Calibration of the LEGEND-200 Experiment to Search for Neutrinoless Double Beta Decay and Searches for Signatures of New Physics with the GERDA Experiment

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Yannick Müller

aus

der Bundesrepublik Deutschland

Promotionskommission

Prof. Dr. Laura Baudis (Vorsitz)

Prof. Dr. Gino Isidori

Dr. Michelle Galloway

Dr. Pin-Jung Chiu

Zürich, 2023

Abstract

Despite the tremendous accuracy of the description of the Universe from tiniest to greatest distances by the combination of the Standard Model of particle physics (SM) and the cosmological standard model (ΛCDM), several puzzles of Nature remain unsolved. Among others, these include the origin of the matter-antimatter asymmetry in the early Universe, preventing a pure existence of radiation, the nature of neutrinos, measured to undergo beyond SM (BSM) oscillations between flavour eigenstates, and the characteristics of the dark matter (DM), comprising more than 80% of the matter content in the Universe.
The GERDA and the LEGEND experiments deploy high-purity germanium (HPGe) detectors enriched in the target isotope ^{76}Ge by up to 90%, in an active liquid argon shield to detect neutrinoless double beta decay ($0\nu\beta\beta$), a process generating two matter particles without the emission of any antimatter. An observation would conclusively demonstrate the Majorana nature of neutrinos, i.e. the equivalence of neutrino and antineutrino, and the realisation of the violation of lepton number conservation in Nature. In the SM double beta decay ($2\nu\beta\beta$), two electrons as well as two antineutrinos are emitted, thus conserving lepton number. But, if the process can also occur without the emission of neutrinos, the induced violation of lepton number conservation may hint towards an explanation of the matter-antimatter asymmetry in the early Universe via leptogenesis.
The main experimental challenges for a detection are the expected ultra-long half-life, already constrained to exceed 10^{26} yr in two isotopes, i.e. around sixteen orders of magnitude longer than the age of the Universe, and the background from natural radioactivity, which can mimic the sought-after signal. Hence, a potential observation requires large source quantities with efficient detection capabilities in an ultra-low background environment. Since no neutrinos carrying energy are emitted, the signature in the experiment is a single peak in the summed electron energy spectrum at a specific energy, the Q value of the decay. Thus, an unavoidable background source is the continuous $2\nu\beta\beta$ spectrum extending over all energies up to Q, necessitating the best possible energy resolution to separate the signal peak from the background continuum.
LEGEND follows the experimental operations of its predecessor GERDA both in the overall experimental design and the calibration strategy. Radioactive ^{228}Th sources, featuring a low neutron emission rate, are regularly, on an approximately weekly schedule, deployed in the vicinity of the HPGe detectors to calibrate their energy response, to measure their energy resolution, and to calibrate the event pulse shape analysis techniques for offline background suppression based on event waveforms.
In this thesis, we present the calibration hardware built and operated to move the radioactive sources, and investigate the uniformity of the illumination of the operated HPGe detectors by the γ-rays emitted by the sources. We further discuss the characterisation measurement of the neutron rate emitted by the sources, a key parameter as neutron emission during calibrations could cause material activation with the production of, among others, ^{77}Ge. Subsequent, delayed decays may mimic potential signals in the region-of-interest for the $0\nu\beta\beta$ analysis during physics

operations. We measured the global neutron emission rate of the seventeen radioactive sources, which yielded $\Psi = (4.30 \pm 0.69_{\text{stat}} \pm 0.93_{\text{syst}}) \times 10^{-4}$ n / (kBq s). Running the corresponding simulations of the neutrons in the experimental setup as well as the subsequent activated isotope transitions, we found a background contribution after cuts of around three orders of magnitude below the background goal of the current stage of the LEGEND experiment, LEGEND-200, of 2×10^{-4} counts / (keV kg yr), allowing for the installation of the sources in the experiment.
Besides the goal of observing $0\nu\beta\beta$, the GERDA and the LEGEND experiments, being specifically designed to perform rare-event searches with a high signal detection efficiency and a high-precision energy resolution in a low background environment, are well suited to also search for other BSM physics phenomena. We analysed the keV range of the available GERDA physics data to search for signatures of bosonic keV-scale dark matter (bDM) and for decay modes of individual electrons, neutrons, and protons in the HPGe detectors, manifesting the violation of charge conservation, or of baryon number conservation. Based on the development of an empirical background model of the GERDA data at low energies motivated by the underlying physical background components, we performed generic searches for monoenergetic peak-like excesses, which, depending on the energy and the event topology, would be induced by these BSM interaction modes. Apart from expected γ-ray transitions, we did not find any statistically globally significant signal excess caused by any of the new physics channel searched for. Consequently, we evaluated constraints on the physical interaction strength of bDM and the particle lifetimes, respectively, in a frequentist statistical framework. We obtained confidence intervals which are among the most stringent results determined with semiconductor detectors, but given the limited data exposure we do not reach the sensitivity of tonne-scale experiments, or indirect constraints in the case of bDM. However, by scaling our empirical background model and the *measured* results from GERDA, we comment on the future sensitivity of the second stage of LEGEND. With more than a tonne of enriched HPGe detector material directly immersed in underground-sourced liquid argon, this LEGEND-1000 experiment will fully cover the inverted neutrino mass ordering regime in the main $0\nu\beta\beta$ search. It addition, it will approach the regime of competitive single particle decay lifetime constraints and further constrain the parameter space for phenomenologically viable bDM models.

The dissertation itself is organised as follows. I describe the theoretical background and the motivation for the search for $0\nu\beta\beta$, followed by the experimental approaches for a detection, in Chapter 1. I then provide a detailed description of the HPGe detector operation and its realisation within the GERDA and the LEGEND experiments, in Chapter 2. Next, I describe the detector calibration strategy, detailing the calibration hardware, and the properties of the radioactive calibration sources, in particular in terms of their neutron emission rate, in Chapter 3. At that point, I will change the focus from the main $0\nu\beta\beta$ search, and present the search for bDM and for electron and nucleon decays in the low energy regime of GERDA in Chapter 4, which I finish with the sensitivity projection for LEGEND-1000. I summarise and conclude on the outcomes of the topics presented in this thesis in Chapter 5. I add supplementary material on the LEGEND-200 calibration source deployment hardware and the related software control in Appendix A. In Appendix B, I provide additional material on the GERDA analyses. I motivate alternative search strategies for the nucleon decay, present an alternative analysis to probe γ-rays mimicking a monoenergetic peak, and discuss an additional application of the presented analysis methods for background investigations.

Author's declaration

The work presented in this thesis has been performed within the GERDA and the LEGEND collaborations, international projects with more than 250 researchers in total. It should thus be obvious that not every detail described here is my work or exclusively my work. Many covered topics have been addressed in small groups, with inputs and supports from many people. Let me thus briefly list my individual contributions to the scope of the projects.

- LEGEND-200: I was the main responsible person and contributor to the testing, assembly, commissioning, and monitoring of the calibration hardware (Sec. 3.2). Given these tasks, I benefitted from strong supports from many people. I further performed the remote monitoring of the neutron flux measurement and the corresponding offline analysis, as well as the related background simulations run to estimate the impact of the neutrons emitted during calibrations (Secs. 3.4, 3.5). I contributed to a related publication describing the calibration source production and characterisation as one of the leading authors [1]. I also investigated the event uniformity and the detector performance study for calibration data taken during the commissioning stage.

- GERDA: I was one of the two leading responsible persons of the sub-analysis team for the searches for new physics in the low energy regime of the GERDA data. I motivated to combine the different search channels detailed in Chapter 4.1 and to reinterpret former studies performed in GERDA based on other physics models. For example, I showed how to interpret the single neutron decay study (Sec. 4.1.2), also as a proton decay study, and implemented the dark Compton scattering channel for bosonic dark matter searches (Sec. 4.1.1). Regarding the analysis methods, I developed an empirical low energy background model for GERDA detailed in Sec. 4.2, a new feature, which had not been achieved before. I also performed the new physics analyses in the frequentist statistical framework, in contrast to the Bayesian framework used in earlier and parallelly performed analyses in GERDA (Sec. 4.4). I did not run the dedicated efficiency simulations for the analyses as introduced in Sec. 4.3, but contributed when incorporating the efficiencies into the final results. Furthermore, I investigated the projected sensitivity reach of the LEGEND-1000 experiment based on the obtained results for GERDA (Sec. 4.6). A publication of the analyses for GERDA, for which I was one of the leading contributors of the writing team, is currently under collaboration-internal review. I further performed the related application of the analysis methods for other searches (Sec. B.3), or the investigation of related methods which are not part of the paper in preparation (Secs. B.1, B.2).

Regarding the writing of the thesis itself, let me state that I wrote the dissertation entirely myself. Wherever I took direct information from others, I cited the corresponding publications, and when following concepts or formulations from the work of others or from published work of myself together with co-authors, I explicitly stated it. I *did* follow suggestions by colleagues

and direct or indirect supervisors, but implemented these solely myself. Let me conclude this preface declaring that I tried to respect and follow the scientific integrity in the entire writing process to my best knowledge.

Contents

Chapter 1

Double beta decay as a probe of lepton number conservation

1.1 The Standard Model of particle physics

For more than half a century now, the Standard Model of particle physics (SM) has proven to provide a tremendously precise description of what the Universe is built upon, and how its building blocks interact with each other. The SM is based on quantum field theory, and its essential building blocks are elementary particles, which represent field excitations. At a fundamental level two classes of particles are distinguished, the actual matter (and antimatter) content, the fermions, with half-integer spin, and the so-called gauge bosons mediating the interactions, or forces between those, having an integer spin. The fermions are again split into quarks and leptons. The former combine to form composite particles such as baryons, which are made of three quarks[1]. For example, the nucleons in the atomic nuclei, i.e. neutrons and protons, belong to these baryons. In contrast, the electrons in the atomic shell belong to the other class of fermions, the leptons, as which also neutrinos are characterised. These particles form the standard matter, referred to as the first particle generation, but all of these have two unstable counterparts with higher masses, referred to as second and third generation, respectively. All fermions also have a counterpart with identical properties except from the electric charge. These antimatter particles form the antileptons and antibaryons. To provide some examples, the siblings of the electron are referred to as muon, and tau, with masses approximately factors of 200 and 2000 higher, but otherwise assumed to be identical. The corresponding antimatter particle of the electron is called positron, which only differs from the matter counterpart by a positive electric charge instead of a negative one, but has the same mass. The bosons consist of gluons mediating the strong force between quarks, the W- and Z-bosons mediating the weak interaction, and the photon, mediating the electromagnetic force. Note that the first two are massive, whereas the latter is assumed to be *perfectly massless*. This is well-understood in the Higgs mechanism. A bosonic field, referred to as the Higgs field, yields the generation of the masses of the W- and Z-boson when the weak and electromagnetic forces, initially unified, decouple in the electroweak phase transition in the early Universe. The Higgs mechanism also generates the masses of the fermions, *but perhaps with the exception of the masses of neutrinos.* An additional boson, called the graviton, is expected to exist to mediate the gravitational force. However, this force is not yet described by the SM, and a complete theory will be needed in

[1]Note that at the composite level, not all matter (or antimatter) particles are actual fermions. Mesons, composite particles made of two quarks, have integer spin, and are hence bosons.

the future. The particle properties and the strengths of how actively the particles participate in those interactions which are described by the SM, are parametrised via charges, with the electric charge being the most famous example. In analogy to the electric charge, in the SM all charges are assumed to be *conserved* in any process. In particular collider-based experiments have probed these processes with enormous accuracy, and have thus established a good confidence in the accuracy of the parameters of the SM, meaning the masses, and the couplings representing the interaction strengths. We provide an illustration of the particle content of the SM in Fig. 1.1.

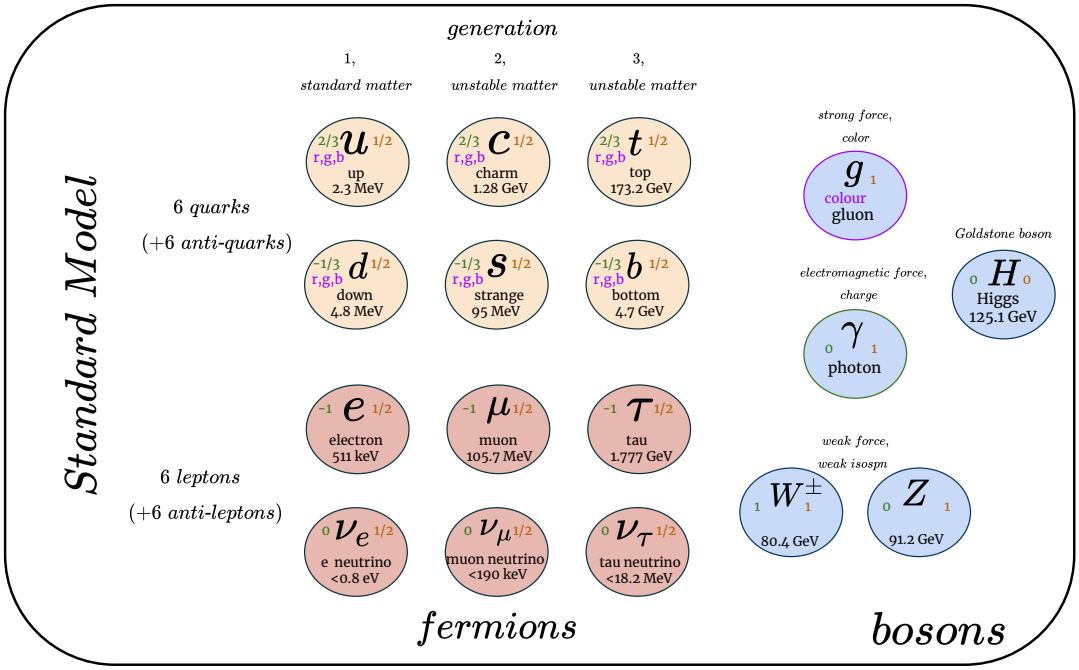

Figure 1.1: Sketch of the fundamental particle content of the Standard Model, including the forces described by it. The particles are separated by spin type, i.e. as fermionic matter, and bosonic force mediators. The electromagnetic charge is indicated in green, the spin in orange, and the colour in violet. The antimatter counterparts of the fermions are explicitly depicted, but differ by opposite charge sign only. The masses of the neutrinos had historically been assumed to be zero, but neutrino oscillations indicate non-zero, yet as of today unknown, masses. The schematic is based on the information provided in Ref. [2].

Above we have emphasised the masses of neutrinos, which were initially assumed to be massless in the SM, as we will motivate in Sec. 1.2. However, the measurements of neutrino oscillations [3, 4] prove that neutrinos change their flavour when propagating over large distances, between electron, muon, and tau neutrino. This behaviour can be explained as neutrinos having mass eigenstates which consist of a combination of the above flavour eigenstates. These oscillations violate the individual leptonic charges, meaning the individual lepton numbers for the electron, muon, and tau families are not conserved. Consequently, the existence of neutrino oscillations, being in contradiction to the SM assumption of charge conservation, demonstrates that the SM is incomplete. In these oscillation measurements, we can identify the important connection between the mass of the particles and the charges associated with them. A natural question to ask is whether the global leptonic charge conservation, referred to as lepton number, i.e. the sum of leptons and antileptons, is violated as well. As a matter of fact, we live in a world made of matter, and not of antimatter, and we may ask ourselves where this matter excess, in

the language of the SM, this violation of the conservation of baryon number, comes from. The main motivation of this work is in fact to probe with highest precision the beyond SM (BSM) violations of the conservation laws of the SM. We will hence consider, as the major investigation, a process that may allow us to understand how the masses of the neutrinos are generated. We will detail this process, referred to as neutrinoless double beta decay, in the reminder of this chapter. In addition, in Chapter 4 we will also probe baryon number and charge (meaning electric) conservation. A collection of some of the predicted BSM processes violating baryon and lepton number conservation is shown in Fig. 1.2, indicating the importance of experimental tests of the SM to provide guidance for the needed extensions of the model.

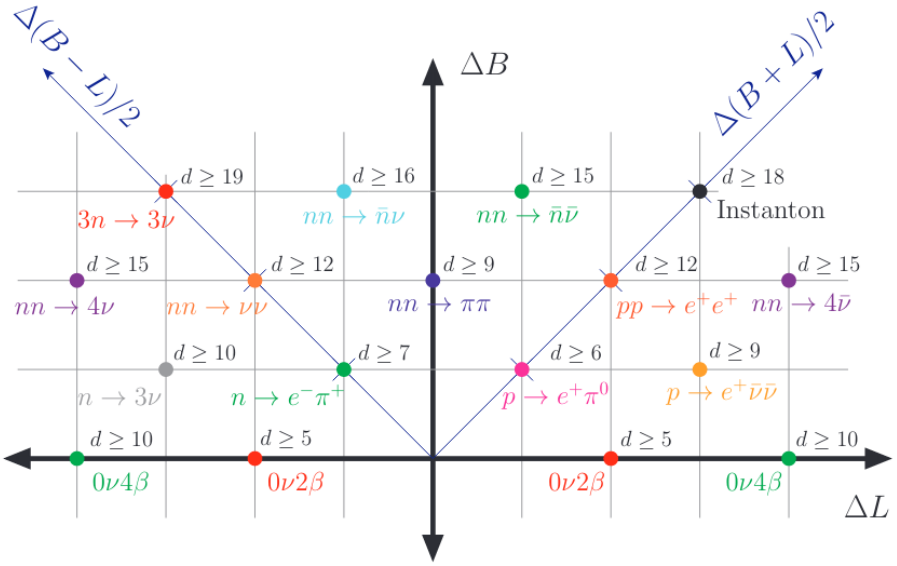

Figure 1.2: Collection of lepton and/or baryon number violating processes as predicted in several BSM extensions. The shown BSM particle decays are categorised by how many units they violate baryon (ΔB) and/or lepton number (ΔL). The diagonal lines indicate the violations of the difference ($\Delta(B-L)/2$) and the sum ($\Delta(B+L)/2$) of B and L, respectively. Different particles are abbreviated as follows. Protons are denoted with p, (anti)neutrons with n ($\bar{n}$), electrons (positrons) with β or e^- (e^+), (anti)neutrinos with ν ($\bar{\nu}$), and pions with π. The figure is taken from Ref. [5].

Another issue of the SM, which may or may not be related to the masses of neutrinos, is the existence of a hidden, unknown form of matter in our Universe, with a total mass contribution exceeding the overall baryonic (plus negligible leptonic) mass content by around a factor of five. This puzzle, referred to as the nature of the dark matter (DM), will also be a part of our work on probes of BSM physics, which we will present in Chapter 4.

1.2 The role of neutrinos

Historically, the study of neutrinos has always played a forefront role in both establishing the rules governing the SM, and, given the incompleteness of the SM, in aiming to find the correct path towards an extension. The neutrino was suggested almost a century ago by W. Pauli as an elusive, massless particle accompanying the electrons emitted in β^- decays [6], as illustrated in Fig. 1.3, left. Considering the continuous electron spectrum as measured by Chadwick [7],

only an additionally emitted particle could preserve the most fundamental laws of energy and momentum conservation.

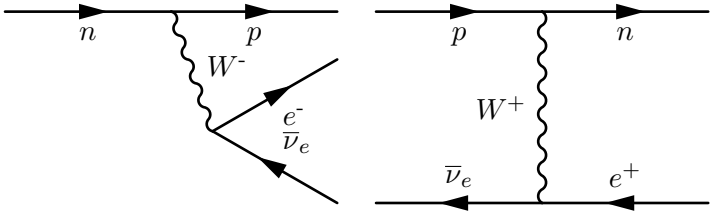

Figure 1.3: Left: Feynman diagram of the SM beta decay (β^-) process, in which a neutron decays into a proton, or more precisely, a down quark inside the neutron decays into an up quark, plus an electron and an electron-antineutrino. Right: Feynman graph of the inverse beta decay, in which an incoming electron antineutrino causes the conversion of a proton into a neutron, or on the quark level of an up quark into a down quark, plus a positron.

With the help of the theoretical foundation led by Fermi in his effective description of beta decays [8], in the 1950s Cowan et al. finally experimentally confirmed its existence via inverse beta decays of electron antineutrinos [9], as depicted in Fig. 1.3, right. The electron neutrino observation was closely followed by the detection of a second neutrino flavour, the muon neutrino, in the 1960s [10]. Even though the third generation of particles had been established in the 1970s and 1980s already, it took until the year 2000 to experimentally confirm the existence of the third neutrino flavour, the tau-neutrino [11]. This long time duration may indicate the experimental difficulties and challenges when probing the properties of these elusive neutrinos. The reason is that they neither interact strongly, nor electromagnetically, but only weakly, which makes a detection a demanding task. Besides their decisive role for the establishment of conservation laws such as the conservation of energy and momentum, neutrinos also let to the insight that some symmetries of Nature are actually violated. In 1956, Wu et al. observed parity violation in weak interactions via the measurements of the directionalities of electrons from β^- decays of ^{60}Co, and of the subsequently emitted γ-rays from the de-excitation of the produced ^{60}Ni nuclei [12]. A subsequent experiment in the same decade by Goldhaber et al. investigated the helicity of the neutrinos by measuring the analogous helicity of the γ particles emitted in the de-excitation of ^{152}Sm upon electron capture on ^{152}Eu [13]. They found that the helicity of neutrinos is equal to their chirality, sometimes referred to as handedness. Exclusively left-handed neutrinos and right-handed antineutrinos participate in weak interactions. Hence, parity violation is said to be maximal in neutrino interactions. Based on these two experiments neutrinos were assumed to be massless in the SM, and correspondingly only the left-handed neutrinos and right-handed antineutrinos were incorporated [14]. However, as we have pointed out already in Sec. 1.1, the observation of neutrino oscillations by the Kamiokande and SNO collaborations [3, 4] proves non-zero neutrino masses. One way of including a neutrino mass into the SM Lagrangian is via a Dirac mass term $\mathcal{L}_\mathrm{D}$ in precisely the same manner as for the other fermions [15]. Before electroweak symmetry breaking, the added term reads

$$\mathcal{L}_\mathrm{D} = -y\bar{L}H\nu_\mathrm{R} + \text{h.c.}\,, \tag{1.1}$$

where L denotes a lepton doublet, H is the Higgs field, y the Yukawa coupling parameter, and ν_R the right-handed neutrino. The abbreviation h.c. stands for the hermitian conjugate term, which has been omitted for brevity. After the phase transition breaking the electroweak symmetry, this term leads to a Dirac mass of the form $m_\mathrm{D} = \frac{v}{\sqrt{2}}y$, where v denotes the Higgs

vacuum expectation value. This straight-forward extension of the SM does not explain why the mass of the electron neutrino is so small compared to the mass of the electron if the generating mechanism is exactly the same. The currently tightest constrained on the yet not measured mass of the electron neutrino from the KATRIN experiment is 0.8 eV [16], which is *more than half a million times lower than the electron mass* of $m_e = 511$ keV. Another option of how to generate the neutrino mass is via a Majorana mass term. Already in 1937, E. Majorana had proposed to understand the neutrino and the antineutrino as the same particle [17]. Based on this interpretation, we can then understand the subsequent experimental observations via the existence of two different chirality eigenstates of the very same particle, the Majorana neutrino. The corresponding mass term to be added to the SM Lagrangian can be written as [18]

$$\mathcal{L}_{\mathrm{M}} = -\frac{1}{2}\bar{\nu}_L M_{\mathrm{M}} \nu_L^C + \mathrm{h.c.}\ , \tag{1.2}$$

where M_{M} denotes a Majorana mass, ν_{L} is the left-handed neutrino, and C refers to the charge conjugation operation. Following the motivation in Ref. [18], after symmetry breaking the total Lagrangian, taking into account both the Dirac and the Majorana mass terms, becomes

$$\mathcal{L}_{\mathrm{D+M}} = \mathcal{L}_{\mathrm{D}} + \mathcal{L}_{\mathrm{M}} = -\frac{1}{2}\bar{n}_{\mathrm{L}} M_{\mathrm{D+M}} n_{\mathrm{L}} + \mathrm{h.c.}\ . \tag{1.3}$$

Here the vector n_{L} is defined as

$$n_{\mathrm{L}} = \left(\nu_{\mathrm{L}}, \nu_{\mathrm{R}}^C\right)^T\ , \tag{1.4}$$

with T denoting the transposition operation, and the mass matrix reads

$$M_{\mathrm{D+M}} = \begin{pmatrix} m_{\mathrm{L}} & m_{\mathrm{D}} \\ m_{\mathrm{D}} & m_{\mathrm{R}} \end{pmatrix}\ . \tag{1.5}$$

Having obtained this 2×2 mass matrix containing both Dirac masses m_{D} and Majorana masses m_{L} and m_{R} for the left- and right-handed neutrino chirality eigenstates, we can now aim at understanding why the neutrino masses are so small compared to the masses of the other fermions. Via the so-called seesaw mechanism [19, 20], we can suppress m_{L} while simultaneously raising m_{R}. We can see this as follows. Let us introduce a new, heavy mass scale m_{N} with a tiny mixing angle $\theta = \frac{m_{\mathrm{D}}}{m_{\mathrm{N}}} \ll 1$, cf. [18]. If we set the masses of the left- and right-handed realisation to $m_{\mathrm{L}} = 0$ and $m_{\mathrm{R}} = m_N$, respectively, we can rewrite the seesaw mass matrix as

$$M_{\mathrm{D+M,\ S.}} = \begin{pmatrix} 0 & m_{\mathrm{D}} \\ m_{\mathrm{D}} & m_{\mathrm{N}} \end{pmatrix}\ . \tag{1.6}$$

The two eigenvalues $m_{1,2}$ of this matrix can be determined from the eigenvalue condition

$$0 \stackrel{!}{=} m_{1,2}^2 - m_{\mathrm{N}}\ m_{1,2} - m_{\mathrm{D}}^2 \tag{1.7}$$

as

$$m_{1,2} = \frac{1}{2}\left(m_{\mathrm{N}} \mp \sqrt{m_{\mathrm{N}}^2 + 4m_{\mathrm{D}}^2}\right) = \frac{1}{2}\left(m_{\mathrm{N}} \mp m_{\mathrm{N}}\sqrt{1 + 4\frac{m_{\mathrm{D}}^2}{m_{\mathrm{N}}^2}}\right)\ . \tag{1.8}$$

A first order Taylor expansion in the mixing strength $\frac{m_{\mathrm{D}}}{m_{\mathrm{N}}}$ yields the approximate solutions $m_1 \approx \frac{m_{\mathrm{D}}^2}{m_{\mathrm{N}}}$ and $m_2 \approx m_{\mathrm{N}}$. In this way, we obtain a light neutrino mass m_1 which is strongly suppressed with respect to the typical quark or lepton masses of order m_{D}, as observed. In

addition, the seesaw mechanism predicts the existence of a very weakly mixing heavy neutrino with an uplifted mass at our new mass scale m_{N}. Let us underline that the mechanism also incorporates lepton number violation, and thus provides a potential input for an explanation of the observed matter-antimatter asymmetry in the Universe. The model describing how to achieve these features based on the seesaw principle and/or extensions thereof, referred to as leptogenesis [21], goes beyond the scope of this experimental thesis, but is explained in e.g. in Refs. [22, 23]. There it is also mentioned how the framework links the lepton asymmetry to the needed baryon asymmetry. For instance, subsequent decays of the heavy neutrinos can generate an excess of baryons over antibaryons. For our needs it is sufficient to understand that probing the seesaw mechanism and a potential Majorana mass contribution to the neutrino mass is of utmost importance for the understanding of the matter-antimatter asymmetry in our Universe.

