THE EUROPEAN
PHYSICAL JOURNAL C

Test of pulse shape analysis using single Compton scattering events

I. Abt, A. Caldwell, K. Kröninger[a], J. Liu, X. Liu[b], B. Majorovits

Max-Planck-Institut für Physik, Föhringer Ring 6, 80805, München, Germany

Abstract. Procedures developed to separate single- and multiple-site events in germanium detector are tested with specially selected event samples provided by an 18-fold segmented prototype germanium detector for phase II of the germanium detector array, GERDA. The single Compton scattering, i.e. single-site, events are tagged by coincidently detecting the scattered photon with a second detector positioned at a defined angle. A neural network is trained to separate such events from events which come from multi-site dominated samples. Identification efficiencies of $\approx 80\%$ are achieved for both single- and multi-site events.

PACS. 23.40.-s; 14.60Pq; 29.40.-n

1 Introduction

Photons of energies around 2 MeV have a high probability to interact in germanium through Compton scattering. The mean free path of the process is ≈ 4.6 cm. If a photon Compton scatters only once inside a germanium detector, the recoiling electron deposits its energy most likely within a 1 mm range, resulting in a so-called single-site event (SSE). If, in contrast, a photon interacts through pair production or scatters multiple times, energy can be deposited at different locations separated by typically a few centimeters, resulting in a so-called multi-site event (MSE). The charge carriers created by the energy deposition in the germanium detector drift towards the anode and cathode of the detector. While the charge amplitude of the induced pulse is determined by the number of carriers (thus by the energy deposited), the time spectrum of the pulse (pulse shape) is determined by the location(s) of the energy deposition(s) and thus the charge drifting times. MSEs are expected to have more involved pulse shapes than SSEs, and thus, pulse shape analysis (PSA) can be used to separate the two classes of events [1–6].

One application of PSA is the background rejection in experiments searching for neutrinoless double-beta decay ($0\nu\beta\beta$) in ^{76}Ge-enriched detectors, such as the GERDA experiment [7]. The expected $0\nu\beta\beta$ signal events have two electrons in the final state with a total energy of 2.039 MeV. These are mostly SSEs. A large fraction of the expected background events are induced by external photons with energy depositions around the Q-value. These events are

expected to be predominantly MSEs which can be rejected by PSA.

In order to study and improve the performance of PSA, SSE- and MSE-dominant data samples have to be collected independently of the pulse shape. In this paper a method to collect single Compton scattering events (SCS) as an SSE-dominant sample is investigated in more detail. The energy of the scattered photon in an SCS event can be calculated given the incoming photon energy and the scattering angle. Therefore, SCS events can be collected by positioning a second germanium detector at a specific angle with respect to the first detector and using it to tag escaped photons with the correct energy [2]. If the incoming photon has an energy of 2.614 MeV as emmitted by a ^{208}Tl source, a photon Compton scattered at 72° has an energy of 575 keV. This signature is used to tag the single recoiling electron inside the first germanium detector. The energy in the event is equal to the germanium $0\nu\beta\beta$ Q-value. The location of the energy deposition of the electron within the detector volume is controlled by positioning the source and the second detector correspondingly.

Another common method to collect an SSE-dominant sample is to select the double-escape events (DEP) [1, 3, 4, 6]. The incoming photon interacts with the germanium detector through pair production and the two 511 keV photons from the positron annihilation escape the detector without further interaction. The electron and positron mostly deposit their energies very locally and result in an SSE. Another useful sample contains so-called single-escape events (SEP) where only one 511 keV photon escapes. The other photon mostly deposits its energy at locations different from those of the electron and positron. Thus, SEP events provide an MSE-dominant sample with energy deposition close to the $0\nu\beta\beta$ Q-value.

[a] *Current address:* II. Physikalisches Institut, Universität Göttingen, Germany
[b] e-mail: xliu@mppmu.mpg.de

However, the DEP events are not a perfect test sample for the expected $0\nu\beta\beta$ events. If the two photons escape the detector, the interaction point is more likely close to the detector surface as compared to SCS events. $0\nu\beta\beta$ events, on the other hand, are distributed evenly within the detector volume. In addition, DEP and $0\nu\beta\beta$ events have different energies. A DEP event induced by a 2.6 MeV photon from a ^{208}Tl source has an energy of 1.59 MeV, quite different from the $0\nu\beta\beta$ Q-value. In these respects studies with SCS samples suffer less from systematic effects.

The experimental setup and the data collection are described in Sect. 2. The Monte Carlo simulation is also included in this chapter. It is used to verify that the collected SCS samples are SSE-dominated. In Sect. 3 a PSA package based on an artificial neural network (ANN) is presented. The training methods are described and the results given.

2 Experimental setup, data selection and MC simulation

2.1 Experimental setup

The experimental setup is illustrated in Fig. 1. The segmented germanium detector under study is a prototype detector for Phase-II of the GERDA experiment [7]. The true coaxial 18-fold segmented n-type HPGe dectector has a weight of 1.63 kg and the dimensions are 69.8 mm height and 75.0 mm diameter; the inner hole has a diameter of 10.0 mm. The segmentation scheme is 3-fold along the vertical axis and 6-fold in the azimuthal angle (see Fig. 1). Signals from the 18 segments and the core of the detector are amplified by charge sensitive pre-amplifiers and read out by a Pixie4 DAQ system [9] with 14-bit ADC's at a sampling rate of 75 MHz. The resolution (FWHM) of the core is ≈ 3.5 keV at 1.3 MeV and those of the segments are between 2.5 and 4.0 keV. A time resolution of roughly 10 ns can be achieved with the sampling rate used. This corresponds to a position resolution of ≈ 1 mm inside the detector volume.[1] More information about the segmented detector and the DAQ system can be found in [10].

