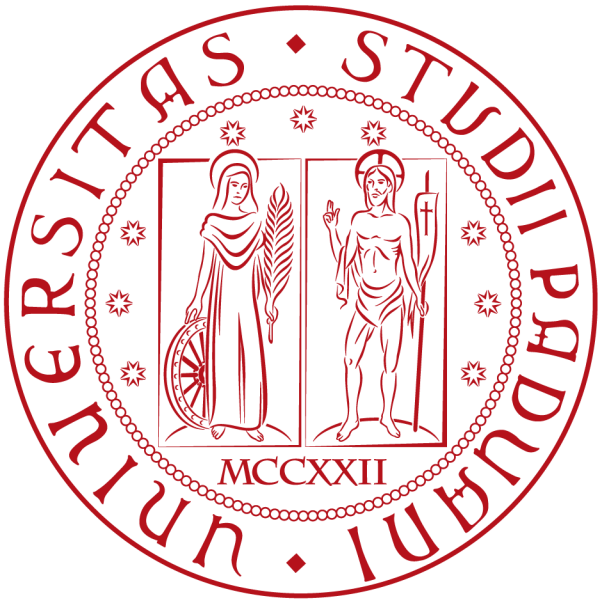


Università degli Studi di Padova

DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"
Corso di Laurea in Fisica

Tesi di laurea

Study and suppression of superficial events in a Germanium detector with artificial neural networks

Candidato:
Michelangelo Pari
Matricola 1027475

Relatore:
Dott. Alberto Garfagnini

Correlatore:
Dott. Eduardo Medinaceli

Anno Accademico 2014-2015

Contents

List of Figures

List of Tables

Chapter 1

Introduction

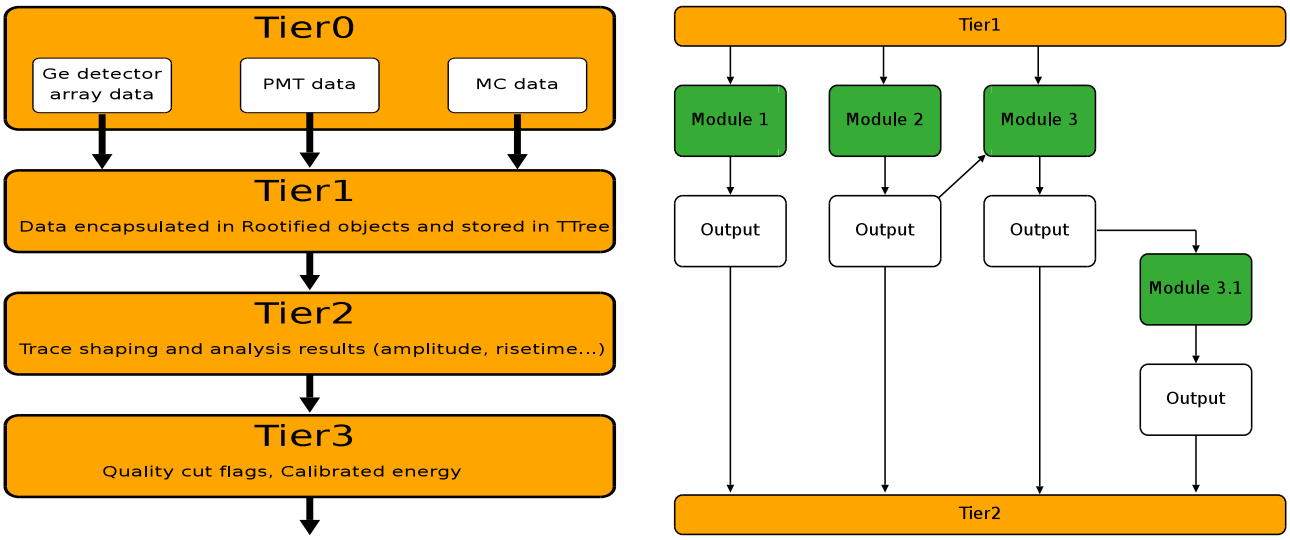
This work has the aim to improve previuos results on background discrimination via pulse shape analysis for the GERDA experiment, focusing on superficial events. Till now pulse shape discrimination was carried out in order to achieve good results in recognition of possible $0\nu\beta\beta$ candidates from γ background candidate events. Study of this type was made through the exploitation of artificial neural networks (ANN), in such a way that, after being trained, networks were able to discriminate signal from background and discard the latter.
In this case superficial events will be considered as signal in order to study their basic characteristics and being able to identify them in the best way possible. The ultimate goal is to create a sample as pure as possible of superficial events: this could be exploited to recognize and discard such events as background events in a GERDA data spectrum.
Before starting this work some introductions are needed.

1.1 The GERDA experiment

The GERmanium Detector Array (GERDA) experiment, at the Laboratori Nazionali del Gran Sasso (LNGS), is a search for the neutrinoless double beta deacay ($0\nu\beta\beta$) of ^{76}Ge. ^{76}Ge is one of double beta $2\nu\beta\beta$ isotopes emitters and so, outgoing from the Standard Model, a possible $0\nu\beta\beta$ emitter [1]. This is very attractive from an experimental point of view, beacause high-purity Germanium detectors (HPGe) can be produced from material enriched in ^{76}Ge. Therefore, in the GERDA experiment, HPGe detectors enriched in ^{76}Ge with an isotopic fraction ups to 86%, act as both sources and detectors. Naked high-purity Germanium detectors (enriched to 86% in ^{76}Ge) are mounted in low-mass copper supports and immersed in a 64 m^3 cryostat filled with liquid argon (LAr). The LAr serves as the cooling medium and shield against external backgrounds. The shielding is complemented by a 590 m^3 steel water tank, to detect muons coming from the outside.
The experiment consists of two phases. GERDA Phase I ended with the establishment of the most stringent $0\nu\beta\beta$ half-life limit for ^{76}Ge ([2]). Past claim for a $0\nu\beta\beta$ signal in ^{76}Ge (Heidelberg-Moscow experiment, [3]) has been strongly disapproved, calling for a further exploration of degenerate Majorana neutrino mass scale. This will be pursued by GERDA phase II aiming for a sensitivity increased by a factor of about 10. In GERDA phase II will be deployed about 25 new custom-made enriched Broad Energy Germanium (BEGe) detectors. These will be added to the previous eight semicoaxial HPGe detectors taken (and reprocessed) from Heidelberg-Moscow (HDM) and IGEX [4] experiments and to five Broad Energy Germanium detectors.

1.2 The GERDA off-line data processing

The charge signals of the HPGe detectors deployed in GERDA are composed by pulses with a rising part of about 1 μs followed by an exponential decay tail ($\simeq$ 100 μs) given by the front-end charge-sensitive preamplifier. The rising part (leading edge) is induced by the drift of electron-hole pairs - generated by γ-ray or charge particle interactions - toward the detector electrodes and it contains information concerning the event topology ([5]). The analog signals are digitized at 100 MHz sampling frequency by 14-bit flash-ADCs (FADC) equipped with integrated anti-aliasing bandwidth filters. For each event, the FADC computes in run-time and writes to disk two traces. The first trace is sampled at 100 MHz and it is 4 μs long (high-frequency-short trace, HFS trace). The second trace is sampled at 25 MHz and it is 160 μs long (low-frequency-long



(a) *Multi-tiers structure of GERDA data.* (b) *Generation of Tier2 file from Tier1 one.*

Figure 1.1: Conceptual schemes of GERDA data storage and manipulation.

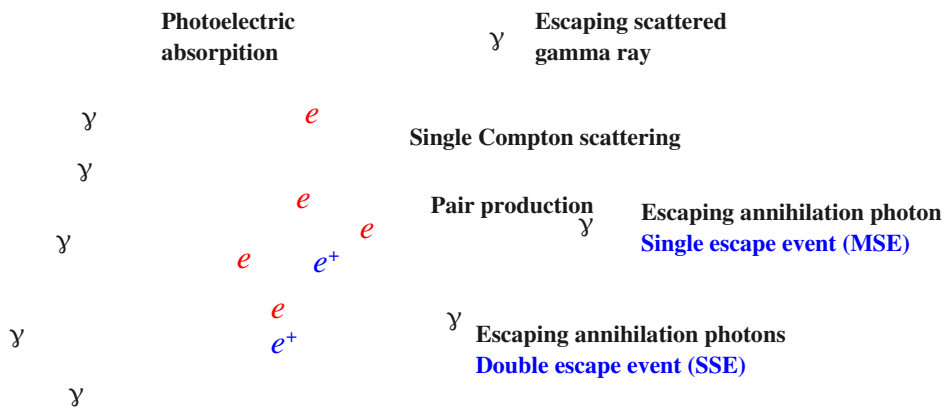
trace, LFL trace). The offline analysis of the digitized charge pulses is performed with a software framework for advanced data analysis and digital signal processing developed for the GERDA experiment: GELATIO (GErda LAyouT for Input/Output). Such a framework, implemented in `C++`, organizes the data into a multi-tier structure, from the raw data up to the condensed analysis parameters, and includes tools and utilities to handle the data stream between the different levels ([6]). The raw data provided by FADC are stored in the lowest level (Tier0). Data are then converted into a new encoding and stored as Tier1. Tier1 file contains exactly the same amount of information of raw data; the difference is that Tier0 is the native data acquisition system (DAQ) format, while the Tier1 is encoded with a different format based on ROOT [7] and MGDO [8]. The latter is a `C++` software library which contains a set of data objects and interfaces to encapsulate, store and manage all the physical quantities of interest. For example, a Tier1 file contains a `TTree` of MGDO object `MGTEvents`, in which are stored all the information acquired by the DAQ. Tier1 files are the starting point for the analysis; Tier2 files, obtained by Tier1 processing, are the first "real data" files: they contain for example all the risetimes of charge pulses, baselines, energies, max amplitude of current pulses etc. Similarly, the Tier3 files store the information extracted from the Tier2 and so on. The core of GELATIO is the digital signal processing which creates the Tier2 files starting from the detector signals stored in the Tier1. To support a highly customizable analysis, the design of GELATIO is based on a modular approach. The analysis is divided into modules, each of them handling a unique and consistent task of the digital data processing. The modules utilized for this work are listed below.

- `GEMDEnergyGauss`. It reconstructs the event energy using an approximate Gaussian filter (uses LFL trace).
- `GEMDRiseTime`. It computes the rise time between 10% and 90% (by default) of the maximum amplitude of the pulse (uses LFL trace).
- `GEMDCurrentPSA`. It computes the current signal as the derivative of the charge signal and then extracts the basic features of the current pulse, like rise time, width and area (uses HFS trace).

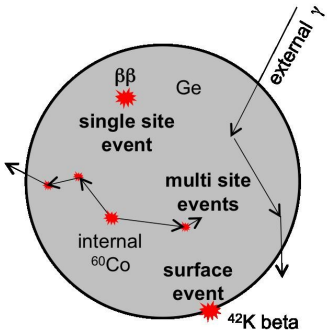
A graphic example of the multi-tier structure of data can be seen in figure 1.1a, while figure 1.1b shows a schematic diagram of a GELATIO run.

1.3 Pulse shape discrimination in Germanium detectors

Pulse shape discrimination analysis is of primary importance in the GERDA experiment. The signature of a $0\nu\beta\beta$ decay of ^{76}Ge is an energy deposition of the two beta particles with an energy sum equal to $Q_{\beta\beta} = 2039\ keV$. Given the energy of the beta particles and the density of germanium crystals, the path



(a) *Different type of events in a detector.*



(b) *Charateristics events in the GERDA experiment.*

Figure 1.2: Type of events used for pulse shape discrimination tecnique.

length is of the order of one millimetre ([9]). On the contrary, in background events near $Q_{\beta\beta}$ (high-energy γ rays), photons interact via multiple Compton scattering and energy is often deposited at several locations well separated by few cm in the detector. Discriminating multiple-site events (MSE) from single-site events (SSE) can be therefore a powerful tool to accept $0\nu\beta\beta$ events and reject backgrounds. The pulse shapes will be, in general, different for the two event classes and this can be used to improve the sensitivity of the experiment, by discriminating between them. Energy depositions from α or β decays near or at the detector surface lead to peculiar pulse shapes as well that allow their identification ([10]).
Figures 1.2a and 1.2b show the graphic representation of the various events types discussed before.

To understand proceedings of this work, it's useful to spend some words about events typology in a energy spectrum acquired with an HPGe detector. First of all, compared with NaI scintillators, the lower atomic number of Germanium leads to a photoelectric cross section smaller by a factor of $10-20$ (and even greater in comparison with higher Z scintillators). The events contributing to the full-energy peak are now much more likely to consist of multiple interactions, such as Compton scattering followed by photoelectric absorption of the scattered photon. Absorption of the full photon energy in a single photoelectric event is relatively rare, as can be seen in figure 1.3 ([11]).