1.3 Experimental probes of neutrino masses

The experimental options to probe neutrino masses comprise the already mentioned neutrino oscillation measurements. The downside of these is that from oscillation results we cannot directly infer the absolute neutrino masses, as the transition probability for neutrinos to alter their flavour eigenstate only depends on the mass-squared differences of the three mass eigenstates, meaning only two mass splittings are currently known. We can understand this behaviour with the help of a toy model, as outlined in Refs. [24, 25]. In such a simplified treatment we consider only two neutrinos with a mixing angle θ and a mass-squared splitting Δm^2. We can now parametrise the mixing matrix as

$$U_{2\nu} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \ . \tag{1.9}$$

In the ultra-relativistic case, where the mass is negligible compared to the kinetic energy, a quantum mechanical flavour eigenstate α propagates in free-space as a plane wave,

$$|\nu_\alpha(L)\rangle = \exp\left(-i\Delta m^2 \frac{L}{2E}\right) |\nu_\alpha(0)\rangle \ , \tag{1.10}$$

where i is the imaginary unit, and E and L denote the energy and the propagation length of the neutrino, respectively. Note that in this formula we have used natural units, which we will also continue to do throughout this work. The transition probability of a neutrino to change its flavour eigenstate then becomes

$$P_{\alpha,\beta} = |\langle\nu_\alpha|\nu_\beta(L)\rangle|^2 = \sin^2 2\theta \sin^2\left(\Delta m^2 \frac{L}{4E}\right) \ , \tag{1.11}$$

i.e. we see that the oscillation probability only depends on the mass-squared difference. The full treatment for three neutrino flavours is based on a similar principle, but contains a more complicated mixing matrix, the Pontecorvo-Maki-Nagakawa-Sakata-matrix $U_{\alpha i}$, usually abbreviated as PMNS-matrix [26, 27], where α denotes the index of the flavour eigenstate of the weak interaction, and i the mass eigenstate. Neutrino oscillation experiments have measured the two mass-squared splittings involved for the three neutrino flavour case to be $\Delta m_{21} \approx 7.6 \times 10^{-5}$ eV^2 and $|\Delta m_{31}| \approx 2.5 \times 10^{-3}$ eV^2 [28]. However, the hierarchy of the involved masses is still unknown [29], meaning whether $m_3 > m_2 > m_1$, or $m_3 < m_1 < m_2$ is the correct hierarchy. The former scenario is referred to as *normal*, the latter as *inverted* neutrino mass ordering.
As we cannot deduce the actual neutrino masses from the oscillation measurements, we will need

to turn our focus to other probes sensitive to the neutrino masses. One option is to investigate the matter mass content and its distribution in the Universe, i.e. to use astrophysical and cosmological observations. These provide a constraint on the sum of the masses of neutrinos, given as

$$\Sigma = \sum_{i=1}^{3} m_i < 0.13\,\text{eV}\ [30]\ . \tag{1.12}$$

This constraint was derived based on observations of the distributions of galaxy clusters and superclusters via weak lensing studies, cf. [30]. During the era of structure formation in the early Universe, the matter density distribution had been affected by the propagation of neutrinos, referred to as free-streaming [31], causing a washing-out of perturbations in the matter distribution. Hence, the distributions of astrophysical and cosmological matter accumulations such as galaxy clusters provide a suitable dataset to constrain the maximal mass of neutrinos. A second option to determine the neutrino masses is via direct measurements in the laboratory, e.g. via a kinematic measurement of the β^- decay spectrum near the endpoint, or via probing neutrinoless double beta decay ($0\nu\beta\beta$). In the former case, a non-vanishing neutrino mass impacts the distribution of the kinetic energy of the electron, causing a kink-like distortion in the β spectrum located $Q_\beta - m_\beta$, where Q_β and m_β denote the Q-value of the decay and the electron neutrino mass emitted in the β decay. The energy spectrum then follows the form, cf. [32],

$$\frac{dN}{dE} \propto p(E)\ (E + m_e)\ (Q_\beta - E)\ \sum_{i=1}^{3} |U_{ei}|^2 \sqrt{(Q_\beta - E)^2 - m_i^2}\ \Theta(Q_\beta - E - m_i)\ , \tag{1.13}$$

where p and E are the momentum and energy of the emitted electron, and Θ is the Heaviside-Theta function. This radicand implies that in the presence of at least one non-vanishing neutrino mass, the maximum energy of the electron is reduced compared to the maximum in the case of vanishing m_i, which would be equal to Q_β. As mentioned in Sec. 1.2, the most sensitive constraint on m_β has been obtained by the KATRIN collaboration as

$$m_\beta = \sqrt{\sum_{i=1}^{3} |U_{ei}|^2 m_i^2} < 0.8\,\text{eV}\ [16], \tag{1.14}$$

where U_{ei} are PMNS-matrix elements introduced above. The other laboratory-based probe mentioned before, $0\nu\beta\beta$, allows us to probe a Majorana mass contribution to the neutrino masses, constrained to be

$$m_{\beta\beta} = \left| \sum_{i=1}^{3} U_{ei}^2 m_i \right| < 36\text{-}156\,\text{meV}\ [33] \tag{1.15}$$

by the KamLAND-Zen collaboration. In Fig. 1.4 we show the allowed parameter space for the neutrino Majorana mass as a function of the minimum neutrino mass m_{min}, of m_β, and of Σ. We indicate both the range of solutions in the normal, and the inverted ordering. The bands were derived by solving Eqs. (1.12), (1.14), (1.15) for free m_{min}, and using the measured best-fit parameters from neutrino oscillation experiments as listed in Ref. [29].

Let us underline the importance of combining all the information on the neutrino masses, the oscillation, the cosmological, the β^- decay, and the $0\nu\beta\beta$ decay data, in order to obtain a full picture of the neutrino properties. In particular the latter problem, the existence of $0\nu\beta\beta$, needs to be solved urgently, given its crucial role for both the understanding of the underlying neutrino mass generation mechanism, and for the cosmological consequences due to its violation of lepton number conservation. We will hence now turn to detailed discussion of this paramount process.

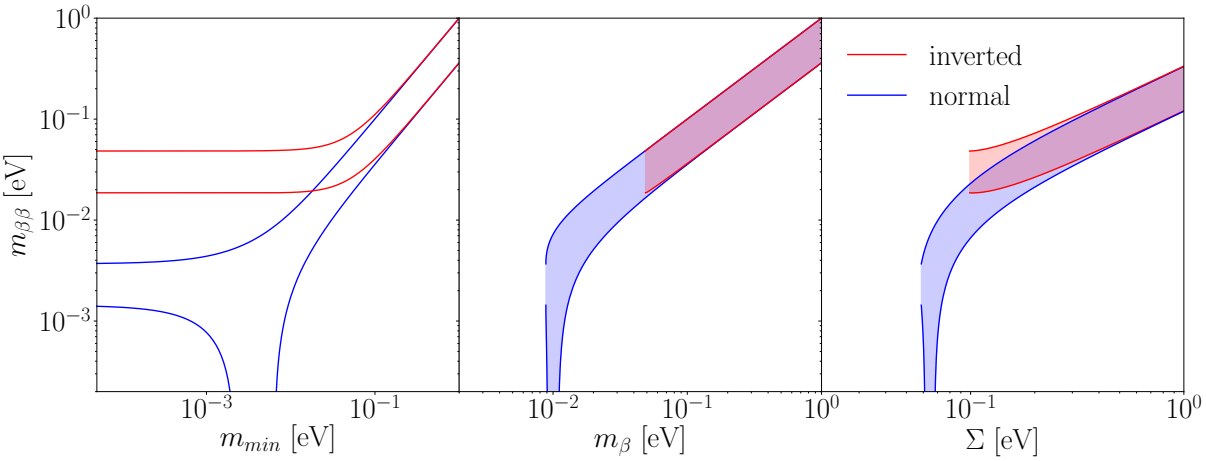

Figure 1.4: The neutrino Majorana mass as a function of m_{min} (left), of m_β (middle), and of Σ (right). The blue and red bands correspond to the normal and inverted mass ordering regimes, respectively, and were evaluated with the measured best-fit neutrino oscillation parameters taken from Ref. [29], neglecting the uncertainties. The bands were derived by solving Eqs. (1.12), (1.14), (1.15), which are satisfied by a range of mass combinations, hence the bands instead of point or line solutions.

1.4 Neutrinoless double beta decay

The single β^- decay occurs if a nucleus has an excess of neutrons, as discussed in Sec. 1.2, the conversion of neutron into a proton under the emission of an electron plus an electron-antineutrino. Already in 1935, M. Goeppert-Mayer suggested that the process may also happen twice simultaneously, with a double neutron-to-proton conversion, accompanied by the emission of two electrons and antineutrinos each, see [34]. We illustrate this process in Fig 1.5, left. We abbreviate this decay as $2\nu\beta\beta$, in correspondence with the particles emitted from the nucleus, namely two $\bar{\nu}_e$ and two electrons. A required condition for this transition to happen is that the single β^- decay is energetically forbidden or highly suppressed. Otherwise this process, being a first order weak process, would always occur before the doubly suppressed second order weak transition $2\nu\beta\beta$.

For 44 even-even nuclei the nuclear mass is below the mass of the final-state isotope in an expected single β^- transition, making the double emission theoretically feasible [35]. We show an illustration of these considerations in Fig. 1.6, left.

Due to the strong suppression of the decay rate, only around a dozen of isotopes are experimentally feasible for a detection. Hence, it took fifteen years after the prediction until the first observation of $2\nu\beta\beta$ decay in the isotope ^{130}Te in geochemical probes [36]. The first direct laboratory-based measurement happened only in 1987, employing the isotope ^{82}Se [37]. Despite the fact that with half-lives typically on the order of $10^{20} - 10^{21}$ yr [38] the decays are extremely rare, $2\nu\beta\beta$ is a SM process which can be described by the weak interaction, and it does not violate any conservation law. It also does not provide knowledge on the light SM neutrino mass on an experimentally feasible level, as the energy scale is around MeV, i.e. too high to identify any kink-like feature at below the eV level around the Q-values, as discussed for the single β^- decay spectrum in Sec. 1.3.
Most importantly, double beta decay may not only occur with the simultaneous emission of the $\bar{\nu}_e$, but if neutrinos have a Majorana mass contribution also without any neutrino emission,

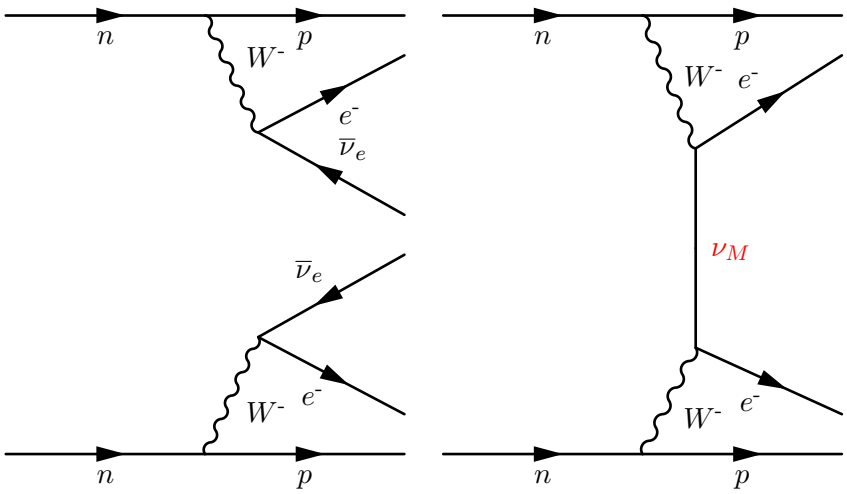

Figure 1.5: Left: Feynman diagram of the established SM process two neutrino double beta decay, with two electrons and two neutrinos in the final state. Right: Feynman diagram of the hypothetical BSM process neutrinoless double beta decay, in which only the two electrons appear in the final state, but no neutrinos, implying both a Majorana mass contribution to the neutrino mass and the violation of the conservation of lepton number.

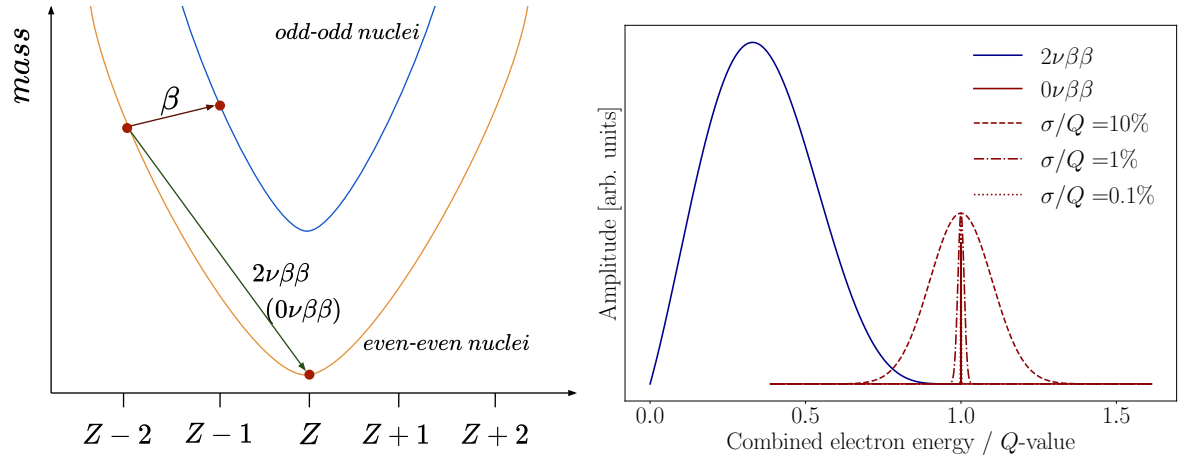

Figure 1.6: Left: illustration of the mass parabola requirement on the nuclear masses of double beta decaying isotopes to make a single beta transition energetically impossible or highly improbable, which is needed to enable double beta decays. Here Z denotes the number of protons in the nucleus. Right: experimental signature of the two double beta decay processes. The $2\nu\beta\beta$ spectrum as a function of the summed electron energy is continuous up to Q, whereas the imprint of $0\nu\beta\beta$ is a single peak located at Q directly, here indicated for different levels of the relative energy resolution σ/Q. Note that the $0\nu\beta\beta$ decay amplitude is unknown.

as indicated in Fig. 1.5, right. In this scenario distinguishing between antineutrino and neutrino would not be an accurate description of neutrinos, enabling the possibility of a Majorana neutrino annihilation. The result would be the emitted final-state particles are only the two electrons, which explains why we introduced the abbreviation $0\nu\beta\beta$ in the previous section, as typically done in the field. The decisive consequence would not only be the Majorana mass contribution, but also the violation of the lepton number by two, indicating that this is intrinsically a BSM process. If $0\nu\beta\beta$ would be observed, information on the neutrino Majorana mass could

be extracted. In this case, the probability of the process to happen would depend on two quantum mechanical quantities. One is the nuclear matrix element $M_{0\nu}$ containing the properties of the nucleus under consideration, and the other the phase space factor $G_{0\nu}$ representing the final states of the leptons. Measuring the process in a certain isotope with a half-life of $T^{0\nu}_{1/2}$, the effective Majorana mass could be determined as [39]

$$m_{\beta\beta} = \frac{m_e}{\sqrt{|M_{0\nu}|^2 G_{0\nu}(Q_{\beta\beta}, Z) T^{0\nu}_{1/2}}} \ . \tag{1.16}$$

Whereas phase space factors have been calculated accurately for the isotopes of experimental interest [40, 41], the nuclear matrix elements, with discrepancies between different calculation methods of around a factor three [42, 39], remain the main source of uncertainty when interpreting experimental $T^{0\nu}_{1/2}$ constraints as detailed in Sec. 1.6. Note that as outlined in Ref. [39], several $0\nu\beta\beta$ exchange mechanisms have been proposed, but here we limit ourselves to the standard case of the light Majorana neutrino exchange. However, independently of the underlying mechanism, an observation of $0\nu\beta\beta$ decay would always imply lepton number violation and a Majorana mass of neutrinos, as proven in Ref. [43].
Having discussed the relevance of the study of $0\nu\beta\beta$, we can now turn to the experimental signature of the two different double beta decay versions. The $2\nu\beta\beta$ decay energy spectrum of the two electrons as seen in an experiment is continuous, similarly to the single β^- decay spectrum, as the neutrinos interact only weakly and thus escape a detection. A convenient description of the $2\nu\beta\beta$ electron energy spectrum is the Primakoff-Rosen parametrisation derived in Ref. [44], which reads

$$f(E) = (E^4 + 10E^3 + 40E^2 + 60E + 30)\frac{E(Q-E)^5}{900} \ . \tag{1.17}$$

We provide a plot of this continuous function in Fig. 1.6, right. In contrast, if no neutrinos are emitted, all the available energy is transferred to the two electrons, implying the theoretical spectrum would be a monoenergetic peak at the Q-value. In a realistic environment the energy resolution with which the energy of the two electrons can be measured is limited, which broadens the peak, as indicated in Fig. 1.6, right. To be able to discriminate the $0\nu\beta\beta$ signal peak from the $2\nu\beta\beta$ continuum, it is thus of utmost importance to achieve the best possible energy resolution at $Q_{\beta\beta}$, which is typically at the MeV scale. In addition, let us keep in mind that the measured half-lives of the SM process exceed the age of the Universe by already around ten orders of magnitude, cf. [45]. The expected rarity of the BSM process with respect to the SM decay hence necessitates the use of large source masses. These considerations lead us towards the experimental key parameters which guide the laboratory searches for $0\nu\beta\beta$, to be discussed in the following section.

1.5 Experimental considerations

From an experimental point of view, Eq. (1.16) implies that the parameter of interest, which can be optimised to increase the sensitivity towards lower $m_{\beta\beta}$, is the half-life $T^{0\nu}_{1/2}$. We will now derive mathematically how this decisive parameter is impacted by the experimental key parameters. These are the source material mass m and the measurement duration t, which are usually combined into the data exposure $\lambda = mt$ for convenience, the background level B, the energy resolution σ_E, the signal detection efficiency ϵ, the isotopic mass M, and the isotope enrichment fraction f. The derivation closely follows the one given in Ref. [46]. Starting from the radioactive decay law, we can estimate the total number of decays observed in an experiment

N as

$$N = \epsilon N_0 \left(1 - e^{-\frac{t \ln 2}{T^{0\nu}_{1/2}}}\right) \approx \epsilon N_0 \frac{t \ln 2}{T^{0\nu}_{1/2}} = \epsilon \frac{\lambda f N_\mathrm{A} \ln 2}{M T^{0\nu}_{1/2}} \ , \tag{1.18}$$

where N_0 denotes the number of initially available nuclei, and N_A is Avogadro's constant. The first order Taylor expansion of the exponential function is well-justified in this case. The reason is the already excluded half-life limits exceed 10^{26} yr [33, 47], an enormous time span compared to the experimental run times of a couple of years, making the restriction to the linear term in Eq. (1.18) an excellent approximation. If we now assume to observe N events in our experiment, the corresponding measured half-life becomes

$$T^{0\nu}_{1/2} = \epsilon \frac{\lambda f N_\mathrm{A} \ln 2}{M N} \ . \tag{1.19}$$

Unfortunately, no convincing, clear evidence of $N > 1$ has been found yet. Hence, we can only substitute $N^{'}$ as a constraint on the number of signal events collected during the physics data exposure at a given confidence level. In the case of sufficiently long half-lifes, the Poisson distribution describes our ultra-rare number of potential decay events, in principle following a Binomial distribution, extremely accurately [48]. We can thus evaluate the expected background event count N_B in the signal region-of-interest (ROI) for an experiment being dominated by background contributions as

$$N^{'} = \sqrt{N_B} = \sqrt{\lambda B \sigma_E} \ . \tag{1.20}$$

Note that B is given in units of counts / (keV kg yr). Here we can identify a square root suppression of the available exposure, mitigating an optimal use of the source material. In an ideal experiment, operating in the so-called background-free regime, where the following condition holds,

$$\lambda B \sigma_E < 1 \ , \tag{1.21}$$

the sensitivity is not limited by the statistical fluctuations of the background, but of the signal itself. Here the linear scaling of $T^{0\nu}_{1/2}$ with λ, see Eq. (1.19), can be kept. In reality a full suppression of any background is, at least in the vast majority of cases, unfeasible. Hence, in the presence of a background which is sufficiently low to satisfy the condition of Eq. (1.21), one refers to operation in the *quasi* background-free regime. The half-life sensitivity for the two cases can be summarised as

$$T^{0\nu,'}_{1/2} = \epsilon \frac{f N_\mathrm{A} \ln 2}{M} \begin{cases} \sqrt{\frac{\lambda}{B\sigma_E}} & \text{background-dominated} \\ \lambda & \text{quasi background-free} \end{cases} \ . \tag{1.22}$$

From Eq. (1.22) we deduce the following desired properties of an experiment aiming to detect $0\nu\beta\beta$. To make optimal use of the available source material, or in other words to reach a linear scaling of the experimental sensitivity with the exposure, the quasi background-free regime should be reached. That is, the background should be sufficiently suppressed to be dominated by the signal fluctuation. The energy resolution essentially behaves similarly as the background suppression, and its improvement should be considered equally important. In particular the suppression of the continuous $2\nu\beta\beta$ component penetrating the $0\nu\beta\beta$ signal region dominantly relies on an optimal energy resolution, suppressing this background contamination as [38]

$$r_{2\nu/0\nu} \propto \frac{T^{0\nu}_{1/2}}{T^{2\nu}_{1/2}} \left(\frac{\sigma_E}{Q_{\beta\beta}}\right)^6 \ . \tag{1.23}$$

This is extremely important considering the identical signal topology of $2\nu\beta\beta$ and $0\nu\beta\beta$ for any experiment which is not able to tag the emitted neutrinos. As neutrinos only interact weakly, this is obviously the case for any feasible $0\nu\beta\beta$ detector, as we will touch upon in Sec. 1.6. From Eqs. (1.22), (1.23) we deduce that besides the background suppression and the resolution performance enhancement, an experiment should measure for long periods, and deploy as much source material as reasonably affordable. Moreover, the efficiency of the signal detection should be increased as much as possible to identify any signal decay in the source material. Lastly, the natural abundance of the target isotope should be enriched to the highest fraction achievable. It is worth mentioning that background contributions other than $2\nu\beta\beta$ can be typically identified with an efficiency different from the signal detection efficiency, and often do not depend on the enrichment fraction. Hence, increasing ϵ and f does not only impact the linear scaling term, but implicitly also the inverse scaling of the background, and is consequently of high relevance. To built an experiment, one may thus choose an isotope with a naturally high abundance of the target isotope, operate the detector system in direct vicinity of the source material, or, where possible, directly deploy the source material as the detection medium, and shield the apparatus from external background. In addition, a high Q-value should be preferably chosen to naturally measure in a region beyond the energies of betas from natural radioactive sources, i.e. beyond $\sim$2.6 MeV, the energy of the ^{208}Tl γ-ray of the ^{232}Th decay chain [49]. Let us remark that a high Q-value also helps in suppressing the background contribution from the $2\nu\beta\beta$ continuum, as implied by Eq. (1.23). In Tab. 1.1 we list the properties of the $2\nu\beta\beta$ isotopes that have been identified to be experimentally feasible to perform a $0\nu\beta\beta$ search [50].

Isotope	$Q_{\beta\beta}$ [keV]	Natural abundance [%]	$M_{0\nu}$	$G_{0\nu}(Q_{\beta\beta}, Z)$ [10^{-14} yr^{-1}]
^{110}Pd	2004.0	11.7	5.33 - 8.91	1.40
^{76}Ge	2039.1	7.8	2.81 - 7.24	0.623
^{124}Sn	2287.7	5.6	2.62 - 4.81	2.55
^{136}Xe	2461.9	8.9	1.71 - 4.2	4.31
^{130}Te	2530.3	34.5	2.65 - 5.50	4.09
^{116}Cd	2809.1	7.6	2.51 - 4.72	4.62
^{82}Se	2995.5	9.2	2.64 - 6.46	2.70
^{100}Mo	3035.0	9.6	3.103 - 7.77	4.36
^{96}Zr	3347.7	2.8	1.56 - 5.65	5.63
^{150}Nd	3367.3	5.6	1.71 - 3.7	19.2
^{48}Ca	4273.7	0.187	0.85 - 2.37	6.35

Table 1.1: Overview of experimentally potentially feasible isotopes undergoing $2\nu\beta\beta$ decay, ordered by their respective Q-value. The natural isotopic abundance, the nuclear matrix elements, and the phase space factors are listed as well. The values are taken from Ref. [50].

Tab. 1.1 also lists the phase space factors of the relevant isotopes. In order to maximise the sensitivity for the lowest neutrino mass possible, from Eq. (1.16) we deduce it should be maximal, as should be the nuclear matrix elements. In Chapter 2 we will see that these generic considerations for an ideal $0\nu\beta\beta$ experiment oversimplify the actual experimental challenges. In particular, germanium based experiments do not necessarily seem to be ideal candidates, according to the naive expectations of suffering from a low natural abundance, phase space factor and Q-value. Yet, these experiments have a long track record of world-leading $0\nu\beta\beta$ sensitivities, for the reasons we will learn below.

1.6 Experimental programs in the field

Several different experimental approaches are being pursued to fullfill the challenging requirements in this rare-event search as implied by Eq. (1.22). A typical common principle is to deploy the source material also as the detector to enhance the signal detection efficiency, and to operate the detectors underground with specifically designed active and passive shields to reduce external background contributions. Here we provide a brief overview of the different strategies currently conducted, and/or planned in the future, closely following the presentation of the experimental program in Section 5 of [51].

Semiconductors Experiments operating semiconductor detectors aim to measure the charge induced by the two electrons emitted in $0\nu\beta\beta$. Usually high-purity germanium (HPGe) detectors enriched in the isotope ^{76}Ge are deployed, as e.g. done by GERDA or MAJORANA DEMONSTRATOR (MJD). These intrinsically clean crystals, being free of any contamination at a measurable level, as determined by GERDA [52], provide the best energy resolution of any technology in the field at a normalised level at around 0.1% at $Q_{\beta\beta}$, as obtained by MJD [53]. Hence, the world-leading half-life sensitivity exceeding 10^{26} yr could be realised by GERDA [47], as shown in Tab. 1.2. The sensitivity will be further improved by around two orders of magnitude by the LEGEND program [54]. We will provide more details on semiconductor detector operation and the corresponding experiments in Chapter 2.

Time-projection chambers Time-projection chambers employ both ionisation and scintillation signal read-out to differentiate the $0\nu\beta\beta$ signal from background events. The detectors can be filled to large masses with either both a liquid and a gaseous phase of the target material, as e.g. done in XENON1T [55] or with a liquid phase, see e.g. EXO-200 [56], or a high-pressure gaseous phase only, see e.g. NEXT-100 [57]. Searching for $0\nu\beta\beta$ in the isotope ^{136}Xe, the next-generation experiments nEXO [58], DARWIN [59], NEXT [60], and PANDAX-III [61] experiments will reach high sensitivities beyond 10^{27} yr, particularly triggered, among others, by the comparatively easy source mass scalability of up to several tonnes.

Bolometers Bolometers are crystals operated at ultra-low temperatures. Depending on the crystal element, phonons and/or scintillation signals are read out. Similarly to the semiconductors, bolometers can be ultra-pure and a very good relative energy resolution at below the % level at $Q_{\beta\beta}$ can be achieved, as shown by CUORE [62] and CUPID-0 [63]. Based on the current experience with crystals containing natural Te (both ^{128}Te and ^{130}Te undergo $2\nu\beta\beta$), ^{82}Se, and ^{100}Mo, the future CUPID program aims to exceed half-life sensitivities of 10^{27} yr [64]. The AMORE program aims to achieve the same order of half-life sensitivity for the isotope ^{100}Mo by operating calcium molybdate crystals [65].

Scintillators Owing to their comparatively simple, large source mass scalability, the liquid scintillator-based experiment KAMLAND-ZEN has obtained the currently leading $T^{0\nu}_{1/2}$ constraint of beyond 10^{26} yr [33]. The experiment loaded an organic liquid scintillator with ^{136}Xe, a strategy that will be further pursued to exceed half-life sensitivities beyond 10^{27} yr [66]. Also the ZICOS [67] and SNO+ [68] experiments will operate liquid scintillators, loaded with ^{96}Zr and ^{130}Te, respectively, to reach sensitivities around that order of magnitude. Inorganic, solid scintillators are investigated for $0\nu\beta\beta$ searches as well. For example, the CANDLES program reached a $T^{0\nu}_{1/2}$ limit beyond 10^{22} yr in ^{48}Ca [69], reading out the scintillation light from both a liquid and a solid target.

Tracking calorimeters The main difference between tracking devices and the approaches mentioned before is the separation between source and detector material. In contrast to the other techniques, tracking experiments deploy a layer of source material which is surrounded by a gaseous tracking system and a calorimeter detector in an onion-shell configuration. In this manner, not only the energy can be measured, but also the angle between the two emitted electrons. Also, the $0\nu\beta\beta$ search can be performed with different source materials. The disadvantages of this approach are the scalability to large source masses and the limited detection efficiency. Based on the experience of the NEMO experiment [70], its successor SuperNEMO [71] aims to reach a sensitivity beyond 10^{26} yr for the $0\nu\beta\beta$ decay in ^{82}Se.

This brief, clearly non-comprehensive overview of the ongoing experimental approaches, indicates how actively the research in this field is pursued due to the important consequences if an observation of $0\nu\beta\beta$ would be made. We list the currently leading measured half-life constraints for the most-common isotope choices employed in the field in Tab. 1.2.

Experiment	Isotope	Exposure [kg yr]	$T^{0\nu}_{1/2}$ [10^{26} yr]	$m_{\beta\beta}$ [meV]
GERDA [47]	^{76}Ge	127.2	1.8	79 - 180
MAJORANA DEMONSTRATOR [53]	^{76}Ge	64.5	0.83	113 - 269
KamLAND-Zen [33]	^{136}Xe	970	2.3	36 - 156
EXO-200 [56]	^{136}Xe	234.1	0.35	93 - 286
CUORE [62]	^{130}Te	1038.4	0.22	90 - 305

Table 1.2: List of currently leading lower half-life limits of the $0\nu\beta\beta$-decay at 90% confidence level (CL), together with the corresponding range of $m_{\beta\beta}$ as provided by the collaborations.

There we also indicate the Majorana mass range covered assuming light Majorana mass exchange. Under this assumption the experimental constraints have just started to reach the inverted neutrino mass ordering scheme, see Fig. 1.4. Ongoing and future experiments such as LEGEND, as introduced in the following chapter, are required to cover this regime entirely.

Chapter 2

The GERDA and the LEGEND experiments

After motivating the experimental requirements to search for $0\nu\beta\beta$ decay and describing the experimental approaches pursued in the field at the end of the previous chapter, we are now prepared to discuss the details of the operation of HPGe detectors. We will start by introducing their general properties and elaborating on why these are well-suited for $0\nu\beta\beta$ searches. Afterwards, we will turn to the description of the GERDA and LEGEND experiments, on which the research projects of this work have been focused.