A 100 kBq ^{228}Th source is positioned at a distance of 23 ± 1 cm from the center of the segmented detector and faces the center point of segment 14, as illustrated in Fig. 1. A second non-segmented and well-type germanium detector, a Canberra reversed germanium detector (REGe) [11], is positioned at the same height with the closed end facing the segmented germanium detector. The distance from the closed end surface to the center of the segmented detector is 23 ± 1 cm. The REGe crystal is 60 mm in height and 65 mm in diameter. It has a resolution (FWHM) of 2.3 keV at 1.3 MeV. It is used to tag the photons scattered mostly in segment 14. The geometrical acceptance of the REGe detector results in recorded SCS events with scattering angles between $\approx 65°$ and $\approx 80°$ corresponding to energy depositions in the segmented detector between
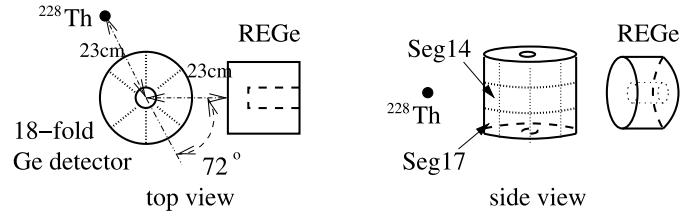
Fig. 1. Schematic of the experimental setup with the 18-fold segmented germanium detector as target and the REGe detector to tag photons at $72°$ (not to scale). The *dotted lines* illustrate the segment boundaries

≈ 1940 keV and ≈ 2110 keV. The precision of the alignment of the REGe detector with respect to the ^{228}Th source and the segmented detector is $\approx 5°$.

The energy thresholds for all channels are set to 100 keV. A coincidence trigger is required between the core of the segmented detector and the REGe with a coincident time window of 500 ns. Due to a technical limitation of the coincidence trigger of the DAQ system, only four channels could be read out. Thus, for each coincidence trigger, only the energies of the core (E_{Core}), segment 14 (E_{Seg14}), segment 17 (E_{Seg17}) (below segment 14, as illustrated in Fig. 1) and the REGe (E_{REGe}) were recorded. 300 time samples were taken for each pulse for all 4 channels. This corresponds to a time window of 4 μs including 1 μs before the arrival of the trigger. In this analysis, however, only the core pulses are used for the PSA.

The actual coincidence trigger rate was ≈ 12 Hz. The independent trigger rates of the segmented detector and of the REGe detector were both ≈ 2000 Hz. This results in an accidental coincidence rate of ≈ 2 Hz. The coincidence trigger rate without the ^{228}Th source is < 0.1 Hz. Therefore, without further cuts, $\approx 20\%$ of all events are expected to originate from accidental coincidences.[2] However, the fraction of accidental coincidence events among the selected SCS events is negligible, as discussed in the next section.

2.2 Event selection

In total 360 000 coincident events were collected. Four different data samples are selected:

- Γ_{SCS}: single-Compton-scattering (SCS) events
  $|E_{\text{Core}} + E_{\text{REGe}} - 2614.5| < 5.0$ keV
  and $1940 < E_{\text{Core}} < 2090$ keV *and* $|E_{\text{REGe}} - 583.2| > 3.0$ keV
- $\Gamma_{2.6}$: events with the 2.6 MeV photon fully absorbed in the segmented detector
  $|E_{\text{Core}} - 2614.5| < 5.0$ keV
- Γ_{DEP}: DEP events
  $|E_{\text{Core}} - 1592.5| < 5.0$ keV (two 511 keV photons escape.)
- Γ_{SEP}: SEP events
  $|E_{\text{Core}} - 2103.5| < 5.0$ keV (one 511 keV photon escapes).

The Γ_{SCS} sample is selected through three cuts. The allowed window of ± 5 keV of the sum energy of both detec-

---

[1] The typical drift velocity of the charge carriers inside a germanium detector is ≈ 1 cm per 100 ns.

[2] The fraction is expected to differ for different energy ranges, as the trigger rate varies.
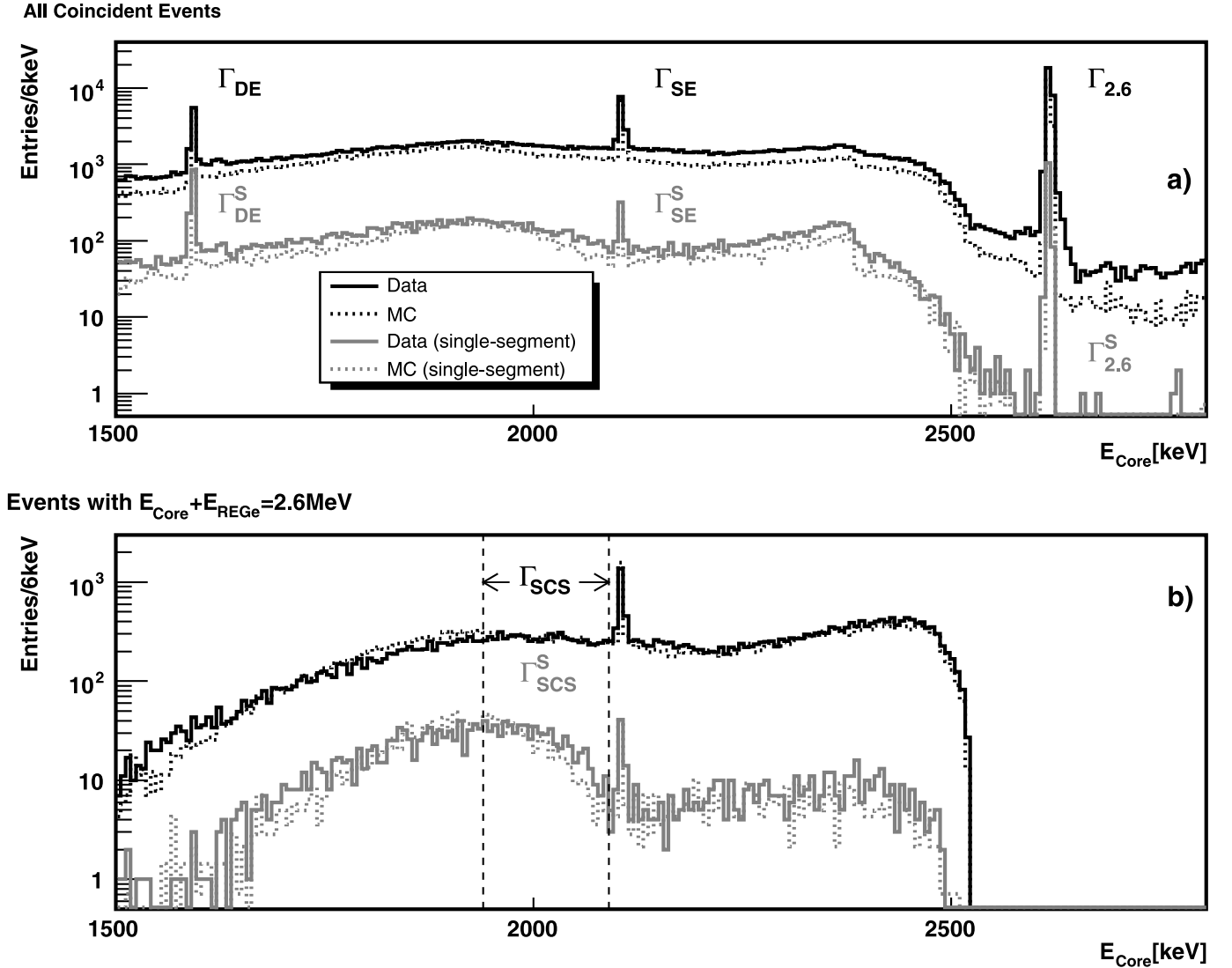
Fig. 2. E_{Core} distributions **a** for all coincident events **b** for events with $E_{\mathrm{Core}} + E_{\mathrm{REGe}} = (2614 \pm 5)$ keV. The 8 selected samples are indicated. The predicted distributions from the Monte Carlo are shown as well

