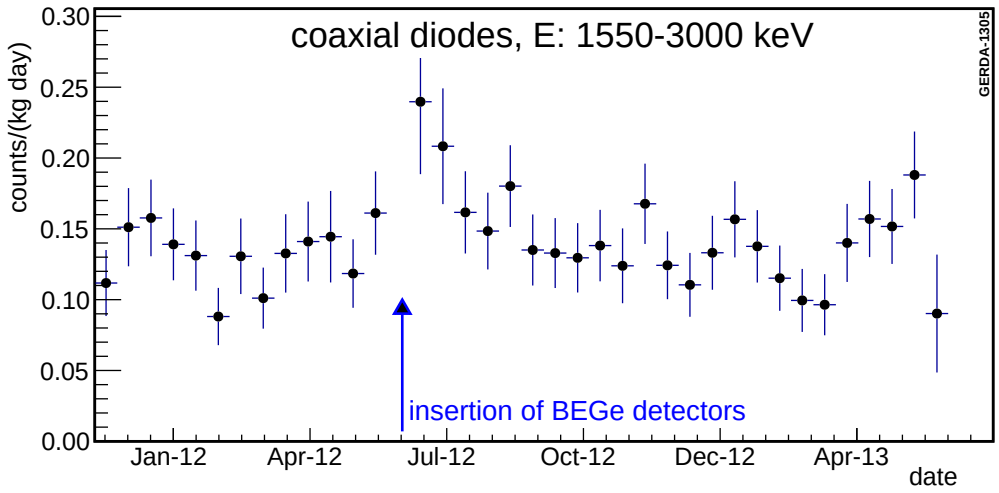


Fig. 1.7: Time distribution of the rate of background events in the 1550 – 3000 keV range for the semi-coaxial dataset. An increase in the count rate is clearly visible after the insertion of the new BEGe detectors [44].

Introduction of the BEGe type detectors.
In July 2012 5 newly produced ^{enr}Ge detectors of the Broad Energy Germanium (BEGe) type were introduced into the tank: GD32B, GD32C, GD32D, GD35B and GD32C. The detectors were manufactured by Canberra Semiconductors N.V., Olen, Belgium from the crystals grown in Canberra Oak Ridge facility in the USA. Due to their highly nonuniform weighting field configuration, they have the enhanced PSD capabilities, when compared to the more traditional semi-coaxial design (see the analysis in Chap. 2). The diodes differ slightly from those that are commercially available – namely, the thin dead layer entrance window is changed into a thick Li-diffused layer. The thin window, desirable in the spectrometry of low energy gamma radiation, would actually expose the active volume of the detector to the residual α/β emitters present at the crystal surface. The Li-diffused dead layer is much thicker (order of 0.5 mm) than the αs range in germanium and the active volume is effectively shielded from the α-induced background. However, some other parts of the detector, namely the p^+ contact and a groove, are still sensitive to the alpha radiation. β particles (e.g. from ^{42}K decay) range in germanium is usually larger than the thickness of the Li layer, so they are not completely stopped. Nevertheless, the Li layer causes the effect of a delayed charge collection in the pulse shape [43] and therefore it is possible to distinguish them by applying the PSD analysis.

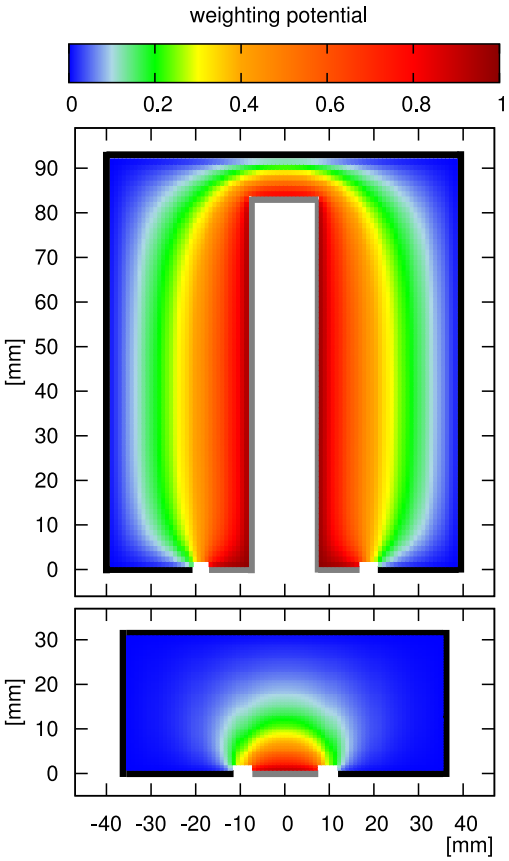


Fig. 1.5: Comparison of weighting potentials in semi-coaxial (top panel) and BEGe-type (bottom panel) detectors. The weighting potential distribution in the BEGe detector provides the enhanced PSD capability – a detailed explanation is included in Chap. 2, which also contains the definition of the weighting potential and the method of its calculation.

$0\nu\beta\beta$ analysis .
An important feature of the experiment was a "blind analysis" concept, which was applied for the first time in the field of $0\nu\beta\beta$ decay search [36]. A $\pm$25 keV energy

window around $Q_{\beta\beta}$ was "blinded" and not available for the analysis during the data taking. Partial unblinding was performed after the background model and the analysis parameters were frozen – but the inner ± 5 keV (± 4 keV for the BEGe detectors) window, where possible signal from the $0\nu\beta\beta$ decay could appear, was still unavailable. Final unblinding took place in Joint Institute for Nuclear Research in Dubna, Russia during a collaboration meeting in June 2013.

The collected data corresponded to 21.6 kg·yr of exposure and live-time of 492.3 days. It was divided into three subsets:

***Golden*:** dataset with the largest exposure and lowest background level, contains major part of the data (only from the semi-coaxial detectors). Exposure: 17.9 kg·yr.

***Silver*:** two short periods of data from semi-coaxial detectors, with higher background level after insertion of the BEGe detectors (see Fig. 1.7). Exposure: 1.3 kg·yr.

***BEGe*:** all data from the BEGe detectors, exposure of 2.4 kg·yr. One of the BEGe detectors (GD35C) has shown an unstable behavior and was excluded from the analysis [7, 45].

Furthermore, a number of cuts were applied to the data before performing the $0\nu\beta\beta$ analysis. First of all, quality cuts were applied to the waveforms to exclude events with the dubious quality e.g. due to high voltage discharges or electromagnetic interference.

Next, the time coincidence cut excluded events with simultaneous signal in the multiple detectors, which clearly cannot originate from the $0\nu\beta\beta$ decay. This step resulted in the background reduction of about 15% around $Q_{\beta\beta}$ with no signal efficiency loss. A muon veto was also used to reject events originating from cosmic radiation. The veto operates on the Cherenkov light generation by the muons in the water tank around the cryostat (Fig. 1.3). Additional plastic scintillator panels are placed over the cleanroom to also veto muons passing through the neck of the cryostat. A 8 µs time window for coincidences between the veto and the detectors is used to remove muon-related events from the analysis set. This provided 7% reduction of the background at $Q_{\beta\beta}$. The last coincidence cut was the one targeted at the fast BiPo[11] events. In this case, the coincidence time window had the length of 1 ms and less than 1% of the events in the ROI are affected by the cut. Therefore, the time coincidence cuts practically do not reduce the detection efficiency. The data-loss due to the cuts is also negligible [7, 46].

Pulse Shape Discrimination. PSD is an important tool for the background reduction used in the GERDA experiment. The main premise behind it is to use the digitized waveform pulse shapes to reject signals that do not have a signature typical to the $0\nu\beta\beta$ events. During the $0\nu\beta\beta$ decay the available energy is transferred to the two electrons (see Eq. 1.2). Taking into account the electron kinetic energy and density of germanium, its mean path in a crystal is approximately 1 mm [47]. Therefore, the whole energy is deposited in a small (when compared to the crystal dimensions) volume. An interaction with such topology is called in short a single-site event (SSE), this is in contrast to the multi-site events (MSEs), in which a γ-ray deposit its energy via multiple interactions (sites) in the detector. The latter constitutes the background with regards to the $0\nu\beta\beta$ decay and should be vetoed. Since a major part of this thesis describes different PSD approaches, the procedure itself is described in detail in the other chapters (e.g. Sec. 2.2). A short summary of the GERDA Phase I data PSD results is only provided here, the full description of the applied methods, together with efficiency values, can be found in [45].

The exact PSD procedure depends on the detector type – different approaches were worked out for BEGe and semi-coaxial detectors. Pulses from the BEGe detectors were analyzed using the A/E approach – the classifier using a ratio of the differentiated current pulse amplitude, A,

[11] BiPo – a short name of the ^{214}Bi-^{214}Po consecutive decays, originating from the ^{222}Rn decay chain. ^{214}Po has a very short half-time of 164 µs, so it follows almost immediately the decay of ^{214}Bi, which in turn has the half-life of 19.7 min.

Table 1.2: Parameters of the analyzed datasets with and without application of Pulse Shape Discrimination procedures [7]. The listed parameters are: exposure $\mathscr{E}$, exposure-weighted average efficiency $\langle\varepsilon\rangle$, counts in 230 keV width window, resulting background index BI and counts in the ROI. See the text for remarks regarding the BI calculation.

data set	$\mathscr{E}$ [kg·yr]	$\langle\varepsilon\rangle$ [%]	cts in $Q_{\beta\beta}\pm\Delta E/2^{\dagger}$	BI in $Q_{\beta\beta}\pm\Delta E/2$ [10^{-3} cts/(keV·kg·yr)]	cts in $Q_{\beta\beta}\pm5$ keV
without PSD					
golden	17.9	68.8 ± 3.1	76	18 ± 2	5
silver	1.3	68.8 ± 3.1	19	63^{+16}_{-14}	1
BEGe	2.4	72.0 ± 1.8	23	42^{+10}_{-8}	1
with PSD					
golden	17.9	$61.9^{+4.4}_{-7.0}$	45	11 ± 2	2
silver	1.3	$61.9^{+4.4}_{-7.0}$	9	30^{+11}_{-9}	1
BEGe	2.4	66.3 ± 2.2	3	5^{+4}_{-3}	0

†) $\Delta E = 230\,\text{keV}$.

to the pulse height E, obtained using an energy filter[12] [47]. The A/E approach did not provide satisfying results for the data from the semi-coaxial detectors and therefore a different technique was needed. In fact, three methods were developed for the semi-coaxial detectors, the classification of the pulses was based on: artificial neural-network (ANN), Projective Likelihood and the current pulse asymmetry. There was a very good agreement between all three of them: more than 90% of the background events in ROI rejected by ANN were also rejected by the other approaches. ANN has been chosen as the leading method.

Signal efficiency for the A/E (BEGe dataset) was calculated using events from the double escape peak (DEP) of the ^{208}Tl line – the acceptance was equal to $(92\pm2)\%$ at $Q_{\beta\beta}$ and 80% of the background events were rejected. The acceptance for the ANN method in the semi-coaxial detectors was determined to be $90^{+5}_{-9}\%$, uncertainties are derived from survival efficiencies of SSEs from the $2\nu\beta\beta$ region and the Compton edge of the 2615 keV line (after subtracting the MSE background) [45].

Results of the $0\nu\beta\beta$ analysis. In the case of observing the $0\nu\beta\beta$ decay, the formula to calculate its half-life is as following [7]:

$$T_{1/2}^{0\nu\beta\beta} = \frac{\ln 2 \cdot N_{Av}}{m_{enr} \cdot N^{0\nu}} \cdot \mathscr{E} \cdot \varepsilon \quad (1.5)$$

where:
N_{Av} – Avogadro number,
$\mathscr{E}$ – total exposure [kg·yr],
m_{enr} – molar mass of ^{76}Ge (75.6 g/mol),
$N^{0\nu}$ – number of observed events in the $Q_{\beta\beta}$ region after subtracting the flat background [cts].

The total detection efficiency ε is a product of the following factors:

$$\varepsilon = f_{76} \cdot f_{av} \cdot \varepsilon_{psd} \cdot \varepsilon_{fep} \quad (1.6)$$

[12]The abbreviation E comes from *energy*, since the pulse height is proportional to the energy deposited in the detector.

Table 1.3: List of all events within the $Q_{\beta\beta} \pm 5$ keV window in GERDA Phase I [7].

Dataset	Detector name	Energy [keV]	Date	PSD vetoed
golden	ANG 5	2041.8	18-Nov-2011 22:52	yes
silver	ANG 5	2036.9	23-Jun-2012 23:02	no
golden	RG 2	2041.3	16-Dec-2012 00:09	no
BEGe	GD32B	2036.6	28-Dec-2012 09:50	yes
golden	RG 1	2035.5	29-Jan-2013 03:35	no
golden	ANG 3	2037.4	02-Mar-2013 08:08	yes
golden	RG 1	2041.7	27-Apr-2013 22:21	yes

and accounts for the enrichment fraction in ^{76}Ge (f_{76}), the active volume fraction (f_{av}), the efficiency of the PSD (the signal acceptance – ε_{psd}) and the efficiency of the $0\nu\beta\beta$ decay depositing its whole energy in the detector, resulting in a full energy peak (FEP) at $Q_{\beta\beta}$ (ε_{fep}). The ε_{fep}, determined by the Monte Carlo (MC) simulations, equals to 0.92 for the semi-coaxial and 0.90 for the BEGe detectors. The analysis parameters of all datasets are summarized in Tab. 1.2. The analysis of the BI values before and after the PSD shows an interesting effect: even though the original BI for BEGes (without PSD) is about a factor 2 higher than for the semi-coaxials, after the application of the PSD it is actually a factor 2 lower. This is because of the very high PSD efficiency for the BEGe detectors. Detailed information concerning the unblinded events (like dataset, date, PSD flag etc.) from the ROI ($Q_{\beta\beta} \pm 5$ keV) is collected in Tab. 1.3.

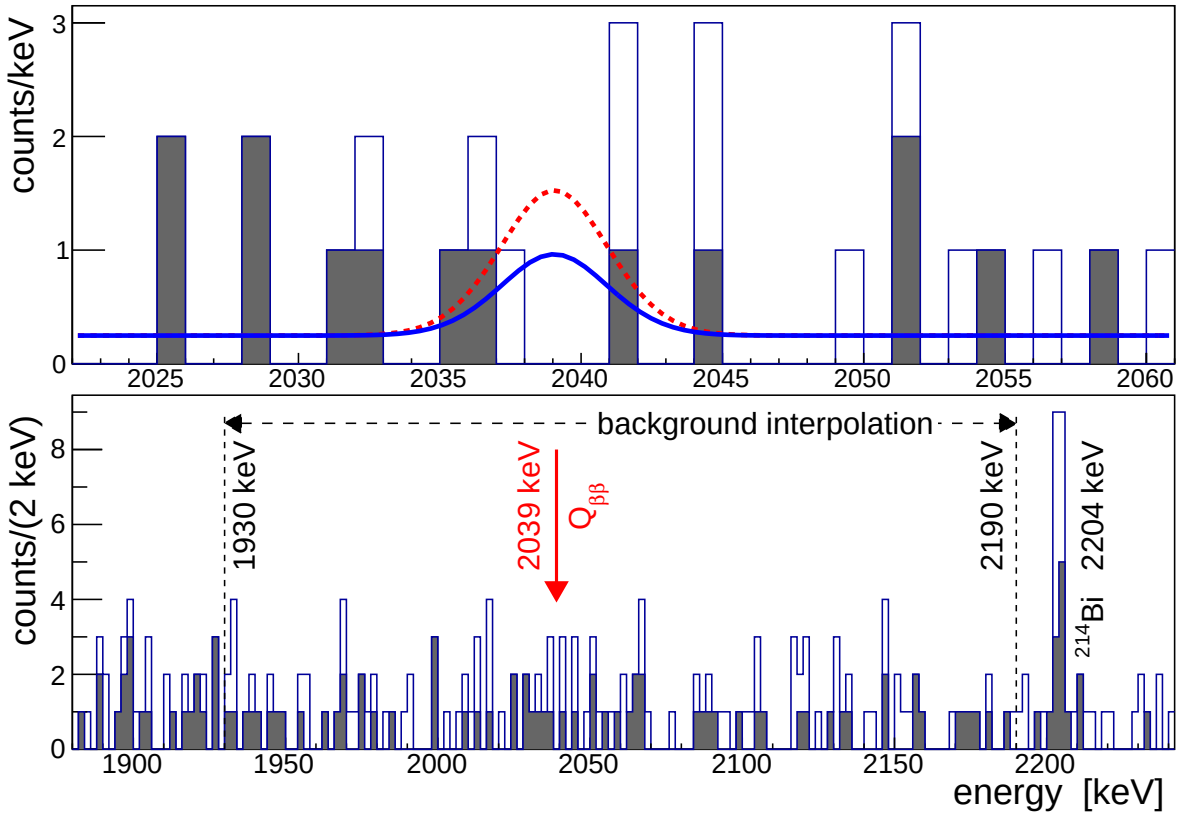


Fig. 1.8: GERDA Phase I energy spectrum, all enriched detectors combined, with and without the application of the PSD (filled and open bars, respectively) [7]. The upper panel shows events around the $Q_{\beta\beta}$. A blue solid line corresponds to the derived limit – expected number of events corresponding to the $T_{1/2}^{0\nu\beta\beta}$ of $2.1 \cdot 10^{25}$ yr. A red dashed line shows the expected $0\nu\beta\beta$ decay events corresponding to the half-life claimed by a part of the Heidelberg-Moscow collaboration $T_{1/2}^{0\nu\beta\beta} = 1.19 \cdot 10^{25}$ yr [48]. The lower panel presents the energy region used for the background interpolation.

A flat distribution of the background events in the $Q_{\beta\beta}$ region is expected from the GERDA background model [44]. The most important components included in the model are:

- Compton continuum events from the ^{208}Tl FEP at 2615 keV and the ^{214}Bi FEPs,
- energy deposition from β particles from ^{42}K decays ($Q = 3.52$ MeV),
- degraded α events from ^{226}Ra, ^{222}Rn and ^{210}Po surface contaminations.

In order to calculate the BI, regions from the known peaks ((2104 ± 5) keV and (2119 ± 5) keV from ^{208}Tl and ^{214}Bi, respectively), as well as the blinded region ((2039 ± 5) keV) are excluded. The energy window width for the calculation is therefore 230 keV (after the exclusion). The background interpolation region is shown in Fig. 1.8.

After the unblinding, it turned out that there were 7 events in the ROI - 6 of them from the semi-coaxials (RG1, RG2, ANG3 and ANG5) and 1 from the BEGe (GD32B). The expected number of events from the background model (flat distribution) was 5.1 ± 0.5. No excess of events over the expected background was observed in any of the datasets. The results after applying the PSD analysis were also consistent – in the semi-coaxial detectors, 3 events were classified as the SSEs for a final result. The single event in the BEGe dataset was vetoed by the A/E analysis.

A profile likelihood fit was used to calculate the number of signal events $N^{0\nu}$ in the data. The fit function is a Gaussian peak, with the mean value at $Q_{\beta\beta}$ and dispertion σ_E. A flat background pedestal is assumed in the fit and σ_E was derived from an energy resolution curve, which in turn was calculated from the calibration data. The analysis resulted in a best fit value of $N^{0\nu} = 0$ and the limit of the $0\nu\beta\beta$ decay half-life was:

$$T_{1/2}^{0\nu} > 2.1 \cdot 10^{25}\,\text{yr} \quad (90\,\%\ \text{C.L.}) \tag{1.7}$$

while the median sensitivity, derived from the parameters in Tab. 1.2, is $2.4 \cdot 10^{25}$ yr (also for the 90% C.L.). The limit value on the half-life can be converted into the limit on the number of counts above the background in the ROI $N^{0\nu} < 3.5$ (the blue line in the top panel of Fig. 1.8).

Combination of the Phase I limit with the results from the previous experiments. GERDA Phase I result is consistent with the limits by previous ^{76}Ge-based experiments, namely Heidelberg-Moscow and IGEX [49, 50]. If their respective results are combined into the single analysis (by the extension of the profile likelihood fit), the limit is strengthened to:

$$T_{1/2}^{0\nu} > 3.0 \cdot 10^{25}\,\text{yr} \quad (90\,\%\ \text{C.L.}) \tag{1.8}$$

Comparison with the discovery claim. Data from GERDA Phase I does not show any indication for the possible existence of the $0\nu\beta\beta$ decay. However, one can compare the claim of the discovery with the half-life from [48] (model H_1 – background + signal with $T_{1/2}^{0\nu\beta\beta} = 1.19 \cdot 10^{25}$ yr) and Phase I result (model H_0 – background only) by calculating the Bayes factor $B = p(data|H_1)/p(data|H_0)$. Taking the $T_{1/2}^{0\nu\beta\beta}$ from H_1, 5.9 ± 1.4 signal events in $Q_{\beta\beta} \pm \sigma_E$ and 2.0 ± 0.3 background events are expected after a PSD cut. Only 3 background events are observed in the GERDA Phase I ROI and none of them within $Q_{\beta\beta} \pm \sigma_E$ energy range. Therefore, the probability of observing no events, assuming H_1, equals to $p(data|H_1) = 0.01$ [42]. $B = 0.024$ and by combining the data from Heidelberg-Moscow and IGEX experiments it is lowered to $B = 2 \cdot 10^{-4}$. This means that it is very unlikely that no signal events are observed with the claimed half-life and the claim can be refuted with a high probability.

1.5 Upgrade to Phase II

The scientific goal of GERDA Phase II is to reach the limit for the half-life of the ^{76}Ge $0\nu\beta\beta$ decay at the level of $T_{1/2}^{0\nu\beta\beta} > 1 \cdot 10^{26}$ yr (90 C.L.) and the limit on the effective Majorana

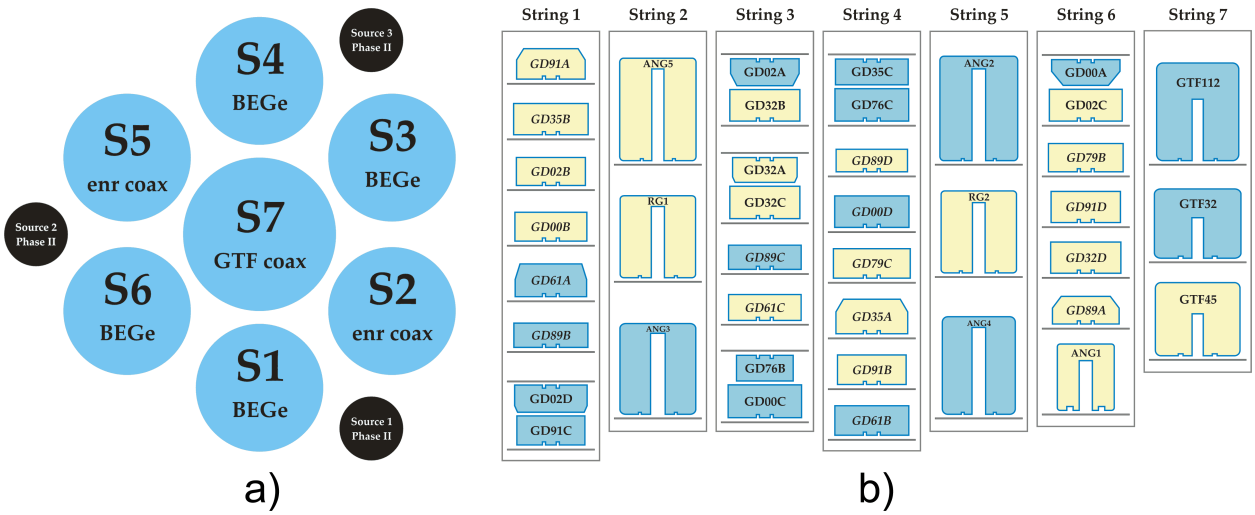


Fig. 1.9: Overview of the detector configuration in the GERDA Phase II array. On the left panel (a) shows the top view of the strings, together with the geometrical location of the calibration sources (black circles). Right panel (b) shows the detector configuration in each string – diodes with a removed passivation layer are drawn in yellow. Drawing by Konstantin Gusev.

mass of $m_{\beta\beta} < 0.1-0.2\,\text{eV}$ (depending on the nuclear matrix elements). It will be possible after accumulating $\approx 100\,\text{kg}\cdot\text{yr}$ of exposure (≈ 3 yr of data taking) with the background index of $1\cdot10^{-3}$ cts/(keV·kg·yr). In Phase II, beside an increase of the active mass by about 20 kg (30 new BEGe detectors [51]), the main goal was to further reduce the background by at least one order of magnitude, with respect to Phase I. Several strategies were incorporated to achieve this goal:

- Reduction of the material mass in the detector's vicinity, namely: holders and contact pins. In Phase I a spring-loaded pin was used to read out the signal from a p^+ contact. It means that a significant amount of the material was very close to the most sensitive part of the detector. Even though the material is very radiopure, the best strategy is to reduce the mass of all hardware elements. The pins were therefore replaced by the very thin bonding wires (diameter $\sim 10\,\mu\text{m}$). Copper-based holders from Phase I were also replaced by the plates made from silicon (an extremely radiopure material [31]). Some of the BEGe detectors are also mounted in pairs, the design which further decrease the required mass of the holders.
- Installation of LAr veto – in Phase I LAr in the cryostat served only as a passive shield and a coolant. The addition of the readout of scintillation light (with Photomultiplier Tubes (PMTs) and Silicon Photomultipliers (SiPMs)) allows to use the LAr as an active veto [52].
- Usage of BEGe-type detectors with the enhanced PSD capabilities. BEGe detectors were also deployed in the GERDA Phase I, but in the Phase II they constitute of $\approx 50\%$ of the total masss of the detectors.

Transition from Phase I to Phase II. The upgrade started in May 2014 with the installation of a new, enlarged lock system, which allowed for the use of an altered germanium array geometry. Assembly of the detector strings and LAr instrumentation was done in a glove-box installed over the lock system. The whole setup was tested with the immersion of the first detectors in July 2014. The LAr veto was deployed by November 2014 and introduction of new Phase II detectors was performed since January untill December 2015.

Data taking in the final configuration of the detectors (40 diodes, shown in Fig. 1.9), started on 20$^{\text{th}}$ of December 2015 (run 53[13]). Due to the initial instabilities a part of the data was later

[13] In GERDA, data taking is divided into runs. Usually a new run is started when there is a some kind of setup modification e.g. high voltage adjustment.

discarded – first stable calibration data is available from 23rd of December at 10:41:51 UTC. Physics data taking started on 25th of December 2015.

New detector array. The detector array had to be changed with respect to Phase I to accommodate new diodes. The array consists of 7 strings – 4 of them contain the BEGe detectors (with the exception of the string 6, which has the ANG2 semi-coaxial detector added at the bottom). Phase I enriched semi-coaxial detectors were put in strings 2 and 5. String 7, placed in the middle, carries non-enriched GTF detectors. A detailed string composition is shown in Fig. 1.9b. Each string is surrounded by a nylon mini-shroud, which reduces volume of LAr from which radioactive ions (mainly ^{42}K) drift due to electric field in the detectors vicinity [44]. The ions near the diodes can be collected by the detector surfaces [9] and contribute to the background by the emission of α/β particles. Similarly to the copper cylinder, the shrouds are covered with the TBP wavelength shifter.

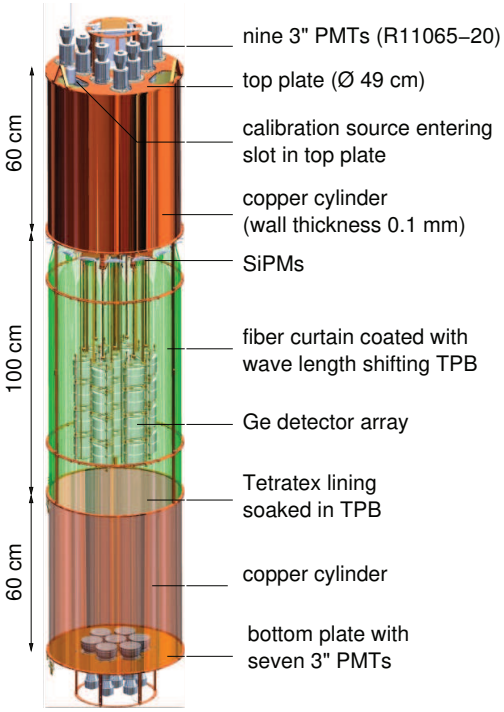


Fig. 1.10: Technical drawing of the LAr veto system of GERDA Phase II [53]. Scintillation light produced by LAr is readout by PMTs at the top and the bottom and by the fiber curtain coupled with SiPMs. Various parts of the system (the copper cylinder, nylon shrouds, fiber curtain) are covered with tetraphenyl butadiene (TBP) – its role is to shift the light from the deep ultraviolet wavelength (128 nm) to about 450 nm (visible part of the spectrum). The shifted light matches the maximal sensitivity of the PMTs and is not absorbed nor reflected by the nylon and glass elements.

LAr veto system. Unlike in Phase I, in GERDA Phase II the LAr is no longer just a passive shield. Two independent systems were installed for the LAr scintillation light readout. The light is produced by ionizing particles by the deexcitating Ar atoms, excited either directly or by the ionization/recombination mechanism [54]. The complete LAr veto setup is shown in Fig. 1.10. A first light readout system consists of two plates equipped with nine and seven 3" PMTs each, at the top and the bottom, respectively. The concept of the LAr readout by PMTs immersed in LAr was tested earlier in the LARGE test facility [52], built specially for that purpose. The copper cylinder is lined with the Tetratex foil, impregnated with TBP to shift the 128 nm wavelength of the scintillation light to about 450 nm, which is visible to PMTs/SiPMs [53]. The second system is using wavelength shifting fibres coupled to SiPMs. They increase the volume of LAr, from which the scintillation light can be collected.

The working principle behind the LAr veto is a readout of the scintillation light coincident to the signal in one of the germanium detectors. Such coincidences are mainly caused by the multiple scattering of the γ-rays from the thorium and uranium chain decays in the construction materials and/or radon daughters deposited on the surfaces. Background from ^{42}K decays (produced by ^{42}Ar present in LAr) is also successfully rejected, since the 1525 keV gamma ray is accompanied with the emission of the β^- particle (decay Q-value of 3525.2 keV), which deposits its energy in LAr. The effect of the LAr veto on a suppression of ^{42}K gamma line can be observed in the inset of Fig. 1.11.

1.6 GERDA Phase II

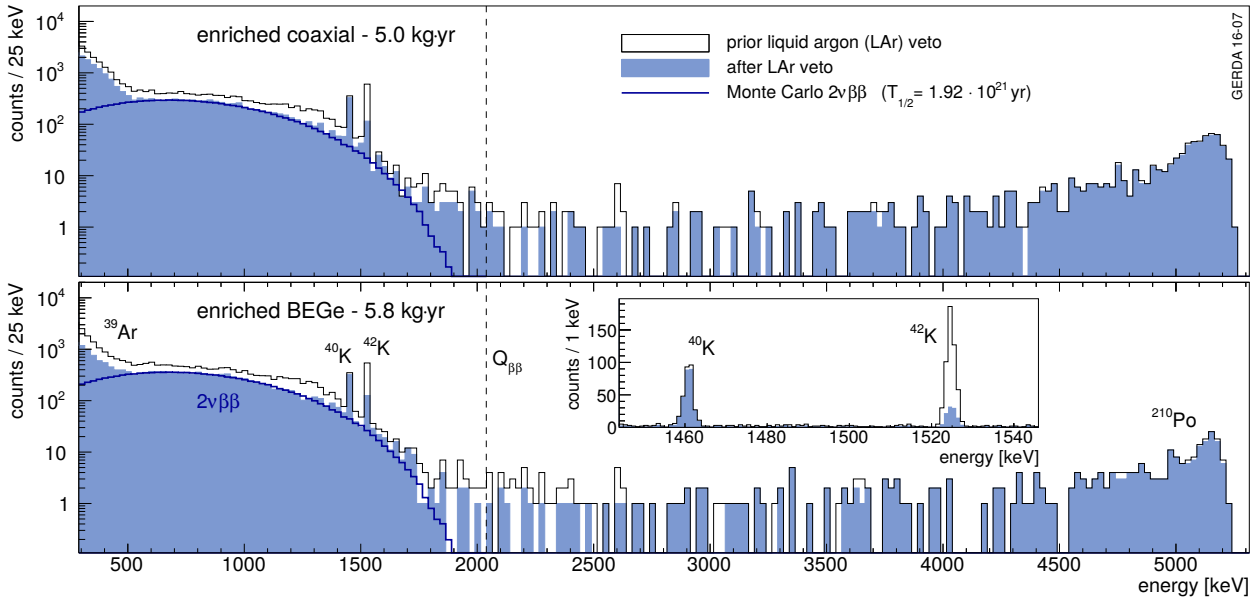


Fig. 1.11: Energy spectra plotted for the enriched semi-coaxial (top panel) and BEGe-type (bottom panel) detectors (Phase II, exposures of 5.0 kg·yr and 5.8 kg·yr for semi-coaxial and BEGe detectors, respectively) [55]. Filled histogram shows the spectra after application of the LAr veto. Simulated events from the $2\nu\beta\beta$ decay (blue solid line) match very well the data in a low energy region (600 – 1400 keV). ^{40}K FEP is not affected (no energy deposition in LAr), while the ^{42}K is suppressed due to the coincident β^- emission.

First results of GERDA Phase II regard the data taken between December 2015 and June 2016. The official unblinding procedure took place on June 17$^{\text{th}}$ 2016 at the collaboration meeting in Ringberg, Germany. The acquired exposure, corresponding to the new data, was equal to 5.0 kg·yr and 5.8 kg·yr for semi-coaxial and BEGe detectors, respectively [9]. The data analysis procedure was almost the same as in Phase I – the only differences were the application of the newly integrated LAr veto and a new PSD method for the discrimination of α events in the semi-coaxial detectors. The method uses an additional artificial neural-network trained on events from two different energy regions: the $2\nu\beta\beta$ decay (1.0 – 1.3 MeV) and the α region (energy over 3.5 MeV). Its signal efficiency, tested on a part of the $2\nu\beta\beta$ decay events, which were not used in the training process, is equal to $(93 \pm 1)\%$. The total efficiency of both neural-network based cuts is $(79 \pm 5)\%$.

Background-free operation. According to the Phase II design goal, the BI was significantly reduced with respect to Phase I (Tab. 1.4). After applying all cuts, it was equal to $3.5^{+2.1}_{-1.5} \cdot 10^{-3}$ cts/(keV·kg·yr) and $0.7^{+1.1}_{-0.5} \cdot 10^{-3}$ cts/(keV·kg·yr) for the semi-coaxial and BEGe datasets, respectively. For the comparison, Phase I values were also included in Tab. 1.4. Regarding the BEGe datasets from Phases I and II, the BI was decreased by almost an order of magnitude and is below 10^{-3} cts/(keV·kg·yr). Therefore, the background goal of Phase II has been achieved.

Such low background level has an important consequence for the projected sensitivity of the experiment. A general relation between the sensitivity for $T_{1/2}^{0\nu\beta\beta}$ and the exposure, assuming the Poisson counting statistics, is given as [1, 56]:

$$T_{1/2}^{0\nu\beta\beta} \propto \sqrt{\frac{\mathcal{E}}{\text{BI} \cdot \Delta}} \tag{1.9}$$

This simple square root relation shows that, at a given background level, the sensitivity growth rate will be slowing down with the increasing exposure. Therefore, the background should be as low as

Table 1.4: Analysis parameteres of the GERDA Phase II data [9]. The parameters of GERDA Phase I data were also included for completeness. "PI extra" denotes the additional Phasa I data, which was not included in the published Phase I results [7].

Dataset	Exposure $\mathscr{E}$ [kg·yr]	FWHM [keV]	Efficiency ε	BI 10^{-3} cts/(keV·kg·yr)
PI golden	17.9	4.3(1)	0.57(3)	11 ± 2
PI silver	1.3	4.3(1)	0.57(3)	30 ± 10
PI BEGe	2.4	2.7(2)	0.66(2)	5^{+4}_{-3}
PI extra	1.9	4.2(2)	0.58(4)	5^{+4}_{-3}
PII coaxial	5.0	4.0(2)	0.53(5)	$3.5^{+2.1}_{-1.5}$
PII BEGe	5.8	3.0(2)	0.60(2)	$0.7^{+1.1}_{-0.5}$

possible to prevent this effect of taking place during the experiment lifetime. In particular, for the background level so low that $\mathscr{E} \cdot \mathrm{BI} \cdot \Delta \lesssim 1$ (i.e. less than one background event is expected for the planned exposure) Eq. 1.9 takes the form:

$$T_{1/2}^{0\nu\beta\beta} \propto \mathscr{E} \tag{1.10}$$

In this case, the sensitivity increases linearly with the exposure – this is so called "background-free" regime. The achieved BI for the BEGe detectors fulfills the requirement and therefore, GERDA is a first background-free experiment in the field [9].

Results. Similarly as in Phase I, the unbinned extended likelihood fit was used in the statistical analysis. The fit was applied to all datasets listed in Tab. 1.4. The analysis includes also an additional dataset from Phase I, which was acquired after the unblinding in 2013, but before the

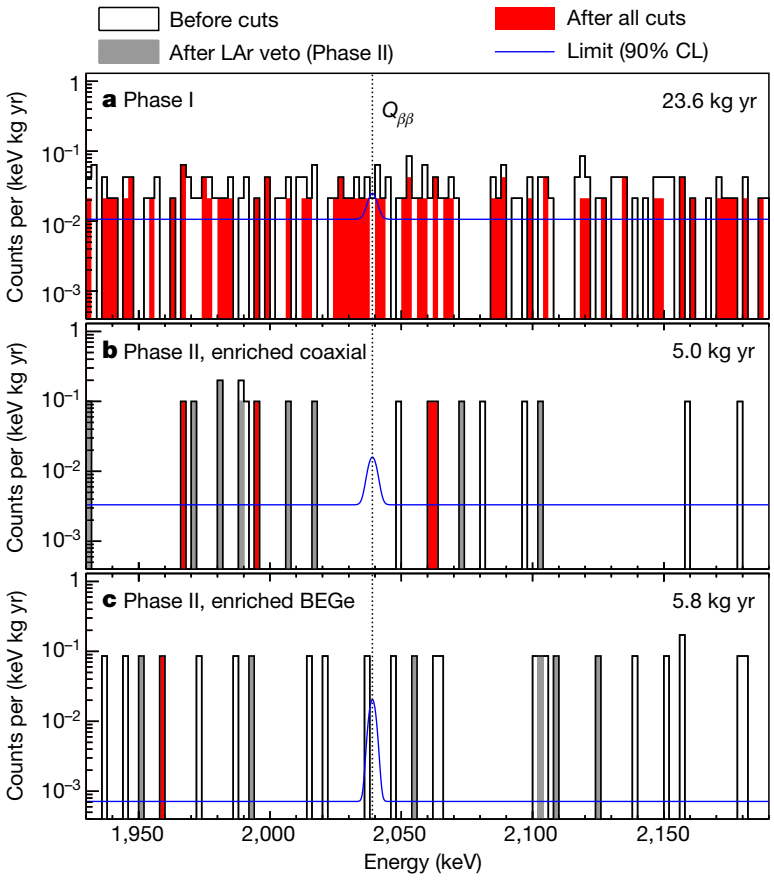


Fig. 1.12: Energy spectra around ROI for: Phase I (upper panel), Phase II enriched semi-coaxial (central panel) and Phase II enriched BEGe (bottom panel) datasets [9]. The exposure values, corresponding to each dataset, are annotated in the upper right corners. Open histograms show the spectra before the cuts, grey ones after LAr veto (only Phase II) and red after all possible cuts (including PSD). Blue lines show the fitted spectrum with the hypothetical signal corresponding to the half-life limit of $T_{1/2}^{0\nu\beta\beta} > 5.3 \cdot 10^{25}$ yr.

upgrade to Phase II. No signal was observed and the extracted limit on the $0\nu\beta\beta$ decay half-life is [9]:

$$T_{1/2}^{0\nu} > 5.3 \cdot 10^{25}\,\text{yr} \quad (90\%\ \text{C.L.}) \tag{1.11}$$

The limit corresponds to 2.0 events in 34.4 kg·yr of total exposure. Energy spectra around ROI are shown in Fig. 1.12. The median sensitivity, assuming no signal, is equal to $4.0 \cdot 10^{25}$ yr.

Second data release. The second unblinding of the Phase II data took place on 30$^{\text{th}}$ of June 2017 at the collaboration meeting in Cracow, Poland. The unblinding of the semi-coaxial dataset was postponed, due to the observed degraded α events in the vicinity of the groove [13]. These events induce faster signals, which are not correctly identified by the neural-network based PSD methods. Therefore, the additional exposure (12.4 kg·yr), with respect to the previous result [9], comes from the BEGe detectors. Again, no $0\nu\beta\beta$ decay signal was observed and the improved limit is [13]:

$$T_{1/2}^{0\nu} > 8.0 \cdot 10^{25}\,\text{yr} \quad (90\%\ \text{C.L.}) \tag{1.12}$$

with the median sensitivity of $5.8 \cdot 10^{25}$ yr. The unblinded spectrum around ROI is shown in Fig. 1.13.

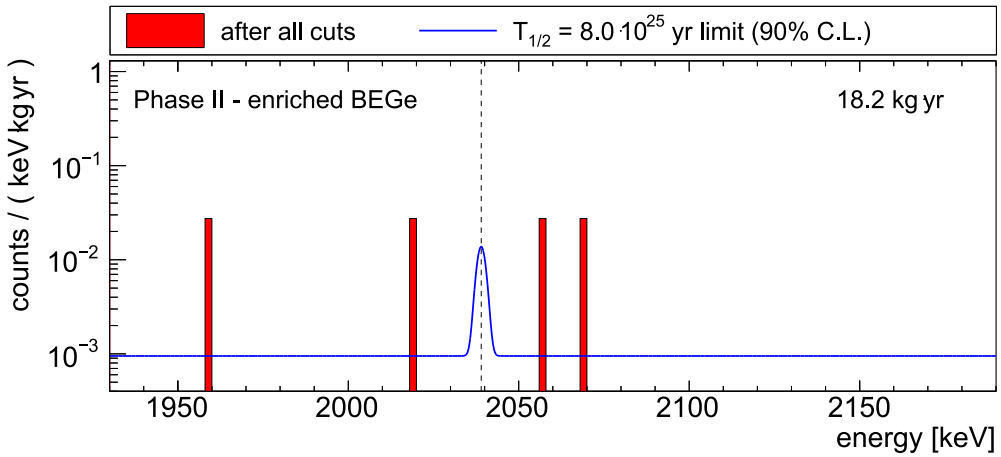


Fig. 1.13: Energy spectrum around ROI for the BEGe dataset after the second data release of Phase II. Four events are present in the background window, corresponding to BI $= 1.0^{+0.6}_{-0.4} \cdot 10^{-3}$ cts/(keV·kg·yr). The blue line show the fitted spectrum with the hypothetical signal corresponding to the half-life limit of $T_{1/2}^{0\nu\beta\beta} > 8.0 \cdot 10^{25}$ yr.

Compared with Fig. 1.12 (bottom panel), 3 additional events are present in the vicinity of $Q_{\beta\beta}$. One of them (furthest to the right in Fig. 1.13, i.e. with the highest energy) appeared during the data taking and was taken into account for the BI estimation before unblinding. Since the exposure was significantly increased with respect to the previous data release, the new BI was actually lower: $0.5^{+0.5}_{-0.3} \cdot 10^{-3}$ cts/(keV·kg·yr). After the unblinding, the two events closest to the $Q_{\beta\beta}$ (on the both sides of the blue line in Fig. 1.13) increased the BI to $1.0^{+0.6}_{-0.4} \cdot 10^{-3}$ cts/(keV·kg·yr). However, it is still within the background goal for Phase II and the experiment remain background-free.

Outlook for the future. Since the experiment is now in the background-free regime and is still collecting data, the projected sensitivity on the $T_{1/2}^{0\nu\beta\beta}$ of $1.0 \cdot 10^{26}$ yr should be achieved in the first half of 2018 [13]. The projected sensitivity for the full design exposure of 100 kg·yr is $1.3 \cdot 10^{26}$ yr for the limit on $T_{1/2}^{0\nu\beta\beta}$ at the 90% C.L. A statistical analysis, presented in [57], shows that in the case that the $0\nu\beta\beta$ decay exists, there is a 50% chance for its discovery, if $T_{1/2}^{0\nu\beta\beta} = 8.0 \cdot 10^{25}$ yr (at the 3σ significance). The achieved background levels, which are lowest in the field, confirm that ^{76}Ge-based approach is a promising technology for the $0\nu\beta\beta$ decay ton-scale experiments.

Chapter 2

Signal formation in germanium radiation detectors and principles of the Pulse Shape Discrimination

One of the advantages offered by the semiconductor detectors is their energy resolution, especially when compared to the other detectors used in the γ-ray spectrometry, namely inorganic scintillating crystals. Crystalline NaI scintillators were a standard tool for this task since early 1950s [58], due to their high detection efficiency, portability and no special requirements in terms of operational conditions. At the beginning of 1960s a new kind of device appeared on the market, namely lithium-drifted germanium detectors (Ge(Li)), which offered the energy resolution comparable only with the silicon detectors. Si, however, is not a particularly well suited material for the γ-ray detectors. First of all, high atomic number Z is beneficial in terms of detection efficiency of γ-rays, and for Si it is rather low when compared to other materials (Si: 14, Ge: 32, I: 53). Another problem is the technical difficulties – the level of impurities in the silicon crystals allows for the fabrication of the detectors with a maximal depletion depth in the order of several mm [58]. Lithium drifting process, applied to germanium (Ge(Li)), allowed for the impurities compensation in large crystals. Large active volume Ge(Li) detectors could then be used for the high resolution (few keV at 1 MeV) γ-ray spectrometry, providing invaluable measurement possibilities. They can be used to determine structure of nuclear levels or to detect minute amounts of radioactive isotopes in environmental samples.

This is possible because the energy needed for the creation of an electron/hole pair (ε_{pair}) is two orders of magnitude smaller than the energy used to create a photon in the scintillator detectors (for Ge $\varepsilon_{pair} = 2.96\,\text{keV}$ [58]). The Fano factor F, governing the fluctuation of a number of produced pairs, is also one of the smallest for this material (≈ 0.08 at 77 K [59]). Ge(Li) technology also allowed for the development of pioneering experimental methods by FIORINI in 1967 for searching for the $0\nu\beta\beta$ decay in ^{76}Ge.

A main drawback of the Ge(Li) technology is the necessity to keep the detectors at liquid nitrogen (LN_2) temperatures at all times to avoid the decompensation of p-type impurities. HPGe technology solved this problem by producing very clean crystals with net impurity concentration levels $|N_A - N_D|$[1] at $\sim 10^{11}\,\frac{1}{\text{cm}^3}$. Electronics grade germanium $|N_A - N_D| \sim 10^{13}\,\frac{1}{\text{cm}^3}$ is firstly cleaned with a zone refining technique, where the purification is achieved primarily through rejection of impurities into the liquid at the crystallization front [60]. The process cleans up material by ≈ 2 orders of magnitude to $10^{11}\,\frac{1}{\text{cm}^3}$. The final refinement is made by growing the crystal with the Czochralski method.

[1] N_A and N_D denote levels of acceptor and donor impurity concentrations, respectively.

Large size of the germanium detectors has an important consequence on the time profile of a generated signal. Considerable drift time of the charge carriers (in order of hundreds of ns) offers a possibility to extract additional information from the pulse shape. Drifting charge carriers from the multiple sites in the detector create a noticeable difference (even after processing by the front-end electronics), which allows distinguishing that pulse from the the one originating only from a single site. To illustrate the difference and describe the signal formation process in detail, two simple semi-analytical models for the pulse shape calculations are presented in the next section.

2.1 Signal formation in semiconductor detectors

In order to simulate the exact response of the germanium detector with deposited energy in a given location, numerical methods have to be applied (one can e.g. use ADL [62] or Siggen [63] packages for this purpose). This is mostly due to the irregular geometry of both semi-coaxial and BEGe type detectors, resulting in the complicated electric field distribution in the crystal. However, there are a few simple geometries, for which the electric field distribution can be determined analytically:

- planar,
- true-coaxial,
- (hemi)spherical.

The first case is already described in numerous semiconductor physics books to illustrate the electric field distribution in a flat p-n junction (especially for silicon detectors [58, 64, 65]) and it will not be discussed here. Common HPGe geometries are illustrated in Fig. 2.1. Left hand side geometries are used to approximate the electric field distributions in their right hand side counterparts. All the presented solutions for the spherical geometry can be applied to the hemispherical one by assuming that the interaction take place far from the top surface, where the electric field can be distorted due to boundary effects. The same assumption regards the true-coaxial case.

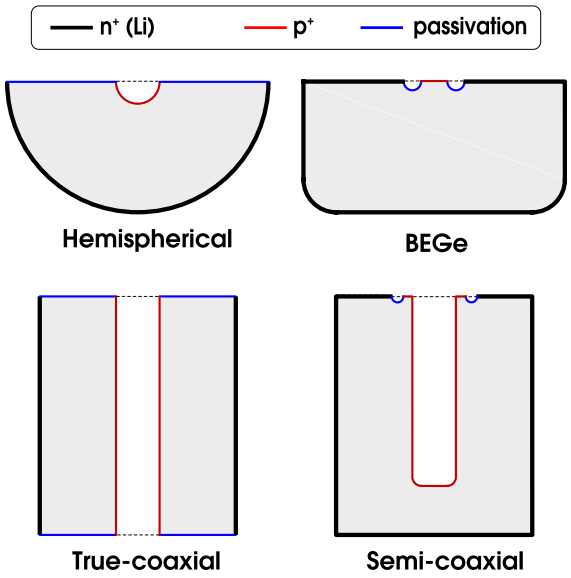


Fig. 2.1: Comparison of HPGe geometries (cross-sectional view). The left hand side geometries are used for model calculations – they approximate the electric field distribution in their right hand side counterparts. Thicker black lines show the Li-diffused n^+ contacts (p-type crystals are assumed), while the red lines represent p^+ contacts, fabricated by either boron implantation or metal evaporation (Schottky barrier) [28, 61]. Passivated surfaces are depicted with blue lines.