2.1 Germanium detector operation

Let us firstly discuss the general properties of Ge semiconductor detectors, their properties and operation, before detailing their performance in calorimetric measurements, particularly in terms of their unparalleled energy resolution.

2.1.1 Ge semiconductor detectors

Germanium detectors are made of Ge crystals, which, from the point of view of electronic band structure, are semiconductors. Modelling the conduction properties of any material via their band structure explains in rather simple terms why upon applying a voltage, currents flow readily in conductors, or no currents flow in insulators. The model works as follows. In its ground state, each material type, i.e. conductor, semiconductor, or insulator, is assigned a particular band structure, with the valence band being fully occupied by electrons. Due to the Pauli exclusion principle [72], only two electrons are allowed to occupy each state, with states of increasing energy up to the Fermi level E_F. In contrast, the conduction band exceeding E_F is free of charge carriers. When applying a voltage, electrons can now flow in conductors, as valence and conduction bands overlap. In insulators, however, the gap between these two bands typically exceeds at least around 10 eV, meaning electrons will not be lifted into the conduction band, and a current can hardly be induced. The band structure of semiconductors is similar to that of insulators, but with the important difference that the band gap is typically $\mathcal{O}(1\,\text{eV})$. For example, the band gap of Ge is 0.67(1) eV at room temperature [73, 74, 75, 76]. Hence, electrons, absorbing deposited energy, can reach a state in the conduction band, resulting in a current flow, or in other words a *potential signal* that can be read out electronically. A schematic of the band theory is pictured in Fig. 2.1.

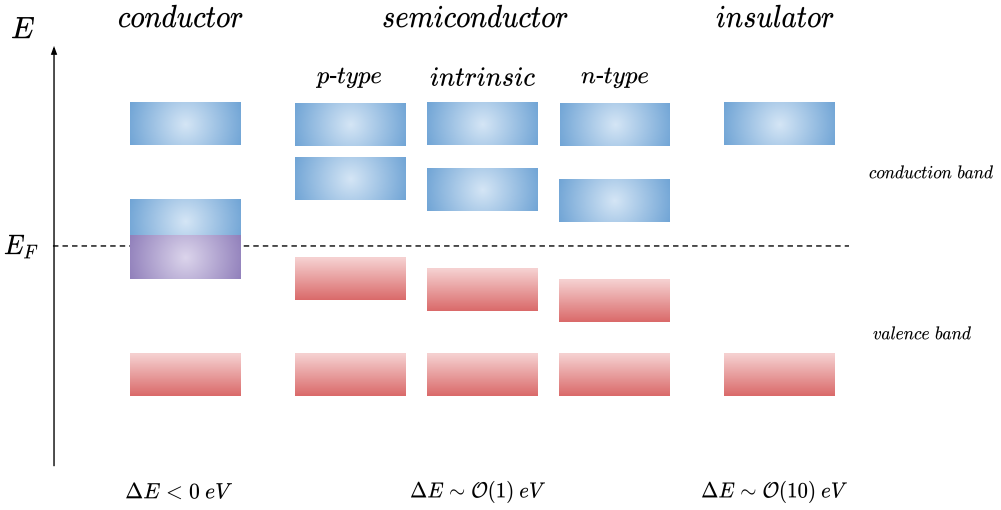

Figure 2.1: Illustration of the band structure of different materials, with a separation between semiconductors of different doping. The energy scale is abbreviated with E, E_F denotes the Fermi energy, and ΔE is the energy gap between valence and conduction bands.

It now becomes clear that semiconductors are highly suitable to perform calorimetric measurements, meaning detecting energy depositions in a material. Most semiconductor detectors benefit from being cooled down to cryogenic temperatures, as thermal excitations can induce noise currents, potentially overlaying relevant energy depositions. In the case of Ge, temperatures around the liquid nitrogen boiling temperature of 77 K are sufficient to suppress the thermal noise. In addition, impurities inside a semiconductor that cannot be fully prevented during the crystal growth process, can induce additional electrons or holes. The latter denote missing electrons in a state which can propagate in an electric field, just as electrons, with an opposite charge. To prevent the currents induced by impurities, excess electrons or holes are introduced into semiconductor detectors intentionally during the fabrication process. This process is referred to as n-type or p-type doping, for negative or positive charge carriers, respectively. The combination of both is referred to as a p-n junction. As an electric field is generated by charge propagation away from the junction in the region between two oppositely doped zones, an inner layer free of charges is established. By applying a reverse bias voltage of typically around 1 to 5 kV for Ge detectors, the charge-depleted centre region can be further enlarged. If now an energy deposition occurs within the depleted area, the external electric field applies a force on the produced electrons and/or holes, causing them to drift towards the n- or p-typed layer, respectively. We depict the formation of the charge-depleted zone in Fig. 2.2.

The current i, or equivalently the charge Q, induced on the readout electrode is then given as

$$i = \frac{dQ}{dt} = -\frac{d}{dt}\frac{q}{V_u}\int_{\vec{x}_i}^{\vec{x}_f} \nabla\phi \cdot \ d\vec{x} = \frac{q}{V_u}\vec{E}_u \cdot \vec{v}\ , \tag{2.1}$$

which is known as the Shockley-Ramo theorem [77, 78]. Here the integral runs over the position inside the drift layer from the initial position $\vec{x}_i$ to the position of the electrode $\vec{x}_f$, q is the charge produced in the initial interaction with instantaneous velocity $\vec{v}$, and ϕ denotes the potential at the interaction point. The voltage applied at the electrode, V_u, can be conveniently set to 1 V if all other electrodes are grounded, as the induced charge is independent of V_u. In this case, the normalised potential is referred to as the weighting potential. Finally, the electric field

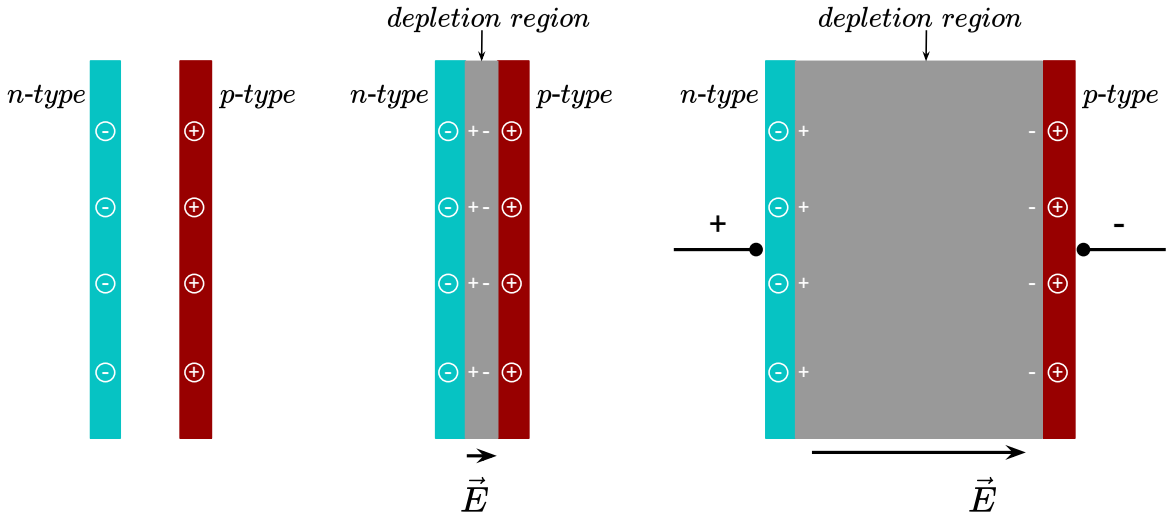

Figure 2.2: Sketch of the formation of the charge-depleted region in a semiconductor. Left: p-typed and n-typed layers have an excess of positive and negative charge carriers, respectively. Centre: combining the two oppositely doped layers induces a drift of free charges, causing an accumulation of electrons at the p-contact, and of positive charges at the n-contact. In the central layer electrons and holes recombine, leaving a zone free of charges across which an electric field is produced. Right: by applying a reverse bias voltage, the depletion region is enlarged.

induced by V_u at the non-grounded electrode is denoted with $\vec{E}_u$. In the Ge detectors operated in our experiments, as described in Secs. 2.2, 2.3, the p-type layer is produced by implanting a $\sim$100 nm thick layer of boron, whereas n-type doping is accomplished via lithium diffusion [79]. The latter produces a dead-layer at the edge of the detectors with an approximate thickness of 1 mm. The remaining bulk material, with typical dimensions of a few cm, is operated as active material sensitive to energy depositions.

2.1.2 Energy estimation and resolution

Energy estimation The energy measurements with the Ge detectors operated in the experiments detailed in this chapter work as follows. Radioactive sources emitting X-rays and γ-rays are typically used for the calibration of the energy scale in the keV to MeV range, as detailed in Chapter 3. These photons undergo different processes inside the Ge detectors, depending on their energy [80]. Below approximately 100 keV, photoelectric absorption [81] is the most probable process to occur. Compton scattering [82] dominates in the range between around 100 keV to 10 MeV. Electron-positron pairs can be created if the energy exceeds the production threshold of $2m_e$, cf. [83]. This process becomes most relevant when the energy goes beyond approximately 10 MeV.
When an interaction of a particle within the depleted zone occurs, the secondary charge induced on the p^+ electrode is collected by a capacitor-resistor feedback loop, and subsequently transported to an analog-to-digital converter (ADC) for digitisation. In GERDA and LEGEND, signal preamplification is required to combat signal attenuation caused by the transmission through several meter long signal cables. Upon digitisation, the resulting waveform then has a typical three-component pattern. First, the baseline, second, a steep rise when the charge signal arrives, typically around $\mathcal{O}(10^{-1})$ µs, and third, a long tail with typically $\mathcal{O}(10^2)$ to $\mathcal{O}(10^3)$ µs

before baseline restoration. The physical energy deposited is proportional to the induced charge, which is itself proportional to the maximum of the waveform pattern. We give an example of a simulated waveform in Fig. 2.3.

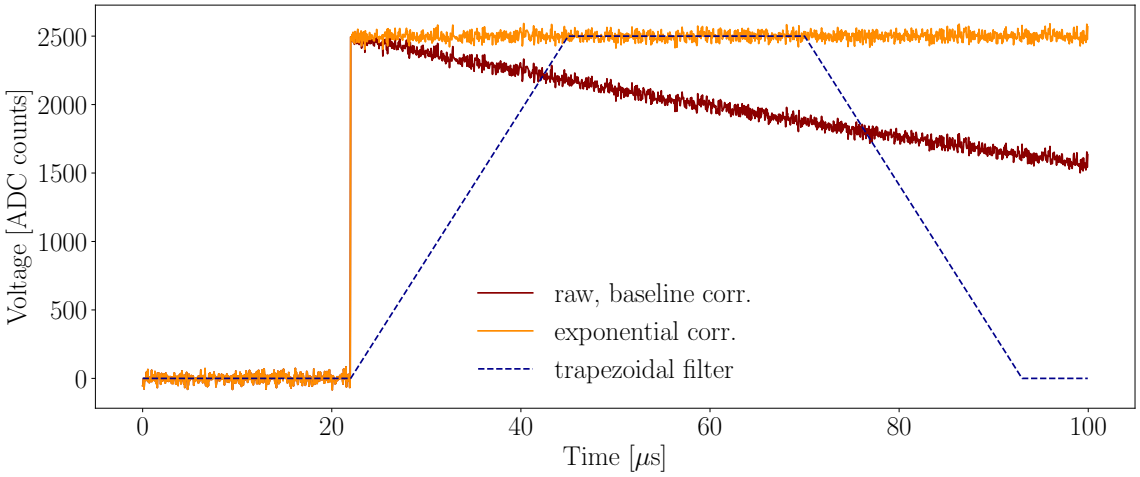

Figure 2.3: Simulated typical waveform for the recorded voltage in ADC counts as a function of time. The waveform follows the typical pattern of a rise from the here subtracted baseline to its maximum value upon charge collection, and the subsequent exponential decay while discharging, overlaid by noise fluctuations (red). An exponential correction is conducted on the time series data to improve the accuracy of the estimation of the maximum via a trapezoidal filter (dashed blue).

The following filtering procedure is applied to estimate this charge accurately. The baseline is subtracted, and an exponential correction for the decrease of the tail can be performed to obtain a flat waveform maximum. In principle, these two steps are optional, as only the difference between maximum and minimum is relevant. In particular the exponential correction however simplifies accurately estimating the maximum, and is hence typically implemented. A basic signal shaping filter, such as a moving average of the time series data, can be subsequently applied to suppress high-frequency noise. Calculating the difference between consecutive averaged amplitudes results in a quasi-Gaussian signal shape. As a next step, a more sophisticated filter can be used to obtain an accurate estimate for the maximal charge. Filters that have been utilised in this step comprise, among others, a Gaussian filter, a trapezoidal filter, a cusp filter based on a sinh and cosh function, or a Zero Area Cusp (ZAC) filter, the latter putting a flat top between the hyperbolic functions [84, 85]. Details on some of these signal filtering techniques can be found in Ref. [86]. We remark that GERDA used a moving average procedure, and then applied a ZAC filter to estimate the energy [87], whereas in LEGEND, trapezoidal, cusp, and ZAC filters are being employed independently. At the current stage of LEGEND, the filter yielding the optimal energy resolution per detector is being taken as energy estimate, although the final filtering procedure has not been decided yet. We show an example application of a trapezoidal filter to a simulated waveform in Fig. 2.3. After signal filtering, only an estimate on the collected charge has been determined, in terms of the maximal voltage value in ADC counts. The translation into a physical energy estimate, i.e. the calibration of the detector energy response, is performed by inducing γ-ray signals of known physical energy, and comparing these to the determined corresponding ADC values. We will discuss the details of the calibration in Chapter 3.

Energy resolution In Sec. 1.5, we discussed the importance of the resolution of the energy scale estimate. Here, we closely follow the explanations from Ref. [76], describing how the limited resolution arises. In general, resolution broadening is caused by both intrinsically physical effects, and experimental parameters such as detector operation conditions and read-out equipment. More precisely, the variance, i.e. the square of the resolution, of the measured energy can be written as

$$\sigma_E^2 = \sigma_{E,I}^2 + \sigma_{E,P}^2 + \sigma_{E,C}^2 + \sigma_{E,N}^2 \ . \tag{2.2}$$

Here $\sigma_{E,I}$ denotes the intrinsic uncertainty of the energy of a γ-ray used for the calibration, $\sigma_{E,P}^2$ and $\sigma_{E,C}^2$ are the uncertainty in the number of created electron-hole pairs, and collected charge, upon an energy deposition, respectively, and $\sigma_{E,N}^2$ is the variance of the noise induced by the readout electronics chain. According to the quantum mechanical Heisenberg uncertainty principle [88], we can estimate the magnitude of the intrinsic resolution via $\sigma_{E,I}\tau \sim \hbar/2$. Typical lifetimes in nuclear processes causing γ emissions range from around 10^{-6} to 10^{-12} s, implying a maximal energy uncertainty of $\mathcal{O}(0.1\,\text{meV})$. As the energy scale of nuclear processes is MeV, where Ge detectors achieve keV level energy resolution, the intrinsic contribution to the total resolution can be safely neglected. Note that further intrinsic processes, such as nuclear recoils induced by γ-ray emission, and a Doppler shift due to the relative thermal motion of the nucleus, do not alter the conclusions for the majority of peaks, since the recoiling nucleus is contained inside the Ge crystal, and no resonance effects are induced, unlike in Mössbauer spectroscopy [89]. In the absence of resonances, the large crystal mass suppresses both recoil and thermal broadening orders of magnitude below the relevant keV scale. We will see an exception to these considerations in Sec. 3.3, where an additional relative motion causes a non-negligible Doppler broadening.

The next contribution in Eq. (2.2), $\sigma_{E,P}$ is associated with the number of created electron-holes pairs. Expecting a number of produced pairs n as E/E_{min}, where E denotes the deposited energy and E_{min} the energy required for individual electron-hole pair production, a Poisson process suggests an uncertainty in the number of pairs of $\sqrt{n}$. Note that at the temperature relevant for GERDA and LEGEND, 87 K as motivated below, $E_{min} = 2.96\,\text{eV}$. For $\mathcal{O}(\text{MeV})$ deposits, we expect very roughly a variance in the number of pairs of 10^3. Multiplied with E_{min} this implies a resolution uncertainty of $\sigma_{E,P} \sim$ keV. However, a pure Poisson fluctuation is not adequate in this case, as the production of the electron-hole pairs is not an independent process, but is caused in a consecutive, correlated manner. A so-called Fano factor F needs to be included to compensate for this correlation, known to be generally $\mathcal{O}(0.1)$, and 0.129 ± 0.003 at the liquid nitrogen temperature of 77 K [90]. The effective number of pairs is then given as $n^{'} = nF$, leading to

$$\sigma_{E,P} = \sqrt{Fn}E_{min} = \sqrt{FE_{min}E} \ , \tag{2.3}$$

which contributes at the 0.1 keV level.

The variation in the charge collected in the detector scales proportionally with the deposited energy as $\sigma_{E,C} = cE$. The linearity has been found empirically, and is motivated by the broadening of a γ-ray peak with a tail-like distortion at the low energy end of the peak, which scales with the presence of impurities. It is hence interpreted as incomplete charge collections. The constant of proportionality has been estimated to be of the order of 10^{-5}-10^{-4} [87]. In GERDA, this contribution was found to be sufficiently low to be considered negligible, whereas during LEGEND commissioning runs, it had been corrected for already at the signal filtering stage, with improvements on the resolution on the order of few 0.1 keV, as we will see in Sec. 3.3. In the physics data taking of LEGEND, similar corrections are performed.

Finally, the electronics component $\sigma_{E,N}$ is caused by noise in the electronics circuit, its compo-

nents, the amplifiers, and the detector itself. The contribution is constant in energy, and cannot be omitted. It can however be reduced to tolerable levels of 1 keV at MeV energies by operating Ge detectors at low temperatures to minimise leakage current and noise from the readout electronics. The use of electronics components with low capacitance and optimal shaping and filtering settings also reduces the noise level. This is particularly important for those electronics components that cannot be operated at low temperature. Overall, the shape of the centroid of a γ-ray peak, in good approximation modelled as a Gaussian profile, will then be widened according to

$$\mathrm{FWHM}(E) = 2\sqrt{2\ln 2}\sqrt{\sigma_{E,N}^2 + FE_{min}E + c^2E^2} \approx 2.355\sqrt{\sigma_{E,N}^2 + FE_{min}E + c^2E^2}\ . \quad (2.4)$$

Here FWHM denotes the full width at half maximum resolution, the prefactor accounts for the conversion from standard deviation to FWHM[1], and the quadratic term in the radicand may or may not be neglected. As mentioned above, in total Ge detectors achieve unparalleled FWHM resolutions at 2 MeV of less than 3 keV. In Sec. 3.3, we will discuss the corrections on the full peak shape with the help of a low-energy tail modelling incomplete charge collections.

2.1.3 Advantages for neutrinoless double beta decay searches

To conclude this general description of Ge semiconductor detectors, let us explicitly state, even at the cost of partial repetitions, the advantages of deploying these in $0\nu\beta\beta$ experiments. As rare-event searches, these measurements, as we have seen in Sec. 1.5, require a maximum amount of the potentially decaying source material, long measurement times, extremely efficient detection capabilities, optimal measurement precision, i.e. energy resolution, and the lowest background levels reasonably achievable. And let us keep in mind all these physical requirements need to be realisable in an actual experiment within reasonable time durations and with limited financial resources. HPGe detector operation is a well-established field, with decades-long experience in rare-event searches. The crystals can be purchased commercially, and the operation conditions, even at cryogenic temperatures, are modest. As the source material operates as detector itself, very high detection efficiencies are reached. The isotopic abundance of ^{76}Ge, naturally around 7.8%, can be enriched up to approximately 90% via centrifugation [91], further contributing to an efficient use of available material. In addition, no internal contaminants have been measured so far, enabling extremely low background conditions [52]. Moreover, the waveforms measured in Ge detectors enables pulse shape discrimination between energy depositions with different event topologies, or in other words different particles and interactions, further reducing the background levels, as demonstrated e.g. in Ref. [92]. We will provide more details on this important feature further below. Furthermore, the energy resolution in HPGe detectors is the best of all currently operated technologies [51]. Considering the combination of all these advantageous experimental features, we conclude that HPGe detectors are highly suitable for $0\nu\beta\beta$-decay searches, contrary to the naive expectation discussed in the previous chapter. We will now discuss which particular Ge diode designs the GERDA and LEGEND programs deployed, or are and will be deploying, respectively, in their experiments.

2.2 GERDA

The Germanium Detector Array, or GERDA, aimed to detect $0\nu\beta\beta$ decay by operating HPGe detectors enriched to approximately 87% in the isotope ^{76}Ge inside a cryostat filled with liquid

[1] The conversion factor can be determined analytically by calculating the distance between the two points at which the Gaussian function takes half of its maximum value.

argon (LAr) [93]. Installed at the underground laboratories of the Istituto Nazionale di Fisica Nucleare (INFN)[2], the Laboratori Nazionale del Gran Sasso (LNGS)[3], GERDA consisted of up to 7 detector strings containing in total 37 enriched detectors, plus 3 of natural composition used for background validation studies [94]. Operating an array of HPGe detectors in two main data-taking phases, the unique approach of GERDA was to directly immerse the bare Ge crystals into the LAr in order to shield the detectors from external radioactivity, additionally using LAr as a coolant. The granular array design also allowed GERDA to reject muon or γ-induced background events if they deposited energy in multiple detector channels simultaneously. The electrons produced in a β or $\beta\beta$ decay instead deposit all their energy at a single, point-like interaction volume of $\mathcal{O}(1\,\mathrm{mm}^3)$ within the bulk of one detector only. This granularity-based background reduction method is hence referred to as detector anti-coincidence cut (AC), or multiplicity one cut (M1). Furthermore, detector-near parts such as nylon mini-shrouds for shielding from particles propagating through the LAr, detector holders or cables and connectors, were carefully selected and screened before they were deployed into the LAr cryostat. The $64\,\mathrm{m}^3$ cryostat was surrounded by an outer tank of $590\,\mathrm{m}^3$ volume filled with ultra-pure water [95], itself covered by layers of rock corresponding to around 3500 metres of water equivalent overburden, reducing the cosmic muon flux by six orders of magnitude compared to surface levels [96]. In fact, as we will see in the following section, the infrastructure is still available onsite. Additional background mitigation was accomplished with the help of an active Cherenkov veto, or muon veto system, consisting of, in total, 64 photomultiplier tubes (PMTs) installed inside the water tank, and a scintillation shield above the experimental facility. The coverage of the water tank led to an overall detection efficiency of the remaining cosmic muons exceeding 99%. The cleanroom, located above the water tank, housed a lock system and a glove box to mount the detectors, as well as other instrumentation, such as the slow control systems used to monitor and regulate the operation conditions. The experiment operated the HPGe detectors in different array configurations during separate data taking phases. During Phase I, running from 2011 to 2013, 21.6 kg yr of valid data exposure was collected. The detectors were of two different types, five of broad energy germanium type (BEGe) with average masses of around 0.7 kg each and average diameters and heights of 73 mm, and 30 mm, respectively, and fourteen of semi-coaxial shape (Coax) with masses of 1 to 3 kg, approximately 60 mm in diameter, and 68 to 108 mm in height [97, 98]. We show a three-dimensional sketch of the different detector types operated in GERDA in Fig. 2.4, where we also indicate the weighting potential introduced in Sec. 2.1.

The main performance feature of the BEGe detectors is their excellent energy resolution of $\sim 3\,\mathrm{keV}$ FWHM at $Q_{\beta\beta}$. In contrast, Coax detectors have an increased surface-to-volume ratio, effectively reducing external background levels normalised to the active volume available to accumulate exposure. The reason is that less materials, such as holders, plastic shield, and cables are required per channel and mass element. As a notable side effect, optimal resolution does not only allow to discriminate the $2\nu\beta\beta$ continuum from the sought-after signal peak, but also enables optimal identification of background γ events due to their different pulse shapes compared to the double electron signal caused by $0\nu\beta\beta$. We will discuss this background mitigation technique, referred to as pulse shape discrimination (PSD) in more detail further below. Measuring a background rate after cuts of 1.1×10^{-2} counts / (keV kg yr) in the region of interest, well approximated as a uniform distribution without any particular features, no signal peak at $Q_{\beta\beta}$ had been identified after unblinding the data of Phase I, and an upper limit on $T_{1/2}$ of $0\nu\beta\beta$ in ^{76}Ge of 2.1×10^{25} yr at 90% CL was found [101]. Before the data taking during Phase II, lasting from 2015 until 2019, a major hardware upgrade was performed [94]. First, additional

[2] https://home.infn.it/it/

[3] https://www.lngs.infn.it/en

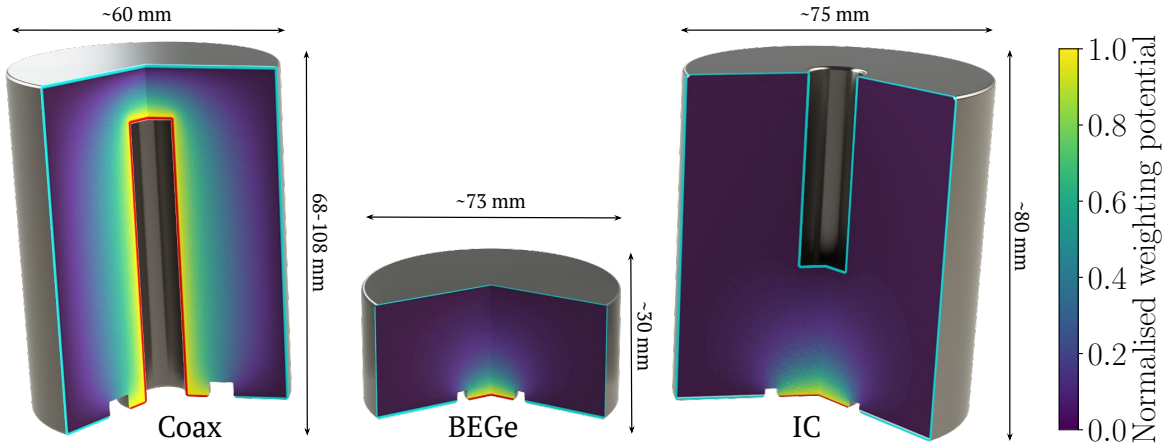

Figure 2.4: Three-dimensional render of the HPGe detector types operated in GERDA. From left to right: semi-coaxial shape (Coax), broad energy germanium (BEGe), and inverted coaxial shape (IC) detectors. Whereas the first two were operated during both main data taking phases of GERDA, the latter were installed at a later stage for the final Phase II+ campaign only. The weighting potential and the average dimensions are shown as well [97, 98, 99]. The red and cyan lines indicate the p^+ and n^+ contacts, respectively. The figure is adapted from Ref. [100] with the help of F. Girard, using SOLIDWORKS[4].

detectors were installed, bringing the total number to 41 detectors overall, 30 BEGe, and 10 of Coax type. Second, an additional active background veto was installed, hereafter referred to as the LAr veto. This hybrid system, consisting of wavelength-shifting fibre curtains and of photomultiplier tubes surrounding the detector array, was another unique component of the GERDA design. The setup allowed the analysts to reject events also depositing energy outside of the detectors, especially those caused upon γ particles escaping a detector after a partial energy deposition. These γ particles, but also muons or external β particles, generate scintillation light in the LAr. As the LAr is itself transparent to its own scintillation light, emitted in the vacuum ultraviolet regime at a wavelength of 128 nm with a bandwidth of around 10 nm [102], the energy deposited outside the HPGe detectors in coincidence with energy depositions within a detector clearly indicates a background event. GERDA collected these external energy depositions with the help of wavelength shifting fibres coated with tetraphenyl butadiene and PMTs and SiPMs instrumented at the upper and lower end of the HPGe detector array. The coating plus the propagation inside the fibres cause a wavelength shift of the collected light into the visible regime, where the readout is most sensitive [103], thus enhancing the overall light yield [104]. As the entire veto chain occurs outside the detector bulk volume, the application of the LAr veto cut did not cause a relevant reduction of the signal efficiency. In fact, it was only dead-time reduced by ∼2% due to random coincidences [47].
In 2018, a minor hardware upgrade was performed, leading to Phase II+. Five inverted coaxial shape detectors (IC) of diameters and heights of around 75 cm and 80 cm replaced the ones of natural composition. These detectors are the state-of-the-art of semiconductor technology, combining the increased surface-to-volume ratio of Coax detectors for external background mitigation with the resolution and pulse shape discrimination performance of BEGe detectors [99]. The latter is accomplished by depositing the boron for p-type doping at a single spot of around

[4] https://www.solidworks.com/

0.1 µm only, see [99]. This p-type point contact approach then also led to an optimised PSD performance, similar to BEGe detectors. We show a render of the IC detectors in Fig. 2.4. A sketch of the experiment is illustrated in Fig. 2.5 top, together with pictures of the detector array and the wavelength-shifting fibres deployed during Phase II (Fig. 2.5 bottom).