tors around 2614.5 keV covers about three times the combined energy resolution (3σ) of the detectors. The geometrical acceptance for SCS events extends to 2110 keV, but SEP events would contaminate the sample, as they have a core energy of $E_{\mathrm{Core}} = 2103.5$ keV in this setup. They are excluded by removing events with the core energy of the segmented detector above 2090 keV. The ^{208}Tl decay also produces 583.2 keV photons with a branching ratio of 84.5%. To avoid coincidences orginating from these photons an energy window of $|E_{\mathrm{REGe}} - 583.2| < 3.0$ keV is excluded.

The single-segment events are selected from each data sample by additionally requiring

– single-segment requirement:
$|E_{\mathrm{Seg14}} - E_{\mathrm{Core}}| < 5.0$ keV or $|E_{\mathrm{Seg17}} - E_{\mathrm{Core}}| < 5.0$ keV .

The single-segment event samples are noted as $\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$, $\Gamma_{2.6}^{\mathrm{S}}$, $\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$ and $\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$, respectively.

The coincidence trigger is only relevant for the Γ_{SCS} sample. However, the other samples are selected out of

the collected coincident events to ensure the same experimental conditions. In principle the REGe detector could also be used to tag 511 keV photons for events in the Γ_{SEP} and Γ_{DEP} samples. However, the statistics available is not sufficient.

The distribution of the energy of the core, E_{Core}, of all coincident events is shown in Fig. 2a. The E_{Core} distribution of all single-segment coincident events is shown in the same plot. Also shown are the simulated spectra which will be discussed in the next section. Figure 2b shows the E_{Core} distribution for all coincident events with $E_{\mathrm{Core}} + E_{\mathrm{REGe}} = (2614.5 \pm 5.0)$keV. The arrows indicate the E_{Core} range corresponding to the acceptance angles for the Γ_{SCS} sample.

The DEP, SEP and 2.6 MeV peaks are all prominant in Fig. 2a. Only the SEP peak is also prominant in Fig. 2b. The 511 keV annihilation photon that escapes the segmented detector is fully absorbed by the REGe in these events. The DEP peak disappears because the two 511 keV
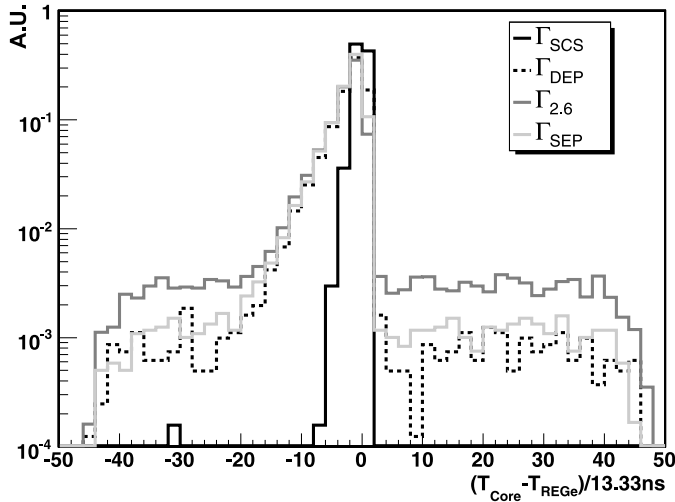
Fig. 3. $T_{\mathrm{Core}} - T_{\mathrm{REGe}}$ distributions. The unit of the x axis is the sampling clock

photons are emitted back to back and only one of the two photons can be tagged by the REGe detector. The numbers of events in all samples are given in the first row of Table 1.

The time between the arrival of the core trigger (T_{Core}) and the REGe trigger (T_{REGe}), $\Delta T = T_{\mathrm{Core}} - T_{\mathrm{REGe}}$, is shown in Fig. 3. The ΔT distribution of the Γ_{SCS} events has a mean value of -9.4 ns with a RMS of 12.4 ns. Only one event falls outside the peak ($|\Delta T|$ more than $8 \times 13.3 = 107$ ns). This confirms that events in the Γ_{SCS} sample are predominantly induced by 2614 keV photons from the ^{208}Tl decay and the fraction of accidental coincidences is negligible at the 10^{-4} level.

The ΔT distributions of the $\Gamma_{2.6}$, Γ_{DEP} and Γ_{SEP} samples are also shown in Fig. 3. These ΔT distributions are composed of "signal" peaks at $\Delta T \approx 0$ and flat distributions of accidental coincidences. The "signal" events in the Γ_{DEP} and Γ_{SEP} samples register the 2.6 MeV photon in the segmented detector through pair production with one annihilation photon reaching the REGe detector. The "signal" events in the $\Gamma_{2.6}$ sample have another photon from the same ^{208}Tl decay registered in the REGe. The numbers of accidental coincidence events can be calculated by fitting the ΔT distributions with 133 ns $< \Delta T < 266$ ns with a constant function. The fractions of "signal" events after subtracting the accidental coincidence events are indicated by f_{c} and given in Table 1. The fractions of accidental co-

incidence events ($1 - f_c$) agree with the rough estimate of $\approx 20\%$ from the trigger rates, as explained in Sect. 2.1. Notice, that most accidental coincidence events in the Γ_{DEP}, Γ_{SEP} and $\Gamma_{2.6}$ samples can be treated as events triggered with only the core of the segmented detector and they are actually classified correctly. This was concluded in [1] where a detailed study of core only triggered events was presented.

The $\Gamma_{2.6}$, Γ_{DEP} and Γ_{SEP} samples have wider ΔT distributions than the Γ_{SCS} sample. This is an artefact of the fixed 100 keV energy threshold applied to the REGe detector. As the overall rise-time of a pulse, see Fig. 5a, does not depend on the energy, the time at which a fixed threshold is reached does. The $\Gamma_{2.6}$, Γ_{DEP} and Γ_{SEP} samples are selected without any cut on E_{REGe}. This results in much wider spreads in E_{REGe} and thus in wider ΔT distributions.