To obtain the induced current on the readout electrode, one can apply Ramo-Shockley theorem [66]:

$$i_{e,h}(t) = \pm q\, \vec{v}_{e,h}(t) \cdot \vec{E}_w(\vec{r}), \qquad \vec{r} = \vec{r}(t) \tag{2.1}$$

where:

$i_{e,h}(t)$ – current induced by a movement of charge carrier with charge $\pm q$,
$\vec{v}_{e,h}(t)$ – charge carrier velocity in a given time moment t,
$\vec{E}_w(\vec{r})$ – weighting field – electric field strength in the point $\vec{r}$, calculated without any charge inside the detector and assuming that the readout electrode is polarized with a unity potential; all other electrodes are grounded.
$\vec{r}(t)$ – position of the drifting charge carrier in a given time moment t.

To find the respective components of Eq. 2.1, one needs to:

$\pm q$ – simulate the position of interactions and deposited energy in the detector e.g. when irradiated with a radioactive source. In the case of the germanium detectors, the most often used programs are GEANT4 [67] and EGS5 [68]. A conversion between the deposited energy and charge represented by the carriers (electrons and holes) is done using the mean energy value for the electron-pair creation (ε_{pair}). The value is greater than the bandgap energy for a given semiconductor, since it also takes into account the energy losses which do not result in a creation of charge carriers (e.g. lattice excitation by phonons).

$\vec{v}_{e,h}(t)$ – in general, the drift velocity is governed by the electric field amplitude. Linear dependence is often used for the silicon detectors using the carrier mobility μ as a scaling parameter ($\vec{v}_{e,h}(\vec{r}) = \pm\mu_{e,h}\vec{E}(\vec{r})$). The linear formula allows for the purely analytical solution of the Ramo-Shockley formula, however it stops being valid for the strong electric fields [64]. On the other hand, for the germanium detectors a phenomenon of the drift velocity saturation appears even for weak fields and the drift velocity/electric field dependence stops being linear[4]. Anyhow, the electric field vector has to be known for a given $\vec{r}$ to determine the drift velocity. It can be found e.g. by solving the Poisson equation.

$E_w(\vec{r})$ – while the actual movement of the charges is governed by the electric field $\vec{E}(\vec{r})$ and the drift velocity, the induced charge/current can be calculated easier with the use of the weighting potential/field [66]. The calculation of $\vec{E}_w(\vec{r})$ is simpler than $\vec{E}(\vec{r})$, since one needs to solve the Laplace equation instead of the Poisson one – this is because of the assumption of an absence of the space charge ρ inside the detector. Boundary conditions assume that the readout electrode is biased with the unity potential and all other electrodes are grounded[5] .

The final induced current pulse $i(t)$ is the superposition of the components $i_{e,h}$ calculated for every moving charge carrier. In practice, this is of course unfeasible and charges are grouped into clusters to minimize the computation time.

2.1.1 Ramo-Shockley theorem

An example of the analysis of induced signals and the electric field inside a p-type germanium detector can be demonstrated by performing the derivation in one dimension (e.g. radius in the polar or cylindrical coordinates). The first step is to solve the Poisson and Laplace equations in order to obtain the electric field $E(r)$ and the weighting field $E_w(r)$ distributions, respectively:

$$\nabla^2\phi(r) = -\frac{\rho}{\varepsilon\varepsilon_0}, \qquad \rho = -qN_A \qquad\qquad \nabla^2\phi_w(r) = 0 \tag{2.2}$$

where: ρ is charge density (assumed to be constant), N_A – acceptor impurities concentration, $\phi(r), \phi_w(r)$ – electric potential and weighting potentials respectively. A short remark about the sign convention: in the p-type semiconductor, acceptor atoms donate holes to the conduction band (or, equivalently, attract the electrons from the band). From the charge equilibrium point of view, the atoms are then negatively charged and counterbalanced by positive holes. In the depletion region free moving charges are swept away and negatively charged atoms remain in the p-type region. Therefore, ρ will have a negative sign. Similar reasoning can be applied (together with the suitable Poisson equation) to the n-type region (Li diffused layer). However, since there is a large disproportionality in the doping levels (p-type $\sim 10^{10}\,\frac{1}{\mathrm{cm}^3}$, n-type $\sim 10^{17}\,\frac{1}{\mathrm{cm}^3}$ [69]), the depletion volume is virtually only in the p-type region [64].

[4] The problem is drift velocity and mobility in the germanium detectors is discussed further in the present section

[5] It should be kept in mind that the weighting field is *not* the actual electric field in the detector. It is rather a mathematical method of to determine the signal shape at the *electrode of interest*. Therefore, the boundary conditions are different than the real bias voltages and even can be of different polarities.

Boundary conditions. Since Eq. 2.2 is a second-order differential equation, the solution will contain two constants from integral calculations. They can be found using the following boundary conditions (r_{p^+} and r_{n^+} denote the radii for p^+ and n^+ electrodes, respectively):

$$\begin{cases} \phi(r_{p^+}) = 0 \\ \phi(r_{n^+}) = V \end{cases} \qquad \begin{cases} \phi_w(r_{p^+}) = 1 \\ \phi_w(r_{n^+}) = 0 \end{cases} \tag{2.3}$$

assuming a p-type device, with the signal readout at the p^+ electrode (boundary conditions for the weighting potential ϕ_w) and a bias voltage V (at the n^+ electrode).

Depletion voltage. Next, the electric field can be calculated from the relation: $E(r) = -\nabla\phi(r)$. For the p-type detectors, the depletion layer usually extends from the outer radius[6], where the p-n junction is formed, to the inner core. When the bias voltage equals to depletion voltage V_{dep}, the electric field finally reaches the p^+ contact at r_{p^+}:

$$E(r_{p+}) = 0 \quad \text{for} \quad V = V_{dep} \tag{2.4}$$

Using the above equation, the value of the depletion voltage V_{dep} can be calculated for the given geometry and impurity concentration distribution .

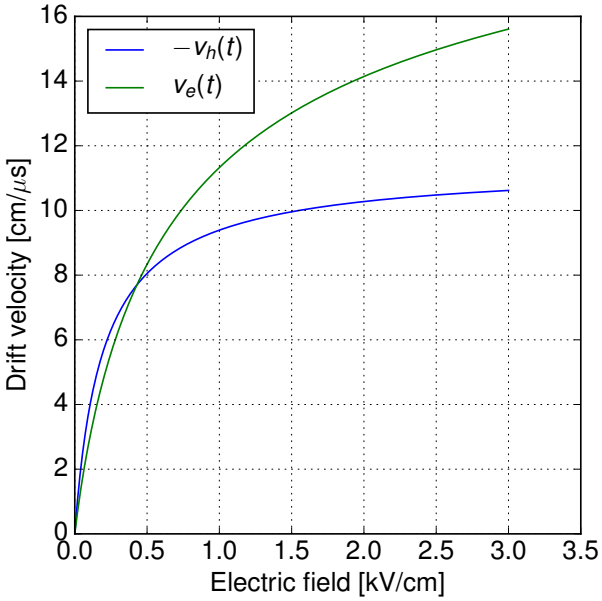


Fig. 2.2: Drift velocity dependency on the electric field in the HPGe detectors. Curves show the model described by Eq. 2.7, model parameters come from [70] for the $\langle 100 \rangle$ crystallographic direction. It is clear that due to the highly different slopes for low and high electric field parts of the curves it is difficult to approximate a whole curve with a linear model.

Linear/constant drift velocity models. After the calculation of $E(r)$ and $E_w(r)$ the only remaining component is the drift velocity $v_{e,h}(r)$. As it was mentioned before, several models can be used for its determination. For the silicon devices the linear velocity dependence on the electric field is often used [64]:

$$v_{e,h}(r) = \pm\mu_{e,h}E(r) \tag{2.5}$$

However, for germanium the constant velocity model is more suitable (for the fields in the order of 1000 V/cm, described e.g. in [71]), although such high fields can be present only with the enough overvoltage (e.g. bias voltage larger by 20-50% than V_{dep}). The constant velocity condition simplifies the final solution of Ramo-Shockley theorem – $r(t)$ can be simply calculated as (it is assumed that v_h has a negative sign):

$$r_{e,h}(t) = r_0 + v_{e,h}(t)\,t \qquad r_0 - \text{interaction position.} \tag{2.6}$$

Therefore, the above relation can be used to eliminate the r dependence in the $E_w(r)$ component of Ramo-Shockley theorem (Eq. 2.1) and write it as a function of time t. However, if the drift velocity is assumed to be constant, it does not depend on the actual electric field $E(r)$ and therefore on the impurity concentration $|N_A - N_D|$. Neither does the weighting field, since by definition it is calculated for $\rho = 0$.

[6] Assuming the n-type doping at the outer radius (by e.g. Li thermodiffusion).

Drift velocity saturation. To take into account the influence of the impurity concentration $|N_A - N_D|$ on the pulse shape, a more realistic drift velocity dependence has to be included in the equations. Proper characterization of $v(E)$ can be performed by using an empirical velocity saturation model, reported for the first time in [72]:

$$v(E) = \frac{\mu_0 E}{\left(1 + (\frac{E}{E_0})^\beta\right)^{1/\beta}} - \mu_n E \tag{2.7}$$

The term containing μ_n was added by Mihailescu et al. [73] to take into account Gunn effect for hot electrons. The expression is plotted in Fig. 2.2, with the equation parameters taken from [70], where the extensive discussion concerning the anisotropy of drift velocity in germanium can be found. Velocity saturation for holes is observed for values just over 500 V/cm and, as will be shown shortly on quantitative examples, these are rather low values of the electric fields, obtain only for the low bias voltages.

Trajectory calculation. Since the velocity in general is a function of the electric field $E(r)$, to obtain $v_{e,h}(t)$ one needs to relate r with t. The trajectory $r(t)$ can be calculated as follows:

$$v(r) = \frac{dr}{dt} = v_{fE}(E(r)) \tag{2.8}$$

v_{fE} is a general expression for the drift velocity as a function of electric field (e.g. Eq. 2.7). The above differential relation, in simple cases of $E(r)$ and $v_{fE}(E(r))$, can be then rearranged and $r(t)$ can be calculated by integration (with the boundary condition of $r(t=0) = r_0$, where r_0 is a initial position of the drift carrier. For the silicon detectors the usually used formula for v_{fE} is the linear relation from Eq. 2.5, the formulas for the electric field in a planar device is also relatively simple [58, 64, 65]. However, to take into account the semi-empirical velocity saturation model $v_{e,h}(E)$ (Eq. 2.7) and electric field distribution for coaxial and spherical geometries (compared in the next section), a numerical algorithm (velocity Verlet [74]) was used to derive the charge carrier trajectory $r(t)$, starting at r_0 (t vector is divided into n Δt steps):

$$\text{initial conditions:} \quad E_0 = E(r_0), \quad v_0 = v_{e,h}(E_0) \tag{2.9}$$

n-th step:

$$v_{n-1/2} = v_{n-1} + \frac{\Delta t}{2} \cdot v_{e,h}(E_{n-1}) \qquad \text{("half" step velocity)} \tag{2.10}$$

$$E_{n-1} = E(r_{n-1}), \quad \Delta r = v_{n-1/2} \cdot \Delta t \tag{2.11}$$

$$r_n = r_{n-1} + \Delta r \tag{2.12}$$

A main feature of the algorithm is the usage of the "half-step" velocity, which stabilizes the solution. This is especially important for the second-order problems, but can be also applied for the first order equations. The procedure is repeated until $r(t)$ is equal to r_{p^+} (holes) or r_{n^+} (electrons). Having calculated $r(t)$ for a given t, $E(t)$ and $v_{e,h}(t)$ are easily derived from $E(r)$ and $v_{e,h}(E)$ by substitution.

Undepleted detector. Another problem is a solution for the undepleted diode. Since the radius of the depletion layer r_{dep} is unknown before solving the Poisson equation, one need to include it in the boundary conditions of Eq. 2.3 and Eq. 2.4 instead of r_{p^+}. A new set of equations allow for the determination of r_{dep} as well as integration constants, which should be applied for the calculations with $V < V_{dep}$. Knowing the solution for the undepleted diode allows for the calculation of capacitance-voltage (C-V) curves. Comparison of the measured curves with the calculated ones can in principle determine the unknown impurity concentration [75].

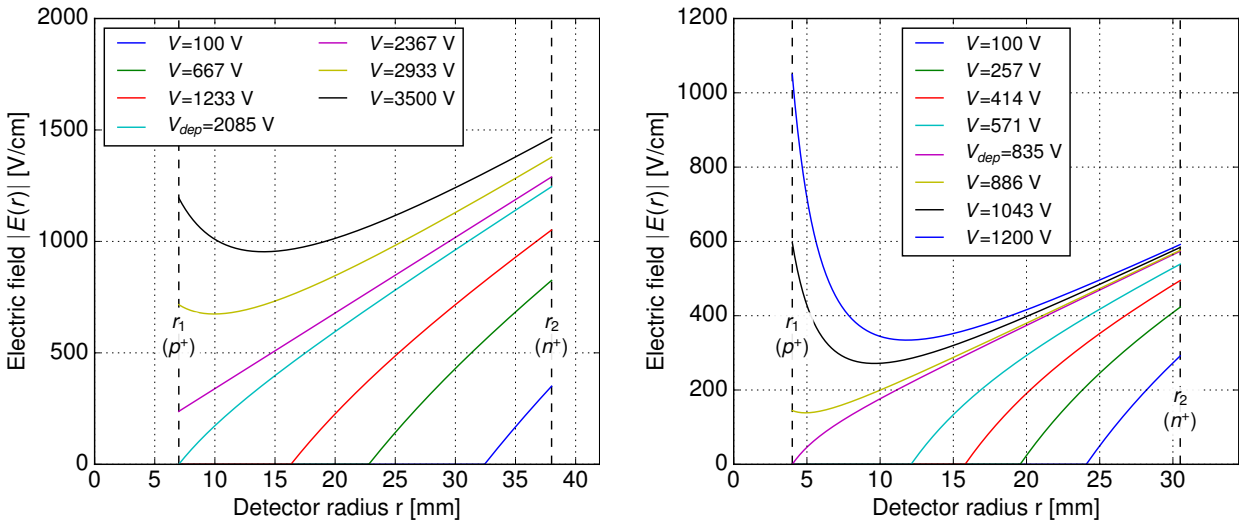
2.1.2 Comparison of true-coaxial and spherical models

Equations derived in the previous section were applied for two model geometries: true-coaxial and spherical, which will be used to illustrate pulse shapes from the semi-coaxial and BEGe type detectors, respectively. Since enriched semi-coaxial detectors in GERDA have a deep inner core, which extends almost to the bottom of the detector 2.1, such approximation should not be far from reality. On the other hand, the spherical approximation has a significant discrepancy in a groove and upper face regions, but the weighting field is still very peaked near the p^+ contact and calculated pulses have very similar shapes to the ones obtained from the purely numerical approach (see e.g. [76]).

Table 2.1: Comparison of equations needed for the calculation of the electric field distribution and the induced current pulses for true-coaxial and spherical models (in the p-type germanium region). Formulas for quantities like depletion voltage or capacitance are also included. The capacitance formula for the spherical model assumes that the detector is of a hemispherical shape (multiplied by a factor $\frac{1}{2}$).

	True-coaxial	Spherical
Poisson equation:	$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\phi}{dr}\right)=\alpha$	$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right)=\alpha$
$\phi(r)$:	$\frac{\alpha}{4}r^2+C_1\ln r+C_2$	$\frac{\alpha}{6}r^2-\frac{C_1}{r}+C_2$
$E(r)$:	$-\frac{\alpha}{2}r-\frac{C_1}{r}$	$-\frac{\alpha}{3}r-\frac{C_1}{r^2}$
C_1, C_2:	$C_1=\frac{V-\frac{\alpha}{4}(r_2^2-r_1^2)}{\ln\frac{r_2}{r_1}}, \quad C_2=\frac{C_1}{r_1}-\frac{\alpha}{6}r_1^2$	$C_1=\frac{\frac{\alpha}{6}(r_2^2-r_1^2)-V}{\frac{1}{r_2}-\frac{1}{r_1}}, \quad C_2=\frac{C_1}{r_1}-\frac{\alpha}{6}r_1^2$
V_{dep}:	$\frac{\alpha}{4}\left(r_2^2-r_1^2\right)-\frac{\alpha}{2}r_1^2\ln\frac{r_2}{r_1}$	$\frac{\alpha}{3}r_1^3\left(\frac{1}{r_2}-\frac{1}{r_1}\right)+\frac{\alpha}{6}\left(r_2^2-r_1^2\right)$
r_{dep} : $(V<V_{dep})$	$r_{dep}^2\left[1+2\ln\frac{r_2}{r_{dep}}\right]+\frac{4}{\alpha}V-r_2^2=0$	$r_{dep}^3-ar_{dep}^2+b=0$ $a=-\frac{3}{2}r_2, \quad b=\frac{1}{2}r_2^3-3\frac{Vr_2}{\alpha}$
$\phi_w(r)$:	$C_{1w}\ln r+C_{2w}$	$-\frac{C_{1w}}{r}+C_{2w}$
$E_w(r)$:	$-\frac{C_{1w}}{r}$	$-\frac{C_{1w}}{r^2}$
C_{1w}, C_{2w}:	$C_{1w}=\frac{1}{\ln\frac{r_1}{r_2}}, \quad C_{2w}=-\ln r_2 C_{1w}$	$C_{1w}=-\frac{1}{\frac{1}{r_1}-\frac{1}{r_2}}, \quad C_{2w}=\frac{C_{1w}}{r_2}$
$i_{e,h}(t)$: (const. $v_{e,h}$)	$-\frac{q\lvert v_{e,h}\rvert}{\ln\frac{r_2}{r_1}(r_0+v_{e,h}t)}$	$-\frac{q\lvert v_{e,h}\rvert}{\left(\frac{1}{r_1}-\frac{1}{r_2}\right)(r_0+v_{e,h}t)^2}$
Capacitance:	$\frac{2\pi\varepsilon\varepsilon_0}{\ln\frac{r_2}{r_{dep}}}z$	$\frac{2\pi\varepsilon\varepsilon_0}{\frac{1}{r_{dep}}-\frac{1}{r_2}}$

$$\alpha=-\frac{\rho}{\varepsilon\varepsilon_0}=\frac{qN_A}{\varepsilon\varepsilon_0}$$



(a) True-coaxial **(b)** Spherical

Fig. 2.3: Radial electric field distributions calculated for the true-coaxial ((a) – left panel) and spherical ((b) – right panel) models.

Solutions of the one dimensional Poisson and Laplace equations, as well as other parameters used for the description of the detector, are collected in Tab. 2.1. Expression for the current pulses $i(t)$ with the constant velocity model are also included for completeness. In both cases the inner and outer radii are denoted with r_1 and r_2, respectively. The equations for the partial depletion radius r_{dep} are in the implicit form – for all the following plots and results they were solved numerically. In the comparison two sets of parameters has been taken into consideration:

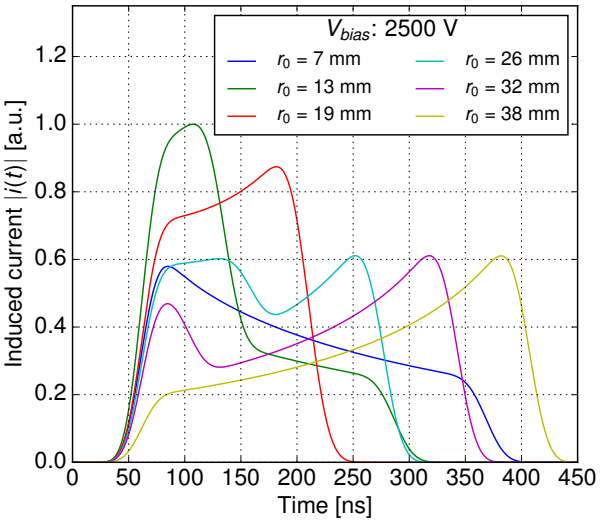
true-coaxial: $r_1 = 7\,\text{mm}$, $r_2 = 38\,\text{mm}$, $z = 90\,\text{mm}$,
$|N_A - N_D| = 0.6 \cdot 10^{10}\,\frac{1}{\text{cm}^3}$, calculated V_{dep}: 2085 V.

spherical: $r_1 = 4\,\text{mm}$, $r_2 = 30.5\,\text{mm}$,
$|N_A - N_D| = 0.5 \cdot 10^{10}\,\frac{1}{\text{cm}^3}$, calculated V_{dep}: 802 V.

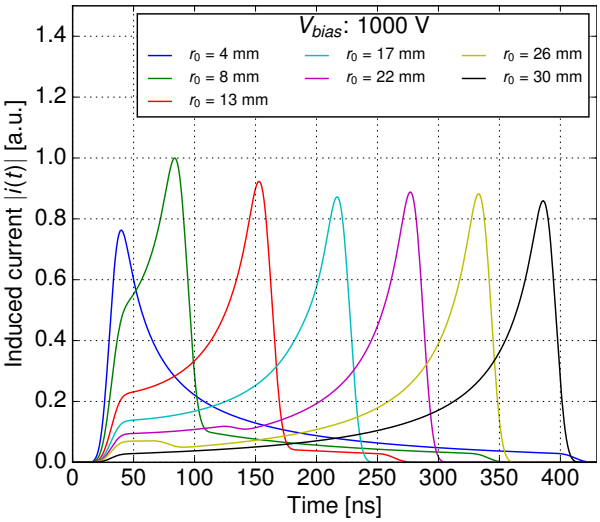
Electric field distribution. Fig. 2.3 shows the radial distribution of the electric field calculated for the true-coaxial (left) and spherical (right) models, for different values of the bias voltage V. Probably the biggest difference between models is the amplitude of the electrics fields at r_2 – for the voltages higher than V_{dep} there is a steady increase for the true-coaxial detector, while in the case of the spherical one the field has almost the same amplitude for all $V > V_{dep}$. This is because for the small anode detectors like BEGe or point-contact ones, the field near the n^+ electrode depends mostly on the impurity concentration gradient [77], rather than on the actual bias voltage. Another distinction is a much larger relative difference between the field amplitudes at the p^+ and n^+ electrodes in the spherical model.

Induced current pulse shape. The most important factor from the PSD point of view are the pulse shapes of the current induced on the readout electrodes. Fig. 2.4 shows calculated pulse shapes for both models, with the inclusion of the drift velocity saturation effect (Eq. 2.7). Pulses were calculated for several values of r_0 to show the dependence of their shapes on the energy deposition locations along the radius. To simulate the effect of the electronics' limited bandwidth, pulses were smoothed with a $5 \times 20\,\text{ns}$ moving average filter.

Current signals from the spherical detector (Fig. 2.4b) have a simple, single-spike structure, preceded by an almost flat pedestal. The shape comes from a very uneven weighting field $E_w(r)$ distribution in the detector volume (compare the $E_w(r)$ distributions for both geometries in Fig. 2.5)



(a) True-coaxial



(b) Spherical

Fig. 2.4: Simulated pulse shapes for single-site energy depositions for the true-coaxial (left panel) and the spherical (right panel) detector models. Interaction sites were selected in the multiple points along the radius. Drift velocity saturation effect was included in the simulation.

– the charge carriers drifting in the low $\vec{E}_w(\vec{r})$ region create only pedestal part and the spiked pulse is induced only when a charge carrier is drifting near the p^+ contact. If an interaction site is far from the p^+ electrode, the pulse is practically only induced by moving holes. This is caused by electrons drift toward n^+ contact, where the weighting field $\vec{E}_w(\vec{r})$ amplitude is low and therefore the induced current is small.

Because the amplitude ratio between the current peak and the pedestal is relatively large, the observed signal will gain a significant amplitude only when the the charge carrier is already close to the p^+ electrode. This fact makes it difficult in real (noisy) conditions to determine the time when the charge started its drift. For the real BEGe detector geometry the ratio between the peak and the pedestal is even larger – for comparison see [76], where a numerical simulation was made that took into account the real geometry, as well as the drift velocity anisotropy.

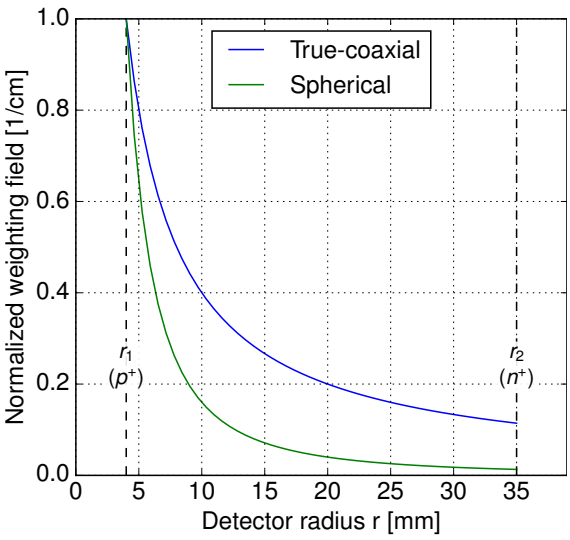


Fig. 2.5: Comparison of the weighting fields in true-coaxial and spherical detectors. Same dimensions were used in both cases ($r_1 = 4$ mm and $r_2 = 35$ mm).

An important property of the signals in the spherical model is a significant drift time in the detector. If the incoming γ-ray interacts with the detector in two sites (e.g. for $r = 25$ and 30 mm), the observed pulse will be a superposition of the individual pulses from each position. The calculated pulse shows (Fig. 2.4b) that a double peaked shape will be obtained, because the delay between the arrivals of charge carriers from both sites is long. This fact makes it possible to discriminate such signal from the one resulting from a single deposition. Furthermore, the maximal current amplitude will be smaller for the multi-site energy deposition, since the area under the current curve is proportional to energy E, same for both pulses. This is a reasoning behind the A/E method, which will be described in detail in the next section.

Pulses in the true-coaxial detectors.

The A/E approach is not effective when applied to the coaxial detectors – from Fig. 2.4a it is clear that pulse shapes have a strong dependence on the interaction position. In general, three types of events can be distinguished:

- $r_0 \approx r_1$ – large electron component contribution, peaked in the initial part (see the top panel of Fig. 2.6),
- r_0 half-way between electrodes – significant electron and hole components, containing two current peaks.
- $r_0 \approx r_2$ – pulse peaked at the end.

Since a single-site energy deposition can have both single- and double-peak structure, the A/E method cannot be applicable. This is illustrated in Fig. 2.7a, where SSE at $r_0 = 26\,\text{mm}$ and MSE with $r_0 = 24$ and 29 mm (same amount of energy is deposited for both events) in the true-coaxial detector can result in the very similar shapes, even if the interaction vertices are separated by 5 mm. The pulse shapes are almost impossible to distinguish by any PSD method, especially in the presence of noise. On the other hand, pulses from the spherical detector are easily distinguishable by both shape (double-peaked structure) and the maximal current amplitude (Fig. 2.7b). They would be indistinguishable only if the energy was deposited in two places, with the similar amount in every site and on the same radius r_0 (but e.g. different azimuthal angles φ). Therefore, more sophisticated methods are needed for the efficient discrimination between SSEs and MSEs in the (true-/semi-)coaxial detectors, which are discussed in the next section.

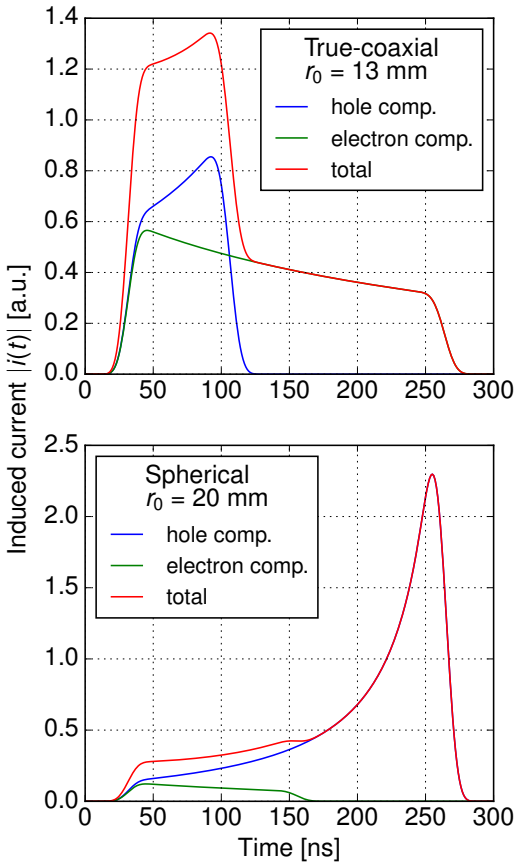
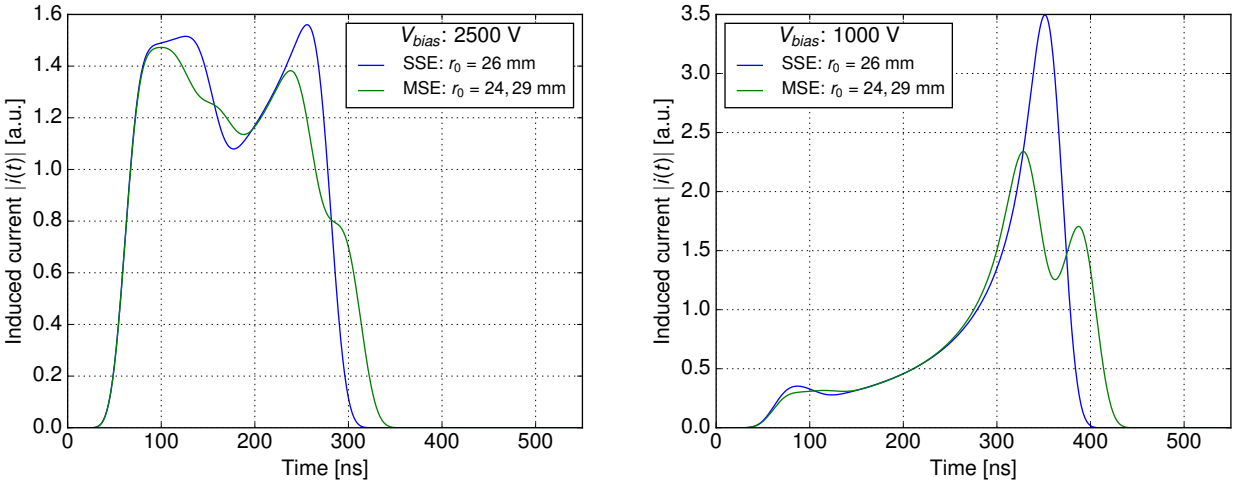


Fig. 2.6: Induced current signals and the respective electron and hole components for the true-coaxial and spherical models.



(a) True-coaxial **(b)** Spherical

Fig. 2.7: Comparison of the possible SSE and MSE in both coaxial (left) and spherical (right) detector models. For some combinations of interaction site positions the MSE pulse shape can be almost indistinguishable from the one resulting from the SSE. Resulting induced charge is the same for both waveforms, for the MSE pulse the individual components are included with equal weights.

2.2 Pulse Shape Discrimination

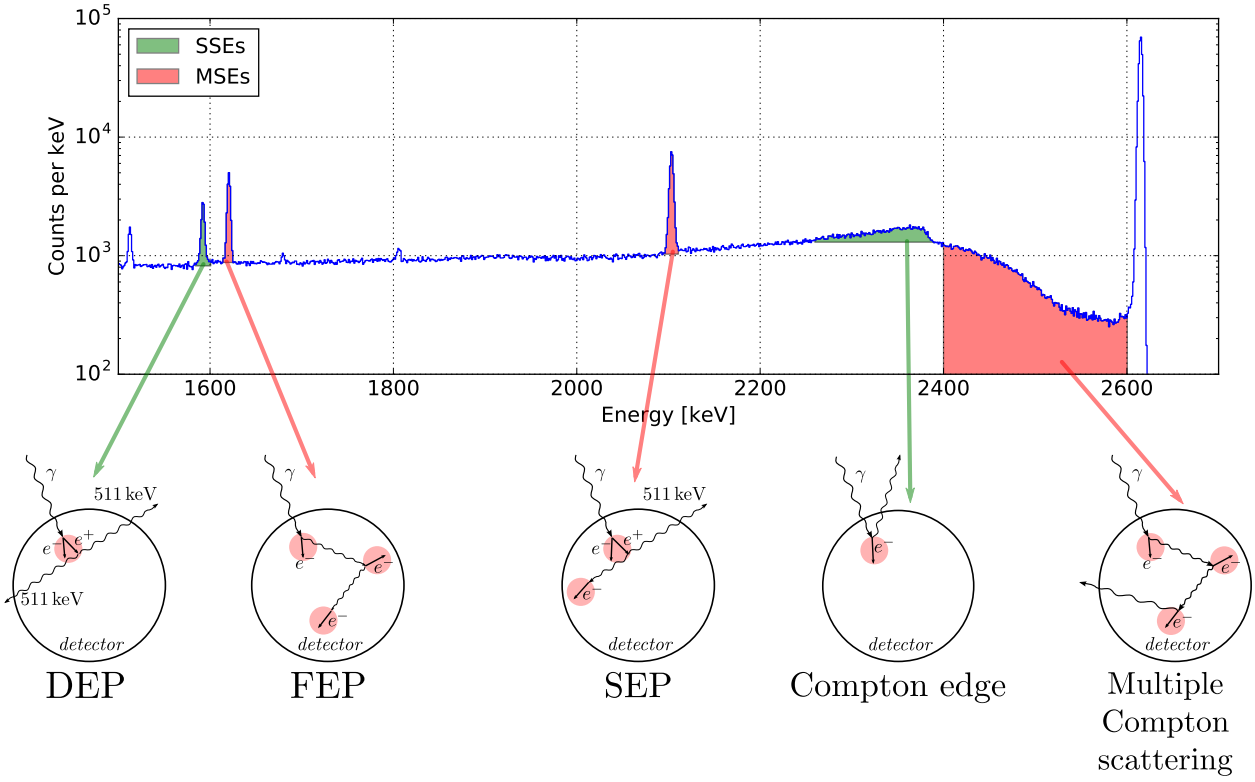


Fig. 2.8: ^{228}Th energy spectrum with the marked energy regions rich in single- (green) and multi-site events (red). Event topology characteristic for each region is shown below. Red circles in the schematic drawing of the detector are the sites of the energy deposition. Abbreviations: DEP – Double Escape Peak, FEP – Full Energy Peak, SEP – Single Escape Peak, SSE – Single-Site Event, MSE – Multi-Site Event.

A main problem in the γ-ray related background suppression via PSD is to differentiate between two types of the events in the detector: SSEs and MSEs. As the name indicates, the main difference between them is a number of interactions inside a detector. The $0\nu\beta\beta$ decay is of SSE type, since both electrons from the decay deposit their energy in a small volume (within a range of less than 1 mm). The same is true for $2\nu\beta\beta$ decay events. Therefore, using the PSD one can reduce the background from MSEs e.g. in the $Q_{\beta\beta}$ energy region. On the other hand, in the case of γ-ray spectrometry, high-energy peaks consist mostly of MSEs[7] and the Compton continuum background is a mixture of both SSEs and MSEs. The inverse procedure, that is eliminating the SSEs, could thus reduce a number of background events in the Compton continuum.

In general, every PSD method needs a "clean" sample of data of both classes (single- and multi-site), either for the training process, the cut position determination or the efficiency estimation. The term "clean" means in this context that ideally the data should contain events from only one of the classes. For γ-ray spectrometers almost clean dataset can be obtained by irradiating the detectors with radioactive sources emitting high-energy γ radiation. This is due to the fact that escape peaks can be formed only if the incoming γ-rays have energy E_γ larger than the e^-e^+ pair creation threshold of 1022 keV. Fig. 2.8 shows a gamma energy spectrum of ^{228}Th registered by the HPGe detector. Colored parts of the spectrum are marked to show energy regions with well defined event topologies (SSE/MSE), like FEPs, escape peaks or Compton edge (CE). ^{228}Th is often chosen for a PSD calibration, since it emits, among others, a γ-ray with the energy of 2614.5 keV, which exceeds the threshold for the pair creation.

[7] It is estimated that FEPs with the energies over 700 keV contain more than 90% MSEs [78, 79].

The kinetic energy of a pair created by the incoming γ-ray is $E_\gamma - 1022\,\text{keV}$. On average, it is equally divided between the electron and the positron [80]. Since electrons (and positrons) have rather short projected range in germanium (under 1 mm), all the energy is deposited in a small volume. From the induced signal point of view it can be treated as a single site. Therefore, the first interaction site (vertex) is where the pair was produced. Since the positron has lost its kinetic energy[8], it annihilates with an electron from the detector material. The created 511 keV γ-rays can e.g. escape from the detector's volume. In this case, the total deposited energy in an event is $E_\gamma - 1022\,\text{keV}$ and can be observed in the energy spectrum as DEP. Since there was only one interaction site, the DEP events can be treated as a clean sample of SSEs.

Alternatively, one or both 511 keV γ-rays can be absorbed in the detector. If the first situation happens, the deposited energy is equal to $E_\gamma - 511\,\text{keV}$ and gives rise to the single escape peak (SEP). In the latter case, both γ-rays undergo absorption and the event appears in the FEP at E_γ in the energy spectrum. Of course, the intermediate topologies are also possible (e.g. with a partial energy deposition from one or both γ-rays) and they contribute to the continuous spectrum in the $(E_\gamma - 1022\,\text{keV})$ to E_γ energy range.

Another sources of the single-site and multi-site events are two continuous parts of the spectrum, namely CE and multiple Compton scattering (MCS) regions. The maximum value of the energy transferred to the electron in a single act of a Compton scattering for a γ-ray with $E_\gamma = 2614.5\,\text{keV}$ is 2381.7 keV (backscattering), therefore, selecting events with energies around this value results in a sample containing a large fraction of SSEs. Events with higher energies have to be of multi-site type, since the energy had to be deposited via a pair-production and an absorption/scattering of the annihilation radiation. Alternatively, the energy can be deposited by the MCS, possibly with a photoelectric absorption as a last step for a total energy deposition. Either way, the event will have the multi-site topology and can be used in the PSD training process.

To set a cut value for a given PSD method, some arbitrary condition has to be selected. In the $0\nu\beta\beta$ decay experiments like GERDA, the signal (SSEs) efficiency is usually defined for 90% acceptance of the DEP events. It is calculated as the ratio of a number of counts in the peak (after the flat background subtraction) after and before the application of the PSD cut. The peak counts can be calculated using a Gaussian fit, but the trapezoidal method [82] also provides good results and is faster from the computational point of view. Similarly, acceptances (or survival probabilities, as those two terms can be used interchangeably in this context) are also calculated for other peaks in the spectrum (e.g. FEPs and SEPs) to asses the PSD efficiency for vetoing MSEs.

2.2.1 BEGe detectors – the A/E classifier

In the analysis of the spherical model (previous section) it was shown that there exists a simple signature in the pulse shape: SSEs are characterized with a single current pulse peak, while MSEs have multiple peaks. Additionally, since the area under the pulse has to be constant in both cases (equal deposited energy), the maximal current amplitude is smaller for the MSEs. This property allows for the discrimination between SSEs and MSEs by calculating the ratio of the maximal current pulse amplitude A and preamplifier output pulse height E. Since the charge-sensitive preamplifier integrates the current pulse from the detector, to obtain back the current pulse the signal has to be differentiated. This approach is called A/E and it is a simple, univariate classifier developed for the small anode detectors, like the ones of BEGe type (produced by Canberra company) or P-type Point Contact (PPC) [83]. The difference between the two types is that PPC detectors have even smaller contact than the BEGe ones and therefore extremely small capacitance ($\approx 1\,\text{pF}$ or lower). However, the readout contacts in both cases are small when compared to the

[8] This is not a definitive requirement, since annihilation can also take place for non-thermalized positrons – annihilation of such positrons has been observed experimentally for example by Deutsch [81].

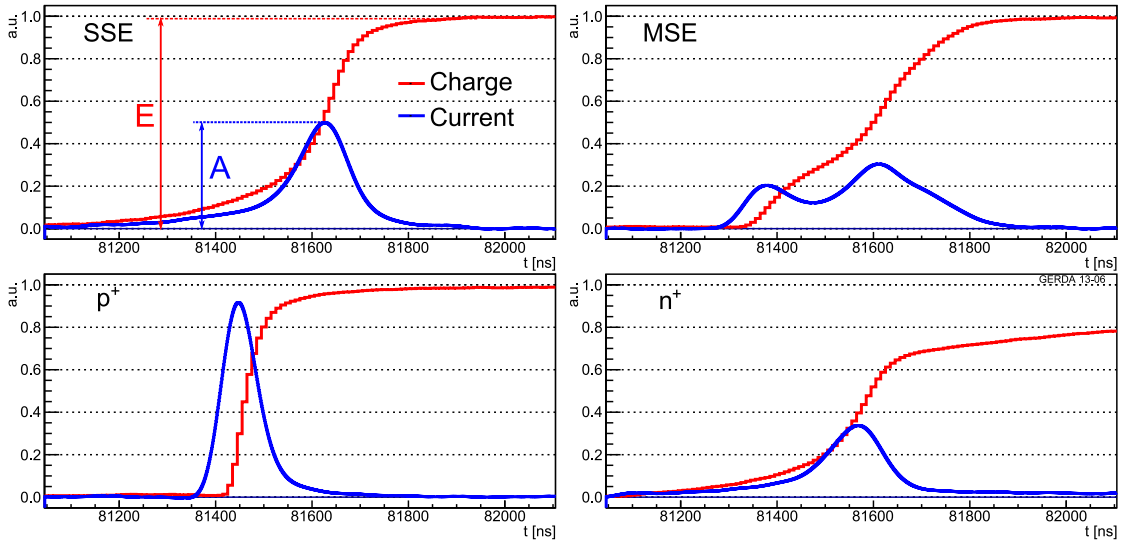


Fig. 2.9: Preamplifier traces characteristic to the different event topologies in the BEGe detector [45].
Top left – Single-Site Event, topology typical to the $0\nu\beta\beta$ decay and pair production interaction with an escape of annihilation γ-rays.
Top right – Multiple-Site Event, multiple peaks in the current signal can be seen, each one from an interaction point in the detector's active volume. Maximal current amplitude, A, is smaller than in its SSE counterpart.
Bottom left – a fast signal typical to the interaction near p^+ contact (caused most probably by α or β emitters). The A/E value is greater than for the SSEs.
Bottom right – a so called "slow pulse", caused by the charge diffusion (rather than drift) from the transitional n^+ layer.

outer dimensions and therefore the weighting field is strongly peaked near the p^+ contact. Because of that, the PSD principle is demonstrated quite well for both of them by the spherical model.

The A/E method was extensively studied as a part of the R&D program for GERDA Phase II BEGe detectors [47, 76, 84, 85], as well as for the PPC detectors in MAJORANA DEMONSTRATOR experiment [86]. The application of the method is illustrated in Fig. 2.9 for four classes of events that one may encounter in the BEGe detector. Aside from SSE and MSE pulses, very fast pulses with higher A/E are produced when energy is deposited on the p^+ contact (e.g. by the α/β emitters). On the other hand, slow pulses are obtained when β particles are registered on the n^+ electrode – the slow component is due to the diffusion of the charge to the active volume from transition layer [43, 87], which is created during the n^+ contact fabrication by lithium thermodiffusion. Since the layer is much thicker than the range of α particles in germanium, only β emitters can induce the pulses on the n^+ electrode.

The A/E classifier, calculated for each event from the ^{228}Th dataset, can be plotted as a function of energy on a 2D histogram (Fig. 2.10, taken from [56], where the extensive description of the A/E based PSD analysis for GERDA Phase II can be found). The distribution is normalized such that the SSE band centroid is equal to 1. Additionally, a small correction due to the linear energy dependence was also applied. As shown in Fig. 2.10 the MSEs have the A/E values in the range of $0.5-0.9$, while the fast pulses from the p^+ contact (α events) have classifier values > 1.05. A double-sided cut can be therefore applied to veto the MSEs and α events, while preserving SSEs. In GERDA Phase II the lower cut (A/E_{low}) was set to obtain 90% DEP acceptance. The upper one was fixed to have $A/E_{high} = 1+2(1-A/E_{low})$. Therefore, the accepted events lie between A/E_{low} and $1+2(1-A/E_{low})$. The upper cut further reduced the DEP acceptance by 2.68% and in total (after applying both cuts) it was equal to $(87.32\pm0.21)\%$ [56].

Similar method is also used for the detectors in the MAJORANA DEMONSTRATOR experiment, yielding less than 10% survival probability of SEP (at 90% DEP acceptance) [86]. The vacuum

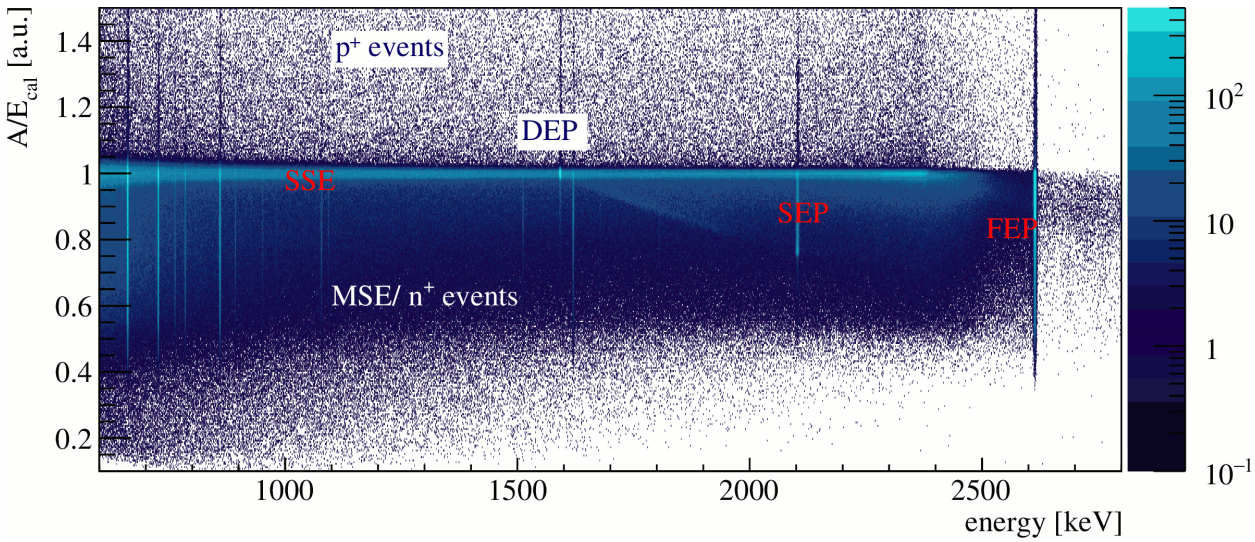


Fig. 2.10: A/E classifier as a function of the energy, plotted for the ^{228}Th events. The parameter was corrected for the linear energy dependence and normalized such that the SSE band centroid is at 1. Figure taken from [56].

cryostat conditions are always favorable in terms of noise and signal quality – in GERDA Phase II the Very Front End, consisting of an input junction field-effect transistor (JFET) and a feedback loop, is not screened from the electromagnetic noise, because putting an opaque, conducting shroud (from e.g. copper) around the detector would block the transmission of the LAr scintillation light. The Very Front End is therefore prone to the pickup of the electromagnetic interference, which can decrease the performance of the PSD, especially if the noise occupy the same frequency bandwidth as the signal. The achieved survival probabilities for the ^{228}Th data (all BEGe detectors included in the analysis) in GERDA Phase II were [56]:

- $(14.90 \pm 0.29)\%$ for the FEP at 1620.5 keV,
- $(10.62 \pm 0.16)\%$ for the SEP at 2103.5 keV,
- $(14.60 \pm 0.02)\%$ for the FEP at 2614.5 keV.

The values were calculated for all available events in the analysis, i.e. averaged over all detectors.

Extensive tests were performed for the BEGe detectors still mounted in the vacuum cryostats, just after receiving them from the manufacturer, in the underground laboratory HADES [51]. Therefore, the A/E efficiency can be compared for the same detectors, but before their installation in the GERDA setup. The cut was again set for 90% DEP acceptance and the achieved survival probability of the MSEs peaks were: 7.1–22% and 5.7–14.7% for 2614.5 keV FEP and 2103.5 keV SEP, respectively. The values are given in ranges because the exact efficiency depends on a given diode. In general, they are better than in GERDA Phase II, for the reasons mentioned before.

2.2.2 Semi-coaxial detectors - multivariate classifiers

As it was pointed out in Sec. 2.1.2, due to differences in signal shapes between the small anode and coaxial detectors, the A/E approach does not yield good results for the latter – similar maximal current amplitude (A) is possible for both SSEs and MSEs, even if the interaction sites happen to be at different radii. Since the semi-coaxial detectors were originally used in the earlier $0\nu\beta\beta$ decay searches like IGEX and Heidelberg-Moscow experiments, the first approaches to the PSD problem were tried as soon as obtaining the digitized pulse was possible. For example, in IGEX the PSD method was based on peaks and lobe counting [88], while the Heidelberg-Moscow analysis used drift times of charge carriers in the crystal [89]. Later, a method based on artificial neural-networks was applied to Heidelberg-Moscow data [79], however, the number of used neurons was rather large (180 in the input layer and 80 in the hidden layer). There was a good agreement with the

previous method – generally, for the $\approx 70\%$ DEP acceptance the 1620.5 keV FEP was reduced to $\approx 30\%$ (the exact value depends on a given detector).