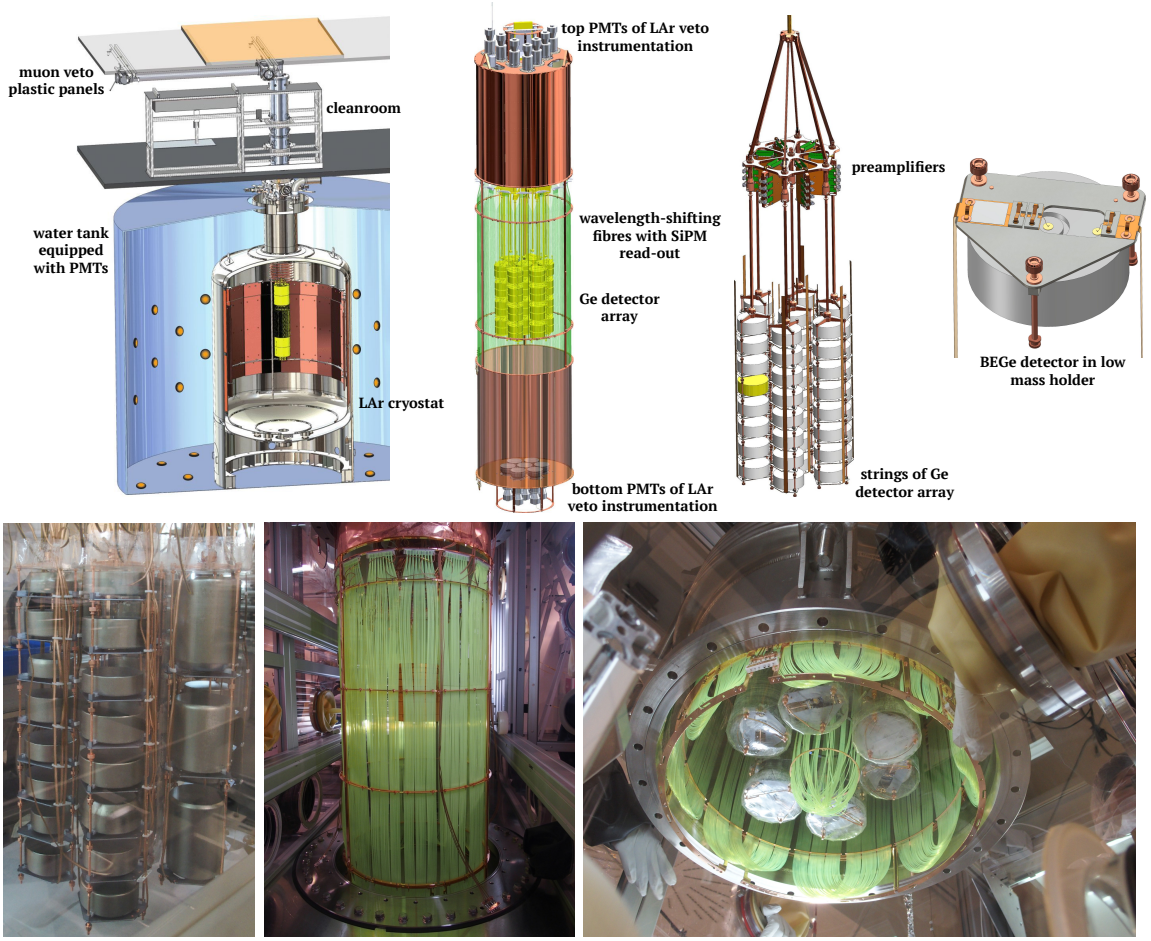

Figure 2.5: Top: sketch of the GERDA experiment, depicting, from left to right, the cryostat infrastructure, the LAr veto instrumentation, the detector array, and a single HPGe diode. Adapted from Ref. [94]. Bottom, left to right: pictures of the Phase II array with the detector strings inside nylon tubes, the fibres of the LAr veto instrumentation, and view into the lock system with the detector array and the LAr veto fibres around. The pictures are taken from the GERDA collaboration.

In all phases, the detectors were calibrated approximately weekly with ^{228}Th sources produced at UZH [87]. This schedule allowed GERDA to continuously monitor the energy scale, to regularly measure the energy resolution, and to frequently tailor and validate the PSD performance of the detectors. We will provide a closer look at the calibrations in Chapter 3. Furthermore, artificial test pulses were injected into the readout chain every 20 s to permanently monitor the detector channel connection, and to evaluate the detector on-times, i.e. the exposures [87].
With the increased exposure and the advanced background reduction techniques, GERDA Phase II was the first experimental operation in the field to achieve the design goal of a measurement in the quasi background-free regime, where the $T_{1/2}$ sensitivity scales linearly with the exposure,

as discussed in Sec. 1.5, [47]. After unblinding the data from all phases, 13 events were observed in the region of interest, 1930 to 2190 keV, which is in good agreement with the background expectation. With only a single event appearing within 2 FWHM around $Q_{\beta\beta}$, no signal evidence had been found above the measured background. With a combined exposure of 127.2 kg yr, a lower limit of $T_{1/2} \leq 1.8 \times 10^{26}$ yr at 90% CL was obtained, coinciding with the sensitivity [47]. The spectrum is shown in Fig. 2.6, where the events in the signal region are highlighted.

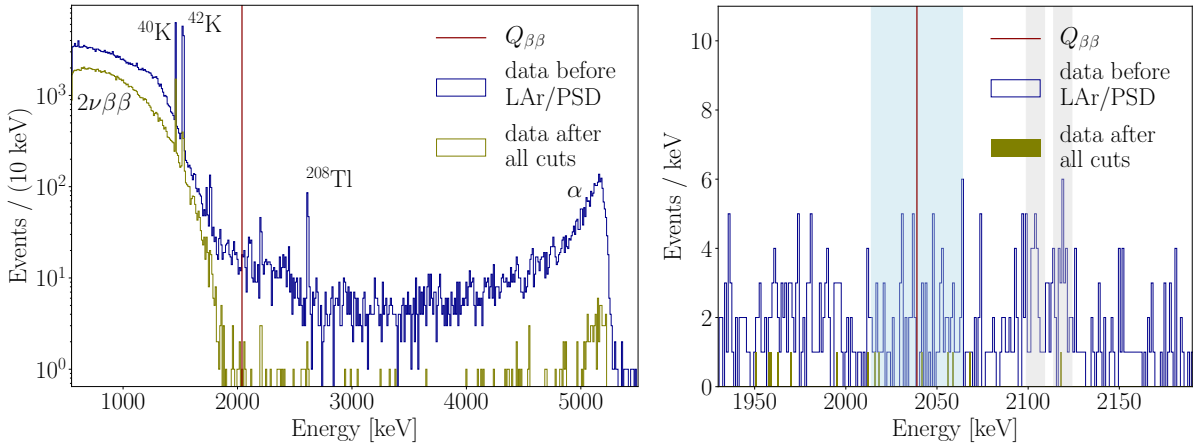

Figure 2.6: Left: measured GERDA energy spectrum before (dark blue) and after (olive) high-level analysis cuts, i.e. after LAr and PSD cuts, with prominent background contributions and $Q_{\beta\beta}$ (dark red) indicated. Right: zoom into the region-of-interest around $Q_{\beta\beta}$. The range in which data was blinded for the analysis is indicated in light blue. The regions excluded from the background evaluation due to the presence of known γ-rays are marked in grey. More details on the data used for the final GERDA analysis can be found in Ref. [47].

The measurement had been the first semiconductor-based result in the field exceeding the 10^{26} yr threshold for $T_{1/2}$. More than three years later, the obtained experimental sensitivity is still world-leading in the field, independently of the isotope. Also, the background level of 5.2×10^{-4} counts / (keV kg yr), again well-described with a flat distribution, poses the lowest value measured in any experiment in the field. Interpreting the result in terms of the neutrino Majorana mass, the obtained upper limit corresponds to a limit on the mass between $m_{\beta\beta} \lesssim$ 79-180 meV, see [47]. This range holds for light Majorana exchange as the dominating mechanism triggering the decay and covers the typical nuclear matrix element ranges discussed in Sec. 1.4. A more detailed description of the experimental instrumentation, of the applied background reduction techniques, and the final analysis can be found in Ref. [94], in Ref. [100], and in Ref. [47], respectively.

2.3 LEGEND

To further increase the sensitivity of HPGe detectors to $0\nu\beta\beta$ decay, the Large Enriched Germanium Experiment for Neutrinoless $\beta\beta$ Decay, or LEGEND, collaboration has been formed as a merger of the two leading $0\nu\beta\beta$ decay ^{76}Ge experiments, namely the MAJORANA DEMONSTRATOR (MJD) and GERDA, plus additional international institutions [105]. With more than 50 institutions involved[5], the ultimate goal of this global project is to surpass a discovery sensitivity

[5] https://legend-exp.org/

of $T_{1/2} \geq 10^{28}$ yr. To reach this goal, LEGEND aims at combining the best strategies of both predecessor experiments. The cornerstones are the direct deployment of HPGe detectors in an active LAr shield like in GERDA, and the use of ultra-low activity detector-near parts and low-noise front-end electronics like in MJD [106, 107]. As the successor of the experiments achieving the best energy resolution (MJD), and the lowest background rates in the field (GERDA), LEGEND has the combined expertise required to achieve the $T_{1/2}$ sensitivities needed to entirely probe the inverted neutrino mass ordering regime. The experiment follows a staged program to keep informing and improving based on available experience and materials. The first phase of the experiment, LEGEND-200 (L-200), is currently in physics data-taking mode, using the existing GERDA infrastructure at LNGS. A total of 142 kg of enriched ^{76}Ge detectors are operated, which will be upgraded to approximately 200 kg in 2024. After five years of operation, sensitivities of $T_{1/2} \geq 10^{27}$ yr can be expected [108]. In the second stage, LEGEND-1000 (L-1000), the final goal will be achieved by an increase of the amount of deployed detector material to at least 1000 kg. In the baseline design, the HPGe detectors will be directly deployed in underground-sourced LAr (UGLAr), depleted in, among others, the isotope ^{42}Ar, which undergoes β decay to ^{42}K. With a suppressed contribution of the daughter isotope decaying via β decay with a Q-value of 3525 MeV [49], i.e. beyond $Q_{\beta\beta}$, the background index will be reduced to a level of 10^{-5} counts / (keV kg yr).

2.3.1 LEGEND-200

In mid 2023, the L-200 experiment operates 101 detector channels in the GERDA cryostat at LNGS, totalling approximately 142 kg of mass distributed over ten strings. The array contains around 70 kg of detector material taken over from MJD and GERDA, with the remainder being new inverted coaxial point contact detectors (ICPC) produced by the two manufacturers, ORTEC[6], Oak Ridge, and Mirion Technologies + Canberra[7], Olen. During this stage, 41 ICPC, 28 BEGe, 6 Coax, and 26 PPC detectors are deployed, where the latter type PPC is the abbreviation for the p-type point contact detectors from MJD. Between late 2023 and middle 2024, approximately 50 kg of additional, new ICPC detector material will be installed to fill the detector array configuration up to twelve strings. With average masses of around 2 kg, implying again an external background reduction optimised from the surface-to-volume ratio, the ICPC detectors will greatly help L-200 to surpass a sensitivity beyond $T_{1/2} \geq 10^{27}$ yr at 90% CL [105] after five years of physics data taking. Assuming again light Majorana mass exchange, this corresponds to $m_{\beta\beta} \lesssim$ 33-71 meV. Having taken over the available infrastructure in early 2020, the collaboration demonstrated promising detector operations during several hardware upgrades and test phases between 2020 and end 2022. In particular, during the Post-GERDA-Test (PGT) in 2020, among others, the front-end electronics and the data acquisition system (DAQ) had been successfully tested with the detectors immersed inside the LAr. On the hardware side, the lock system on top of the cryostat had to be upgraded to accommodate additional space to mount a larger detector array, also requiring to lift the cleanroom roof with the help of a crane. After the PGT, the cryostat was emptied and refilled with newly purified LAr. Deploying an upgraded LAr veto system, combining a freshly coated TPB layer evaporated on new fibre curtains with pure SiPM readout, the γ-tagging efficiency of the LAr veto was greatly improved compared to GERDA. All the front-end electronics and charge-sensitive amplifiers were immersed into the LAr and are now situated next to the detectors or connected to a copper plate around 30 cm above the detector array referred to as bird's nest, due to the large number of cables. Also, the

[6] https://www.ortec-online.com/
[7] https://www.mirion.com/

water tank had been refilled with ultra-pure water, after repairing, or replacing where needed, the operated PMTs of the muon veto. A schematic of the experimental infrastructure, as well as pictures of the current detector array configuration, are shown in Fig. 2.7.

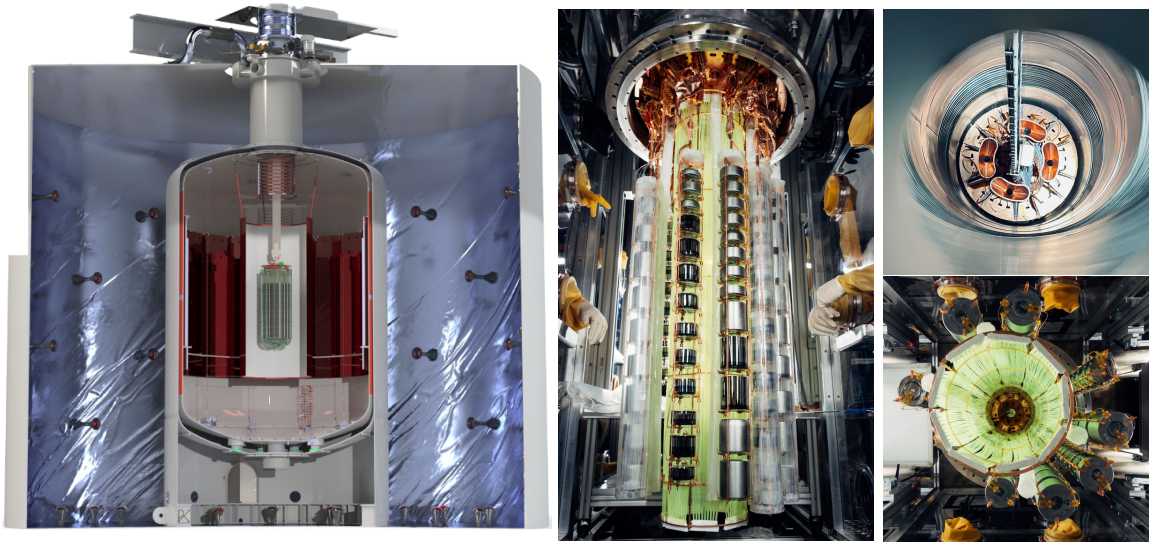

Figure 2.7: Left: render of the LEGEND-200 experiment. The detector array is immersed in a LAr filled cryostat, which is surrounded by a water tank equipped with PMTs. Centre: picture of a part of the detector array inside the glovebox. Right, top: image of the top part of the array when lowered into the empty cryostat. The Cu funnels on top guide the calibration sources down to their desired positions next to the detectors. Right, bottom: view of a part of the detector array and the inner fibre barrel of the LAr veto instrumentation when lifted into the glovebox. All figures and images are collaboration material. Render credit: P. Krause and others, image credit: M. Willers and others.

As an additional upgrade, the source insertion systems developed in our group at UZH was installed and commissioned, allowing for flexible and automised calibration runs, as discussed in Sec. 3.2. These enabled to achieve a further major milestone, running a commissioning phase during summer 2022 in a standard operation scheme, meaning regularly switching between calibrations and physics data taking periods. With 28 ICPC detectors of 60 kg total mass distributed over four strings, this LEGEND-60 (L-60) phase performed the calibrations with ^{228}Th sources approximately weekly, followed by physics run times until the start of a subsequent calibration. Moreover, the corresponding data analysis pipeline was applied successfully. An example spectrum of physics data taken during run 27 of this period is shown in Fig. 2.8.

From the L-60 calibration runs, excellent resolution and PSD performance could be deduced. Measuring ICPC FWHM resolution values at $Q_{\beta\beta}$ of $\lesssim 3$ keV, a measurement precision better than 0.15% was achieved already at the commissioning stage. Optimal resolution and PSD performance will be needed long-term to achieve the background goal of L-200. To exceed the sensitivity of 10^{27} yr, a background level of around 2×10^{-4} counts / (keV kg yr) will be required. This is only a modest background reduction of approximately a factor 2.5 compared to the measured value in GERDA, which is mainly based on the same principles as in the previous experiments. More precisely, these comprise the passive and active background mitigation strategies applied in GERDA and MJD, namely, the different shielding layers and careful material selections, and the muon, AC, and PSD, plus in addition the LAr veto cut as in GERDA. A

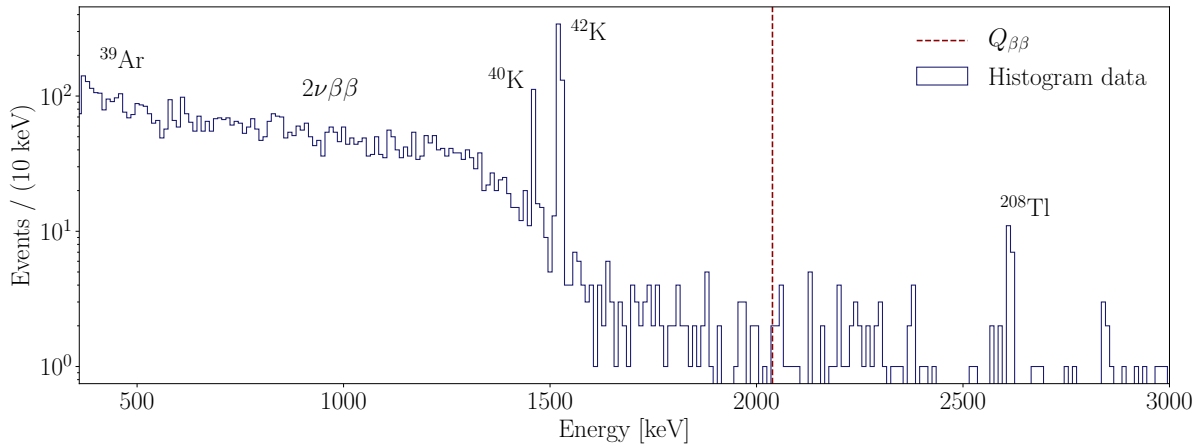

Figure 2.8: Example energy spectrum as measured in one full week of physics data taking during the L-60 commissioning phase in 2022. Identified background contributions are indicated.

sketch of the background reduction strategies is illustrated in Fig. 2.9, left.

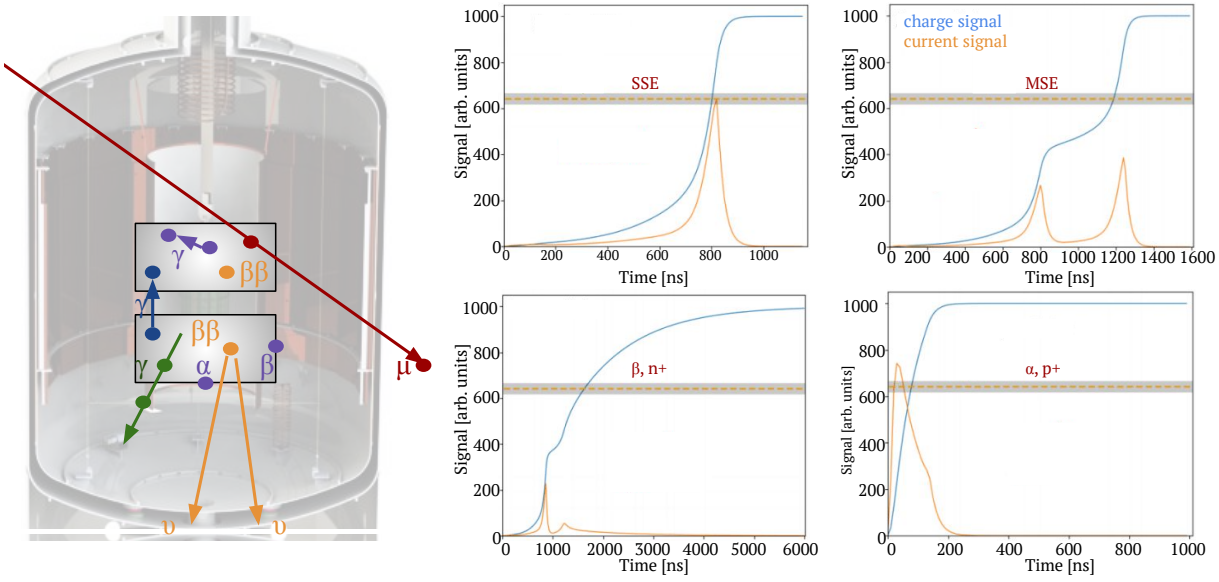

Figure 2.9: Left: illustration of the different topologies of simulated signal and background events for background suppression. Whereas $\beta\beta$ events occur in the bulk volume of one detector, muons, γ-, and surface α- or β particles can be rejected based induced Cherenkov light, coincident events and/or scintillation light, or the pulse shapes of the events, respectively. The illustration is adapted from a render made by P. Krause and other LEGEND collaborators. Right: Examples of simulated charge and current waveforms of a signal-like single-site event (SSE), a multi-site event (MSE), a β event near the n-contact, and an α event near the p-type contact (from left top to bottom right). The signal acceptance region for the current amplitude normalised to energy, which is proportional to the charge, is indicated in grey bands. All four plots are adapted from [109].

Given the larger average detector masses, the low levels of radioactivity of the detector near-parts, and the expected improved average resolution and LAr veto plus PSD cut performance,

the modest background goal seems reasonably achievable. Additionally, several studies are ongoing within the collaboration to further reduce the impact of background contributions to provide more contingency towards the goal. Some of the main background contributions in GERDA were identified to be caused by the primordial decay chains of ^{238}U and ^{232}Th, or ^{42}Ar present in the LAr [110]. Careful material selection, and an improved shielding with nylon shrouds, helps in reducing these contaminants. Others were interpreted as degraded α events. A subdominant contribution came from detector-internal decays of the metastable or ground state isotopes ^{77m}Ge and ^{77}Ge, subsequently summarised as $^{77(m)}$Ge. The α particles only cause relevant depositions if they appear at the detector surface, at the transition between the n- and p-doped layers, leading to partial energy depositions surviving the cuts. Full energy depositions at the p-contact are of too high energy to be relevant, and the thickness of the n-contact is sufficiently large to prevent the α particles from reaching the active volume. The $^{77(m)}$Ge isotopes can be produced via the capture of neutrons, which are themselves produced by cosmic muons penetrating the cryostat. As they undergo β decays with Q-values of 2.7 MeV [49], i.e. beyond $Q_{\beta\beta}$, they can potentially mimic an electron signal in the region of interest.
To understand how these background contributions can be reduced further, it is vital to finally provide more details on the PSD analysis. Energy depositions caused by electrons, like in the $0\nu\beta\beta$ signal case, produce a continuous waveform, i.e. charge or current time profile, as the energy is deposited in a single site in the detector bulk of around 1 mm^3 volume. A 2 MeV γ hit however, potentially causing multiple energy depositions due to a mean free path of around 2 cm [111], will cause edges or peaks in the rising voltage or current signal, respectively. Additionally, α events occurring near the p-type contact of a detector lead to a fast signal rise, and a high induced charge, exceeding the signal expectation. Lastly, an external event occurring near the n-type contact will only lead to incomplete charge collections. All these type of pulse shapes can be distinguished from the clean, continuous waveform pattern of a single-site signal event. Simulated example waveforms for the different event classes are illustrated in Fig. 2.9, right. The PSD classification is usually evaluated as the so-called A/E distribution of an event, where A is the current amplitude, and E the energy, proportional to the charge. The signal acceptance window for this classifier is then defined via a 90% survival fraction for signal-like events collected during calibrations, as detailed in Chapter 3. For BEGe detectors, the resulting signal efficiency in physics data-taking mode was around 88% in GERDA, and for Coax around 69%, mostly due to the limited PSD performance. For IC detectors, the efficiency was around 90% [47]. In L-200, the overall accuracy of the energy estimation, i.e. the detector resolution, is expected to be much better when normalised to overall active mass, since only six Coax detectors are operated. This implies the collective PSD cut performance will be improved, reducing the potential α-induced background contribution. Note that the cleaner detector-near parts already imply reduced radioactivity levels, i.e. a further expected background gain.
To understand the case of the subdominant $^{77(m)}$Ge, we emphasise that the ground state β decay is accompanied by several γ-rays emitted in the de-excitation of the daughter isotope ^{77}As. This means with an improved PSD cut performance, combined with a better LAr scintillation light collection, the generated multi-site events can be identified more efficiently, and thus be suppressed further. However, even with optimised analyses, the contribution from the metastable state ^{77m}Ge will not be reduced further easily, as its β decay to the ground state ^{77}As is not accompanied by γ-emission. Yet, with a half-life of only 54(1) s [112], as opposed to 11.2 h for ^{77}Ge [49], a delayed coincidence method has been developed. If a single-site energy deposition is observed in the region of interest within e.g. 3 min after a muon veto trigger, corresponding to the time range of 90% of ^{77m}Ge β decays, this event may have been caused by a neutron capture in a HPGe detector, and can hence be flagged [113]. Detailed studies are ongoing in the

collaboration to identify the precise time window for the delayed coincidence cut, to ensure the optimum compromise between background sensitivity and minimisation of the induced signal dead-time. The latter may be kept in a reasonable range by applying the cut only to certain detectors such as the fired channel plus its direct neighbouring detectors. Due to a more stringent background level requirement, these studies are particularly important for the second stage of LEGEND, L-1000, to which we now turn our attention.

2.3.2 LEGEND-1000

In the subsequent stage, L-1000, at least 1000 kg of enriched, HPGe detectors will be operated over a full decade. The final goal is to reach a discovery sensitivity beyond $T_{1/2} \geq 10^{28}$ yr, corresponding to $m_{\beta\beta} \lesssim$ 10-20 meV, which fully covers the inverted neutrino mass ordering regime [54]. A background index of 1×10^{-5} counts / (keV kg yr) will be required to reach the target sensitivity, corresponding to a rather ambitious factor of 20 background reduction compared to L-200. To put this goal into perspective, let us underline that this corresponds to expecting less than a single background event in 1 FWHM window around $Q_{\beta\beta}$ when taking data with one tonne of material for ten years! It becomes immediately evident that already a handful of events at $Q_{\beta\beta}$ would pose a very convincing discovery, visible to the eye. We list the main parameter goals of L-200 and L-1000 together with the values obtained in GERDA in Tab. 2.1. A plot of the $T_{1/2}$ discovery sensitivity at the 3σ level as a function of exposure for different background levels is shown in Fig. 2.10, left. Also there we indicate the corresponding values of GERDA, L-200, and L-1000.

Experiment	Exposure [kg yr]	Background level [10^{-4} counts / (keV kg yr)]	$T_{1/2}$ sensitivity [10^{26} yr]	$m_{\beta\beta}$ range [meV]
GERDA	127.2	5.2	≥ 1.8	$\leq$ 79 - 180
LEGEND-200	1000	2	≥ 10	$\leq$ 33 - 71
LEGEND-1000	10000	0.1	≥ 120	$\leq$ 10 - 20

Table 2.1: Overview of the main experimental parameters as achieved in GERDA, and projected for the L-200 and L-1000 experiments. The spread in the $m_{\beta\beta}$ range is caused by the uncertainties in the nuclear matrix elements, cf. Sec. 1.4.