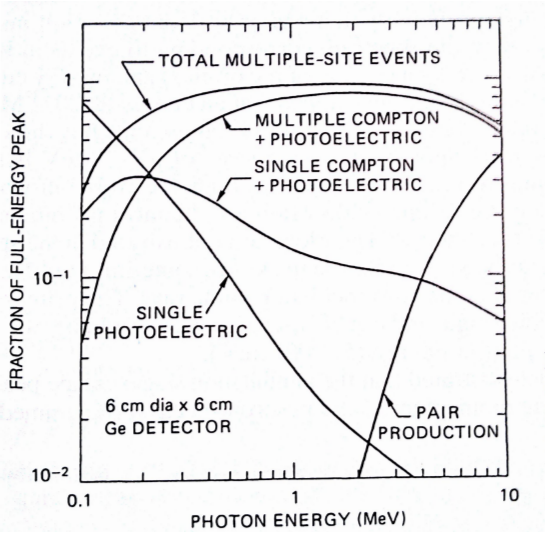


Figure 1.3: Fractions of the FEP contributed by different energy loss mechanisms in a 6 cm × 6 cm coaxial HPGe detector, as predicted by Monte Carlo simulation. Absorption of the gamma-ray photon in a single photoelectric interaction predominates only for energies below about 140 keV (taken from [11]).

At high gamma-ray energies pair production becomes predominant. Electron-positron pair is created at the site of the original gamma-ray interaction. Electron's energy is rapidly absorbed (SSE), while positron annihilate creating two 511 keV gamma-rays. So, all the events belonging to the double escape peak (DEP)

will be SSE type, while all those belonging to the single escape peak (SEP) will be MSE type. Regarding superficial events, they will be discussed in the following section, as they depend on the particular employed detector.

1.4 Broad Energy Germanium detectors

Recent developments in pulse shape analysis show that detectors with a small read-out electrode exhibit pulse shape discrimination performance superior to coaxial HPGe detectors ([12]). A commercial detector with such characteristics is the p-type HPGe broad energy Germanium (BEGe) produced by Canberra [13]. A modified model of this detector was successfully used in the frame of the GERDA experiment (see [9]). Detectors of this type are used in this field of research for the first time, and a massive exploitation of them is provided in GERDA phase II, as said before. Detector features are briefly described in the following.

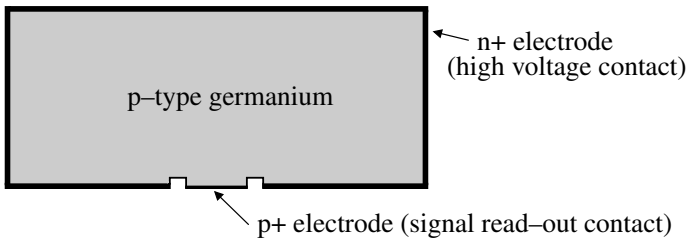


Figure 1.4: Sketch of a BEGe detector. The signal read-out electrode and the groove are not to scale. The plot shows a vertical section of the detector passing through the symmetry axis (taken from [12]).

The crystal is made of p-type HPGe with Li-drifted n^+ contact covering the whole outer surface including most of the bottom part. The small p^+ contact is located in the middle of the bottom side (see figure 1.4). By Shockley-Ramo theorem, the istantaneous current induced on a given electrode (by each single charge carrier) is equal to:

$$i = q\vec{v} \cdot \vec{E}_w \tag{1.1}$$

Where q is the charge of the carrier, $\vec{v}$ is its istantaneous velocity and $\vec{E}_w$ the weighting field in its actual position. The induced charge on the electrode is given by the product of the charge on the carrier multiplied by the difference in the weighting potential ϕ_w from the beginning to the end of the carrier path:

$$Q = q\Delta\phi_w \tag{1.2}$$

In this special case of BEGe detectors we are in presence of the so called "small pixel effect" [11], where weghting field and weighting potential have the shape reported in figure 1.5.

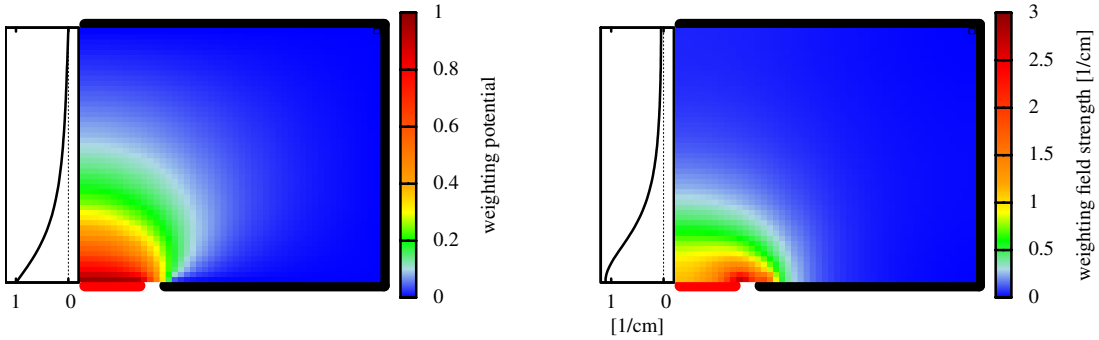


Figure 1.5: Weighting potential (ϕ_w) and weighting field strength ($\|\vec{E}_w\|$) of the small read-out electrode computed for half of a vertical section of the BEGe detector (taken from [12]).

In such a situation, looking at equation 1.2, one can see that over most of the detector volume only the holes will contribute to the induced charge signal on the electrode. Again, looking at equation 1.1, it can be seen that over most of the detector volume only holes are responsible for the current peak and the contribution of electrons is small ([14]). It can be noticed that for all these events, the last part of the hole collection happens along a common path which is independent of the starting position. Accordingly, the last part of

the induced signal is identical for the different events. Since the first part of the signal, induced by the holes and electrons in weak $\|\vec{E}_w\|$ (and ϕ_w) regions, is comparatively small, the signal shapes are essentially independent of the interaction position ([12]). Consequently, for a localized deposition, the maximum of the current pulse is nearly always directly proportional to the total charge q of the moving charge cluster (seen as a single charge carrier) and so directly proportional to the energy of the event. This translates in a powerful A/E discrimination, discussed below.

- For SSE, as just said, A/E is independent of the energy and the interaction location inside mostly detector active volume. So, for SSE in the bulk detector volume one expects a nearly Gaussian distribution of A/E with a width dominated by the noise in the readout electronics ([10]).
- For MSE the current pulses of the charge clusters from the different locations will have different drift times and hence two or more time-separated current pulses are visible. For the same total energy E, the maximum current amplitude A will be smaller in this case.
- For surface events near the p^+ electrode, the electrons contribution to the current pulse will be no more negligible, but comparable with that of holes. So, for the same energy E, the maximum current amplitude will be quite greater in this case, in such a way to be able to recognize efficiently these signals.

Before describing n^+ events characteristics, n^+ region of detector has to be described. The n^+ electrode is formed by infusion of lithium, which diffuses inwards resulting in a fast falling concentration profile starting from saturation at the surface. The $p-n$ juction is below the n^+ contact surface, and, from the junction, electric field decrease till reaching 0 at the edge of n^+ electrode dead layer (DL) (0.8-1 mm thick).

- For an event occurred in the dead layer, holes (electrons do not contribute at all) can be transferred from the dead layer into the active volume via diffusion. This is true only for the DL region in wich Li concentration is not so high to results in a significant recombination probability. The slow nature of diffusion compared to the charge carrier drift in the active volume results in a increased rise time of signals for these events. Charge recombination instead causes a ballistic deficit loss in the energy reconstruction. The final result is a reduced A/E ratio ([10]).

It's important to remember that, unlike the current amplitude, the amplitude of the induced charge (and so the energy) is independent of the event type and of the region where charges are formed, provided all carriers are fully collected ([11]).

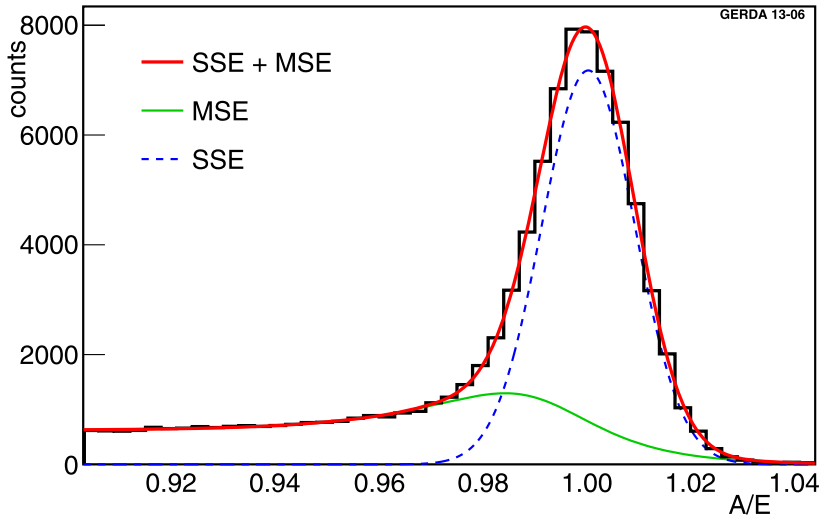


Figure 1.6: A/E distribution for Compton continuum data. The dashed blue curve is the Gaussian component and the green curve is the component approximating the MSE contribution (taken from [10]).

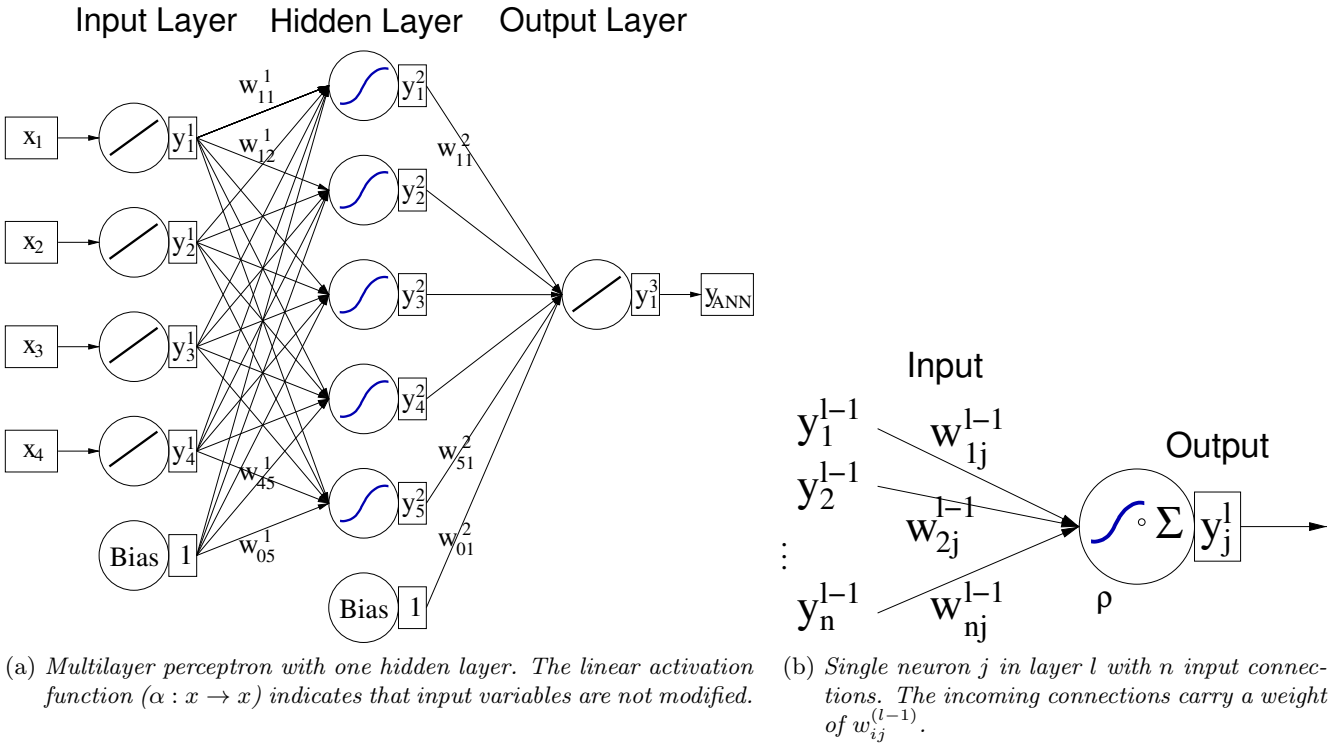
In figure 1.6 can be seen what said before about A/E distribution of SSE and MSE. In particular MSE term is approximated by the function:

$$f(x = A/E) = A_0 \cdot \frac{e^{A_1(x-A_2)} + A_3}{e^{(x-A_2)/A_4} + A_2} \tag{1.3}$$

MSE term is parameterized empirically by parameters A_0, A_1, A_2, A_3 and A_4.