In GERDA Phase I three independent PSD methods were used for the semi-coaxial detectors data [45]: the artificial neural-network (TMlpANN algorithm), the Projective Likelihood classifier and the one based on the current pulse asymmetry. The first two are implemented in the Toolkit for Multivariate Data Analysis (TMVA) package [90], which is included in the ROOT framework [91]. Since GERDA data is stored in the ROOT binary format, the package is very convenient to use. The leading method used for the $0\nu\beta\beta$ analysis was TMlpANN, which used the input variables from entire rising edge of a charge pulse. First, the amplitude of the pulse was normalized to 1 and horizontally shifted such that $t = 0$ corresponds to the half of its maximal amplitude. Then, time values when the pulse reaches $1, 3, 5, \ldots, 99\%$ of the normalized amplitude are determined (since the sampling period is 10 ns, linear interpolation is needed between the sampled points). The calculated 50 time values are then used as the input data to the artificial neural-network. Before processing the pulse, its baseline is subtracted and it is smoothed with a moving window average filter (80 ns integration time). A training sample of SSEs is taken from the DEP (^{228}Th calibration) and MSEs pulses come from the FEP at 1620.5 keV.

The second method, based on the Projective Likelihood classifier, uses a different pulse pre-processing procedure – for each pulse, a current maximum is found by the differentiation, which centers the time position for the pulse (maximal current value at $t = 0$). Similarly to the previous method, a baseline of the pulse is subtracted and then it is normalized amplitude-wise. 8 input variables are used in total, each is a sum of four consecutive pulse height values. The energy regions used for the training are different – CE events are used as a sample of SSEs and MCS region as a sample of MSEs. Since both training regions and the method are vastly different, Projective Likelihood and TMlpANN can be treated as independent methods.

The last method combines the A/E classifier with the additional parameter, based on the current pulse asymmetry (A_s) [45]. The asymmetry A_s is defined as:

$$As = \frac{\sum_{i=0}^{i=n_m} i(t) - \sum_{i=n_m}^{i<200} i(t)}{\sum_{i=n_m}^{i<200} i(t)} \tag{2.13}$$

where: $i(t)$ is the current pulse and n_m is the index of the sample corresponding to the current maximum value. Noise-removing moving window average filter with length L is applied to the pulse before differentiation. The parameter A_S is combined with the A/E into a single classifier:

$$q_{AS} = A/E \cdot (c + A_S) \tag{2.14}$$

Contrary to the previous methods, which use the fixed cut, here the cut value is optimized using the following quantity S:

$$S = \frac{\varepsilon_{\mathrm{DEP}}}{\sqrt{f_{bkg} + 3/N_{bkg}}} \tag{2.15}$$

where $\varepsilon_{\mathrm{DEP}}$ is the DEP survival probability, f_{bkg} is a survival fraction of the events from the physics dataset in the $1700 - 2200$ keV. For each detector a set of optimal parameters c and L is found that maximizes S. The term with N_{bkg} (total number of background events) is used so that the parameters are not optimized for the zero background.

2.2.3 TMVA package and multivariate classifiers

The TMVA package provides an environment for the processing, parallel evaluation and application of multivariate classification. It is closely integrated with the ROOT framework [91], which is a standard tool for the data processing in high energy and nuclear physics. Input dataset, used for the method training and testing, can be directly provided as a ROOT tree or an ASCII file. Two of the classification methods from TMVA were studied in this work: Projective Likelihood and Multi-Layer Perceptron (MLP) artificial neural-networks and will described in the next sections. Apart from them, other classifiers like Boosted Decision Trees (BDT), Support Vector Machine (SVM) and k-Nearest Neighbour (KNN) etc. are also available in TMVA.

2.2.3.1 Projective Likelihood

Projective Likelihood, also known as the naïve Bayes classifier, can provide an optimal separation between the signal and the background, provided that no correlations between the input variables are present. However, in most practical cases the input variables are correlated to some degree, either in a linear or non-linear manner.

The likelihood ratio for a given event, represented by a vector of input variables $x_1, x_2 \ldots x_{Nvar}$, is defined as:

$$y_{\mathscr{L}} = \frac{\mathscr{L}_S}{\mathscr{L}_S + \mathscr{L}_B}, \quad \mathscr{L}_{S,B} = \prod_{i=1}^{N_{var}} p_{i\,S,B}(x_i) \tag{2.16}$$

where: S (B) index denotes the signal (background), $p_{i\,S,B}$ is the Probability Density Function (PDF) for i-th variable from vector x . The $y_{\mathscr{L}}$ value defines the likelihood of the event belonging to the signal vs the background class, but in order to calculate it, probability density function (PDF) for each input variable has to be known.

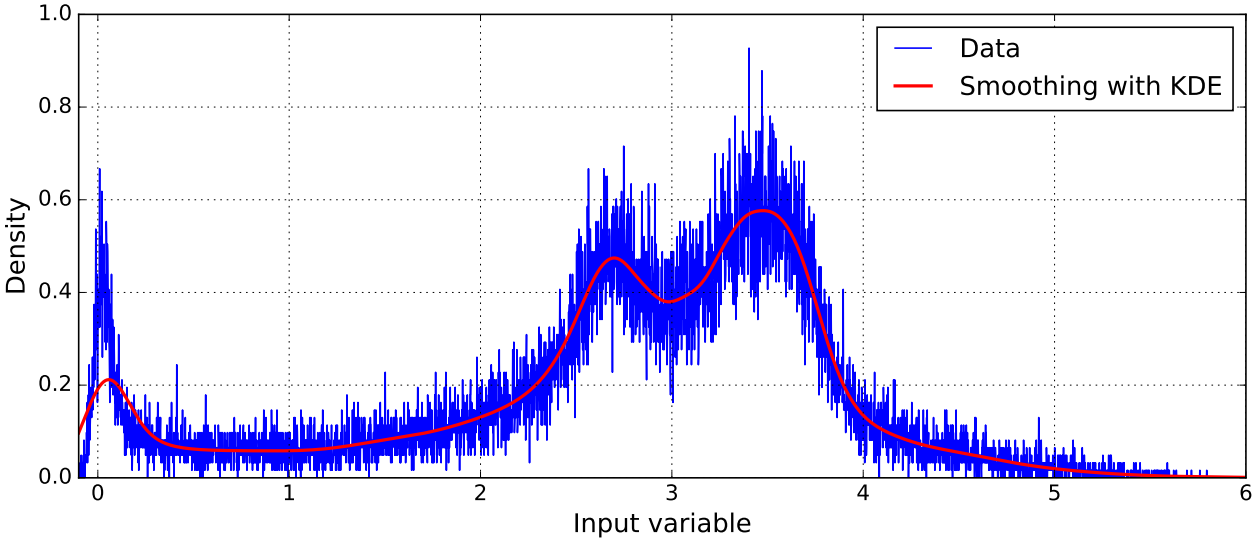


Fig. 2.11: Illustration of the gaussian KDE smoothing algorithm, used for in the "training" process of the Projective Likelihood classifier. The blue histogram shows the distribution of raw input data (e.g. summed amplitudes extracted from a waveform – compare with Fig. 2.12). The red line shows the new, smoothed distribution, extracting the featured from the input data, but eliminating the statistical variations. Binning density and a smoothness factor of the new distribution are provided by the user and can be adjusted to accomplish the best . The red distribution is saved and later used in the classification process.

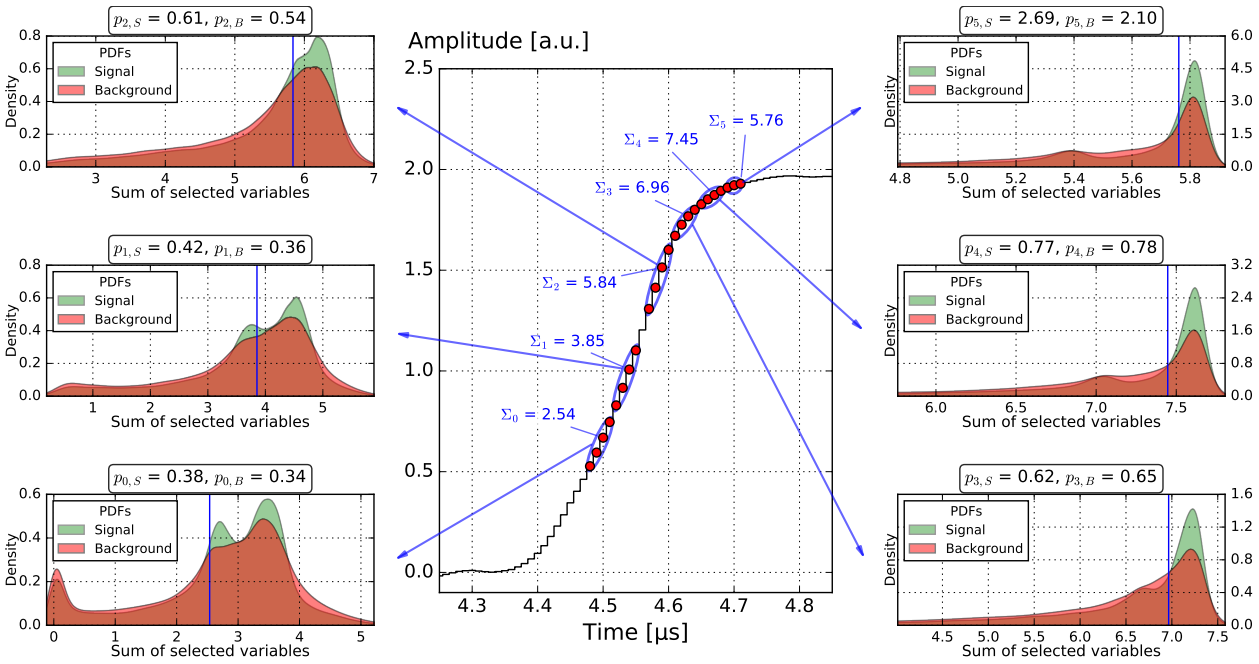


Fig. 2.12: Example of the classification process using the Projective Likelihood classifier.
Central panel – rising edge part of a waveform under classification. Red points show the selected input variables – amplitudes of the normalized pulse. Every 4 (in the exception of the last group, which consists of 3) points are summed together to reduce the number of variables, as seen by the Projective Likelihood algorithm. Sums of variables are annotated next to the ellipses.
Insets – PDF distributions of input variables, smoothed with the gaussian KDE method (implemented in TMVA software). Signal PDFs (green) correspond to the CE events, the background ones their MCS counterparts. Vertical blue lines show the position of the input variables of waveform under classification. The calculated PDFs values ($p_{i,S}$ and $p_{i,B}$), corresponding to sums of the input variables extracted from the waveform, are shown in the boxes over the insets. The values are multiplies to obtain the Projective Likelihood classifier $\mathscr{L}$.

Training process. The PDFs are calculated using finely binned histograms of the training data. They are then smoothed using the Kernel Density Estimation (KDE) method (Fig. 2.11) and are later normalized, so that the total integral of each PDF is 1. The PDFs are calculated for each input variable, separately for the signal and background training data. Apart from the gaussian KDE method of the PDF approximation, an approach based on fitting splines is also available in TMVA, but it was observed that the gaussian KDE method performs best. Obtained PDFs are written to the XML file (so called "weights") for a later use in a classification phase. A typical disk size of the Projective Likelihood weights file is about 1 MB and is significantly larger than for the MLP method.

Classification. After loading the previously generated PDFs, a given event can be classified using Eq. 2.16. The likelihood classifier value is then computed by iterating over all variables in the input vector and by calculating partial probabilities ($p_{i\,S,B}(x_i)$) using the individual PDFs. The product of partial values ($\mathscr{L}_S$) for the signal PSDs is then normalized using its background counterpart ($\mathscr{L}_B$) to obtain a classifier value between 0 and 1. Event with a value close to 0 is classified as background-like (MSE), while a value close to 1 means it is signal-like (SSE).

The classification procedure is illustrated with an example in Fig. 2.12. The PDFs of the selected groups of input variables are shown in the insets. They were previously prepared using the KDE smoothing procedure described in the previous paragraph, for the signal (SSEs) and the background (MSEs) samples taken from the CE and MCS regions, respectively. The waveform to be classified is shown in the central panel. First step is to calculate the $p_{i,S}$ and $p_{i,B}$ values for

every group of the input variables. In the example shown in the figure, the sum of the amplitudes in the first group is equal to 2.54, which corresponds to PDF values of $p_{0,S} = 0.38$ and $p_{0,B} = 0.34$. These values mean that there is a slight indication that the waveform is of signal type. PDF values $p_{i,S}$ and $p_{i,B}$ are calculated for all groups and then multiplied to get $\mathscr{L}_S$ (0.13) and $\mathscr{L}_B$ (0.07), respectively, which are then used to calculate the final Projective Likelihood classifier value $y_{\mathscr{L}}$ (0.65 – see Eq. 2.16).

2.2.3.2 Multi-Layer Perceptron Neural-Networks

The MLP neural-network consists usually of 2 (input and output) + N_l hidden layers of "neurons". A neuron is defined as $R^n \rightarrow R$ transformation (where n is the number of variables present in the previous layer – this depends on the network structure and is usually different from the number of input variables) with a given set of weights. A general MLP architecture is illustrated in Fig. 2.13. The weights are parameters used with a neuron response function ρ, which maps the input value of the neuron to the output value, which is then passed over to a next layer and so on. Training is a process of weights selection, for which the output classifier provides the best separation between signal and background training samples. In theory, one hidden layer can approximate any function, assuming that the number of neurons is large enough. However, more layers with a smaller total number of neurons can in practice be more robust and computed more efficiently, while still providing good separation between the classified events [90].

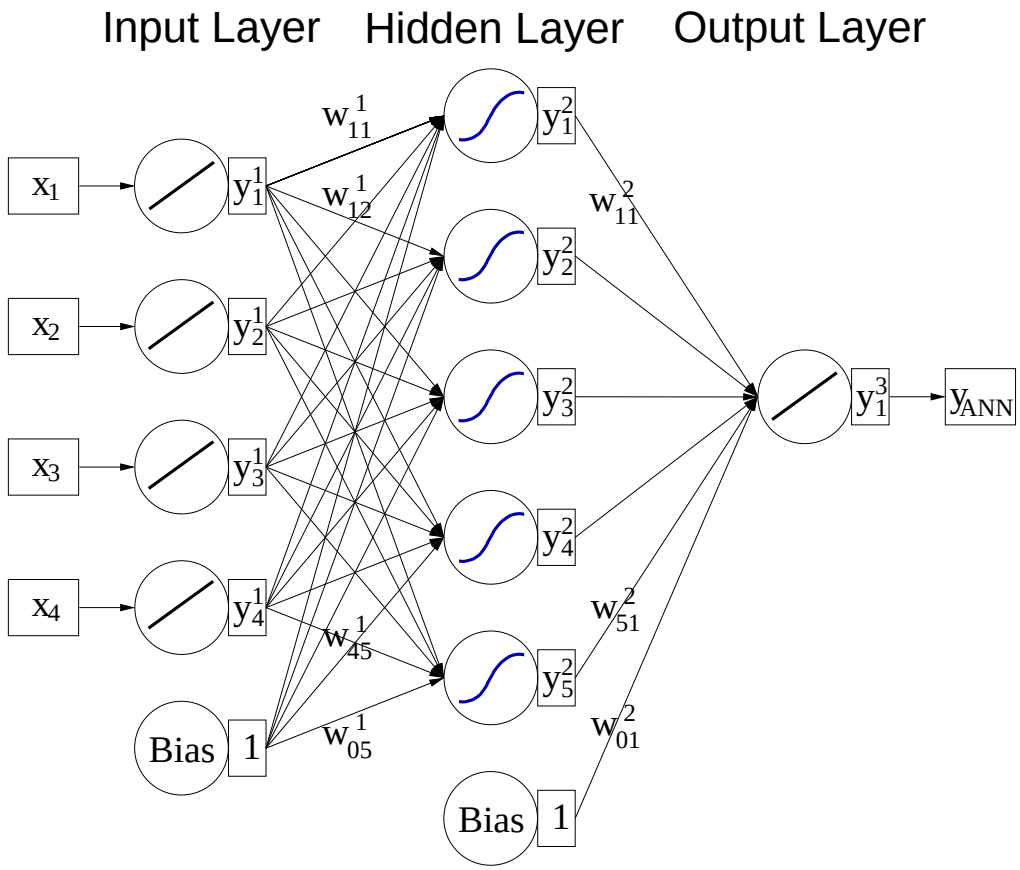


Fig. 2.13: Illustration of the MLP neural-network architecture [92]. The network consists of three parts: the input layer, one or more hidden layers, responsible for the classification and the output layer, which normalized the final classifier value. x_i are the input variables extracted from the waveform. w_{ij} are the weights – coefficients determined in the training process. Each one is associated with every connection between the given neuron and all the neuron in a next layer. Curvy lines in circles represent non-linear neuron response functions (e.g. hyperbolic tangent or sigmoid).

Training and classification process. During the training phase of the method a back-propagation algorithm is used. The main concept behind the algorithm is to calculate the so-called *error function*, that is the measure of a difference between the expected classifier value (either 1 or 0, depending if the event is tagged as the signal or the background, respectively) and the computed output of the neutral network with a current set of weights. Using a gradient descent method, such set of weights is found for which the *error function* achieves its minimal value. The set is later saved in a small XML file for the later use. Aside from neural-network weights, the file also contains all information about the training process (statistics, date, length etc.), as well as parameters like the type of the response function in the neurons.

Chapter 3

Development of the multivariate Pulse Shape Discrimination methods for germanium detectors in vacuum cryostats

Before the application to the data from the GERDA experiment in frames of this thesis, the multivariate PSD methods were tested on the data from the detectors in the vacuum cryostats. The idea was to apply the methods to the high quality data (in terms of noise and time stability) to recognize the efficiency limits due to the physical properties of the detectors. From this point of view the vacuum cryostat provides the best possible conditions. First of all, the HPGe detector and the JFET are shielded by the cryostat wall (usually 1 – 2 mm of Al) from the possible electromagnetic interference. Since the JFET is very close to the detector the input capacitance C_{in} is minimized (especially important for a low capacitance detectors like BEGes). This fact also helps with the noise reduction, because the series and the flicker noise components scale with C_{in} [82].

In the chapter several test measurements are described – they were performed for both semi-coaxial and BEGe detectors. The first measurement was carried out with a pair of two small (10% relative efficiency[1]) semi-coaxial detectors. They are housed in portable cryostats and can be freely rearranged to obtain different measurement geometries. The tests were carried out in the coincidence mode to study the effect of the PSD on events depositing energy in both detectors. In case of e.g. Compton edge events the backscattered γ-rays can be selected using this technique. Obtained efficiencies also served as the numbers for comparison with other detector geometries/types. One of the detector had the excessive leakage current rate (worse energy resolution), which allowed to also compare the PSD efficiency for more noisy signals.

More data for the testing of the PSD methods was gathered using an n-type semi-coaxial detector based spectrometer (60% rel. efficiency). The detector is installed in a radiopure lead shield and is routinely used for the low-background materials screening. For this set of measurements the PSD method was also used to reject the Compton continuum background and increase its sensitivity, using the inversed cut (that is accepting MSEs, rejecting SSEs). Apart from the analysis of the standard ^{228}Th dataset, additional cross-check with a ^{56}Co radioactive source was carried out. The ^{56}Co isotope emits high energy γ-rays, which give rise to the DEPs in different energy regions than ^{228}Th. This provided the opportunity to test the PSD efficiencies for SSEs in a wider range of energies and not only at 1592.5 keV (DEP from 2614.5 keV). The worked out method was also applied to the data from the BEGe-based spectrometer, operated in the underground laboratory HADES [93], achieving a high improvement of its sensitivity.

[1] The efficiency is quoted with regards to a 3"×3" NaI detector-based γ-ray spectrometer [82].

3.1 Coincidence measurements with small semi-coaxial detectors

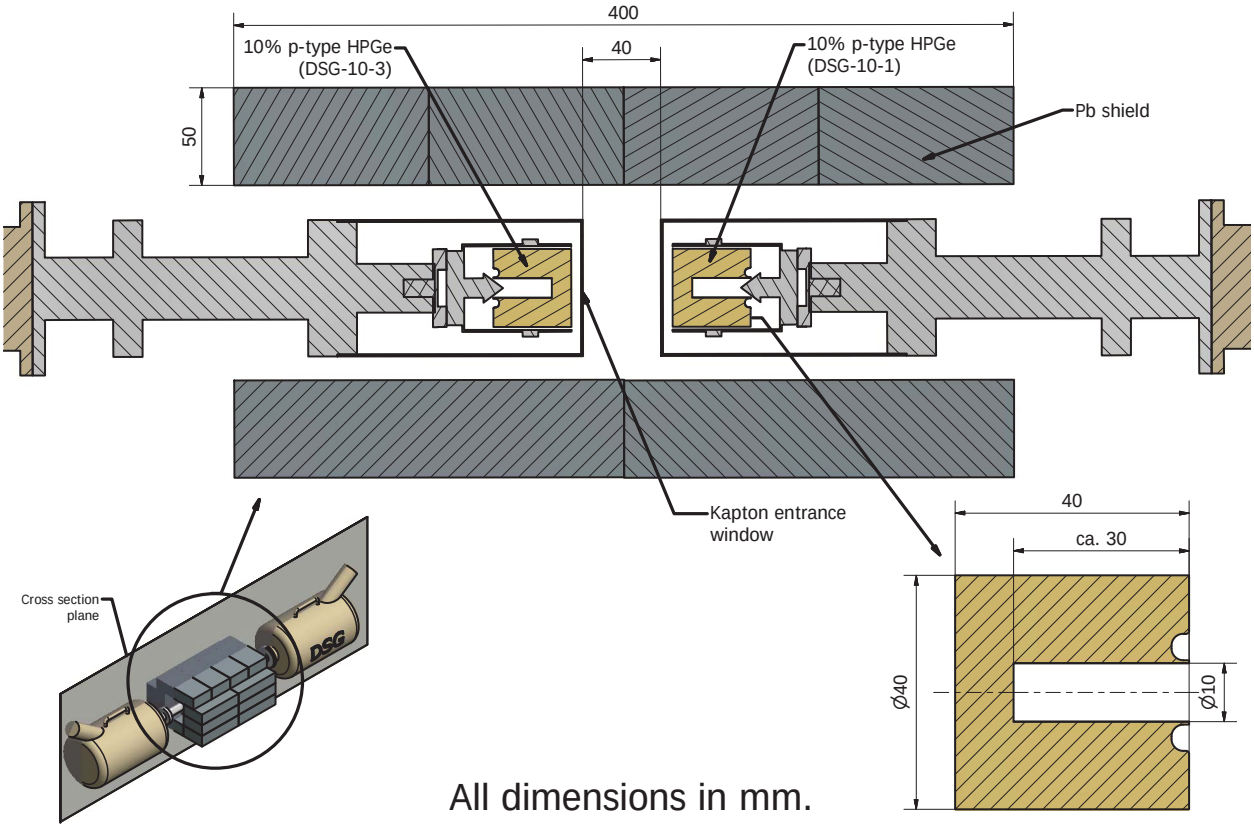


Fig. 3.1: Schematic drawing of the coincidence measurement setup. Two 10% relative efficiency p-type semi-coaxial detectors were placed inside the lead shield. The central part of the figure shown the cross-sectional view of the setup. The distance between the detectors is 40 mm and allows for the placement of large calibration sources. Dimensions of the crystal are shown in the bottom right corner.

3.1.1 Description of the setup

The setup shown in Fig. 3.1 consisted of two, very similar, HPGe p-type detectors. Both are characterized by 10% relative efficiency and were manufactured by DSG Detector Systems GmbH, Mainz, Germany. They are placed in portable vacuum cryostats, for which a LN_2 refilling period is around 4 days. Detailed information about the detectors is collected in Tab. 3.1.

A lead shield was constructed to reduce the background radiation while acquiring the data with relatively weak calibration sources. Since the cryostats are not designed for low-background operation[2], the LN_2 dewar side was not shielded. Also, the copper lining, usually used as an inner shield against Pb X-rays in low-background spectrometers, was not applied here, since the waveforms used for training of PSD methods come from events in the 1500 – 2700 keV range.

Preamplifier waveforms were digitized using a SIS3302 100MHz/16bit flash analog to digital converter (FADC) card. A detailed scheme of the analysis chain is shown in Fig. 3.2. Data from the FADC was saved on a PC via an optical link, which provides 1 Gbit/s readout speed[3]. One of the

[2] Usually, low-background cryostats have U-shaped vacuum chambers, so there is no "line of sight" between the detector and possible external radiation sources after mounting in a lead shield.

[3] http://www.struck.de/sis1100-ecmc.htm

Table 3.1: Parameters of the detectors used in the coincidence measurement.

Detector Name	DSG10-1	DSG10-3
Type	p	p
Relative efficiency	10%	10%
Nominal voltage [V]	2000	2000
Operating voltage [V]	2000	2000
Energy resolution at 1332 keV [keV]	2.65	1.75
Entrance window material	Al	Kapton

detectors, namely DSG10-3, had a much better energy resolution than the other (1.75 vs 2.65 keV), therefore it was used as a primary source of data. To utilize the whole dynamic range of the FADC, signals from DSG10-3 were amplified using a fast linear amplifier. The second detector, due to the worse energy resolution, served mostly for registering coincidence events. However, the same analysis steps were applied to them to compare the effects of the dynamic range adjustment and the energy resolution on the PSD performance.

3.1.2 Data processing

Prior to the PSD analysis, processing of "raw" FADC data is needed. It consists of:

- energy reconstruction – filtering the noise from the waveform to obtain the pulse height and create an energy spectrum,
- timing information – extracted by applying a trigger procedure (e.g. the waveform's amplitude crossing the threshold value),
- rejection of invalid events – recognizing pile-ups, false coincidences, overflow events etc.

All of the above parameters can be extracted from the preamplifier signal using a GELATIO framework [94], developed for the digital signal processing of the data from germanium detectors. The software performs operations like signal smoothing, baseline extraction, determining the timestamp of a trigger, as well as digital pulse height filtering. The two most common methods of energy reconstruction are implemented in GELATIO: the gaussian and trapezoidal digital filters [95]. They transform the input waveform to either a gaussian or trapezoid pulse, which is characterized by a much lower noise level. This step is analogous to the electronic filtering performed by the shaping amplifier in a traditional γ spectrometry setup.

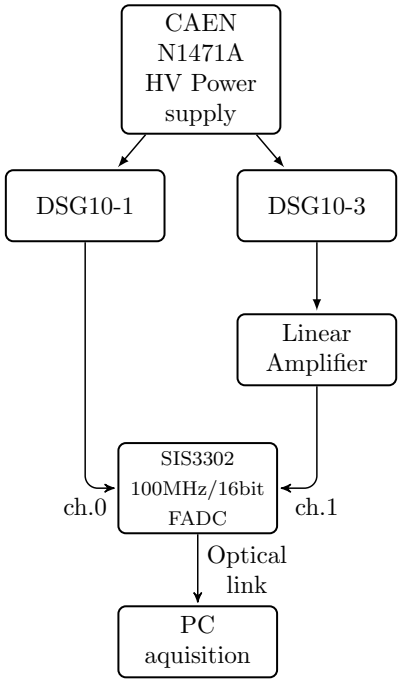
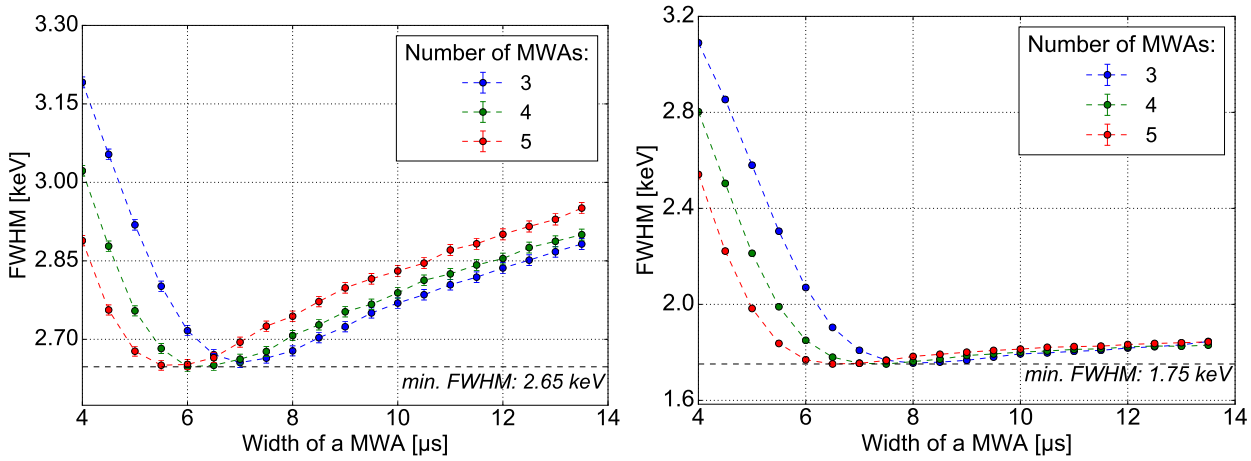


Fig. 3.2: Block diagram of the acquisition scheme used in the coincidence measurements. A linear amplifier was used to match the FADC's input range for the waveforms from channel 1. Channel 0 served for registering coincidences (and had the worse energy resolution), therefore it was connected directly to the FADC.

A separate run with a ^{60}Co source was performed to obtain data needed for the optimization of the energy filter. In principle, one could use any γ radiation source, however the value of the energy resolution of 1332.5 keV peak from ^{60}Co is traditionally used to compare the resolutions of



(a) Channel 0 (DSG-10-1)

(b) Channel 1 (DSG-10-3)

Fig. 3.3: Visualization of the energy resolution optimization procedure. Semi-gaussian filter was applied to the data from both spectroscopy channels (numbered 0 and 1 – panels a and b, respectively) by the multiple application of the MWA filter of width shown on the horizontal axis. Number of the applied MWA filters has been varied as well to find optimal parameters (see the legend). For channel 1 (right panel) minimal value is obtained for 5 MWAs with a width of 6.5 µs. Optimal energy resolution for channel 0 is actually obtained for a slightly shorter filter width, however, due to the software limitation optimal parameters for channel 1 are used for both cases. For the explanation of the different shapes of minimization curves see the text.

gamma spectrometers [82]. In the described measurements, data was processed using a gaussian filter. First, the filter applies a moving average deconvolution of the preamplifier waveform, which corrects for the falling exponential tail of the pulse. The decay constant, determined from an exponential fit, was equal to 47 µs. After the deconvolution, the Moving Window Average (MWA) filter with a given width is applied n times to the pulse. Therefore, a filter is characterized by those two parameters, which can be optimized to give the best energy resolution. Results of the optimization are plotted in Fig. 3.3 (left panel: DSG10-1 – channel 0, right panel: DSG10-3 – channel 1) – both plots show the FWHM of the ^{60}Co 1332.5 keV peak as a function of filter parameters.

One can easily notice different shapes of both curves, especially for the larger MWA width values. The shape can be explained by the analogy to the traditional spectroscopic shaping amplifiers (e.g. Ortec 672), where a similar curvature can be observed for the FWHM plotted versus the amplifier shaping time constant τ_s. The analogy is well justified, since a frequency response of both CR-RC^n and multiple MWA filters converge to an equivalent gaussian filter kernel with increasing n [96]. From an analysis of the preamplifier noise and shaping amplifier equivalent circuit 3 main noise components can be distinguished in such system [82]:

- *series noise* – its main source is an input JFET and its magnitude increases with both the input capacitance C_{in} and the JFET temperature T. Its contribution decreases for longer shaping time constants τ_s,
- *parallel noise* – increases with the detector's leakage current I_{leak} and τ_s, decreases with larger feedback resistance values R_f,
- *flicker noise* – also called $1/f$ noise, independent of the shaping time constant, but similarly to the series noise, it increases with C_{in} [97].

The total noise can be therefore characterized as the quadratic sum of three terms: one increasing with τ_s, one with the decreasing relationship and the last one independent of it. Since both detectors are equipped with similar front-end electronics (JFET, feedback resistors and capacitors) and have similar capacitance, it can be then concluded that the increasing FWHM with increasing MWA

width for channel 0 is due to the larger leakage current I_{leak}. The reason for the increased I_{leak} is the observed vacuum deterioration in the cryostat, which in turn results in a worse insulation and the temperature increase (I_{leak} is directly related to the thermally generated carriers in the diode).

The compromised energy resolution could be in principle also linked with the faulty detector, however, earlier measurements indicated same energy resolution as for channel 1 ($\approx 1.8\,\text{keV}$). Vacuum insulation of channel 1 cryostat was regenerated[4] just before taking the measurements described in this work. Indeed, the optimization curve shows that there is almost no increase in the FWHM for larger MWA widths – a sign of a good vacuum insulation and the low temperature of the diode.

GELATIO does not allow for the channel-dependent filter parameters, therefore, only one set of them can be chosen. Curves in Fig. 3.3b indicate that the minimal value of FWHM (1.75 keV) for channel 1 detector is obtained for 5 MWAs (red curve) with the width of 6.5 µs. Therefore, these values were chosen as final parameters. On the other hand, for channel 0 detector the FWHM reaches its minimal value (Fig. 3.3a) for a slightly shorter MWA width (around 6 µs). However, signals from channel 1 detector were the primary data used in the analysis and its parameters had higher priority. Anyhow, there is a nebligable difference in the FWHM between MWA widths of 6.0 and 6.5 µs.

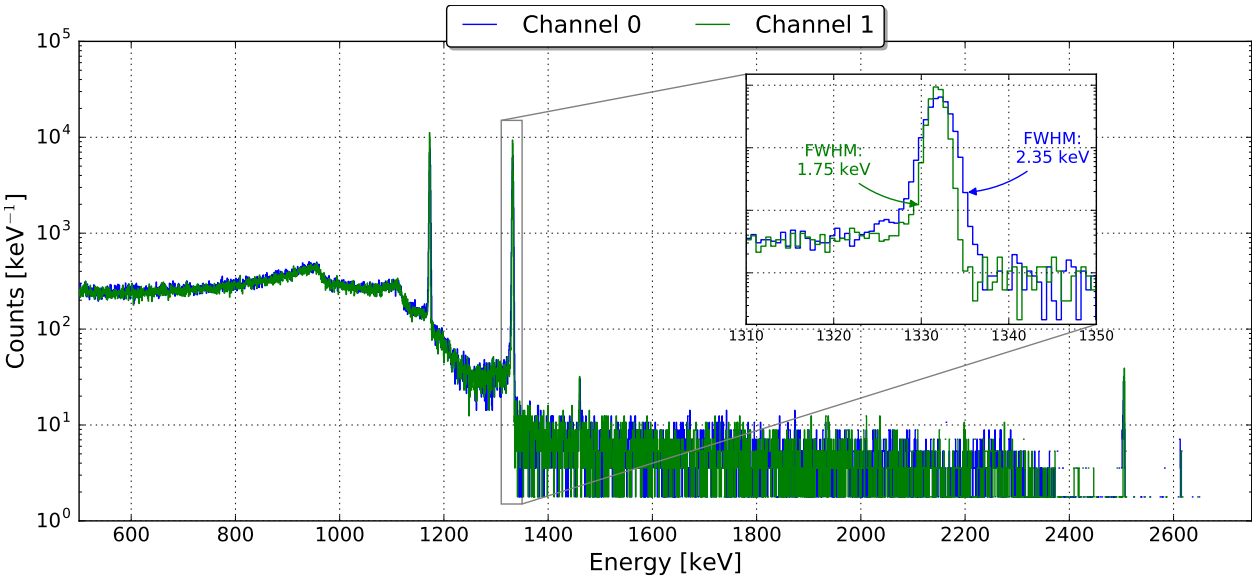


Fig. 3.4: ^{60}Co energy spectra obtained from the analyzed detectors (channel 0 – blue line, channel 1 – green line). The inset shows the energy resolution of the 1332 keV peak for both detectors.

After the energy filter optimization, quality cuts were applied to the data to reject events of dubious quality. Examples of such events are:

- *pile-ups*, detected by the presence of multiple triggers in a single waveform or the slope of the baseline (from the previous event),
- abnormal baseline sigma value, indicating temporal rise of the noise and/or gain changes,
- coincidence events outside the expected time window (few µs).

Rejected events were not included in the further analysis. The obtained energy spectrum, obtained by the application of the gaussian filter, was then calibrated from ADC counts unit to keVs using a linear function.

[4] By heating the absorbent and pumping the cryostat using an ultra high vacuum pump.

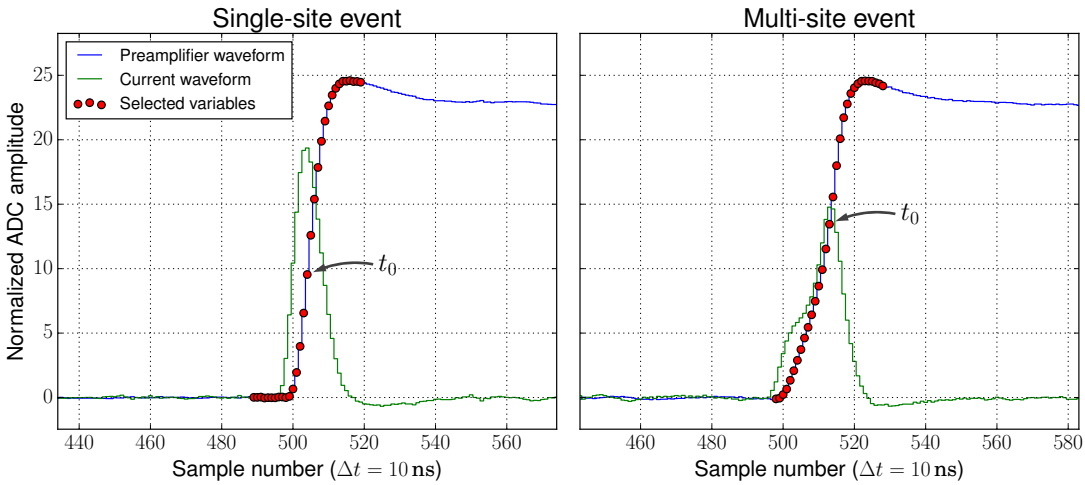


Fig. 3.5: Example of pulse shapes from typical single-site (left panel) and multi-site (right) events from the semi-coaxial detectors. Current pulses (green) were obtained by differentiating the preamplifier output (blue) and applying a 50 ns MWA filter. Red points represent the amplitudes selected to perform the PSD. t_0 indicates the voltage amplitude corresponding to the maximal current of the pulse and is used for the alignment of the selected samples (15 before and 15 after t_0).

3.1.3 PSD input variables extraction

For every multivariate classification problem the crucial aspect is the selection of the input variables. In the case of ^{76}Ge-based $0\nu\beta\beta$ decay experiments, the usual parameters are those extracted from the waveform's shape. Therefore, a procedure must be worked out to chose the most important waveform parameters. Feeding the whole waveform ($\approx$ 1000 samples) is suboptimal – the training statistics requirement is getting bigger with e.g. the number of neurons in the network [90]. It would be also challenging from the computational point of view and be very prone to overtraining. The overtraining phenomenon happens when the method works *only* on the events used for training, instead of recognizing the general features of the waveform.

A waveform extraction algorithm, illustrated in Fig. 3.5, initially creates 31 input variables. Firstly, the waveform amplitude is normalized, that is all samples are divided by the energy value, obtained from the energy filter. After the first step, the pulse is differentiated and smoothed with a single 40 ns MAW filter, and since the preamplifer acts like an integrating circuit, differentiation creates back the current signal from the detector (for details refer to Sec. 2.1 for the current signal shapes obtained from the model geometries). The exact amplitude of the current signal is not crucial, since the current pulse maximum is used only for time alignment of the input variables. The maximal value of the current pulse is chosen as a central point (t_0), around which all the input variables are selected from the *preamplier's* waveform (blue line in Fig. 3.5). 31 variables are selected in total: the normalized amplitude at t_0, 15 samples before and 15 after. The obtained data is then used in the next analysis step, which is a dimensionality reduction.

3.1.4 Dimensionality reduction with Principal Component Analysis

Since the input variables were selected in a rather arbitrary manner, probably not all of them are crucial to achieve the high efficiency of the PSD method. For example, the first input variable (Fig. 3.5) will have an amplitude of around 0 for a majority of the pulses[5], regardless if it is a SSE or a MSE. Similar behavior can be observed for the last variable, since it is mostly unaffected by

[5] In fact, more than 70% of all traces corresponding to events with high energy deposition have their first input variable contained within $-3\sigma \ldots 3\sigma$ of a Gaussian distribution with a mean value of 0.

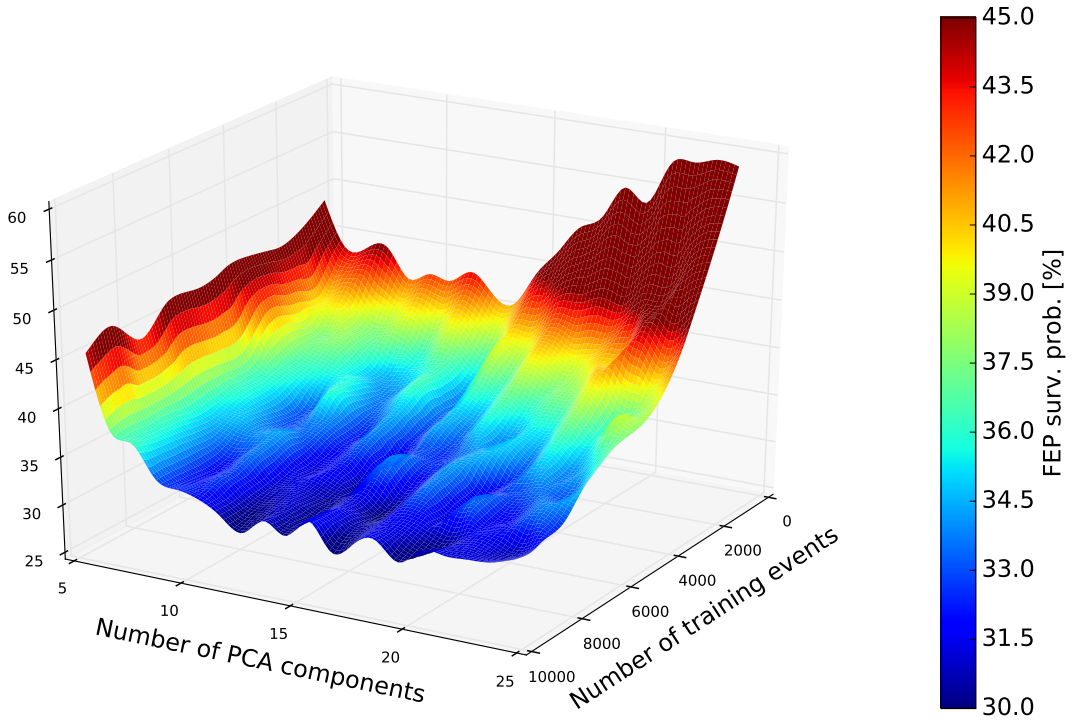


Fig. 3.6: Visualization of the PSD classifier efficiency as a function of the number of training events and the PCA components (data is interpolated for the visual purposes). The z-axis shows the survival probability of the 1620 keV FEP for the fixed 90% DEP acceptance. For the additional explanation see the description of Fig. 3.7a and the text.

the rising part of the pulse (after the pulse height normalization). Additionally, adjacent variables (e.g. the first and the second) can have a similar value, since the signal amplitude changes are slow when compared to the sampling frequency. It would be therefore desirable to apply a procedure which would select only these input variables, which bring the most important information from the PSD point of view.

This can be achieved by applying Principal Component Analysis (PCA) to the data. PCA is an unsupervised algorithm allowing for the capturing of the essential patterns in the data [98]. Its application consists of several steps:

- determination of a mean value of each input variable,
- computation of the covariance matrix, its eigenvalues and eigenvectors, for all input variables after the subtraction of the mean value,
- creation of the transformation matrix using the eigenvectors. The eigenvectors are sorted by their eigenvalues, that is the eigenvector with the largest eigenvalue corresponds to the first PCA component, and so on.

After the transformation matrix is obtained, it can be used to make $R^n \rightarrow R^n$ (in this case $n = 31$ is the number of variables) linear transformation to get a new set of the PCA variables (called "PCA components"). The new variables are characterized with several advantageous properties, from the machine-learning point of view: the variables are sorted with a decreasing variance and the linear correlations between them are removed. A main premise for using PCA for the dimensionality reduction is that with variables with large variance also have the large discrimination power. Therefore, one can discard m last components with low variance and end up with a smaller $(n-m)$ number of the input variables.

On the other hand, the usefulness of the removed linear correlations strongly depends on the type of classifier used in the analysis. To give an example, the Projective Likelihood classifier suffers loss of performance for the correlated input variables, while MLP neural-network deals well even with the non-linear correlations [90]. On the other hand, the benefit from a dimensionality

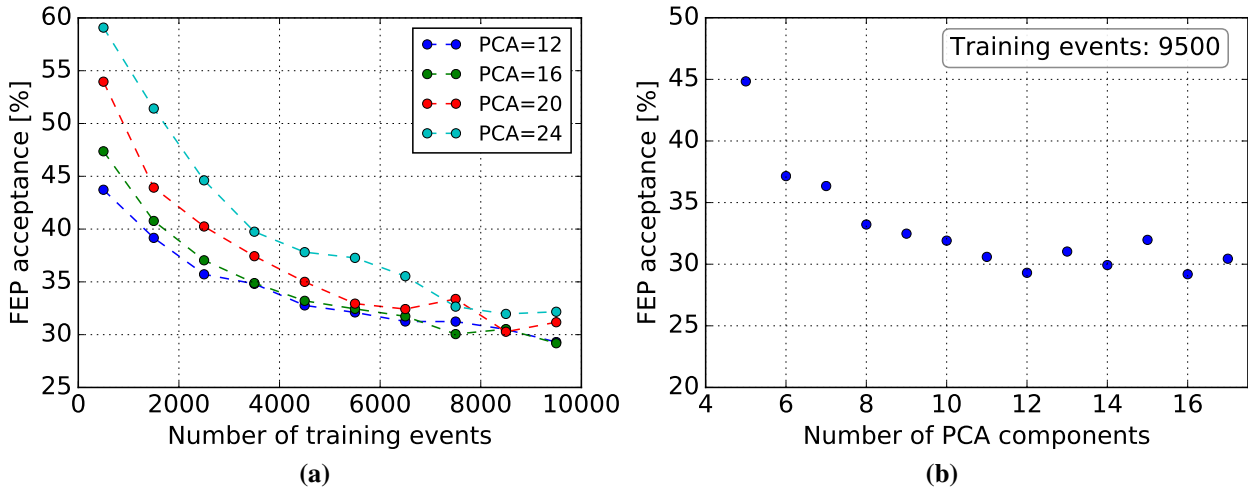


Fig. 3.7: *Left panel (a)* – acceptance of the events from 1620 keV FEP as a function of the number of training events and PCA components (same data as in Fig. 3.6). For the large training sample (bigger than e.g. 8000 events), good efficiency (low MSEs acceptance) is achieved regardless of a number of PCA components. However, for small training samples (2500 events), there is a clear efficiency improvement after the PCA dimensionality reduction.
Right panel (b) – acceptance of the 1620 keV FEP events as a function of the number of PCA components for the fixed number of training events (9500). There is no significant improvement of the efficiency after increasing the number of PCA components over 12.

reduction is that it leads to a simplification of the multivariate classifier structure (e.g. smaller neuron number in case of artificial neural-networks). Also, the smaller statistics can be used for the training, since it takes a smaller number of events to "populate" volume of phase space with a smaller number of dimensions [99]. The effect on the dimensionality reduction on the required training statistics is shown in Fig. 3.6, 3.7a and 3.7b, where the MLP neural-network was trained using the DEP and FEP events, with varied number of both training events and a number of PCA components.

Naturally, an instant question appears: how many components can be rejected without actually diminishing the separation efficiency of the classifier? To determine the minimal number of the variables, the PSD training and classification process was performed for a different number of PCA components. Results of the optimization are plotted in Fig. 3.7b. They show that using the first 12 PCA components is sufficient to achieve a maximal reduction of the 1620 keV FEP. This number of components is therefore used in the further analysis.

3.1.5 PSD training with ^{228}Th data

To provide the training data for the multivariate methods, one needs samples of both SSEs and MSEs. This can be done by irradiating the detector with a high energy γ source, which produces regions rich in events of both types (see Sec. 2.2 for a detailed explanation). In this case, the ^{228}Th source was chosen, because of a 2614.5 keV γ-ray, emitted by its daughter ^{208}Tl.

