**THE EUROPEAN
PHYSICAL JOURNAL C**

CrossMark

Regular Article - Experimental Physics

Improving sensitivity of a BEGe-based high-purity germanium spectrometer through pulse shape analysis

M. Misiaszek[1], K. Panas[1,a] 🔵, M. Wojcik[1], G. Zuzel[1], M. Hult[2]

[1] M. Smoluchowski Institute of Physics, Jagiellonian University, Lojasiewicza 11, 30-348 Kraków, Poland
[2] European Commission, DG JRC, JRC-Geel, Retieseweg 111, 2440 Geel, Belgium

Abstract We performed Pulse Shape Analysis to separate single-scattered gamma energy deposition events from multiple-scattered photons in a high-sensitivity γ-ray spectrometer. The spectrometer is based on a Broad Energy High Purity Germanium detector and the developed technique uses multivariate analysis by an application of the Multi-Layer Perceptron Neural Network. A very good separation of the single-site- and multi-site events was achieved leading to a significant reduction of the background level of the investigated spectrometer – the double escape peak, rich in single-site events, was reduced by 95%, while the full energy peaks lost at most 25% of their counts. The peak to Compton ratio, calculated for the 2614.5 keV gamma line from ^{208}Tl, was improved by 114.3%.

1 Introduction

For germanium semiconductor detectors a single energy deposition by γ-ray is often described as a single-site event (SSE). Multi-site events (MSEs), like e.g. multiple Compton scatterings, have several interaction sites separated by a distance of about 1 cm. The differences and discrimination between SSEs and MSEs is of primary interest for experiments like GERDA [1,2] looking for the neutrinoless double beta ($0\nu\beta\beta$) decay with application of the High Purity Germanium (HPGe) detectors. Events from the hypothetical $0\nu\beta\beta$ decay would ionize the detector's active volume by means of two electrons (with a range of less than 1 mm in germanium) and therefore belong to the first category. The background due to external γ-rays is typically of multi-site type, because γ-rays with energies in the range of ~ 1 MeV deposit their energy mainly via multiple Compton scattering with a mean free path of a few centimeters.

For applications in γ-ray spectrometry the situation is reversed: SSEs should be rejected and MSEs preserved. The SSEs are of background type, because they are mostly single Compton scattered events with the scattered photon escaping the crystal, while the full energy peaks (FEPs) registered by the detector (and used to evaluate the activities of radionuclides) contain mainly MSEs.[1] The developed procedure shall therefore make peaks more distinguished by reducing the flat Compton continuum and make possible to evaluate peaks from radioisotopes, which would be otherwise under spectrometer's minimal detectable activity (sensitivity).

Discrimination between SSEs and MSEs for the Broad Energy Germanium (BEGe) detectors was extensively studied in the frame of GERDA experiment [3–7], also with respect to applications in ultra-low background γ-ray spectrometry [8]. The approach that was worked out was to look for a ratio of the maximal current signal amplitude (A) to the corresponding event energy (E) – the so-called A/E cut. This is a one-parameter method, where the cut value is determined according to the calibration data, obtained usually by irradiating the spectrometer with a ^{228}Th source (due to the emitted high energy γ-ray from the decay of ^{208}Tl – a more detailed explanation can be found in Sect. 3.2). Measurements carried out for various BEGe detectors demonstrated very good performance of the A/E method, however, its applications to the coaxial detectors did not provide satisfactory results.

The main goal of our work was to find an alternative, efficient and stable Pulse Shape Analysis (PSA) procedure to distinguish between MSEs and SSEs in BEGe type detectors. As it will be shown, our method does not need any corrections like e.g. the ones related to the dependence of the A/E classi-

---

a e-mail: krzysztof.panas@doctoral.uj.edu.pl

[1] For typical detectors used nowadays in γ-ray spectroscopy (crystal volume of several hundreds of cm^3), and for energies of some hundred keV and higher, the full energy peaks contain mostly events, which were caused by a few Compton scatterings followed by the photoelectric effect.

🌱 Springer
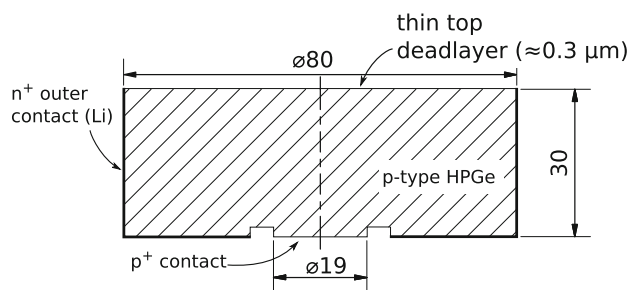
Fig. 1 Schematic drawing of the 50% relative efficiency BEGe detector applied in the Ge-5 spectrometer and used in the presented study

fier on energy [5]. To achieve this we applied methods implemented in the Toolkit for Multivariate Analysis (TMVA) [9]. TMVA provides a ROOT-integrated machine-learning environment for the processing and parallel evaluation of multivariate classification and regression techniques [10]. TMVA includes classifiers like Projective Likelihood (PL), Multi-Layer Perceptron Neural Network (MLP, see Sect. 4.1 for more details), Boosted Decision Trees (BDT) and Support Vector Machine (SVM). Each of the implemented methods provides training, testing, performance evaluation algorithms and visualization scripts. The training and testing is performed with the use of user-supplied data sets in the form of ROOT trees or text files.

2 Experimental setup

To study the new PSA technique we collected data using the BEGe-based γ-ray spectrometer (Ge-5) operated by the JRC-Geel (formerly Institute for Reference Materials and Measurements) in the HADES underground laboratory. The rock overburden of 500 m water equivalent provides attenuation of the muon flux by about four orders of magnitude [11]. The spectrometer is based on a 50% relative efficiency p-type BEGe diode installed in a standard vacuum cryostat produced by Canberra (model BE5030). A massive low-radioactivity shield surrounds the detector in order to reduce the environmental background and makes the spectrometer sensitive to very weak radioactivity. A sketch of the Ge-5 crystal is shown in Fig. 1.