2.3 MC simulation

The GEANT4 based Monte Carlo package MaGe [12] is used to simulate the setup. In order to speed up the computation only the ^{208}Tl decay is simulated and not the complete decay chain of the ^{228}Th source. The energies as deposited in the germanium detectors are smeared event by event according to the detector resolutions. The same energy thresholds and the coincidence trigger as for the measured data are applied to the simulated events. The MC is normalized to the data by counting the number of events within the energy region of $E_{\mathrm{Core}} + E_{\mathrm{REGe}} = 2614 \pm 5$ keV, since events satisfying this requirement are almost exclusively induced by the 2614 keV photon from the ^{208}Tl decay (see previous section).

The simulated distributions of E_{Core} are shown in Fig. 2a and b. The same selection cuts as required for the 8 data samples are applied to the MC events. The data to MC ratios are given in Table 1. They agree with the fractions of events with true coincident triggers (f_{c}) within $\approx 10\%$. The overall excess of data of $\approx 20\%$ for all but the SCS samples agrees well with the accidental coincidence rate.

2.4 Distinction between MSE and SSE in MC

The variable R_{90} is defined as the radius of the volume that contains 90% of the total energy deposition in a ger-

Table 1. The numbers of events in all data samples are presented in the first row. For the Γ_{SCS} sample, f_c in the second row corresponds to the fraction of events with $|\Delta T| < 107$ ns. The error on f_{c} comes from statistics only. For the $\Gamma_{2.6}$, Γ_{DEP} and Γ_{SEP}, it corresponds to the fraction of events in the central peaks of the $|\Delta T|$ distributions after subtracting the accidental coincidence contribution. The ratios of event numbers for data and MC are given in the third row with statistical errors only

sample	Γ_{SCS}	$\Gamma_{2.6}$	Γ_{DEP}	Γ_{SEP}	$\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$	$\Gamma_{2.6}^{\mathrm{S}}$	$\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$	$\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$
#events	6716	25 780	6898	10 093	642	1131	1059	411
f_{c} [%]	>99	$78_{\pm 1}$	$87_{\pm 1}$	$85_{\pm 1}$	$97_{\pm 4}$	$78_{\pm 2}$	$87_{\pm 3}$	$82_{\pm 4}$
# MC/data [%]	$103_{\pm 1}$	$66_{\pm 1}$	$80_{\pm 1}$	$79_{\pm 1}$	$88_{\pm 3}$	$70_{\pm 2}$	$78_{\pm 2}$	$73_{\pm 4}$
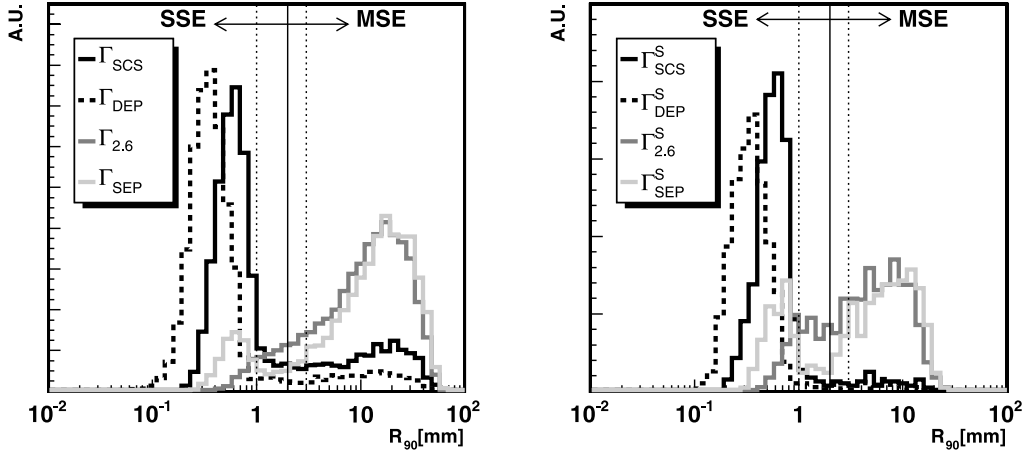
Fig. 4. R_{90} distributions of the 4 selected samples without the single-segment requirement (*left*) and with the single-segment requirement (*right*). The 3 *vertical lines* on each plot indicate the R_{90} values of 1, 2 and 3 mm

manium detector. It is used to study the size of the volume within which the energy is distributed. Details are described in [8]. The distributions of R_{90} as calculated using MC information are shown in Fig. 4 for the 8 selected samples. Events from Γ_{DEP} and Γ_{SCS} samples mostly have much smaller R_{90} than those from Γ_{SEP} and $\Gamma_{2.6}$ samples. Γ_{SCS} events have slightly larger R_{90} than Γ_{DEP} events due to the higher energy of the recoiling electron.

A fraction of the SCS events have relatively large R_{90} (> 2 mm). In most of these events the 2.6 MeV photon Compton scatters several times inside the segmented detector before reaching the REGe detector. They still survive the Γ_{SCS} cuts due to the relatively large geometrical acceptance of the REGe detector. Events with $R_{90} > 2$ mm in the Γ_{DEP} sample originate from photons not interacting with the detector through pair production, but through multiple Compton scattering, and still depositing the same amount of energy as in DEP events. These events are significantly reduced by applying a single-segment cut, as shown in Fig. 4. The fraction of events from the $\Gamma_{2.6}$ and Γ_{SEP} samples with $R_{90} < 2$ mm have the high energy photon depositing energy very locally. These fractions of events increase after applying a single-segment cut.

The "position resolution" of the DAQ is ≈ 1 mm, as explained in Sect. 2.1. However, a conservative cut of $R_{90} < 2$ mm is used to distinguish SSEs from MSEs [1]. The fractions of SSEs (f_{SSE}) in each sample are listed in Table 2. The errors on f_{SSE} are estimated by varying the R_{90} cut value between 1 and 3 mm. Γ_{SCS} has a smaller fraction of SSEs than Γ_{DEP}, due to the relatively large selection window. The f_{SSE} fractions for the Γ^S samples are larger than for the Γ samples, since the single-segment cut already removes most MSE events.