Considering the ambitious program goals, the collaboration actively investigates several potential upgrades on both the hardware and the analysis sides. Most importantly, the main novel feature of L-1000 will be to immerse the Ge detectors in tubes filled with UGLAr instead of natural LAr directly. This operation environment will reduce the ^{39}Ar, and ^{42}Ar background contributions by more than three orders of magnitude, as measured by the DarkSide collaboration [114]. Whereas the former, with a β decay Q-value of 565 keV [49], is mostly relevant for dead-time reduction upon causing scintillation light, or for BSM physics searches in the low energy range, the latter is of utmost importance for the $0\nu\beta\beta$ decay search. As mentioned above, the isotope ^{42}Ar β decays to ^{42}K, which itself undergoes a β decay with a Q-value of 3.5 MeV [49], i.e. above the region of interest. The operation in UGLAr thus strongly suppresses this contribution.
Equally, or maybe even more important, is the plan to deploy ICPC detectors only. As we had seen above, combining optimal resolution, implying pulse shape, performance, with increased masses is a promising avenue to reduce background levels normalised to active mass. The crystal growth processes have been and will be further refined, to produce ICPC detectors of more than

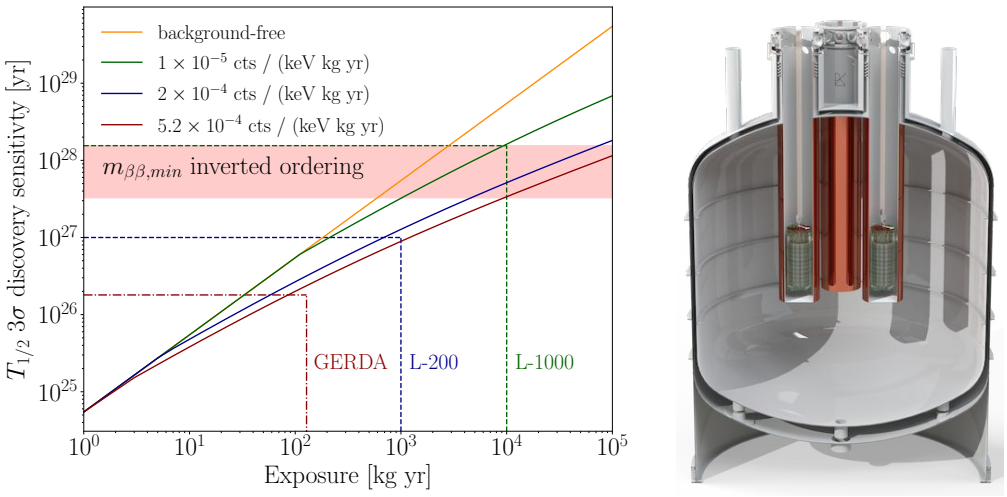

Figure 2.10: Left: projected $3\,\sigma$ discovery sensitivity of L-200 and L-1000 as a function of data exposure, compared to the measured 90% CL sensitivity of GERDA. The scaling of the sensitivity for different background levels normalised to FWHM resolution and exposure are indicated. The red band denotes the approximate minimum $m_{\beta\beta}$ range under the assumption of inverted neutrino mass ordering. The figure assumes the numerical performance parameters published in Ref. [54] and the statistical method outlined in Ref. [39]. Right: sketch of the L-1000 setup at an SNOLAB-like facility, with four detector modules operated inside Cu tubes filled with UGLAr. The latter separate UGLAr from natural LAr contained in the cryostat. The sketch is taken from LEGEND collaboration material, made by P. Krause[8].

double the average mass of the detectors in GERDA and MJD, without the cost of losing the needed cleanliness [115]. Novel schemes on the signal readout chain from front-end electronics to the amplifiers are tested to operate with low-noise conditions also in the low energy regime at keV levels, as was realised in MJD [107]. This would help in rejecting anti-coincident events among different detector channels more effectively. Furthermore, discussions are ongoing on the final location of the experiment. Laboratories deeper underground than LNGS, e.g. SNOLAB[9], or the Sanford Underground Research Facility (SURF)[10], inherit better shielding against cosmic muons [116, 117], and consequently against secondary neutron activation of ^{76}Ge. Yet the mentioned studies of the delayed muon veto coincide cut seem to suggest that even at LNGS depths, the ambitious background goal will be achievable without inducing more than about 10% dead-time [113].
Other active research on the hardware side is focused on novel materials for an improved LAr veto efficiency, such as polyethylene naphthalate (PEN). This active, scintillating plastic is already in use for detector holders in L-200, and could potentially be used for more detector-near parts [118]. Novel designs of nylon shrouds surrounding the detectors are being investigated as well. These latter studies would become particularly important if UGLAr will not be available in the required amount. To reduce this risk, the UGLAr-containing copper tubes itself are planned to be shielded by a cryostat filled with natural LAr. Increasing the scintillation light yield within this volume of natural LAr via doping with a very small, ppm level, xenon concentration is another active research topic [119]. Additionally, simulations are being run to identify the minimum needed dimensions of the cryostat, and to investigate alternative cryostat configura-

[9] https://www.snolab.ca/
[10] https://sanfordlab.org/

tions, e.g. with surrounding vacuum space. Based on the cryostat design studies, and taking into account different hardware constraints at different laboratories, the collaboration has identified two possible configuration schemes. In a SNOLAB-like environment, four equal detector modules of 250 kg each, similar to the L-200 array, would be immersed in four tubes filled with UGLAr. At LNGS instead, a single tube would be filled with around 56 individual, closely packed detector strings, which would boost the detector anti-coincidence cut. Both design options allow for a staged installation without severe interruptions of the physics data taking. In this way, data collection can already be started with available detectors, and subsequently further modules could be added depending on availability. This staged mounting could start around 2031, and last until around 2035. We illustrate a sketch of the four module version in Fig. 2.10, right.
Independently of the experimental layout, also on the analysis side several studies are being pursued to enhance the sensitivity. As an example, the collaboration is exploring the impact of machine-learning algorithms on the data analysis chain at any stage and on the data from different subsystems. These comprise applications such as the basic data cleaning for data quality cuts, PSD analyses of HPGe detector waveforms, or the SiPM triggers of the LAr instrumentation, as partially already applied by MJD and GERDA [120, 100]. MJD also conducted already advanced pulse shape simulations and charge trapping correction algorithms to understand and correct for incomplete charge collections, boosting both the energy resolution and the PSD performance [120], which will be further refined by LEGEND. Cuts removing late charge collections with partial energy depositions, which can be generated if external electrons reach a detector, but do not fully penetrate its active volume, are explored, as done by MJD [53]. Although this list is clearly non-comprehensive, it indicates the active research within the LEGEND collaboration.
To conclude, LEGEND strongly benefits from the experience and success of both the GERDA and MJD experiments in HPGe detector operation and the related analysis. Further aiming to improve the unparalleled energy resolution and background levels of its predecessors, LEGEND is on a promising path towards achieving the ultimate goal of fully covering the inverted mass ordering regime. In the next chapter, we will discuss the details of the calibration procedure in L-200. This process is crucial to understand the energy scale, the resolution, and the PSD performance of the HPGe detectors to perform these extremely sensitive probes of $0\nu\beta\beta$ decay.

Chapter 3

Calibration of the LEGEND-200 Experiment

Given the precision required to detect a $0\nu\beta\beta$ decay signal convincingly, the energy scale and resolution of each operated HPGe detector needs to be calibrated extremely accurately. To monitor the stability of the energy response and the measured resolution, the calibration should be performed frequently. Preferably, only a short duration should be needed to reduce the time lost for the physics data-taking mode. In LEGEND, radioactive sources are thus deployed regularly to irradiate the detector array in a controlled manner. In addition, any background induced during calibration must remain below the appropriate level for the physics analysis. In our case, we refer to radioactive isotope production via the capture of neutrons emitted by the sources. In this chapter, we first introduce the general approach of how detector calibrations are performed in LEGEND. We continue with a description of the hardware used to deploy the sources into the cryostat, and of the testing that we have performed to ensure the long-term operability of the deployment systems. We will also report on the performance of the calibration hardware in terms of the uniformity of the detector array irradiation during the L-60 commissioning phase. We then shift our focus towards the calibration-induced background contribution for the main $0\nu\beta\beta$ analysis, starting from analysing the source characterisation measurements done to determine the neutron flux. This is followed by a detailed study of the Monte Carlo simulations modelling the effects of neutrons emitted inside the cryostat. We will then assess the impact of energy calibrations on the overall background budget for L-200.

3.1 Calibration procedure

3.1.1 Energy scale and resolution determination

Any electronic measurement device must be calibrated to translate a measured electrical signal into the physical quantity under study, such as the number of detected particles, their charge, or their energy. We had motivated the latter, the energy response of a HPGe semiconductor detector, already in Chapter 2. Here we will explain the calibration procedure in our experiments in detail. As the energy range of $0\nu\beta\beta$ decay searches is at the MeV scale, γ-radiation is a powerful tool for the detector calibration. With discrete energies in the keV to MeV range, the detection of γ-rays from well-known isotope transitions provides characteristic imprints in the recorded ADC signals corresponding to energy quanta, as detailed in Sec. 2.1. The strategy for L-200, similar to that of GERDA before it, is thus to position radioactive sources close to the Ge detector array. The deposited energies of the γ particles emitted by the sources are measured

as charge-proportional voltages recorded with the ADC. By making a histogram of the recorded signals, we can identify the known γ-ray locations in the spectrum. We can then easily convert ADC units to energy by applying a linear conversion. We present a complete example in the following section.
Note that in Ge detectors, the energy response is linear to a very good approximation. However, some detectors have a slightly non-linear response, which can be treated via quadratic corrections. We only considered the linear term in our analyses described in Sec. 3.3. The interested reader may refer to Ref. [87] to see how these detector channels have been handled in GERDA.
Having determined the conversion from ADC to, e.g. keV, we can now translate a measured ADC voltage for a given detector to an energy estimate for each energy deposition individually. Considering the importance of a precise understanding of the energy scale and energy response for the measurement of a monoenergetic peak like in the $0\nu\beta\beta$ decay analysis, let us underline that the systematic uncertainties on the energy scale are below approximately 0.2 keV at $Q_{\beta\beta}$ (less than 0.01%) in GERDA across the array [87], proving the remarkable measurement accuracy of Ge detectors.
The role of calibrations does not stop here. Another crucial parameter in any search for a rare monoenergetic event is the measurement resolution, as motivated in Sec. 1.5. We remind the reader of the role of the $2\nu\beta\beta$ continuum potentially fluctuating into the $0\nu\beta\beta$ search region if the resolution is highly limited. Investigating the width of the identified γ-peaks in the now converted energy histogram via a fit of the peak allows us to estimate the broadening induced by the finite measurement resolution. With a linear energy response, an ensemble of fully absorbed γ particles causes a symmetric peak profile which can be modelled as a Gaussian distribution. The standard deviation of the fitted Gaussian σ provides an estimator for the energy resolution, or more commonly for Ge detectors, the full-width at half maximum width estimator (FWHM), which is given as FWHM$= 2\sqrt{2\ln 2}\,\sigma \approx 2.355\,\sigma$ for a Gaussian distribution. Compton scattered events or other backgrounds cause a continuum surrounding a γ-line, and thus need to be fitted in addition to the Gaussian peak. As the resolution in Ge detectors is excellent, we can limit the fit windows to sufficiently small regions around the peak. This means that the fit of the continuum can conveniently be accomplished with a simple linear function. If the Compton scatters cause a strong transition between the regimes above and below the mean, a step function may be added to fit the underlying pedestal. Note that incomplete charge collection in the detectors may alter the peak shape, causing a low tail in the spectrum towards lower energies, and pile-up events may alter the peak shape towards higher energies, where energies are interpreted with respect to the centroid position, i.e. the Gaussian mean. If these features are seen in the data, additional functions need to be added to achieve an accurate fit of the data, e.g. via a tail function. In Sec. 3.3, we will show some example peak fits applied to L-60 commissioning data. The measured standard deviations at each identified line can then be fitted with a square root function as motivated in Sec. 2.1, to interpolate between the measured values to determine the final energy resolution of each detector at $Q_{\beta\beta}$.

3.1.2 Advantages of ^{228}Th

Having understood how we can calibrate the energy scale of the detectors, and how to determine the measurement precision, we are now set to decide which radioactive source material provides optimal properties for the detector calibration. The radioactive material of choice in both GERDA and L-200 is ^{228}Th. Its decay chain leads to multiple γ-rays with energies around 2 MeV. The most relevant lines are induced by the decay of ^{208}Tl, an isotope present in the decay chain of ^{228}Th with a branching ratio of 33%, cf. Fig. 3.1. This γ-radiation allows us to calibrate the

detectors at the energy scale of $Q_{\beta\beta}$.

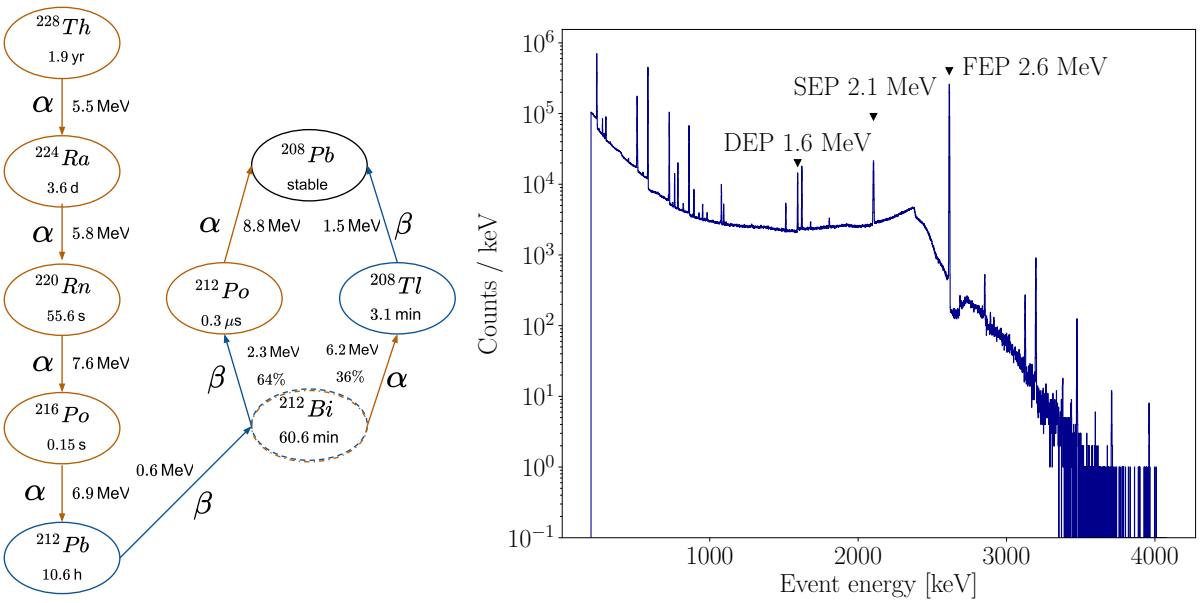

Figure 3.1: Left: schematic of the ^{228}Th decay chain, with α and β transitions indicated in gold and blue, respectively. The γ-rays accompanying the beta decays are not shown, but the double escape peak (DEP), single escape peak (SEP), and full energy peak (FEP) mostly relevant for calibration, are caused by the ^{208}Tl decay to ^{208}Pb. The numerical values are adapted from Ref. [49]. Right: simulated ^{228}Th calibration spectrum as expected to be observed in the L-200 detector array, with the ^{208}Tl DEP, SEP, FEP indicated. The details of the simulation setup are discussed in Sec. 3.5.

With a half-life $T_{1/2} \approx 1.9\,\mathrm{yr}$, ^{228}Th does not completely decay away during the experimental run time of around 5 yr, and therefore does not require frequent source replacements. ^{208}Tl has a γ-transition with an energy of 2614.5 keV, which can create an e^+e^- pair when it interacts with an internal electron, as the energy exceeds the threshold of twice the electron mass m_e. With a mean free path of $\sim$1 mm, the e^- causes a charge deposition immediately, whereas the e^+ annihilates with neighbouring electrons, resulting in two subsequent γ emissions. These may now escape the detector, as at the MeV scale, the mean free path of γ particles in Ge is $\sim$2 cm. In this case, we only measure the initial e^- in a single location, as in the case of a $0\nu\beta\beta$ event. The measured energy is then 1592.5 keV, as an energy of precisely $2m_e$ has been lost due to the escaping γ-pair. This peak is referred to as double-escape peak (DEP). The cases where both secondary γ particles are contained within the detector, or one is contained and one escapes, are referred to as full energy peak (FEP), and single-escape peak (SEP), respectively. The energies of these two are then simply the initial incoming energy of the primary γ, and 2113.5 keV, i.e., a reduction by one electron mass. We show a sketch of the three different interaction patterns in Fig. 3.2.

The three peaks clearly cover $Q_{\beta\beta}$ well from both sides, providing an excellent pattern for the energy calibration at 2 MeV. These three ^{208}Tl peaks provide an additional feature. The DEP peak events deposit their energy at a single location within a detector, and are thus referred to as single-site events, producing unique waveform patterns. These patterns serve as a reference for identifying the waveform signatures of actual $0\nu\beta\beta$ events. The SEP and FEP instead cause multi-site events, as the secondary, non-escaping γ particle, or γ particles in the case of the

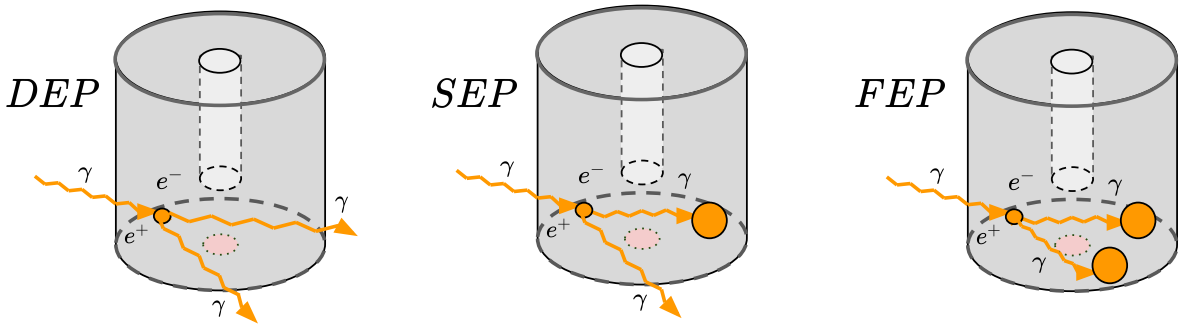

Figure 3.2: Sketches of the event features of γ particles interacting inside an HPGe detector, causing, among partial energy deposits via Compton scattering, the DEP, the SEP, or the FEP. The sketches are adapted from a schematic made by C. Scaffidi.

FEP, propagate away from the initial interaction point within the detector. Such events lead to altered waveforms, as discussed in Sec. 2.1, and thus indicate which kind of waveform patterns background events produce in the signal region. Given the different pulse shape features of the corresponding events in the calibration data, the pulse shape discrimination cut for background suppression in the $0\nu\beta\beta$ search is commonly tailored and validated with calibrations, as discussed in Ref. [100]. We remark, however, that the energies between DEP and $Q_{\beta\beta}$ do not align precisely. Additional bremsstrahlung effects alter the more energetic $0\nu\beta\beta$ events, implying the correspondence is approximate, and not one-to-one. This leads to some systematic uncertainties in the pulse shape analysis, which we do not cover here. Let us mention that charge drift time analyses and simulations can guide the compensation for the bremsstrahlung effects. We refer the interested reader to Ref. [121] for a detailed treatment. As another side remark, the reader may have noted the use of ^{228}Th, and not of its mother isotope ^{232}Th. There are two reasons for this. The first is the activity needed to illuminate the array within a reasonable time frame of order 1 h, without overshooting the DAQ specifications required not to exceed rates of around 300 Hz per detector channel. The constraints can be satisfied with activities of around 5 kBq, as shown in GERDA, and simulated for L-200, as reported in Ref. [122]. With $T_{1/2} \sim \mathcal{O}(10^{10})$ yr, it seems implausible to use the amount of isotopic material needed, compared to the small source containers with dimensions of around a few mm we can deploy with ^{228}Th, as we will discuss in Sec. 3.2. The second reason is that ^{228}Ac appears in the decay chain of ^{232}Th, which causes a FEP in the direct vicinity of the ^{208}Tl DEP, as pointed out in Ref. [1]. This may cause ambiguous energy estimates around the DEP, and consequently, the different waveforms may affect the accuracy of the PSD studies.

To summarise, the three pillars of calibrations in our experiments are, in order of importance, the calibration of the energy response of the detectors, the measurement of the energy resolution, and the PSD cut analysis. In addition, not mentioned yet, calibrations are performed regularly, approximately on a weekly schedule, by immersing the radioactive ^{228}Th sources into the cryostat for typically 2 to 6 h. Following this schedule allows us to monitor the stability of energy scale, resolution, and PSD efficiency, and to compensate for any potential shifts or instabilities within short response times. Let us emphasise that the calibration hardware and software frameworks of GERDA and L-200 are designed to run calibrations as frequently as required without any need for long interruptions of the physics data taking causing relevant loss of physics data. In the following section, we describe the radioactive sources, and the calibration hardware we operate in order to move the sources frequently, and automatically, without manual operations.

3.2 Calibration hardware

3.2.1 General design

At the University of Zurich (UZH), we are the co-leading calibration hardware task group of the LEGEND collaboration, together with the Los Alamos National Laboratories (LANL). Whereas our colleagues produced the sources, as briefly touched in Sec. 3.4, we manufactured, tested, and commissioned five new source insertion systems (SIS). The main performance features these L-200 calibration systems need to deliver, as requested in Ref. [54], are the safe deployment of radioactive material, flexible and simple handling in terms of calibration data taking operations, long-term stability and operability, source position accuracy and repeatability, and mitigation of background contributions for the physics data taking mode. The design of the L-200 SIS is an upgrade of the version developed and constructed for GERDA [123], which has proven to satisfy all the above requirements during a full experimental campaign already. A detailed description of the GERDA version can be found in Ref. [124], with additional information and updates discussed in Refs. [125, 126]. A preliminary design concept was described in Refs.[54, 122]. We provide a schematic of the L-200 SIS, and an illustration of the location and range of movement of the sources in the experiment, in Fig. 3.3.

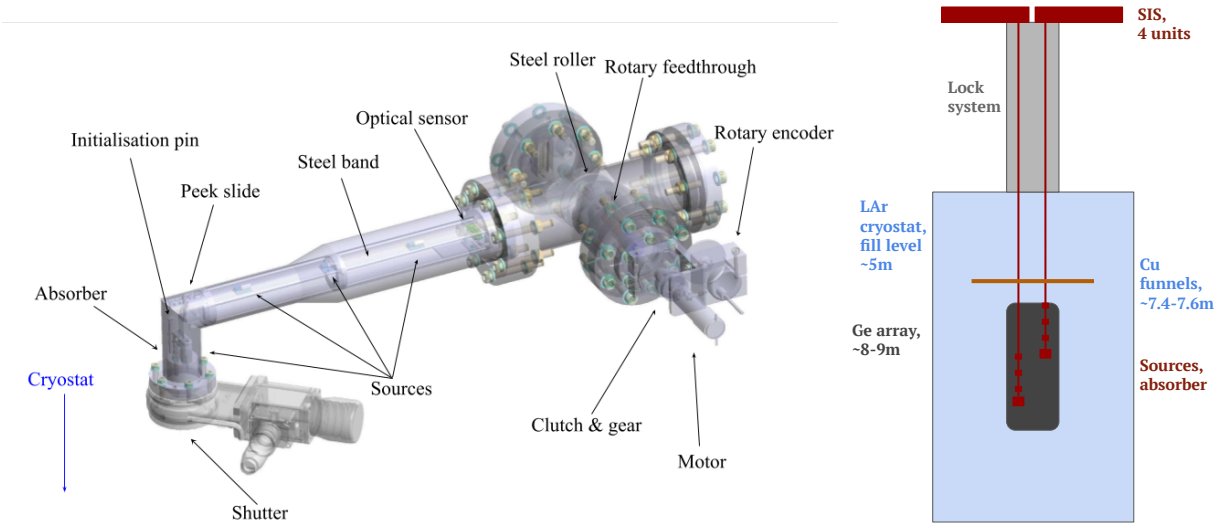

Figure 3.3: Left: drawing of a single L-200 SIS, with the main components labelled. A DC motor is powered to control the movement of a stainless steel band, to which four radioactive sources are attached. An optical sensor and a rotary encoder positioning system monitor the positions of the sources independently. A detailed description of the SIS and its positioning systems is given in the text. Right: illustration of the experimental layout, in which the SIS is mounted on top of the lock system. During calibrations, the sources are lowered into the LAr cryostat, and guided through Cu funnels above the HPGe detector array to their calibrations positions next to the detectors. The indicated dimensions are approximate, and given in the reference frame of the SIS, where the parking position of the sources is defined as 0 m.

The main principle is to power a DC motor to control the movement of a 12.7 mm wide, 11.2 m long, stainless steel band to which a radioactive source and a tantalum absorber are attached. With a motor operation voltage of 24 V, the speed of movement is approximately 10 mm/s. A microcontroller communicates with two position sensors for independent position-

ing control. The main measurement system is an optical sensor system with two photodetectors at a horizontal offset of 0.5 mm tracking the intensity of infrared light emitted by a light-emitting diode (LED) during movement. With equally spaced holes of 2 mm diameter separated by 4 mm, darkness-brightness transitions occur. Due to the offset between the two sensors, the direction of movement can be checked, by detecting rising and falling edges in the light intensity. This system, incrementally counting dark-bright transitions, is called the incremental encoder. The achieved position resolution is 1 mm, as measured with a laser measuring tool using a 6.2 m high test stand at UZH. The secondary positioning sensor is a revolution counter with optical storage, counting 64 multi-turn values, and single-turn angles with a resolution of 264 over 360 degrees. As the turn counts are permanently stored even if the motor is not powered, we refer to this system as the absolute encoder. Knowing the thickness of the stainless steel band, the turn count can be converted to a position with a theoretical accuracy of also approximately 1 mm. Due to the holes in the band, and its imperfect enrollment, the band thickness slightly varies, causing the position accuracy of the second system to worsen to $\mathcal{O}(10\,\text{mm})$. In the following section, we explain how the sensor can be calibrated to achieve a few mm resolution.
Compared to the GERDA version, some major changes were in order given the different configurations of the two experiments. Most importantly, we modified the SIS to accommodate four sources per unit instead of a single source in GERDA, to improve the event homogeneity among the larger L-200 detector array without the need for questionably many stopping positions. Note that the array extends from around positions 7900 mm to 8900 mm in the reference frame of the SIS, indicating that too many stops would be needed. Here the origin is the parking position on top of the lock system, outside the cryostat. Considering the horizontal plane, we installed four SIS onsite at LNGS, with strong support from the LEGEND hardware team, compared to three as in GERDA. The additional system serves as a backup at UZH for training and remote live support in the case of onsite malfunctions. The five SIS can be seen in Fig. 3.4.

Not only the horizontal dimensions of the SIS had to be changed to accommodate three additional sources, but also a vertical height reduction was needed to fit the SIS in between the limited space between cleanroom ceiling and the flange of the lock system on which the SIS are mounted. This was achieved by reducing the height of the tantalum absorber below the bottom-most source from 60 to 37.5 mm. As a third major change, the absorber had to be reduced in diameter from 36 to 32 mm, and equipped with a 3 mm high PTFE plate with rounded edges at its bottom. These adjustments help moving the sources smoothly through copper funnels inside the LAr cryostat, directly above the detector array between the positions 7440 to 7650 mm, into nylon cylinder tubes next to the array. The tubes are needed to guide the sources through the front-end amplifiers, cables, and connectors at the birds' nest without interference. The funnel was needed as the detector array was found to rotate when being deployed into the LAr. Without funnel guidance, the sources do not hit the holes in the bird's nest, and can thus not be positioned next to the detectors. An additional novel feature of the L-200 SIS is that the LED is only powered while the motor turns. This allows us to keep the microcontroller turned on during special calibration runs in which the SiPMs of the LAr veto setup are kept powered on. We summarise the main characteristics of the L-200 SIS and the main design differences compared to the GERDA SIS in the two lists below.

Main characteristics

- Power: custom microcontroller powers a DC motor
- Sources: attached to a stainless steel band, connected to a motor

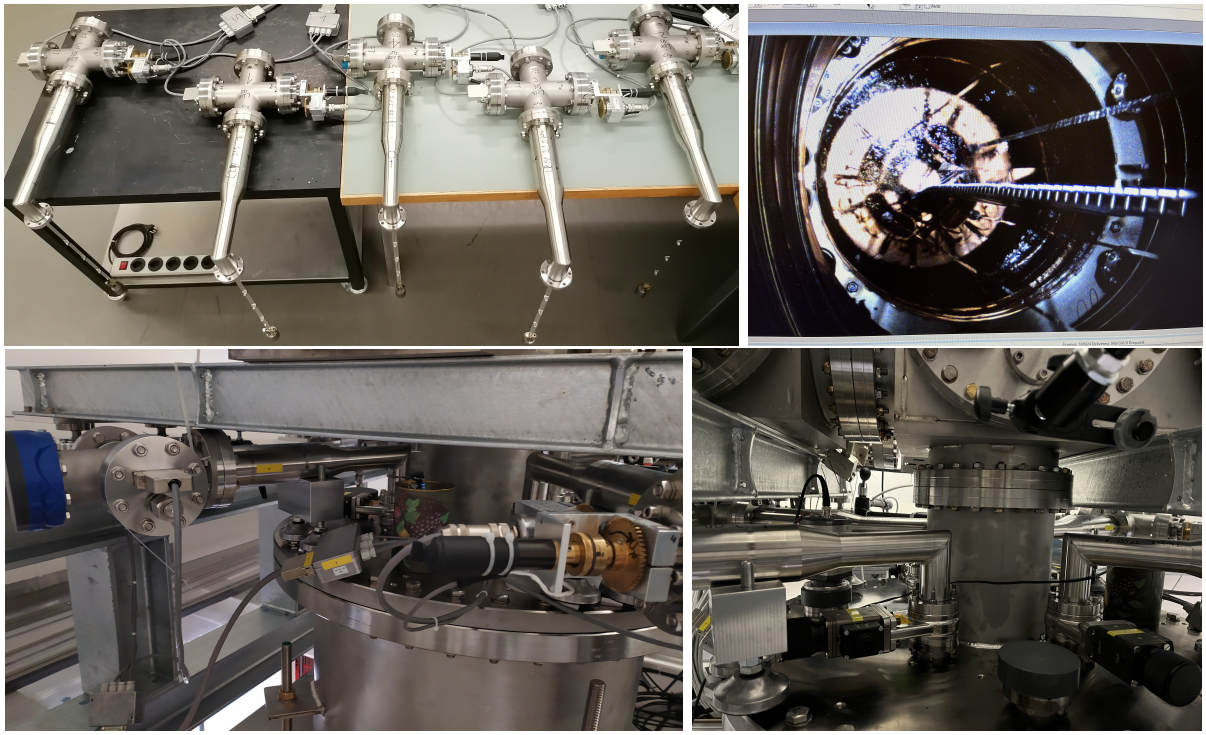

Figure 3.4: Top left: picture of the five L-200 calibration source insertion system after assembly at UZH. Top right: view into the L-200 cryostat when the sources are lowered to the detector array. The photo was taken with a camera installed above the cryostat. Bottom: pictures of the SIS mounted on top of the lock system at LNGS.