1.5 TMVA and Artificial Neural Networks

The Toolkit for Multivariate Data Analysis (TMVA [15]) provides a ROOT-integrated machine learning environment for the processing and parallel evaluation of multivariate classification (and regression) techniques. It contains the object oriented implementation in C++/ROOT of many multivariate analysis methods, including artificial neural networks (ANN).
A neural network is in general a computational model, inspired by the architecture and operation principles of central nervous systems, capable of machine learning and pattern recognition.



(a) *Multilayer perceptron with one hidden layer. The linear activation function ($\alpha : x \to x$) indicates that input variables are not modified.*

(b) *Single neuron j in layer l with n input connections. The incoming connections carry a weight of $w_{ij}^{(l-1)}$.*

Figure 1.7: Schematic view of a one-hidden layer MLP and a generic neuron (taken from [15]).

Figure 1.7b shows the general structure of a neuron. Each neuron receives input information from the others through inter-neuron connections, that are weighted with an appropriate connection weights ($w_{ij}^{(l-1)}$) based on the connection's importance. Each neuron is also associated with a bias value, that in case of figure 1.7b we should call $w_{0j}^{(l-1)}$. The neuron processes the input information (and the bias) through the neuron responce function (ρ) returning an output value. The neuron responce function can be seen as a function $\rho : \mathbb{R}^n \to \mathbb{R}$ that maps the neuron input $(y_1^{(l-1)}, \ldots, y_n^{(l-1)})$ onto the neuron output. This responce function can be separated into two functions: the synaps function

$$k : (y_1^{(l)}, \ldots, y_n^{(l)} | w_{0j}^{(l)}, \ldots, w_{nj}^{(l)}) \to w_{0j}^{(l)} + \sum_{i=1}^{n} y_i^{(l)} w_{ij}^{(l)} \tag{1.4}$$

and the activation function

$$\alpha : x \to \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{1.5}$$

so that $\rho = \alpha \circ k$. The most simple activation function is the linear function $\alpha : x \to x$, that doesn't modify input variables (all information are always transmitted). Functions 1.4 and 1.5 are those we are going to use in the analysis.
Figure 1.7a shows an example of a multi layer perceptron (MLP) with one hidden layer. Multi layer perceptron is a particular kind of neural network where neurons are organized in layers and connections are allowed only between neurons belonging to different layers. The input layer consists of the chosen variables for the

classification, while output layer of one single neuron that holds the output variable: the neural network estimator y_{ANN}.

Artificial neural networks have to be trained for every specific problem before the usage. The training phase is aimed to properly set the weights of each connection according (in our case) to the "backpropagation algorithm" (BP). The weights are chosen in order to minimize the error function E, defined as:

$$E(\vec{x}_1, \ldots, \vec{x}_N|\vec{w}) = \sum_{a=1}^{N} \frac{1}{2}(y_{ANN,a} - \hat{y}_a)^2 \tag{1.6}$$

Where a represents the single processed event from the training sample, N is the number of events in the training sample and $\hat{y}_a$ is the known expected result for that particular event: 0 for background and 1 for signal. As said before $y_{ANN,a}$ is the output of the output layer for the event a, that is, in case of figure 1.7a (and all biases set to 0):

$$y_{ANN} = \sum_{j=1}^{n_{hl}} y_j^{(2)} w_{j1}^{(2)} = \sum_{j=1}^{n_{hl}} \tanh\left(\sum_{i=1}^{n_{var}} x_i w_{ij}^{(1)}\right) w_{j1}^{(2)} \tag{1.7}$$

Where n_{hl} is the number of neurons in the hidden layer, n_{var} is the number of discriminating variables (number of neurons in the input layer), $w_{ij}^{(k)}$ are the connections weights and x_i are the single variables. In equation 1.6 each vector $\vec{x}_i$ represents all discriminating variable for the event i and $\vec{w}$ is the vector of all network weights. Starting from a random set of weights $\vec{w}^{\,(i)}$, according to the BP algorithm, the new vector of weights will be:

$$\vec{w}^{\,(i+1)} = \vec{w}^{\,(i)} - \eta \nabla_w E\big|_{\vec{w}^{(i)}} \tag{1.8}$$

η is a positive number and it's called "learning rate". Such an algorithm will converge to a stationary point of the error function, with a speed and precision stated by the numerical value of η. In this particular case in which η is positive, the algorithm will converge to a minimum: E is never negative, and so it will have a global minimum value at or near 0 (0 only if it's possible for the neuron to represent the target function perfectly).

In order to allow consistency checks, the training algorithm described above is performed using only a subsample of the whole training data set. The other subsample becomes the test sample, to which is applied the just trained network. This guarantees a statistically independent evaluation of the MVA algorithms based on the test sample ([15]). Therefore, the whole TMVA training phase consists of a proper training and a testing phase. Just after these two steps it is always possible to access and check the results regarding the reliability of trained network, computed by the application of such a network on the test subsample.

Chapter 2

Phase 1

For the following analyis it will be used ANN configuration parameters and discrimination variables which led to the best results in [16]. The first phase of this work consists therefore in training ANN to discriminate SSEs from MSEs in the active volume of detector, making sure that the chosen parameters of the networks are able to return results in agreement with previuos ones.

2.1 Settings and data samples

Discriminating variables are:

1. Rise time of the charge pulse, from 10% to 90% of the maximum height.

2. A/E ratio.

3. Parameter characterizing the asymmetry of the current pulse.

The latter parameter is chosen as:

$$Asym = (A/E) \cdot \frac{L - R}{L + R} \tag{2.1}$$

Where L and R are the left and right current pulse integrals computed from the maximum amplitude to the 3% of the maximum amplitude.
The chosen neural network is the MLP with a single hidden layer containing 18 neurons. The data for the training phase is taken from the "*N28_ Th228_ calibration*" measurement of Thorium spectrum with the $GD91C$ detector (see section 3.1 for more information). The DEP of ^{228}Th at 1592 keV has been chosen as SSE prototype, while ^{214}Bi FEP at 1620 keV as MSE prototype. Taking into account the shape of spectral background [17], these peaks has been fitted with:

$$f(x) = \frac{A_0}{A_1\sqrt{2\pi}} e^{-\frac{(x-A_2)^2}{2A_1^2}} + A_3 + \frac{A_4}{e^{\frac{x-A_2}{2A_1}} + 1} \tag{2.2}$$

Where A_i are fit parameters and represent respectively: gaussian area, sigma and mean value, background right, background left - background right.
Selected events are all those lying into 2σ of the gaussian part of fitting function 2.2.
From the DEP fit it's also possible estimating the background percentage in the SSE sample, that was computed taking mean value of the fitted function integrals from 6σ to 8σ at the left and right of the centroid position. Background percentage in the DEP sample results $(15.5 \pm 0.3)\%$. Figure 2.1 shows the two fitted peaks.

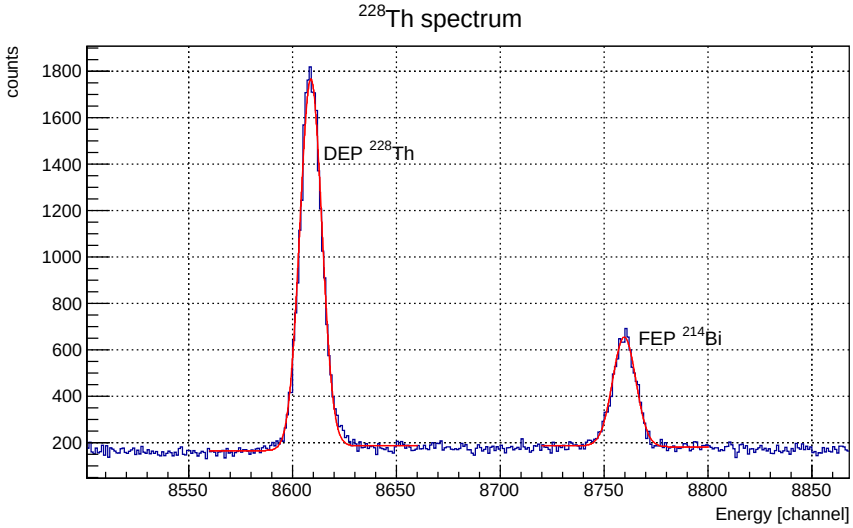
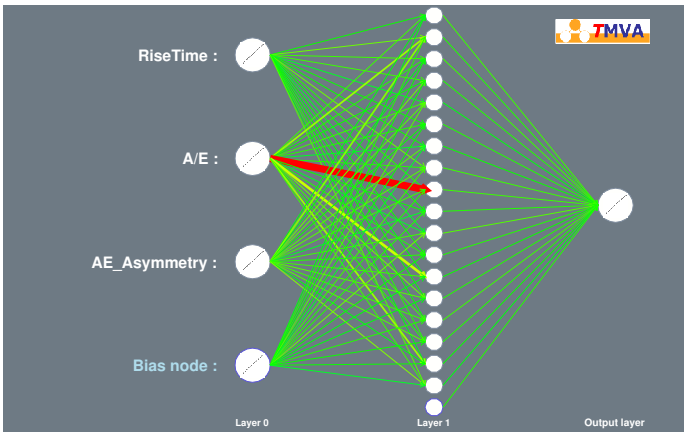


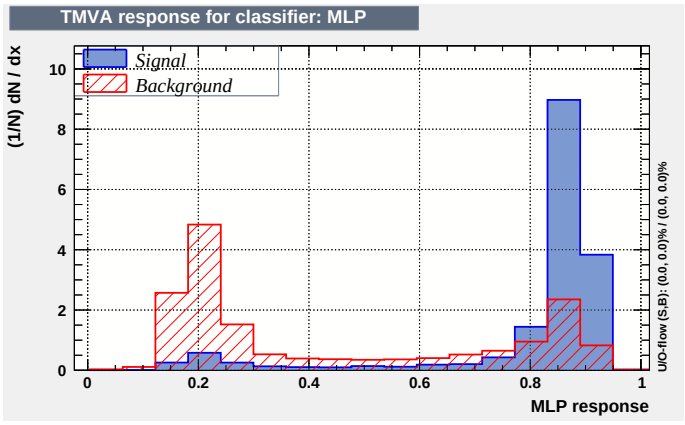
Figure 2.1: DEP of ^{228}Th and FEP of ^{214}Bi fitted with function 2.2.

2.2 Networks training results and application

All relevant information are enclosed in the classifier output distribution, that can be seen in figure 2.2b. From here the optimal cut value can be chosen according to signal efficiency and background rejection. Signal efficiency for a given cut value is computed calculating signal fraction above that value. Background rejection computing background fraction before cut value.



(a) *Network architecture. The colour and width of the arrows represent the value of the weight of every connection.*



(b) *Output of classifier.*

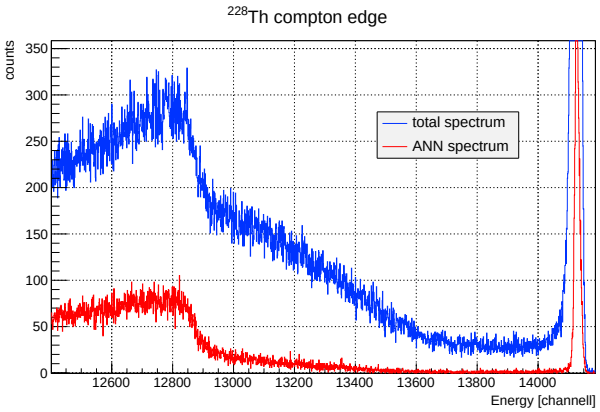
Figure 2.2: Network's responses for training phase

Optimal cut value has been chosen as that output classifier value which maximizes signal efficiency and background rejection. Values are in table 2.1.

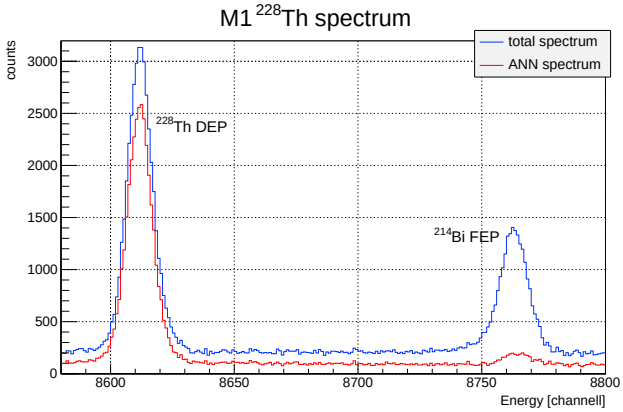
Cut Value	Signal Efficiency	Background Rejection
0.82	0.78	0.78

Table 2.1: Training phase evaluation values that will be utilized for network application phase.