If only the segmented detector is used for triggering, $f_{SSE} \approx 78\%$ for the Γ_{DEP} sample, and $\approx 12\%$ for the $\Gamma_{2.6}$ sample [1] (89% and 30% for Γ_{DEP}^S and $\Gamma_{2.6}^S$ samples, respectively). These values are similar to the ones for coincident events. Therefore, even though accidental coincidences are not simulated by the MC, the f_{SSE} values as presented in Table 2 can be used to evaluate the data samples collected with the coincidence trigger.

If the estimated 1 mm position resolution can be achieved through PSA, the SSEs from each sample should

Table 2. Fractions f_{SSE} of events with $R_{90} < 2$ mm in each sample

sample	Γ_{SCS}	$\Gamma_{2.6}$	Γ_{DEP}	Γ_{SEP}
f_{SSE}	72^{+3}_{-6} %	10^{+6}_{-7} %	88^{+1}_{-2} %	15^{+3}_{-3} %
sample	Γ_{SCS}^S	$\Gamma_{2.6}^S$	Γ_{DEP}^S	Γ_{SEP}^S
f_{SSE}	92^{+1}_{-3} %	26^{+12}_{-15} %	96^{+1}_{-1} %	31^{+6}_{-5} %

be correctly identified. The PSA procedure is described in the following section.

3 Pulse shape analysis

The same artificial neural network (ANN) package as used in [1] is used here to perform the pulse shape analysis. The ANN is trained with an SSE sample against an MSE sample. In [1] Γ_{DEP} (without coincidence trigger) was used as the SSE-dominant sample and events in the 1620 keV line (with the 1620 keV photon from ^{212}Bi decay fully absorbed in the segmented detector) as the MSE-dominant sample. The trained ANN was able to identify both SSE and MSE events with $\approx 85\%$ efficiencies.

In this study, a similar analysis is performed. The ANN is trained with the Γ_{DEP} sample (SSE-dominant) against the Γ_{SEP} sample (MSE-dominant). The trained ANN is used to verify that the collected events in the Γ_{SCS} sample are SSE-dominant. The results are shown in Sect. 3.2 after a general description in Sect. 3.1.

In a second analysis the ANN is trained with the Γ_{SCS} against the $\Gamma_{2.6}$ sample. It is shown in Sect. 3.3 that the results are consistent.

3.1 General features of the ANN

The core pulse of the segmented detector of a typical Γ_{DEP} event is shown in Fig. 5a. The rising part of the pulse contains information about the event structure as explained in
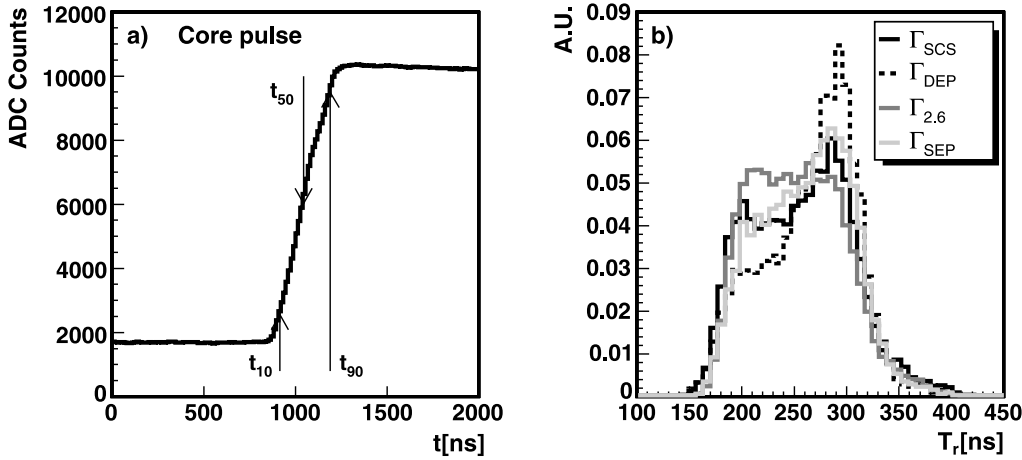
Fig. 5. a Core pulse of one DEP event, t_{10}, t_{50} and t_{90} are indicated by *arrows*. **b** distributions of the rise time $T_r = t_{90} - t_{10}$ for the 4 samples under consideration

in Sect. 1. The time t_{50} is defined as the time at which the pulse has reached 50% of its maximum.[3] The 20 values before and the 20 after t_{50} are used for PSA. Thus, the selection of the 40 values is independent of the absolute amplitude of the pulse and thus independent of the energy.

t_{10} and t_{90} are defined as the times when the pulse reaches 10% and 90% of its maximum, respectively. The distributions of the pulse rise time, $T_r = t_{90} - t_{10}$, are shown in Fig. 5b. T_r is fully covered by the 40 values which cover a time window of 533 ns. The dominance of long rise-times in the Γ_{DEP} sample reflects the dominance of events close to the detector surface.

The ANN package as used here has 40 input neurons for the 40 pulse values. It has two hidden layers with 8 and 2 neurons each and 1 output neuron. The ANN is trained such that a large ANN output (NN_{out}) indicates that the event is SSE-like and a small NN_{out} indicates that it is MSE-like.

Since both NN_{out} and R_{90} are related to the size of the energy deposition in the detector, a correlation between R_{90} and NN_{out} is predicted. On average events with small R_{90} should have large NN_{out} and vice versa. It is clear that R_{90} is not the only variable that determines the pulse shape. Other, second order effects like the drift anisotropies caused by the crystal structure and inhomogenious doping concentrations also modify the pulse shapes. Therefore, a 100% correlation between NN_{out} and R_{90} is not expected. The details of this correlation can only be studied with a detailed pulse shape simulation which is beyond the scope of this paper.

3.2 Verification of ANN training with single Compton scattering events

The ANN is trained with the Γ_{DEP} sample as SSE-dominant (signal-like) and the Γ_{SEP} sample as MSE-dominant (background-like). The training takes 300 itera-

tions.[4] The trained ANN is then applied to all Γ_{SCS} and $\Gamma_{2.6}$ events. It should correctly identify them as single–site and the multi–site events. The NN_{out} distributions for all 4 samples are shown in Fig. 6a. The NN_{out} distributions from the ANN trained with the $\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$ and $\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$ samples are shown in Fig. 6b.