To train the TMVA-based methods and to evaluate their performance, apart from the energy information (as obtained usually from the system based on multi channel analyzers), the waveforms from the preamplifier have to also be registered. They were acquired with the Struck SIS3302, a 16-bit 100 MHz Flash Analog to Digital Converter (FADC) – the total length of each waveform was 30 μs. Noise and events with dubious quality (trigger position much more before/after pre-trigger, pile-upped events with multiple slopes) were excluded form further analysis. For pulses, which survived
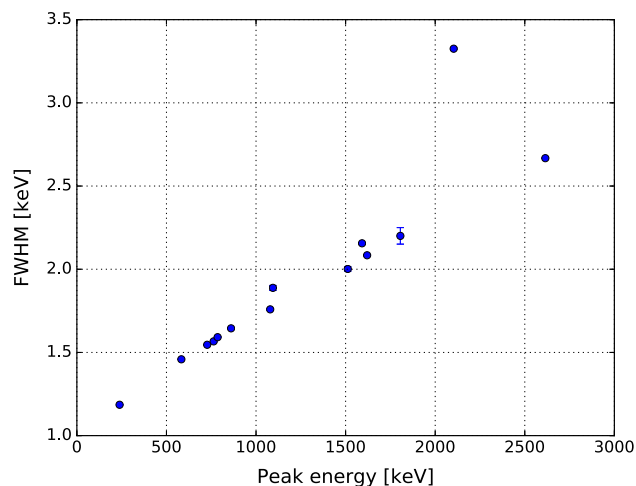

Fig. 2 Energy resolution of the peaks in the ^{228}Th spectrum. The outlying peak at 2103 keV is a single escape peak from 2615 keV ^{208}Tl line. Its abnormally high FWMH value is due to the Doppler broadening [12]

the quality cuts, the first step was to perform the energy reconstruction. The energy value for each pulse was determined using the trapezoidal filter described in [13], followed by a calibration using a linear function. The energy resolution values for the selected peaks from the ^{228}Th decay chain spectrum are plotted in Fig. 2.

3 Pulse shape analysis

The data features that are used to train machine learning models have a huge influence on the final performance. In our case of the PSA, employing neural networks procedure to the ^{228}Th data, the task of feature selection was limited to a vector of the sampled waveforms of the preamplifier output voltage. After series of dedicated tests, we found that considering only samples close to the point of the maximal current as the input parameters give us substantial discrimination effect. The shape of the pulse in the range of limited samples manifests the type of interaction in the detector due to the differences in speed and possible fluctuations of the charge collection over time.

In the first step, after initial data pre-processing and energy reconstruction, we were performing pulse shape normalization. The sampled amplitudes were divided by the corresponding reconstructed energy values in order to remove potential energy dependency of the signal. Next, normalized input variables were extracted from each rising edge and the pulses (edges) were digitally differentiated and smoothed in order to find the moment of the maximal values of the current signals. 31 single amplitudes in total were chosen for each pulse for the PSA: 15 before, 1 at and 15 after the moment
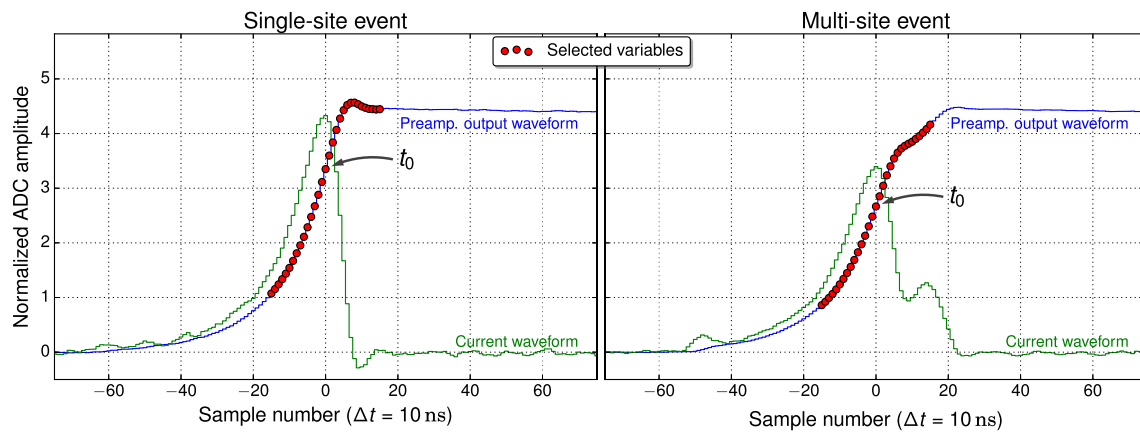
Fig. 3 Pulse shapes from typical single-site (left) and multi-site (right) events. The current pulses (in green) were obtained by digital differentiation of the preamplifier output (in blue) and smoothing with a moving average filter (width of 50 ns). Red points represent the amplitudes selected to perform the PSA. t_0 indicates the amplitude, which corresponds to the maximal current of the pulse

corresponding to the maximal current (time t_0). The extraction procedure is illustrated in Fig. 3.

In the course of this study we also found that the 31 samples can be divided into 4 groups and consecutively summed together. This is most likely due to the fact that sampling rate of the used FADC is $4\times$ faster than preamplifier bandwidth. The operation of samples summing also effectively reduces the dimensionality of the problem (from 31 dimension to 4), which in turn decreases the level of the classifier complexity (e.g. number of neurons in the MLP). Lower complexity classifiers require less computation time and also are less prone to overtraining.

3.1 Multi-layer perceptron neural network

In the course of this study we found that a method based on an MLP neural network has shown the best separation efficiency. Many detailed descriptions of this kind of neural network are available in the literature (e.g. [9,14]), a short summary will be provided here for a better understanding of the PSD method.

A conceptual drawing of an MLP neural network is shown in Fig. 4 (the so-called "bias neurons" were not drawn). The network consists of several layers. The first one is an input layer – in the described method its role was to normalize (values between -1 and 1) each of the input variables. A simple linear transformation was used for this purpose (denoted with a straight line in a circle in Fig. 4).

The normalized values ($y_1, y_2, \dots y_n$, where n is a number of input variables) are then "fed" to the next layer. This is the so-called "hidden layer" and this is where the classification process really takes place. Each connection between the neurons in the input and the hidden layers has a number w_{ij}^k associated with it, called the "weight" (i is an index of neu-

ron in the previous layer, j is a similar index, but in a current layer, and k is the index of the previous layer). Additionally, each neuron is characterized with a (usually) non-linear function (called the "activation function"). To calculate the output value of a given neuron, we consider all connections between neurons from the previous layer, where the output values are already calculated (e.g. the normalized values from the input layer). Then, each weight is multiplied by the previous layer output values and the products are summed together. After this operation, the activation function is applied to the sum (in this case it was hyperbolic tangent). The single neuron response is summarized in the inset in Fig. 4.