A training algorithm of every classifier in the TMVA package requires samples of events for two classes. Therefore, events from the 1592 keV DEP $\pm$ 1/2 FWHM represented the sample of SSEs (class "signal"), while 1620 keV FEP $\pm$ 1/2 FWHM events were selected as the sample of MSEs (class "background"). The classification process tags the event resembling the "signal" class with a value close to 1 and those resembling the "background" with a value close to 0. The FWHMs of DEP and FEP were 1.96 keV and 2.05 keV, respectively. An equal number of 9500 events was used in both cases (SSEs/MSEs), even though more events from DEP were available. All the results,

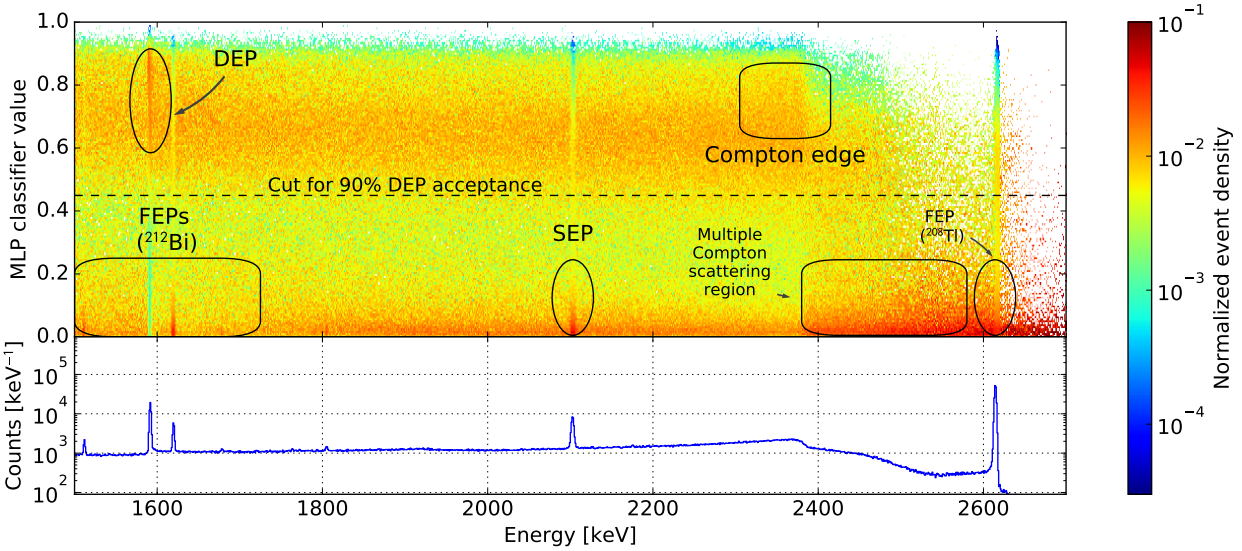


Fig. 3.8: *Top panel* – 2D normalized histogram of classified events from the ^{228}Th measurement for channel 1 detector. Two very well separated bands are visible – they correspond to the MSEs and SSEs for classifier values around 0 and 0.7, respectively.
Bottom panel – high energy part of the registered ^{228}Th energy spectrum.

presented in this section, refer to the above mentioned training conditions, that is the selection of signal and background regions, as well as the training sample size, unless indicated otherwise.

Metadata obtained from the previous step was used to classify all events from the ^{228}Th dataset. The results are plotted in Fig. 3.8 in the form of a 2D histogram (top panel). It is normalized with respect to a vertical axis, such that an integral of the event density for each vertical bin column is equal to 1. The normalization is done for the visualization purposes only (it removes the effect of different peak intensities).

Two main bands are visible on the histogram in Fig. 3.8 - one near MLP ≈ 0 and another for MLP ≈ 0.7. Since MSEs from FEP were used as class "background" training sample and SSEs from DEP were used as class "signal" sample, waveforms resembling MSEs are classifier with a classifier value close to 0 and SSE-like events have classifier value close to 1. Therefore, all events with multiple energy deposition in the detector (like FEPs, SEPs, MCS) should lie in the lower band, while single energy deposition events (DEPs, Compton edge, single Compton scatterings) should like in the upper band.

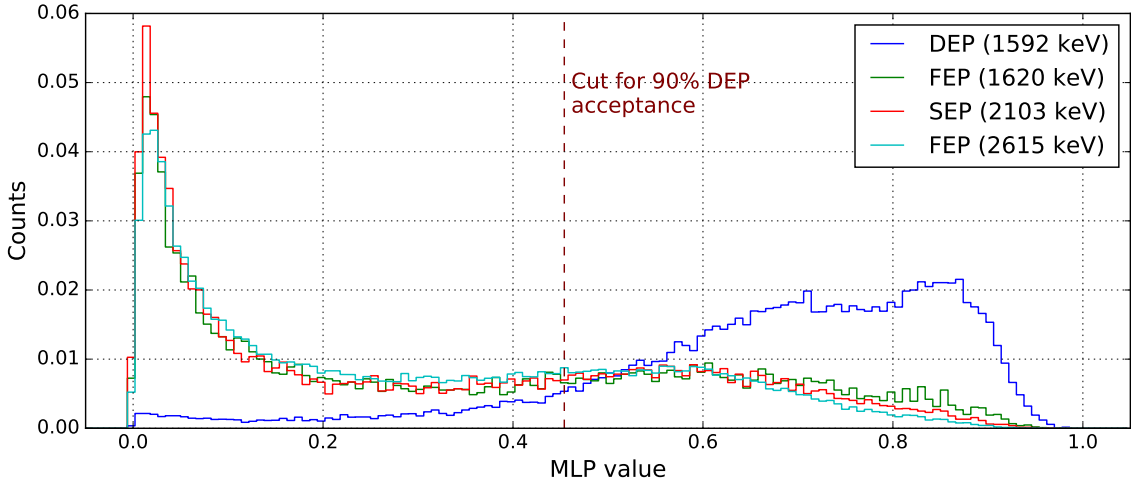


Fig. 3.9: Normalized classifier distribution of events from various peaks from ^{228}Th spectrum, as registered by the small semi-coaxial detector (channel 1). For each peak the events are taken from $E \pm 1/2$FWHM, where E equals to the peak centroid.

Table 3.2: Acceptances of the events from high energy peaks from the ^{228}Th spectrum after applying the PSD cut (90% DEP acceptance).

Energy [keV]	Isotope	Acceptance [%]
1460.8	^{40}K	27.1 ± 2.4
1512.7	^{212}Bi	26.8 ± 2.7
1592.5	^{208}Tl (DEP)	90.1 ± 0.2
1620.5	^{212}Bi	28.9 ± 0.8
2103.5	^{208}Tl (SEP)	26.2 ± 0.5
2614.5	^{208}Tl	28.9 ± 0.1

Table 3.3: Acceptances of the events from the selected Compton continuum regions of ^{228}Th spectrum for a fixed 90% DEP acceptance cut.

Energy range [keV]	Acceptance [%]
1520 – 1580	55.3 ± 0.2
1635 – 1950	54.8 ± 0.1
1950 – 2085	53.6 ± 0.1
2110 – 2235	54.5 ± 0.1
2200 – 2350	56.1 ± 0.1
2350 – 2380	57.5 ± 0.2
2400 – 2500	35.0 ± 0.2
2500 – 2600	18.1 ± 0.2

This is indeed the case, as both the DEP and the Compton edge from 2614.5 keV FEP from ^{208}Tl are clearly visible in the upper band of the 2D histogram. ^{212}Bi, ^{208}Tl FEPs and SEP are localized in the lower band, together with the multiple Compton scattered events with energies within the 2400 – 2600 keV range. Since only few events lie in the central part of MLP classifier distribution (0.2–0.5 range), both events types are very well separated. This can be also seen in Fig. 3.9, where the classifier distribution for different peaks is shown (events with energies $E \pm 1/2$ FWHM, where E equals to the peak centroid). There is a very good agreement between the distributions for various MSEs peaks (namely the FEPs and the SEP), even though they are localized in different parts of the energy spectrum.

As it can be seen in Tab. 3.2, survival probabilities for events from the FEPs and the SEP are under 30%, while preserving 90% acceptance of DEP events. There is a small difference in acceptance ($\approx 4\%$) between FEPs and the SEP events and it is probably caused by the different topology of the events in the SEP and the FEPs (see Fig. 2.8), which results in a slightly different pulse shape. FEP events are mostly created by multiple Compton scattering of the incoming γ-ray, while events in SEP originate from pair production + absorption of one of the 511 keV gammas. Acceptances for the events from Compton continuum regions of the ^{228}Th spectrum are gathered in Tab. 3.3. The mean value of the acceptance is $\approx 55\%$ for the events with energies < 2380 keV. The highest value is for the CE events (57.5% for the 2350 – 2380 keV range). Higher energy regions contain almost exclusive MSEs, therefore the survival probability drops significantly to 35.0% in MCS region (2400 – 2500 keV) and finally to 18.1% (2500 – 2600 keV).

There is some discrepancy between the 2400 – 2500 keV, 2500 – 2600 keV and 2614.5 keV FEP survival probabilities. Again, this can be explained by the event topology. The energy region between Compton edge (2381.7 keV) and 2500 keV contains a large fraction of *double* scattered events. The backscattered 2614.5 keV γ-ray has a new, lower energy (232.8 keV) and can scatter one more time in the detector. This time, a maximal energy deposition (again, from the Compton scattering kinematics) to the electron is ≈ 110 keV. The sum of deposited energy in the detector for both scatterings equals to ≈ 2490 keV. On the other hand, 2500 – 2600 keV events cannot be only two times Compton scattered, due to the kinematic constrains stated above. Therefore, their lower survival efficiency probably results from the larger number of scattering sites. A volumetric effect can also have some influence on the acceptance, because the last scattering site has to be near the edge of the detector for the γ-ray to escape. This is in contrast with the full energy deposition (2614.5 keV) for which the last step has to undergo via photoelectric effect and it can be localized deeper in the crystal.

Of course, in the 2500 – 2600 keV range, the energy deposition can be caused also by a Compton scattering and a photoelectric effect sequence. However, this is unlikely since the free mean

path (derived from the photoelectric effect cross-section in germanium) for the $\approx 250\,\text{keV}$ γ-ray is about 8.4 cm, a factor two larger than the crystal dimensions.

3.1.6 Analysis of the coincidence measurements

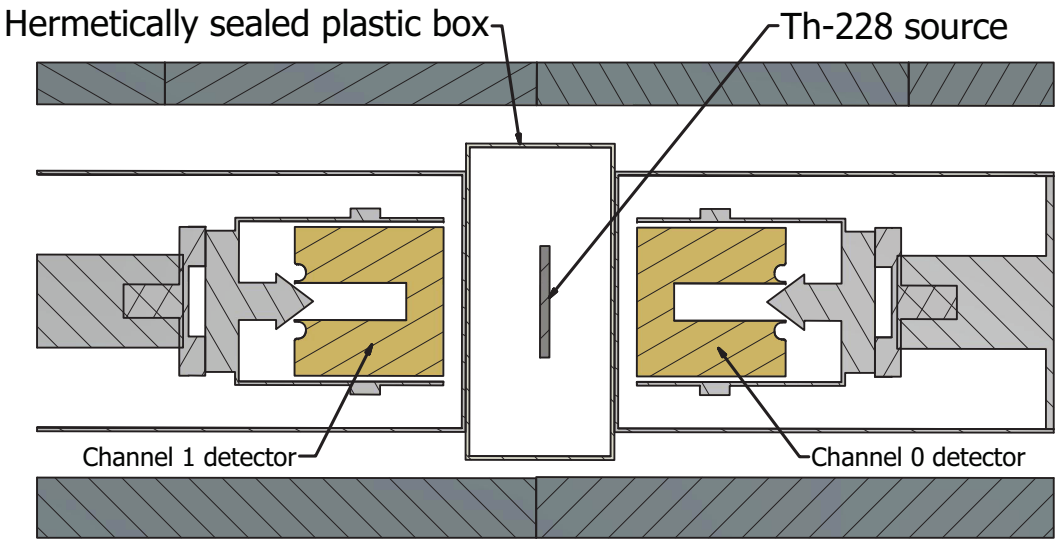


Fig. 3.10: Schematic drawing of the measurement setup with an indicated position of the ^{228}Th source. The plastic box seals the radon, generated in the $^{228}\text{Th} \rightarrow {}^{224}\text{Ra} \rightarrow {}^{220}\text{Rn}$ decay chain.

For some events in the detector, like the ones involving a pair production or Compton scattering, a portion of incoming radiation energy is escaping detector's volume. Such events have usually a well defined topology inside the detector, e.g. SSEs from DEP have two 511 keV γ-rays associated with the positron annihilation. If the escaping γ-rays can be registered in a second detector, then in principle one could use this information to select a cleaner sample of either single-site or multi-site events.

Fig. 3.10 shows the measurement setup with a marked location of a ^{228}Th source. Since the main purpose of the source was to use it as an alpha radiation emitter, the radioisotope was electrodeposited on a surface of a steel plate ($\varnothing = 30\,\text{mm}$). One of the ^{228}Th daughters produced in the decay chain is ^{220}Rn, which can escape the source's surface. Then it can deposit its own daughters on the surfaces in the measurement chamber or on the detector's endcap and increase the background in the later measurements. Therefore, the source was contained in a hermetically sealed plastic box.

Fig. 3.11 (left panel) shows the coincidence events between two detectors. Channel 1 detector was the primary one used in the analysis, therefore, for the most data, it was used as the triggering channel. This means that the pulses from channel 0 detector were saved only when in coincidence with those from channel 1. A small coincidence window of 150 ns was chosen to reduce the number of random coincidences. First two vertical lines are the events from the DEP and the SEP, registered when one of the escaping 511 keV γ-rays was partially absorbed in the second detector. Bottom right panel of Fig. 3.11 shows a resulting energy spectrum in channel 1 detector after applying the cut $E_0 \in [100, 512]$ (E_0 being the energy deposited in channel 0 detector). Both the DEP and the SEP are clearly visible, while e.g. 1620 keV FEP is suppressed. The FEP at 2614.5 keV is still visible and it is due to the random coincidences. The reason of much larger suppression of events from 1620 keV, when compared with those from 2614.5 keV one, is due to the differences in intensities. For every ^{228}Th decay, there is almost a 36% percent chance that 2614.5 keV γ-ray will be emitted vs a $< 2\%$ chance for 1620 keV line.

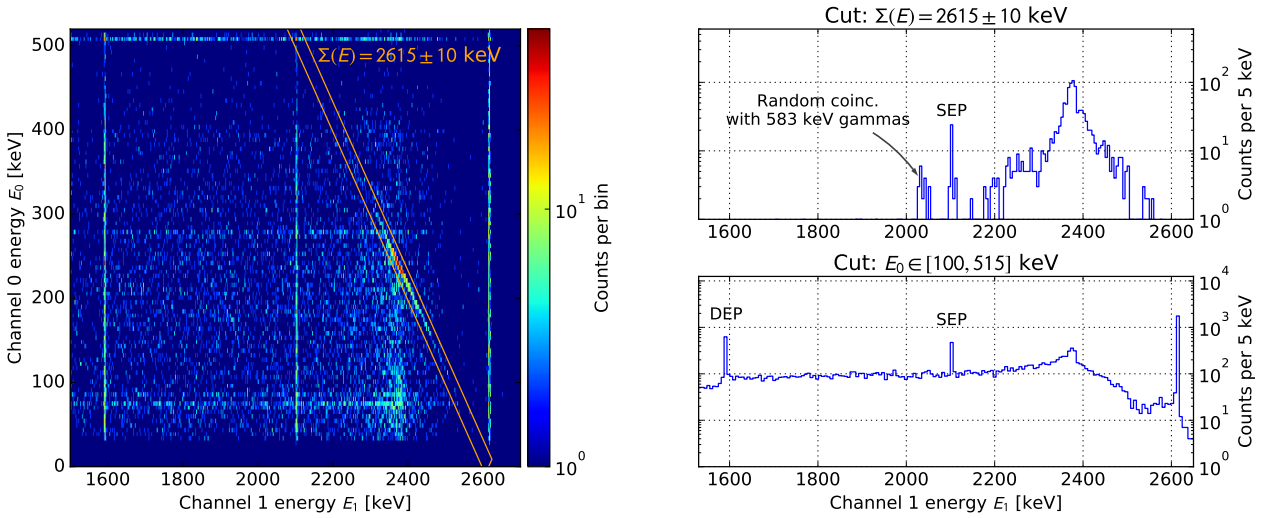


Fig. 3.11: *Left panel* – visualization of the coincidence measurements on a 2D histogram. The horizontal axis indicates the deposited energy in channel 1 detector, while the vertical one regards the energy in channel 0 detector. A cut showing events with a sum of energies from both detectors being equal to (2615 ± 10) keV is indicated with an orange rectangle.
Right top panel – the energy spectrum in the channel 1 detector after an application of the $\sum(E) = (2615 \pm 10)$ keV cut.
Right bottom panel – the energy spectrum in channel 1 detector after application of a cut requiring simultaneous energy deposition also in channel 0 detector.

Another cut that can be applied to the data is to restrict the sum of energy values in both channels to be (2615 ± 10) keV. The cut is illustrated by an orange rectangle in Fig. 3.11. Data with the applied cut contains Compton scattered events, which deposit their energy in both detectors. Since the scattering angle is $\approx 180°$, most events have the energy close to a value predicted by the formulas for the energy transfer in a Compton scattering, i.e. 2381.7 keV for the incoming 2614.5 keV γ-ray[6]. A resulting energy spectrum in channel 1 detector is shown on the top right panel of Fig. 3.11. A few events from the SEP are also visible, however, the statistics is very low, since the cut requires 511 ± 10 keV to be deposited in the second detector. If the range of energy deposition in the second channel is extended, like in bottom right spectrum in Fig. 3.11, the continuum under the peaks also increases. A perfect setup for obtaining the clean sample of DEP events could in principle use triple coincidences in three detectors (1592.5 keV + 2×511 keV), but an extra detector was unavailable when the measurements were made.

Since the selected Compton edge events are almost a pure sample of SSEs, their classifier distribution can be compared with the DEP events (without the coincidence cut), which are always used to fix the cut position – the comparison is visualized in Fig. 3.12. Distributions are quite similar, however, there is an important difference, namely that the peak of the Compton edge events distribution is slightly shifted to the lower classifier values. Also, the relative contribution of the "false MSEs" (SSEs misclassified as MSEs - assuming that there is few random coincidence events after the cut) in the distribution is larger (the flat part in Fig. 3.12 - MLP values smaller than 0.4). Both effects result in the decreased acceptance at the level of 72.6%, compared to 90% of DEP events.

The above mentioned discrepancy was investigated in terms of the number of used PCA components. The possible connection would be that the less degrees of freedom are "allowed" to the neural-network, the smaller the possibility that it would learn to preferentially distinguish DEP events over Compton edge ones (since both types should be treated as SSEs). For the extreme

[6] $E_e = \dfrac{E_\gamma}{1 + \frac{E_\gamma}{511\,\mathrm{keV}}(1 - \cos\theta)}$, where θ is a γ-ray scattering angle

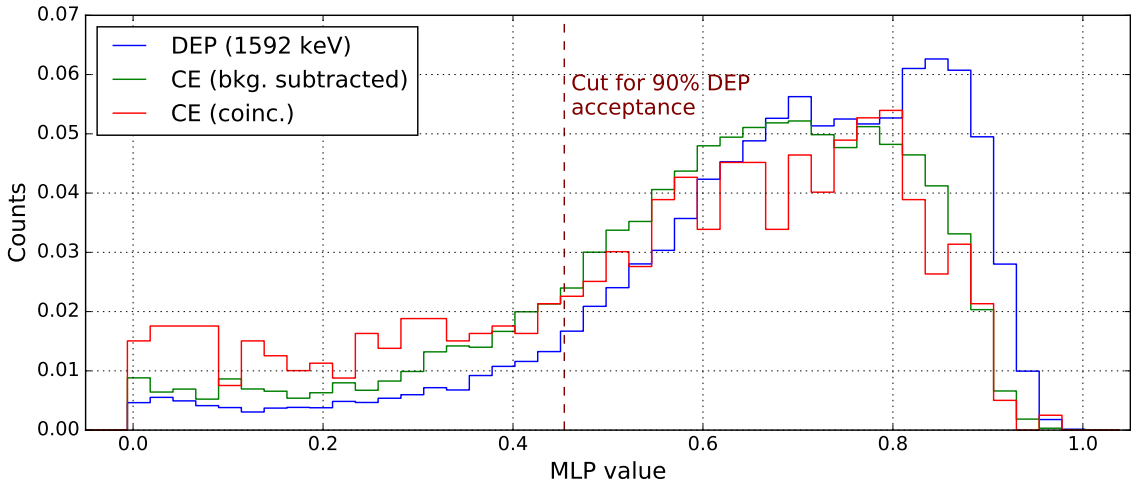


Fig. 3.12: Comparison of the normalized classifier distribution of DEP events (without coincidence cuts) and Compton edge events (energy range 2300 – 2382 keV), filtered with the coincidence cut ($\Sigma(E) = 2614.5 \pm 10$ keV). Large bin widths are due to the a much small number of events available in the coincidence dataset. 72.6% of coincident Compton scattered events are accepted for the cut fixed at 90% acceptance of DEP.

case of only 6 PCA components, the acceptance for Compton edge events with the coincidence cut rises to 81% (at 90% DEP acceptance). However, as shown in Fig. 3.7b, there is a penalty associated with such extreme dimensionality reduction, i.e. the survival probability for MSEs is larger than in a case of using a higher number of PCA components ($\approx 37\%$ vs $\approx 30\%$ at 1620 keV for 6 and 12 PCA components, respectively). Further dimensionality reduction (e.g. to only 5 PCA components), aside from yielding even worse MSE acceptance, produces non-physical effects, e.g. additional energy dependent band of MSEs appears on the classifier vs energy histogram).

An alternative explanation of why decreasing the number of PCA components (and thus the discriminating power between SSEs and MSEs) reduces the discrepancy between DEP and Compton edge acceptances could be that some random coincidence events, which are in fact MSEs, get misclassified as SSEs. To exclude this possibility, a more stringent coincidence setup would be needed, preferably with both collimated beams from the source to the first detector and a collimated entrance to the second one. Unfortunately, the more collimation is used the stronger (higher activity) radioactive source has to be, to obtain considerable count rate in the measurements. Two examples can be given: in [84] 250 kBq ^{228}Th was used with the second detector moved further apart to avoid random coincidences, lead shield was used only to avoid registering radiation from a source in the secondary detector. A setup described in [100] contains true slit collimators and four secondary detectors, but a very strong 780 MBq ^{137}Cs source is needed in the data taking campaigns. Both solutions are not feasible for the relatively weak (few kBqs) ^{228}Th source, therefore repeating the presented measurements with a much stronger source could exclude the hypothesis presented in this paragraph. However, no strong γ-ray sources were available in the department for that purpose.

3.1.7 PSD efficiency on data with suboptimal noise conditions

As was pointed out in Sec. 3.1.2, due to the excessive leakage current signals from channel 0 detector contained higher noise levels. Since both crystals are characterized by the same relative efficiency (10%), operating voltage and the same model symbol, one can assume that crystal geometries and the impurity concentrations[7] are very similar in both cases. It is therefore possible

[7] A resulting depletion voltage of the germanium diode mostly depends on the diode geometry and impurity concentration. For the two exemplary crystals, the one with a smaller impurity concentration will be characterized by a smaller depletion voltage. The depletion voltage of both detectors was measured to be ≈ 1500 V. For the explanation of the dependence between impurity concentration and depletion voltage see the models in the section 2.1.2.

Table 3.4: Acceptance of the events from high energy peaks (^{228}Th spectrum) after applying a PSD cut for the 90% DEP acceptance. The training sample of 4900 events (same statistics for signal and background classes) was used for both channels. HS stands for "High Statistics" and denotes the column with the results for the training sample of 9500 events instead of 4500 (same as in Tab. 3.2). Channel 0 had worse energy resolution than channel 1 due to the increased leakage current (2.65 vs 1.75 keV at 1332.5 keV, respectively). Provided uncertainties are statistical only.

Peak energy [keV]	Isotope	Acceptance [%]		
		Channel 0	Channel 1	Channel 1 (HS)
1460.8	^{40}K	30.5 ± 5.2	30.9 ± 2.4	27.1 ± 2.4
1512.7	^{212}Bi	28.5 ± 4.8	33.7 ± 2.5	26.8 ± 2.7
1592.5	^{208}Tl (DEP)	90.4 ± 0.4	90.1 ± 0.2	90.1 ± 0.2
1620.5	^{212}Bi	29.3 ± 1.4	33.2 ± 0.8	28.9 ± 0.8
2103.5	^{208}Tl (SEP)	32.2 ± 0.8	30.8 ± 0.5	26.2 ± 0.5
2614.5	^{208}Tl	32.9 ± 0.2	33.4 ± 0.1	28.9 ± 0.1

to compare the effect of a leakage current on the PSD performance, while excluding other possible factors listed above. From the measurements of the "test point" voltage[8], the leakage current value for channel 0 detector was approximated to be at around (450 ± 50) pA. The same type of measurement performed for channel 1 detector yields about (25 ± 10) pA.

Results of the comparison are gathered in Tab. 3.4. Columns 3 and 4 show the PSD efficiency for both channels. The training was performed on 4500 events from DEP and 1620 keV FEP regions each, the training statistics for channel 0 detector is smaller due to readout triggered by only channel 1 detector. To make a fair comparison, the number of events available for training was equalized for both channels. Additionally, the table contains a fifth column with the results for channel 1 with the enlarged training statistics of 9500 events (same data as in Tab. 3.2). The results are rather surprising - the survival probability for the events from ^{40}K and ^{212}Bi peaks are actually smaller for channel 0 detector, which has higher noise level than channel 1 detector. For the lower energy FEPs (1460.8 keV and 1512.7 keV) the differences are within the statistical errors, but the difference in the survival probability for 1620.5 keV FEP is beyond the errors. On the other hand, acceptance for the SEP events is smaller for channel 1 detector, but in both cases the difference is not greater than 4%. Even then compared to the optimal conditions (channel 1 HS, last column), the survival probability differences are smaller than 6%.

It can be concluded that even a quite large increase in a leakage current (hundreds of picoamperes) does not significantly impair the PSD performance in semi-coaxial detectors for the presented method. This can be expected if one takes into account the spectral properties of the leakage current induced noise. Even though it has a flat spectral distribution at the input of the preamplifier, at the output its magnitude decreases with the increasing frequency. This is due to the transfer function of the preamplifier in a charge-sensitive configuration. Therefore, the effect of noise is largest for signals which occupy low frequency part of the spectrum (or, in other words, the ones associated with large time constants). This is the reason why large shaping times yield bad energy resolution in systems with large leakage currents (e.g. room temperature semiconductor detectors). On the other hand, a rising edge of the preamplifier signal is much faster (hundreds of

[8] The test point is a voltage node in the preamplifier, which allows for the leakage current measurements. It is placed just after the feedback loop resistor R_f. Since the leakage current does not flow through the JFET's gate due to its very high impedance, it has to flow through the feedback resistance R_f. Knowing the value of R_f and the test point voltage without the HV applied, it is possible to estimate the leakage current from the voltage drop on R_f.

nanoseconds), so it is more affected by its high-frequency counterpart. However, low-frequency noise can still propagate to the PSD input variables via pulse height normalization[9], which is used in order to remove energy dependency from the analyzed pulse. It can be concluded that the PSD performance is affected by the combination of both low- and high-frequency noise components, but the results from the presented analysis show that low frequency component has small influence of the final PSD efficiency. This fact is quite fortunate from the points of view of the passivated HPGe detectors operated in the LAr, for which the slow increase of the leakage current was reported [101].

3.2 Application of the Pulse Shape Discrimination to the n-type semi-coaxial detector

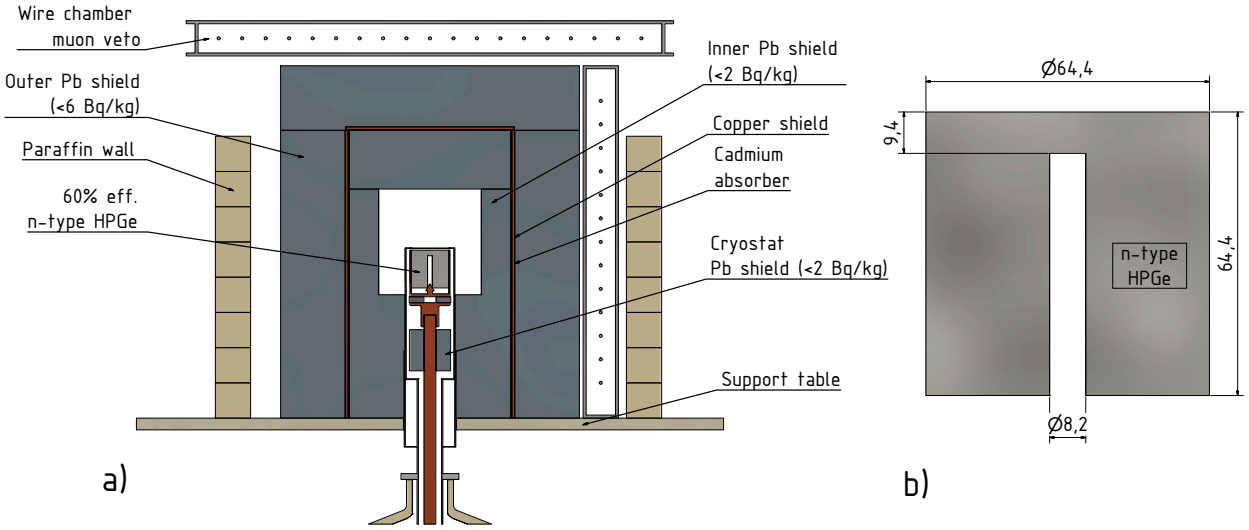


Fig. 3.13: *Left panel a)* – schematic drawing of the detector and the shield. Lead with the lowest available ^{210}Pb content was put the closest to the detector, while lead with slightly higher radioactivity level was used as an outside shield. Cadmium neutron absorber and copper sheets are placed in between the lead rings. Three gasous wire chambers, outside the shield, are working as a muon veto (right, top and on the back – not shown on the drawing).
Right panel b) – detailed cross sectional view of the crystal (n-type HPGe, all dimensions in mm).

The next step in the study of the multivariate PSD methods was to test the efficiency of the algorithm (section 3.1) on a larger n-type detector. A 60% relative efficiency detector (Fig. 3.13), produced by DSG (Detectors System GmbH. in Mainz, Germany), is contained in a shield made of low radioactivity lead and cadmium neutron absorber. Three muon veto wire chambers are placed outside the shield, together with paraffin bricks used for neutron thermalization. A measurement chamber inside the shield is continuously flushed with the nitrogen gas, produced by the LN_2 boiloff from a dewar, to remove the radon gas. The spectrometer is operated in the surface laboratory for low activity measurements and it is equipped with a traditional readout chain – an analogue shaping amplifier (ORTEC 672) and a multi-channel analyzer (FastComtec MCA-3). The acquisition setup was modified for the pulse shape measurements: the output of a resistive feedback preamplifier was coupled with a SIS3302 100 MHz/16 bit FADC (same model that was used in the measurements described in Sec. 3.1). Unlike the previous measurements for smaller semi-coaxial detectors, the external amplification of the signal was not needed, since the preamplifier circuit has a built-in linear amplifier stage.

[9]This is not the feature of the presented method – all PSD methods known to the author use some kind of pulse height normalization.

Data was processed in the same way as it was described in Sec. 3.1, the energy filter optimization procedure yielded 2.15 keV FWHM of 1332.5 keV ^{60}Co peak. Apart from the ^{60}Co isotope, the measurements were performed with the following radioactive sources:

- ^{228}Th – data for the training sample selection and efficiency assessment,
- ^{56}Co – the source emits several high energy gammas – 2598.5 keV (intensity: 17.3%), 3201.9 keV (3.2%), 3253.4 keV (7.9%) – which allow for the independent PSD efficiency cross-check in multiple energy regions. The lines also give rise to double- and single-escape peaks, therefore a relatively clean sample of both SSEs and MSEs is available.

3.2.1 Evaluation of the ^{228}Th data

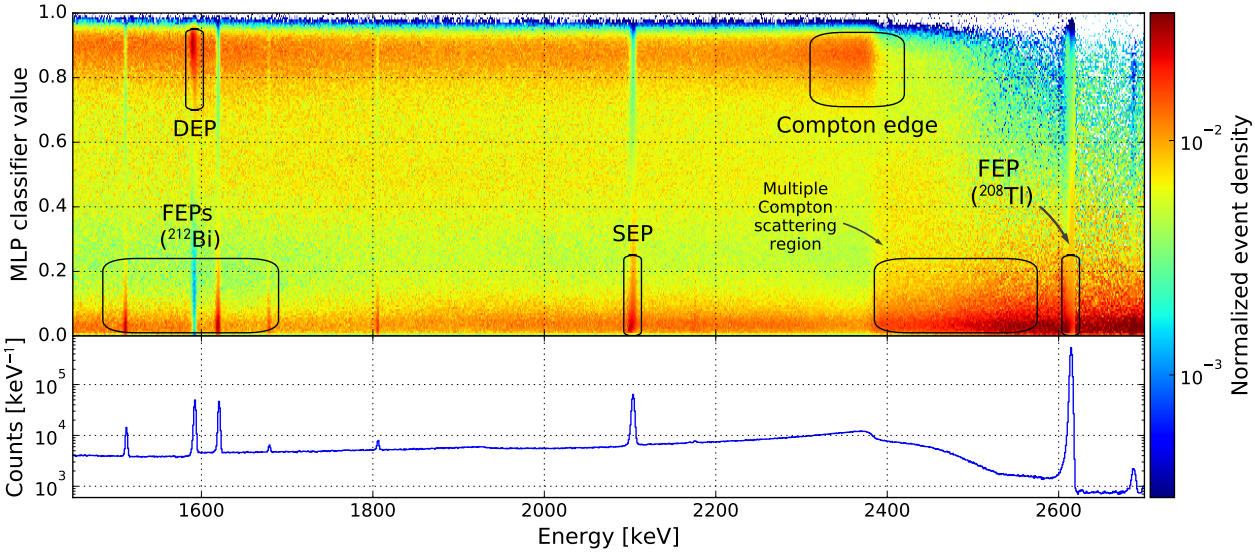


Fig. 3.14: *Top panel* – 2D normalized histogram of classified events from the ^{228}Th measurement for the n-type DSG60 semi-coaxial detector. Similar structure of the classifier distribution is observed as in the measurements for other detectors (compare with e.g. Fig. 3.8).
Bottom panel – high energy part of ^{228}Th energy spectrum.

After a short ^{60}Co run, to collect data for energy optimization and calibration, the detector was irradiated with the a ^{228}Th source (still sealed in a plastic box). Due to its low activity ($\approx$ 2 kBq) it was placed directly on the detector's endcap.

When the data was gathered, a training procedure characterized the same parameters as in Sec. 3.1.5. For clarity, the parameters can be shortly summarized below:

- *training sample*: the DEP and FEP at 1620 keV were used as a source of single-site and multi-site event samples, respectively. However, due to longer measurement times more statistics was available. 20000 events were used from both DEP and FEP energy regions.
- *dimensionality reduction*: the input variables were reduced to 12 PCA components. The same behavior was observed as in the case of a smaller p-type semi-coaxial detector: using more than 12 components did not improve the PSD efficiency.

Table 3.5: Acceptance of the events from high energy peaks from the DSG60 n-type semi-coaxial detector (^{228}Th spectrum), after applying the PSD cut for 90% DEP survival probability.

Energy [keV]	Peak type	Acceptance [%]
1512.7	FEP (^{212}Bi)	25.0 ± 0.6
1592.5	DEP (^{208}Tl)	90.1 ± 0.1
1620.5	FEP (^{212}Bi)	27.2 ± 0.2
2103.5	SEP (^{208}Tl)	27.4 ± 0.1
2614.5	FEP (^{208}Tl)	33.1 ± 0.1

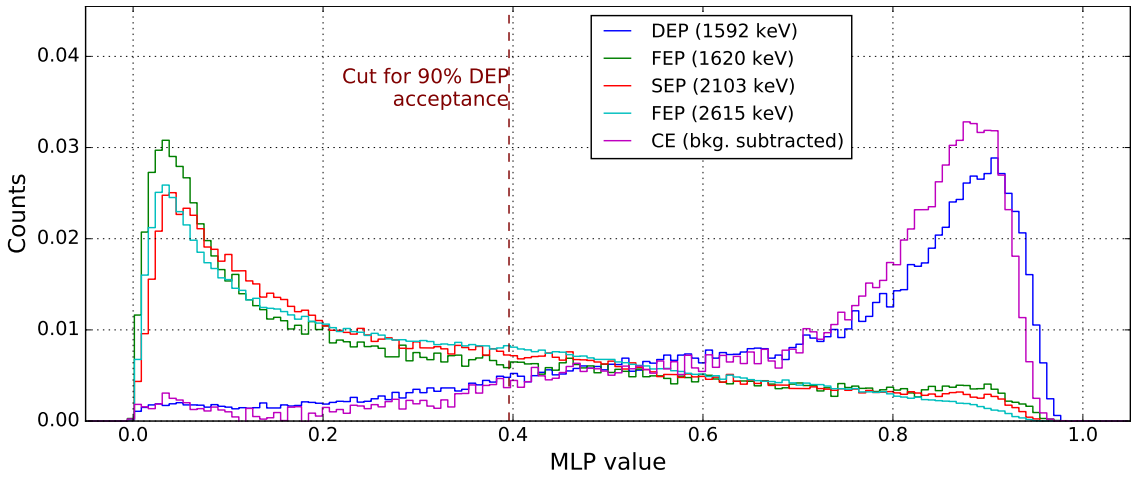


Fig. 3.15: Normalized classifier distributions of events from various peaks from ^{228}Th spectrum, as registered by the n-type semi-coaxial detector (60% relative efficiency). For each peak the events for the calculation are selected from the $E \pm 1/2$FWHM range, where E equals to the peak centroid. The survival probability of the Compton edge events (with a subtracted background) is 91.1%.

The data was analyzed following the same steps as in the previous measurements, classified ^{228}Th events were used to set a cut achieving 90% DEP survival probability. Acceptances of events from high energy γ peaks present in the ^{228}Th spectrum are collected in Tab. 3.5, while the classified events are plotted on the 2D histogram in Fig. 3.14.

When compared to the PSD efficiency for the small semi-coaxial detector (Tab. 3.2) similar levels are achieved. The main difference would be the acceptance for the FEP at 2614.5 keV, which is slightly worse than for the other peaks (by $\approx 5\%$). This is due to its classifier distribution (Fig. 3.15), which has a surplus of events in the MLP range of 0.35–0.5, when compared with other FEPs. On the other hand, the acceptance for SSEs from the DEP is almost the same as for the events from the Compton edge with a subtracted background (90.1% vs 91.1%). Classifier distributions for DEP and Compton edge regions are also much more similar than in the case of DSG10 (Fig. 3.12), for which only 72.6% of Compton edge events survived the cut. Observed difference cannot be explained by a different number of events used for the training process. It was repeated with a limited statistics of 10000 events and virtually the same efficiency was obtained. Three possible explanations can be provided:

- statistics used for training the method for DSG10-3 detector (only 9500 events) may still be cause the energy dependent behavior in the SSEs classification,
- in the smaller semi-coaxial detector signal signature is vastly different for the events from Compton edge than from their DEP counterparts,
- in the smaller detector the Compton edge events can penetrate it relatively deeper (when compared with its dimensions). Therefore, the effect may be due to the volumetric dependency of the classifier.

The SSE acceptance should be therefore always cross-checked with events from other energy regions, than the one used in the training process. This can be achieved by either using Compton edge events (or DEP events if Compton edge region was used as the SSEs sample) or by irradiating the detector with a different radioactive source, emitting high energy γ-rays.

3.2.2 Evaluation of ^{56}Co data

A next step in the PSD efficiency evaluation was to irradiate the detector with a ^{56}Co source. The source was produced in Henryk Niewodniczanski Institute of Nuclear Physics PAN in Cracow by the proton irradiation of thin iron foils in the cyclotron. The data processing procedure was

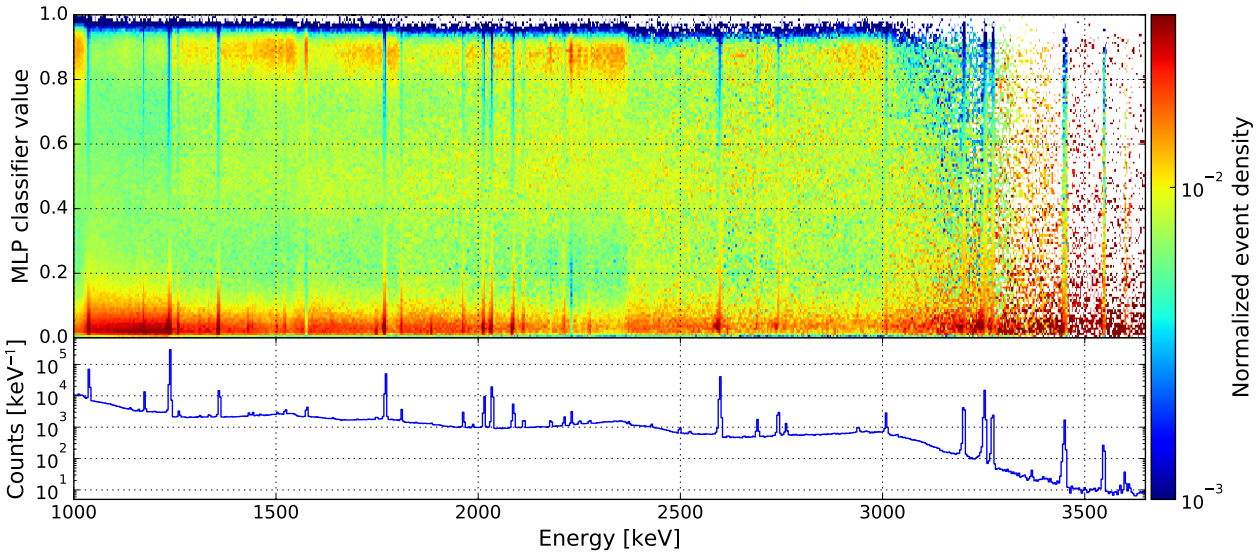


Fig. 3.16: *Top panel* – 2D normalized histogram of classified events from the ^{56}Co dataset. Unlike for the ^{228}Th histogram (Fig. 3.14), several Compton edge structures are visible for classifier value ≈ 0.9, as well as two DEPs: at 1576.5 and 2231.5 keV.
Bottom panel – high energy part of ^{56}Co energy spectrum.

carried out in the same way as for the ^{228}Th dataset (same quality cuts and energy filter parameters). Training metadata from ^{228}Th measurement (both the PCA matrix and MLP weights) was then used to classify the events from ^{56}Co dataset (no separate training was performed). High energy FEPs from ^{56}Co give rise to the DEPs, SEPs and Compton edge regions for vastly different energy values than in the case of ^{228}Th. A normalized 2D histogram of the classifier distribution as a function of energy is shown in Fig. 3.16. A similar two-band structure is obtained like for the ^{228}Th dataset.

PSD efficiency values for events from peaks in the ^{56}Co spectrum are collected in Tab. 3.6 (only peaks with statistics over 2000 counts). The table is divided into three parts, showing FEPs, DEPs and SEPs survival probabilities separately. Best efficiency (largest suppression) is achieved for the relatively low energy FEPs, in the range of 1000 – 1400 keV ($\approx 23\%$ acceptance). In the

Table 3.6: A summary of the acceptances of the events constituting the peaks in ^{56}Co energy spectrum (energy $>$ 1000 keV and more than 3000 counts). The cut was set to the 90% 1592 keV DEP (^{228}Th) acceptance. Provided uncertainties are statistical only.

Energy [keV]	Peak type (intensity [%])	Acceptance [%]	Energy [keV]	Peak type (intensity [%])	Acceptance [%]
1037.8	FEP (14.1)	22.0 ± 0.1	3202.0	FEP (3.2)	30.6 ± 0.3
1175.1	FEP (2.3)	22.3 ± 0.4	3253.5	FEP (7.9)	31.3 ± 0.2
1238.3	FEP (66.5)	22.5 ± 0.1	3273.1	FEP (1.9)	31.0 ± 0.4
1360.2	FEP (4.3)	22.2 ± 0.2	3451.2	FEP (0.9)	31.0 ± 0.6
1771.4	FEP (15.4)	24.5 ± 0.1			
1810.8	FEP (0.6)	23.1 ± 1.4	1576.5	DEP	87.5 ± 1.0
1963.7	FEP (0.7)	25.6 ± 1.1	2231.5	DEP	88.3 ± 1.3
2015.2	FEP (3.0)	26.0 ± 0.3			
2034.8	FEP (7.8)	25.8 ± 0.2	2087.5	SEP	21.7 ± 0.5
2598.5	FEP (17.0)	28.6 ± 0.1	2742.5	SEP	34.2 ± 0.6
3009.6	FEP (1.0)	31.9 ± 0.8			

1700 – 2100 keV range, the survival probability is equal or better than for the FEP (1620.5 keV) and SEP (2103.5 keV) from ^{228}Th spectrum. This fact proves that no overtraining takes place – the efficiency is virtually the same for the peaks at different energy regions. FEPs from 2600 – 3450 keV range have slightly higher survival probability (= worse PSD efficiency), at the level of $\approx 31\%$. The FEP at 3009.6 keV is the exception with even larger survival probability, however it lies on the Compton edge of 3202.5 keV, 3253.5 keV and 3273.1 keV peaks, therefore its abnormal background composition may influence its acceptance[10]. FEPs from this region are characterized with the similar acceptance value as the 2614.5 keV peak from the ^{228}Th dataset (33.1%).

Although in principle every FEP can give rise to both DEPs and SEPs, the ones with the considerable statistics can be observed only for FEPs with high intensity. In the analyzed dataset, only two DEPs were available for analysis: 1576.5 keV (from 2598.5 keV FEP – 17.0%) and 2231.5 keV (3253.5 keV 7.9%). Both of them are characterized by high acceptance, 87.5% and 88.3%, respectively, and the values are very close to the 90% survival probability, set for the 1592.0 keV DEP from the ^{228}Th dataset. A very good agreement between the DEPs confirms that the method shows no signs of the overtraining.

3.2.3 Increasing the FEP detection sensitivity with the PSD

While in the $0\nu\beta\beta$ experiments the expected signal is of a single-site type, in γ spectrometers, based on the HPGe detectors, the registered FEPs with the energy of a few hundred keV contain events of mostly multi-site type [78]. On the other hand, the observed Compton continuum background contains a large fraction of SSEs. Therefore, in principle, the PSD cut presented in previous sections can be inverted to accept MSEs from FEPs and suppress SSE part of the Compton continuum background.

The question remains: what is the optimal cut value to obtain the highest background suppression and, at the same time, a maximal acceptance of FEPs? A commonly used formula for comparison of different gamma spectrometry systems is the Figure of Merit (FoM), described e.g. in [93]. It combines the FEP detection efficiency and the background level into a single parameter and it is calculated as follows:

$$\mathrm{FoM} = \frac{\varepsilon(E)}{\sqrt{R(E)B(E)}} \tag{3.1}$$

where:
E – γ ray energy,
$\varepsilon(E)$ – detection efficiency,
$R(E)$ – energy resolution,
$B(E)$ – background level.

Since the efficiency $\varepsilon(E)$ is directly proportional to the acceptance of a peak $\mathrm{A_{PSD}}$ and the resolution $R(E)$ has the same value before and after the PSD cut, the FoM ratio before and after the cut depends only on the obtained acceptance $\mathrm{A_{PSD}}$ and a ratio of background levels before and after the cut:

$$\frac{\mathrm{FoM_{PSD}}}{\mathrm{FoM}} = \frac{\varepsilon_{\mathrm{PSD}}(E)}{\varepsilon(E)}\sqrt{\frac{B(E)}{B_{\mathrm{PSD}}(E)}} = A_{\mathrm{PSD}}(E)\sqrt{\frac{B(E)}{B_{\mathrm{PSD}}(E)}} \tag{3.2}$$

$$\mathrm{FoM}_{impr} = \left(\frac{\mathrm{FoM_{PSD}}}{\mathrm{FoM}} - 1\right) 100\% \tag{3.3}$$

[10] Due to the energy spectrum being densely populated by the peaks and their corresponding Compton edges, the SSE/MSE ratio in the background continuum may be energy dependent. Therefore some discrepancies may result from the subtraction of the linear background under the peak before and after the application of the PSD cut.

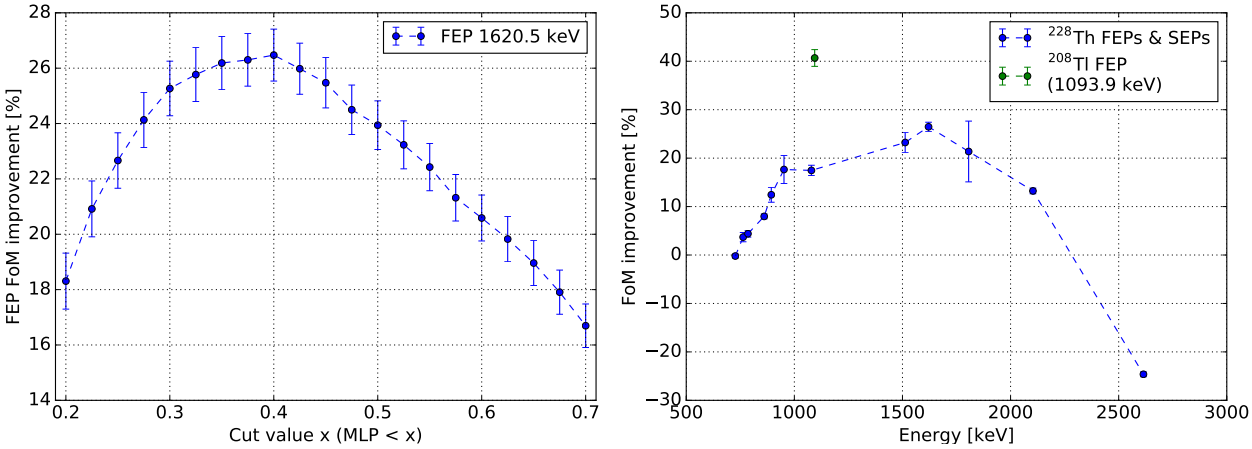


Fig. 3.17: *Left panel* – optimization of the MSE cut for the a maximal FoM$_{impr}$ of the 1620.5 keV peak from ^{212}Bi. The best improvement is achieved for the MLP < 0.4 cut. Dashed lines are only for a visual guidance.
Right panel – FoM$_{impr}$ obtained for MSE peaks from the ^{228}Th spectrum after the application of an optimal PSD cut (maximal FoM$_{impr}$). Dashed lines are only for a visual guidance.

The FoM$_{impr}$ parameter is convenient from the optimization point of view, since it constrains the FEP events acceptance and the background reduction rate. Both values depend on a given cut value, therefore, an optimization procedure consists of cut value variation and calculation of FoM$_{impr}$ for every step. The procedure is illustrated on the left panel of Fig. 3.17. In this case, FoM$_{impr}$ was calculated for the 1620 keV FEP from ^{228}Th dataset, but the same trend was observed for almost all other peaks. An optimal cut value in this case is MLP < 0.4. The right panel shows the FoM$_{impr}$ for other peaks in a ^{228}Th spectrum, an abnormal value of the FoM$_{impr}$ for 1093.9 keV (green point) is explained later in this section. More detailed analysis results, together with the obtained acceptances after a cut application, are collected in Tab. 3.7.