The Γ_{SCS} events have in average larger NN_{out} values than the $\Gamma_{2.6}$ events. The peaks of the distributions are well separated. However, while the distribution for $\Gamma_{2.6}$ events is quite similar to the one for Γ_{SEP} events, the distribution for the Γ_{SCS} events looks different from the one for Γ_{DEP} events. A shift of the peak is expected from the MC simulation, since there is a higher percentage of Γ_{SCS} events with R_{90} values above 2 mm indicating an MSE-like structure of the events, see Fig. 4. The Γ_{DEP} distribution in Fig. 6 a in addition features a plateau towards high NN_{out} values. This is probably an artefact of the spatial distribution of the events which are predominantly close to the surface which also influenced the ANN training. This feature does not show up in Fig. 6b, most probably due to the much lower statistics in the $\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$ sample.

The classification of events using the distributions depicted in Fig. 6a (and b) is based on a cut in NN_{out}, $NN_{\mathrm{out}}^{\mathrm{CUT}}$. An event is classified as SSE-like, if $NN_{\mathrm{out}} > NN_{\mathrm{out}}^{\mathrm{CUT}}$, or MSE-like, if $NN_{\mathrm{out}} < NN_{\mathrm{out}}^{\mathrm{CUT}}$. For a given value of $NN_{\mathrm{out}}^{\mathrm{CUT}}$, the survival efficiency for any data sample, ϵ^{ANN}, is defined as the fraction of events in that sample that are identified by the ANN as SSE-like events.

The probabilities to correctly identify SSE- and MSE-like events, $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$, are calculated using the Monte Carlo predictions for the purities f_{SSE} of the samples used, see Table 2, and using the measured ϵ^{ANN} for the data samples. A linear dependence $\epsilon^{\mathrm{ANN}} = a \times f_{\mathrm{SSE}} + b$ is assumed. For a given $NN_{\mathrm{out}}^{\mathrm{CUT}}$, the values for ϵ^{ANN} are calculated for all samples, a linear fit is performed to obtain the slope and the line is extrapolated to $f_{\mathrm{SSE}}=1$ to obtain $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$. It is extrapolated to $f_{\mathrm{SSE}} = 0$ to determine $1 - \eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ (see Fig. 6d for two fits). The fit procedure takes errors into account. The errors on f_{SSE} are listed in Table 2 and those on ϵ^{ANN} are statistical only. The resulting $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$

---

[3] Pedestals are subtracted by using the information during the 1 μs interval before the trigger.

[4] The ANN trained with 500 iterations gives similar results.
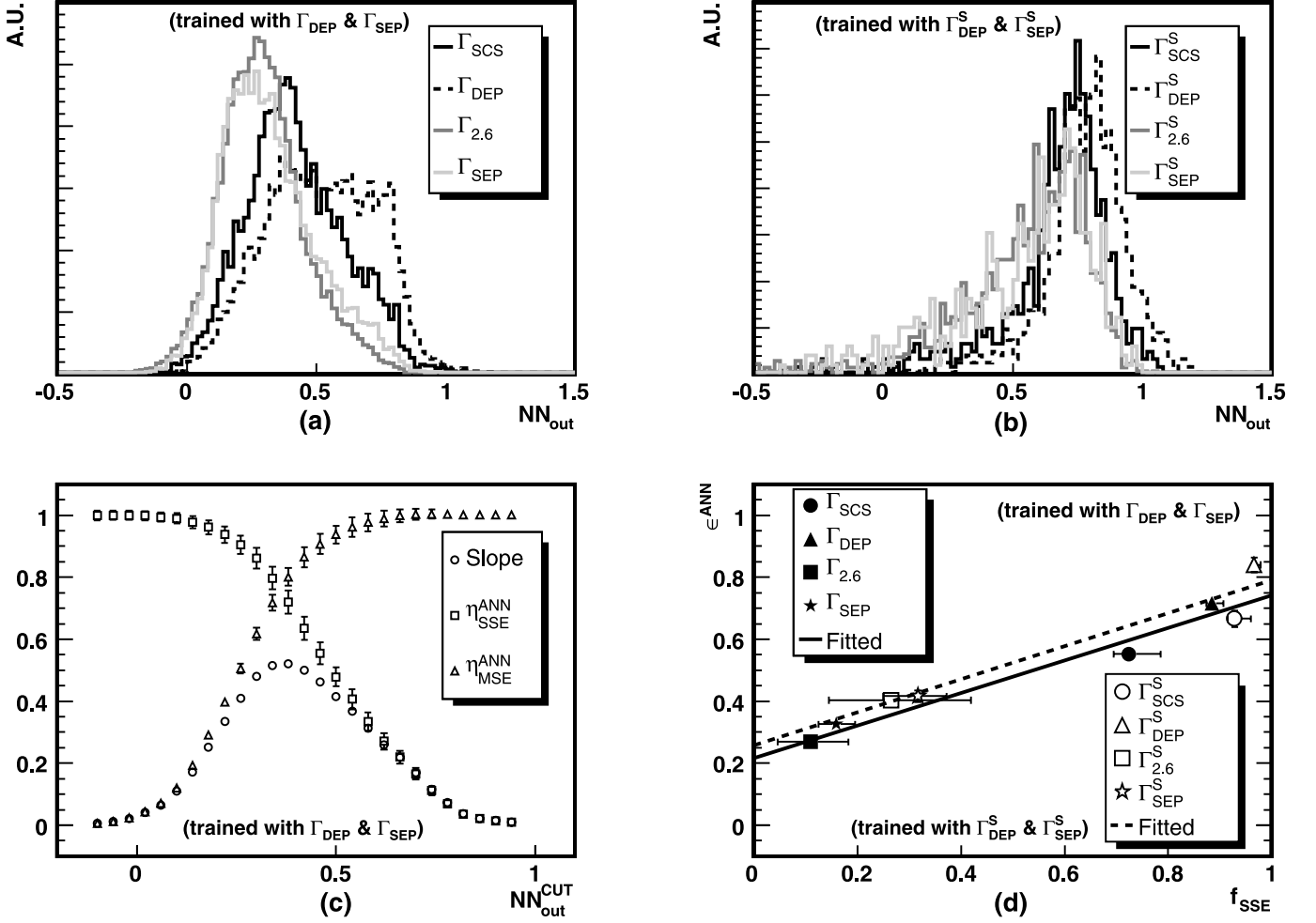
Fig. 6. Results of the ANN analysis, if the ANN is trained with the Γ_{DEP} ($\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$) sample as SSE-dominant and the Γ_{SEP} ($\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$) sample as MSE-dominant: **a** NN_{out} distributions for all four samples. (ANN trained with the Γ_{DEP} and Γ_{SEP} samples.) **b** NN_{out} distributions for all four samples. (ANN trained with the $\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$ and $\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$ samples.) **c** $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ vs. $NN_{\mathrm{out}}^{\mathrm{CUT}}$. The fitted slope a, see text, is shown as well. Errors are taken from the MINUIT fit. (ANN trained with the Γ_{DEP} and Γ_{SEP} samples. Results with the $\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$ and $\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$ samples are not shown here.) **d** ϵ^{ANN} vs. f_{SSE}; ϵ^{ANN} values correspond to the value of $NN_{\mathrm{out}}^{\mathrm{CUT}}$ giving the maximum fitted slope a. Also given are results for the single segment samples indicated by $^{\mathrm{S}}$ (*open points*). See text for details

and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ as a function of $NN_{\mathrm{out}}^{\mathrm{CUT}}$ are shown in Fig. 6c. The fitted slope a is shown in Fig. 6c as a function of $NN_{\mathrm{out}}^{\mathrm{CUT}}$ as well. A clear maximum for a is visible.