The first application phase of this network is done on the Compton edge of ^{228}Th FEP. In this region the majority of events is multiple site and I'm expecting a susbstantial reduction. Another test to be applied to just trained network consists of processing the spectral region of ^{228}Th DEP and ^{214}Bi FEP taken from another measurement of Thorium spectrum (result of application on training measurement is already known). It has been chosen measurement "*M1_Th228_top_v2*", taken with same detector of training sample.



(a) *Application on N28 measurement: Compton edge.*



(b) *Application on M1 measurement: ^{228}Th DEP and ^{214}Bi FEP region.*

Figure 2.3: Application results of trained network: energy spectra.

	Compton Edge	^{228}Th DEP	^{214}Bi FEP
Discarded events (%)	86.8	19.8	83.3
Survived events (%)	13.2	80.2	16.7

Table 2.2: Ratio of discarded and survived events in application phase.

Results of these applications can be seen in figure 2.3a for application on Compton edge of training file, and in figure 2.3b for application on Thorium DEP and Bismuth FEP of measurement "*M1*". The first figure shows that network successfully brings to zero the "pure" multiple Compton scattering region; the latter that network is clearly able to discriminate between multiple site and single site. Numeric values indexes of the application goodness are displayed in table 2.2.
Values obtained in this first phase are in a quite good agreement with previuos values of [16]. Neural network set parameters and discriminating variables have returned the positive expected results, therefore these settings will be maintained for the newest part of this analysis.

Chapter 3

Phase 2

Here starts the real part of this work. ANN with the same configuration and discriminating variables chosen in the previous chapter will be employed trying to discriminate superficial events from all other types. In such a way, if discriminating parameters are satisfactory and classification is feasible, it will be possible to isolate superficial events from the others and to create a "pure" sample of superficial events. If this will be successful, such a method could be adapted to the GERDA experiment for the removal of superficial events from GERDA spectra (using the final obtained pure sample as background for ANN).

3.1 The data: Upside-Down BEGe GD91C

The whole data set exploited for this part of analysis comes from Bjoern Lehnert (et al.) experiment on the BEGe detector GD91C at LNGS, first presented at the *GERDA Collaboration Meeting* on 4th November 2013 at LNGS ([18], [19]). Experimental setup and objectives of this experiment are listed below.

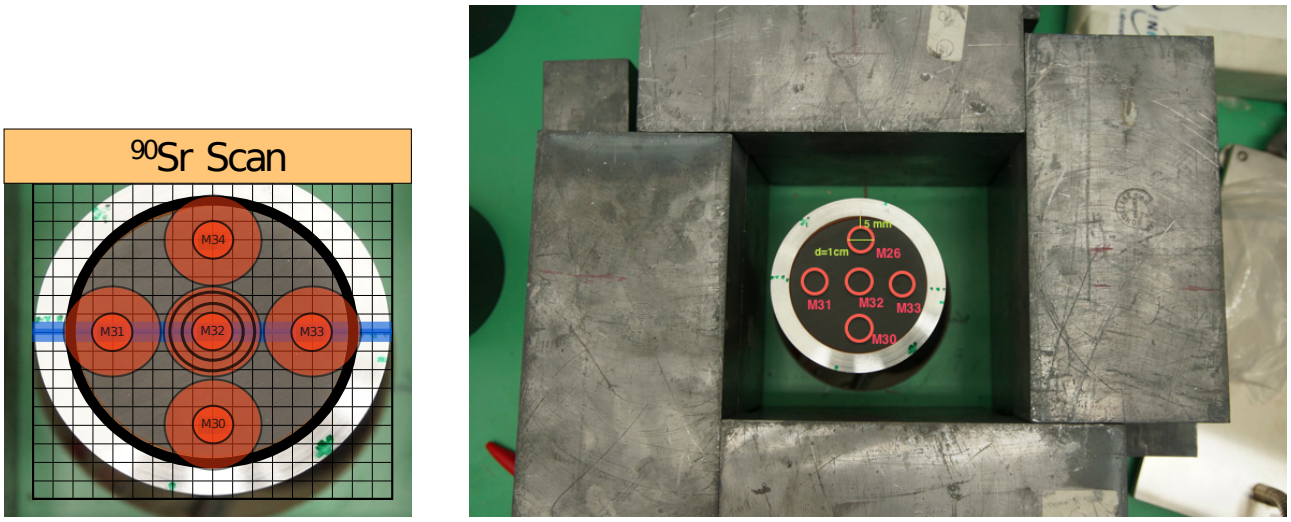
Experimental setup

- BEGe detector GD91C was mounted upside-down, with the p^+ contact side upwards.
- 0.6 mm carbon-epoxy window.
- No passivation layer.
- Reverse bias voltage 3800 V (3600 V only if specified).

Objectives

- Transition layer investigation.
- Pulse shapes at p^+ contact.
- DEP pulse shapes without passivation.
- Full Charge Collection Depth on p^+ contact side, Active Volume without passivation.

Measurements with several sources have been acquired, each of them for a different purpose. A detailed description of Bjoern's experiment is not our purpose, so here it ends.
What is of interest to this work is the ^{90}Sr set of measurements. $^{90}_{38}$Sr β^- decays in ground state (g.s.) of ^{90}Y with a $Q_\beta = 545.9\ keV$ and $T_{1/2} = 28.80\ y$. In his turn $^{90}_{39}$Y β^- decays in g.s. of the stable nucleus ^{90}Zr, with a $Q_\beta = 2279.8\ keV$ and $T_{1/2} = 2.6684\ d$ [20].
All the measurements have been taken on the upside-down detector's top surface. Experimental apparatus can be seen in figure 3.1a and 3.1b. In figure 3.1a are shown the ^{90}Sr measurement sites: five measurements have been taken, named from "M30" to "M34". The big circular red areas represent the beam size on detector's surface and the small inner one the collimator size. Diameter of these two regions are respectively 24 mm and 10 mm. Central region of detector corresponds to the p^+ contact, where the space enclosed by two black circumferences just out of M32 collimator is the groove between p^+ contact and n^+ one. In figure 3.1b can be seen the so called "lead castle" (LC), used as shielding from external background. During measurements top of lead castle was covered by the ^{90}Sr collimator, an Al plate and a PE plate. Bremsstrahlung (BS) and background measurements for ^{90}Sr were acquired too (with closed collimator).
Another measurement that will be taken into account will be that of a ^{60}Co source, acquired in similar conditions to ^{90}Sr one. $^{60}_{27}$Co β^- decays into ^{60}Ni excited states ($Q_\beta = 2823.07\ keV$) then γ, giving two spectral lines at 1173.228 and 1332.492 keV (most probable ones) [20].



(a) *Detailed view of GD91C upside-down data taking.*

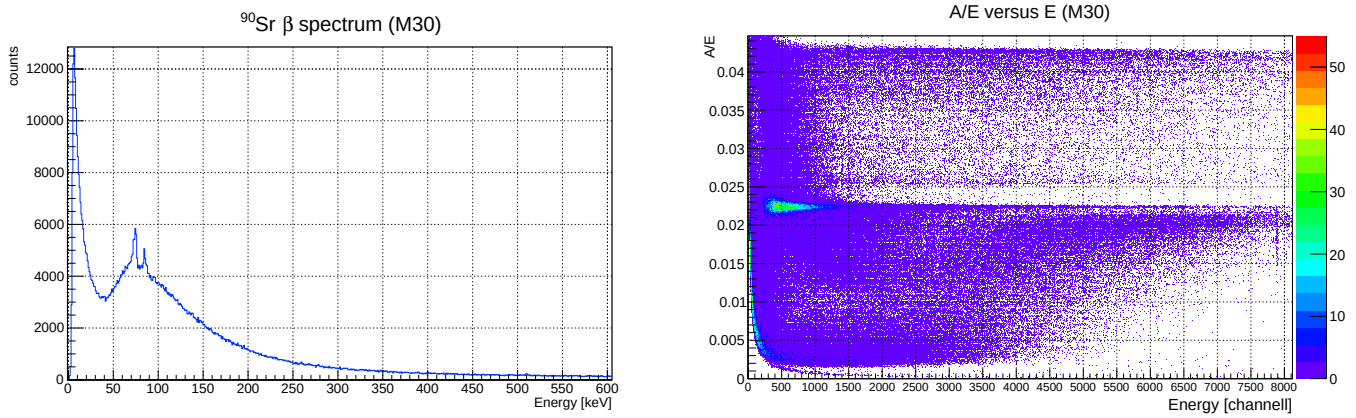
(b) *View of GD91C inside the Lead Castle (LC). M26 acquired in the same M34 site.*

Figure 3.1: BEGe detector GD91C in the ^{90}Sr measurements. Taken from [21] and [19]

.

3.2 Analysis

Beta surface events of ^{90}Sr are those superficial events we intend to study, while ^{60}Co can be exploited as source of internal events (high energy γ rays interacting in the detector's active volume).
First of all, we can observe general features of a raw ^{90}Sr spectrum.

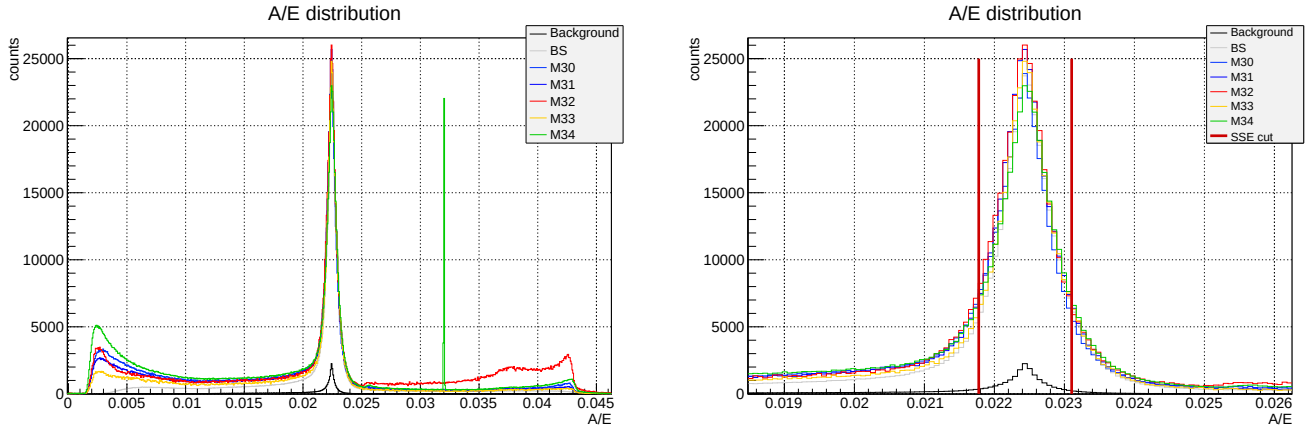


(a) *Calibrated energy spectrum of ^{90}Sr*

(b) *A/E as function of energy (in channel)*

Figure 3.2: Graphics for ^{90}Sr M30 measurement.

Figure 3.2a shows the calibrated energy spectrum of ^{90}Sr M30 measurement. Calibration was carried out through the Thorium measurement "M1_Th228_top_v2". Calibration clearly allows us to identify Kα1 and Kβ1 Pb fluorescence peaks, respectively at energies of 74.97 and 84.94 keV [22]. In figure 3.2a, before the proper ^{90}Sr beta spectrum, near zero, we notice the presence of a very sharp and high peak.
Figure 3.2b is a bright example of the BEGe physics discussed in section 1.4: about between 0.025 and 0.02 of A/E axis lies the single site band, while the whole region above SSE band is entirely populated by p^+ events. Figure 3.2b clearly shows the ease of identification and removal of p^+ events from the others exploiting the powerful discriminating parameter A/E (see section 1.4). For such a reason, we'll never consider p^+ events in this study, cutting off them a priori. Thus, we have at first to identificate the SSE band: once obtained such an interval, all events above its upper limit can be discarded as p^+.



(a) ^{90}Sr *A/E distributions. The high outlayer line in M34 measurement is an* α *particle. Notice that M32 measurement (the red one) has a higher* p^+ *region, as expected.*

(b) *SSE band first cut at* $3RMS$ *of bakground distribution.*

Figure 3.3: A/E normalized distribution of all ^{90}Sr measurements.