It should be mentioned that in principle more hidden layers can be used. However, the Stone–Weierstrass theorem states that the feedforward perceptron neural network can approximate any non-linear function (the precision depends on a number of neurons used in the layer) using just one hidden layer [9]. The only requirement is that the neuron activation should be non-linear itself [15].

The last step in the response calculation takes place in the output layer. The principle of a neuron response calculation is the same as in previous layers. In this specific case (the neural network used in this work), the activation function is a sigmoid, in contrast to the hyperbolic tangent functions from the hidden layer. Since the output of the sigmoid can only take values between 0 to 1, the response is normalized to this range.

For the training process of the neural network, sets of background and signal events must be selected. Their purpose is to set the neuron values in the network in such way that after classification of the training events the network will output the value close to 1 for events from signal set and 0 for their background counterparts.
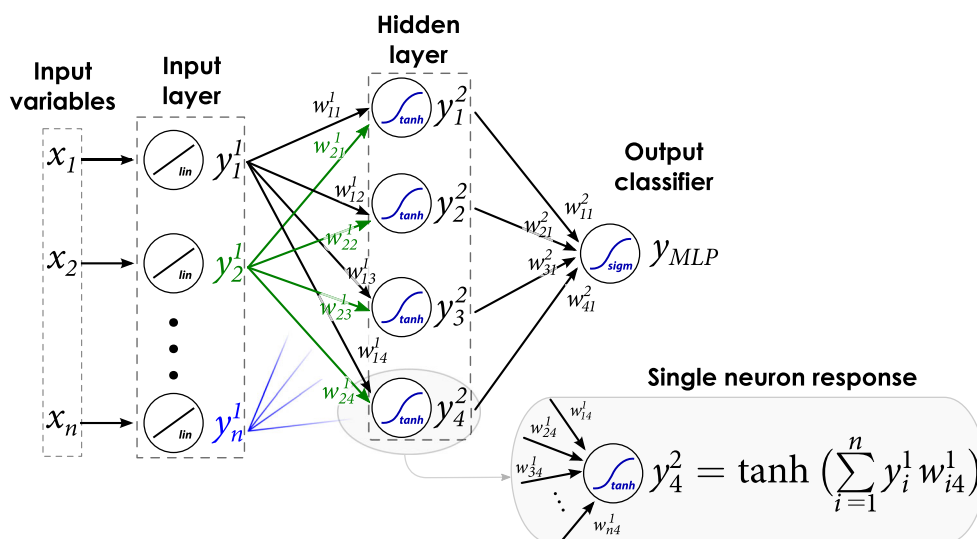
Fig. 4 Schematic drawing of multi-layer perceptron neural network (bias neurons were not shown for simplicity). The inset shows how a single neuron response is calculated

3.2 Topologies of the events in the HPGe detector

As mentioned in the previous section, samples of the signal and the background events are needed for the training procedure. In this case, they were obtained from a different energy regions of the ^{228}Th chain spectrum. It was acquired using a strong (~ 100 kBq) source, which was placed inside the measurement chamber, about 20 cm above the detector's end-cap. ^{228}Th was selected due to its high energy (2614.5 keV) γ line from the ^{208}Tl decay (one of the last daughters in the chain). The high energy of the γ-ray is needed, since otherwise single escape and double escape peaks (SEP and DEP, respectively) would not be observed. As it will be explained shortly, they are crucial for the PSA efficiency evaluation.

All peak types (SEP, DEP and FEP), Compton edge (at 2386 keV) and the multiple Compton scattering region (between the Compton edge and the 2614.5 keV FEP) are of special interest in the PSA studies because they represent different events topologies inside the detector – they are schematically shown in Fig. 5. If the γ-ray is absorbed in the detector through a electron-positron pair production mechanism, both 511 keV γ-rays produced in the positron annihilation have to be absorbed to observe a full energy deposition. If one or two of the annihilation γ-rays escape from the active detector volume, they give rise to the SEP or to the DEP, respectively. Detector events giving rise to the DEP are characterized by the energy deposition in a small volume, which size is typically 1 mm^3 for germanium [8] - it is much smaller than detector's dimensions. Therefore, DEP events represent the SSEs. On the other hand, events

from SEPs and FEPs[2] require energy deposition in multiple vertices in the detector, thus, they are proxies for the MSEs.

The reason that the MSEs and SSEs can be distinguished by the analysis of the preamplifier signal is the weighting field distribution in the detector [6]. The biggest contribution to the current signal amplitude comes from the charge carriers (holes – electrons drift towards the n$^+$ HV electrode) drifting near the p$^+$ contact electrode (see Fig. 1). In the first approximation, we can assume that current peak (which can be seen on the left panel of Fig. 3) is induced only when the hole arrives in the vicinity of the p$^+$ electrode.

Let's now analyse what happens if there are many interaction points (e.g. from the scattered γ-ray). In a simple example, we can assume that the first one takes place near the p$^+$ contact and the other one near the top right/left edge of the detector. In this case, a distance between interaction sites is about 5 cm (dimensions taken from Fig. 1). Since the carrier drift velocity is close to saturation for the electric fields present in a typical germanium detector and is about 10 cm/μs [16], the drift time of the holes would be about 500 ns. Therefore, the two current peaks (superposition of current signals from both sites) will be separated by a considered delay (which can be seen in the current trace on the right panel of Fig. 3). This property of the electric field inside a BEGe detector allows to detect a strong pulse shape signature of MSEs, since the interaction vertices are usually spaced by a few centimeters. The two categories (SSEs and MSEs) can be distinguished by their pulse shapes observed at the preamplifier's output. Since HPGe detectors typically