The correlations between the values of ϵ^{ANN} and f_{SSE} are shown in Fig. 6d for the value of $NN_{\mathrm{out}}^{\mathrm{CUT}}$ which maximizes the slope a. For the case of ANN trained with Γ_{DEP} and Γ_{SEP} samples, the maximum value of $a = 0.524 \pm 0.015$ is achieved at $NN_{\mathrm{out}}^{\mathrm{CUT}} = 0.37$. The slope a does not approach the ideal value of 1, indicating that f_{SSE} and ϵ^{ANN} are not fully correlated. This is expected as the predictions for f_{SSE} are entirely based on the simple variable R_{90} as discussed in Sect. 3.1. The results for the single segment samples are also shown. They were subjected to the identical analysis using the equivalent samples for training. The results of the fits are indicated for both single segment and unrestricted event samples.

The results for $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ are given in the first two rows of Table 3 with errors deduced from the fits. The ANN can correctly identify both SSE and MSE events at the 75% to 80% level. The results for the single segment data sets

Table 3. $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ with the ANN trained with various SSE-dominant samples against various MSE-dominant samples

ANN Training		Analysis	
SSE-dominant	MSE-dominant	$\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$	$\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$
Γ_{DEP}	Γ_{SEP}	$74.1_{\pm 2.7}$ %	$78.3_{\pm 2.8}$ %
$\Gamma_{\mathrm{DEP}}^{\mathrm{S}}$	$\Gamma_{\mathrm{SEP}}^{\mathrm{S}}$	$79.1_{\pm 7.2}$ %	$74.3_{\pm 6.8}$ %
Γ_{SCS}	$\Gamma_{2.6}$	$69.0_{\pm 2.1}$ %	$81.5_{\pm 2.5}$ %
$\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$	$\Gamma_{2.6}^{\mathrm{S}}$	$70.2_{\pm 4.3}$ %	$84.2_{\pm 5.1}$ %

are similar to ones for the unrestricted samples. These results agree in general with the values of $\approx 85\%$ as achieved in [1].

The compatibility of the points with the linear fits in Fig. 6d leads to the conclusion that the SSE-like events in the Γ_{SCS} sample are identified with about the same efficiency as in the other samples. This is the most important result of this study indicating that tagged SCS events can indeed be used to further study pulse-shapes in more detail.

As a cross check, the ANN was also trained with the Γ_{DEP} sample as SSE-dominant and the $\Gamma_{2.6}$ sample as MSE-dominant. The same procedure with the trained ANN was then applied. The results are $\eta_{\mathrm{SSE}}^{\mathrm{ANN}} = 80.5 \pm 2.9\%$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}} = 77.9 \pm 2.8\%$. These results are very similar to the values shown in Table 3, indicating that the training procedure is reliable and achieves consistent results.

3.3 Cross-check using SCS events for ANN training

The same procedure as described in the previous section is repeated with the ANN trained using the Γ_{SCS} ($\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$) as the SSE-dominant and the $\Gamma_{2.6}$ ($\Gamma_{2.6}^{\mathrm{S}}$) as the MSE-dominant samples. The results are shown in Fig. 7. The plateau feature in the NN_{out} distribution from the Γ_{DEP} sample as seen in Fig. 6 a is much reduced, as shown in Fig. 7a. This is most probably due to the fact that events in the Γ_{SCS} sample have a relatively more uniform spatial distribution than the Γ_{DEP} events, resulting in a trained ANN which is insensitive to the spatial information.