As can be seen in figure 3.3a all A/E distributions have the same central peak, i.e. SSE peak. So an A/E cut is the most simple way to reject all SSE band and all p^+ events. Having a good statistic, cut is chosen by taking mean value and RMS of background A/E distribution, and cutting at ± 3RMS. SSE cutting interval results:

$$\text{SSE Range} = [0.0218, 0.0231] \tag{3.1}$$

This could appear quite small for non background distributions (see figure 3.3b), but the point is that for ^{90}Sr measurement a considerable part of (mainly BS) MSE events clusters on the left side of SSE peak with a shape expressed by equation 1.3 (see figure 1.6). In such a way cutting too much could lead to a considerable loss of MSE events and probably of some superficial events. Anyway, residual presence of SSEs is not a problem for us, because ANN could easily identify them, as seen in section 2.2.
Cutting in this way, however, could lead to problems due to the linear A/E systematic small energy dependence discussed in [10]. This small observed dependence of A/E from energy is thought to be an artefact of data acquisition and/or signal processing. Explicit correction was carried out in [9] exploiting the whole ^{228}Th calibration data to compute the sistematic (negative) slope of this linear dependence. In our case, deciding to cut off all the data exceeding the ^{90}Sr Q_β value, such a correction is negligible. So, evaluating calibration function at ^{90}Sr Q_β, energy cut value in channel turns out to be:

$$Q_\beta^{ch} = 2954 \text{ ADC channels} \tag{3.2}$$

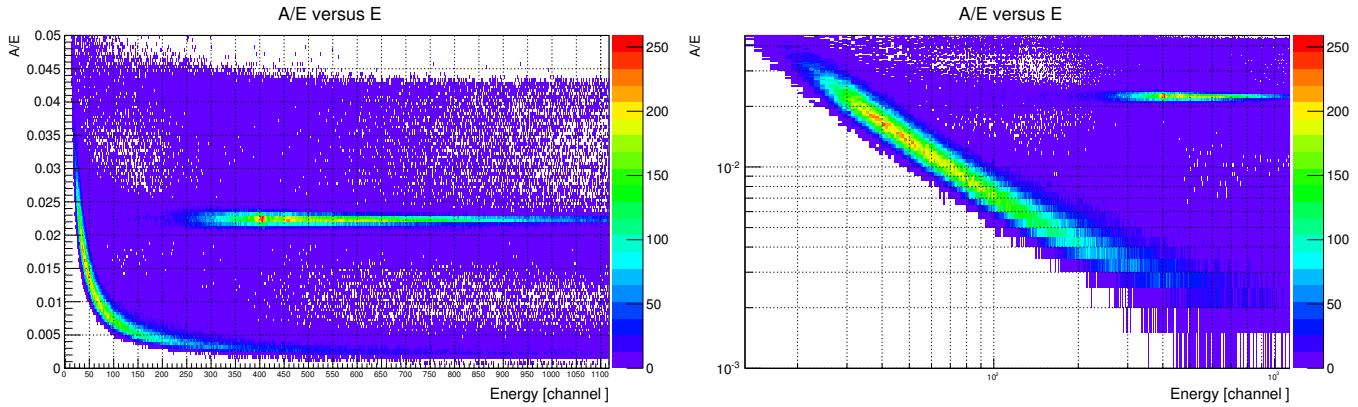
In such a way we are focusing on the whole ^{90}Sr low-energy beta.

We have to do some considerations about the events survived the cuts 3.1 and 3.2. Survived sample now contains mainly MSE and n^+ events. The latter are those events we want to focus on. It is not possible to identify n^+ events from MSE by simple A/E or energy cuts. According to the theory of BEGe detector discussed in section 1.4 MSE and n^+ events have quite comparable A/E ratio. Thus, in discrimination of these two types of events, parameter A/E, that has proved to be the most important discriminating variable between SSE and MSE (see figure 2.2a), should have to loose almost all its importance. Another discriminating variable should rise above the other ones: current signal risetime. Indeed, due to holes diffusion from dead layer, n^+ are the slowest events among all those who are considered. The purpose then, is to training an artificial neural network that will be able to distinguish between n^+ and MSEs. To do this we have to make a careful choice of background and signal training samples. As previously done in section 2.2 "pure" background and signal sample are needed for training phase. But if a pure background sample can be identified with whole events inside a gamma line, it is not possible to obtain a sample consisting only of n^+ events.
Thus, our purpose is to prepare the purest n^+ training signal sample as we can do. Till now selected cuts for signal sample are:

1. $A/E < 0.0218$, for SSE and p^+ events.

2. $E < 2954$, for events after the ^{90}Sr Q_β.

Looking at figure 3.2a and 3.2b one can notice that the events survived the latter cuts are not just n^+ and MSE, but also all those "not classified" events lying under the sharp peak near zero in figure 3.2a. In the same energy range of this peak, in figure 3.2b, can be seen an high-density region with a well definite shape. The situation is the same for all the other ^{90}Sr measurements.



(a) *No log_{10} scale.* (b) *XY log_{10} scale.*

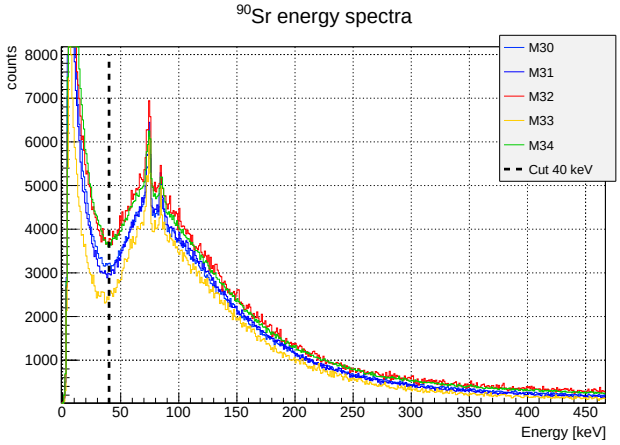
Figure 3.4: A/E as function of E for M30 measurement: view of low-energy region.

In figure 3.4a is shown a detailed view of low energy region of A/E versus E plot. It's clearly visible the horizontal SSE band, and the other "unknown" high density region, that seems to have a shape $\propto 1/E$. This first impression is confirmed by looking at figure 3.4b, that is the same previous plot with bilogarithmic scale set on. Now the inspected events distribution takes the form of a straight line, that is the shape that assumes a power law in bilogarithmic scale. In particular, if the shape is in a good approximation $\propto 1/E$, the slope of that linear region should be about -1.
The most important question now is what types of events are hidden in such a region. At first sight, looking at the energy spectrum, they seem to be some type of pathological events, due to, for example, electronics. Therefore the idea would be to cut off the "pathological" peak from the energy spectrum, with all the events enclosed by this. But looking at figure 3.4a one can think that "pathological" events could be enclosed only in the "hyperbolic" high-density region, so cutting on energy would lead to a loss of all possible physical events between "hyperbolic" region and SSE band. More information are needed.

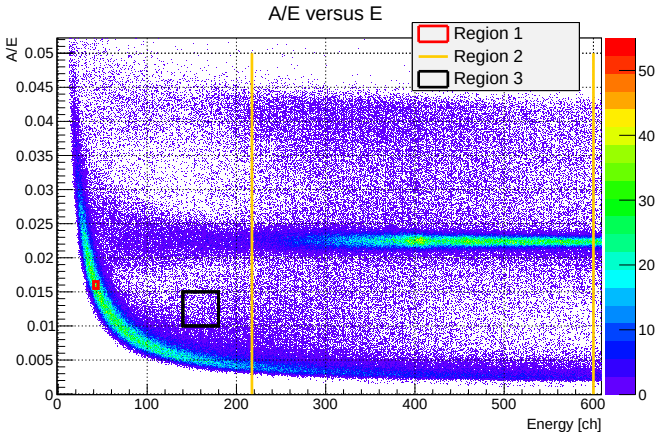
The most simple way to learn more about these events is "looking inside" the regions of interest. It has been written a software framework that's able to put in parallel a Tier1 and Tier2 file, selecting from both of them only the events specified by a user's cut (both at Tier1 that Tier2 level), and extracting from them any desired variable or object (through a method that has only to be implemented as algorithm). All the extracted information are saved in two `TTree`s, one for Tier1 and one for Tier2 methods. The variables and objects used for this inspection are the discriminating variables chosen for ANN (see section 2.1) plus the waveforms of all selected events. Through the writing of another program, it becomes possible to "see" any selected event. The inspection has been made on three regions, described in table 3.1.

	min E [ch]	max E [ch]	min A/E	max A/E
Region 1	40	45	0.0155	0.0165
Region 2	217	600	/	/
Region 3	140	180	0.010	0.015

Table 3.1: Ranges of inspection regions



(a) *β energy spectra with energy cut at 40 keV*

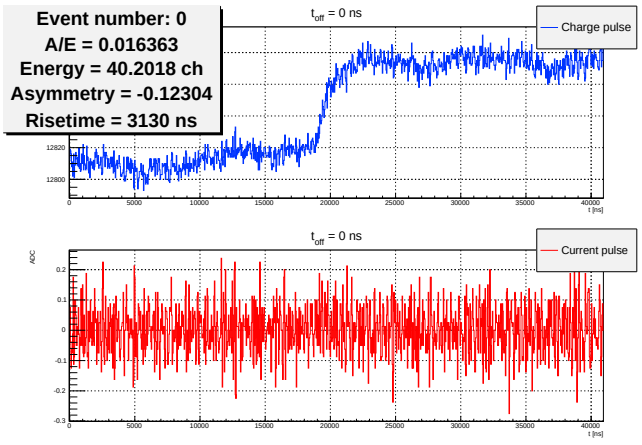


(b) *Selected regions for the inspection. Ranges at table 3.1.*

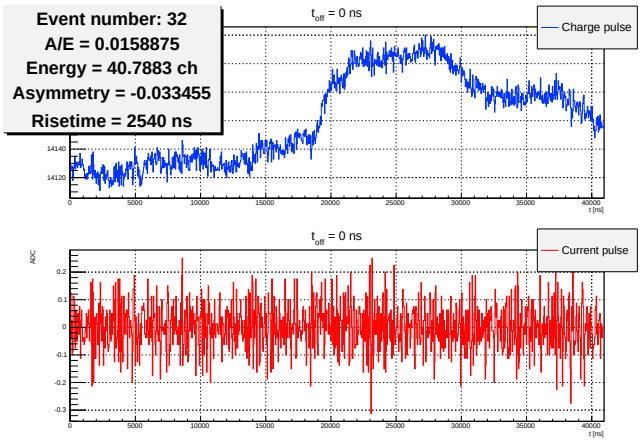
Figure 3.5: Graphics to inspect the suspicious events.

The three selected regions are visible in figure 3.5b. Motivation for such selections follows. First of all it has been chosen an energy cut at 40 keV to roughly discriminate between "suspicious" and physical events, as shown in figure 3.5a. Then, from the A/E versus E plot, it has been chosen one region (Region 1) entirely enclosed in the "hyperbolic" region, another big one (Region 2) as a sample of the whole physical events, and a last region (Region 3) to inspect events situated between the hyperbolic and SSE regions.

By observation of events in these regions, it becomes possible to make assumptions on the "hyperbolic" region. First of all, almost all events belonging to Region 1 are like those reported in figure 3.6a and 3.6b. Comparing these events with those of Region 2, one can notice that there is a very big difference between them. Events shown in figures 3.7a, 3.7b, 3.7c and 3.7d are all physical events (like all those events founded in this region), respectively they are SSE, MSE, p^+ and n^+ candidates. Focusing on the n^+ candidate it is possible to see that it's a very slow event in comparison with SSE, MSE and p^+ candidates. The majority of events in Region 3 is very similar to those shown in figures 3.8a and 3.8b. They are probably n^+ lower in energy and time, however they are very different from Region 1 events too. Another information that turns out from this inspection is that, focusing only on the ANN discriminating variables, events from Region 1 are dangerously similar to those slow n^+ events we have seen in Region 2. Therefore, taking assumption that the "hyperbolic" events are not of interest for us, it is very important to remove as much as possible of them. Moreover, we must be careful not to lose all events between pathological region and SSE one (i.e. events shown in Region 3). An optimal cut has to be designed.