According to the Monte Carlo simulation results from [78], the multi-site component[11] in FEPs becomes greater than the single-site one (photoelectric effect) for the γ-ray energies over $\approx$150 keV. Therefore, the FoM$_{impr}$ by the PSD should already by observed from these energy values. However, FoM$_{impr}$ for 727.1 keV ^{212}Bi FEP is $\approx 0\%$ – the peak area reduction ($\approx 30\%$) effectively cancels out the background suppression effect. This happens even though the peak energy is much higher than the 150 keV value discussed above. The only explanations for the discrepancy

Table 3.7: Acceptances and FoM$_{impr}$ for the peaks in ^{228}Th dataset after the application of the optimized FoM$_{impr}$ PSD cut (MLP < 0.4). Energy values are taken from the gaussian fit and may differ from the ones provided in nuclear data tables.

Energy [keV]	Isotope	Acceptance [%]	FoM improvement [%]
727.1	^{212}Bi	71.7 ± 0.1	-0.2 ± 0.4
763.2	^{208}Tl	73.3 ± 0.6	3.7 ± 1.0
785.4	^{212}Bi	73.1 ± 0.4	4.4 ± 0.7
860.6	^{208}Tl	74.0 ± 0.2	8.0 ± 0.5
893.4	^{212}Bi	75.3 ± 1.0	12.4 ± 1.5
952.1	^{212}Bi	76.2 ± 1.8	17.7 ± 2.9
1078.9	^{212}Bi	74.3 ± 0.6	17.5 ± 1.1
1094.1	^{208}Tl	88.3 ± 1.0	40.7 ± 1.7
1512.6	^{212}Bi	74.4 ± 1.2	23.2 ± 2.0
1592.3	^{208}Tl*	9.3 ± 0.2	-84.1 ± 0.3
1620.6	^{212}Bi	72.9 ± 0.4	26.5 ± 0.9
1806.0	^{212}Bi	74.8 ± 3.9	21.4 ± 6.3
2103.5	^{208}Tl**	72.3 ± 0.3	13.2 ± 0.6
2614.5	^{208}Tl	66.4 ± 0.1	-24.6 ± 0.5

* – Double Escape Peak
** – Single Escape Peak

[11] Single/multiple Compton scattering + photoelectric effect.

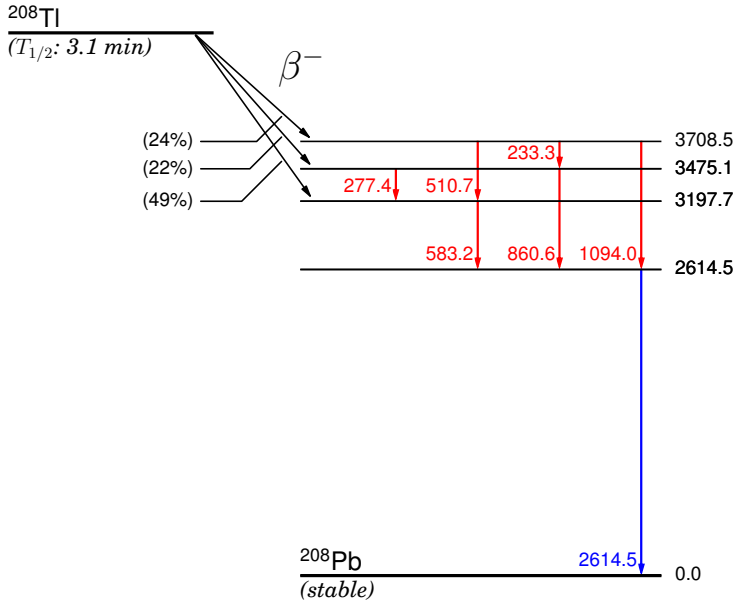


Fig. 3.18: Decay scheme with the most intensive ($> 20\%$) β^- decays of ^{208}Tl. Red lines are the gamma transitions which can give rise to the 1093.9 keV peak.

would be a limited PSD efficiency for low energy events or the background containing too many MSEs, which are not distinguishable from the FEP events. From the analytical considerations regarding the shape of the Compton continuum (e.g. from [102]), it is known that the content of the single-scattered Compton events[12] in the spectrum starts rising with the deposited energy in the detector higher than ≈1200 keV.

For the higher energy peaks (1500 – 1800 keV) the FoM$_{impr}$ reaches the value of 21–26.5%. However, the 1093.9 keV peak has an abnormal value of FoM$_{impr}$ of 40.7%, while for its neighbor at 1078.9 keV it is only 17.5%. Its acceptance is also much higher than the for other peaks (88.3%). This observation can be explained by analyzing the ^{208}Tl decay scheme (Fig. 3.18) – apart from the direct transition between 3708.5 and 2614.5 keV levels ($E_\gamma = 1094$ keV), the indirect transitions through intermediate levels are also possible. Emitted multiple γ-rays have lower energy values[13], but of course the total energy is the same. For the close source geometry measurements, like the one performed for the ^{228}Th dataset, those multiple γ-rays can be emitted and absorbed faster

[12] From the incident γ-ray with energy ≈ 2.6 MeV.
[13] Namely: $233.3 + 860.6$ keV, $510.7 + 583.2$ keV, $233.3 + 277.4 + 583.2$ keV.

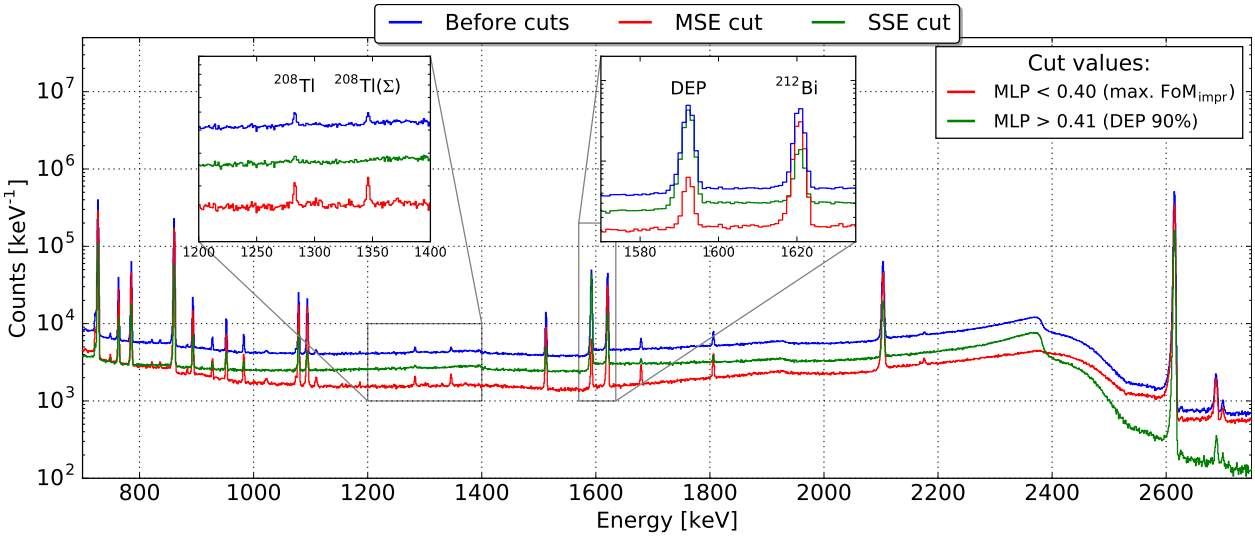


Fig. 3.19: ^{228}Th spectrum, with the visualized effects of the SSE (fixed at 90% DEP acceptance) and the MSE (optimized for a highest FoM$_{impr}$, see text for details) PSD cuts. Two highlighted spectrum ranges show the effect of the PSD on small intensity MSEs peaks (full energy and summation) from ^{208}Tl (left inset) and DEP/FEP doublet (right inset).

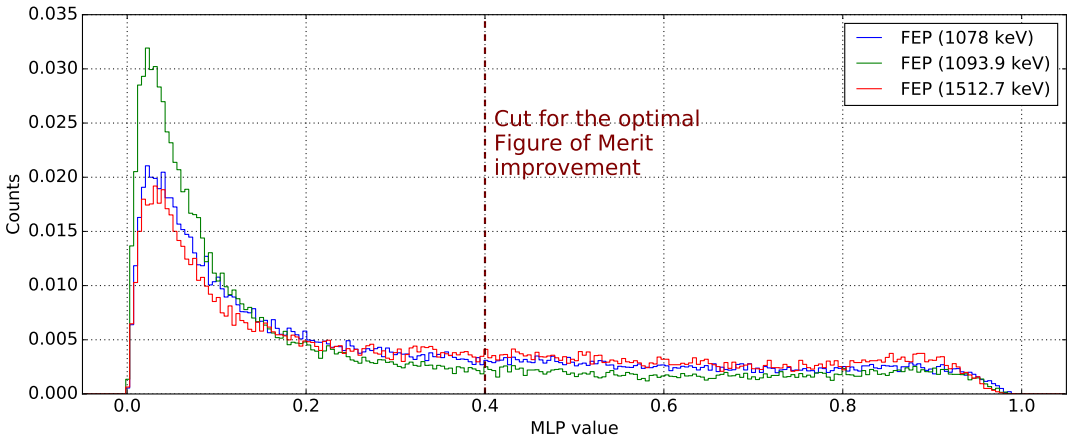


Fig. 3.20: Classifier distribution for 1093.9 keV peak from ^{228}Th spectrum, as well as for neighbouring FEPs. The contribution from summation events (see text for details) make the distribution much more peaked at the value corresponding to the MSEs band (≈ 0.05), which in turn boosts both the peak acceptance and its FoM$_{impr}$.

than the detector's resolving time, i.e. they cause a signal with the same pulse height that the one originating from absorbed 1093.9 keV γ-ray from the direct transition. If so, 1093.9 keV peak contains events from both types of transitions, corresponding to a FEP and a summation peak. However, their pulse shapes may be actually different – γ-rays causing the summation peak can be registered in sites quite far apart, since cascade gammas are emitted with considerable angle between their directions [103]. Therefore, there is a high probability that they will deposit energy in different parts of the detector and resulting pulse shape will be classified as a MSE. On the other hand, a scattered 1093.9 keV γ-ray will probably travel a smaller distance inside the detector, since low energy gammas after the scatterings have a short mean free path in germanium. In that case, the interaction sites will be closer to each other than for the summation peak events. This behavior considerably affects the classifier distribution for the 1093.9 keV peak – it is plotted in Fig. 3.20. Its shape, normalized to the total counts, indicates that it contains much more MSEs than the neighboring FEPs (also included in the plot).

In the case of the 2614.5 keV peak the achieved FoM$_{impr}$ is actually negative. This is expected if one takes into account that both the peak and the background in this energy range is of multi-site type, since no possible SSE with deposited energy over Compton edge (2382 keV) can occur if no gammas over 2614.5 keV are emitted. And if the peak and the background are both composed of MSEs, they are affected by the PSD cut in a similar fashion and no reduction of the single-site Compton continuum takes place. It is worth noting that even though the gamma have to interact in multiple places, the resulting signal shape can still be indistinguishable from the one originating from a single-site interaction. This is due to the fact that if the second interaction happens at the very short distance from the first one ($\sim$ 1 mm) the resulting superposition of the individual pulse shapes may not considerably differ from the one resulting from SSE. Due to the cylindrical symmetry of the germanium detectors similar effect can also occur if sites of energy deposition are localized on the same radius.

3.2.4 Evaluation of the background run

Apart from the measurements with the radioactive sources, a special run to measure the spectrometer's background was also performed. All muon veto chambers were turned on during the measurement and the veto signal was fed to the second channel of the FADC. This way the data can be analysed with or without the muon vetoed events. This is in contrast with the traditional analogue acquisition chain, where the vetoed events are not registered in the spectrum and infor-

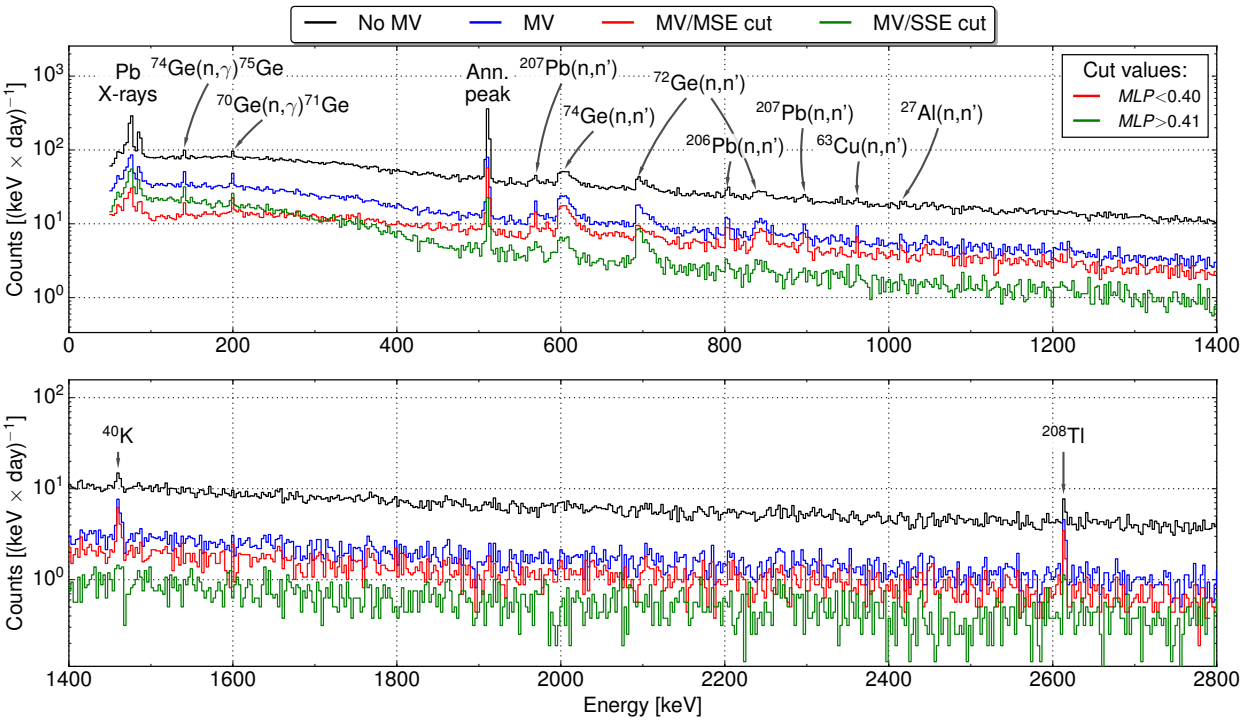


Fig. 3.21: Energy spectra of the ≈ 6 d background run in the DSG60 n-type semi-coaxial detector. Various peaks, originating mostly from fast neutron interaction can be seen in the 500 – 1000 keV energy range – for the detailed explanation of background components see the text. Abbreviations stand for: MV – muon veto (see. Fig. 3.13, MSE cut – multi-site events cut (optimal FoM_{impr}), SSE cut – single-site events cut (90% 1592.5 keV DEP acceptance).

mation about them is lost. The live-time of the measurement was 5 days and 16.5 hours – obtained spectra are plotted in Fig. 3.21.

Multiple peaks can be observed in the spectra, especially after application of the muon veto, which considerably lowers the continuous background level. Looking from the left (lower energies) several X-ray lines, originating from the lead shielding, can be recognized (72.8 keV, 75.0 keV, 84.9 keV and 87.4 keV). The X-rays can be emitted by lead atoms exited by the electrons, which are either muon-induced or originating from the ^{210}Bi decay[14] [104]. Commonly observed peak at 511 keV is mostly due to the annihilation of positrons, created by the decay of μ^+ in the shield [31].

Most of the observed peaks are due to neutron interactions with the detector and shielding material nuclei. Two peaks at 139.7 keV and 198.4 keV are caused by a neutron capture in ^{74}Ge and ^{70}Ge, respectively. The capture leads to a formation of metastable states in $^{75\mathrm{m}}$Ge (47.7 s) and $^{71\mathrm{m}}$Ge (20.4 ms), which later deexcitate via the γ-rays emission. Since the metastable states have considerably longer half-life than the spectrometer resolving time ($\sim$ µs), energy brought by the neutron interaction does not sum up with the deexcitation energy. On the other hand, the 500 – 1000 keV range contain numerous peaks resulting from inelastic fast neutron scattering on germanium in the detector. Peaks from the interaction with ^{72,74}Ge exhibit an anomalous asymmetric and broad shape, with the broadening to the higher energy values. They are caused by the sum of a deexcitation from a given energy level (eg. 691.6 keV in ^{72}Ge) and some extra recoil energy, deposited by the inelastic scattering of the neutron. Half-lives of excited level states are shorter or comparable with the detector resolving time (444.2 ns and 12.4 ps for ^{72}Ge and ^{74}Ge respectively), otherwise, a pulse with summed energy could not be observed by the acquisition chain. Two more (less distinctive) neutron induced peaks from cryostat materials are also visible

[14] A note regarding the ^{210}Bi presence in lead: ^{210}Bi is a daughter of ^{210}Pb, which is present in small amounts (ranging from <Bq/kg to $\approx 2.5\,\mathrm{kBq/kg}$ [31]) in natural lead ($^{210}\mathrm{Pb} \xrightarrow{\beta^-} {}^{210}\mathrm{Bi} \xrightarrow{\beta^-} e^- + \bar{\nu}_e + {}^{210}\mathrm{Po}$, $E^{Bi}_{endpoint} = 1.16\,\mathrm{MeV}$).

at 962.1 keV and 1014.4 keV (Cu and Al respectively). In the high energy part of the spectrum (bottom panel of Fig. 3.21) only two peaks are visible: ^{40}K at 1460.8 keV and ^{208}Tl at 2614.5 keV.

Fig. 3.21 shows also the effect of PSD cuts on the obtained energy spectrum (red and green lines). The same cut values were used as in the previous sections (SSE cut: MLP > 0.41, 90% DEP acceptance; MSE cut: MLP < 0.4, best FoM$_{impr}$). As expected, the MSE cut has high acceptance of FEPs like ^{40}K or ^{208}Tl. Neutron induces peaks in germanium are clearly of the multi-site type, since the deexcitated γ-ray usually deposits its energy further from the primary interaction site. The exception here is the peak at 691.6 keV from ^{72}Ge(n,n'). When compared to its neighboring peak at 595.85 keV, it is clearly more suppressed by the MSE cut. This unusual behavior can be explained by taking into account that a transition from the first exited state in ^{72}Ge (691.6 keV) is a forbidden one ($0^+ \rightarrow 0^+$) and therefore the nucleus must deexitate in an internal conversion process with an emission of the electron [105]. Since it has a much smaller range in the detector than its γ-ray counterpart, all energy is deposited in a small volume, resulting in a SSE. The SSE cut reduces the 595.85 keV peak to about 24% and 691.6 keV to $\approx 60\%$ (exact values are difficult estimate due to the very low statistics).

3.3 Development of the Pulse Shape Discrimination method for the BEGe-type spectrometer

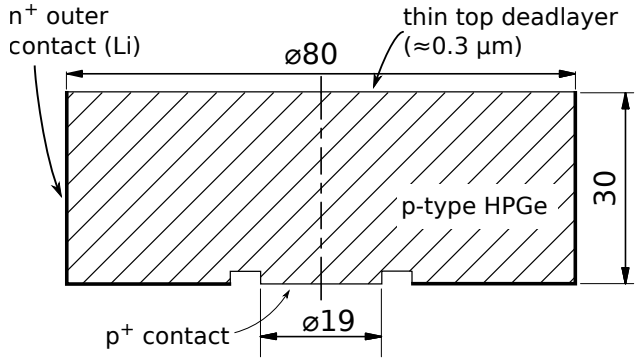


Fig. 3.22: Schematic drawing of the 50% relative efficiency BEGe detector used in the measurements in HADES underground laboratory. Unlike the semi-coaxial or BEGe type detectors used in GERDA, BEGe detectors used for the γ-ray spectrometry purposed are manufactured with a thin entrance window ($\approx 0.3\,\mu$m), increasing the efficiency for low energy γ-rays.

PSD methods presented in Sec. 3.1 and Sec. 3.2 did not use any prior assumptions on the shape of the signal – the input information, that the algorithm takes into account, is just a sampled, normalized waveform. Therefore, the method should also work when applied to the other detector types, providing that the pulses contain a signature that even theoretically allows for the SSE/MSE discrimination. The possibility of applying the same method to both semi-coaxial and BEGe type detectors would have certain advantages, especially from the point of view of software integration in large $0\nu\beta\beta$ decay experiments like GERDA, as well as in studies of systematic uncertainties e.g. using MC simulations.

Data analyzed in this PSD study was gathered in European Commission's Joint Research Centre Geel (formerly Institute for Reference Materials and Measurements). It was registered by the underground (HADES underground laboratory) γ-ray spectrometer (internal name Ge-5), based on the BEGe HPGe detector. The crystal is of p-type and produced by Canberra (model BE5030) – a schematic drawing is shown in Fig. 3.22. Muon flux in the laboratory is reduced by four orders of magnitude by a rock overburden (500 m w.e.) [106] and the detector is surrounded by a low-radioactivity lead shield in order to further reduce its background.

Similarly to the procedure carried out for the semi-coaxial detectors, the training data containing both SSE and MSE samples is needed. ≈ 100 kBq ^{228}Th source was used for this purpose, which was placed inside the measurement chamber. The waveforms were registered using SIS3302, a 16-bit 100 MHz FADC. Typical normalized pulses are shown in Fig. 3.23. Multiple current peaks in the MSE pulse are easily seen, while its SSE counterpart contain only one. This

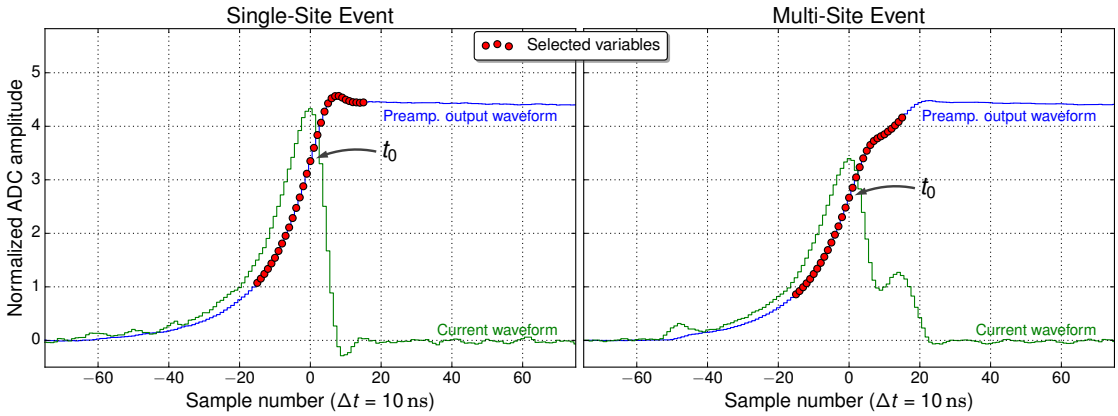


Fig. 3.23: Pulse shapes (blue) from typical single-site (left panel) and multi-site (right panel) events. The current pulses (green) were obtained by differentiating the preamplifier output and smoothing with a 50 ns moving average filter. Red points represent the amplitudes selected as the input variables to the PSD algorithm. t_0 indicates the amplitude, which corresponds to the maximal current of the pulse.

is not the case for signals from semi-coaxial detectors, for which the multiple current peaks are visible also for the SSEs. Trapezoid energy filter was used for the energy reconstruction – the obtained energy resolution values at DEP (1592.5 keV) and ^{208}Tl FEP (2614.5 keV) are 2.14 keV and 2.65 keV respectively.

3.3.1 Evaluation of the ^{228}Th data

Input variables were reduced to 8 PCA components before applying the training process. SSEs and MSEs samples have been taken from the DEP at (1592.5 keV) and ^{212}Bi at 1620.5 keV, respectively. Centroid $\pm 1/2$ FWHM energy range was used for both DEP and FEP, resulting in 20000 training pulses for each event type. Classified ^{228}Th data is plotted on a normalized 2D histogram in Fig. 3.24 – two very well separated bands corresponding to MSEs (MLP $\approx$ 0.05) and SSEs (MLP $\approx$ 0.05) are clearly visible. Similar structure can be observed for the classifier applied

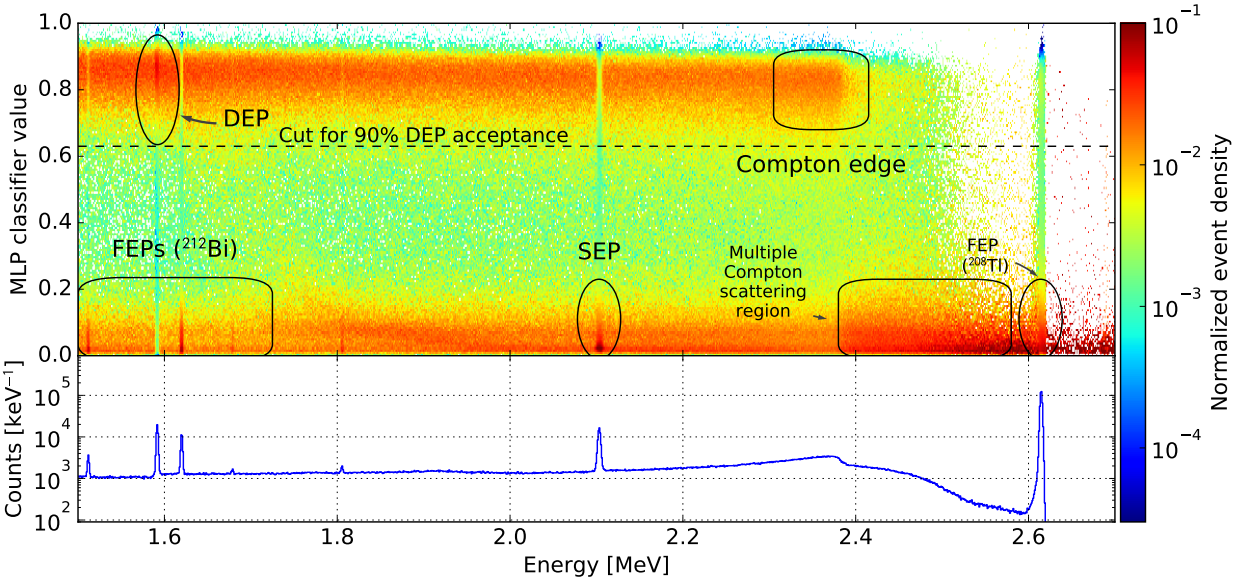


Fig. 3.24: *Top panel* – normalized 2D histogram of classified events from the ^{228}Th measurement for the BEGe type detector, operated in the underground laboratory HADES. A similar structure of the classifier distribution is observed as in the measurements for other detectors (compare e.g. Fig. 3.14) – single- and multi-site bands are clearly visible.
Bottom panel – ^{228}Th energy spectrum, corresponding to the events distribution in the top panel.

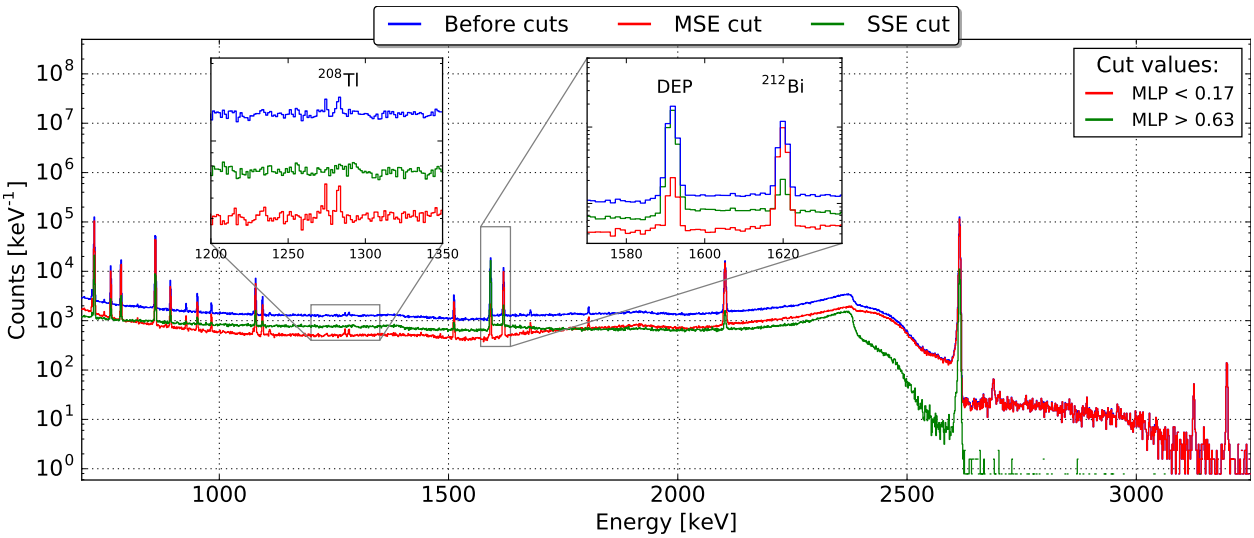


Fig. 3.25: Energy spectrum reconstructed from the ^{228}Th dataset (blue line). Red and green lines (MSE and SSE cuts, respectively) show an effect of the MLP neural-network classifier. The SSE cut, promoting events with the single-site energy deposition, was fixed at the 90% DEP acceptance, while the MSE cut was optimized for the maximum FoM$_{impr}$ (for details see the description in sec. 3.3.2). The insets show the regions of low intensity ^{208}Tl peaks (left) and DEP/^{212}Bi peaks (right, with the large content of SSEs/MSEs, respectively).

to the data from semi-coaxial detectors (for comparison see Fig. 3.8 and Fig. 3.14), however, in the case of the BEGe detectors the separation is much stronger – fraction of the events classifier in the "valley" between the bands is much smaller than for the semi-coaxial detectors.

In the BEGe type detectors, MSEs and SSEs can be distinguished very efficiently due to the weighting field distribution in the detector – the biggest signal contribution (the current peak seen e.g. at left panel of Fig. 3.23) comes from the charges drifting near the p^+ contact (for the p-type detector hole component dominates in the signal – the electrons drift towards the n^+ HV electrode, so their contribution is much smaller) [76]. Therefore, if the charge carriers are created far from the p^+ contact, their contribution to the signal can be signal after a considerable delay, in order of hundreds of nanoseconds. The delay is position dependent (longer for the charge drifting farther from the p^+ contact), so the induced pulse from the MSE is a superposition of time delayed SSE pulses (see right panel of Fig. 3.23). For a more detailed analysis see Sec. 2.

Table 3.8: Acceptance of the events from high energy peaks from BEGe type detector (^{228}Th spectrum) after applying the PSD cut for fixed 90% DEP survival probability.

Energy [keV]	Peak type	Acceptance [%]
952.1	FEP (^{212}Bi)	14.0 ± 1.9
1078.6	FEP (^{212}Bi)	13.8 ± 0.8
1093.9	FEP (^{208}Tl)	8.7 ± 2.3
1512.7	FEP (^{212}Bi)	10.5 ± 1.6
1592.5	DEP (^{208}Tl)	90.0 ± 0.2
1620.5	FEP (^{212}Bi)	11.1 ± 0.4
2103.5	SEP (^{208}Tl)	6.3 ± 0.2
2614.5	FEP (^{208}Tl)	8.7 ± 0.1

Table 3.9: Comparison of the PSD efficiency with the A/E method, the analysis was done for the same detector [85]. The cut was fixed at 91% DEP acceptance for both methods.

Energy [keV]	Isotope	Acceptance [%]	
		MLP	AoE
1592.5	^{208}Tl*	91.0 ± 0.3	91
1620.5	^{212}Bi	12.0 ± 0.4	24
2103.5	^{208}Tl**	6.7 ± 0.3	19
2614.5	^{208}Tl	9.6 ± 0.1	31

* – Double Escape Peak

** – Single Escape Peak

Once again the survival probabilities of the events from high energy peaks in the ^{228}Th spectrum were calculated – the results are collected in Tab. 3.8. For the fixed cut at 90% DEP events acceptance the survival probability of all SEPs and FEPs is under 15%. Also, the same effect, present in analysis of DSG60 data, was observed for the BEGe detector – namely the lowered acceptance of 1093.9 keV peak due to the summation events. Even though the source was further away from the detector, its high activity still contributed to the summation effects. High efficiency in suppressing summing events is confirmed by the vetoed events from 2700 – 3500 keV energy range – the MLP classifier rejects over 99% of them (see Fig. 3.25).

Results of the analysis can be compared with the A/E classifier (for the description see Sec. 2.2.1), because the same detector was analysed in both cases [85]. Comparison data for two methods is collected in Tab. 3.9, the cut for the MLP classifier was moved for 91% DEP acceptance, to obtain the same SSE efficiency for both methods. It can be seen that the survival fraction for the MSEs from peaks is two (^{212}Bi FEP) to three times (SEP) smaller. Some differences may be due to the different source used in the measurement (^{232}Th vs ^{228}Th), for which additional peaks are present in the spectrum near DEP at 1592.5 keV (namely 1588.19 keV from ^{228}Ac). For a smaller BEGe detector (Ge-8) described in [85] the obtained efficiency with A/E is comparable to the one obtained with MLP (for 92% DEP acceptance, ^{212}Bi FEP was reduced to 19% and SEP to 7%). The comparison proves that the PSD analysis based on MLP with the PCA reduction is a good alternative for the detectors where A/E analysis becomes problematic. On the other hand, in optimal conditions similar results can be obtained with both methods. The big advantage of the presented method is that, contrary to A/E, it does not need to perform the energy dependence corrections of the classifier, at least in the vacuum cryostat conditions (compare with the results from GERDA detectors in 4).

3.3.2 Improving the sensitivity of the BEGe-based γ spectrometer via Pulse Shape Discrimination

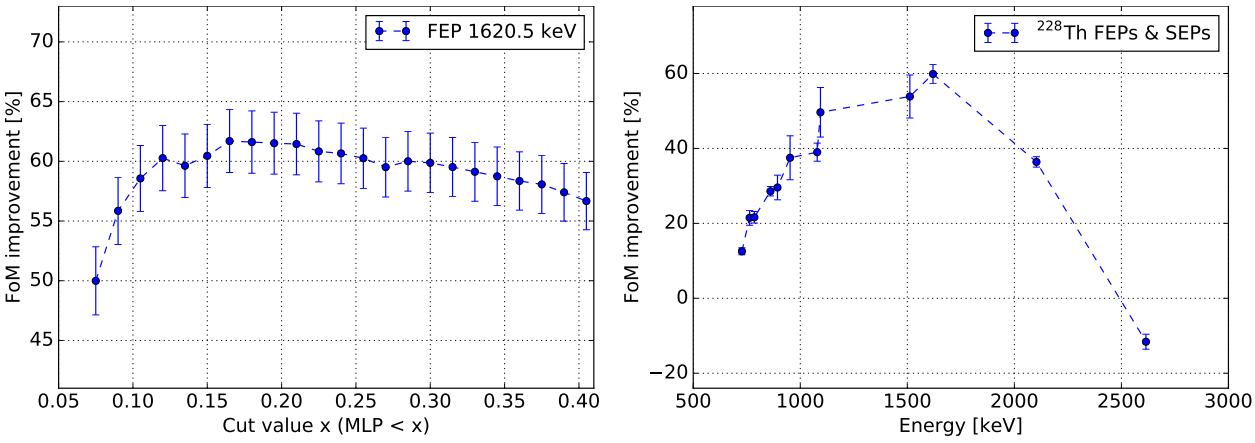


Fig. 3.26: *Left panel* – optimization results of the MSE cut for the maximal FoM$_{impr}$ of the 1620.5 keV peak (^{212}Bi). The best improvement is achieved for a cut value of MLP < 0.4 cut. Dashed lines are for visual guidance only.
Right panel – FoM$_{impr}$ obtained for the MSE peaks from ^{228}Th spectrum after application optimal PSD cut. Dashed lines are for a visual guidance only.

Applying a similar procedure as the one described in Sec. 3.2.3, the background reduction using PSD was also studied for the BEGe-based spectrometer. An optimization procedure for the 1620.5 keV peak is shown in the left panel of Fig. 3.26. The optimal cut value was found to be MLP < 0.165. The energy spectrum after application of the cut is plotted in Fig. 3.25. From the γ spectrometry point of view probably the most interesting feature is the background reduction for

Energy [keV]	Isotope	Acceptance [%]	FoM improvement [%]
727.4	^{212}Bi	69.8 ± 0.3	8.8 ± 0.9
763.4	^{208}Tl	72.1 ± 1.1	17.4 ± 2.0
785.6	^{212}Bi	71.2 ± 0.7	18.6 ± 1.5
860.7	^{208}Tl	71.7 ± 0.3	25.5 ± 1.3
893.5	^{212}Bi	72.7 ± 1.7	28.0 ± 3.3
952.3	^{212}Bi	73.9 ± 3.1	34.6 ± 5.8
1079.0	^{212}Bi	74.0 ± 1.2	37.7 ± 2.5
1094.2	^{208}Tl	77.6 ± 3.3	48.1 ± 6.6
1511.9	^{212}Bi	77.9 ± 2.8	53.0 ± 5.8
1591.7	^{208}Tl*	3.1 ± 0.2	-93.5 ± 0.4
1620.0	^{212}Bi	75.9 ± 0.9	61.7 ± 2.6
2103.2	^{208}Tl**	81.6 ± 0.6	39.5 ± 1.5
2614.6	^{208}Tl	72.2 ± 0.2	-17.3 ± 1.9

* – Double Escape Peak

** – Single Escape Peak

Table 3.10: Acceptance and FoM$_{impr}$ values for the events from the peaks in the ^{228}Th dataset. The applied PSD cut (MLP < 0.17) was optimized for the maximal FoM$_{impr}$ of the 1620.5 keV peak (left panel of Fig. 3.26).

low intensity peaks, like the ones in the left inset in the Fig. 3.25. The peak at 1282.8 keV (on the right hand side) has the intensity of 0.052%, the lower energy one is the summation peak from ^{208}Tl (583.2 + 763.1 keV).

The best result of FoM$_{impr}$ is obtained for the 1620.5 keV FEP (Fig. 3.26 and Tab. 3.10) – almost 62% FoM$_{impr}$ is obtained. In the case of SEP, the background continuum contains a large fraction of MSE events from the pair production + scattering of annihilation radiation, which would explain a low FoM$_{impr}$ value, even with the quite high peak acceptance (5.7% higher than for 1620.5 keV FEP). One of the reason of high FEP FoM$_{impr}$ could be that it was used as the training sample, however, to test this hypothesis a separate training was done with the SEP instead of the FEP as the MSEs sample. The obtained result was the lower FoM$_{impr}$ values for both peaks ($(56.6 \pm 2.6)\%$ and $(38.9 \pm 1.5)\%$ for FEP and SEP, respectively), therefore such good FoM$_{impr}$ (compared to other peaks in the spectrum) must be caused by a favorable background composition (high SSEs fraction) for this energy range.

Chapter 4

Development of the Pulse Shape Discrimination methods for the external background suppression in GERDA Phase II

This chapter contains a description of the data structure in GERDA Phase II experiment, as well as a process of the development and application of the PSD methods. The analysis was performed for both types of detectors (semi-coaxial and BEGe) with the application of the PCA dimensionality reduction methods. The MLP neural-network classifier was used for the BEGE data, while the semi-coaxial dataset was analysed using the Projective Likelihood classifier. A second version of the Projective Likelihood based method was also included, i.e. the one using the variable summing instead of the PCA transformation.

Before the explanation of the analysis procedure it is mandatory to introduce the data structure of the GERDA experiment. To give an example, the concept of "tiers", while encountered also in other experiments, may not fully understandable for the reader not involved in the experiment. Afterwards, the chapter contains the description of the training procedure and the results for the calibration data (^{228}Th). After the PSD metadata has been obtained, it was applied to the so called "physics" data (without the presence of the calibration sources). The adjective "physics" is used since this is the part of the data on which the $0\nu\beta\beta$ analysis is performed. In the following analysis, the term "Phase II data" means the data from runs 53–64 (collected untill the unblinding in June 2016), unless stated otherwise. Lastly, the signal efficiency of the methods applied to the semi-coaxial detectors was also compared with the available MC simulated data.

Data cycles. All analysis steps in GERDA are done off-line[1], starting from the raw waveforms from FADC. The blinded data is available from the beginning of the experiment. Therefore, the analysis parameters, used for the energy reconstruction or energy calibration, can be in principle changed later if e.g. a new quality cut is introduced or a new energy filter improves the energy resolution. The data is divided into "cycles" – the cycle is the data released after reprocessing with the given calibration curves, cross talk matrices, quality cuts etc. The selection of the above mentioned parameters can influence the waveform energy reconstruction or even reject events that would be otherwise used in the training process. Analysis results presented in this work regard the data cycle v02.02, which was the most recent one before the unblinding on 17$^{\text{th}}$ of June 2016.

[1] Except for the physics data blinding, which uses energy values reconstructed by the FADC software on-the-flight, while digitizing the signal.

Table 4.1: Mapping of the GERDA Phase II detectors and the FADC channels (in parentheses). The mapping was not changed during all of the GERDA Phase II data taking. The prefix "GD" was omitted in front of the names of BEGe detectors (e.g. GD91A → 91A).

String1	String2	String3	String4	String5	String6	String7
91A(0)	ANG5(8)	02A+32B (11,12)	35C+76C (19,20)	ANG2(27)	00A+02C (30,31)	GTF112(37)
35B(1)	RG1(9)	32A+32C (13,14)	89D(21)	RG2(28)	79B(32)	GTF32(38)
02B(2)	ANG3(10)	89C(15)	00D(22)	ANG4(29)	91D(33)	GTF45(39)
00B(3)		61C(16)	79C(23)		32D(34)	
61A(4)		76B+00C (17,18)	35A(24)		89A(35)	
89B(5)			91B(25)		ANG1(36)	
02D+91C (6,7)			61B(26)			

Data structure. Data is divided in the so called "tiers" – the idea is that the higher the tier, the higher compression of the information. E.g. waveforms are included only in tier 1 – higher tiers contain only the extracted parameters like the pulse height (energy), baseline level etc. A detailed description of the analysis chain in GERDA can be found in [94], however, the most important aspects needed to understand the presented analysis are summarized below:

Tier 0 Raw data containing: germanium detector signals, LAr veto traces or MC-generated pulses is stored in tier 0. Its binary format depends on the source of the data e.g. FADC manufacturer or the type of the MC software. The class can in principle accommodate for every sampling frequency of the signals, but in the case of the GERDA data, the sampling period is 10 ns.

Tier 1 Tier 0 is converted to tier 1, which is a standardized format consisting of the custom C++ class (MGTEvent). Two versions of the germanium signals exists in tier 1. The first one is the low frequency (LF) trace, where every 4 samples of the original signal are summed up (sampling period is therefore also increased 4 times to 40 ns). This signal is used e.g. for the energy reconstruction with the gaussian filter. A second high frequency (HF) signal, is the fragment of the original waveform containing its rising edge at the original sampling period of 10 ns. The HF signal is of primary interest for the PSD analysis.

Tier 2 Output of the energy filter procedures, as well as information like the baseline value, the time trigger etc. are stored in tier 2. Tier 2 is obtained by processing the tier 1 data with GELATIO software [94]. Since no waveforms are stored in this tier, the file size of the data is significantly lower (by a factor of ~ 10).

Tier 3 Tier 3 is created by using the tier 2 binary data and "metadata" stored in the ASCII format. The metadata contains the information like calibration curves, cross-talk corrections etc. The energy information from this tier is used for the waveform normalization in the PSD procedure.

Tier 4 Most of the data from tier 3, like quality flags and energy is passed to the tier 4. It also includes the PSD and LAr veto flags. Compression factor of the file size for tiers 3 and 4, with regards to tier 2, is ~ 10 and ~ 20, respectively.

From the $0\nu\beta\beta$ analysis point of view, tier 4 contains all important information extracted from the data. However, if for some reason additional data is needed e.g. from tier 1 or tier 3, the tiers can be easily connected using the ROOT tree's friend mechanism. All tiers contain the same number of events and files, further more, they follow a simple naming scheme. Because of this, a single, unique "key" allows to easily identify files location. This is is in contrast with the Phase I data, where there was no 1:1 correspondence between tier files.

4.1 Analysis of Phase II calibration data

Run	BEGe				Semi-coaxial			
	DEP	FEP	CE	MCS	DEP	FEP	CE	MCS
53	636	365	1009	1173	875	1244	2428	3298
54	641	397	1076	1246	780	1175	2287	3142
55	561	336	953	1108	992	1494	2919	4039
56	1452	885	2444	2805	2809	4283	8263	11375
57	613	371	1031	1188	852	1278	2451	3415
58	1108	673	1834	2114	1447	2161	4159	5715
59	294	181	503	571	437	657	1247	1750
60	1263	766	2106	2402	2062	3051	5855	8029
61	620	375	1032	1199	752	1081	2178	2962
62	834	509	1402	1590	1158	1719	3271	4535
63	1002	629	1703	1953	1207	1735	3303	4602
64	306	186	514	580	393	571	1111	1548
Total:	9336	5678	15612	17934	13769	20453	39477	54415

Table 4.2:
Average number of events, from the calibration dataset, available for a training of the PSD methods. Four energy ranges are included:
DEP, FEP: centroids $\pm$ $^1/_2$FWHM,
CE: 2350 – 2375 keV,
MCS: 2350 – 2375 keV.
For a given dataset (either BEGe or semi-coaxial) events are summed up and divided by a number of detectors.

The GERDA detectors are subjected to frequent calibrations using 3 ^{228}Th radioactive sources. The calibrations are done $\approx$ weekly or when any instabilities are observed in the system. Beside the energy calibration, pulse shapes of events from the ^{228}Th spectrum can be used for a training of the PSD methods.

Training statistics. Tests of the TMVA classifiers for data from detectors in the vacuum cryostat conditions (see Chap. 3) have shown that a considerable statistics is needed for the training process to obtain satisfactory results. Tab. 4.2 contains available average[2] number of training events, for both BEGe and semi-coaxial datasets, in four energy ranges with relatively clean samples of either SSEs or MSEs. Since the single calibration is usually only a few hours long, the obtained training statistics is rather small (note: for some runs there are more than one calibration session). It is therefore needed to join together the respective calibrations for a given detector and use them in a single training process. However, this can be only done for the stable (in time) input data and its stability has to be investigated prior to the training.

Data stability. Changes of the measurement conditions can result in a variation of pulses' gain, shapes or both. They can be caused by: gain fluctuations (e.g. of the linear amplifier), high voltage (HV) variations, alternations in the operating point of

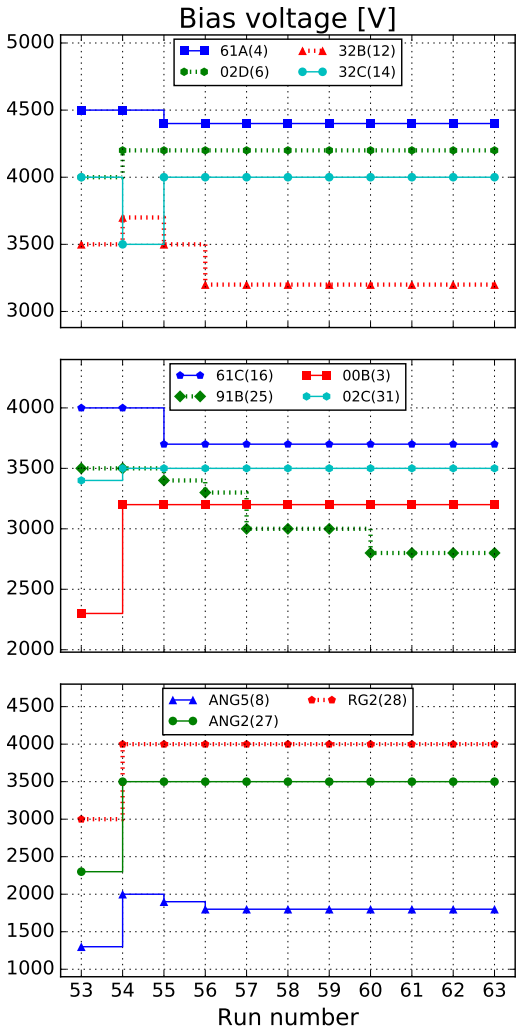
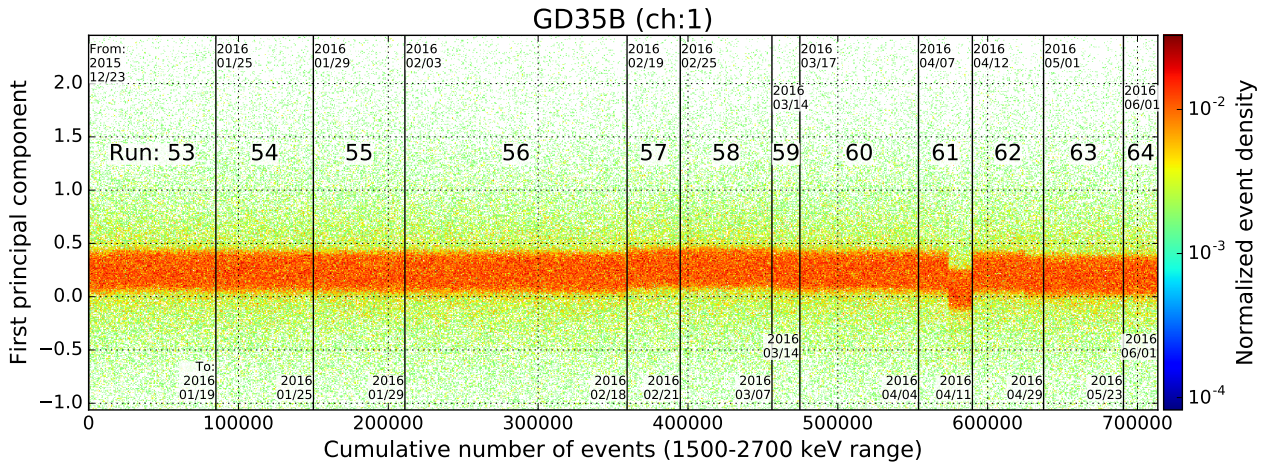
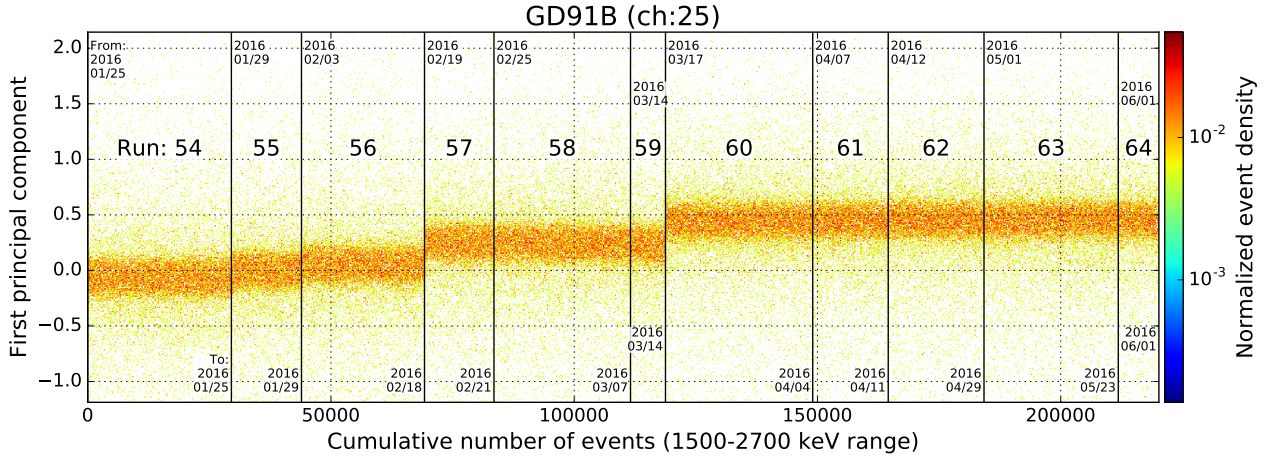


Fig. 4.1: Changes of bias voltages (HV) of the detectors in GERDA Phase II (stable channels were not plotted). HV adjustments are usually done to reduce the leakage current.