The resulting identification probabilities $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ are given in the last two rows of Table 3. The ANN can correctly identify SSE-like events at the 70% and MSE-like events at the 80% level. This confirms again that the selected SCS samples are enriched in SSE-like events and can be used to train the ANN package.
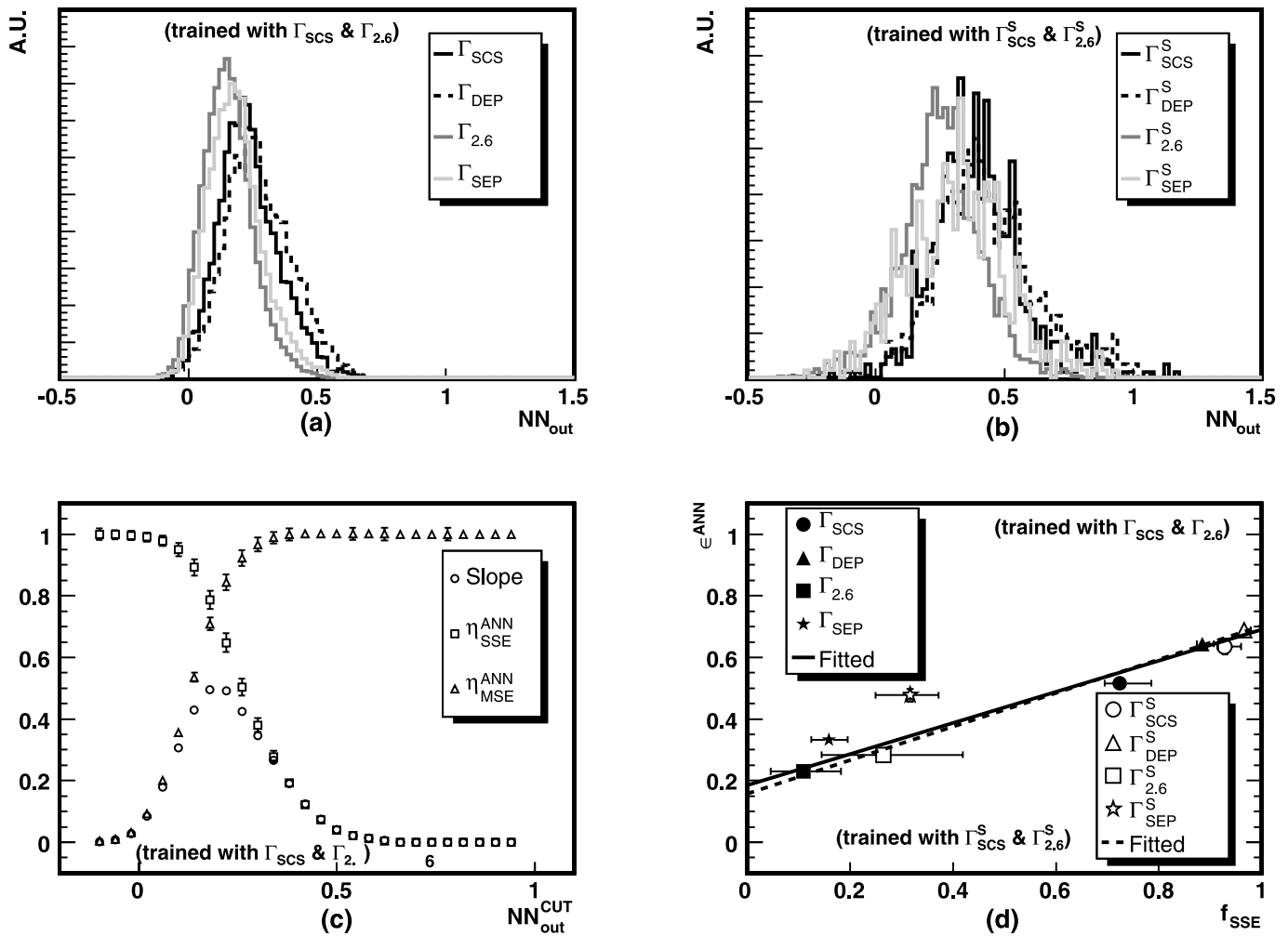
Fig. 7. Results of the ANN analysis, if the ANN is trained with the Γ_{SCS} ($\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$) sample as SSE-dominant and the $\Gamma_{2.6}$ ($\Gamma_{2.6}^{\mathrm{S}}$) sample as MSE-dominant: **a** NN_{out} distributions for all four samples. (ANN trained with the Γ_{SCS} and $\Gamma_{2.6}$ samples.) **b** NN_{out} distributions for all four samples. (ANN trained with the $\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$ and $\Gamma_{2.6}^{\mathrm{S}}$ samples.) **c** $\eta_{\mathrm{SSE}}^{\mathrm{ANN}}$ and $\eta_{\mathrm{MSE}}^{\mathrm{ANN}}$ vs. $NN_{\mathrm{out}}^{\mathrm{CUT}}$. The fitted slope a, see text, is shown as well. Errors are taken from the MINUIT fit. (ANN trained with the Γ_{SCS} and $\Gamma_{2.6}$ samples. Results with the $\Gamma_{\mathrm{SCS}}^{\mathrm{S}}$ and $\Gamma_{2.6}^{\mathrm{S}}$ samples are not shown here.) **d** ϵ^{ANN} vs. f_{SSE}; ϵ^{ANN} values correspond to the value of $NN_{\mathrm{out}}^{\mathrm{CUT}}$ giving the maximum fitted slope a. Also given are results for the single segment samples indicated by $^{\mathrm{S}}$ (*open points*)

4 Conclusions and outlook

Events with photons Compton scattering only once inside a germanium detector, SCS events, can be selected by tagging the scattered photon with a second germanium detector. The pulse shapes of these events can be studied and used to test methods that distinguish between single-site and multi-site events.

In order to collect SCS events and perform pulse shape analysis, an 18-fold segmented prototype detector for the Phase-II of the GERDA experiment was positioned in front of a ^{228}Th source. A second germanium detector was positioned to record the escaped photons at $72°$, corresponding to 2040 keV energy deposit in the segmented detector, close to the Q-value of the $0\nu\beta\beta$ decay of ^{76}Ge.

According to the MC simulation $\approx 72\%$ of the collected SCS events are true SSE events. The SSE-dominance is verified by an artifical neural network (ANN) trained in an independent way. These SCS events are then themselves used to train the pulse shape analysis package and thus the trained PSA is able to identify single- and multi-site events with efficiencies at the $\approx 80\%$ level.

The studies can be improved in many aspects. The fraction of SSE events in the collected SCS sample can be increased by further improving the tagging method. For example, the whole experimental setup can be shielded from external photons and collimators can be positioned between the two detectors. In order to study the effect of the identification efficiencies on the sensitivity of double beta decay experiments (GERDA) the background events and their corresponding SSE and MSE fractions for the experiment under investigation need to be known in detail. This requires a detailed MC simulation of the (GERDA) experiment. A pulse shape simulation can improve the understanding of the correlation between NN_{out} and R_{90}. The use of individual pulses from neighbouring segments is expected to improve the identification efficiency. This was not considered in these studies. Detailed Monte Carlo simulations of the GERDA experiment and pulse shape simulations for the detectors are under way.

References

1. I. Abt et al., Eur. Phys. J. C **52**, 19 (2007)
2. F. Petry et al., Nucl. Instrum. Methods A **332**, 107 (1993)
3. J. Hellmig, H.V. Klapdor-Kleingrothaus, Nucl. Instrum. Methods A **455**, 638 (2000)
4. B. Majorovits, H.V. Klapdor-Kleingrothaus, Eur. Phys. J. A **6**, 463 (1999)
5. D. Gonzãlez et al., Nucl. Instrum. Methods A **515**, 634 (2003)
6. S.R. Elliott, V.M. Gehman, K. Kazkaz, D.-M. Mei, A.R. Yong, Nucl. Instrum. Methods A **558**, 504 (2006)
7. GERDA Collaboration, S. Schönert et al., Nucl. Phys. Proc. Suppl. **145**, 242 (2005)
8. I. Abt et al., Nucl. Instrum. Methods A **570**, 479 (2007)
9. User's Manual Digital Gamma Finder (DGF) PIXIE-4, XIA LLC, http://www.xia.com
10. I. Abt et al., Nucl. Instrum. Methods A **577**, 574 (2007)
11. Canberra Reverse-Electrode Coaxial Ge Detector, http://www.canberra.com/Products/494.asp
12. M. Bauer et al., J. Phys.: Conf. Ser. **39**, 362 (2006)