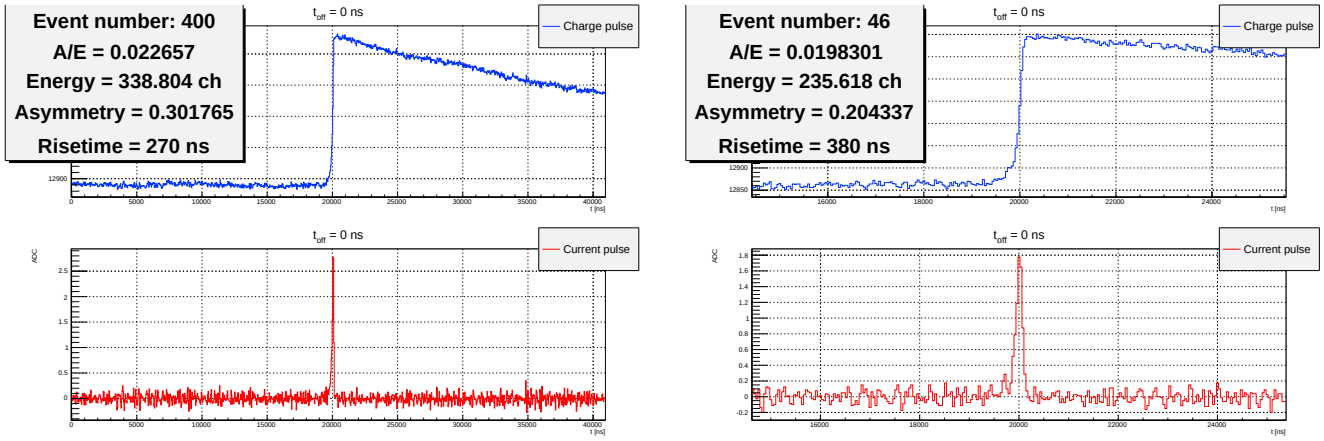


(a)

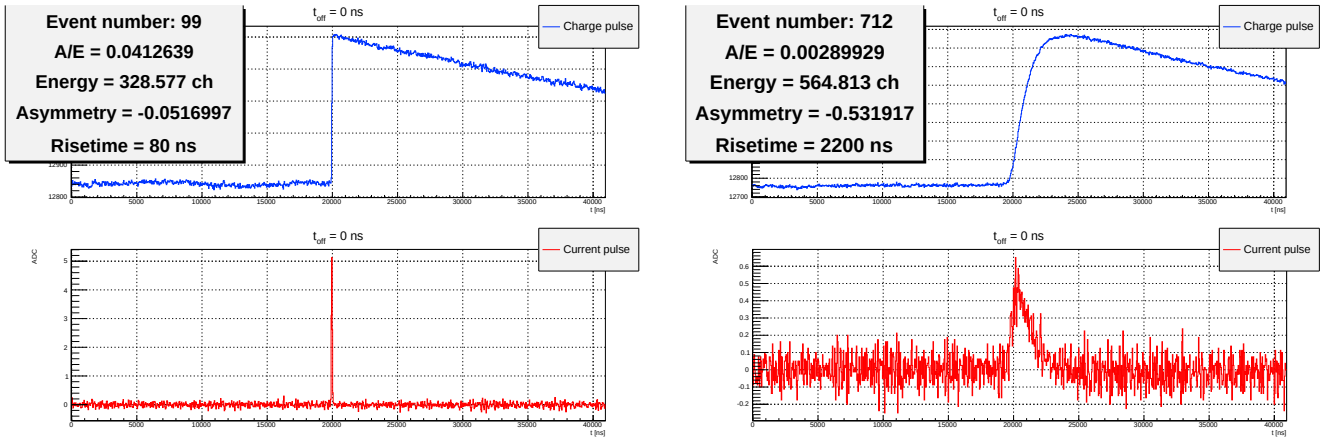


(b)

Figure 3.6: Typical events from Region 1

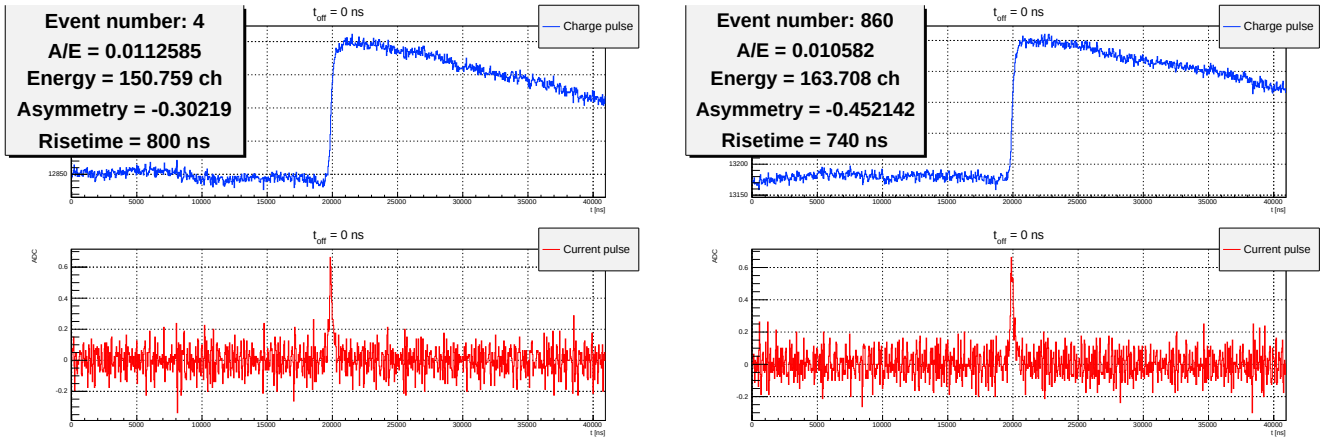


(a) *Candidate SSE event* (b) *Candidate MSE event*



(c) *Candidate p^+ event* (d) *Candidate n^+ event*

Figure 3.7: Events from Region 2



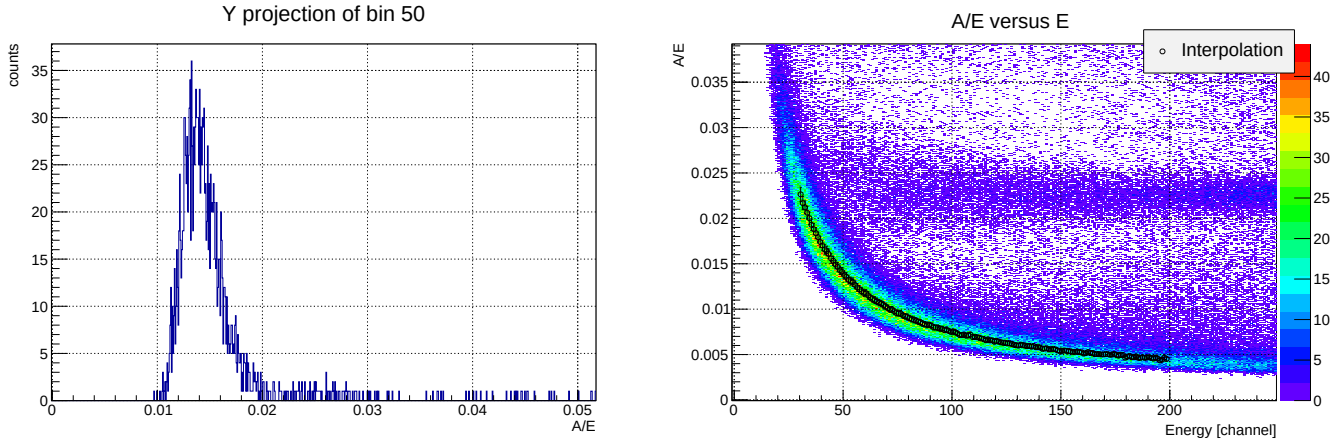
(a) (b)

Figure 3.8: Typical events from Region 3

Before starting to decide the cut for pathological events, it is very useful trying to understand the physics behind this "hyperbolic" region. From figure 3.6a and 3.6b it can be seen that the charge pulse derivative (current pulse) is almost entirely a background signal. That is, for all the events of this type the rise of the charge pulse is so slow and background dominated that the derivative is only affected by background fluctuations and do not perceive the real rise of the pulse. Thus, the current signal lacks a true peak and appears as a background noise. In such a way, the current amplitude A will be about the same for all events

that behave like this, and so the hyperbole in the A/E versus E plot is perfectly explained.
All of this, however, is only a first approximation, indeed, first of all, the derivative of charge pulses shown in previous figures is a "raw" derivative, while the derivative computed by the `GEMDCurrentPSA` module of `GELATIO` follows the application of some noise filters, that leads to a better current pulse signal. Anyway, what has been just shown is the basic and main concept to understand the hyperbolic-like distribution of such events.

Once got here, it's possible to proceed with greater awarness than before preparing the cut for the pathological region. Exploiting the simple bilogarithmic shape of the region, the idea is to interpolate the center of the bi-logged region with a straight line, then shift the latter of a certain parameter that has to be an index of the region's dispersion. All the events under this shifted stright line will be cut off.
We start estimating the center of such a region. To do this it has been used the Y projection: for each X bin of A/E vs. E plot (see for example figure 3.4a), Y projection was computed for all the X bins where the hyperbolic region showed an high density (in the M30 case from channel 30 to 200).



(a) *Y projection of bin 50 of A/E versus E plot, M30 measurement.*

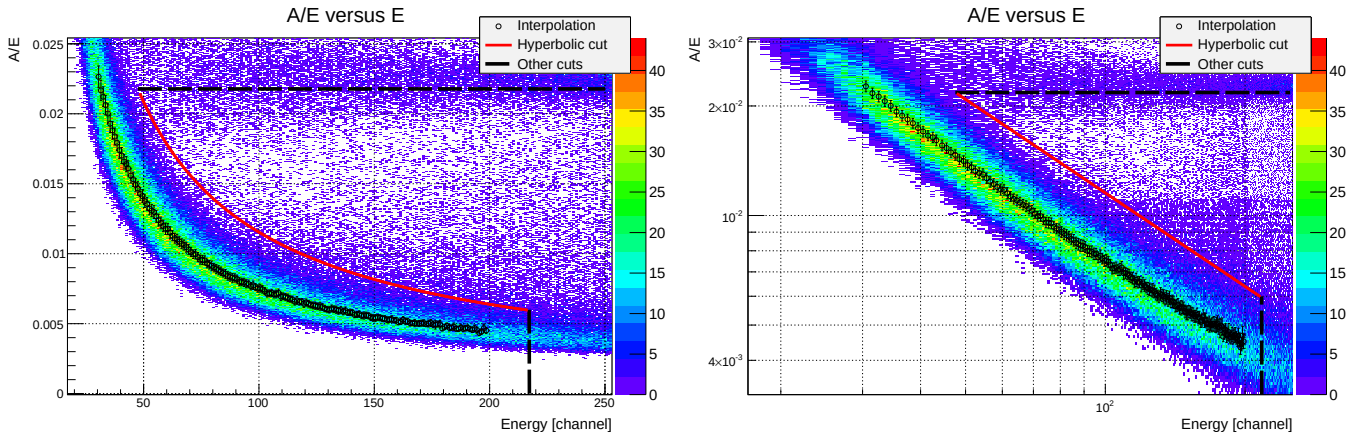
(b) *Interpolation of M30 hyperbolic region.*

Figure 3.9: View of one Y projection and hyperbolic interpolation of M30 measurement.

Output of an Y projection can be seen in figure 3.9a. For each Y projection in the established range has been computed a gaussian fit on the main peak. The graph of centroid position versus the X projected bin center returns the desired interpolation of the hyperbolic region, as shown in figure 3.9b. Then, after linear fitting the bi-logged interpolating graph, it has been computed the quadratic sum of all distances of graph's points from the linear function. Call σ_M the square root of this value. As an order of magnitude, σ_M represents the bi-log linear law dispersion of the region's mean values. Thus, if we are interested in the dispersion of the total region, its value will be about $\sqrt{N} \cdot \sigma_M := \sigma_L$, where N is the number of region interpolating points. The bi-log linear function is perpendicularly up-shifted by $2\sigma_L$: all that's located below is cut off. Examples of result can be seen in figure 3.10a and 3.10b. All linear fit results and σ_L values are in table 3.2.

	m	q	Range $[ch]$	σ_L
M30	-0.854 ± 0.006	-0.41 ± 0.01	$[30, 200]$	0.116
M31	-0.872 ± 0.007	-0.38 ± 0.01	$[32, 169]$	0.090
M32	-0.889 ± 0.007	-0.35 ± 0.01	$[30, 170]$	0.085
M33	-0.886 ± 0.007	-0.36 ± 0.01	$[30, 164]$	0.110
M34	-0.859 ± 0.005	-0.40 ± 0.01	$[25, 217]$	0.139

Table 3.2: Values of linear bilogarithmic fits. Fitted function is $y = m \cdot x + q$.



(a) *No log_{10} scale.* (b) *log_{10} scale.*

Figure 3.10: Applied cuts for M30 measurement (last and not visible cut is that on Q_β of ^{90}Sr).

As shown in figure 3.10a, the "hyperbolic" cut is simply added to the others. Hyperbolic cut must be stopped by looking at energy spectrum of figure 3.5a. Pathological events, indeed, are characteristics only of low energy region. Moreover, looking at figure 3.5b and 3.7d, it is possible to notice that the typical n^+ candidate founded in Region 2 lies exactly in the tail of hyperbole. In such a way, do not stop the hyperbolic cut would lead to an high loss of signal. Energy stop value was chosen at 40 keV, that correspond to 217 ADC channels. The other stop value of hyperbolic cut is the lower limit of SSE band.

Before going on with ANN training phase it's useful to compute some percentages.

	M30	M31	M32	M33	M34
Events in Q_β (%)	92.33	94.21	83.50	95.32	88.69

Table 3.3: Percentage of events that lie in the interval $[0, Q_\beta]$.