[2] Averaged over the number of detectors in the dataset.



(a) Short term instability



(b) Long term instability

Fig. 4.2: Distributions of the first PCA components in the calibration data for the BEGe detectors: a) GD35B (channel 1) and b) GD91B (channel 25), for the energy region 1500 – 2700 keV. The plots are normalized to take into account different counting rates due to e.g. source position changes during the calibrations – for each column the number of counts is equal to 1. For each run the start and stop dates are annotated in the left top and right bottom corners, respectively.
Top panel a) – a short term instability in GD35B, a sudden shift in the distribution can be seen for the second half of the run 61. Data affected by the shift was excluded from the training sample.
Bottom panel b) – long term shifts in the input data from GD91B, caused by the HV adjustment (see Fig. 4.1).

the JFETs (drain current and voltage adjustments) and so on. The contact impedances on the signal feedthroughs in the cryostat can also affect the observed waveforms.

The question is: how to monitor the changes in the pulse shape of the rising edge? Typically, the edge consists of about 100 samples and clearly a lower dimensional parameter is needed, which would contain this information. In this work, the first PCA component was used for the stability monitoring, since it is characterized with the largest variance – its distribution provides a good measure of an overall pulse shape. It is also used in the training process and its variation will immediately change the output classifier distribution.

After the analysis of the time dependence of a first PCA component, the analyzed detectors were divided into 3 classes:

- Stable channels
- Short-term instabilities
- Long-term shifts

In the case of the stable channels calibration data from all runs (53–64) was summed up together. The short term instabilities usually included only a single "faulty" calibration, which can be excluded from the training procedure. Because the *physics data* taken after this calibration can also be affected by the same problem, it has to be excluded from the classification. An example of a such instability is shown in Fig. 4.2a, where the first PCA component is plotted in time for the detector GD35B (channel 1). Run 61 was calibrated twice: first time on April 7^{th} and the second one on April 11^{th}. A clear shift can be seen in the distribution for the calibration taken on April 11^{th} (a second half of run 61 data), however the distribution is back to its original shape in run 62.

Fig. 4.2b shows a different situation – several shifts are present, which affect the data distribution for longer periods. They are caused by the detector bias voltage changes (Fig. 4.1), applied to counter the increases of the leakage current. For such cases, if the number of shifts was not greater than 3, the PSD analysis was made for each stable part separately. The number of sub-runs, to which the Phase II data was divided, is stated below for each detector:

- GD91A (0) - 3 parts,
- GD89B (5) - 2 parts ,
- GD91C (7) - 2 parts,
- GD32B (12) - 2 parts,
- GD32C (14) - 2 parts,
- GD89C (15) - 3 parts,
- GD61C (16) - 2 parts,
- GD76B (17) - 2 parts,
- GD35C (19) - 2 parts,
- GD79C (23) - 2 parts,
- GD35A (24) - 2 parts,
- GD91B (25)- 2 parts,
- GD91D (33) - 2 parts,
- GD32D (34) - 2 parts,
- GTF32 (38) - 2 parts,
- GTF45 (39) - 2 parts.

Since dividing into parts means that in principal the classifier distribution may be different for each period, if the same acceptance has to be obtained (e.g. 90% DEP), the cut value needs to also be calculated separately. However, the time dependent cut was not possible to implement in the Phase II software at the time of unblinding, therefore, the channels with the biggest differences with regard to cut value had to be excluded from the analysis. They were annotated with the red font in the above list. Other channels shown in the list also had the time dependent cut values, but their variation was so small that the average values could be used for all runs instead.

Excluded data. Due to the technical problems during the datataking, not all collected data was used in the final analysis. Detectors with unstable behavior, due to the increased leakage current or lack of test pulser, were either excluded entirely or put into the anti-coincidence (AC) only mode. Data from the channels in the AC-only mode has either energy value of 0 or 10 000 keV, well over the FADC dynamic range. Using the above procedure, data cannot be mistakenly taken into analysis, but on the other hand, it can be used for the multiplicity cut to reject non-$0\nu\beta\beta$ decay events. The decision to put a channel into AC-only mode is taken after initial analysis of the channel stability and is realized by changing the energy value on the tier 3 level. Since only selected detectors are put in the AC-only mode, the training statistics are in principle different for all channels – this can be observed in Fig. 4.2b, where the cumulative number of events for entire Phase II data is smaller for detector GD91B (bottom panel) than for GD35B (top panel).

4.1.1 Training of the PSD classifiers and efficiency calculations for the calibration data

TMVA classifiers. GERDA Phase II data was analyzed using two classifiers, available from the TMVA package: Projective Likelihood for coaxial detectors and MLP for BEGe detectors. The CE events of the 2614.5 keV peak (SSEs) were chosen as a "signal" sample (energy between 2350 and 2375 keV) and MCS (MSEs) events as a "background" sample (energy between 2450 and 2550 keV). The MLP classifier was not applied to the semi-coaxial detectors since the method TMlpANN, based on a similar training algorithm [90], was already used (see Sec. 2.2.2). It was

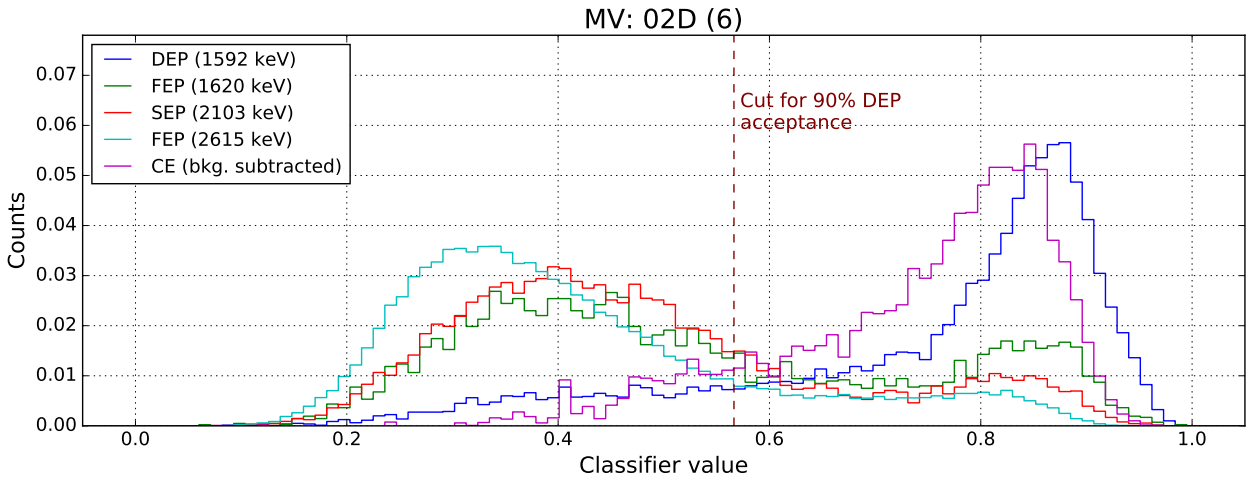


Fig. 4.3: Classifier distribution for events from the peaks in ^{228}Th spectrum and CE events (background subtracted) for BEGe detector GD02D (channel 6). A shift between the DEP and CE events distributions is due to the energy dependence of the SSE band (see Fig. 4.4).

therefore more beneficial to develop a second method, which uses the different algorithm (Projective Likelihood) and energy regions for training.

Dimensionality reduction. For comparison, two approaches to reduce the number of input variables were applied to the GERDA Phase II data: PCA transformation and variable summing. The PCA procedure was described in Sec. 3.1.4. The second approach (used in the Phase I analysis [45]) is to sum up the neighboring input variables extracted from the signal. It was also found that first six samples can be rejected entirely and the remaining variables are then:

- $\text{sum}_0 = \sum_{i=7}^{n=10} x_i$
- $\text{sum}_1 = \sum_{i=11}^{n=14} x_i$
- $\text{sum}_2 = \sum_{i=16}^{n=19} x_i$
- $\text{sum}_3 = \sum_{i=20}^{n=23} x_i$
- $\text{sum}_4 = \sum_{i=24}^{n=27} x_i$
- $\text{sum}_5 = \sum_{i=28}^{n=30} x_i$

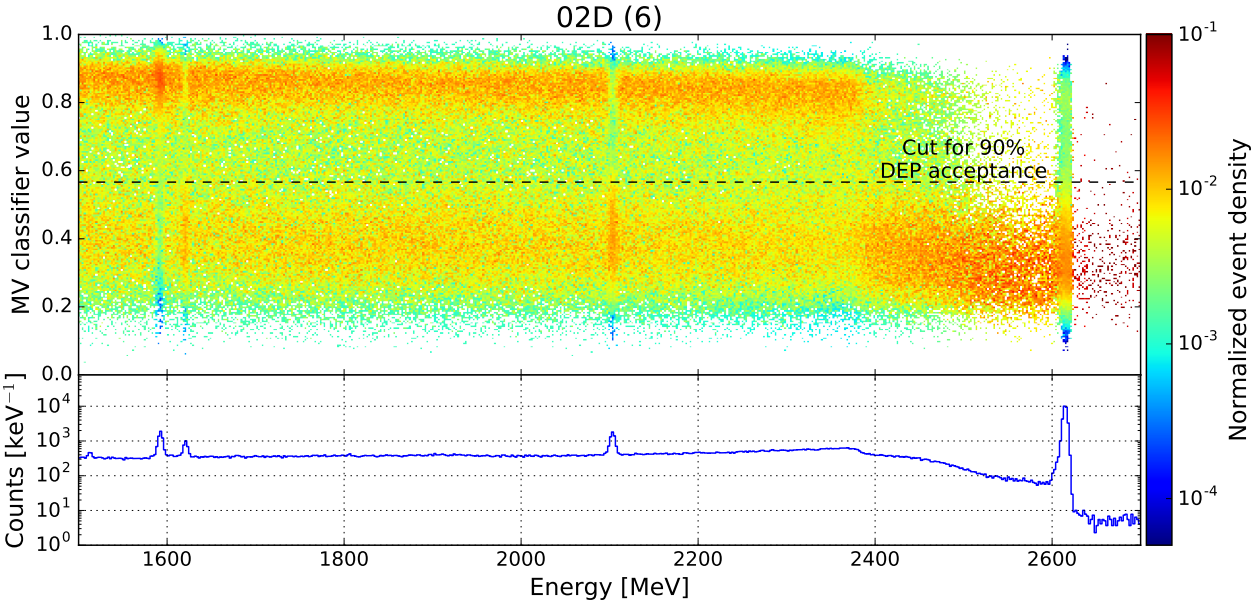


Fig. 4.4: Two dimensional, normalized histogram showing the classifier (MLP with the PCA dimensionality reduction) distribution as a function of energy for BEGe detector GD02D (channel 6).

Table 4.3: Summary of the resulting acceptance values for events from the ^{228}Th calibration data for all analyzed datasets. The Compton edge acceptance was calculated in the 2325 – 2375 keV energy region after subtracting MSE background taken from the 2400 – 2450 keV range. The table contains additional row for the MV classifier applied to the semi-coaxial detectors with the acceptance moved to 80% DEP to allow for a direct comparison with the MV2 classifier (also 80% DEP acceptance). Uncertainties are statistical only.

Dataset	DEP [%]	^{212}Bi [%]	$Q_{\beta\beta}$ [%]	SEP [%]	C. edge [%]	^{208}Tl [%]
BEGe (MV: 90% DEP)	89.7 ± 0.2	19.3 ± 0.4	45.2 ± 0.1	11.6 ± 0.2	79.6 ± 0.1	11.3 ± 0.1
Semi-coaxial (MV: 90% DEP)	89.8 ± 0.4	59.8 ± 0.3	74.0 ± 0.1	58.1 ± 0.2	88.8 ± 0.1	62.0 ± 0.1
Semi-coaxial (MV: 80% DEP)	81.2 ± 0.5	44.2 ± 0.3	61.3 ± 0.1	44.5 ± 0.2	80.1 ± 0.1	46.2 ± 0.1
Semi-coaxial (MV2: 80% DEP)	79.4 ± 0.5	39.2 ± 0.3	57.7 ± 0.1	53.9 ± 0.2	75.1 ± 0.1	43.8 ± 0.1

where x is the vector of 31 input samples (red points in Fig. 3.5). Every new variable is the sum of 4 original amplitudes (except for sum_5, which is the sum of the last 3 samples). Different dimensionality reduction schemes were applied to the semi-coaxial and BEGe detectors:

Semi-coaxial (MV): The first one has a designation **MV** in the tier 4 data. It used the PCA transformation (a reduction to 8 variables) and the acceptance is set to a 90% DEP survival probability.

Semi-coaxial (MV2): The second implemented classifier, called in short **MV2**, is trained on the same dataset as **MV**, but the difference is that it uses the summing method for the dimensionality reduction. Basically, it is the same classifier that was used in the Phase I PSD analysis. The acceptance is set to 80% DEP events survival probability.

BEGe: Data from BEGe detectors was reduced using the PCA transformation to 12 variables and classified using the MLP neural-network. The classifier is stored in tier 4 data as **MV**. The cut is set to 90% DEP events survival probability.

For both dimensionality reduction variants (MV and MV2) the semi-coaxial dataset classifier was Projective Likelihood.

Tab. 4.3 summarizes results of the analysis the both semi-coaxial and BEGe detectors. The cut values were calculated for the DEP acceptance of 90% – slight discrepancies are due to the exclusion of small fragments from the data after the integration of the PSD into tier 4. In the case of the MV2 classifier (Projective Likelihood + variables summing) the cut was set to the 80% DEP acceptance. Detailed results of the analysis (calculated for every detector separately) can be found in the appendix A – in this chapter, only the average (calculated for all events in the dataset) results are presented.

Results for the BEGe detectors. Acceptances achieved for the BEGe detectors are $(19.3 \pm 0.4)\%$ for the 1614.5 keV FEP and about 11.5% for the SEP/2614.5 keV FEP. The low acceptance of events from high energy peaks is due to the energy dependence in the classifier, which can be seen by comparing the acceptance of the CE events (bkg. subtracted), $(79.60.1)\%$, with the acceptance for the DEP, set to 90%. This fact makes the assertion of the acceptance for $0\nu\beta\beta$ region problematic.

Acceptance of the CE events is different for every detector, but for the detectors where the cut happens to lie in the middle of the bands, the CE acceptance is closer to the one for DEP (e.g.

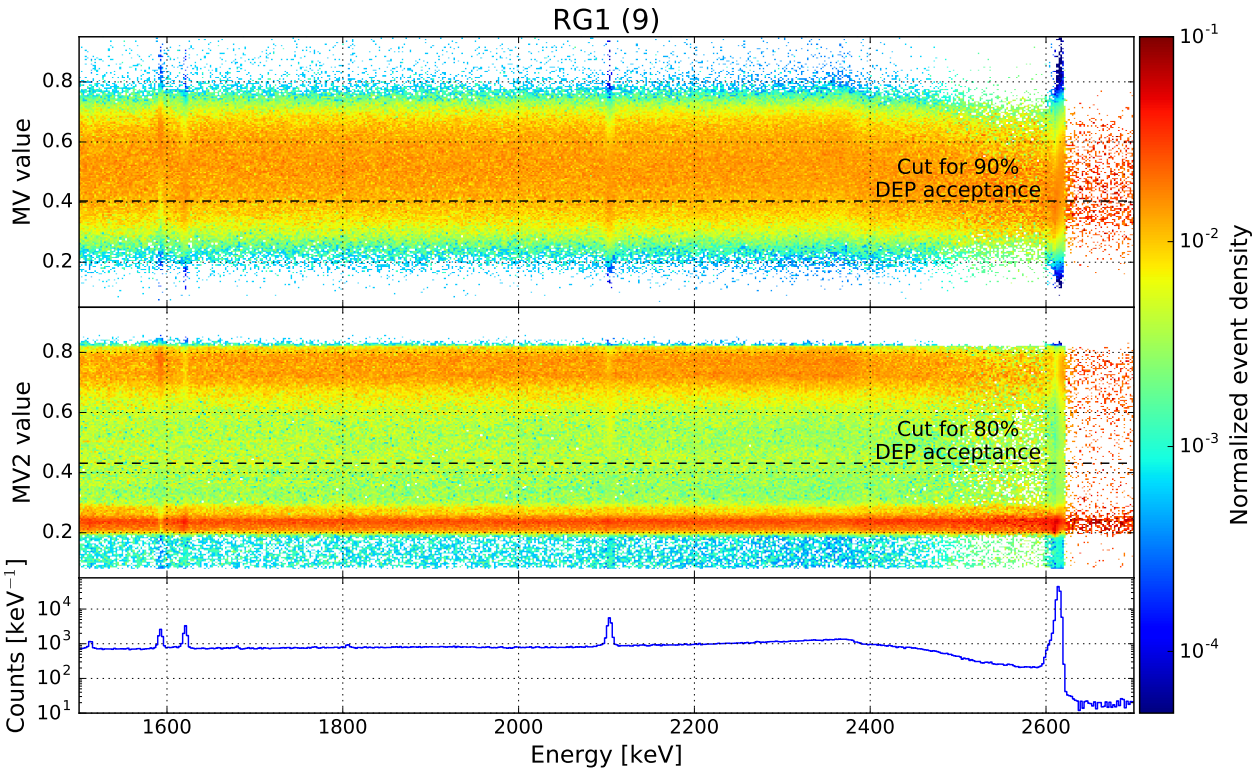


Fig. 4.5: Comparison of the MV (top) and MV2 (center panel) calibration data classifier distributions for the semi-coaxial detector RG1. The most distinctive difference is the existence of MSEs and SSEs bands in the case of the MV2 classifier.

detector GD02D, channel 6 – see the 2D histogram in Fig. 4.4) – in this case it is $(90.8 \pm 0.3)\%$. From the stability point of view, it would be therefore more beneficial to put the cut between the bands than to fix it to the 90% DEP acceptance. The effect of the PSD cut position on the events from FEP, SEP and DEP is shown in Fig. 4.6.

Results for the semi-coaxial detectors. Since semi-coaxial detectors were analyzed using different Projective Likelihood based classifiers, both of them were included in Tab. 4.3. Apart from the acceptance values, also the distribution shape of the classifier is vastly different for the two methods (see Fig. 4.5). In the case of MV2, two separated bands are visible – on the first sight it may look like the perfect separation exists between MSEs and SSEs. However, a closer inspection shows that a substantial fraction of the SSEs are classified as MSEs and vice versa (see e.g. MCS region or the SEP on the central panel of Fig. 4.5). This can be compared with e.g. classifier distribution for the MLP classifier for the BEGe detector data (Fig. 4.4). Therefore the existence of well defined bands alone does not necessarily mean high separation efficiency of the classifier.

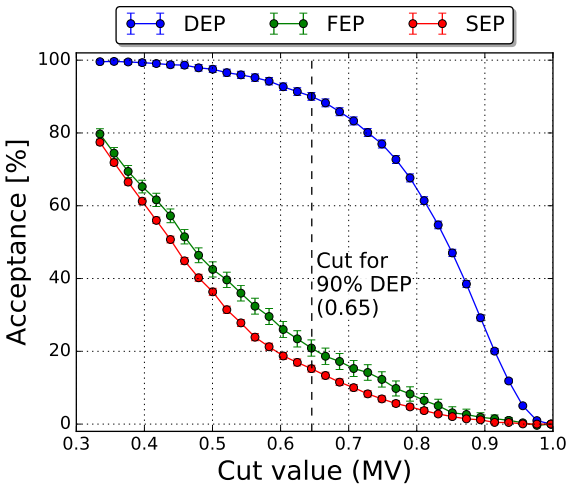


Fig. 4.6: Acceptance curves (as a function of the MV cut value) for the detector GD89A. Survival probabilities of the DEP, FEP and SEP were calculated for each cut value (filled circles on the plot – continuous lines are provided only for the visual reference).

Comparison between MV and MV2 classifiers. Since the MV and MV2 classifiers were implemented in tier4 with the different DEP acceptance levels (90 and 80%, respectively), to

compare their efficiencies Tab. 4.3 contains additional row with calculated acceptance values for the MV method with the DEP acceptance shifted to 80%. This way the direct it is possible to compare both analyses for semi-coaxial detectors. As it can be concluded from the data in the table, the efficiency in rejecting MSEs is slightly better for MV2 in the case of FEPs. SEP, on the other hand, is reduced more by the MV classifier. For MV2 SSEs from Compton edge, the acceptance is lower than in the DEP for about 4%, which is not the case for the MV, where the SSEs acceptance is almost the same for these two energy regions. The MSE background subtraction for Compton edge evens was performed by calculating the difference of classifier distributions for 2325 – 2375 keV and 2400 – 2450 keV energy ranges. Since the latter contains only MSEs, the resulting new distribution contains only SSEs.

To conclude, the actual efficiency on the ^{228}Th data is very similar for the MV and MV2 methods applied to the semi-coaxial detectors. Larger difference in the classifier behavior were observed after the application to the physics data, which is described in the next section.

4.2 Application of the PSD to the GERDA Phase II physics data

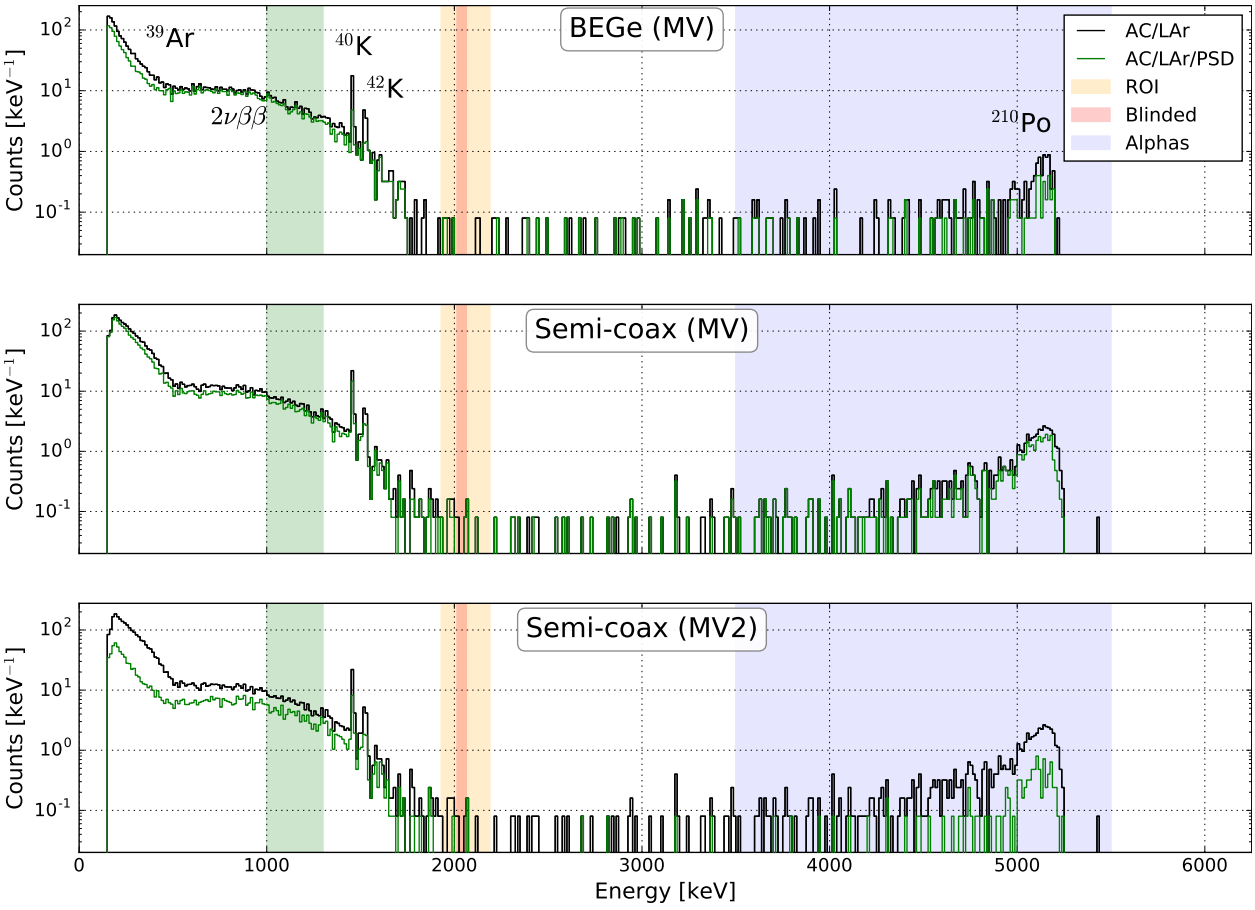


Fig. 4.7: Energy spectra showing unblinded physics data from GERDA Phase II (runs 53–64) (only the data for which the classifiers were applicable). Events after the muon veto and anti-coincidence (AC) cuts are plotted in black, events left after PSD cuts are plotted in green. Known background components (like e.g. ^{39}Ar) were annotated on the top panel. Energy regions used for the PSD efficiency calculations (Tab. 4.4) were marked with the color strips (see legend for details).

PSD metadata ("weights"), obtained by training the methods on the calibration data, was later applied to classify the waveforms from the "physics" dataset. The energy spectra of GERDA

Table 4.4: Summary of the resulting acceptances for the GERDA physics data for all analyzed datasets. The $2\nu\beta\beta$ acceptance was calculated in the 1000 – 1300 keV energy range for all events in this region, as well as ones prefiltered the with LAr veto flag, which should be almost exclusive SSEs. Events in the BW/ROI are tallied after applying the LAr veto. The table contains additional row for the MV classifier, applied to the semi-coaxial detectors with the DEP acceptance moved to 80%, to allow for a direct comparison with the MV2 classifier.

Dataset	$2\nu\beta\beta$ [%]	$2\nu\beta\beta$ LAr [%]	BW ($\pm$200 keV)	BW (190 keV)	Unblinded $Q_{\beta\beta}$ [$\pm$25 keV]	α (3.5–5.5 MeV)
BEGe (MV: 90% DEP)	75.5 ± 0.8	87.4 ± 0.8	4/13	3/7	0/0	76/184
Semi-coaxial (MV: 90% DEP)	79.3 ± 0.8	83.2 ± 0.9	14/24	6/9	2/3	462/668
Semi-coaxial (MV: 80% DEP)	66.5 ± 0.9	70.2 ± 1.1	11/24	4/9	2/3	375/668
Semi-coaxial (MV2: 80% DEP)	58.0 ± 0.9	62.1 ± 1.1	7/24	2/9	1/3	118/668

Phase II events are shown in Fig. 4.7. All the facts concerning the $0\nu\beta\beta$ analysis are included in the GERDA internal report [107], most important of them are also mentioned here. The calendar timespan of the gathered dataset is 159.29 days (25/12/2015 – 1/06/2016). Due to the breaks for the calibrations and rejected data of bad quality, the obtained "duty cycle" is at 82.0%.

Tab. 4.4 contains the summary of the PSD methods efficiency, applied to the physics data. Two background window (BW) definitions are included in the table: the first one is simple $Q_{\beta\beta} \pm 200$ keV, while the second one in used in the BI calculation and excludes the possible peaks in this energy region (see Fig. 4.8) and the blinded energy range (ROI) around the $Q_{\beta\beta}$. It will referred to as 190 keV BW. Both definitions were used during the development of the PSD methods, also by other groups working on this topic, and they were therefore included in this work for the comparison purposes.

BEGe dataset. The PSD efficiency for the BEGe dataset is summerized in a first row of Tab. 4.4. The detailed detector-wise table can be found in App. A. Acceptance of $2\nu\beta\beta$ events, prefiltered with the LAr veto to reject possible MSEs, is very close to the set DEP acceptance of 90% (Tab. 4.3). The MV classifier is not particularly effective in rejecting the alpha events (energy range 3.5 – 5.5 MeV) – more than half of them survive the cut. This probably due to the fact that the signals originating from α emitters, with the energy deposited in the groove or p^+ contact,

Dataset	All events	Events left after:		
		LAr	PSD	PSD && LAr
BEGe (MV: MLP/PCA)	22	7	5	3
Semi-coaxial (MV: PL/PCA)	19	12	10	8
Semi-coaxial (MV2: PL/Summ.)	19	12	6	3

Table 4.5: Summary of the events in 190 keV ROI (see Fig. 4.8), including the blinded region and taking into account the effect of LAr and PSD cuts. Columns show respectively: the number of events not affected by the LAr cut, PSD cut or ones not vetoed by neither of them. Cut of the MV classifier for the BEGe and semi-coaxial datasets was set to 90% DEP acceptance and its MV2 counterpart to 80%. PL stands for Projective Likelihood.

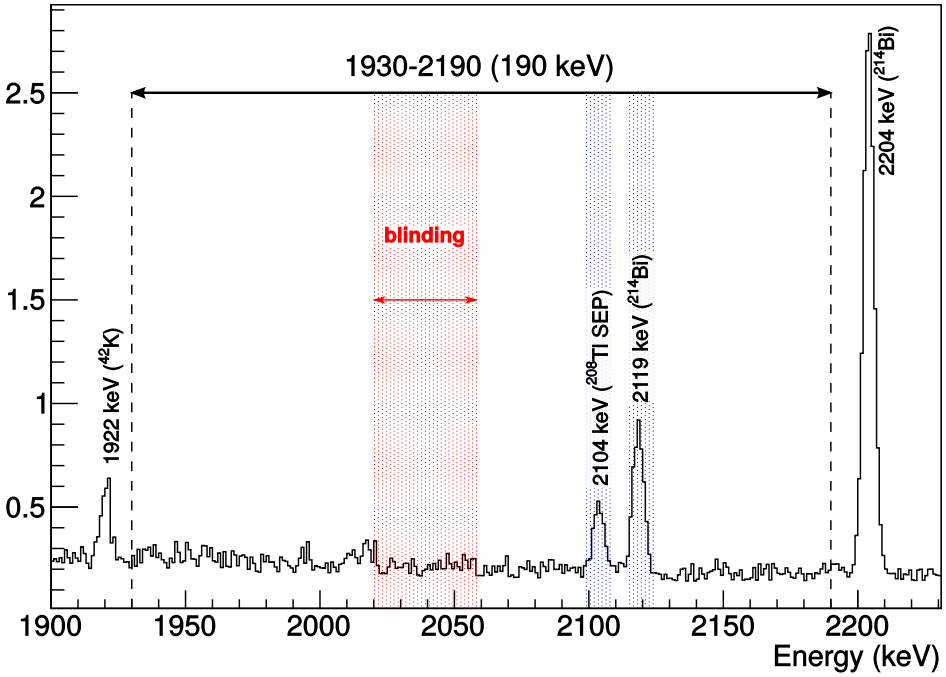


Fig. 4.8: Background window (190 keV wide) shown on the energy spectrum from the GERDA background model. Blinded region around $Q_{\beta\beta}$ (±25 keV) is shown in red, blue regions with the known background peaks (2104 keV from ^{208}Tl SEP and 2119 keV from ^{212}Bi FEP) ±5 keV are excluded from the background index calculation. Source: GSTR-13-004.

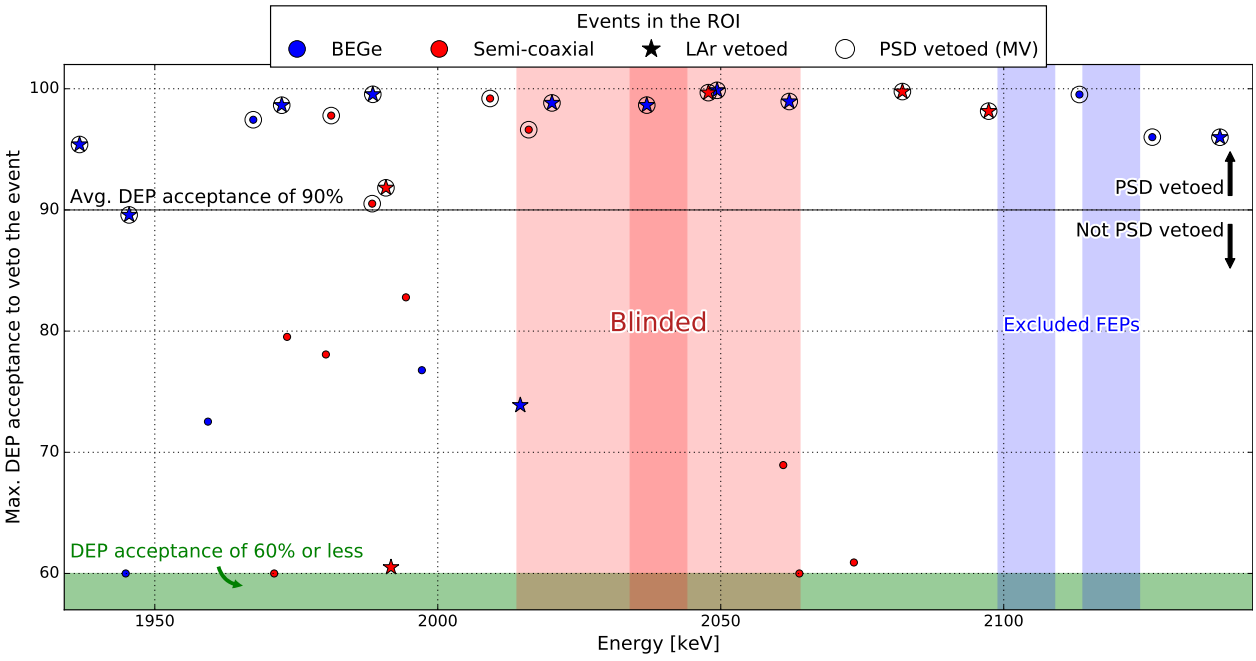


Fig. 4.9: Visualization of the BW/ROI events for both unblinded semi-coaxial (red dots) and BEGe (blue dots) datasets. Events vetoed by the MV PSD are marked with circles and the ones vetoed with LAr are marked with a star symbol. Classifier values were transformed into the DEP acceptance levels (see text for the explanation). Lower values of the DEP acceptance were limited to 60% for the readability purposes (only 3 events were affected by this, with the DEP acceptances of 36%, 38% and 0%). Light-red band shows the blinded region ($Q_{\beta\beta} \pm 25$ keV) and darker red strip inside it denotes signal region ($Q_{\beta\beta} \pm 5$ keV).

have a very fast rising edge [76]. Therefore, they may not resemble the signals from the calibration dataset, for which the PCA matrix was calculated.

Events from the BW/ROI are plotted in Fig. 4.9, for both semi-coaxial and BEGe datasets. For the BEGe dataset, three events are left in the 190 keV wide BW after applying both PSD and LAr veto cuts. None of them is from the blinded $Q_{\beta\beta}$ ±25 keV part of the spectrum. Using the cut curve, (like the one in Fig. 4.6, i.e. the dependence of DEP acceptance on the cut value) it is possible to transform the classifier value of the event to the DEP acceptance value, for which the event is vetoed. Translating the classifier value into the DEP acceptance allows for the direct comparison of events from all detectors, as can be seen in Fig. 4.9.

To give an example: the third non-vetoed event from the BEGe dataset (detector GD89A, 1997.2 keV, May 4$^{\text{th}}$ 2016) has the classifier value of 0.75, while the cut threshold for the target DEP acceptance of 90% is 0.65. Therefore, it survives the PSD cut for that acceptance level. However, lowering the acceptance to 77% would result in the cut threshold of 0.755 and the event would be rejected. Applying the following procedure to all non-vetoed events, it can be concluded that lowering the DEP acceptance to 73% would veto all events but one, which is classified with the unusually high classifier value of 0.985. This event cannot be rejected using a low edge cut with the reasonable signal efficiency, since it lies higher than the SSE band. Two sided cut would be needed in this case, but it was not introduced since it would require a separate study of its effects on the overall PSD performance.

Semi-coaxial dataset. For the semi-coaxial dataset the summary of the PSD performance on the physics data is collected in Tab. 4.4. Events from the BW/ROI are plotted in Fig. 4.9 as red dots. In the case of the MV classifier 8 events survive both LAr and PSD cuts. Large differences can be noticed in Tab. 4.4 between MV and MV2 classifiers, especially regarding the number of events surviving in the BW region, as well as $2\nu\beta\beta$ events efficiency prefiltered with the LAr veto. Even then the cuts for both MV and MV2 are fixed at 80% DEP acceptance, the MV2 efficiency for almost pure sample of SSEs from the $2\nu\beta\beta$ region is still $\approx 8\%$ lower. On the other hand, MV2 is much more effective in vetoing both BW/ROI events and high energy surface events, originating from the α emitters (^{210}Po and ^{226}Ra [44]). The MV2 classifier vetoes all but 118 out of 668 α events (3.5 – 5.5 MeV energy range after the LAr veto), which translates to $\approx 17\%$ survival probability.

Quite large discrepancies exist between the DEP and the $2\nu\beta\beta$ (after the LAr veto) events acceptances in the semi-coaxial dataset. For the target DEP events acceptance of 90%, the MV classifier accepts $(83.2 \pm 0.9)\%$ of the LAr prefiltered $2\nu\beta\beta$ events. This is an expected value when compared to the GERDA Phase I efficiency for ANN with background model corrections [108]. After lowering the DEP acceptance to 80%, the acceptance of the $2\nu\beta\beta$ events (after LAr veto) is at 70% level. It can be therefore said that in general is about 10% lower than the target DEP acceptance. The question remains: how to estimate the PSD efficiency for the $0\nu\beta\beta$ decay? DEP acceptance is not a good proxy, since it is affected by the volumetric effects - DEP events are mostly from the detector corners and surfaces, because of the lower probability of annihilation γ-rays absorption. The $2\nu\beta\beta$ decay events have the same events topology as their $0\nu\beta\beta$ counterparts, but the difference in energy (≈ 1 MeV) amplifies possible energy dependent effects, as well as possible lowering of the efficiency due to the lower signal to noise ratio. This is especially true in Phase II, where the copper shrouds were removed to allow for the LAr scintillation light detection and they do not longer shield the front-end electronics from the electromagnetic interference. To independently estimate the $0\nu\beta\beta$ efficiency the MC simulations are therefore needed. The detailed description of this approach is presented in the next section.

4.3 Evaluation of the signal efficiency for the semi-coaxial detectors with simulated pulses

Even thought the calibration data provides training samples for the PSD methods, the evaluation of the $0\nu\beta\beta$ signal efficiency is not straightforward – the only energy regions rich with SSEs are DEP and CE. DEP events are localized mostly on the detector edge (especially corners) due to the increased probability of the escape of 511 keV γ-rays, the same volumetric effect is true for the CE events. This means that the PSD acceptance for the SSEs near the core of the detector is still uncertain if the cut is set based on the DEP survival probability. Since the $0\nu\beta\beta$ decay events are distributed homogeneously in the detector, a sample with similar distribution would be needed to calculate the signal efficiency. The only candidate could be the $2\nu\beta\beta$ decay region (e.g. 1.0 – 1.3 MeV), preferably in anticoincidence with the LAr veto to exclude background events. However, since the $2\nu\beta\beta$ decay events in this region have roughly half of the $Q_{\beta\beta}$ energy, the another uncertainty is introduced due to the lower signal to noise ratio in the analyzed pulses. It is therefore necessary to test the PSD methods on the pulses which are characterized with the same event topology (point-like energy deposition), same energy (pulse height) and noise levels. This can be achieved only with the MC simulation of the pulses and similar approach was used in Phase I of the GERDA experiment [45].

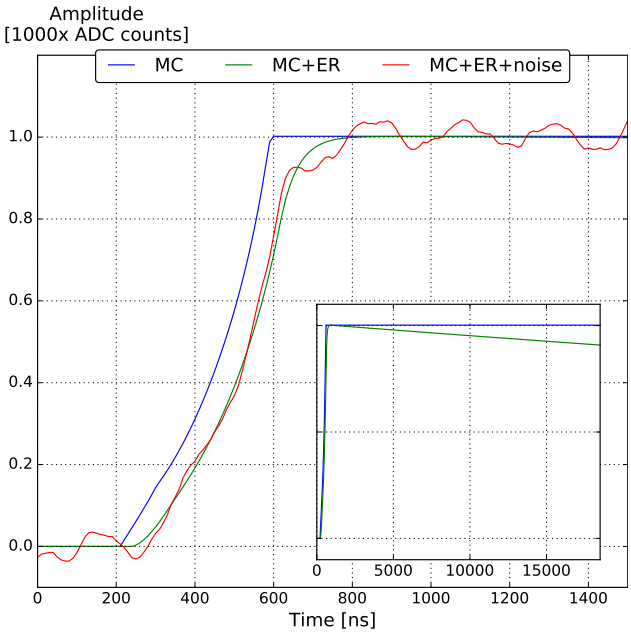


Fig. 4.10: Waveforms show a 500 keV simulated $2\nu\beta\beta$ decay event from the detector ANG3. The plot compares: a raw MC simulated charge pulse, an MC pulse with the applied electronics response (ER) and an MC pulse with applied ER and superimposed noise. The noise was taken from the real recorded baseline waveform (in an absence of a signal). The inset shows the same pulses (except for the noisy pulse for a visual clarity), but in a wider time scale – the effect of the decaying exponential tail was also taken into account in the ER.

It should be pointed out that the BEGe detectors are much less affected by this problem, because the drifting charge induces most of the signal when it is near a readout electrode. The biggest difference between pulses is due to the drift time in the low weighting potential zone and results in the longer pedestal of the current pulse. Therefore, the uncertainty of the signal efficiency assessed on the DEP events in the A/E method for the Phase II data is only 2.4% [56]. Signals from the semi-coaxial detectors are much more position dependent and the discrepancy between signal efficiency and DEP acceptance is expected to be more severe.

The MC dataset used for the signal efficiency evaluation was the same as the one used to test other PSD method, namely the ANN_mse neural-networks. Pulses were simulated by the MPIK Heidelberg group and available to the author for the reprocessing. The general principles of the simulation process were also included here, more detailed information can be found in [109].

Energy deposition simulation. The first step is to simulate the energy deposition locations by the γ-rays from the ^{228}Th source (MAGE [110]), as well as the $2\nu\beta\beta$ and $0\nu\beta\beta$ decays (DECAY0 package [111]). MAGE is a GEANT4-based MC software developed by the GERDA

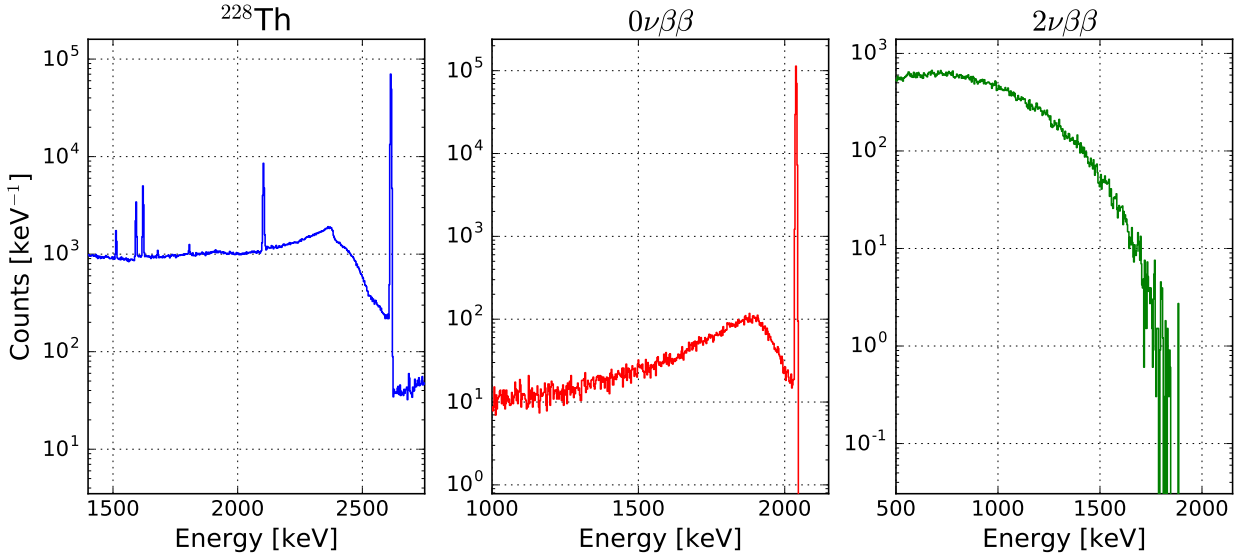


Fig. 4.11: Energy spectra of the MC datasets: ^{228}Th (left panel – events used for the PSD training), $0\nu\beta\beta$ (central panel) and $2\nu\beta\beta$ (right panel) decays, as simulated for the ANG5 detector. The thorium dataset was obtained by simulating the detector array irradiation with a point-like calibration source (source position was adjusted to match the Phase II energy spectra). The $0\nu\beta\beta$ and $2\nu\beta\beta$ decays were simulated homogeneously in the detector's volume. Apart from the $0\nu\beta\beta$ peak (central panel), lower energy events are also visible – they are originating at the edges of the detector and incomplete energy deposition results from escaping electrons and/or Bremmstrahlung radiation.

and MAJORANA DEMONSTRATOR collaborations, which includes detailed geometries of both experiments, as well as cross-sections for all relevant interaction types of $\alpha/\beta/\gamma$ radiation. Because the entire detector array is simulated in the same timee, the effects like e.g. the anti-coincidence cut can be taken into account in the MC dataset. Almost any radioactive source can be simulated with the proper treatment of the decay chain branching, coincidence summing effects etc. On the other hand, DECAY0 package was written to precisely simulate the $2\nu\beta\beta$ and $0\nu\beta\beta$ decays with all associated effects like e.g. angular and energy correlations between the electrons.

Waveforms simulation. Detailed description of the signal formation process in the germanium detector was described earlier in Sec. 2.1. ADL 3.0 software [62] was adopted in GERDA for MAGE output for that purpose [112]. ADL solves the Poisson (Laplace) equations numerically for the given detector geometry and impurity concentration to obtain electrical (weighting) fields. The fields are then used to simulate the drift of the charge carriers (electrons/holes) and an induced current signal. In the case of the described MC dataset, homogenous impurity concentration (without the geometrical gradient) is assumed in the simulation. Since the exact impurity concentrations in the crystals are unknown, the concentrations were deducted from the depletion voltages, supplied by the manufacturer in the detector datasheets. For a given geometry, a series of simulations can be performed with the impurity concentration $|N_A - N_D|$ and bias voltages as parameters – a correct value of $|N_A - N_D|$ is found when the minimal bias voltage to achieve depletion matches the voltage from the datasheet[3]. Simulation with ADL results in the library of SSE pulses for all possible locations in the detector. Waveforms for MSEs are a superposition of SSE pulses from all interaction sites, weighted with the energy deposited in the given site.