---

[2] This is true for the γ-rays with an energy greater than ≈ 500 keV or so, for which the cross-section for the full energy deposition by the photoelectric effect is very low.
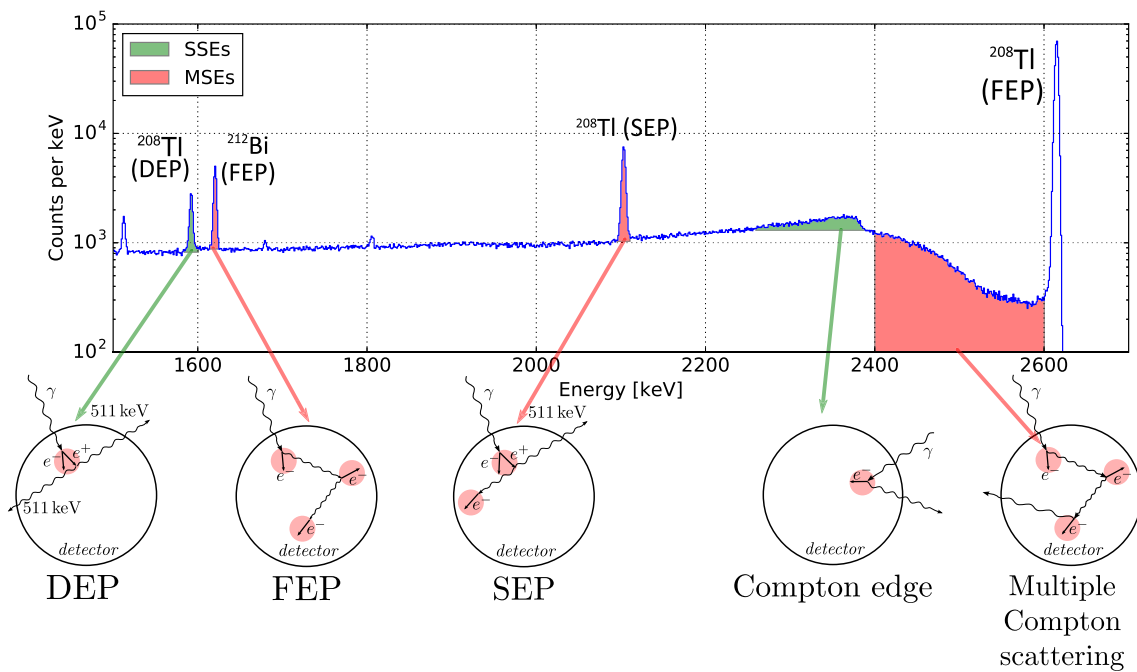
Fig. 5 ^{228}Th energy spectrum with the marked energy regions rich in single- (green) and multi-site events (red). Event topology characteristic for each energy region is sketched below. All peaks and the Compton edge region have always a small admixture of events of opposite type (e.g. DEP has an admixture of MSEs and FEPs have a SSEs component). Practically pure sample of MSEs may be selected only from the multiple Compton scattering region

use charge-sensitive preamplifiers, their output is simply an integrated current pulse (in a time scale of a few hundred ns).

3.3 Analysis of the ^{228}Th calibration data

After the reconstruction of the events in the ^{228}Th spectrum and having the corresponding event topologies in mind, one can distinguish several energy regions, rich either in SSEs or MSEs (see also Fig. 5), which can be used in the training process of the multivariate classifiers:

1. DEP, SEP and FEPs – peaks contain high density of events mostly of a single type, either SSEs (DEP) or MSEs (FEPs, SEP) due to the reasons described in the previous section.
2. Compton edge (CE) region at 2300–2375 keV is a good candidate for the SSEs sample. The Compton continuum from the single scattered γ-rays contains solely SSEs. Furthermore, its shape indicates that the density of events is higher in the CE region compared to the rest of the continuum [12]. If other background components are present in the spectrum (e.g. flat background from multiple scattered photons with the initial energy of 2615 keV – MSEs), the most favourable ratio of SSEs to MSEs (SSE "purity") is therefore in the CE region. The purity is not generally as good as in the case of the DEP, but on the other hand, the number of events available for training is larger.
3. Multiple Compton scattering region (MCS) between 2400 and 2600 keV. The events this energy range can result only from multiple scattered photons, they are therefore almost perfectly pure sample of the MSEs. Other types of events are possible here only due to muon or neutron interactions. SSEs with such energies are not possible due to the kinematic constraints of the Compton scattering mechanism if the largest FEP present in the spectrum is at 2614.5 keV, or lower.

The considered training variants are therefore, the following:

1. DEP/FEP – events from the peaks at 1592 and 1620 keV. The events considered in the analysis come only from the peak region within $\pm \frac{1}{2}$ FWHM on either side.
2. DEP/MCS – the sample from FEP is replaced with events from the 2450–2600 keV region.
3. CE/MCS – the DEP sample is replaced by events from the energy window of 2350–2370 keV, the sample of MSEs is as in the previous variant.
4. CE/FEP – variants 1 and 3 mixed together.

After the training of MLP neural network, we classified all available events – this means that every event in the spectrum
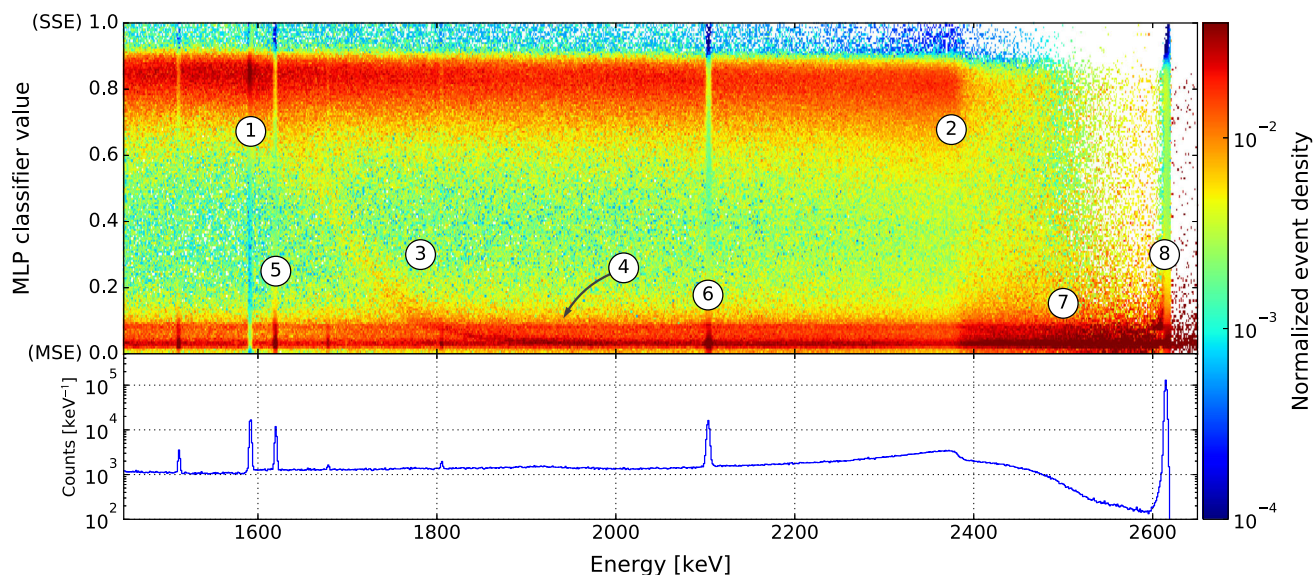
Fig. 6 2D histogram (top panel) showing the distribution of the MLP classifier value (vertical axis) vs. energy (horizontal axis) for the data from ^{228}Th decay chain. The MLP neural network was trained using DEP and FEP training data (see text for details). The energy spectrum corresponding to the upper histogram is shown in the bottom panel. The 2D histogram shows clearly separated bands corresponding to SSEs (MLP close to 0.8) and MSEs (MLP close to 0.05). Several features may be noticed on the scatter plot: "1" – DEP, "2" – Compton edge, "3" – scattered annihilation γ-rays, "4" – Compton edge of the 511 keV annihilation γ-rays, "5" – ^{212}Bi FEP, "6" – SEP, "7" – multiple Compton scattering region, "8" – ^{208}Tl FEP