In table 3.3 are reported percentages of events before Q_β of ^{90}Sr. For M34 measurement at about 20000 ch there is an high peak due to an α particle. Events percentage is recomputed excluding this peak: the new percentage is of 90.95%.
Now we'll assume that the whole considered data sample always goes from 0 to Q_β.
All percentages of table 3.4 are computed on this interval for each data sample.

	M30	M31	M32	M33	M34
p^+ (%)	15.85	15.73	26.28	16.62	16.69
SSE (%)	33.68	37.62	32.60	41.75	30.15
Pathological (%)	18.63	16.43	14.36	14.39	20.14
Total cut (%)	68.16	69.78	73.24	72.76	66.98

Table 3.4: Cut percentages of ^{90}Sr from 0 to Q_β.

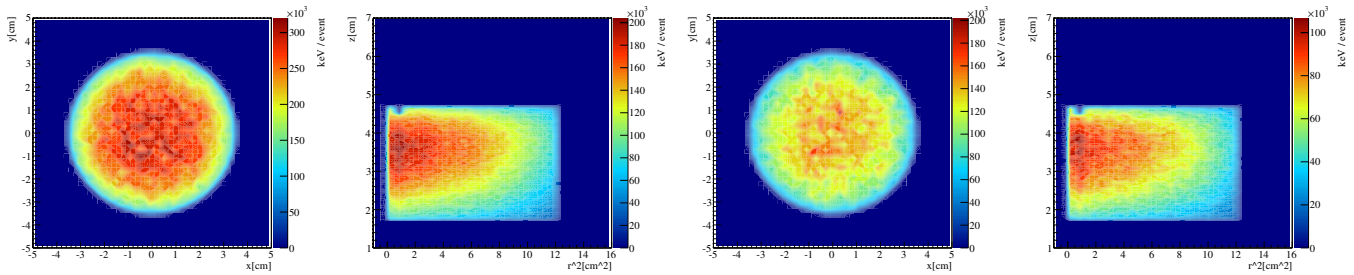
In table 3.4 are shown every cut's events percentages. Looking at p^+ percentages, is easy to understand that measurement M32 is that with source on p^+ contact.
Cut used to compute values of table 3.4 are those previously discussed:

- $A/E > 0.0231$ for p^+.
- $A/E \in [0.0218, 0.0231]$ for SSE.
- $\left(A/E < 0.0218\right)$ `AND` $\left(E < 217\right)$ `AND` $\left(log_{10}(A/E) < m \cdot log_{10}(E) + q'\right)$ for pathological events. Where m is the slope of the bilogarithmic linear fit and $q' = q + 2 \cdot \sigma_L \cos(\arctan m)$ (see table 3.2).

The total number of events cut off (in the energy region of interest) is always higher than 65% for each data sample. All survived events are mainly n^+ and MSEs.

3.3 Training networks

So far we have chosen a smart cut to prepare an n^+ data sample as pure as possible. However, it's clear that the n^+ purity to which we aspire needs the exploiting of more intelligent method than a simple cut. We also know that all events that survived the cut are both n^+ and MSEs. So, now, the purpose is to exploit artificial neural networks to purify the cut data sample, making a distinction between n^+ and MSEs.
What we have for the neural network is a "dirty" signal sample, with a "dirt" percentage (MSE) that we don't know precisely. In such a situation it's absolutely needed an high purity background sample. A pure background sample can be chosen as a point of reference for selecting the cut value and helps to identify the background component of signal sample looking at the signal and background overlap in the output of classifier (see figure 2.2b).
The chosen background sample meets our requirements very well. As previously said in section 3.1, chosen backround sample is ^{60}Co: it has two γ lines at about 1173 and 1332 keV, that are certainly internal events. A further proof of this is given by the Monte Carlo simulation of such a measurement made by Katharina Von Sturm.



(a) 1173 keV (b) 1332 keV

Figure 3.11: Monte Carlo simulation of the ^{60}Co source on the top surface of upside-down BEGe GD91C. Taken from [23].

Results of simulation can be seen in figure 3.11a and 3.11b: they are respectively referred to 1173 and 1332 keV ^{60}Co peaks. For each figure, the first graphic is a top view of the upside-down detector GD91C upper surface (p^+ surface: x vs. y), the second is the graphic of $r^2 = x^2 + y^2$ versus z (height). Colour indicates the number of hits, i.e. number of interactions. From these figures it is possible to see that all the events are located into the inner part of detector.
We choose as background sample these two γ peaks of ^{60}Co. Sample is selected by fitting γ peaks with function 2.2 and selecting all events lying into 3σ of fitting function.

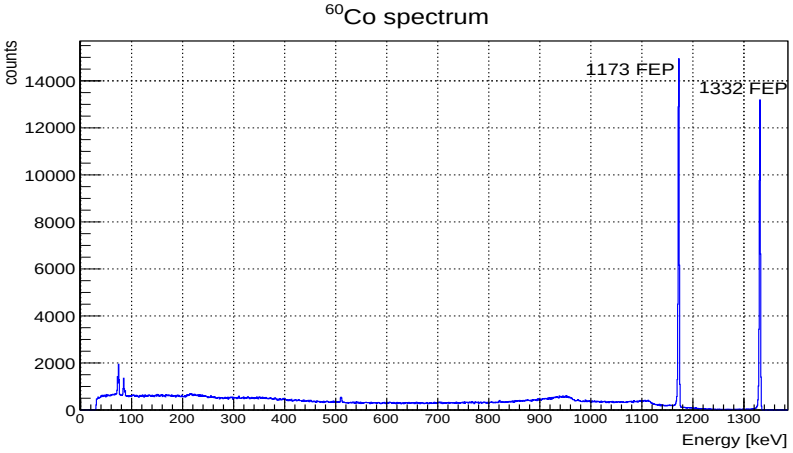


Figure 3.12: ^{60}Co calibrated spectrum.

In figure 3.12 is shown ^{60}Co energy spectrum. The chosen background sample, however, is incomplete. The two ^{60}Co peaks are both full energy peaks, so they are made up almost entirely by multiple site events (see figure 1.3). As said in section 3.2, the decided SSE cut could be underestimated, leaving a certain amount of SSEs in the prepared signal sample. By using only ^{60}Co FEPs as background sample these survived SSE events won't be recognized and discriminated by network. On the contrary, adding to background sample a pure SSE peak, it will be possible to recognize also these survived SSEs, as successfully done in section 2.2. The chosen background SSE peak is the DEP peak of M1 Thorium measurement, that we can see in figure 2.3b. The DEP of ^{228}Th has been fitted with function 2.2, selecting all the events lying into 2σ of fitting function.
Now we are going to proceed as follows. Signal sample used for the ANN training phase will be the cut sample of M30 measurement. Then, the application phase will be carried out applying the trained network on the other ^{90}Sr cut data samples.

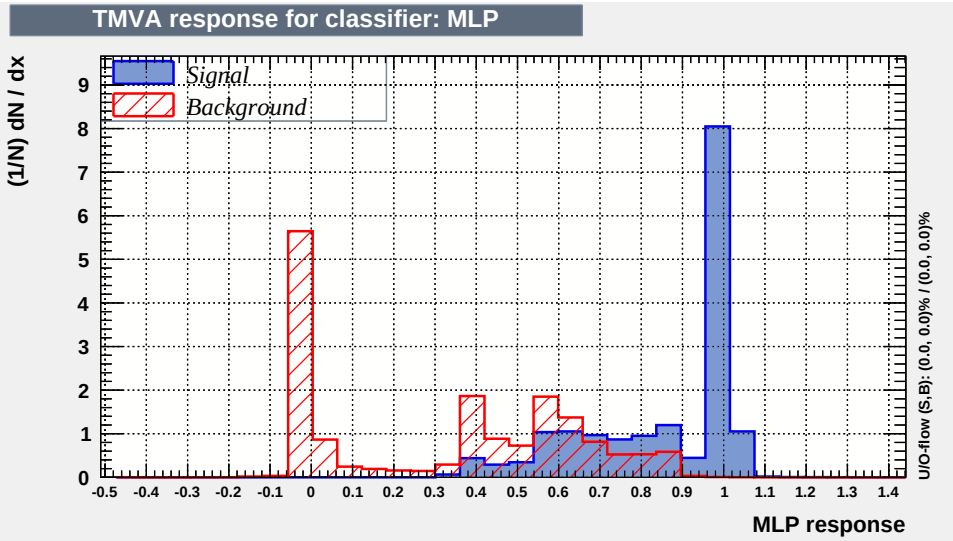


Figure 3.13: Output of classifier. To notice the intersection region between signal and background.

Looking at the output of classifier in figure 3.13, it can be seen that there is a considerable overlap region between signal and background. Remembering that the training signal sample is polluted by the same type of events that populate the background sample, an overlap region is expected. It may seem as if the neural network has understood that there is a considerable number of events shared by signal and background samples, and so it has put them in the region between "pure signal" and "pure background", leaving to us the choice of classifying these events as "signal in background" or "background in signal".
The choice of the cut value follows from the consideration that background sample employed for the training is very pure, while this is false for the signal one. Thus, we select the cut value trying to maximize the background rejection. The cut is chosen at 0.9 of output classifier. Background rejection and signal efficiency are shown in table 3.5.

Cut Value	Signal Efficiency	Background Rejection
0.90	0.568	0.998

Table 3.5: Chosen cut and related parameters.

3.4 Results

Networks has been trained using as signal the M30 data sample (cutted as described in section 3.2) and as background the two ^{60}Co FEPs and the ^{228}Th DEP. Application has been made on the other four ^{90}Sr measurements opportunely prepared as done for training signal sample, in such a way to be able to discriminate mostly only between n^+ and MSE. Cutting percentages are shown in table 3.6.

	M31	**M32**	**M33**	**M34**
Discarded events (%)	47.82	49.48	56.14	36.65
Survived events (%)	52.18	50.52	43.86	63.35

Table 3.6: Results of ANN application phase.

To test the goodness of ANN application results the best way is to look at significant plots as done in figures 2.3a and 2.3b for ^{228}Th energy spectrum. But now the situation is different: energy spectrum doesn't contain any useful information. Firstly, it is possible to check the result on the A/E vs. E graph.