- Positioning: optical sensor and rotary encoder systems provide independent measurements
- Location: mounted in cleanroom, sources deployed into cryostat during calibrations

Differences compared to GERDA

- Multi-source approach: four sources per SIS instead of one
- Number of systems: four SIS units instead of three
- Dimensions: longer stainless steel bands; extension of the horizontal tube; shrinking of the vertical part, accompanied by shortening the Ta absorber
- Funnel passage: change of deployment materials and smoothening of deployment parts
- LED powering: optical sensor system only powered while moving

We show pictures of the SIS mounted inside the cleanroom at LNGS, of the nylon tubes next to the array, and of the sources being moved down into the cryostat in Fig. 3.4. In Fig. 3.5, we show images of source capsules in the multi-source configuration, of the absorber with the corresponding Torlon® holders for fixation, the PTFE plate for sliding properties, and a Torlon® pin for positioning system initialisation purposes as described below.

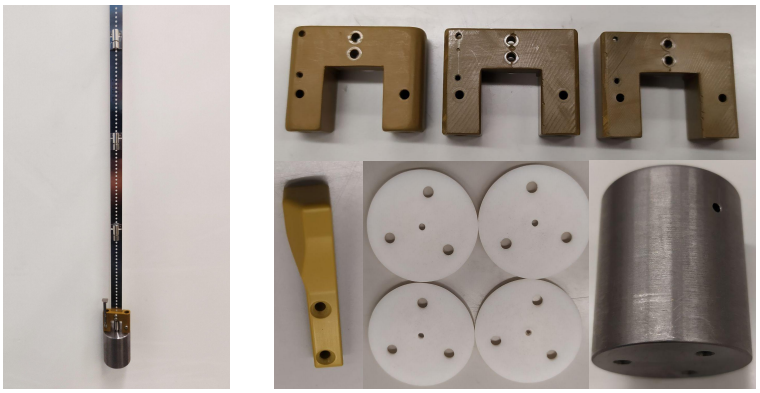

Figure 3.5: Left: photo of four mock source containers. Each SIS system is equipped with four sources, three mounted into spot-welded holders attached to the stainless steel band, and one mounted into the tantalum absorber at the bottom of the band. The picture is adapted from Ref. [1]. Right: pictures of the hardware components lowered into the cryostat; Torlon® holders for attaching the parts to the band (top), a Torlon® pin for system initialisation (bottom left), a PTFE plate at the bottom of the absorber (bottom centre) to improve the sliding through Cu funnels on top of the L-200 array, and the Ta absorber (bottom right).

3.2.2 Details of the hardware features

We now turn to a detailed motivation of how the main requirements that guided the design, testing, assembly, and commissioning are fulfilled.

Safety The radioactive material is contained in welded, sealed stainless steel containers, or source capsules. These are located inside the SIS both during calibrations and physics data taking, without the need for any intervention, i.e., no user needs to get in any contact with the sources, apart from long-term source replacements due to, e.g. activity decrease over time, or special calibrations for LAr veto characterisations. Three capsules per SIS were mounted into M4 tapped holes on source holders, which were resistance spot-welded onto the movable stainless steel band inside the SIS. The fourth source was mounted directly into an M4 tapped hole on top of the Ta absorber, which is attached to the band with the help of a Torlon® holder, i.e., it is also permanently inside the SIS, see Fig. 3.5, left. We append technical drawings of the source containers and the deployment parts in Appendix A. We tested the stability of the new holders under realistic experimental conditions by performing a destructive test of one band with the help of a crane. Before the destruction, we repeatedly immersed the holders with and without mock containers more than 290 times in liquid nitrogen (LN_2), followed by more than 300 iterations moving the holders over a Torlon® plastic slide located in the angle of the SIS tube. For these tests, a brush, a fan, and an air gun were used to reheat the setup when too much ice had accumulated, which cannot occur inside the pressure-tight cryostat, and thus had to be prevented not to alter the conditions in our mock setup. A picture of the test setup is shown in Fig. 3.6, left.

After the tests, which proceeded without any malfunction, the measured weights (converted to kg) at which the crane caused the band to break were measured to be 35, 45, and 50 kg, with a systematic uncertainty of 5 kg each due to the precision of the Newton meter scale connected to the crane. The breaking points were always near the spot-welding points, as indicated in Fig. 3.6. Given the carried load of approximately 600 g in total, dominated by the Ta absorber of 500 g weight, the measured breaking force is sufficiently high to ensure the stability of the

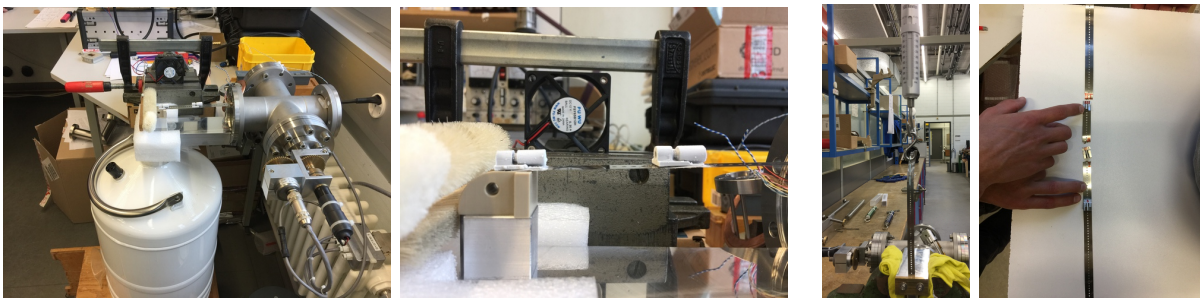

Figure 3.6: Left: photos of the test setup of the long-term stability of the source holders spot-welded onto the steel band under cryogenic conditions, and of the stability of the plastic slide located at the edge of the SIS. The holders were repeatedly immersed into a Dewar filled with LN_2, and air fan, brush, and at times a heat gun were used to remove accummulated ice. Right: pictures of the destructive test using a crane to enforce disruption of the stainless steel band, and of the locations where the band broke upon applying the force with a crane. In three iterations, one for each source holder, the band always broke in the direct vicinity of the spot-welding points, but at forces exceeding the required stability.

hardware components deployed for a safe operation in the real experiment, even in the case of a major, forced stop of motion. Let us remark that during the PGT, see Sec. 2.3, a prototype SIS, relying on the GERDA design but with two additional source holders attached to the steel band, was operated successfully. In the very early phase of the PGT, a stop of the movement of the band was observed, caused by a path interference between SIS and a copper cylinder inside the cryostat. Even in this situation, no damage to the SIS had occurred, proving the stability of the hardware. However, we concluded that installing nylon tubes to guide the SIS absorber may prevent such interferences, which was consequently done for L-200.

Handling The firmware of the microcontrollers powering the motors to rotate the stainless steel bands during deployments runs a custom code developed at UZH. Each controller powers two motors, but in principle, the housing of the microcontroller allows to control three SIS, providing two backup control slots onsite in case of communication errors. Remote communication with the microcontrollers is set up within the L-200 slow control system, via the universal serial bus (USB) and recommended standard 422 (RS-422) communication protocols. The software interface used for the control, C++ and Java based, was developed at the University of Padua (UniPD), relying on a stand-alone Python3-based software we wrote locally at UZH for hardware testing. We underline that the latter code was thoroughly used for testing at UZH and onsite at LNGS during the PGT. In Appendix A, we provide an overview of the working principles of our Python3-based code framework. The slow control software, as well as the slow control data monitoring is accessible remotely. This communication chain allows for flexible calibrations to be run anytime and controlled from anywhere. In case of any malfunction, the system should further indicate the cause. The GERDA SIS firmware, featuring a dedicated number-based encryption system, provided valuable feedback information to identify and consecutively resolve any issue. The details of the coding scheme are described in Ref. [124]. The scheme has proven to work well, and is already incorporated into the microcontroller communication chain, such that all information can be accessed remotely. We hence kept it for L-200 without changes.

Long-term stability Each new, changed, adjusted, or upgraded hardware component of the L-200 SIS was tested carefully at UZH before the final assembly. Apart from testing the new

source holders described above, we tested the side effects of the reduced Ta absorber weight compared to the GERDA version. After assembly, we tested the modified absorber under cryogenic conditions by immersing the stainless steel band into a Dewar filled with LN_2. We investigated the twisting angles and the oscillation amplitudes during the boiling of the liquid upon immersion. Even though the amplitude was enhanced from on average 3.5 cm to 4.9 cm compared to the old absorber version, and the average maximum twisting angle from 175 degree to 188 degree, we did not encounter any forced stop of the movement of the band, or any other malfunction. We also equipped the bottom of the absorber with a PTFE plate to reduce the friction when sliding down the funnel into the nylon cylinder and edge-polished all hardware parts, including the plastic parts attached to the absorber, to prevent stops inside the nylon tube. We tested the sliding on aluminium and copper mock funnels, and on a copper plate with different tilt angles, see Fig. 3.7, to reproduce as many scenarios which could be encountered inside the cryostat as possible.

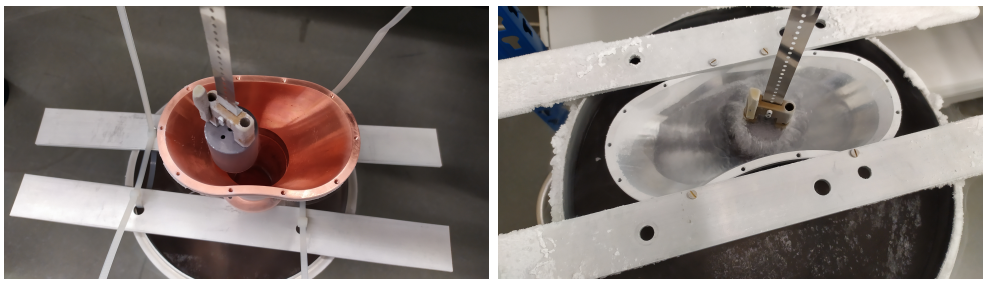

Figure 3.7: Pictures of the SIS movement through different mock funnels meant to guide the sources down to the detector array in L-200. Even down to tilt angles of approximately 25 to 30 degree, the absorber slided well when equipped with plastic plates at the bottom.

After filing and polishing all edges of all immersed parts, the sliding worked without interruptions. At LNGS, the absorber got stuck a few times at the funnels during the commissioning phase, most likely caused by a higher resistance when moving from the nylon tubes into the bottom segment of the copper funnels. Tightening the clutches connecting the motor axis and the mechanical feedthrough from, on average, 3.5 kg to 5.0 kg for the SIS affected by the stops might help to overcome these issues. Not based on the path of the SIS, but on the vertical height reduction, we reduced the initialisation offset from 10 to 4 mm. Initialisation means the band is moved upwards until the Torlon® pin hits an electronic switch, indicating the top-most position. Upon switch activation, the band is moved down 4 mm into its zero, or parking, position to ensure the release of the switch. Measuring the release distance of all switches of each SIS, we found 2 mm as a conservative upper bound, which we doubled to consider potential long-term loosening of tightness. We iteratively initialised a switch in a prototype SIS 5000 times, without obtaining any error. When mounting the SIS, we carefully levelled the inner stainless steel band guidance structure before, and where possible while moving it into the SIS tube, to prevent toroidal deformations. During the assembly we found that such deformations may cause a deflection of the stainless steel band in a near top position, implying the Torlon® pin may not activate the end switch, but fails to hit, which we could resolve with the levelling procedure, see Fig. 3.8.

One of the final systems was tested twice for half the expected L-200 run time, estimated as follows. Assuming a weekly calibration over 5 yr with an approximate deployment distance of almost 10 m, the first test was to run the SIS $0.5 \times 5 \times 52 \times 10 \times 2$ m, i.e., 2.6 km without interruption. The second test was an iterative procedure with multiple stops and changes of

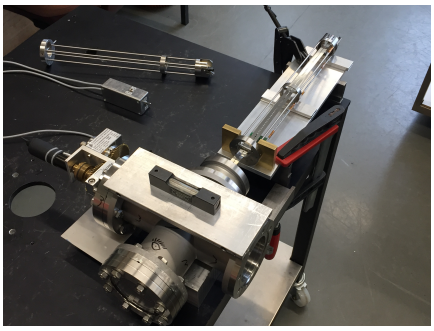
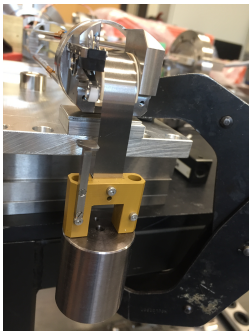

Figure 3.8: Left: picture of the inner structure of the SIS guiding the steel band movement. Right: image of mock deployment parts including the Ta absorber attached to a test band moving over the plastic slide at the edge of the SIS. The inner SIS structure and the end-switch for system initialisation can be seen in the upper left.

direction. Both test did not lead to any unexpected stop of movement or malfunction. This system is now kept as the backup SIS. The four installed systems were tested for an expected 20% of the full cycle, successfully moving these SIS 500 m without stopping. Even in consideration of all the carefulness taken when testing and assembling the SIS, we strongly recommend to routinely monitor and investigate the functionality of the SIS onsite. For example, the unexpected array rotation caused short-term changes to the SIS deployment hardware. Hence, some, at least small-scale malfunctions need to be expected.
We provide a list summarising the performed run tests below.

Long-term tests

- Stability of the source holders in air and in a cryogenic liquid
- Usability of the redesigned Ta absorber in a cryogenic environment
- Sliding behaviour of plastic parts through guiding funnels
- Alignment of the movement of the stainless steel band
- Repeatability of system initialisations
- Life time distance tests, continuous and iterative
- Onsite moving and funnel transition tests

Source positioning As introduced above, the source positions are regulated by two independent sensors, the optical sensor system referred to as incremental encoder, and the revolution counter called absolute encoder. We crosschecked the position accuracy of both sensor systems with a laser measurement device at the 6.2 m high test stand in the assembly hall at UZH. We identified a very stable, long-term position accuracy of the incremental system, with deviations compared to the laser-measured positions of on average 1 mm, which aligns with the systematic uncertainty of the laser position itself. The position deviations of the laser measurement and the absolute positions compared to the incremental positions taken at the test stand at UZH are plotted in Fig. 3.9, left.

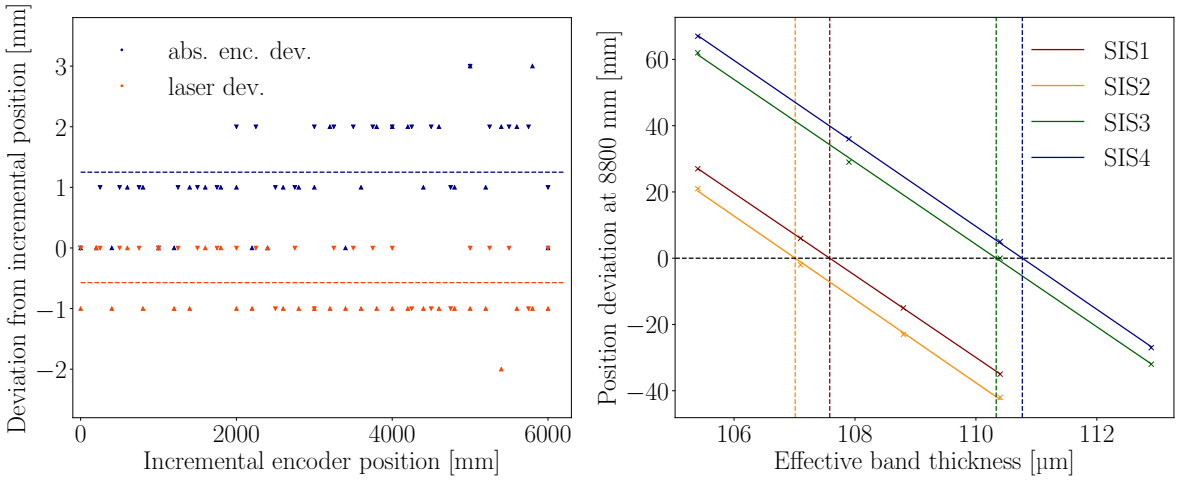

Figure 3.9: Left: comparison of the position deviation between a laser measurement and the incremental position (orange), and between the absolute position after correction and the incremental position (blue), both indicating a reliable source position precision of few mm only on the scale of several meters. Right: deviation between the two position sensor measurements at the maximal deployment position in the cryostat at LNGS, as a function of the effective band thickness parameter assumed for the position determination of the absolute sensor, for all four SIS. We determined the optimal band thickness as the value where the linear fit function of the measured deviations vanishes. In both plots, the position differences are defined as incremental minus absolute or laser-measured positions.

Note that this value corresponds to the intrinsic systematic position resolution, limited by detecting the rising and falling edges in the light intensity curve upon dark-brightness transitions caused by the holes in the steel band during movement. As a side remark, it also corresponds to the systematic uncertainty of the operated laser system at the 10 m scale. We interpret the incremental position value as accurate, and use it as the main position. Note that inside the cryostat, we need to apply a thermal correction due to the temperature gradient between LAr and lock system, which is done via a linear thermal correction term as in the past [125]. The part of the stainless steel band which is immersed in the LAr is expected to be contracted by around 0.27%, meaning a compensation distance of approximately 8 to 11 mm is added to the measured source position during calibrations, depending on the precise absolute position inside the HPGe detector array. The regions with a strong thermal gradient above the surface of the liquid are sufficiently small to be negligible within our precision range of around 1 mm.
The correction induces an additional systematic uncertainty on the position. The main uncertainty then comes from the precise array rotation angle, which causes a horizontal offset of the band when traversing the funnel, and in consequence an unknown vertical offset. We hence use the bottom-most position, visible in a camera view from the top of the lock system via the bending in the steel band, as a reference point to set the source positions with respect to the detector array, with an estimated uncertainty of 4 mm, evaluated as the standard deviation of all measured bottom positions.
The precision of the absolute position system is less accurate. In fact, it even strongly depends on the respective SIS unit. Slightly different steel band thicknesses may cause the deviations, potentially due to inhomogeneities in, e.g. the roll on which the band is enrolled, or in the band itself due to the laser hole punching procedure. To compensate for these effects, we calibrated

the absolute position value with respect to the main position value during the onsite commissioning at LNGS in June 2022. In a first step, the position deviation at the maximum position of 8800 mm was measured for each SIS separately, for four different assumed band thickness parameters, see Fig. 3.10. Note that in autumn of the same year, the maximum position was increased to 8895 mm by extending the nylon tubes below the copper funnels. This marginal enlargement does not alter the absolute encoder calibration. We then determined the band thickness parameter per absolute system for which the deviation vanishes by applying a linear fit, see Fig. 3.9, right. Using the identified parameter, we then monitored the position deviations over the entire range of movement, which we fit with a fifth order polynomial. A polynomial correction value is then applied to each of the 64 multiturn values, independently for each SIS. The achieved position accuracy of each SIS over multiple deployment cycles is shown in Fig. 3.11, where the deviation is plotted with respect to the incremental, i.e., the main, position. As a side remark, let us mention that neither marginally changing the degree of the polynomial, nor expanding the correction scheme from the multiturn-only to a more-refined angle-based correction does not improve the position accuracy relevantly, as the systematic uncertainties mentioned above are already of similar magnitude to the achieved precision. We underline that during the previous long-term testings at UZH, we found that the position discrepancies remained within a few mm over the 2.5 km of the expected L-200 operation length, proving the long-term reliability of the several mm position precision over the $\mathcal{O}(10\,\mathrm{m})$ range. Furthermore, we emphasise that no indications of any band stretching over time were found, as checked by comparing the positions as measured with the incremental encoder to the laser measured positions also *after* the long-term test run at UZH.

Background considerations During physics data-taking mode, the sources are parked outside the cryostat, approximately 7.9 m above the top-most detector to prevent any background contribution from both γ particles emitted in the ^{228}Th decay chain, and neutrons emitted due to (α, n)-reactions. The corresponding solid angle is thus reduced by roughly three orders of magnitude. Stainless steel material from the lock system top flange, the cryostat neck, and the water tank, plus the copper of the bird's nest plate, and around 3 m of LAr, provide additional shielding. Three of the four sources per SIS are parked horizontally, mitigating any direct line-of-sight for γ-rays to reach the detector array. The bottom source, mounted into the absorber, is still shielded by 37.5 mm of Ta, i.e., more than nine times the corresponding mean free path of MeV-γ particles of Ta of around 4 mm [127]. The absorber functions more as a weight to ensure the stable movement of the band than as a needed background shield. It still mitigates any scintillation light triggers in the LAr that otherwise may induce a dead-time from triggering the LAr veto. The probability of neutrons reaching the detectors is highly suppressed, due to the small solid angle from the parking position, and also because of the materials and the LAr between the sources and the detectors. We remark that the background contribution in GERDA Phase I induced by the radioactive sources, which had been parked *inside*, at the top of the cryostat, was estimated to be $\mathcal{O}(10^{-5})$ counts / (keV kg yr). Already in Phase II, and also now in L-200, the enlarged distance, together with the corresponding reduced solid angle, the additional material in between, and the operation of the LAr veto, ensure that the background induced by the sources during physics mode becomes fully irrelevant compared to the L-200 background goal of 2×10^{-4} counts / (keV kg yr). However, the neutrons emitted during calibrations can, upon capture, lead to potentially dangerous radioactive isotope productions. Suppose these isotopes undergo decay processes causing single-site energy depositions with energies beyond $Q_{\beta\beta}$. If the half-lives of these decays are longer than the source removal times after calibrations, potential background events may arise during the physics data taking. We will discuss this scenario in

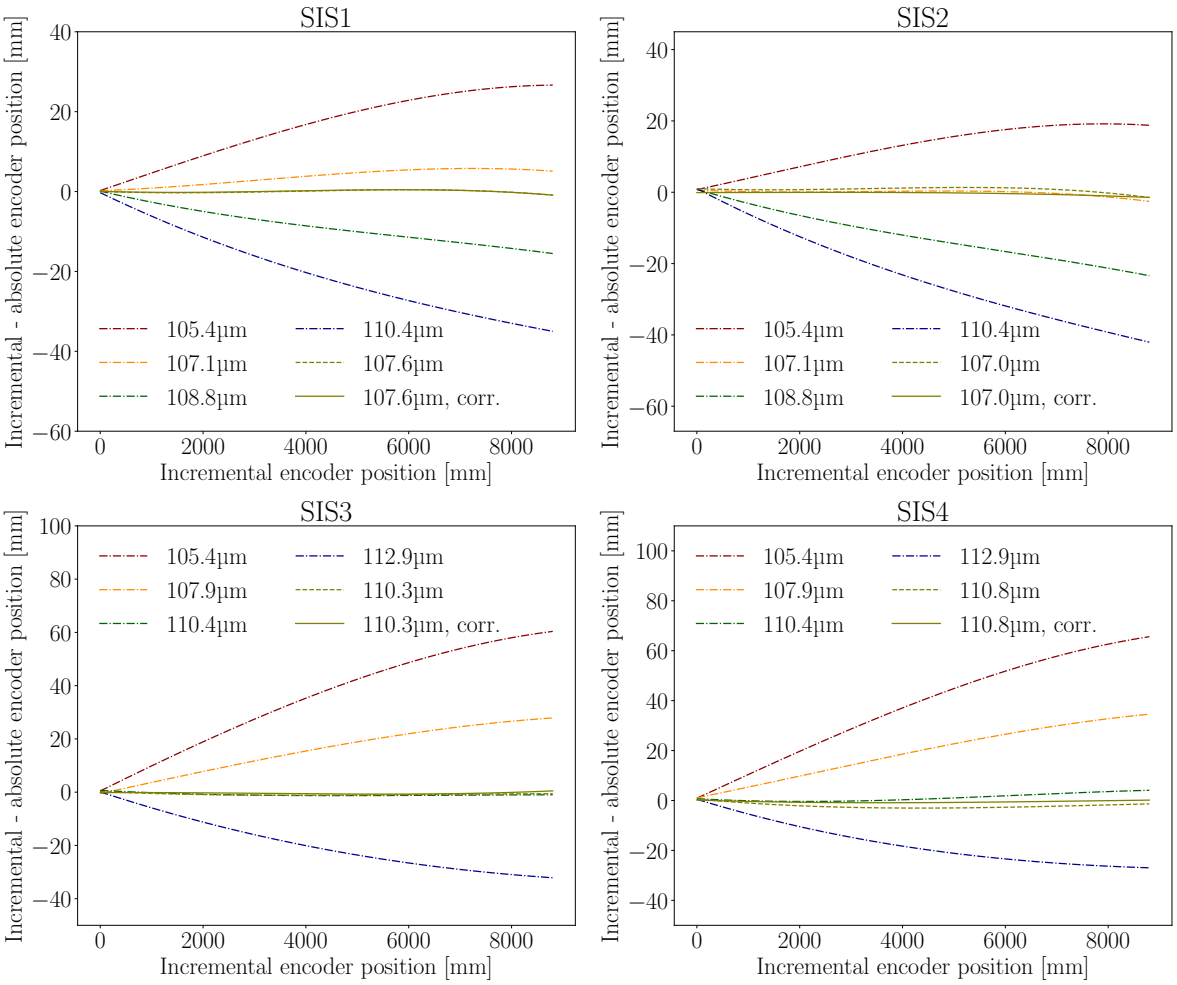

Figure 3.10: Position discrepancy between the two source positions as a function of the main position, as measured with the incremental encoder, during deployment of the sources into the cryostat at LNGS. The different lines represent a polynomal fit of the position deviations as monitored every second for different band thickness parameters assumed in the determination of the absolute position, in correspondence with the measurement shown in Fig. 3.9, right. For the final step of the calibration of the absolute encoder positions, correction values were applied (lines indicated with corr.).

detail in Secs. 3.4 and 3.5, but mention already here that we find a negligible contribution for L-200.

Thoughts on the future L-1000 SIS Let us conclude this section by sharing brief thoughts on the L-1000 SIS design. The general concept of moving multiple sources on a single source string was proven to work. A microcontroller-communication-based powering scheme allows for a flexible and simple calibration procedure. Also, operating two independent positioning systems seem a very useful feature in the case of a malfunction, though the position calibration method may be automated to allow for remote commissioning. Regarding the mechanical situation, more input on the general L-1000 setup is needed for appropriate planning. For instance, if the average detector string length increases in L-1000 or other mechanical constraints, e.g.

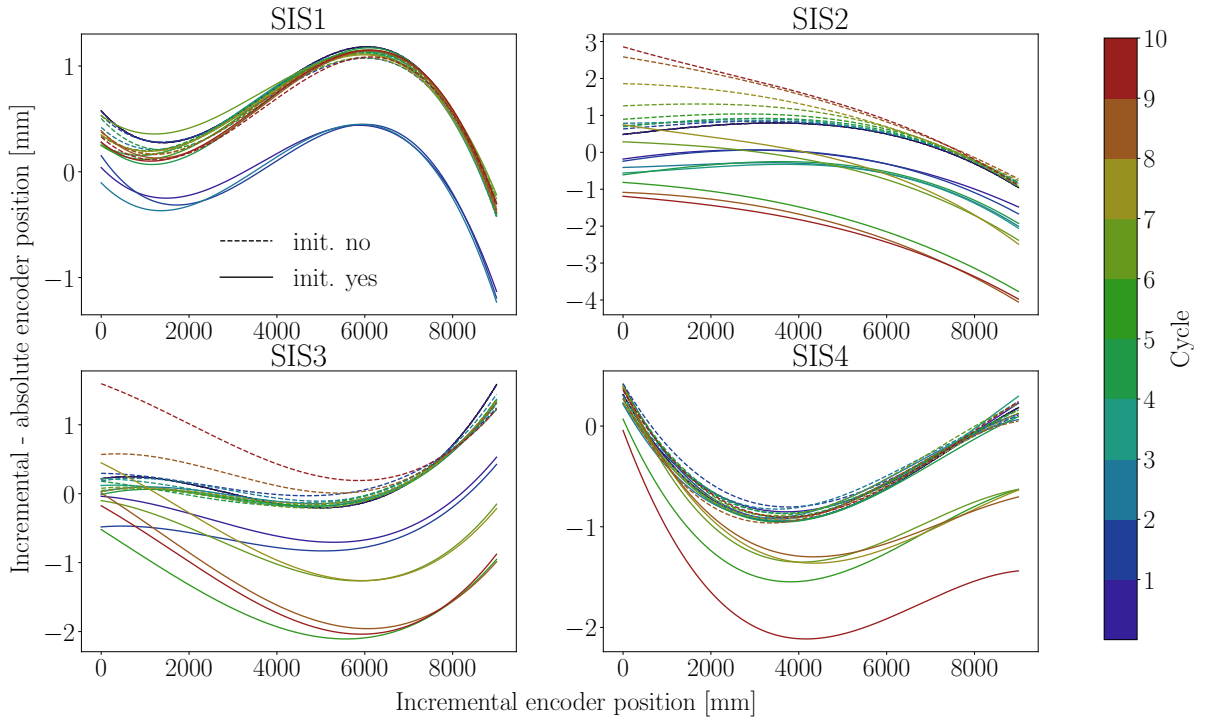

Figure 3.11: Deviation between the source positions determined by the absolute and the incremental positioning systems for each of the four operated SIS. The position deviations over ten deployment cycles with (solid lines) and without (dashed lines) initialisation of the position systems at the parking position are fitted with a polynomial function, indicating position discrepancies of only a few mm.

at the flanges, necessitate design changes, an extension of the SIS to accommodate even more sources will require careful planning of the mechanical sections. Simply extending the horizontal components, inside and outside, does not seem feasible. The outer arm may become too heavy to allow for easy handling, and the inner structure may suffer from strong deformation during the assembly. We suggest redesigning the horizontal sections to ensure better stability in the case of extension.