contains its own classifier value, a number between 0 and 1. The results of the analysis can be presented on an events density plot, which is a 2D histogram, with the classifier value on the y-axis and the event energy on the x-axis (Fig. 6 on the top, on the bottom the corresponding energy spectrum is shown). The scatter plot is normalized with respect to the vertical axis such that the integral of the event density for each bin column is equal to 1. Since the normalization removes the effect of different peak intensities it is done for the visualization purposes only – it does not affect the efficiency results. Generally, the events with the MLP values close to 1 are of single-site type, while those with the MLP values close to 0 are of the multi-site type. Probably the most interesting feature from the background reduction point of view is a clear separation of the continuous background bands.

The most distinctive features of the ^{228}Th 2D histogram are marked with numbers in Fig 6. SSEs, like the ones from DEP ("1"), lie in the SSE band with the classifier value ≈ 0.8, while MSEs, like the ones from FEPs ("5" and "8") or ^{208}Tl SEP ("6"), are localized in the MSE band with MLP ≈ 0.05. Other interesting features are: separated CE (around 2385 keV, "2") and the MCS region (energy value over 2385 keV and lower than the energy of ^{228}Tl line at 2614.5 keV - "7"). The Compton edge of the annihilation gammas (DEP energy – 1592 keV – summed up with the energy of the electron produced by the Compton backscattering of the 511 keV gamma – 340 keV) is marked with "4".

There is also a diagonal band with energy just above the DEP ("3") with a higher density of events – their topology composes of pair production and the scattering or absorption of the 511 keV annihilation quanta. The events are transitional between SSE- and MSE-like, their classifier value decreasing with the energy can be explained by the increasing energy contribution of the second interaction point. For the scattering with a low energy deposition, the contribution to the pulse shape is negligible and the events receive similar classification like the SSEs. On the other hand, the amplitude contribution of events with energy deposition of a few hundred keV is significantly larger. Furthermore, the mean free path is also significantly larger (order of a few cm) – it means that the distance between interaction points is also longer and the observed delay (as explained in the previous section) is more pronounced.

For the rejection of single Compton scattered background events one can make a cut at e.g. MLP < 0.4 – this effectively removes events from the whole SSEs band. Since FEPs are strongly separated, their counts loss is negligible. Simultaneously, most of the SSEs from the Compton continuum can be rejected. The effect of the PSA cut on the energy spectrum is shown in Fig. 7. For completeness, we also included the inverted cut (only events with MLP value *larger* than the cut value), for which the MSEs are suppressed. This mode is of interest in $0\nu\beta\beta$ decay searches, where, as mentioned before, the $0\nu\beta\beta$ decay signal is of the single-site type.
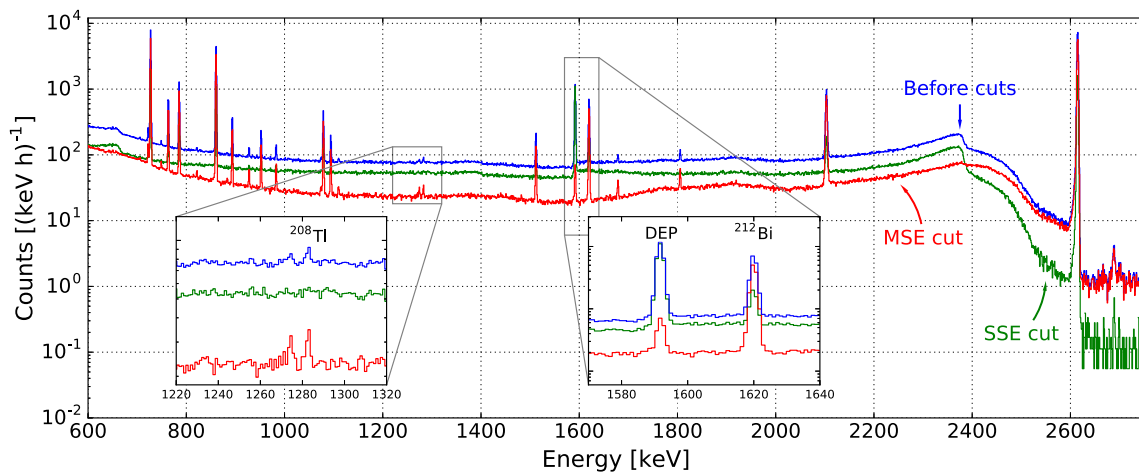
Fig. 7 Effect of the PSA cut on the ^{228}Th spectrum (DEP/FEP training variant). The exact values for the SSE cut (rejecting MSE) and MSE cut (rejecting SSE) are described in the text. For the SSE cut, the DEP is left almost intact and the Compton edge is much more distinctive. On the other hand, the MSE cut reject Compton continuum background, improving detectability of weak FEPs (e.g. the 1282 keV ^{208}Tl peak)