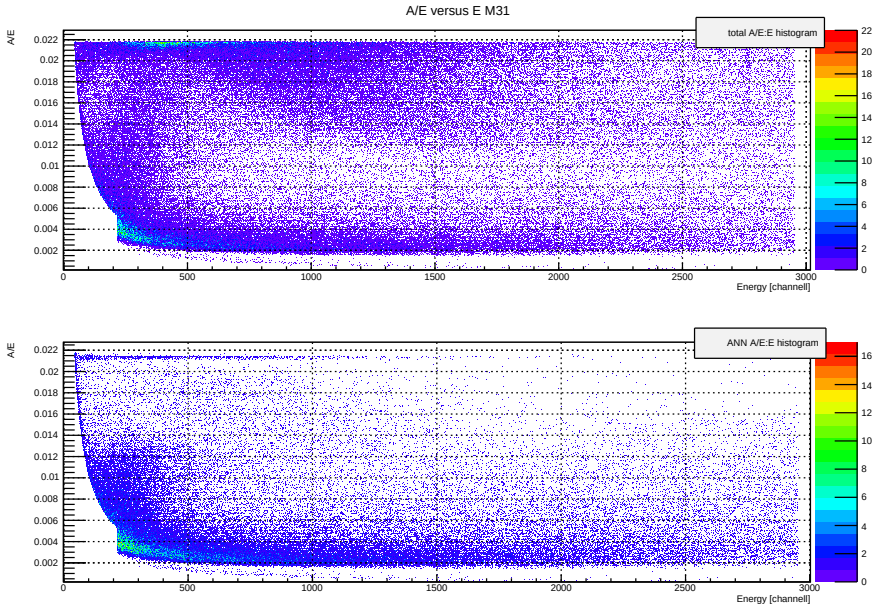
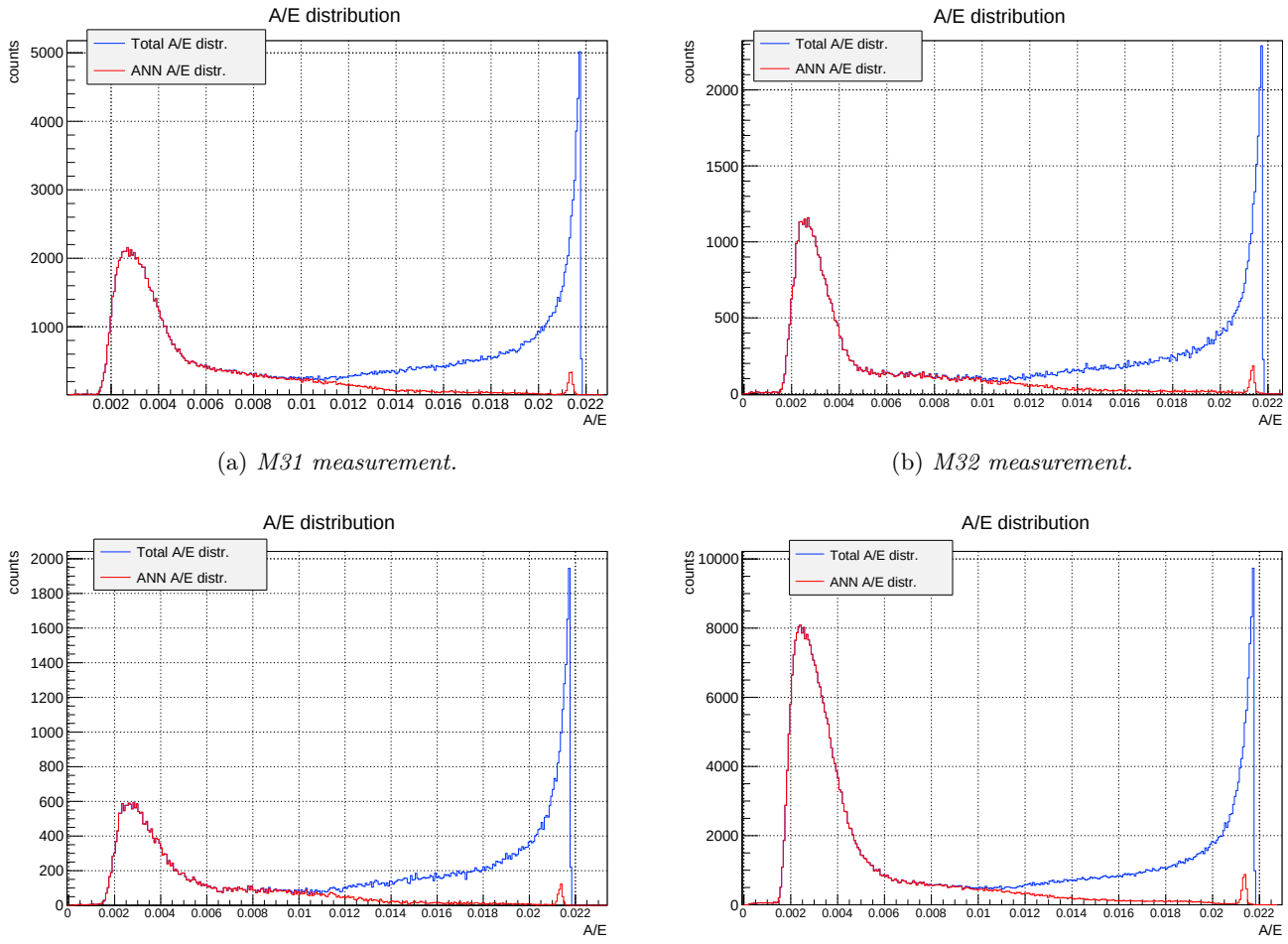


Figure 3.14: A/E vs. E plots for M31 measurement before and after ANN application. Notice that colour legend is changed.

Looking at figure 3.14 it can be seen that ANN finds and removes the SSE events survived to the sample cutting phase. Also the region just below SSE band is greatly reduced in number of events. This is the main MSE region that, in the A/E distribution, clusters below SSE peak with a shape described by the equation 1.3 (see figure 1.6). The region least affected by the cuts is the low A/E region and low energy region. This is in agreement with the shape of beta ^{90}Sr spectrum and with the fact that n^+ have a ballistic energy defect due to charges recombination. A clearly vision of this SSE and MSE cut is given by the A/E distribution histograms before and after network application.
From figures 3.15a, 3.15b, 3.15c and 3.15d it's visible what it has just been said. Survived SSE peak is completely cut down, so as the exponential MSE tail described in figure 1.6 by equation 1.3. From here we can see that network recognizes both MSEs and SSEs.
However, the most important variable to discriminate n^+ events from the others is the pulse risetime. Looking at risetime histograms give us important information on the network application goodness.



(a) *M31 measurement.*

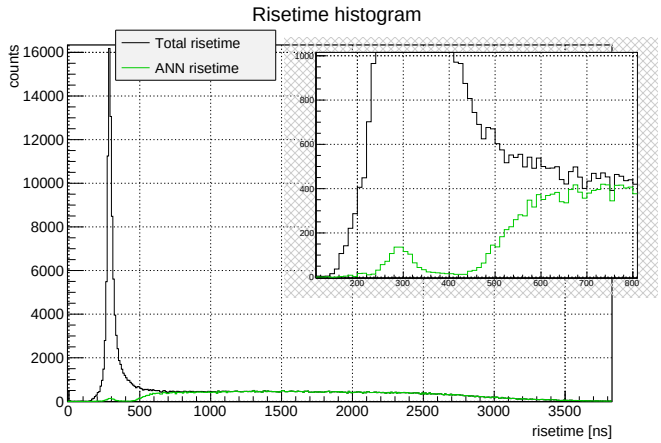
(b) *M32 measurement.*

(c) *M33 measurement.*

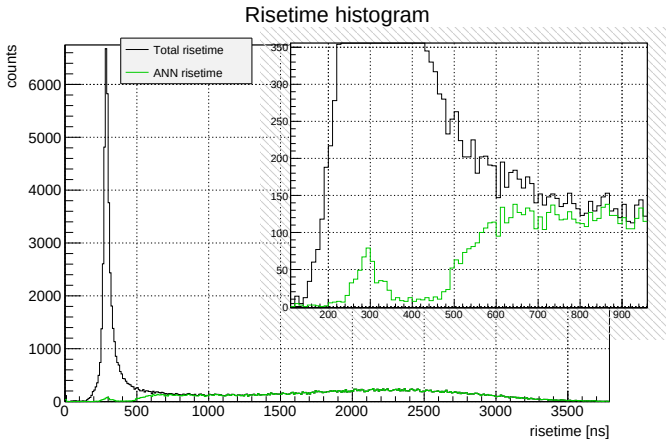
(d) *M34 measurement.*

Figure 3.15: A/E distributions before and after neural network application.

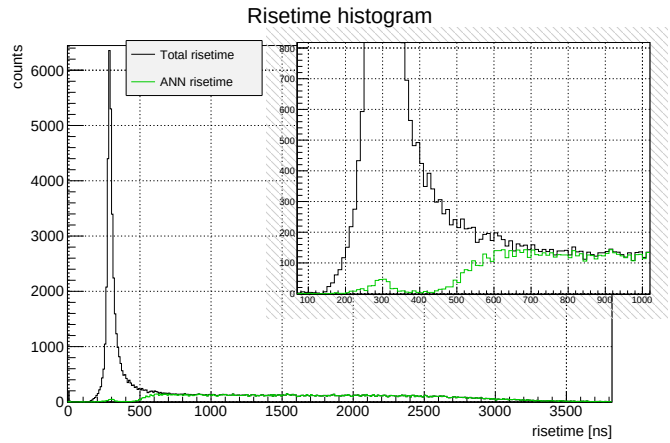
Figures 3.16a, 3.16b, 3.16c and 3.16d prove that network is able to discriminate between internal and superficial (n^+) events. All the cut off events belong to the low-risetime peak. The best feature of these plots is not only that all cut off events belong to risetime peak, but also that the risetime peak cut is a smart cut. It can be seen in the zoomed histograms that cutting is not brutal nor simple, but smooth. The region between risetime peak and risetime tail is smoothly brought to zero, and smooth is the arising of the survived risetime tail. This could not be reproduced by a simple risetime cut.
A problem may be given by the hypothetical survived pathological events. The fact that the high-risetime region has been left untouched by network application, says us that if some (not negligible) number of pathological events has survived the preparing cut, it has not been removed by application phase. This would be in agreement with what we noticed in section 3.2: n^+ and pathological events are very similar regarding the discriminating variables used for the network.



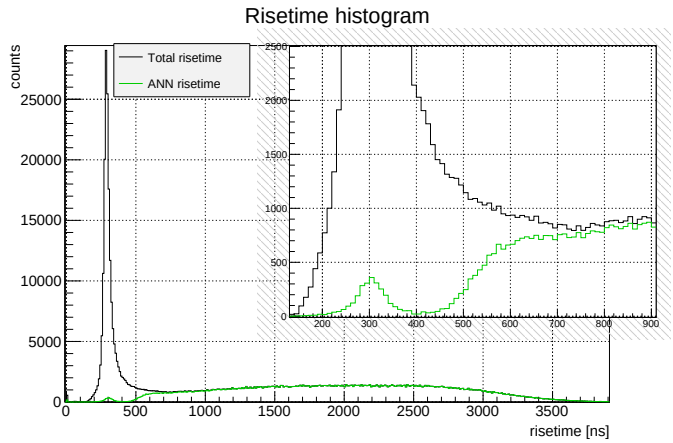
(a) *M31 measurement.*



(b) *M32 measurement.*



(c) *M33 measurement.*



(d) *M34 measurement.*

Figure 3.16: Risetime histograms before and after neural network application. The smaller histogram reported on each canvas is the zoom on the cut down risetime peak.

3.5 Conclusions

Discrimination by Artificial Neural Networks proves to be a correct method to reach the point of this work. Chosen settings and variables confirm their validity also in discriminating between superficial and internal events. After preparing the best signal sample as possible, the network was able to discriminate between superficial n^+ events and internal MSE (and SSE) ones. Discrimination has confirmed that n^+ events are characterized by a much higher ristime and a lower A/E ratio compared to internal events. Four high-pure samples of superficial events are available now.

Possible improvements and future developments:

- More detailed study of the so called "pathological events", trying to remove all of them from the data sample (their presence is probable in the obtained final samples).
- Study of new discriminating variables for a best discrimination of n^+ events.
- Employment of others TMVA classifiers instead of ANN. Study and comparison of the results to determine the best discriminator for such a problem.
- Adapt the method to the GERDA experiment. ANN application on GERDA spectra, using the obtained final n^+ samples as background for the training phase.

References

[1] K. H. Ackermann et al. "The GERDA experiment for the search of $0\nu\beta\beta$ decay in ^{76}Ge". In: *Eur. Phys. J. C.* 2330.73 (2013).

[2] M. Agostini et others (GERDA collaboration). "Results on neutrinoless double-β decay of ^{76}Ge from Phase I of the GERDA experiment". In: *PRL* 122503.111 (2013).

[3] A. Dietz H.V. Klapdor-Kleingrothaus I.V. Krivosheina and O. Chkvorets. "Search for neutrinoless double beta decay with enriched ^{76}Ge in Gran Sasso 1990-2003". In: *Phys. Lett. B* 198.586 (2004).

[4] C. E. Aalseth et al. "The IGEX Ge-76 neutrinoless double-beta decay experiment: Prospects for next generation experiments". In: *Phys. Rev.* D65 (2002), p. 092007. arXiv:`hep-ex/0202026`.

[5] L. Pandola M. Agostini and P. Zavarise. "Off-line data processing and analysis for the GERDA experiment". In: *J. Phys.: Conf. Ser.* 012047.368 (2012).

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

[7] Rene Brun and Fons Rademakers. "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 (1997), pp. 81–86.

[8] M. Agostini et others. "The MGDO software library for data analysis in Ge neutrinoless double-beta decay experiments". In: *J. Phys.: Conf. Ser.* 042027.375 (2012).

[9] D. Budjas et al. "Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double β decay experiment". In: *J. Instrum.* 4 (2009).

[10] M. Agostini et al. "Pulse Shape Discrimination for GERDA Phase I Data". In: *Eur. Phys. J. C.* 2583.73 (2013).

[11] G. F. Knoll. *Radiation detection and measurement.* 4th ed. Wiley & Sons, 2010.

[12] 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 ^{76}Ge". In: *JINST* P03005.6 (2011).

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

[14] D. Budjas. "Germanium detector studies in the framework of the GERDA experiment". Dissertation. MPI for Nuc. Phys., 2009.

[15] Andreas Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv:`physics/0703039`.

[16] M. Perini. "Germanium detectors pulse shape analysis using artificial neural networks". LT thesis. Unipd, 2013.

[17] K. Debertin and R. G. Helmer. *Gamma- and X-ray spectrometry with semiconductors detectors.* North-Holland, 1989.

[18] K. Gusev T. Bode D. Budjas and B. Lehnert. *Upside-Down BEGe: GD91C.* 2013.

[19] Dusan Budjas Bjoern Lehnert. *Updates on GD91C LNGS.* 2013.

[20] Laboratoire National Henri Becquerel. *Recommended Data.* URL: `http://www.nucleide.org/DDEP_WG/DDEPdata.htm`.

[21] Bjoern Lehnert. *Priv. Com.* 2013.

[22] G. R. Gilmore. *Pratical Gamma-ray Spectrometry.* 2nd ed. Wiley & Sons, 2008.

[23] K. Von Sturm. *Priv. Com.* 2013.