Regarding the vertical design, the final deployment length is a crucial parameter that has not yet been determined. The current L-200 SIS, with its 11.2 m long stainless steel band, allows for a maximum movement of approximately 10.5 m. The steel roll inside the steel cross could accommodate 12 m of band material. If these dimensions are sufficient, in principle, no changes would need to be made on this side. We recommend redesigning the vertical dimensions, to accommodate a redesigned Ta absorber. Having a thinner version, consequently longer to maintain the same weight needed for the stable movement of the string, provides more contingency in terms of the space inside the cryostat, and inside the detector array, in particular with respect to the detector-near parts such as detector holders or cables.

Additionally, redesigning the initialisation switch to obtain a more azimuthally symmetric configuration would have been a beneficial adjustment from the retrospective. Having the switch positioned at the side instead of in the centre of the vertical part of the SIS tube caused an asymmetric configuration of the deployment parts. Otherwise, the Torlon® pin could have been replaced with a more concentric version, potentially simplifying the transition through the funnels. Alternatively, an induction-based switch instead of a mechanical switch may help to

prevent an asymmetric pin configuration.
On the physics side, careful consideration of the source-induced background is essential due to potentially different array designs and source parking positions during data-taking compared to L-200. The neutron background induced *during calibrations* is very important in this regard. As mentioned above, we will comment on this background contribution in Sec. 3.4 when describing the source characterisation measurements.

3.3 Calibration data analysis

3.3.1 Detector performance

During the L-60 commissioning phase, several test calibrations were conducted. We analysed selected calibration data to investigate two key features of the calibration: the detectors' performance in terms of energy resolution, and the uniformity of the event distribution among detectors and detector strings in the array. We performed the resolution analysis by studying the broadening of the main γ-ray-induced peaks, and the uniformity analysis by comparing the number of hits in the DEP, SEP, and FEP among different detector channels. Using the general LEGEND software stack *pygama*[1], an available Python3 code was used to identify the locations of the γ-peaks of known energy in each detector's ADC spectrum. We then applied a custom code –not developed for the primary analysis chain of the collaboration– to fit a linear response function of the peak positions to determine the energy scale in physical units. We focused on the data of six calibration runs, of which two runs each were representative for three different schemes. In runs 012 and 014, all 4 SIS were deployed to move all 16 sources to 4 stop positions, $(8200, 8400, 8600, 8800)$ mm, with approximately 30 min duration each. In runs 022 and 023, 13 sources were used. SIS1 had only been equipped with one source, as tests of the LAr veto performance required a lower activity. In addition, the positioning scheme was changed to 3 positions, $(8200, 8500, 8800)$ mm, with 2 h duration each (4 h at 8500 mm in run 023). Finally, in runs 025 and 026, the same scheme of three positions with 2 h duration each was used, but a charge trapping correction algorithm had been applied by the analysis group of the collaboration at the data processing stage.
Investigating the drift time of the charge upon an energy deposition, meaning the time difference from the start of the rise until reaching the maximum of a waveform (Sec. 2.1), helps to increase the separation between events with full charge collection and incomplete charge collection. The latter can be caused by residual impurities, or crystal lattice dislocations [128]. Technically, the discrimination between both classes of events was simplified by linearly scaling the estimated uncorrected energy in ADC units as $E_{\text{corr}} = E_{\text{uncorr}}(1 + \delta t\, \alpha)$, where δt is the drift time estimated with the help of the trapezoidal filter (Sec. 2.1), and α is the parameter adjusted to correct for the trapping of charges. Subsequently, the relative energy resolution at the FEP peak was minimised as a function of α. The optimal charge trapping parameter α_{opt} was then applied to all energies in ADC units in the spectrum. Using this algorithm, the energy resolution was improved, as we will see below. More details of the principles of charge trapping corrections are described in Ref. [128]. Note that the analysis described in the following paragraphs directly used the energy estimates in ADC units as provided by the analysis group of LEGEND, with or without a previous application of the charge trapping correction algorithm, depending on the considered run. We compare the different properties of the runs selected for our study in Tab. 3.1. A full example of a peak identification in the spectrum is shown in Fig. 3.12, together with an energy scale fit.

[1] `https://github.com/legend-exp/pygama`

Runs	Number of sources	Stop positions [mm]	Duration per position [h]	Charge trapping correction
012, 014	16	8200, 8400, 8600, 8800	0.5	no
022, 023	13	8200, 8500, 8800	2	no
025, 026	13	8200, 8500, 8800	2	yes

Table 3.1: Overview of the properties of the commissioning runs selected for the detector performance and event distribution uniformity study. Note that the stop positions are stated in the reference frame of the SIS.

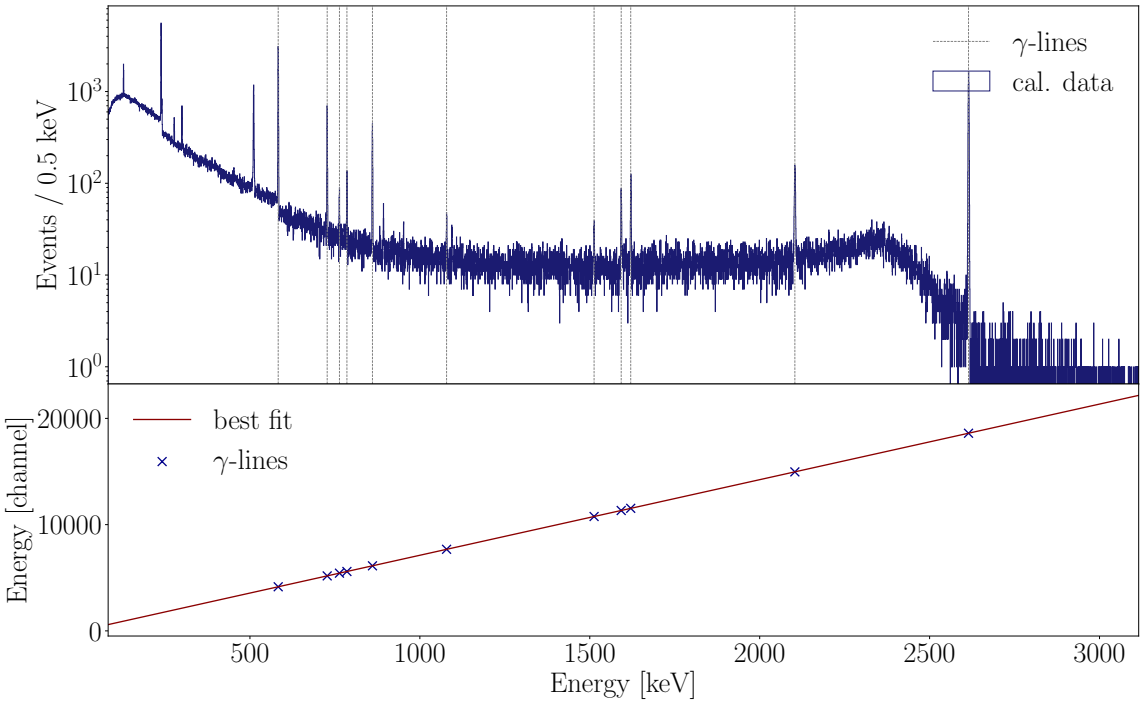

Figure 3.12: Top: calibrated energy spectrum for the detector channel V02160B during the L-60 commissioning calibration data taking run 012. The dashed grey lines indicate the detected peaks. Bottom: linear calibration curve for the corresponding detector as determined by a linear fit of the mean positions of the detected peaks.

Based on the fits of each identified γ-line in the calibrated spectra, we precisely determined the count strength amplitude and the resolution of each peak. We modelled the peak shape as a crystalball peak, located above of a background pedestal modelled as a linear function. The crystallball function consists of a Gaussian profile combined with a power-law tail, which allows for fitting the full energy events in the peak and those events in the low energy peak caused by imperfect charge collections simultaneously. The power-law tail is constrained not to exceed the FWHM level, and thus does not alter the resolution estimate. To consider a potential step in the continuum caused by Compton scattering, we also used a step function parametrised as a complementary error function, with mean and standard deviation aligned with the parameters of the Gaussian part of the crystallball peak. The event number in the actual peak can then be evaluated as the symmetrised amplitude of the Gaussian component, meaning no events from the background affect the measured hit count. We remark that the counting method based on a

full fit of the peak shape leads to estimates of the real signal event number with only marginal systematic uncertainties, caused by the fit uncertainties. Commonly used counting methods in a signal window surrounding the peak, even if taking into account background subtraction from a side-band estimate, cannot distinguish accurately between full energy absorption and incomplete charge collections. In such a standard analysis, the relevant number of available signal-like waveforms for the PSD analysis can thus not be counted as precisely as with the method presented here.
To investigate the detector performance, we interpolate between all measured resolution values at the different energies within each detector's spectrum to obtain the energy resolution at $Q_{\beta\beta}$. As motivated in Sec. 2.1, we fit a square root function of the form $\sqrt{a + bE}$ to the energy E, where a and b are the fit parameters. We give an example of such a resolution curve in Fig. 3.13.

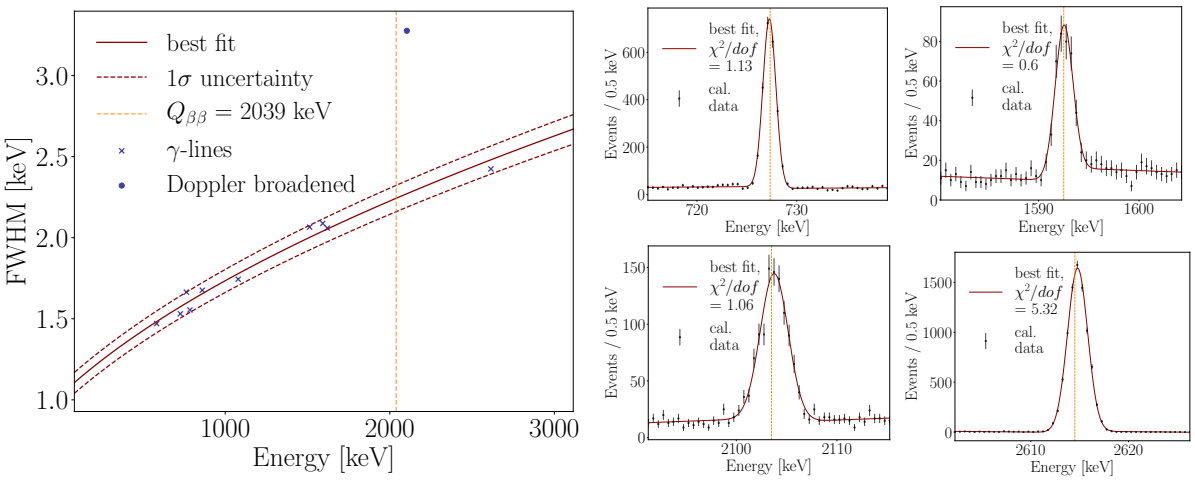

Figure 3.13: Left: energy resolution at FWHM for the detector channel V02160B during the L-60 commissioning calibration data taking run 012. The best-fit curve and its 1 σ uncertainty (dark red) are based on the square root of a linear function in energy fitted to the widths of the identified γ-peaks from Fig. 3.12 (blue data points). The dashed golden line indicates $Q_{\beta\beta}$. Note that we did not include the width of the SEP in the fit due to the non-negligible Doppler broadening of around 1 keV, caused by the loss of energy information if only one of the two secondary γ particles escapes. Right: examples of selected fits of γ-ray-induced peaks for the same detector and run, using a crystalball function to model the peak, and a linear plus step function to constrain the Compton continuum, respectively.

In Fig. 3.14, we show the individual and exposure-weighted energy resolutions in selected L-60 calibration runs. We are interested in exposure-weighted estimators, as detectors of larger mass and longer on-time contribute more to the overall exposure for the $0\nu\beta\beta$ analysis, i.e., their resolution has an increased impact on the overall performance of the experiment. Given their advantageous surface-to-volume ratio, these detectors may also achieve lower background levels. From Fig. 3.14, we deduce excellent energy resolutions at $Q_{\beta\beta}$. We achieved exposure-weighted averaged FWHM resolution values below 3 keV, without any visible performance loss correlated with larger detector masses. In particular, with increasing experience, i.e., run number in the plot, the overall resolution improved. This is due to the gained experience in optimising the detector operation conditions, meaning the optimal voltage settings, and the application of charge trapping correction algorithms.

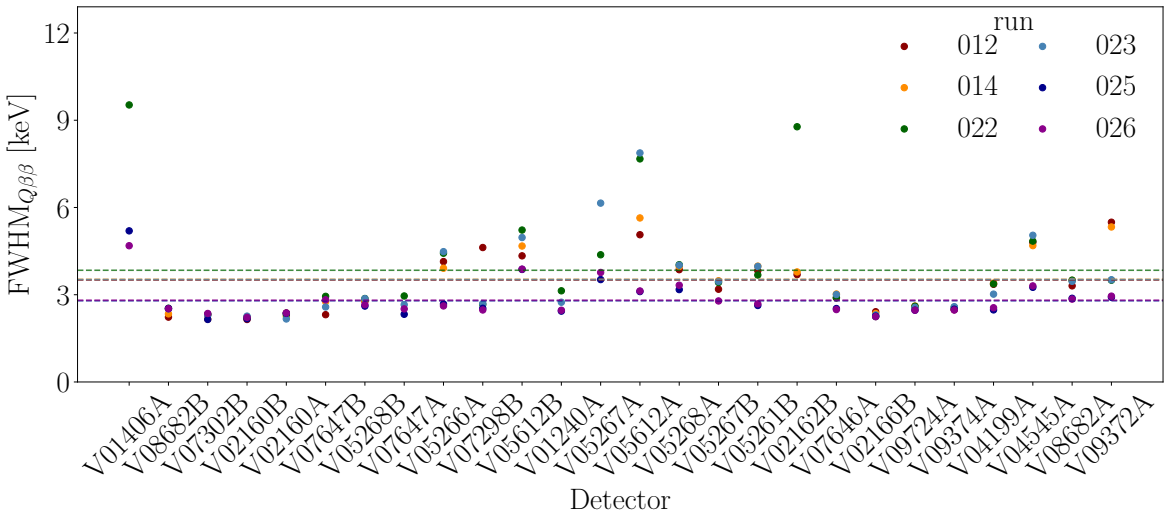

Figure 3.14: The FWHM resolution at $Q_{\beta\beta}$ for the individual HPGe detector channels for selected calibration runs with different operation settings during the L-60 commissioning phase. The detectors are ordered by increasing mass, indicating no performance loss for higher-mass detectors. The dashed lines show the exposure-weighted average per run. Excellent detector performances, with energy resolution below 3 keV, were achieved.

3.3.2 Event uniformity

Regarding the homogeneity of the illumination, we checked the event rates in the DEP, SEP, and FEP across the detectors. Most importantly, we note that it is possible to calibrate detectors over the entire array within a few hours, allowing for frequent calibrations without long interruptions of the physics data-taking. As indicated in Fig. 3.15, the exposure-weighted rates are also relatively uniformly distributed in the individual detectors, where relatively is to be interpreted as a qualitative statement.

We also statistically analysed the uniformity of the event counts in the three peaks of interest, to strengthen the qualitative observations. We binned the count rates in the detectors' DEP, SEP, and FEP for the six different runs. In the case of uniformity, statistical fluctuations would cause a symmetric histogram centred at the mean rate. Looking at Fig. 3.16, we readily see that the distributions are mostly asymmetric.

We calculated the respective unbiased mean μ, standard deviation σ, skewness γ, and Fisher's kurtosis κ, to quantify the moments of the distributions. We further performed a Shapiro-Wilk test [129], evaluating the corresponding p-value p_{SW} under the assumption of normality. This test statistic is more powerful in identifying deviations from the normal distribution than comparable non-parametric statistical tests [130], making it a suitable tool for our investigations. The numerical results of these estimators are listed in Tab. 3.2, indicating an acceptable but far from the optimal level of homogeneity.

The observed rate reduction and enlarged variance for the later runs can be easily attributed to the three missing sources in SIS1. We also studied the event uniformity among groups of detectors, more precisely among different detector strings and vertical detector positions. The evaluated rates are plotted in Fig. 3.17.

Interestingly, we do not only see the already identified pattern of a reduction for later runs, but also slightly reduced event rates in the detectors at the top and the bottom of the strings,

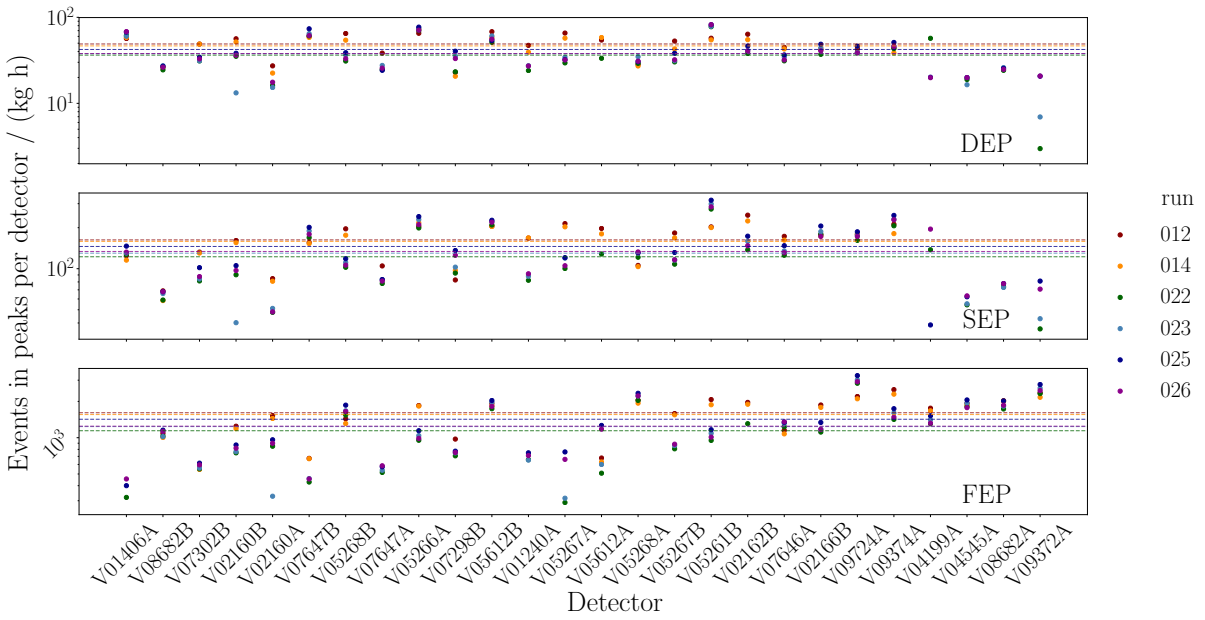

Figure 3.15: Event rates in the DEP, SEP, and FEP when normalised to detector exposure for the HPGe detectors deployed during L-60. The dashed lines indicate the exposure-weighted average per run.

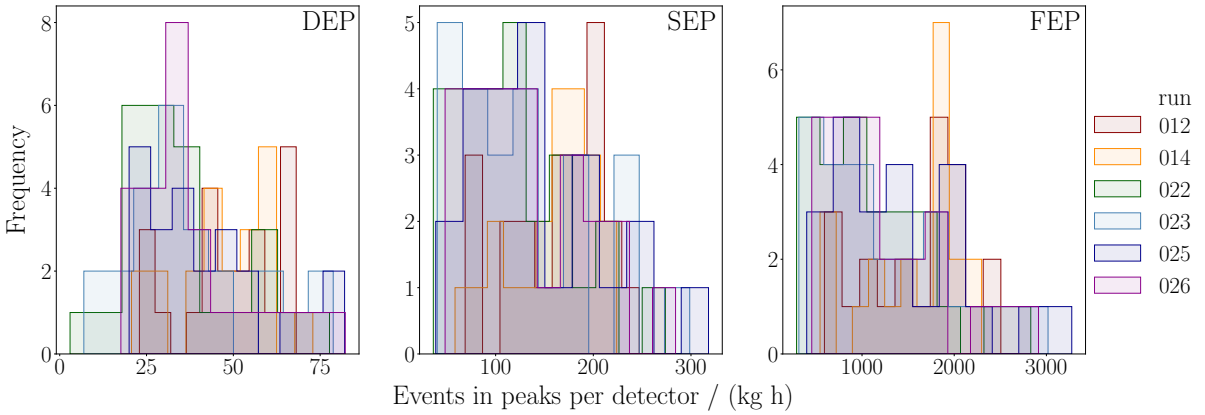

Figure 3.16: Histogram of the exposure-normalised event rates in the DEP (left), SEP (middle), and FEP (right) plotted in Fig. 3.15. The distributions indicate a deviation from normality for all three peaks.

equally for all positioning schemes. For the topmost channels we recommend extending the calibration duration compared to the middle positions. However, the reduction towards lower detectors was caused by reducing the number of detector strings from the planned 14 down to 12 to accommodate space for new, wider detectors. It resulted in an extension of the detector string lengths, surpassing the length of the nylon tubes guiding the calibration sources, i.e., the sources are not sufficiently low to irradiate these detectors. To compensate for the observed lower event rates, the nylon tubes were extended in autumn 2022, now allowing for more homogeneous illumination of the array even for the lowest detectors. The difference in the rates between the strings is caused by differing radial positions with respect to the source locations, implying

Estimator	Run 012	Run 014	Run 022	Run 023	Run 025	Run 026
	49.0	46.5	36.4	37.8	42.3	37.9
μ	162.7	158.6	122.4	130.0	145.6	133.8
	1610.8	1556.2	1142.6	1236.9	1422.5	1246.1
	13.6	13.4	17.1	18.2	17.9	16.8
σ	48.2	44.3	56.5	67.4	68.2	59.2
	547.5	512.1	633.0	722.5	722.5	632.6
	-0.48	-0.43	0.74	0.65	0.84	1.22
γ	-0.49	-0.71	0.79	0.74	0.70	0.76
	-0.45	-0.68	0.87	0.80	0.86	0.98
	-0.82	-0.31	0.38	0.03	-0.09	0.80
κ	-0.73	-0.21	0.40	-0.06	0.11	-0.02
	-0.83	-0.65	0.42	-0.00	0.35	0.49
	0.14	0.17	0.09	0.18	0.05	0.00
p_{SW}	0.19	0.19	0.26	0.16	0.40	0.17
	0.21	0.07	0.13	0.11	0.16	0.05

Table 3.2: Numerical values for the statistical estimators characterising the homogeneity of the exposure-weighted rates seen in the DEP (top), SEP (middle), and FEP (bottom) for different L-60 calibration runs. Assuming optimal conditions, the means should be large, the standard deviations, the skewnesses, and the kurtoses vanishingly small, and the p-values close to unity.

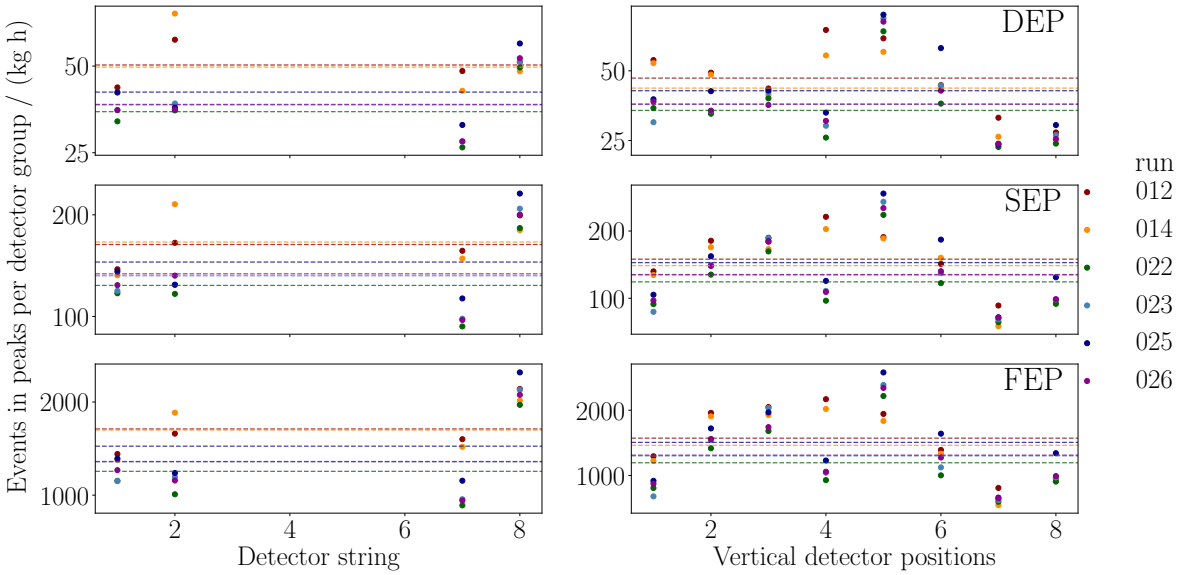

Figure 3.17: Event rates in the DEP, SEP, and FEP when normalised to exposure for grouped detectors, i.e., for detector strings (left) or detector positions within a string (right). Note that an equal position in a string does not immediately imply an equal physical position, as the detectors have different heights. The dashed lines indicate the exposure-weighted average per run.

different propagation lengths of the γ particles through the LAr and, thus, different event rates. The geometry of the flange where the SIS are mounted causes this effect, making it difficult

to counteract. For the following reason a certain level of inhomogeneity needs to be expected independently of any geometrical effect, positioning scheme, or source configuration. Even in the case of optimal homogeneity, the rates in the DEP and SEP are reduced for larger detectors, whereas the situation is flipped for the FEP. Larger detectors will be hit by the γ-rays more frequently, providing a larger volume to contain the secondary γ-pair. In contrast, for small detectors, the two γ particles escape more often, implying the FEP rates are reduced. It is important to note that the only single-site-like events are DEP events, while SEP and FEP events cause multi-site scatters. It may thus be advantageous to position the strings with *larger* detectors azimuthally closer to a flange where a SIS is mounted, in contradiction with the common intuition.

To conclude our event distribution study, we emphasise that the measured count rates validate the approach of deploying multiple calibration sources vertically distributed along the stainless steel band of a SIS, as described in Sec. 3.2. Nevertheless, the irradiation scheme does not lead to perfect event homogeneity, which we cannot expect to achieve. We suggest again monitoring the event homogeneity during later L-200 calibration runs to at least slightly improve the event uniformity, particularly in the DEP. Given the different string configurations between L-60 and L-200, such a study should be performed when all detectors are installed. Improving the uniformity of the detector irradiation by adjusting stop positions and positioning schemes may reduce dead-times during calibrations and the overall calibration duration. These studies may also inform the design of the L-1000 SIS and array configuration, and the corresponding source positioning scheme. While accumulating data in the present phase of L-200 with around 142 kg of HPGe detectors (as of summer 2023), we encourage successive analysts to conduct preliminary studies of the detector performance and the event uniformity for the same reasons. As a motivation, we used our code framework developed above to compare the achieved FWHM resolutions at $Q_{\beta\beta}$ for the four detector types operated in L-200, namely ICPC, BEGe, Coax, and PPC detectors (Secs. 2.2, 2.3), in Fig. 3.18

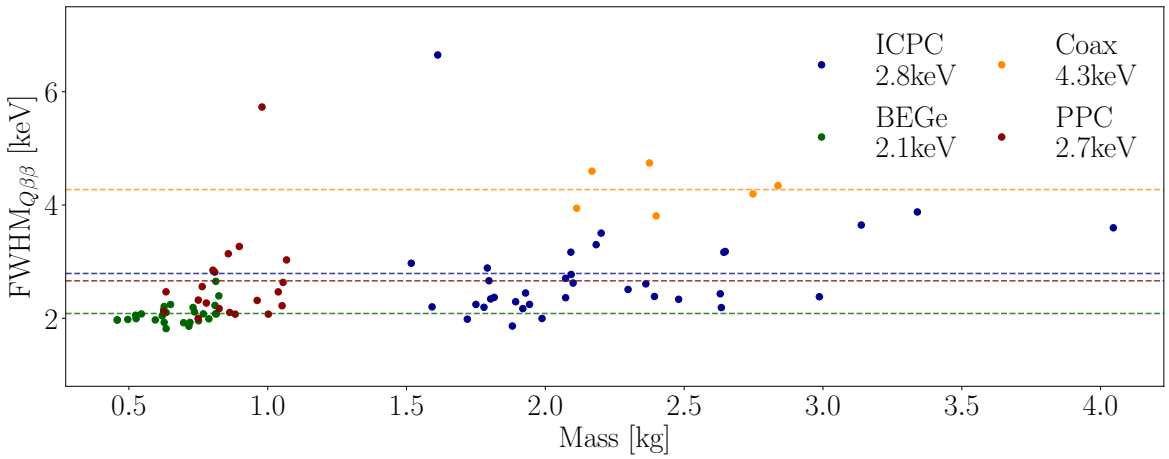

Figure 3.18: The FWHM resolution at $Q_{\beta\beta}$ as a function of the detector mass for the individual HPGe detector channels for calibration run r001 of period p06, performed in June 2023, with the different detector types indicated by the colour scheme. The corresponding dashed lines show the exposure-weighted average within each type, with the corresponding numerical values included in the plot legend. Neither a strong performance degradation with increasing mass, nor for the ICPC detectors is observed.