Table 1 Reduction of peaks area and Figure of Merit improvement after application of the MSE PSA cut (for each training variant the DEP events acceptance was set to 5%)

Peak energy [keV]	Isotope	Acceptance [%]				FoM improvement [%]			
		MLP (DEP/FEP)	MLP (DEP/MCS)	MLP (CE/MCS)	MLP (CE/FEP)	MLP (DEP/FEP)	MLP (DEP/MCS)	MLP (CE/MCS)	MLP (CE/FEP)
727.3	^{212}Bi	74.9	74.4	74.2	70.8	5.9	4.7	4.2	5.4
763	^{208}Tl	75.9	75.1	74.9	71.8	16.0	14.2	13.6	15.9
785	^{212}Bi	75.9	74.9	74.6	71.5	17.5	15.3	14.7	17.0
860.6	^{208}Tl	76.3	75.2	74.9	72.1	23.2	21.0	20.3	22.4
1078	^{212}Bi	76.7	76.0	75.1	71.8	32.9	30.4	29.3	33.5
1093	^{208}Tl	79.7	79.1	78.6	75.7	42.3	40.0	39.7	46.1
1512	^{212}Bi	79.4	78.2	77.7	72.5	44.7	41.3	41.1	44.6
1592	^{208}Tl (DEP)	4.9	4.9	4.8	4.9	−90.6	−90.7	−90.8	−89.6
1620.5	^{212}Bi	77.7	77.1	75.3	69.1	49.6	44.4	42.8	46.7
1805	^{212}Bi	83.8	82.9	80.9	74.2	37.6	33.4	32.1	41.1
2103	^{208}Tl (SEP)	84.8	84.4	82.5	76.1	33.5	30.0	29.5	36.3
2614.5	^{208}Tl	77.3	78.4	75.6	62.3	−15.1	−15.0	−17.6	−27.1

To quantify the efficiency of our method, we calculated the peaks area reduction, defined as a ratio of the area before and after the PSA cut. The peaks were fitted with the Gaussian curve with a linear function to subtract the background underneath. In every variant the cut was set to obtain 95% reduction of the DEP events (e.g. this requirement translated to cut value of MLP < 0.28 for the DEP/FEP training variant). The efficiency results on FEPs are summarized in Table 1. For the sake of completeness we also included the acceptance values of the Compton continuum parts of the spectrum, collected in Table 2.

The biggest acceptance of the FEPs is obtained for the first training variant, with DEP events taken as the signal sample and events from the 1620 keV FEP treated as the background sample. Acceptances as high as ≈ 80% can be obtained for FEPs with the energies above 1000 keV. Comparable performance is also obtained for the variant with the DEP/MCS training set. Variants using CE instead of DEP as SSEs sample reduce FEPs more, which is not desirable from the γ spectrometry point of view, since the detection efficiency is reduced as well.

The above results take into account only the peak reduction for an arbitrarily set PSA cut (5% acceptance of the DEP). However, one should have in mind both FEPs reduction (as small as possible) and the Compton continuum suppression (as large as possible). To take both effects into

Table 2 Reduction of background continuum in the selected energy ranges (^{228}Th spectrum) after applying MSE PSA cut (DEP events acceptance set to 5% for all training scenarios)

Energy range [keV]	Acceptance [%]			
	MLP (DEP/FEP)	MLP (DEP/MCS)	MLP (CE/MCS)	MLP (CE/FEP)
630–660	44.3	45.2	46.1	39.7
990–1065	32.4	33.0	32.9	28.2
1140–1500	31.0	31.9	31.5	25.9
1525–1575	29.8	30.7	30.3	25.3
1635–1700	29.4	32.3	31.6	22.4
2010–2085	37.4	39.1	37.9	28.9
2110–2235	40.2	42.3	40.6	30.4
2300–2385	37.0	39.1	37.0	26.8
2400–2500	63.8	67.3	64.6	45.7
2500–2600	79.9	83.1	81.3	58.9

account, we calculated a single parameter called the Figure of Merit (FoM), described in detail in [17]:

$$\text{FoM} = \frac{\varepsilon(E)}{\sqrt{R(E)B(E)}} \qquad (1)$$

where:

- E – γ-ray energy,
- ε – detection efficiency,
- R – energy resolution,
- B – background level.

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

$$\frac{\text{FoM}_{\text{PSA}}}{\text{FoM}} = \frac{\varepsilon_{\text{PSA}}(E)}{\varepsilon(E)} \sqrt{\frac{B(E)}{B_{\text{PSA}}(E)}} = A_{\text{PSA}}(E) \sqrt{\frac{B(E)}{B_{\text{PSA}}(E)}}. \qquad (2)$$

An improvement of FoM (FoM$_{impr}$) after the application of the cut may be expressed as follows:

$$\text{FoM}_{impr}[\%] = \left(\frac{\text{FoM}_{\text{PSA}}}{\text{FoM}} - 1\right)100\%. \qquad (3)$$

Interestingly, the best FoM improvement (right part of Table 1) was observed for the CE/FEP training variant (at least for most peaks). This means that even with the training sample containing some MSEs (CE), the single-site Compton continuum events are recognized better than when using DEP events (less MSEs). Since the FEPs acceptances are actually worse for this variant (left part of Table 1), this means in order to achieve a better FoM improvement more of the Compton

continuum events must be removed (when compared with DEP/FEP training variant).

We have also calculated the peak to Compton ratio (PtC) for the 2614.5 keV ^{208}Tl FEP line, after application of the MSE cut. According to the IEEE standard [18] the PtC is defined as the ratio of counts in the channel corresponding to the maximum of the 1332.5 keV peak of ^{60}Co to the average number of counts of the Compton continuum, in the energy range of 1040–1096 keV. Since we did not use ^{60}Co source in our measurements, we need to rescale the energy ranges for the ^{208}Tl FEP (by the 2614.5/1332.5 ratio). The Compton continuum range will hence become 2040.5–2150.5 keV. However, a simply prescaled region will contain a SEP at 2103.5 keV. The range had to be therefore shifted to the left side of the SEP in order to contain only the flat part of the Continuum. We kept the width of the region the same and finally the counts were averaged in the 2000–2090 keV energy range. Before the cut the PtC was equal to 97.5 and after application of the cut on MLP classifier (DEP/FEP training variant, 5% DEP events acceptance) it was increased to 208.9. This means an improvement by 114.3%. It should be kept in mind that the absolute values may be far off when compared to PtC calculated according to IEEE standard (i.e. for ^{60}Co peak), but we were interested in the relative effect on the peak height/Compton continuum ratio in our measurement.