[3]The relation between the impurity concentration N_A and the depletion voltage V_{dep} for the simple coaxial and spherical geometries can be found in Sec. 2.1.2

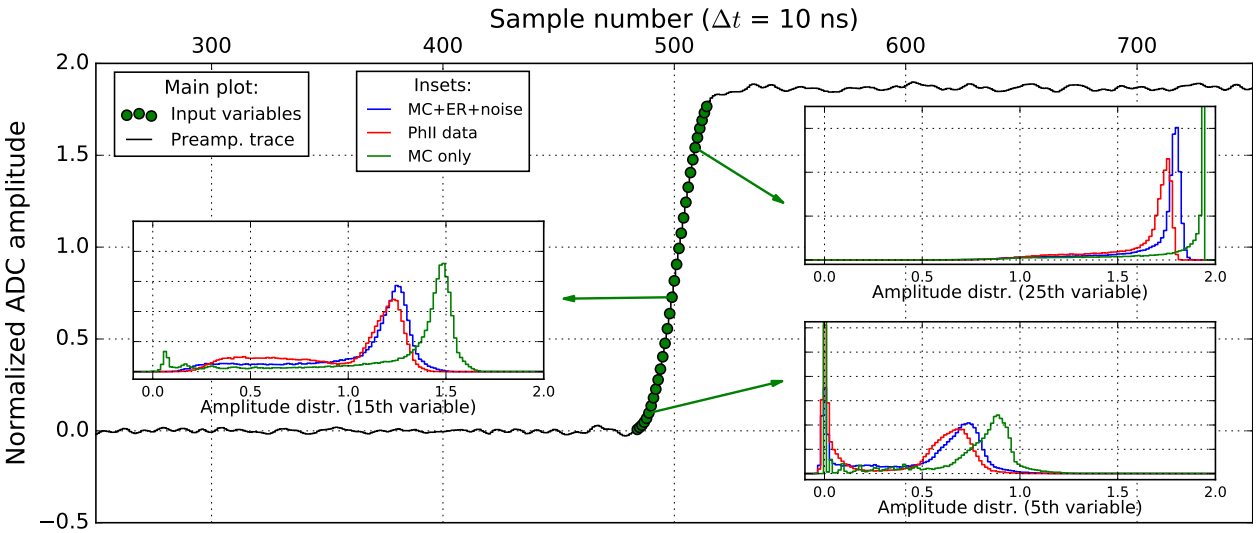


Fig. 4.12: Input variables (green points) extracted from an example pulses. Every inset shows distributions of the given variable for events from the Compton edge region (2300 – 2375 keV) for the detector Amplitudes number 5, 15 and 25 were selected, as representative for the beginning, middle and upper part of the rising edge. Distributions are shown for the 3 datasets: as obtained from ADL (solid green line - "MC only"), after application of the electronics response (ER) and superimposing the noise (solid blue line - "MC+ER+noise") and from the GERDA Phase II data (solid red line - "PhII data"). Parameters of the ER function were selected to obtain best agreement with the Phase II data, especially in the middle part of the pulse (15^{th} variable), which corresponds to the maximal current pulse amplitude.

Electronics response. Before the noise addition, simulated pulses have to be modified by applying the electronic response of the readout chain. A preamplifier has a finite bandwidth, which affects the speed of a rising edge of the pulse. It also includes a resistive feedback technique to discharge a feedback capacitor C_f, therefore the preamplifier pulse is characterized with the exponentially decaying tail. Both effects were taken into account by calculating a transfer function $T(s)$ of a generic preamplifier model, consisting of the $C_f R_f$ feedback loop and an ideal operational amplifier (behaving like a one pole filter with an amplification). An impulse response $h(t)$ is then calculated by applying an inverse Laplace transform of the transfer function:

$$h(t) = \mathscr{L}^{-1}\{T(s)\} \tag{4.1}$$

The response is then applied to the simulated signal by a numerical convolution[4]:

$$v_{\text{out}}(t) = v_{\text{MC}}(t) * h'(t). \tag{4.2}$$

Since the MC signal is sampled with a 10 ns period (same as in the data), the convolution is done in the discrete domain:

$$v_{out}[i] = \sum_{j=0}^{N} v_{MC}[i] - h'[i-j], \tag{4.3}$$

where N denotes length of the signal (number of samples). The functions are the discretized versions of the ones in Eq. 4.2.

A more detailed description and a derivation of the electronics model can be found in App. B. It should be mentioned that the presented electronics response is different than the one used by MPIK Heidelberg group, which is realized by applying 3 times a 70 – 80 ns MWA filter (the value

[4]To account for the fact that ADL provides a charge pulse (an integrated current pulse), while the transfer function $T(s)$ is calculated for the current pulse, the derivative of the impulse response is used in the convolution.

Table 4.6: Summary of survival probabilities for the MV classifier, applied to the MC dataset and GERDA Phase II (runs 53–64) data. Training procedure was performed separately for each dataset. Good agreement (e.g. between $2\nu\beta\beta$ efficiencies) is observed only for some detectors (namely ANG5, ANG3 – under 5.3% for all regions), while for other the differences can be as high as 8.6% (FEP for RG1). Δ columns show the differences in acceptances between MC and data. The differences were colored for readability: green – $|\Delta| < 2.5\%$, orange – $2.5\% < |\Delta| < 7.5\%$, red – $|\Delta| > 7.5\%$.

	MV														
Detector	DEP MC [%]	DEP [%]	$0\nu\beta\beta$ MC [%]	$2\nu\beta\beta$ MC [%]	$2\nu\beta\beta$ LAr [%]	Δ	FEP MC [%]	FEP [%]	Δ	SEP MC [%]	SEP [%]	Δ	$Q_{\beta\beta}$ MC [%]	$Q_{\beta\beta}$ [%]	Δ
8 (ANG5)	89.7	90.3	83.9	81.3	79.8	1.5	56.2	53.7	2.5	55.0	49.7	5.3	73.1	68.3	4.8
9 (RG1)	90.0	90.0	82.1	81.5	85.0	-3.5	58.9	67.5	-8.6	58.7	65.8	-7.0	74.5	77.8	-3.3
10 (ANG3)	90.0	89.9	81.9	80.0	79.8	0.2	58.2	58.5	-0.3	57.4	56.9	0.6	74.0	72.6	1.4
27 (ANG2)	90.3	89.5	81.0	76.7	81.5	-4.8	51.3	58.5	-7.2	53.9	58.8	-4.9	70.5	74.3	-3.8
28 (RG2)	89.9	90.3	79.6	77.1	79.8	-2.8	60.0	55.9	4.1	60.2	53.4	6.8	76.5	72.0	4.5
29 (ANG4)	90.0	90.3	82.0	78.1	83.1	-5.0	58.8	61.9	-3.1	58.2	59.1	-0.9	74.6	75.7	-1.0
36 (ANG1)	90.5	90.0	80.0	80.3	86.8	-6.5	65.8	67.5	-1.7	66.8	68.4	-1.6	79.1	80.2	-1.1

Mean active mass weighted $0\nu\beta\beta$ efficiency: $(81.7 \pm 2.0)\%$

depends on a given channel). While this method includes the effect of a limited bandwidth of the preamplifier, it does not take into account the exponential decay of the pulse. This means that the last input variables, extracted from a pulse, can have a higher mean value than in the data and therefore it could affect the PSD efficiency . It is also true that the severity of the effect depends on a method of input variables extraction and other methods may be less sensitive to it.

Presented results regard only the MC dataset – the training, classification and survival probability calculation was done only on the simulated data. In principle, one could use the TMVA weights trained on the real data and classify the $0\nu\beta\beta$ decay simulated events to obtain the signal efficiency. But since the real impurity concentrations in the detectors, as well as detailed electronic response, are not known, it would introduce further systematic uncertainties.

Training and classification sets. Events used for the PSD training on the MC dataset were selected in a same manner as for the data (Sec. 4.1.1). Generally, a larger number of events was available in the MC dataset that in the GERDA Phase II data. Therefore, the training statistics in the MC dataset were normalized with respect to the real data, i.e. the same number of events was used for training in both cases. However, all available events were used for the classification process to reduce statistical uncertainties in survival probability determination.

Results of the analysis are collected in Tab. 4.6 and Tab. 4.7 for the MV and MV2 classifiers, respectively. The tables contain comparison of survival probabilities for the GERDA Phase II and MC datasets. The cut values of the MV (MV2) classifier for MC and Phase II data were set to match 90% (80%) DEP acceptance. Acceptance values were calculated for the events from peaks (DEP, FEP and SEP), as well as the Compton continuum region around the $Q_{\beta\beta}$ energy from the ^{228}Th spectrum. The $0\nu\beta\beta$ decay efficiency was calculated as the reduction factor of simulated $0\nu\beta\beta$ decay events in the range $Q_{\beta\beta} \pm \mathrm{FWHM}(Q_{\beta\beta})$. $\mathrm{FWHM}(Q_{\beta\beta})$ was calculated separately for each detector using a resolution curve obtained from the Phase II calibration data (^{228}Th spectra). The $2\nu\beta\beta$ decay efficiency was derived from the survival fraction of the events in the 1000 – 1300 keV energy range. It can be directly compared with the Phase II results, since after the application of LAr veto to the physics data, in the above mentioned region, it is almost only the $2\nu\beta\beta$ decay events [9], with virtually no background.

Table 4.7: Summary of survival probabilities for the MV2 classifier, applied to both MC dataset and GERDA Phase II (runs 53–64) data. Training procedure was performed separately for each dataset. Much greater differences can be observed between MC and Phase II data than for the MV classifier (Tab. 4.6), e.g. in the $0\nu\beta\beta$ events survival probability. Other energy regions are also affected by the discrepancies. Good agreement with the data (under 5%) can be seen only for the detector ANG3 (channel 10). Δ columns shows the differences in acceptances between MC and data. The differences were colored for readability: green – $|\Delta| < 2.5\%$, orange – $2.5\% < |\Delta| < 7.5\%$, red – $|\Delta| > 7.5\%$.

	MV2														
Detector	DEP MC [%]	DEP [%]	$0\nu\beta\beta$ MC [%]	$2\nu\beta\beta$ MC [%]	$2\nu\beta\beta$ LAr [%]	Δ	FEP MC [%]	FEP [%]	Δ	SEP MC [%]	SEP [%]	Δ	$Q_{\beta\beta}$ MC [%]	$Q_{\beta\beta}$ [%]	Δ
8 (ANG5)	80.0	80.2	63.8	60.1	63.9	-3.8	40.1	37.7	2.3	49.2	35.9	13.2	60.0	54.6	5.4
9 (RG1)	80.0	79.9	63.0	59.1	73.6	-14.5	38.9	50.2	-11.3	49.6	49.9	-0.3	59.7	64.7	-5.0
10 (ANG3)	79.9	80.0	64.5	61.3	66.0	-4.8	41.1	39.7	1.5	50.2	41.0	9.3	60.6	57.5	3.1
27 (ANG2)	80.1	80.1	59.7	56.6	65.0	-8.4	35.7	39.4	-3.7	44.2	42.2	2.0	56.6	58.8	-2.3
28 (RG2)	79.9	80.1	78.3	73.6	66.0	7.6	34.8	37.3	-2.5	46.9	37.3	9.6	58.0	56.3	1.7
29 (ANG4)	80.1	80.1	66.9	60.7	66.5	-5.8	41.6	39.5	2.1	52.3	39.3	13.0	62.1	57.4	4.7
36 (ANG1)	80.4	80.5	64.7	65.5	75.3	-9.8	50.6	51.1	-0.5	56.8	52.4	4.4	66.1	66.8	-0.7

Mean active mass weighted $0\nu\beta\beta$ efficiency: $(65.5 \pm 1.6)\%$

The MV classifier (Tab. 4.6) trained on the MC dataset shows generally a better agreement than MV2 (Tab. 4.7). For the $0\nu\beta\beta$ decay efficiency, almost all channels show rather small discrepancy of 5% or under, with the exception of channel 36 (ANG1) yielding 6.5%. Other energy regions are characterized with higher discrepancies discrepancies are also present – e.g. the difference of the FEP acceptance for RG1 equals to 8.6%. Channels 8 (ANG5) and 29 (ANG4) show the overall best agreement, with all acceptance differences under or equal to 5.3% and 6.5%, respectively. The total $0\nu\beta\beta$ decay efficiency is equal to 81.7% and was calculated as the average weighed over the active masses of the detectors [113]. Therefore, a difference of 8.3% between DEP and $0\nu\beta\beta$ decay acceptances is observed.

Much bigger discrepancies between the real data and the simulation can be seen for the MV2 classifier (Tab. 4.7), with the cut set to DEP survival probability of 80%. Differences for the $0\nu\beta\beta$ acceptances are as large as 14.5% (channel 9, detector RG1). MC results for RG1 in particular show very poor consistency with the data – large discrepancies are also observed for the FEP and the $Q_{\beta\beta}$ regions. All differences have a negative sign. This does not not necessarily mean that the PSD efficiency is compromised, since the efficiency is lower for both "signal" events (SSEs from $2\nu\beta\beta$ decay) and "background events" (MSEs from FEP). Therefore the effect can be understood as a shift of the cut to reject both type of events, resulting in lower acceptances. The mean $0\nu\beta\beta$ decay events PSD efficiency, weighted over detector's active masses, equals to $(65.5 \pm 1.6)\%$, therefore the deviation from the DEP acceptance of 80% is even larger than for the MV classifier (14.5%).

Volumetric effect in the PSD performance. The discrepancy between the DEP and the $0\nu\beta\beta$ decay acceptances could have several possible causes. First of all, there is a well known issue of the volumetric dependence for the PSD for semi-coaxial detectors [114, 115]. The DEP and $0\nu\beta\beta$ decay events have in principle the same single-site signature. MC simulation allows not only for the signal ($0\nu\beta\beta$) efficiency determination, but also for the studies of the volumetric effects of a given classifier. This was done for example in [115] for the assessment of the spacial effects for GERDA Phase I neural-network based classifier TMlpANN, same tests were performed

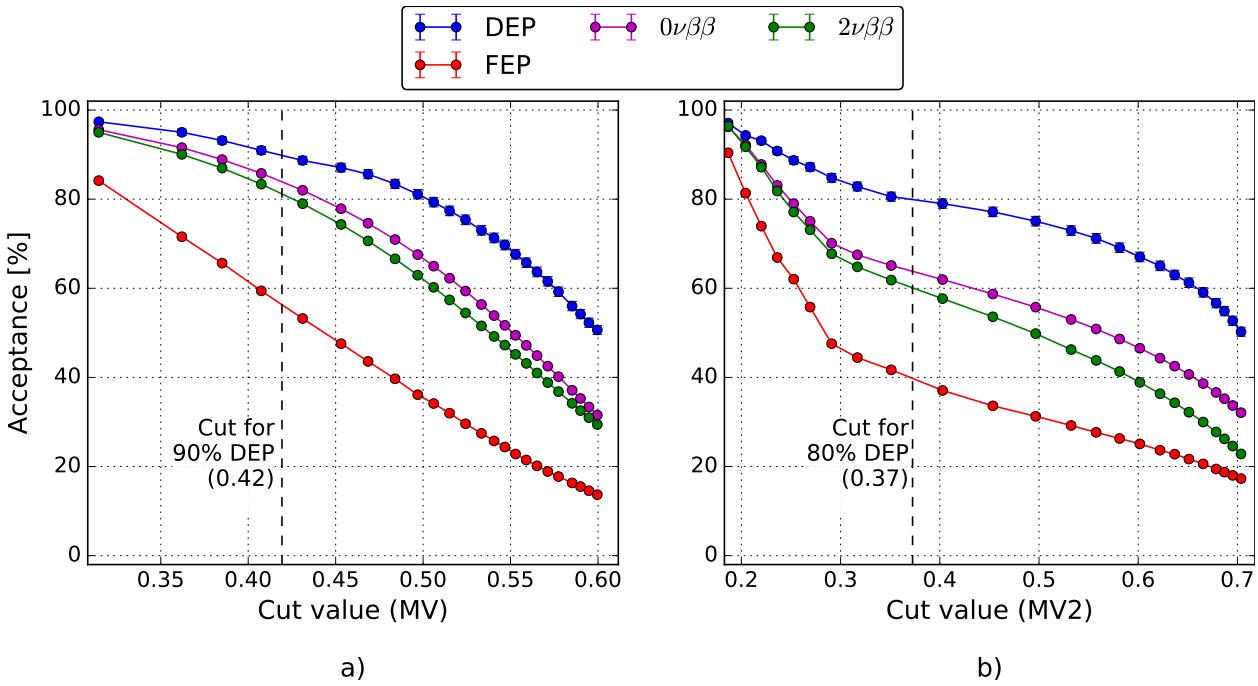


Fig. 4.13: Relation between the acceptance and the cut value of simulated SSEs for different energy regions and topologies: DEP (^{228}Th dataset), $0\nu\beta\beta$ and $2\nu\beta\beta$ decay events in the ANG5 detector (channel 8). FEP curve was included to show also the MSEs acceptance (see also Fig. 4.6 for the similar plot for a BEGe detector in GERDA). For both MV (a) and MV2 classifiers (b) the acceptance is higher for the DEP than for the $0\nu\beta\beta$ and $2\nu\beta\beta$ events. The difference is not constant, but is getting bigger with the increasing cut (decreasing DEP acceptance). Therefore, to avoid the discrepancy between the DEP and the $0\nu\beta\beta$ acceptances, cut should be set for high DEP survival probability, which in turn is not favorable from the background suppression point of view (FEP curve).

here for the MV and MV2 classifiers. Results can be observed on the plot in Fig. 4.14, where the survival efficiency of the $0\nu\beta\beta$ decay events is shown as a function of $r-z$ and $x-y$ coordinate pairs. Like before, the cut values for MV and MV2 classifiers were set for 90% and 80% DEP survival probabilities, respectively. One of the conclusions, that can be drawn from the plot, is that the acceptance of the $0\nu\beta\beta$ decay events has abrupt changes – it is either very high ($\sim$ 100%, yellow zones) or very low ($0-25\%$, black/purple zones). Thus, the reduction of the $0\nu\beta\beta$ decay events acceptance with the increasing cut value is caused by almost complete suppression of consecutive "zones" in the detector, at least in the first step (that is lowering the cut to 85-95%). Further increase of the cut value causes gradual development of the suppressed zones further into the detector volume. This was confirmed by observing the series of similar plots, but with a varied cut value. Furthermore, the "strips" of the low $0\nu\beta\beta$ decay events acceptance were visible also for the MV2 classifier, if the cut was lowered (which in turn increased DEP acceptance). The pattern is also visible in Fig. 4.15. The figure shows a distribution of mean classifier value for each bin (see the caption for details), along with shapes of current signals from the selected regions of the detector.

Alpha events suppression efficiency. In Sec. 4.2 very good efficiency of the α events suppression was reported for the MV2 classifier (Tab. 4.4, $\approx$ 18% survival probability, in contrast with $\approx$ 56% for MV). Comparing the values in Tab. 4.6 and Tab. 4.7 it is evident that $0\nu\beta\beta$ decay events acceptance is lower for the MV2 classifier (65.5% vs 81.7%) and therefore the reason of the better performance of α events suppression may be simply lower overall signal efficiency, especially since DEP is not a good proxy for the $0\nu\beta\beta$ decay (see Fig. 4.13). A following procedure was

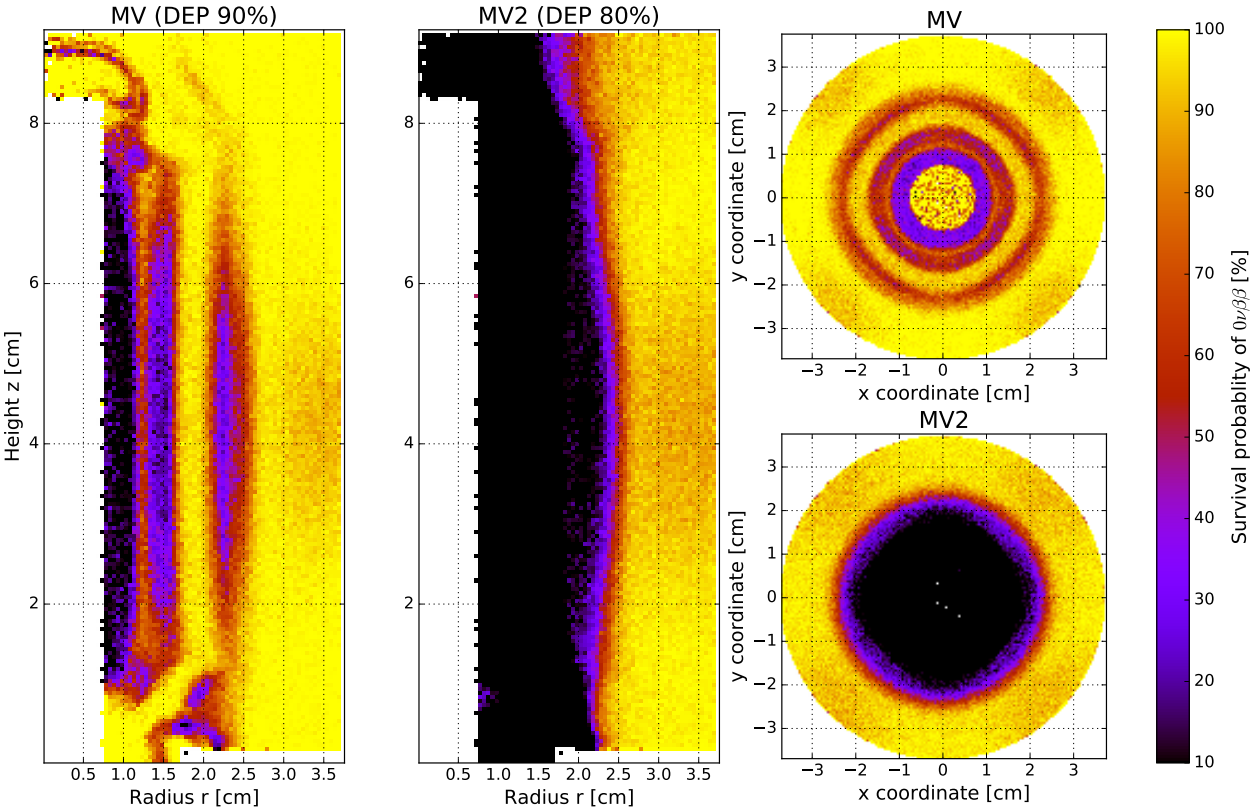


Fig. 4.14: Volumetric survival probability distribution of the $0\nu\beta\beta$ decay (single-site) events in the ANG3 detector.
Two panels on the left – the distributions in the $r-z$ coordinates for the MV and MV2 classifiers, respectively. The cut, set to 90% DEP acceptance, excludes certain zones in the detector, where the survival probability drops abruptly to $\approx 30\%$. This can be observed for radii of 0.7 (p^+ contact), 1.5 and 2.25 cm. MV2, on the other hand, vetoes a large portion of the inner detector's volume, including the regions near the p^+ contact and the groove, where the α-induced background originates.
Two plots on the right-hand side – the same distributions, but in the $x-y$ coordinates. Angular dependence in the PSD performance can be observed – the survival probability is slightly ($\approx 5\%$) lower for the diagonals and outer radii of the detector (both MV and MV2 classifiers).

performed to test this hypothesis (as an example numerical values for RG1 detector were provided in parentheses for the illustrative purposes):

1. Using the interpolated acceptance curves (DEP vs the $0\nu\beta\beta$ decay events acceptances, similar as in Fig. 4.13a) for the *MV classifier*, a new cut value was found making the MV and MV2 $0\nu\beta\beta$ decay efficiencies the same. MV2 classifier $0\nu\beta\beta$ decay efficiency was taken from Tab. 4.7 (69.8%).
2. The DEP events acceptance by the *MV classifier* had to be lowered for this new cut (77%). At this point the $0\nu\beta\beta$ decay events acceptances are the same for both classifiers.
3. The *new cut* was also applied to the GERDA Phase II data to calculate a new α survival probability (46.6%). Since the cut was higher (DEP acceptance lower) than before, the α rejection was better than for the 90% DEP acceptance level (76.6%).

The procedure was repeated for every detector and new average α survival probability of 39.6% was calculated for the MV classifier. Still, it is much higher than the value of 17.7% for the MV2 classifier. It can be therefore concluded that for the same $0\nu\beta\beta$ decay efficiency the MV2 classifier is still better for the α events suppression.

Analysis of the survival probability of the $0\nu\beta\beta$ events near the detector's core (central panel of Fig. 4.14, $r \approx 0.7$ cm) shows that they are completely vetoed by the MV2 classifier. This is

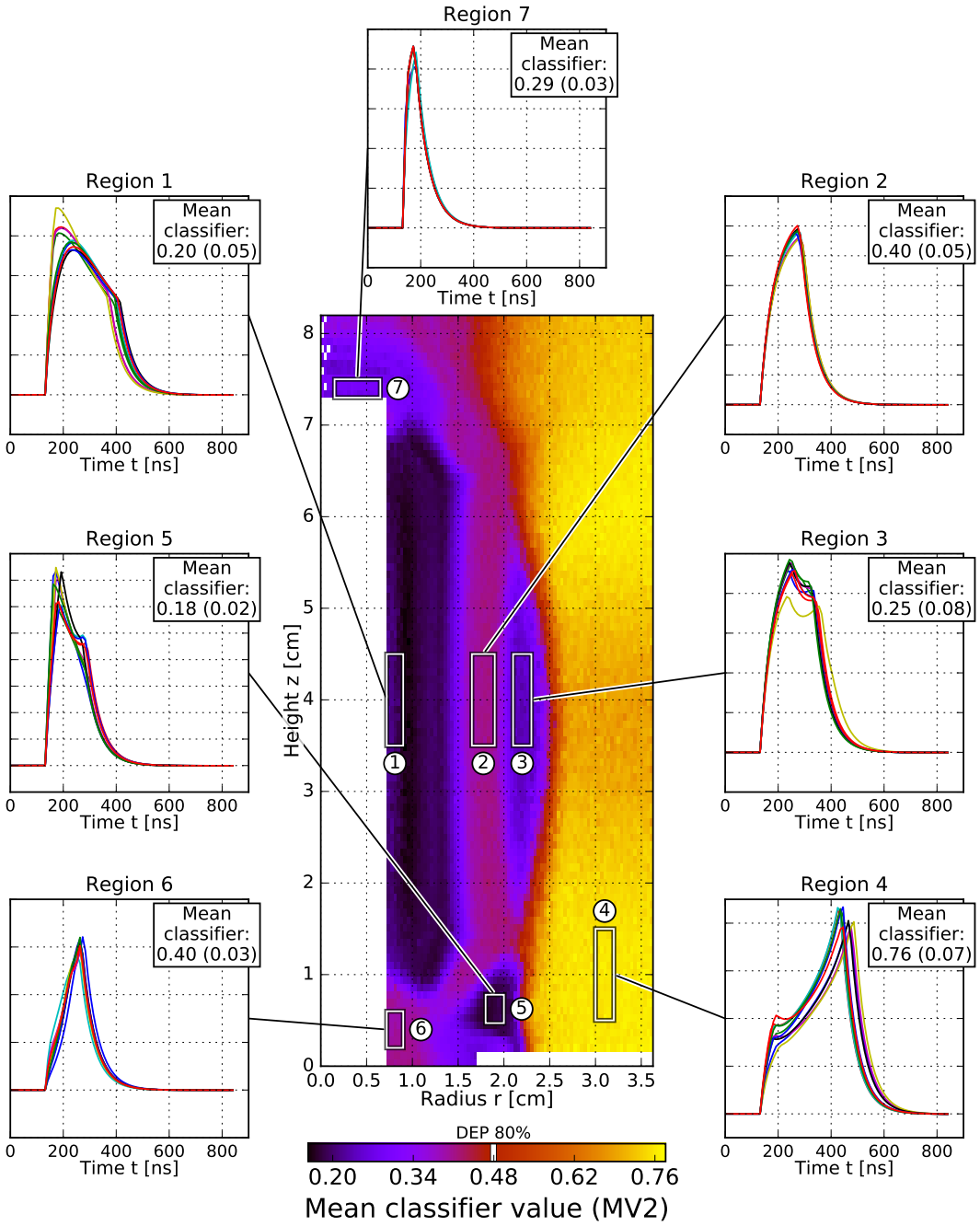


Fig. 4.15:

Central panel – mean classifier (MV2) distribution in the RG1 detector (channel 9), as a function of the $r-z$ coordinates. Each bin selects a number of the $0\nu\beta\beta$ decay events, for which the mean value of the classifier is calculated and shown on the plot. A classifier value corresponding to the 80% DEP acceptance cut is marked on the colorbar with a white stripe. Several detector regions (white rectangles) were selected to investigate the pulse shapes.

Insets – every inset shows 10 examples of current traces, characteristic to a given region in a detector. The noiseless current pulses were shown, instead of standard noisy charge traces, to visually expose the differences in pulse shapes. The mean classifier value (along with a standard deviation in the parentheses) was calculated and noted in the text box.

Plots visualize pulse shapes in the regions of the detector with different mean classifier values, in order to explain the abrupt changes in the classifier distribution. Low classifier values (events vetoed by the PSD) have a fast rising initial part of the pulse (e.g near p^+ contact – region 1). Waveforms corresponding to the central regions (2 and 3) rise slower than in region 1, but are still under 80% DEP threshold and therefore are rejected. The only events which survive the cut are in the outer parts of the detector (right bottom inset – region 4) – they are characterized with the small contribution of the electron-induced part (small bump around 200 ns) and slow rising of its hole-induced counterpart.

probably due to the fast rising edge of the signals originating there (Fig. 4.15, region 1), which is taken into the account in a training process of MV2. MV, on the other hand, uses the PCA transformation, calculated on the calibration data (γ events), and consequently is more "focused" on the central part of the pulse. This is because most of the γ events have their energy deposited in the middle and outer part of the detector, since they constitute the majority of the detector's volume. Consequently, the fast rising edge has a diminished effect on the classifier, since the input variables carrying this information are transformed by PCA with small weights, in contrast to the variables closer to the current maximum (Fig. 3.5). Lastly, it should be noted that there is a variety of signal shapes that can be caused by the α radiation – regions 1, 6 and 7 in Fig. 4.15 are all near the p^+ contact. However, signals from e.g. region 7 are much shorter than the ones from region 5. This fact is recognized by the classifier, since the events have different mean classifier values (0.29 vs 0.18, for regions 7 and 5 respectively). On the other hand, region 6 waveforms do not rise as fast as those from region 7 and the mean classifier value is much higher here (0.4) – the waveforms also have the different shape than those from region 1. Still, all above mentioned regions are rejected, because the classifier values are well below the 80% DEP acceptance cut value (0.48).

Since the surface induced events, like the ones from the α and β radiation cannot be vetoed by the LAr scintillation light signal[5], the MV2 classifier offers a powerful possibility to reduce their contribution to the background. It is worth noting that the method *was not* trained using any signals caused by the α particles, but only the ^{228}Th calibration data. Therefore, the classifier trained on the γ-ray data can offer additional background reducing capabilities, virtually "for free".

The only drawback is a rather low signal ($0\nu\beta\beta$ decay) efficiency, which can reduce the limit setting possibilities of the experiment. Presented results are the first assessment of the Projective Likelihood based PSD methods on the MC generated signals and provide means for better understanding the classifier's performance. The very nature of the Projective Likelihood based classifier makes it very well suited for the rejection of so called "outliers" – events with the traits outside of the training set and not matching the signal nor the background datasets. On the other hand, neural-network based methods, like TMlpANN (a leading method used for the $0\nu\beta\beta$ analysis in GERDA) or MLP (presented in the earlier chapters with the application for the γ spectrometers) provide better separation between the signal and background sets – in this case SSEs and MSEs. But because they do not veto the surface events as efficiently, a second training has to be made on SSEs from $2\nu\beta\beta$ decay region from physics data and high energy α region. This is problematic, since the α population is quite small from the multivariate PSD training point of view.

It would be therefore ideal to have a single classifier, being able to veto both background types, that is from γ and α radiation, while preserving as much SSEs as possible. Taking into account the results concerning the α events suppression efficiency, the MV2 classifier fulfills the first requirement. The only drawback is its signal efficiency – it should be further optimized to preserve more $0\nu\beta\beta$ events. This can be achieved by e.g. modifying the training procedure so that the cut rejects smaller detector's volume, but still just enough to exclude the thin layer of the detector around the core. Although there are still problems with the agreement between the MC and Phase II data efficiencies, the qualitative behavior of the classifier should be very similar.

[5] Assuming that the particle is emitted at the surface and directed into the crystal.

Conclusions

New generation of the $0\nu\beta\beta$ decay experiments will require unprecedently low background levels to achieve their discovery potential or to be able to exclude the inverted neutrino mass hierarchy scenario [57]. GERDA Phase I, due to the background lower by an order of magnitude with respect to previous IGEX and Heidelberg-Moscow experiments, could refute with high probability, even with the much lower exposure, the long-standing claim of the decay observation by a part of the Heidelberg-Moscow collaboration [7]. In GERDA Phase II the experiment is virtually "background-free" (the first one in the field) and its sensitivity is increasing linearly with the acquired exposure [9]. The low background level in Phase I was achieved by the application of ultra-clean materials and liquid argon shielding. In the second phase of the experiment, the active liquid argon veto allowed for the further background reduction. However, Pulse Shape Discrimination was also a crucial tool to obtain the designed sensitivities.

According to the GERDA background model [44], one of the biggest background issues is related to the so called surface events - α and β decays, mostly from ^{210}Po (radon decay chain) and ^{42}K (a daughter of ^{42}Ar), respectively. Since the particles are emitted in the detector's direction, they do not produce scintillation light in liquid argon. The contribution of the surface events to the GERDA Phase I background index was estimated to be as high as 35% [44]. Therefore, the Pulse Shape Discrimination is the only way capable of lowering the surface-related background. The leading Pulse Shape Discrimination method in the experiment for the semi-coaxial detectors, based on the neural-network classifier does not exclude all α originating events [9, 45]. This is due to the fact that it is trained on the ^{228}Th calibration data and is therefore efficient mostly for the single- and multi-site events discrimination. Consequently, a second neural-network had to be applied to efficiently exclude the degraded α events from the $Q_{\beta\beta}$ region [13]. However, training of the network is problematic due to the limited number of α events (training sample) registered in the experiment. The presented method, based on the Projective Likelihood classifier can veto them with high efficiency without additional training and can potentially be a good alternative to the existing approach. In the case of the BEGe detectors, the method based on the Multi Layer Perceptron neural-network has shown a good Pulse Shape Discrimination efficiency, when applied to data from a detector in the vacuum cryostat (Sec. 3.3, [8]). Since the method's training process is highly automated and does not require application of a large number of corrections, it can be potentially used in the future ^{76}Ge-based experiments (like LEGEND [15], employing the BEGe/PPC type detectors [86]) with a large number of the readout channels.

Another important filed of application of the Pulse Shape Discrimination is γ-ray spectrometers. Is is possible by increase their efficiency by rejecting single-site events from the Compton continuum, while preserving the multi-site events from Full Energy Peaks. The germanium-based spectrometers are routinely used for the prescreening of construction materials for the low-background experiments, with the specific sensitivities lower than 40 µBq/kg and 15 µBq/kg, for the ^{40}K and ^{238}U/^{232}Th isotopes, respectively (GeMPI detector [116]). The improvement provided by Pulse Shape Discrimination comes without any shielding nor detector modifications – the only requirement is the digitization of preamplifier signals . The method was tested on the semi-coaxial and BEGe based γ-ray spectrometers and allowed for the sensitivity (Figure of Merit, as defined in

Chap. 3) improvement by $\approx 25\%$ and $\approx 40\%$, respectively. The improved sensitivity is needed to find even more radiopure materials for the future $0\nu\beta\beta$ decay/dark matter experiments.

A new approach to the dimensionality reduction in Pulse Shape Discrimination methods, namely Principal Component Analysis [98], provides an unbiased selection of the input variables with the highest discriminating power. Moreover, the presented optimization results for the small p-type HPGe detectors show that the reduction increases the multi-site events rejection efficiency for the limited training statistics. While in the laboratory settings gathering enough statistics is usually not a problem, it can be challenging in the case of large-scale experiments, like GERDA or LEGEND. Due to the modifications of experimental conditions (gain or bias voltage changes) some parts of the data show abruptly different input variables distributions. In these cases, the data has to treated separately and the resulting training statistics is low. The dimensionality reduction can be also helpful for the classifiers prone to overtraining [90]. It was again demonstrated (Sec. 3.1) that limiting the number of principal components used for the training equalized the acceptance of single-site events in the different energy regions, namely double-escape peak and Compton edge of 2614.5 keV γ line.

Some drawbacks of the presented methods were also observed and discussed. Further work on the Projective Likelihood method with the variable summing (Sec. 4.3) can possibly increase its acceptance of the $0\nu\beta\beta$ decay events, while keeping the high rejection efficiency of the surface events. The compact electronics model, described in the thesis, provides a good approximation of the first order effects of the preamplifier on the pulse shape. This includes the limited bandwidth and decaying tail due to the discharge of the integrating capacitor in the preamplifier circuitry. However, some discrepancies were still observed between the Pulse Shape Discrimination efficiency on the data and Monte Carlo simulated pulses. More laboratory tests would be required to decide what is the reason for them, preferably with the detectors with known crystal impurity concentration profiles, which are needed for the electric field and pulse shape simulation. Until then, the origin of discrepancies will be challenging to determine, since both the electronic and the pulse simulation models may require improvements. The presented electronics model can be easily modified by adding more terms to the circuit, which could be responsible for the different pulse shapes due to e.g. long coaxial cables, connecting the preamplifier with the data acquisition hardware. It was also observed that the discrepancies are smaller if the Projective Likelihood classifier uses input variables after the Principal Component Analysis dimensionality reduction.

Lastly, the constantly developing machine-learning field provides good outlooks for the future. The interesting candidate here are the "deep-learning" neural-network algorithms, which use a more complicated network structure, that is: more hidden layers. This fact allows them to extract more abstract features from the raw data and they are successfully applied e.g. in the image recognition [117]. The reason that they started to be used only recently is that their traning phase is numerically difficult. With the widespread of the Graphic Processing Unit based computation the training speed could be increased by the factor of 50 or more [118]. Recently presented methods [119] by the Max Planck Institute in Munich group, applying the deep-learning algorithms to the GERDA data, show very promising results. It seems that the deep-learning is a natural extension of the Multi-Layer Perceptron multivariate methods, presented in this thesis, which would be currently classified as "shallow network" due to a single hidden layer [118].

Acronyms and abbreviations

ANN artificial neural-network.

B.C. boundary condition.
BI background index.
BEGe Broad Energy Germanium.
BW background window.

CE Compton edge.
C.L. confidence level.
cts counts.

DEP double escape peak.
$2\nu\beta\beta$ double-beta.

FADC flash analog to digital converter.
FEP full energy peak.
FoM Figure of Merit.
FWHM full width at half maximum.

GTF Genius Test Facility.

HPGe High Purity Germanium.
HV high voltage.

IGEX International Germanium EXperiment.

JFET junction field-effect transistor.

KDE Kernel Density Estimation.

LAr liquid argon.
LN_2 liquid nitrogen.
LNGS Laboratori Nazionali del Gran Sasso.

MC Monte Carlo.
MCS multiple Compton scattering.
MLP Multi-Layer Perceptron.
MSE multi-site event.
MWA Moving Window Average.

$0\nu\beta\beta$ neutrinoless double-beta.

PCA Principal Component Analysis.
PCB printed circuit board.
PDF probability density function.

PMT Photomultiplier Tube.
PPC P-type Point Contact.
PSD Pulse Shape Discrimination.

ROI region of interest.

SEP single escape peak.
SiPM Silicon Photomultiplier.
SSE single-site event.

TBP tetraphenyl butadiene.
TMVA Toolkit for Multivariate Data Analysis.

Appendix A

Detector-wise results of the PSD analysis in GERDA Phase II

A.1 BEGe detectors

Table A.1: PSD efficiency calculated for the BEGe detectors in the ^{228}Th dataset (MV classifier: MLP neural-networks and PCA dimensionality reduction). Efficiency in the "Total" row is calculated by taking all available events from the calibration data (i.e. statistics-weighted average from the acceptances of the individual detectors).

Channel number	Cut value	DEP [%]	^{212}Bi [%]	$Q_{\beta\beta}$ [%]	SEP [%]	C. edge [%]	^{208}Tl [%]
0 (91A)	0.687	89.6 ± 1.0	16.1 ± 1.8	43.0 ± 0.3	8.9 ± 0.8	75.1 ± 0.4	11.1 ± 0.1
1 (35B)	0.667	89.8 ± 0.9	16.3 ± 1.4	45.2 ± 0.2	10.7 ± 0.6	80.9 ± 0.3	10.2 ± 0.1
2 (02B)	0.679	88.8 ± 1.1	16.4 ± 2.1	44.2 ± 0.3	9.2 ± 0.8	77.4 ± 0.4	11.0 ± 0.1
3 (00B)	0.674	88.4 ± 1.2	16.4 ± 2.1	42.3 ± 0.3	11.6 ± 0.9	79.1 ± 0.4	11.2 ± 0.1
4 (61A)	0.591	89.5 ± 1.0	22.5 ± 1.5	49.4 ± 0.3	15.6 ± 0.7	87.1 ± 0.3	15.4 ± 0.1
6 (02D)	0.566	89.8 ± 1.0	24.8 ± 1.8	50.9 ± 0.3	16.9 ± 0.8	90.8 ± 0.3	17.3 ± 0.2
11 (02A)	0.662	90.4 ± 0.9	17.5 ± 1.6	44.2 ± 0.3	9.0 ± 0.8	81.7 ± 0.4	11.5 ± 0.2
13 (32A)	0.648	89.7 ± 1.0	27.4 ± 2.3	47.1 ± 0.4	11.7 ± 1.2	79.1 ± 0.5	14.6 ± 0.2
16 (61C)	0.621	87.7 ± 0.8	24.5 ± 1.4	48.4 ± 0.3	14.1 ± 0.7	81.8 ± 0.3	13.5 ± 0.1
17 (76B)	0.662	86.7 ± 1.0	24.3 ± 2.4	46.6 ± 0.4	13.6 ± 1.1	72.0 ± 0.5	12.6 ± 0.2
18 (00C)	0.709	88.8 ± 0.9	14.7 ± 1.1	41.2 ± 0.3	7.7 ± 0.6	75.5 ± 0.4	8.5 ± 0.1
19 (35C)	0.720	89.2 ± 0.9	17.1 ± 1.6	41.4 ± 0.4	8.5 ± 0.7	72.6 ± 0.5	8.3 ± 0.1
20 (76C)	0.704	87.9 ± 0.8	15.9 ± 1.1	40.5 ± 0.3	7.1 ± 0.5	73.1 ± 0.4	7.0 ± 0.1
21 (89D)	0.664	89.3 ± 0.8	20.6 ± 1.9	47.0 ± 0.4	9.3 ± 1.0	80.5 ± 0.4	12.1 ± 0.2
22 (00D)	0.671	88.2 ± 0.9	15.1 ± 1.5	43.9 ± 0.3	10.3 ± 0.7	77.8 ± 0.4	9.8 ± 0.1
24 (35A)	0.635	91.6 ± 0.9	23.4 ± 1.2	46.8 ± 0.3	16.3 ± 0.6	81.7 ± 0.3	12.9 ± 0.1
26 (61B)	0.667	88.9 ± 0.9	19.5 ± 1.5	45.7 ± 0.3	11.2 ± 0.7	79.9 ± 0.4	10.1 ± 0.1
30 (00A)	0.677	91.4 ± 0.9	18.4 ± 2.1	45.0 ± 0.4	10.9 ± 0.9	82.5 ± 0.4	11.1 ± 0.2
31 (02C)	0.668	91.0 ± 0.8	19.5 ± 1.2	45.2 ± 0.3	12.3 ± 0.6	84.3 ± 0.3	9.5 ± 0.1
32 (79B)	0.690	88.6 ± 1.0	18.3 ± 1.6	41.9 ± 0.3	8.7 ± 0.8	74.0 ± 0.5	8.5 ± 0.1
33 (91D)	0.633	89.1 ± 1.0	23.7 ± 1.7	47.7 ± 0.3	12.9 ± 0.8	80.8 ± 0.4	12.8 ± 0.1
35 (89A)	0.650	88.8 ± 0.9	21.9 ± 1.9	45.2 ± 0.3	13.8 ± 0.8	76.2 ± 0.4	12.6 ± 0.2
Total:		89.7 ± 0.2	19.3 ± 0.4	45.2 ± 0.1	11.6 ± 0.2	79.6 ± 0.1	11.3 ± 0.0

Table A.2: Summary of the resulting acceptances for the GERDA physics data for the individual BEGe detectors, using the MV classifier. The $2\nu\beta\beta$ events acceptance was calculated in the 1000 – 1300 keV energy range for all events in this region, as well as ones prefiltered the with LAr veto flag, which should be almost exclusive SSEs. Events in the BW/ROI are tallied after applying the LAr veto.

Channel number	Cut value	$2\nu\beta\beta$ [%]	$2\nu\beta\beta$ LAr [%]	BW ($\pm$200 keV)	BW (190 keV)	Unblinded $Q_{\beta\beta}$ [$\pm$25 keV]	α (3.5–5.5 MeV)
0 (91A)	0.687	75.9 ± 3.1	93.0 ± 2.7	0/0	0/0	0/0	4/6
1 (35B)	0.667	70.9 ± 3.5	88.3 ± 3.3	0/0	0/0	0/0	2/7
2 (02B)	0.679	71.1 ± 4.0	82.1 ± 4.2	1/2	0/1	0/0	10/21
3 (00B)	0.674	73.1 ± 4.1	83.6 ± 4.3	0/0	0/0	0/0	4/10
4 (61A)	0.591	81.6 ± 3.3	91.9 ± 2.7	0/2	0/0	0/0	3/10
6 (02D)	0.566	81.2 ± 3.4	90.4 ± 3.0	0/1	0/1	0/0	2/6
11 (02A)	0.662	67.4 ± 3.9	83.3 ± 4.8	0/0	0/0	0/0	1/5
13 (32A)	0.648	83.7 ± 3.9	91.5 ± 3.6	0/0	0/0	0/0	4/8
16 (61C)	0.621	77.9 ± 4.3	82.9 ± 4.5	2/3	2/2	0/0	2/7
17 (76B)	0.662	73.3 ± 5.1	88.2 ± 4.5	0/0	0/0	0/0	6/8
18 (00C)	0.709	74.3 ± 3.7	90.3 ± 3.1	0/1	0/1	0/0	1/1
19 (35C)	0.720	73.1 ± 3.6	87.5 ± 3.9	0/0	0/0	0/0	3/6
20 (76C)	0.704	68.4 ± 3.5	80.4 ± 3.8	0/0	0/0	0/0	0/0
21 (89D)	0.664	79.6 ± 4.1	86.8 ± 4.1	0/0	0/0	0/0	4/13
22 (00D)	0.671	77.0 ± 3.1	86.4 ± 3.0	0/0	0/0	0/0	1/2
24 (35A)	0.635	76.4 ± 3.5	87.9 ± 3.4	0/2	0/0	0/0	5/19
26 (61B)	0.667	78.6 ± 3.3	91.7 ± 2.8	0/0	0/0	0/0	0/3
30 (00A)	0.677	73.3 ± 4.3	91.1 ± 3.8	0/0	0/0	0/0	0/5
31 (02C)	0.668	75.4 ± 3.7	85.7 ± 4.0	0/0	0/0	0/0	12/20
32 (79B)	0.690	86.8 ± 3.9	92.9 ± 3.4	0/1	0/1	0/0	1/4
33 (91D)	0.633	76.1 ± 3.9	83.8 ± 4.1	0/0	0/0	0/0	8/13
35 (89A)	0.650	74.5 ± 4.2	85.4 ± 3.9	1/1	1/1	0/0	3/10
Total:		75.5 ± 0.8	87.4 ± 0.8	4/13	3/7	0/0	76/184

Table A.3: List of the events in the $Q_{\beta\beta}$ ROI of BEGe detectors (MV classifier), with the specified date, energy, classifier value, cut threshold and the LAr veto/PSD flags. Rows with the yellow background show the unblinded events.

Detector name (channel)	Energy	Event date (UTC)	Classifier value vs threshold	Is LAr vetoed	Is PSD vetoed	Is LAr or PSD vetoed
02D (06)	2153.7	2015-Dec-25 19:55:21	0.243 vs 0.566	No	Yes	Yes
02B (02)	2049.4	2016-Jan-03 13:16:28	0.270 vs 0.679	Yes	Yes	Yes
61A (04)	2187.2	2016-Jan-12 14:30:16	0.378 vs 0.591	Yes	Yes	Yes
61C (16)	1944.9	2016-Jan-28 01:00:55	0.985 vs 0.621	No	No	No
79B (32)	2113.4	2016-Feb-01 11:26:23	0.218 vs 0.690	No	Yes	Yes
02A (11)	2179.8	2016-Feb-12 12:45:16	0.434 vs 0.662	Yes	Yes	Yes
00C (18)	2126.3	2016-Feb-19 23:03:02	0.533 vs 0.709	No	Yes	Yes
89D (21)	2020.2	2016-Feb-21 18:33:39	0.368 vs 0.664	Yes	Yes	Yes
00D (22)	2036.9	2016-Mar-04 10:18:06	0.379 vs 0.671	Yes	Yes	Yes
00B (03)	2014.6	2016-Mar-08 04:15:05	0.774 vs 0.674	Yes	No	Yes
61C (16)	1959.4	2016-Mar-13 05:40:59	0.743 vs 0.621	No	No	No
61A (04)	1972.4	2016-Mar-17 17:44:32	0.377 vs 0.591	Yes	Yes	Yes
61B (26)	2138.2	2016-Mar-23 07:32:46	0.458 vs 0.667	Yes	Yes	Yes
35B (01)	2157.3	2016-Apr-04 06:44:37	0.801 vs 0.667	Yes	No	Yes
61C (16)	1988.5	2016-Apr-05 15:09:28	0.229 vs 0.621	Yes	Yes	Yes
35A (24)	2159.7	2016-Apr-27 02:42:51	0.563 vs 0.635	Yes	Yes	Yes
89A (35)	1997.2	2016-May-04 02:47:12	0.750 vs 0.650	No	No	No
35A (24)	1936.7	2016-May-04 10:42:39	0.494 vs 0.635	Yes	Yes	Yes
61C (16)	2062.1	2016-May-18 00:48:49	0.347 vs 0.621	Yes	Yes	Yes
02B (02)	1967.4	2016-May-24 17:56:16	0.380 vs 0.679	No	Yes	Yes
00C (18)	1945.5	2016-May-30 17:29:45	0.705 vs 0.709	Yes	Yes	Yes
02B (02)	2157.6	2016-May-31 06:47:00	0.320 vs 0.679	Yes	Yes	Yes
			Events left:	7	5	3/22

A.2 Semi-coaxial detectors

A.2.1 MV classifier

Table A.4: PSD efficiency calculated for the semi-coaxial detectors in the ^{228}Th dataset (MV classifier: Projective Likelihood and the PCA dimensionality reduction). Efficiency in the "Total" row is calculated by taking all available events from the calibration data (i.e. statistics-weighted average from the acceptances of the individual detectors).