3.4 Stability of the cut

Depending of the type of the cut, various features of the energy spectrum are suppressed or amplified. Figure 7 visualizes the effect of the PSA cuts, set for either 5 or 95% of the DEP acceptance (SSE cut: MLP > 0.28, MSE cut: MLP < 0.28). One can notice that e.g. the SSE cut makes the Compton edge much more distinctive, while removing FEPs. On the other hand, the MSE cut reduces the DEP, which is mostly full of single-site like events, while the low-intensity peak from ^{208}Tl at 1282 keV is more visible.

The acceptance values are not particularly sensitive to the exact cut value, as long as the latter is in the "valley" of the classifier distribution (between the bands visible on the scatter plot in Fig. 6). Figure 8 illustrates the effect of the different MSE cut values on acceptances of selected peaks. The large suppression of the DEP takes place for the cut value of about ≈ 0.15, what is just above the lower band in Fig. 6. FEP and SEP still retain more than 70% of their area. Moving the cut from 0.28 to 0.58 changes the DEP acceptance only by 5% (from 5 to 10%, respectively). We can therefore conclude that the acceptance values are quite stable as long as the cut value does not lie in neither band.

3.5 Background data analysis

We also performed a measurement of the detector's background (19.7 d) and the weight files obtained by the DEP/FEP
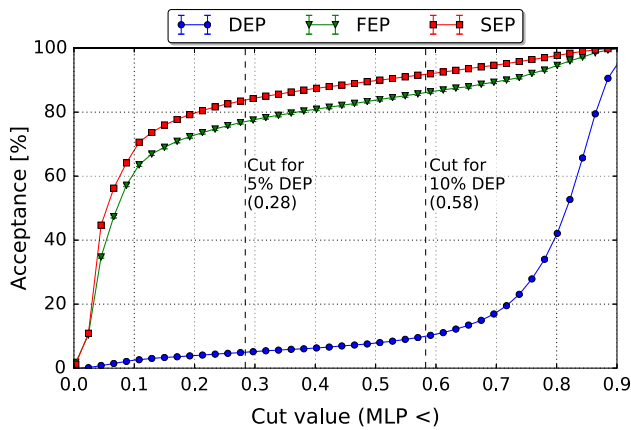
Fig. 8 Acceptance of the DEP (1592 keV), ^{212}Bi FEP (1620 keV) and SEP (2103 keV) events as a function of the MSE PSA cut (MLP smaller than a given value, DEP/FEP training scenario). Two dashed vertical lines show the cut values for 5 and 10% DEP events acceptance

training with ^{228}Th data were applied to classify the background events. Figure 9 shows the effect of the cut on the background energy spectrum. The cuts applied are the same as the ones in the Fig. 7. Apart from the 511 keV annihilation line, only two background peaks are visible: 1460.8 keV from ^{40}K and 2614.5 keV from ^{208}Tl. The noticeable effect of the MSE PSA cut is that the peaks were hardly affected, while the SSE cut almost completely suppressed them. We estimated the acceptances in the ^{40}K and ^{208}Tl lines for the MSE cut to be about 85 and 90%, respectively (due to the low statistics we can only provide approximated values). Thus, the PSA efficiency is the same for the peaks in the calibration and for the background spectra. Due to the very low background of the detector a longer measurement with a weak sample (source) would be required to determine the cut efficiencies with higher statistics.
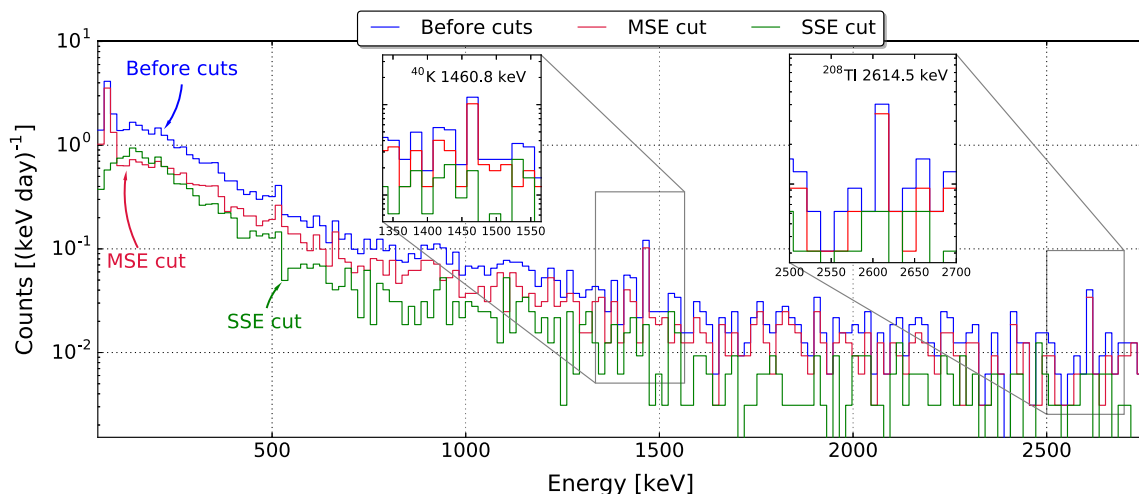
4 Discussion

4.1 Classifiers available in the TMVA package

From the classifiers available in the TMVA package, apart from the MLP, we also studied the Projective Likelihood and Boosted Decision Trees. It turned out that the performance of the Projective Likelihood classifier is subpar when compared to the MLP – the obtained acceptance of the SEP was 21.2% for 90% DEP acceptance. In the case of the Boosted Decision Trees, the performance is actually similar to the MLP, however, the classification is more demanding numerically, which results in a longer computation time. This is due to the large number of trees needed to perform a proper classification.

4.2 Comparison with other methods

The presented method was compared with the previously used A/E PSA, described in [8]. The analysis was performed using the same detector (Ge-5, underground laboratory HADES). To obtain 91% acceptance of the DEP, allowing for a direct comparison, we moved the cut to MLP > 0.59. The achieved reduction factors are compared in Table 3. For the same acceptance of the DEP (91%), our method rejects more events from the FEPs and SEPs in the SSE cut mode. This can be explained by a clearly better separation of the SSE and MSE bands using the MLP classifier. Also, the classifier does not have the energy dependence, which is the case for the A/E classifier, where a linear correction has to be applied [5]. This makes the analysis process simpler. Another difference is related to the presence of bands in the MLP classifier distribution (top panel of Fig. 6). In the A/E method the histogram has only one band, corresponding to the SSEs [8]. To achieve

Fig. 9 The effect of the PSA cut on the background spectrum. The only visible peaks in the spectrum, i.e. the 1460.8 keV peak from ^{40}K and 2614.5 keV from ^{208}Tl are shown in detail

Table 3 Acceptance of the selected peaks from the ^{228}Th spectrum in comparison with A/E method from [8] (cut applied: MLP value > 0.56, DEP events acceptance at $\approx 91\%$, DEP/FEP training scenario)

Peak energy [keV]	Isotope	Acceptance [%] (this work)	Acceptance [%] (A/E [8])
1592.5	^{208}Tl (DEP)	91.1	91
1620.5	^{212}Bi	14.8	24
2103.5	^{208}Tl (SEP)	8.7	19
2614.5	^{208}Tl	12.5	31

the desired efficiency, the A/E cut has to be made on the edge of the band. The consequence is that the cut is very sensitive to e.g. shifts of the bands or numerical instabilities. With this respect, the presence of two bands is desirable – the stability of the MLP cut was discussed in the previous sections.

It should be also pointed out that in [8] the detector was irradiated with a ^{232}Th source instead of ^{228}Th, what caused presence of other peaks in the spectrum. In particular, there is another peak near the DEP, namely at 1588 keV from ^{228}Ac and their interference might have been a source of some uncertainties in the evaluation of the A/E method.

A different method uses a single parameter based on the combination of rise-times of the signal to reduce the background in BEGe-based spectrometers [19]. However, the obtained results (e.g. PtC improvement) are calculated for the data from ^{60}Co radioactive source and cannot be directly compared with the presented results in this work, based on the ^{228}Th data.

5 Conclusions

The presented PSA method shows very good results with an application to a BEGe-based high-sensitivity HPGe spectrometer. Due to the high non-uniformity of the weighting field inside the detector, SSEs and MSEs can be distinguished with high efficiency. The multivariate algorithm was trained with the pulse shape data resulted from irradiation of the detector with a ^{228}Th source. The efficiency of the reduction of the Compton continuum was assessed and resulted in a significant improvement of the sensitivity of the investigated low-background γ-ray spectrometer. The presented method may be applied to other BEGe-based spectrometers, improving their performance only by changing the data acquisition system from a standard multichannel analyzer (MCA) to a fast ADC card.

The worked out technique with a reversed problem (reduction of MSEs and optimization of survival probability for SSEs) may be also applied to BEGe detectors used in the searches for rare nuclear processes at low energies, like neu-

trinoless double beta decay. Experiments like GERDA [1,2] or LEGEND [20] use (will use) a large numbers of such detectors. A clear advantage is the stability (clear separation of MSEs and SSEs – see Fig. 6) and the automation of the method, which is important in the case of large number of detectors. Simultaneously, it provides similar, or better results compared to the A/E method.

References

1. K.-H. Ackermann et al., The Gerda experiment for the search of $0\nu\beta\beta$ decay in ^{76}Ge. Eur. Phys. J. C **73**(3), 2330 (2013)
2. M. Agostini et al., Background free search for neutrinoless double-β decay with Gerda. Nature **544**, 47–52 (2017)
3. D. Budjáš et al., in *Pulse Shape Analysis with a Broad-Energy Germanium Detector for the Gerda Experiment*. Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE, New York (2008)
4. D. Budjáš et al., in *Background Suppression Using Pulse Shape Analysis with a BEGe Detector for Neutrinoless Double Beta Decay Search with* GERDA, ed. by M.L. Marshak. AIP Conference Proceedings, vol. 1182, no. 1. AIP, San Diego (2009)
5. D. Budjáš 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. J. Instrum. **4**(10), P10007 (2009)
6. M. Agostini 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. J. Instrum. **6**(03), P03005 (2011)
7. M. Agostini et al., Characterization of a broad energy germanium detector and application to neutrinoless double beta decay search in ^{76}Ge. J. Instrum. **6**(04), P04005 (2011)
8. R.G. de Orduna et al., Pulse shape analysis for background reduction in BEGe detectors. EUR 24521EN, European Union (2010)
9. A. Hoecker et al., TMVA—toolkit for multivariate data analysis. PoS ACAT **040**, 92–98 (2007). arXiv:physics/0703039
10. R. Brun, R. Fons, ROOT—an object oriented data analysis framework. Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. **389**(1), 81–86 (1997)
11. E. Andreotti et al., in *Status of Underground Radioactivity Measurements in HADES*. 3rd International Conference on Current Problems in Nuclear Physics and Atomic Energy. Publishing Department of KINR, June 2010, Kiev, Ukraine, pp. 601–605
12. G. Gilmore, *Practical Gamma-Ray Spectroscopy* (Wiley, New York, 2011)
13. A. Georgiev, G. Werner, Digital pulse processing in high resolution, high throughput, gamma-ray spectroscopy. IEEE Trans. Nucl. Sci. **40**(4), 770–779 (1993)

14. M.B. Christopher, *Pattern Recognition and Machine Learning* (Springer, New York, 2006) **(Print)**
15. A.N. Gorban, D.C. Wunsch, in *The General Approximation Theorem*. The 1998 IEEE International Joint Conference on Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence, vol. 2. IEEE (1998)
16. B. Bruyneel, P. Reiter, G. Pascovici, Characterization of large volume HPGe detectors. Part II: Experimental results. Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. **569**(3), 774–789 (2006)
17. M. Hult et al., Underground gamma-ray spectrometry. Acta Chim. Slov. **53**(1), 1 (2006)
18. E. Fairstein, S. Wagner, IEEE standard test procedures for germanium gamma-ray detectors. IEEE Std 325-1996 (1996)
19. Y.-H. Mi et al., Compton suppression in BEGe detectors by digital pulse shape analysis. Appl. Radiat. Isot. **121**, 96–100 (2017)
20. N. Abgrall et al., The large enriched germanium experiment for neutrinoless double beta decay (LEGEND) (2017). arXiv:1709.01980v1