Channel number	Cut value	DEP [%]	^{212}Bi [%]	$Q_{\beta\beta}$ [%]	SEP [%]	C. edge [%]
8 (ANG5)	0.436	91.5 ± 1.3	53.5 ± 0.6	68.5 ± 0.2	49.8 ± 0.4	88.8 ± 0.2
9 (RG1)	0.404	90.2 ± 1.0	67.7 ± 0.7	78.4 ± 0.2	66.4 ± 0.4	90.8 ± 0.2
10 (ANG3)	0.429	91.3 ± 1.1	58.5 ± 0.6	72.3 ± 0.2	56.7 ± 0.4	86.2 ± 0.2
27 (ANG2)	0.416	88.7 ± 1.2	60.3 ± 0.6	75.3 ± 0.2	60.5 ± 0.4	91.3 ± 0.2
28 (RG2)	0.420	88.2 ± 1.1	55.6 ± 0.7	71.2 ± 0.2	53.8 ± 0.4	88.9 ± 0.2
29 (ANG4)	0.417	91.5 ± 0.9	61.6 ± 0.5	75.1 ± 0.2	58.6 ± 0.4	89.8 ± 0.2
36 (ANG1)	0.406	92.0 ± 0.7	67.4 ± 0.9	80.6 ± 0.2	68.6 ± 0.5	93.1 ± 0.2
Total:		89.8 ± 0.4	59.8 ± 0.3	74.0 ± 0.1	58.1 ± 0.2	88.8 ± 0.1

Table A.5: Summary of the resulting acceptances for the GERDA physics data for the individual semi-coaxial detectors, using the MV classifier. The $2\nu\beta\beta$ decay events acceptance was calculated in the 1000 – 1300 keV energy range for all events in this region, as well as ones prefiltered the with LAr veto flag, which should be almost exclusive SSEs. Events in the BW/ROI are tallied after applying the LAr veto.

Channel number	Cut value	$2\nu\beta\beta$ [%]	$2\nu\beta\beta$ LAr [%]	ROI ($\pm$200 keV)	ROI (190 keV)	Unblinded $Q_{\beta\beta}$ [$\pm$25 keV]	α (3.5–5.5 MeV)
8 (ANG5)	0.436	75.5 ± 2.0	82.8 ± 2.3	1/6	1/2	0/1	22/59
9 (RG1)	0.404	85.7 ± 1.9	87.1 ± 2.2	1/1	1/1	0/0	64/85
10 (ANG3)	0.429	78.4 ± 1.9	82.5 ± 2.2	6/6	2/2	1/1	105/120
27 (ANG2)	0.416	77.2 ± 1.8	81.8 ± 2.2	0/3	0/1	0/0	10/31
28 (RG2)	0.420	76.7 ± 2.1	81.1 ± 2.3	1/1	0/0	0/0	28/50
29 (ANG4)	0.417	81.0 ± 1.8	81.6 ± 2.2	4/5	2/2	0/0	221/287
36 (ANG1)	0.406	86.0 ± 2.4	90.3 ± 2.6	1/2	0/1	1/1	12/36
Total:		79.3 ± 0.8	83.2 ± 0.9	14/24	6/9	2/3	462/668

Table A.6: List of the events in the $Q_{\beta\beta}$ ROI of semi-coaxial detectors (MV classifier), with the specified date, energy, classifier value, cut threshold and the LAr veto/PSD flags. Rows with the yellow background show the unblinded events.

Detector name (channel)	Energy	Event date (UTC)	Classifier value vs threshold	Is LAr vetoed	Is PSD vetoed	Is LAr or PSD vetoed
ANG3 (10)	1991.7	2015-Dec-31 15:16:48	0.560 vs 0.429	Yes	No	Yes
ANG1 (36)	1988.4	2016-Jan-27 15:34:06	0.402 vs 0.406	No	Yes	Yes
ANG3 (10)	2047.8	2016-Feb-10 09:54:09	0.227 vs 0.429	Yes	Yes	Yes
ANG4 (29)	1994.4	2016-Feb-10 13:04:08	0.471 vs 0.417	No	No	No
ANG4 (29)	1931.9	2016-Feb-23 03:04:09	0.513 vs 0.417	No	No	No
ANG4 (29)	2097.3	2016-Mar-02 19:46:00	0.288 vs 0.417	Yes	Yes	Yes
ANG3 (10)	1980.2	2016-Mar-05 10:21:25	0.491 vs 0.429	No	No	No
ANG3 (10)	1971.1	2016-Mar-13 04:42:33	0.628 vs 0.429	No	No	No
RG1 (09)	2160.7	2016-Mar-18 12:39:46	0.367 vs 0.404	Yes	Yes	Yes
ANG3 (10)	2178.9	2016-Mar-23 09:53:50	0.482 vs 0.429	Yes	No	Yes
ANG5 (08)	2016.1	2016-Mar-24 22:37:36	0.344 vs 0.436	No	Yes	Yes
ANG3 (10)	2063.9	2016-Mar-28 16:00:18	0.625 vs 0.429	No	No	No
ANG5 (08)	1981.2	2016-Apr-20 04:26:22	0.307 vs 0.436	No	Yes	Yes
ANG5 (08)	1990.8	2016-Apr-26 04:20:12	0.407 vs 0.436	Yes	Yes	Yes
ANG2 (27)	2009.2	2016-Apr-26 13:48:54	0.213 vs 0.416	No	Yes	Yes
RG1 (09)	2073.5	2016-May-01 13:09:52	0.541 vs 0.404	No	No	No
ANG1 (36)	2061.0	2016-May-22 11:44:40	0.500 vs 0.406	No	No	No
ANG5 (08)	2082.1	2016-May-29 06:42:10	0.214 vs 0.436	Yes	Yes	Yes
ANG5 (08)	1973.4	2016-May-30 17:45:53	0.500 vs 0.436	No	No	No
			Events left:	12	10	8/19

A.2.2 MV2 classifier

Table A.7: PSD efficiency calculated for the semi-coaxial detectors in the ^{228}Th dataset (MV2 classifier: Projective Likelihood and variable summing). Efficiency in the "Total" row is calculated by taking all available events from the calibration data (i.e. statistics-weighted average from the acceptances of the individual detectors).

Channel number	Cut value	DEP [%]	^{212}Bi [%]	$Q_{\beta\beta}$ [%]	SEP [%]	C. edge [%]
8 (ANG5)	0.405	78.3 ± 1.5	40.4 ± 0.7	56.4 ± 0.2	62.0 ± 0.4	76.4 ± 0.3
9 (RG1)	0.431	78.5 ± 1.4	40.1 ± 0.8	59.5 ± 0.2	55.7 ± 0.5	72.6 ± 0.3
10 (ANG3)	0.430	82.7 ± 1.3	40.1 ± 0.7	57.9 ± 0.2	51.6 ± 0.4	75.2 ± 0.3
27 (ANG2)	0.411	78.0 ± 1.5	36.9 ± 0.6	55.8 ± 0.2	54.1 ± 0.4	75.3 ± 0.2
28 (RG2)	0.436	78.9 ± 1.2	36.4 ± 0.7	56.1 ± 0.2	47.5 ± 0.4	72.6 ± 0.3
29 (ANG4)	0.432	81.3 ± 1.1	38.9 ± 0.6	58.1 ± 0.2	52.0 ± 0.4	76.8 ± 0.2
36 (ANG1)	0.398	80.7 ± 0.9	47.6 ± 1.0	63.0 ± 0.3	55.2 ± 0.6	77.8 ± 0.3
Total:		79.4 ± 0.5	39.2 ± 0.3	57.7 ± 0.1	53.9 ± 0.2	75.1 ± 0.1

Table A.8: Summary of the resulting acceptances for the GERDA physics data for the individual semi-coaxial detectors, using the MV2 classifier. The $2\nu\beta\beta$ decay events acceptance was calculated in the 1000 – 1300 keV energy range for all events in this region, as well as ones prefiltered the with LAr veto flag, which should be almost exclusive SSEs. Events in the BW/ROI are tallied after applying the LAr veto.

Channel number	Cut value	$2\nu\beta\beta$ [%]	$2\nu\beta\beta$ LAr [%]	ROI ($\pm$200 keV)	ROI (190 keV)	Unblinded $Q_{\beta\beta}$ [$\pm$25 keV]	α (3.5–5.5 MeV)
8 (ANG5)	0.405	60.1 ± 2.3	68.9 ± 2.8	1/6	0/2	0/1	0/59
9 (RG1)	0.431	58.2 ± 2.7	60.4 ± 3.2	1/1	1/1	0/0	24/85
10 (ANG3)	0.430	56.4 ± 2.3	59.4 ± 2.8	3/6	0/2	1/1	46/120
27 (ANG2)	0.411	53.6 ± 2.2	59.2 ± 2.8	0/3	0/1	0/0	4/31
28 (RG2)	0.436	58.2 ± 2.5	59.6 ± 2.9	0/1	0/0	0/0	0/50
29 (ANG4)	0.432	59.0 ± 2.3	61.9 ± 2.8	2/5	1/2	0/0	44/287
36 (ANG1)	0.398	65.2 ± 3.3	70.1 ± 4.0	0/2	0/1	0/1	0/36
Total:		58.0 ± 0.9	62.1 ± 1.1	7/24	2/9	1/3	118/668

Table A.9: List of the events in the $Q_{\beta\beta}$ ROI of semi-coaxial detectors (MV2 classifier), with the specified date, energy, classifier value, cut threshold and the LAr veto/PSD flags. Rows with the yellow background show the unblinded events.

Detector name (channel)	Energy	Event date (UTC)	Classifier value vs threshold	Is LAr vetoed	Is PSD vetoed	Is LAr or PSD vetoed
ANG3 (10)	1991.7	2015-Dec-31 15:16:48	0.768 vs 0.430	Yes	No	Yes
ANG1 (36)	1988.4	2016-Jan-27 15:34:06	0.189 vs 0.398	No	Yes	Yes
ANG3 (10)	2047.8	2016-Feb-10 09:54:09	0.446 vs 0.430	Yes	No	Yes
ANG4 (29)	1994.4	2016-Feb-10 13:04:08	0.481 vs 0.432	No	No	No
ANG4 (29)	1931.9	2016-Feb-23 03:04:09	0.306 vs 0.432	No	Yes	Yes
ANG4 (29)	2097.3	2016-Mar-02 19:46:00	0.345 vs 0.432	Yes	Yes	Yes
ANG3 (10)	1980.2	2016-Mar-05 10:21:25	0.189 vs 0.430	No	Yes	Yes
ANG3 (10)	1971.1	2016-Mar-13 04:42:33	0.329 vs 0.430	No	Yes	Yes
RG1 (09)	2160.7	2016-Mar-18 12:39:46	0.270 vs 0.431	Yes	Yes	Yes
ANG3 (10)	2178.9	2016-Mar-23 09:53:50	0.560 vs 0.430	Yes	No	Yes
ANG5 (08)	2016.1	2016-Mar-24 22:37:36	0.346 vs 0.405	No	Yes	Yes
ANG3 (10)	2063.9	2016-Mar-28 16:00:18	0.471 vs 0.430	No	No	No
ANG5 (08)	1981.2	2016-Apr-20 04:26:22	0.187 vs 0.405	No	Yes	Yes
ANG5 (08)	1990.8	2016-Apr-26 04:20:12	0.269 vs 0.405	Yes	Yes	Yes
ANG2 (27)	2009.2	2016-Apr-26 13:48:54	0.137 vs 0.411	No	Yes	Yes
RG1 (09)	2073.5	2016-May-01 13:09:52	0.465 vs 0.431	No	No	No
ANG1 (36)	2061.0	2016-May-22 11:44:40	0.269 vs 0.398	No	Yes	Yes
ANG5 (08)	2082.1	2016-May-29 06:42:10	0.122 vs 0.405	Yes	Yes	Yes
ANG5 (08)	1973.4	2016-May-30 17:45:53	0.296 vs 0.405	No	Yes	Yes
			Events left:	12	6	3/19

Appendix B

Impulse response of the preamplifier

B.1 Mathematical model of the electronics response

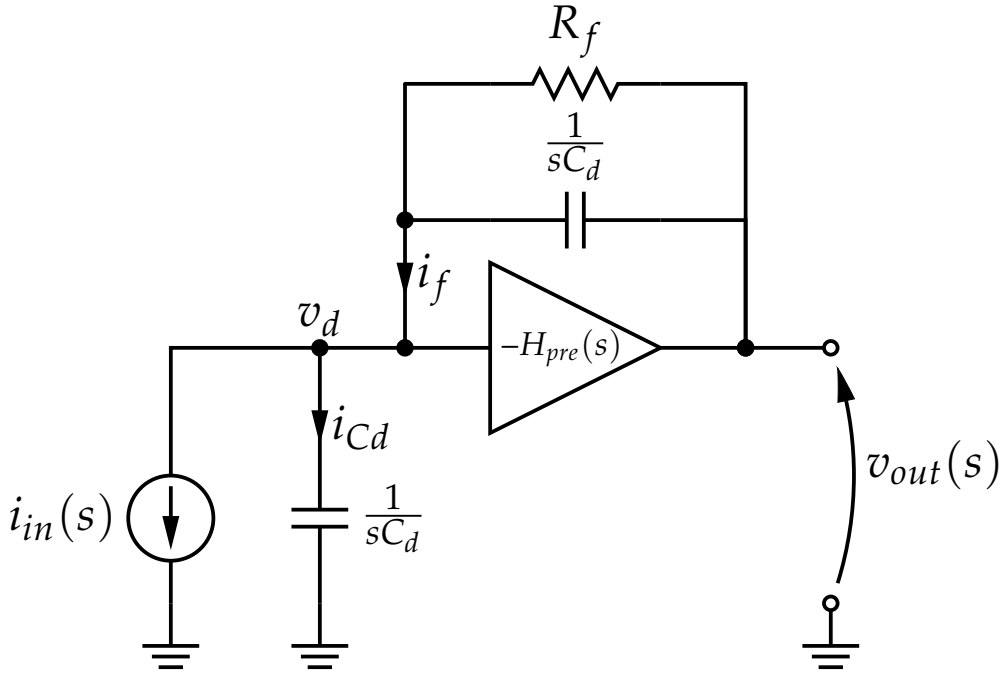


Fig. B.1: Simplified schematics of a charge-sensitive preamplifier configuration.

Waveforms obtained from the Monte Carlo simulation (in this case, energy deposition information from Geant4 and pulse shape simulation from ADL software) have to be further processed in order to include the effects of the electronics in the PSD analysis. Some of these effects are:

- bandwidth limiting, resulting from the limited speed of the charge-sensitive preamplifier,
- exponential tail, a consequence of the feedback capacitor (C_f) discharge by the feedback resistor (R_f),
- capacitance of the detector (C_d), responsible for further limiting of the bandwidth of the front-end system.

One of the approaches to take into account the above mentioned effects is to convolve a current pulse with the impulse response of the preamplifier, which can be analytically obtained with the Laplace analysis of the circuit shown on Fig. B.1. The circuit is the simplest version of a "generic" preamplifier and even though it is not the same circuit that is used in the GERDA experiment, all the above mentioned effects are included in the impulse response. The bandwidth limiting of the operational amplifier is realized by assuming single-pole transfer function with the voltage

gain H_{pre}). Aside from the bandwidth, for the needs of the Laplace analysis, all other features are assumed being ideal (e.g. infinite input impedance).

The transfer function ($T(s) = v_{out}(s)/i_{in}(s)$) of the preamplifier can be described with a following equation:

$$T(s) = \frac{1}{C_f} \frac{1}{\alpha s^2 + s\left[1 + \alpha(\omega_{sum} + \omega_{pre})\right] + \alpha\omega_{pre}\omega_{sum} + \omega_f} \tag{B.1}$$

where:
$C_{sum} = C_f + C_d$
$\alpha = \frac{C_{sum}}{C_f \mathrm{GBP}}$
$\omega_f = \frac{1}{R_f C_f}$
$\omega_{sum} = \frac{1}{R_f C_{sum}}$

The impulse response is then obtained by calculating the inverse Laplace transform of $T(s)$:

$$h(t) = \mathscr{L}^{-1}\{T(s)\} \tag{B.2}$$

It is theoretically possible to use the analytical formula for $h(t)$, however, due to its complication it is very impractical to use. It was therefore calculated using Scipy library for Python programming language [120] after providing numerical values of the parameters. The impulse response $h(t)$ is illustrated at the top panel of Fig. B.2. While almost all parameters were the same for the all coaxial detectors, the GBP parameter was optimized to match the real data from the experiment.

B.2 Derivation of the impulse response

$$i_{in} + i_{Cd} = i_f \qquad i_f = \frac{v_{out} - v_d}{R_f \parallel \frac{1}{sC_d}} \qquad v_{out} = -H_{pre} v_d \qquad v_d = i_{Cd}\left(\frac{1}{sC_d}\right) \tag{B.3}$$

$$T(s) = \frac{v_{out}(s)}{i_{in}(s)} = \frac{1}{C_f} \frac{1}{s + \frac{1}{C_f R_f} + \frac{1}{H_{pre}(s)}\left(\frac{C_f + C_d}{C_f}\right)\left(s + \frac{1}{R_f(C_d + C_f)}\right)} \tag{B.4}$$

$$H_{pre}(s) = \frac{\mathrm{GBP}}{(s + \omega_{pre})} \qquad \omega_{pre} = \frac{\mathrm{GBP}}{K_{pre}} \tag{B.5}$$

after substitution:

$$T(s) = \frac{1}{C_f} \frac{1}{\alpha s^2 + s\left(1 + \alpha(\omega_{sum} + \omega_{pre})\right) + \alpha\omega_{pre}\omega_{sum} + \omega_f} \tag{B.6}$$

$C_{sum} = C_f + C_d$
$\alpha = \frac{C_{sum}}{C_f \mathrm{GBP}}$
$\omega_f = \frac{1}{R_f C_f}$
$\omega_{sum} = \frac{1}{R_f C_{sum}}$

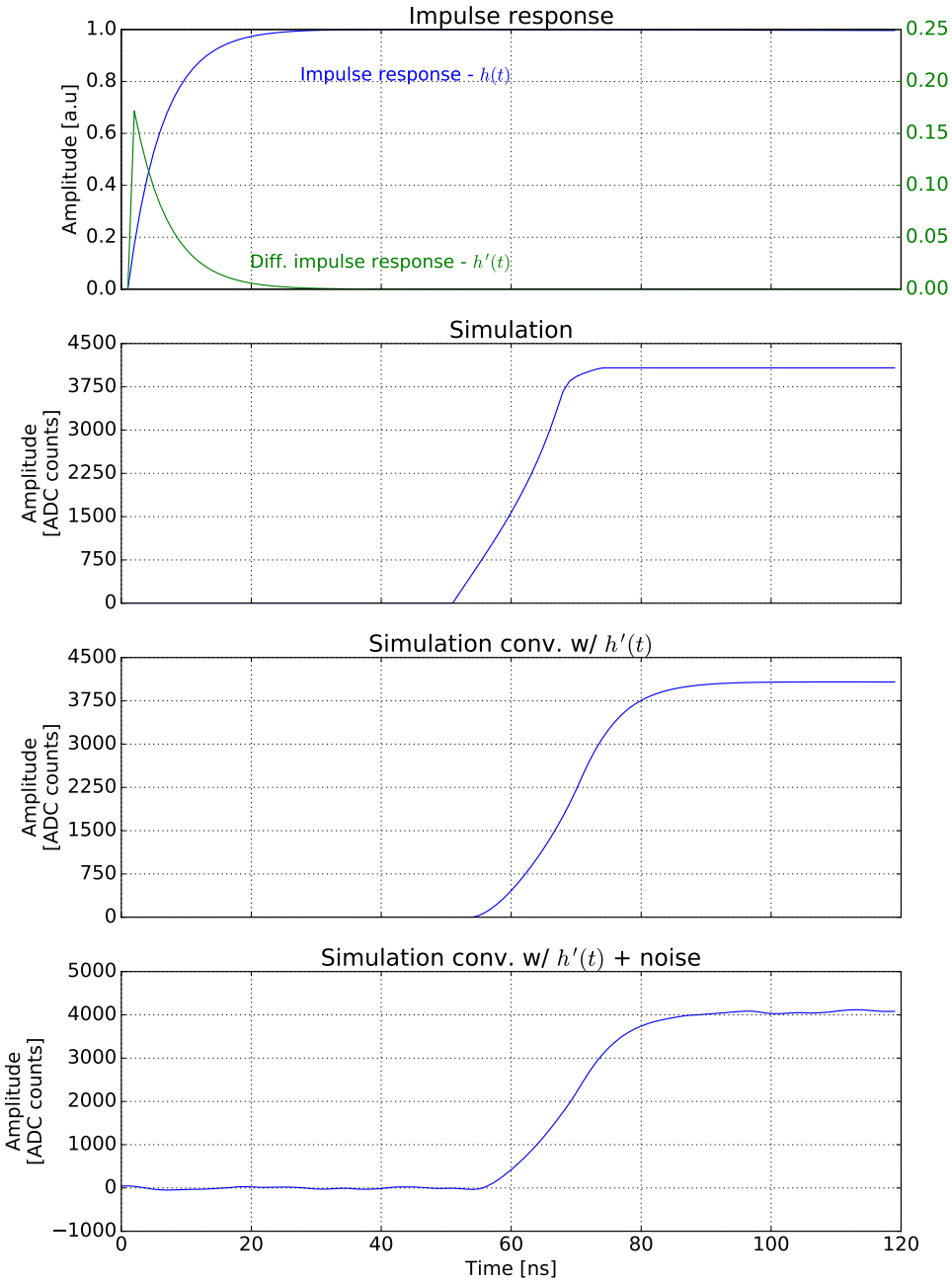


Fig. B.2: Illustration of the waveform convolution with the electronic response. The top panel shows the impulse response $h(t)$, obtained by calculating the inverse Laplace transform of the preamplifier transfer function $T(s)$. Next three panels show the sequential procedure of obtaining the final pulse used in the PSD analysis. The last step involves the addition of a noise waveform, registered from the data when no event signal was present (the so called "baseline" waveform). Since the transfer function $T(s)$ is calculated for the current signal input $i_d(s)$ and the simulation output is the charge waveform $Q(t)$, the impulse response $h(t)$ has to be additionally differentiated ($h'(t)$) to take this into account.

Acknowledgements/Podziękowania

I would like to express my gratitude to my supervisor Professor Marcin Wójcik for the guidance and introducing me to the GERDA experiment. Many thanks to dr Grzegorz Zuzel, who introduced me to the low-background measurements, experimental aspects of the $0\nu\beta\beta$ experiments and helped me with preparing this thesis.

In the everyday work I could always count on my colleagues from the Experimental Computer Physics Department – I would like to thank in particular dr Marcin Misiaszek for introducing me to the pulse shape and data analysis, but also for the fruitful discussions about neutrino physics and life in general. I am also very grateful to dr hab. Zenon Nieckarz for constant encouragement during the writing process, as well as to dr Krzysztof Pelczar, from whom I learned a lot about electronics and numerical methods. Eng. Marek Adamczyk always provided me with instant solutions to all the electronics-related problems I encountered during my measurements, which I am sure would take me ages to solve on my own. Thanks to all my officemates with whom I shared the workspace during my days at the Institute of Physics.

I would like to also acknowledge the help I received from my colleagues from the Data Analysis group of the GERDA experiment, as well as other collaboration members. I would like to especially thank dr Matteo Agostini and dr Luciano Pandola, who were always available to answer my questions. The results concerning the $0\nu\beta\beta$ decay efficiency of the presented methods would not be possible without the help of dr Andrea Kirsch from the Max Planck Institute in Heidelberg, who provided me the official Monte Carlo data of GERDA Phase II.

I am also grateful to dr David Radford from Oak Ridge National Laboratory for specially implementing the true-coaxial and hemispherical HPGe geometries in the `fieldgen` software, for the purpose of verifying my analytical calculations.

Serdecznie podziękowania dla pana Jana Jurkowskiego oraz mgr. inż. Dariusza Borowicza, od których nauczyłem się wiele na temat technologii wytwarzania detektorów germanowych – znacznie ułatwiło mi to zrozumienie przeprowadzonych pomiarów na potrzeby tej pracy.

Chciałbym szczególnie podziękować moim Rodzicom oraz całej Rodzinie za nieustające wsparcie podczas całych studiów doktoranckich.

This work was supported by the grants: no. UMO-2012/05/E/ST2/02333 (Polish National Science Centre) and no. 7150/E-338/M/2015 (Faculty of Physics, Astronomy and Applied Computer Science of the Jagiellonian University). Polish Ministry of Science and Higher Education, through the cooperation of Polish scientific centers with JINR Dubna, financed from the Polish contribution to the JINR common funds, supported this work as well.

References

[1] Dell'Oro, S. et al. "Neutrinoless double beta decay: 2015 review". In: *Advances in High Energy Physics* 2016 (2016) (cit. on pp. 1, 5–8, 19).

[2] Avignone III, F. T. et al. "Double beta decay, Majorana neutrinos, and neutrino mass". In: *Reviews of Modern Physics* 80.2 (2008), p. 481 (cit. on p. 1).

[3] Abe, S et al. "Precision measurement of neutrino oscillation parameters with KamLAND". In: *Physical Review Letters* 100.22 (2008), p. 221803 (cit. on p. 1).

[4] Gando, A. et al. "Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen". In: *Phys. Rev. Lett.* 117 (8 2016), p. 082503 (cit. on p. 1).

[5] Agostini, M et al. "Results on $\beta\beta$ decay with emission of two neutrinos or Majorons in ^{76}Ge from GERDA Phase I". In: *The European Physical Journal C* 75.9 (2015), p. 416 (cit. on pp. 1, 8).

[6] Pandola, L. "Searching for the neutrinoless double beta decay with GERDA". XV International Conference on Topics in Astroparticle and Underground Physics (TAUP). 2017 (cit. on p. 1).

[7] Agostini, M et al. "Results on Neutrinoless Double-β Decay of ^{76}Ge from Phase I of the GERDA Experiment". In: *Physical Review Letters* 111.12 (2013), p. 122503 (cit. on pp. 1, 13–15, 20, 91).

[8] Panas, K. et al. "Improving Sensitivity of a BEGe-based High-Purity Germanium Spectrometer Through Pulse Shape Analysis". Submitted to European Journal of Physics, C (cit. on pp. 2, 91).

[9] Collaboration, G. et al. "Background-free search for neutrinoless double-β decay of ^{76}Ge with GERDA". In: *Nature* 544.7648 (2017), pp. 47–52 (cit. on pp. 2, 18–21, 84, 91).

[10] Beringer, J. et al. "Review of particle physics particle data group". In: *Physical Review D (Particles, Fields, Gravitation and Cosmology)* 86.1 (2012), p. 010001 (cit. on p. 3).

[11] Klapdor-Kleingrothaus, H. V. et al. "Evidence for neutrinoless double beta decay". In: *Modern Physics Letters A* 16.37 (2001), pp. 2409–2420 (cit. on p. 3).

[12] Schönert, S et al. "The GERmanium Detector Array (GERDA) for the search of neutrinoless $\beta\beta$ decays of 76Ge at LNGS". In: *Nuclear Physics B-Proceedings Supplements* 145 (2005), pp. 242–245 (cit. on p. 3).

[13] Pandola, L. "Searching for the neutrinoless double beta decay with GERDA". Presentation at 15th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2017), Sudburry, Canada. https://indico.cern.ch/event/606690/contributions/2591397/, accessed on 2018-01-21. 2017 (cit. on pp. 3, 21, 91).

[14] Website of the LEGEND experiment: http://legend-exp.org/. Accessed: 2018-01-21 (cit. on p. 3).

[15] Abgrall, N et al. "The large enriched germanium experiment for neutrinoless double beta decay (LEGEND)". In: *AIP Conference Proceedings*. Vol. 1894. 1. AIP Publishing. 2017, p. 020027 (cit. on pp. 3, 91).

[16] Schönert, S. "Results from GERDA and prospects for LEGEND: background-free search for neutrinoless double beta decay of ^{76}Ge". Presentation at XVII International Workshop on Neutrino Telescopes XVII International Workshop on Neutrino Telescopes, Venice, Italy. https://agenda.infn.it/internalPage.py?pageId=0&confId=11857, accessed on 2018-01-21. 2017 (cit. on p. 3).

[17] Giuliani, A. and Poves, A. "Neutrinoless double-beta decay". In: *Advances in High Energy Physics* 2012 (2012) (cit. on p. 4).

[18] Goeppert-Mayer, M. "Double beta-disintegration". In: *Physical Review* 48.6 (1935), p. 512 (cit. on p. 4).

[19] Faessler, A. and Simkovic, F. "Double beta decay". In: *Journal of Physics G: Nuclear and Particle Physics* 24.12 (1998), p. 2139 (cit. on p. 4).

[20] Grotz, K. and Klapdor, H. V. "Predictions of 2ν and 0ν double beta decay rates for nuclei with A$\geq$ 70". In: *Physics Letters B* 157.4 (1985), pp. 242–246 (cit. on p. 4).

[21] Furry, W. H. "On transition probabilities in double beta-disintegration". In: *Physical Review* 56.12 (1939), p. 1184 (cit. on p. 4).

[22] Inghram, M. G. and Reynolds, J. H. "Double beta-decay of ^{130}Te". In: *Physical Review* 78.6 (1950), p. 822 (cit. on p. 5).

[23] Kirsten, T. et al. "Massenspektrometrischer Nachweis von $\beta\beta$-Zerfallsprodukten". In: *Zeitschrift für Physik* 202.1 (1967), pp. 273–292 (cit. on p. 5).

[24] Barabash, A. S. "Experiment double beta decay: historical review of 75 years of research". In: *Physics of Atomic Nuclei* 74.4 (2011), pp. 603–613 (cit. on pp. 5, 6).

[25] Elliott, S. R. et al. "Direct evidence for two-neutrino double-beta decay in ^{82}Se". In: *Physical Review Letters* 59.18 (1987), p. 2020 (cit. on p. 5).

[26] Fiorini, E. et al. "A search for lepton non-conservation in double beta decay with a germanium detector". In: *Physics Letters B* 25 (Nov. 1967), pp. 602–603 (cit. on p. 5).

[27] Caldwell, D. et al. "Recent results from the UCSB/LBL double beta decay experiment". In: *Nuclear Physics B - Proceedings Supplements* 13 (1990), pp. 547 –550. ISSN: 0920-5632 (cit. on pp. 5, 6).

[28] Darken, L. S. and Cox, C. E. "High-Purity Germanium Detectors". In: *Semiconductors and Semimetals* 43 (1995), pp. 23–23 (cit. on pp. 6, 24).

[29] Cremonesi, O. and Pavan, M. "Challenges in double beta decay". In: *Advances in High Energy Physics* 2014 (2014) (cit. on pp. 6, 8).

[30] Mount, B. J. et al. "Double-β-decay Q values of Se 74 and Ge 76". In: *Physical Review C* 81.3 (2010), p. 032501 (cit. on p. 7).

[31] Heusser, G. "Low-radioactivity background techniques". In: *Annual Review of Nuclear and Particle Science* 45.1 (1995), pp. 543–590 (cit. on pp. 7, 8, 17, 63).

[32] Agostini, M et al. "Limits on uranium and thorium bulk content in Gerda Phase I detectors". In: *Astroparticle Physics* 91 (2017), pp. 15–21 (cit. on p. 7).

[33] Barabash, A. "The new generation of double beta decay experiments: are there any limitations?" In: *Journal of Physics G: Nuclear and Particle Physics* 39.8 (2012), p. 085103 (cit. on p. 7).

[34] Gómez-Cadenas, J. and Martín-Albo, J. "Phenomenology of neutrinoless double beta decay". In: *arXiv preprint arXiv:1502.00581* (2015) (cit. on p. 7).

[35] Elliott, S. R. and Vogel, P. "Double beta decay". In: *Annual Review of Nuclear and Particle Science* 52.1 (2002), pp. 115–151 (cit. on p. 8).

[36] Ackermann, K.-H. et al. "The GERDA experiment for the search of $0\nu\beta\beta$ decay in ^{76}Ge". In: *The European Physical Journal C* 73.3 (2013), pp. 1–29 (cit. on pp. 9–12).

[37] Günther, M. et al. “Heidelberg-Moscow $\beta\beta$ experiment with ^{76}Ge: Full setup with five detectors”. In: *Physical Review D* 55.1 (1997), p. 54 (cit. on p. 8).

[38] Bode, T. “The neutrinoless double beta decay experiment GERDA Phase II: A novel ultra-low background contacting technique for germanium detectors and first background data”. An optional note. PhD thesis. TUM: Technische Universität München, July 2016 (cit. on p. 10).

[39] Riboldi, S. et al. “A low-noise charge sensitive preamplifier for Ge spectroscopy operating at cryogenic temperature in the GERDA experiment”. In: *IEEE Nuclear Science Symposium & Medical Imaging Conference*. IEEE. 2010, pp. 1386–1388 (cit. on p. 9).

[40] Aharonov, Y et al. “New laboratory bounds on the stability of the electron”. In: *Physical Review D* 52.7 (1995), p. 3785 (cit. on p. 10).

[41] Klapdor-Kleingrothaus, H. et al. “GENIUS-TF: a test facility for the GENIUS project”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 481.1 (2002), pp. 149–159 (cit. on p. 10).

[42] Lehnert, B. “Search for $2\nu\beta\beta$ Excited State Transitions and HPGe Characterization for Surface Events in GERDA Phase II”. PhD thesis. 2015 (cit. on pp. 11, 16).

[43] Lehnert, B. “Background rejection of n+ surface events in GERDA Phase II”. In: *Journal of Physics: Conference Series*. Vol. 718. 6. IOP Publishing. 2016, pp. 62035–62040 (cit. on pp. 12, 34).

[44] Agostini, M et al. “The background in the $0\nu\beta\beta$ experiment Gerda”. In: *The European Physical Journal C* 74.4 (2014), p. 2764 (cit. on pp. 12, 16, 18, 80, 91).

[45] Agostini, M et al. “Pulse shape discrimination for Gerda Phase I data”. In: *The European Physical Journal C* 73.10 (2013), p. 2583 (cit. on pp. 13, 14, 34, 36, 74, 81, 91).

[46] Lehnert, B. “The GERDA Experiment and the Search for Neutrinoless Double Beta Decay”. In: *Proceedings, 49th Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 15-22, 2014*. 2014, pp. 155–160 (cit. on p. 13).

[47] Budjáš, D. et al. “Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double beta decay experiment”. In: *Journal of Instrumentation* 4.10 (2009), P10007 (cit. on pp. 13, 14, 34).

[48] Klapdor-Kleingrothaus, H. et al. “Search for neutrinoless double beta decay with enriched ^{76}Ge in Gran Sasso 1990–2003”. In: *Physics Letters B* 586.3 (2004), pp. 198–212 (cit. on pp. 15, 16).

[49] Klapdor-Kleingrothaus, H. V. et al. “Latest results from the Heidelberg-Moscow double beta decay experiment”. In: *The European Physical Journal A-Hadrons and Nuclei* 12.2 (2001), pp. 147–154 (cit. on p. 16).

[50] Aalseth, C. et al. “IGEX ^{76}Ge neutrinoless double-beta decay experiment: prospects for next generation experiments”. In: *Physical Review D* 65.9 (2002), p. 092007 (cit. on p. 16).

[51] Agostini, M et al. “Production, characterization and operation of ^{76}Ge enriched BEGe detectors in GERDA”. In: *The European Physical Journal C* 75.2 (2015), p. 39 (cit. on pp. 17, 35).

[52] Agostini, M et al. “LArGe: active background suppression using argon scintillation for the Gerda $0\nu\beta\beta$ experiment”. In: *The European Physical Journal C* 75.10 (2015), p. 506 (cit. on pp. 17, 18).

[53] Agostini, M et al. “Upgrade for Phase II of the GERDA Experiment”. In: *arXiv preprint arXiv:1711.01452* (2017) (cit. on p. 18).

[54] Cennini, P et al. “Detection of scintillation light in coincidence with ionizing tracks in a liquid argon time projection chamber”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 432.2 (1999), pp. 240–248 (cit. on p. 18).

[55] Agostini, M et al. “First results from GERDA Phase II”. In: *Journal of Physics: Conference Series*. Vol. 888. 1. IOP Publishing. 2017, p. 012030 (cit. on p. 19).

[56] Wagner, V. E. “Pulse Shape Analysis for the GERDA Experiment to Set a New Limit on the Half-life of Neutrinoless Double Beta Decay of Ge-76”. PhD thesis. 2017 (cit. on pp. 19, 34, 35, 81).

[57] Agostini, M. et al. “Discovery probability of next-generation neutrinoless double-β decay experiments”. In: *arXiv preprint arXiv:1705.02996* (2017) (cit. on pp. 21, 91).

[58] Knoll, G. *Radiation Detection and Measurement*. John Wiley & Sons, 1989. ISBN: 0-471-81504-7 (cit. on pp. 23, 24, 27).

[59] Pehl, R. and Goulding, F. “Recent observations on the Fano factor in germanium”. In: *Nuclear Instruments and Methods* 81.2 (1970), pp. 329 –330. ISSN: 0029-554X (cit. on p. 23).

[60] Darken, L. and Cox, C. “High-purity germanium technology for gamma-ray and X-ray spectroscopy”. In: *MRS Online Proceedings Library Archive* 302 (1993) (cit. on p. 23).

[61] Looker, Q. *Fabrication process development for high-purity germanium radiation detectors with amorphous semiconductor contacts*. University of California, Berkeley, 2014 (cit. on p. 24).

[62] Bruyneel, B et al. “Pulse shape analysis and position determination in segmented HPGe detectors: The AGATA detector library”. In: *The European Physical Journal A* 52.3 (2016), pp. 1–11 (cit. on pp. 24, 82).

[63] D.C. Radford, Siggen/Fieldgen, Oak Ridge, 2004. URL: http://radware.phy.ornl.gov/gretina/ (cit. on p. 24).

[64] Lutz, G. et al. *Semiconductor radiation detectors*. Vol. 40. Springer, 1999 (cit. on pp. 24–27).

[65] Spieler, H. *Semiconductor detector systems*. Vol. 12. Oxford university press, 2005 (cit. on pp. 24, 27).

[66] He, Z. “Review of the Shockley–Ramo theorem and its application in semiconductor gamma-ray detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 463.1 (2001), pp. 250–267 (cit. on pp. 24, 25).

[67] Agostinelli, S. et al. “GEANT4—a simulation toolkit”. In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303 (cit. on p. 25).

[68] Hirayama, H. et al. *The EGS5 code system*. Tech. rep. Stanford Linear Accelerator Center (SLAC), 2005 (cit. on p. 25).

[69] Vénos, D et al. “Performance of HPGe detectors in the temperature region 2–77 K”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 365.2-3 (1995), pp. 419–423 (cit. on p. 25).

[70] Bruyneel, B. et al. “Characterization of large volume HPGe detectors. Part II: Experimental results”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 569.3 (2006), pp. 774 –789. ISSN: 0168-9002 (cit. on pp. 26, 27).

[71] Gadeken, L. and Robertson, B. “The influence of Ge(Li) detector pulse shape variations on constant fraction and snap-off timing discriminators”. In: *Nuclear Instruments and Methods* 136.2 (1976), pp. 255–259 (cit. on p. 26).

[72] Caughey, D. M. and Thomas, R. E. “Carrier mobilities in silicon empirically related to doping and field”. In: *Proceedings of the IEEE* 55.12 (1967), pp. 2192–2193. ISSN: 0018-9219 (cit. on p. 27).

[73] Mihailescu, L et al. “The influence of anisotropic electron drift velocity on the signal shapes of closed-end HPGe detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 447.3 (2000), pp. 350–360 (cit. on p. 27).

[74] Swope, W. C. et al. “A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters”. In: *The Journal of Chemical Physics* 76.1 (1982), pp. 637–649 (cit. on p. 27).

[75] Birkenbach, B et al. “Determination of space charge distributions in highly segmented large volume HPGe detectors from capacitance–voltage measurements”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 640.1 (2011), pp. 176–184 (cit. on p. 27).

[76] Agostini, M et al. “Signal modeling of high-purity Ge detectors with a small read-out electrode and application to neutrinoless double beta decay search in Ge-76”. In: *Journal of Instrumentation* 6.03 (2011), P03005 (cit. on pp. 28, 30, 34, 66, 80).

[77] Luke, P. et al. “Low capacitance large volume shaped-field germanium detector”. In: *IEEE Transactions on Nuclear Science* 36.1 (1989), pp. 926–930 (cit. on p. 29).

[78] Roth, J. et al. “Segmentation and pulse shape discrimination techniques for rejecting background in germanium detectors”. In: *IEEE Transactions on Nuclear Science* 31.1 (1984), pp. 367–371 (cit. on pp. 32, 59, 60).

[79] Majorovits, B and Klapdor-Kleingrothaus, H. “Digital pulseshape analysis by neural networks for the Heidelberg-Moscow-Double-Beta-Decay-Experiment”. In: *The European Physical Journal A-Hadrons and Nuclei* 6.4 (1999), pp. 463–469 (cit. on pp. 32, 35).

[80] Attix, F. H. *Introduction to radiological physics and radiation dosimetry*. John Wiley & Sons, 2008 (cit. on p. 33).

[81] Deutsch, M. “Gamma-Rays from ^{64}Cu, Annihilation of Swift Positrons, and Experiments on Orbital-Electron Capture”. In: *Physical Review* 72.8 (1947), p. 729 (cit. on p. 33).

[82] Gilmore, G. *Practical gamma-ray spectroscopy*. John Wiley & Sons, 2011 (cit. on pp. 33, 41, 44).

[83] Barbeau, P. et al. “Large-mass ultralow noise germanium detectors: performance and applications in neutrino and astroparticle physics”. In: *Journal of Cosmology and Astroparticle Physics* 2007.09 (2007), p. 009 (cit. on p. 33).

[84] Budjáš, D. “Germanium detector studies in the framework of the GERDA experiment”. PhD thesis. 2009 (cit. on pp. 34, 53).

[85] Orduna, R. G. de et al. “Pulse shape analysis for background reduction in BEge detectors”. In: *EUR 24521EN, European Union* (2010) (cit. on pp. 34, 66, 67).

[86] Mertens, S et al. “MAJORANA Collaboration’s experience with germanium detectors”. In: *Journal of Physics: Conference Series*. Vol. 606. 1. IOP Publishing. 2015, p. 012005 (cit. on pp. 34, 91).

[87] Aguayo, E et al. “Characteristics of signals originating near the lithium-diffused N+ contact of high purity germanium p-type point contact detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 701 (2013), pp. 176–185 (cit. on p. 34).

[88] Gonzalez, D et al. “Pulse-shape discrimination in the IGEX experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 515.3 (2003), pp. 634–643 (cit. on p. 35).

[89] Baudis, L et al. “The Heidelberg-Moscow experiment: improved sensitivity for ^{76}Ge neutrinoless double beta decay”. In: *Physics Letters B* 407.3-4 (1997), pp. 219–224 (cit. on p. 35).

[90] Hoecker, A. et al. “TMVA: Toolkit for Multivariate Data Analysis”. In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039 (cit. on pp. 36, 39, 46, 47, 73, 92).

[91] Brun, R. and Rademakers, F. “ROOT—an object oriented data analysis framework”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1-2 (1997), pp. 81–86 (cit. on pp. 36, 37).

[92] Höcker, A et al. TMVA *Users Guide*. Online, accesed 8.01.2018: https://github.com/root-project/root/blob/v6-10-00/documentation/tmva/UsersGuide/TMVAUsersGuide.pdf. 2018 (cit. on p. 39).

[93] Hult, M. et al. “Underground gamma-ray spectrometry”. In: *Acta Chimica Slovenica* 53.1 (2006), p. 1 (cit. on pp. 41, 59).

[94] Agostini, M et al. “GELATIO: a general framework for modular digital analysis of high-purity Ge detector signals”. In: *Journal of Instrumentation* 6.08 (2011), P08013 (cit. on pp. 43, 70).

[95] Stein, J et al. “X-ray detectors with digitized preamplifiers”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 113.1-4 (1996), pp. 141–145 (cit. on p. 43).

[96] Smith, S. W. et al. “The scientist and engineer’s guide to digital signal processing”. In: (1997) (cit. on p. 44).

[97] Radeka, V. “Low-noise techniques in detectors”. In: *Annual Review of Nuclear and Particle Science* 38.1 (1988), pp. 217–277 (cit. on p. 44).

[98] Wold, S. et al. “Principal component analysis”. In: *Chemometrics and intelligent laboratory systems* 2.1-3 (1987), pp. 37–52 (cit. on pp. 47, 92).

[99] Keogh, E. and Mueen, A. “Curse of Dimensionality”. In: *Encyclopedia of Machine Learning*. Ed. by Sammut, C. and Webb, G. I. Boston, MA: Springer US, 2010, pp. 257–258. ISBN: 978-0-387-30164-8 (cit. on p. 48).

[100] Sturm, K von et al. “A Compton scattering setup for pulse shape discrimination studies in germanium detectors”. In: *Applied Radiation and Isotopes* 125 (2017), pp. 163–168 (cit. on p. 53).

[101] Heider, M. B. et al. “Operation and performance of a bare broad-energy germanium detector in liquid argon”. In: *Journal of Instrumentation* 5.10 (2010), P10007 (cit. on p. 55).

[102] Shafroth, S. M. “Scintillation Spectroscopy of Gamma Radiation, Volume 1”. In: *Scintillation Spectroscopy of Gamma Radiation*. 1967 (cit. on p. 61).

[103] Wood, G. and Jastram, P. “Direction and polarization angular correlations of gamma-ray cascades in ^{208}Pb”. In: *Nuclear Physics* 32 (1962), pp. 411–429 (cit. on p. 62).

[104] Hausser, G. “Cosmic ray-induced background in Ge-spectrometry”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 83.1-2 (1993), pp. 223–228 (cit. on p. 63).

[105] Aalseth, C. and Miley, H. “Developments in gamma-ray spectrometry: systems, software, and methods-II. 6. Digital Pulse-Shape Discrimination for HPGe: Fast Neutron Sensitivity”. In: *Transactions of the American Nuclear Society* 84 (2001), pp. 338–339 (cit. on p. 64).

[106] Andreotti, E et al. “Status of underground radioactivity measurements in HADES”. In: *Proceedings from 3rd International Conference" Current Problems in Nuclear Physics and Atomic Energy", Kiev*. 2010, pp. 7–12 (cit. on p. 64).

[107] GERDA Analysis Group. “Parameters and facts for the analysis dataset of the $0\nu\beta\beta$ decay of GERDA Phase II”. GERDA Scientific/Technical Report GSTR16-007. 2016 (cit. on p. 78).

[108] Kirsch, A. “ANN COAX PSD for Phase II Run 53–64 and for Phase I Run 47 + 49”. Presentation at GERDA Collaboration Meeting in Ringberg. 2016 (cit. on p. 80).

[109] Kirsch, A. “ANN COAX PSD simulation of $0\nu\beta\beta$ efficiency”. Presentation at GERDA Analysis Call. 2016 (cit. on p. 81).

[110] Bauer, M et al. “MaGe: a Monte Carlo framework for the Gerda and Majorana double beta decay experiments”. In: *Journal of Physics: Conference Series*. Vol. 39. 1. IOP Publishing. 2006, p. 362 (cit. on p. 81).

[111] Ponkratenko, O. et al. “Event generator DECAY4 for simulating double-beta processes and decays of radioactive nuclei”. In: *Physics of Atomic Nuclei* 63.7 (2000), pp. 1282–1287 (cit. on p. 81).

[112] Salathe, M. et al. “Study on modified point contact germanium detectors for low background applications”. PhD thesis. Ruprecht-Karls-Universität Heidelberg, 2015 (cit. on p. 82).

[113] Heider, M. B. “Active masses of the Phase I enriched detectors”. GERDA Scientific/Technical Report GSTR-12-014. 2012 (cit. on p. 85).

[114] Volynets, O. “Methods to improve and understand the sensitivity of high purity germanium detectors for searches of rare events”. PhD thesis. Technische Universität München, Max-Planck-Institut für Physik, Germany, 2012 (cit. on p. 85).

[115] Kirsch, A. et al. “Search for the neutrinoless double ß-decay in Gerda Phase I using a Pulse Shape Discrimination technique”. PhD thesis. Ruprecht-Karls-Universität Heidelberg, 2014 (cit. on p. 85).

[116] Heusser, G et al. “Low-level germanium gamma-ray spectrometry at the μBq/kg level and future developments towards higher sensitivity”. In: *Radioactivity in the Environment* 8 (2006), pp. 495–510 (cit. on p. 91).

[117] LeCun, Y. et al. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444 (cit. on p. 92).

[118] Schmidhuber, J. “Deep learning in neural networks: An overview”. In: *Neural networks* 61 (2015), pp. 85–117 (cit. on p. 92).

[119] Holl, P. “Deep Learning Possibilities for GERDA”. Presentation at GERDA Collaboration Meeting in Cracow. 2017 (cit. on p. 92).

[120] Jones, E. et al. “SciPy: open source scientific tools for Python”. SciPy library webpage: https://www.scipy.org/. Accessed: 2018-01-21. 2014 (cit. on p. 104).