.

Here we focused on one example calibration run, r001 of the data taking period p06, performed recently in June 2023. We neither identified a performance degradation for the ICPC detectors compared to other detector types, nor for increasing detector masses. This poses promising information for L-1000, which will only deploy medium- to large-mass ICPC detectors. In addition, we display the exposure-weighted event rates for the DEP, SEP, and FEP for the different detector types in Fig. 3.19, indicating an imperfect, but acceptable level of uniformity, similarly to the outcome of our study for L-60.

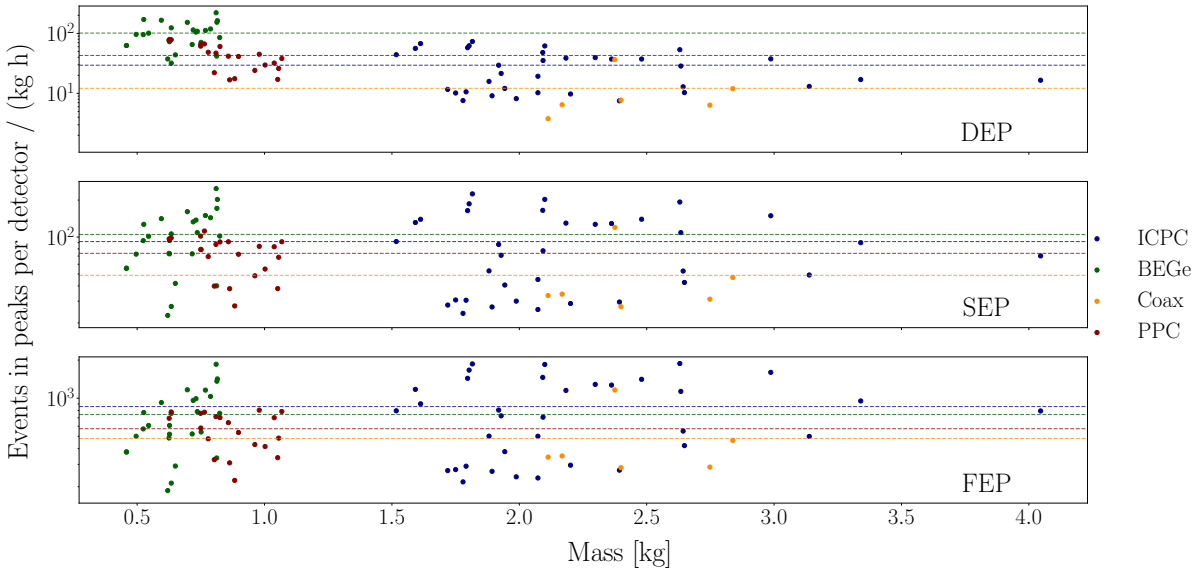

Figure 3.19: Event rates in the DEP, SEP, and FEP when normalised to detector exposure for the HPGe detectors of the L-200 calibration r001 of period p06. The colours show the deployed detector types, and the dashed lines indicate the exposure-weighted average within each type. A detailed study of the event distributions for L-200 calibrations is left to future work.

We leave the detailed statistical investigation of both the detector performance and the event distribution over extended periods, e.g. for current and upcoming calibration runs, to future work.

3.4 Source characterisation

3.4.1 Source production and characterisation

Following the GERDA I and II source design developed at UZH, [124, 131, 132], the ^{228}Th calibration sources for L-200 were produced by our collaborators at LANL. Simulations reported in Ref. [122] revealed that an average source activity of 5 kBq provides sufficient statistics within the capabilities of the DAQ and the readout electronics systems of L-200. At LANL, radioactive Th material from Eckert and Ziegler[2], dissolved as thorium-chloride in a hydrogen chloride solution [1], was electrodeposited onto 50 µm thick gold foils. The rolled foils were subsequently filled into 17 stainless steel containers, 16 to equip the four SIS onsite, and one to serve as a backup source. After welding the containers, the sealing of the capsules was tested successfully in both LN_2 and heated water. Material inhomogeneities during deployment causing slight

[2] www.ezag.com

variations in the activity compared to the desired value. Hence, the specific activity of each source was measured by our collaborators at LANL, using α counting and γ-ray spectrometry. With the former, an average activity of 4.3 kBq, with a standard deviation of 2.4 kBq, was obtained in September 2020. The latter measurement, conducted in February 2021, yielded an average activity of 4.4 kBq, again with a standard deviation of 2.4 kBq, confirming the first result. The details of the production process, welding tests, activity determination, and activity A_i of each source i, including uncertainties, can be found in Ref. [1].
The deposition of the radioactive material onto gold foil is motivated by a particular purpose, namely to reduce the neutron flux emitted by the ^{228}Th sources. These neutrons are caused mainly by (α, n) reactions. Gold has an (α, n)-production energy threshold of 9.94 MeV, which exceeds the ^{228}Th decay chain's maximal α-decay energy of 8.8 MeV, and thus mitigates these neutron emissions [131]. Remaining neutrons are produced by (α, n) reactions on other materials or potential impurities. In Ref. [84], it was reported that the neutron rate of the GERDA I and II sources, also embedded in gold foil, was reduced by approximately one order of magnitude compared to commercial sources. These results align well with what we found for the L-200 sources [1], as we will detail in the following section. Let us underline that a low neutron flux is a crucial parameter which impacts the $0\nu\beta\beta$ decay search, as the neutron capture inside the HPGe detectors may cause the production of ^{77}Ge. Its ground state beta decay to ^{77}As has a Q-value of 2.7 MeV, i.e., above $Q_{\beta\beta}$, and a half-life of approximately 11.3 h, i.e., much longer than the calibration source removal time of around 14 min. Hence, these decays pose a risk of causing background events potentially indistinguishable from $0\nu\beta\beta$ events inside the region-of-interest during the physics data taking mode following a calibration. A detailed analysis of this contribution, and also of other potentially dangerous neutron-induced background contributors, is given in Sec. 3.5.

3.4.2 Neutron emission

Considering the importance of a low background level, any part deployed near the HPGe detectors must be screened beforehand to understand its potential background contribution. Of course, the same holds for the radioactive sources due to the potential activation of ^{76}Ge by emitted neutrons. We published the description of the neutron screening measurement already in Ref. [1], together with colleagues from LANL. Here, we closely follow our text as published together with the other authors of the corresponding paper, but provide additional information on the preparation of the detector measurements and on the calculations conducted to determine the neutron flux.

Preparation of the neutron measurement

We performed a full measurement campaign to measure the neutron flux of the ^{228}Th sources [1], starting from the detector calibration, the detection efficiency determination, and a background measurement. We used a low-background LiI(Eu) detector system from SCIONIX[3] underground in the Gator facility [133] at LNGS, operated by our group at UZH. The LiI(Eu) detector houses a Li crystal enriched to 96% in ^{6}Li, and a modified R8250 PMT from Hamamatsu[4], which is connected to a high-voltage (HV) supply. The HV supply and a multi-channel analyser DAQ system from ORTEC[5] are located outside a detector shield made of 200 mm thick borated

[3] https://scionix.nl/
[4] https://www.hamamatsu.com/eu/en.html
[5] https://www.ortec-online.com/products/electronics/multichannel-analyzers-mca

polyethylen walls, with a fraction of natural boron of around 5% [124]. The modifications made to the PMT were described in Ref. [84]. A PVC source holder, with 20 source threads, is located around 80 mm from the detector, separated with a Pb block of 20 mm thickness to stop γ particles, and with five PE disks of 10 mm each to moderate the emitted neutrons. We tested different moderation thicknesses to find the best thermalisation, i.e., detection efficiency, for this thickness. Thicker shields attenuate too many neutrons, and a lower PE width does not slow down neutrons sufficiently to enhance the neutron capture cross section, keeping in mind that the cross section scales inversely with the neutron velocity v. A detailed study of the optimal moderation width was reported in Ref. [124]. We show a schematic image of the detection system and a photo of the PVC holder with mounted 17 calibration sources in Fig. 3.20.

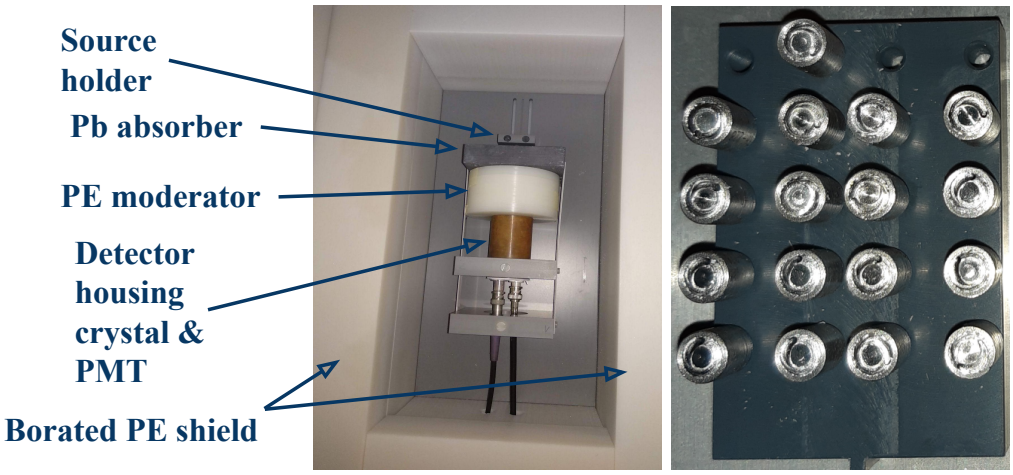

Figure 3.20: Left: schematic image of the LiI(Eu) detector setup used to measure the n flux emitted by the ^{228}Th sources. Incoming neutrons cause a ^{6}Li isotope to split into an α particle and a tritium ion, which then excite surrounding atoms. A PMT inside the crystal housing (copper cylinder in the middle) detects secondary de-excitation. Right: picture of the source holder with 17 calibration sources mounted for the neutron flux measurement. Figure adapted from Ref. [1].

The detection principle works as follows. When an impeding neutron is captured inside the crystal, a ^{6}Li isotope is activated to form ^{7}Li, which splits into an α particle and a tritium ion. The Q-value of the process is 4.78 MeV, shared between the final state ions. The full reaction thus reads as follows,

$$^{6}_{3}\mathrm{Li} \; + \; \mathrm{n} \; \rightarrow \; ^{7}_{3}\mathrm{Li} \; \rightarrow \; ^{3}_{1}\mathrm{H} \; + \; ^{4}_{2}\mathrm{He} \; + \; 4.78 \text{ MeV}. \tag{3.1}$$

Subsequently, the accelerated ions excite neighbouring atoms, which then de-excite via photon emission, with a maximum intensity at a wavelength of 475 nm [134]. This scintillation light of the secondary processes is then collected by the PMT, operated at 750 V, similarly to measurements presented in Ref. [131].
We tested different voltage settings and found sufficient gain and resolution performance for this value, as discussed in the following. The tests were done with an ^{241}Am^{9}Be neutron source with an activity of (160 ± 4) neutrons / s, measured in 2013 [135], and with an ^{152}Eu γ source. First we took preliminary data with the neutron source at different HV settings of 750, 775, 800, 850 V. We filtered incoming voltage rises from PMT signals with a trapezoidal filter as described in Sec. 2.1, fixing both the rise time of the edge and the flat-top width of the trapezoid

to 2 µs. In the next step, we fitted the neutron peak in each resulting channel spectrum with a binned fit of a Gaussian distribution, parallelly constraining the background continuum with a linear function and a double-step function modelling the Compton continuum. We used a complementary error function in combination with a logistic curve to describe the double-step function. This allowed us to estimate the rate in the neutron peak R as the Gaussian amplitude over the measurement duration, and the normalised resolution as the standard deviation of the Gaussian σ divided by its mean μ. The performance parameter p_{opt} as an estimator for the optimal operation settings could then be determined as the maximum among of the product of the efficiency and the inverse normalised resolution, i.e., $p_{\text{opt}} = R \times \mu/\sigma$. In Tab. 3.3, we provide an overview these preliminary measurements.

HV setting [V]	Measurement duration [d]	Peak rate R [cts / s]	Normalised resolution σ/μ	Performance parameter p_{opt} [cts / s]
750	5.9	0.072	0.094	0.77
775	4.5	0.0615	0.092	0.67
800	2.0	0.0389	0.083	0.47
850	1.0	no peak visible	no peak visible	none

Table 3.3: Overview of the preliminary HV testings used to estimate the gain as a measure of the efficiency and of the neutron peak resolution. The optimal HV setting of 750 V was chosen in correspondence with the maximum of the performance parameter, defined as the product of the peak detection efficiency and the inverse normalised resolution. Note that at 850 V, the neutron events were widely spread, and consequently no clear peak was visible.

Obtaining 750 V as our optimal setting, we measured 5 d with the γ source to calibrate the energy response of the detector accurately. We fitted the identified, visible γ-ray peaks in the resulting channel spectrum with the same fit function as above. We then fitted the means of the Gaussian profiles with a linear curve to convert ADC units into physical energy estimates, as plotted in Fig. 3.21. We obtained a goodness-of-fit (gof) estimate of $\chi^2/dof = 1.74$, where dof denotes the number of degrees-of-freedom in this fit, indicating good linearity of the detector response. We show the spectrum in calibrated energy in Fig. 3.21, top. From a similar fit of the neutron peak, obtained with a prolonged 17 d measurement with the neutron source, see Fig. 3.21, we deduced a precise resolution at the thermal neutron peak energy of 9.5%. This is sufficient to separate the events in the signal peak from the Compton background continuum at lower energy. Here we obtained a gof of $\chi^2/dof = 1.43$, again indicating an appropriate model. Converting the count amplitude of the Gaussian component of the neutron peak again into a rate, and normalising with respect to the flux, we obtained a thermal neutron detection efficiency of $\epsilon = 4.55(3) \times 10^{-4}$, where the statistical uncertainty is propagated from the fit uncertainty. In addition, varying the fit range induces a systematic effect of 6%. The determined efficiency value is 14 % lower than the value found in Ref. [84]. Let us underline that the former measurement was performed above ground, therefore with a higher background rate, and for only around a third of the measurement time. As the neutron peak fit does not accurately model potential tail events, we assign an additional systematic uncertainty on the efficiency. We compared our fit evaluation with a counting statistic based on the measured spectrum and a background-only measurement of 17 d. Using a 2σ wide window, where σ denotes the standard deviation of the Gaussian component, we observed only 12 background events. These are negligible compared to the $\mathcal{O}(10^5)$ events in the signal peak. We will provide the details on the counting analysis in the following subsection. Here we remark that we tested $2, 3, 4\,\sigma$ intervals to estimate the impact of the counting range, which yields a maximum discrepancy to the efficiency value as

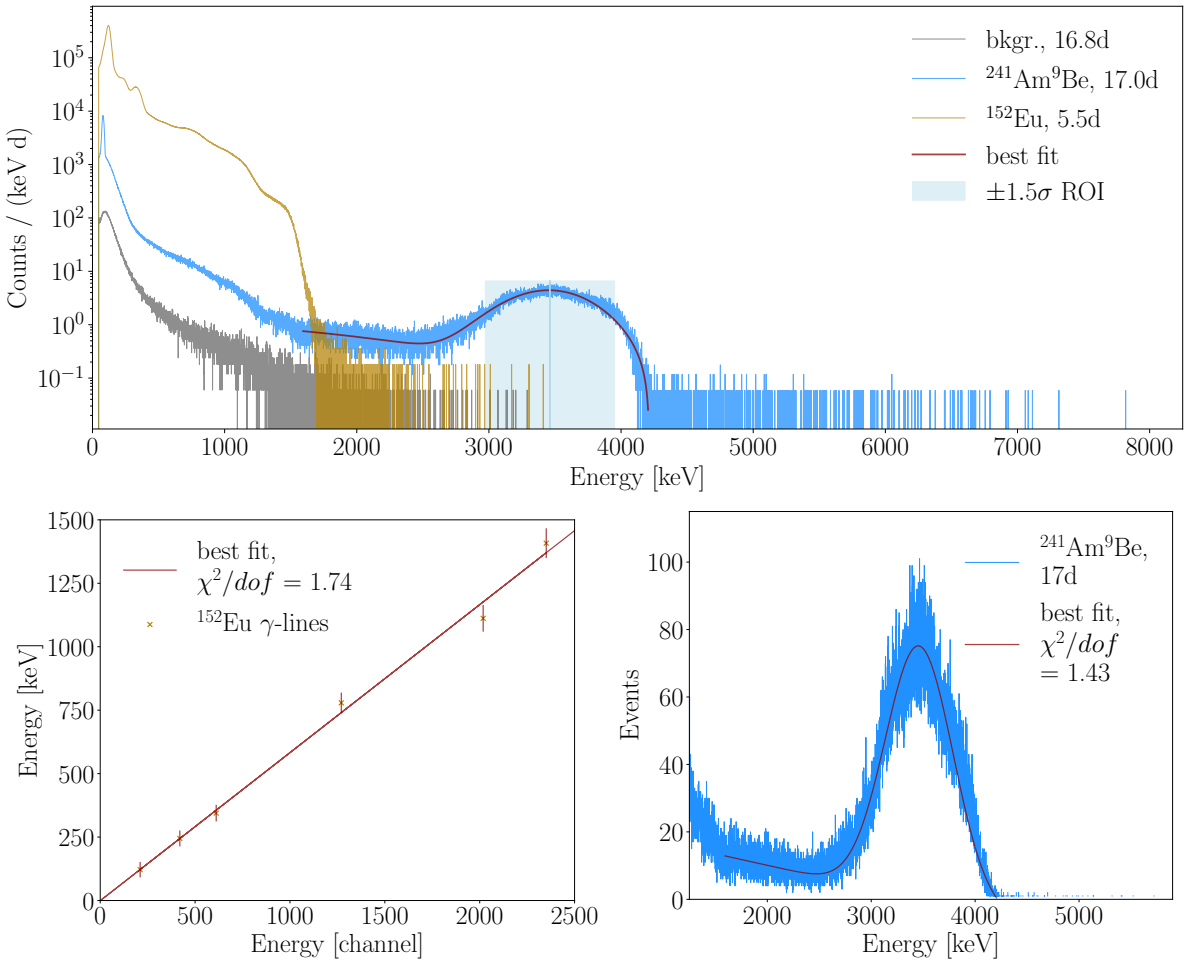

Figure 3.21: Top: energy spectra taken with an ^{152}Eu γ source to calibrate the detector's energy response (gold), with an ^{241}Am^{9}Be neutron source to evaluate the neutron detection efficiency (blue), and without a source to measure the background (grey). The energy region highlighted in light blue indicates a $\pm 1.5\sigma$ region of interest for thermalised neutrons. The standard deviation of the Gaussian component of a fit of the peak caused by the neutrons emitted by the ^{241}Am^{9}Be source (red curve) was used to estimate the width of the signal region. Bottom left: calibration curve deduced from γ-line signals via measurement of a ^{152}Eu source to convert from uncalibrated energy hits into physical units, here keV. Bottom right: zoom into the fit region for the efficiency determination. The efficiency estimation was crosschecked via a counting analysis, cf. text for details. The figure is adapted from Ref. [1].

estimated from the fit of 9%. We consider this value as the systematic uncertainty of the efficiency induced by the fit model approach. Let us remark that a complete overview of all investigated contributions to the uncertainty on the efficiency is provided at the end of this section. As a last preparation step, we increased the background statistics to 188.8 d, without any source deployed in between.

Neutron data taking

Proceeding with our measurement campaign, we took neutron data in a combined measurement with the 17 calibration sources for 79.9 d. After removing two strong sources for hardware testings in L-60, we continued with the 15 remaining sources for another 58.6 d. The corresponding spectra and the environmental neutron background are shown in Fig. 3.22.

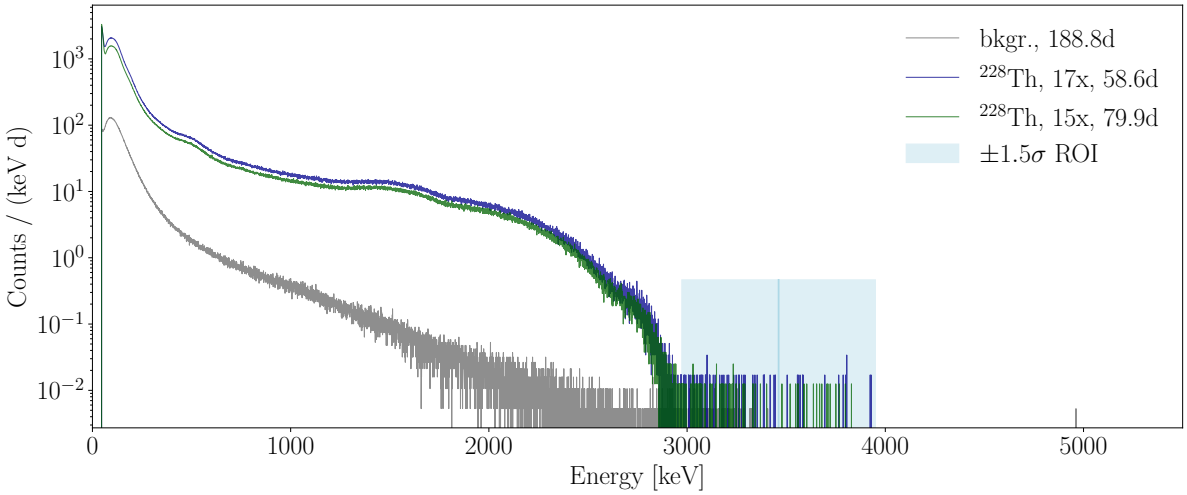

Figure 3.22: The obtained energy spectra from the two neutron measurements with 17 (blue) and 15 (green) sources, respectively, and from the long-term background measurement (grey). The light blue region corresponds to the thermal neutron signal region as determined in Fig. 3.21. The figure is adapted from Ref. [1].

We again perform a counting analysis to analyse the neutron flux emitted by the radioactive sources, as, in this case, the event numbers are too low to apply a binned fit. As the source slots of the individual sources differ from the position of the thread at which the neutron source was placed, we weigh the efficiency with the normalised solid angle and the activity of each sources [1]. At the end of this section, we will provide the details of the weighing procedure and the corresponding counting evaluation. The individual solid angles and activities in the weights are not precisely known. Thus, we also took these uncertainties into account in our analysis. We took the individual uncertainties on the source activities directly from the measurements of our colleagues at LANL. We estimate the effective solid angle when normalised to the position of the efficiency measurement to 0.97, with a standard deviation of 0.05, where we applied the formulae for the solid angle correction in a cylindrical geometry as provided in Ref. [136].
Furthermore, we assign to each solid angle a systematic effect of 3%, which we estimated by varying the geometrical distances by 4 mm. This corresponds to the dimension over which the radioactive ^{228}Th is distributed. Additionally, not all the dimensions are measured within mm precision, further motivating to conservatively include a systematic uncertainty due to this deviation from a point-like geometry. Moreover, the different source slots cause a deviation of the neutron moderation length from the nominal 50 mm. We hence compared the amplitude of the neutron peak of the main efficiency measurement to the amplitude of the peak obtained during a 5 d long data taking period with 30 mm PE only. The peak height drops by 11%, which we assign as a conservative estimate of the systematic contribution associated with the different moderation lengths. As a final systematic effect, we need to consider the different neutron emission spectra of our target isotope ^{228}Th, and of ^{241}Am^{9}Be used for the efficiency

evaluation. The impact of this deviation had been investigated via simulations in [137], stating a 12.1% effect on the uncertainty of the neutron flux of the sources directly.
After discussing the systematic effects, we can now perform the counting experiment of the actual neutron measurement. Separating by 17 and 15 source data taking, respectively, we obtain estimates for the combined neutron flux as emitted by the calibration sources as

$$\Psi = \begin{cases} (4.00 \pm 0.88_{\text{stat}} \pm 0.90_{\text{sys}}) \ \times 10^{-4} \ \text{n} \,/\, (\text{kBq}\,\text{s}) & 17 \text{ sources} \\ (4.96 \pm 1.02_{\text{stat}} \pm 01.24_{\text{sys}}) \ \times 10^{-4} \ \text{n} \,/\, (\text{kBq}\,\text{s}) & 15 \text{ sources} \end{cases} . \tag{3.2}$$

We evaluate the statistical uncertainty as the interval bound of a symmetrised 1 σ confidence interval approximated as a Gaussian distribution. The systematic uncertainty contains all analysis-related and geometrical effects and the spectral effect, as discussed above. As we performed the two measurements with the same set of sources, except for those taken out earlier, it is more intuitive to determine a global estimate. As a simple, intuitive estimation, we can calculate a weighted average of the two flux results, weighing each contribution by its *inverse* variance, which yields a global estimate of

$$\Psi = (4.36 \pm 0.67_{\text{stat}} \pm 0.73_{\text{syst}}) \times 10^{-4} \ \text{n} \,/\, (\text{kBq}\,\text{s}) \ . \tag{3.3}$$

With this method, we enhance the contribution of the more precise measurement, clearly a reasonable choice of weights. However, having largely overlapping source configurations within the same measurement apparatus, we expect a strong correlation between the results. To consider this, we follow the prescription for combining two measurements with correlated systematic uncertainties as outlined in Ref. [138]. We need to estimate the correlation coefficient between the measurements to evaluate the combined, correlated flux and its uncertainty. Considering the data were taken with the same sources at the same source holder threads, implying the same moderation lengths and solid angles, we assume a maximum correlation between the systematic uncertainties s_i, where i increments over the two data sets. Consequently, we evaluate the global contribution in the systematic uncertainty of both measurements c as the minimum of the two systematic uncertainties s_{min}, i.e., we set $c = s_{\text{min}}$. We then calculate the correlation coefficient following Ref. [138] as

$$\rho = \frac{c}{t_1 t_2} = \frac{s_{\text{min}}}{\sqrt{s_1^2 + r_1^2}\sqrt{s_2^2 + r_2^2}} = 0.40 \ , \tag{3.4}$$

where r and t denote the statistical and the total uncertainty, respectively. Assuming Gaussianity for the uncertainties, as motivated below as a reasonable approximation, we obtain a global estimate respecting correlations as

$$\Psi = (4.30 \pm 0.69_{\text{stat}} \pm 0.93_{\text{syst}}) \times 10^{-4} \ \text{n} \,/\, (\text{kBq}\,\text{s}) \ . \tag{3.5}$$

The estimated flux is similar in magnitude to the weighted average above, but the increased statistic and particularly systematic uncertainty intervals are now more appropriately estimated. Note that in the analysis presented here, the numerical results deviate by a few percent from the values published in Ref. [1]. The reason is a more refined estimation of the statistical uncertainties, considering the influence of the background fluctuations. The interpretation of the results remains unchanged. As noted above, we give a complete list of all sources of uncertainty on the measurement at the end of the section. As an important outcome, let us underline that the neutron flux emitted by the L-200 calibration sources is more than an order of magnitude below the flux emitted by commercial sources [84, 124]. When comparing the flux to the measured results for the sources operated in GERDA Phase II and II+, the flux is of similar magnitude,

but reduced by approximately 48 and 45%, respectively [84, 132].
In our measurement, we performed a dedicated analysis of the systematic uncertainties induced by the geometrical setup, whereas such a detailed investigation of these effects had not been done in the past. Overall, we see that the dominant uncertainty on our result is of systematic nature. This implies that the uncertainty on the global neutron emission of the radioactive sources cannot be strongly reduced by simply extending the data-taking duration. Instead, an improvement on the uncertainty induced by the efficiency measurement from the different source threads in the main measurement could be obtained by taking data with each radioactive source individually. This would induce a questionably long total duration, strongly delaying the possible deployment of the sources in L-200 on a reasonable time scale. As an alternative option, which would definitely be quicker, one could determine the efficiency for each source slot individually. This was not possible due to the limited time frame the ^{241}Am^{9}Be source was provided by LNGS. A third option, running dedicated simulations of the neutron propagation, may tighten the impact of the moderation length and solid angle, and also of the different emission spectra. But it will not reduce the contribution of the PMT response on the overall detection efficiency, which will not be modelled easily. Also the contribution of the uncertainty of the source activities would not be taken care of. We further feel that having a dedicated measurement suggests an efficiency estimate which is more reliable and convincing than the result deduced from a simulation, which itself underlies other effects such as the accuracy of the geometrical model. To summarise, obtaining a more precise global neutron flux estimate was not feasible, but the determined result is an improvement compared to similar measurements for GERDA II and II+. To quantify the precise impact of the measured neutron flux as emitted by the sources on the $0\nu\beta\beta$ analysis in L-200, we ran detailed simulations as described in the Sec. 3.5. We will see that the estimated induced effect mitigates any need for a more precise measurement result.

Formulae for the analysis of the measurements

Here we provide a detailed overview of the formulae used for the neutron efficiency and the neutron flux measurement, as discussed above and as published in Ref. [1]. We closely follow the analysis of counting statistics in the presence of systematic uncertainties as provided in Refs. [139, 140]. To perform a counting statistic to evaluate the detection efficiency from the data taken with the ^{241}Am^{9}Be neutron source, we write the likelihood function for a Poisson measurement in the presence of a background B, a signal strength M, and with a Gaussian systematic on the flux Φ as

$$\begin{aligned} \mathcal{L}(n, b; \phi | M, B; \Phi) &= \mathcal{P}(b|B) \times \mathcal{P}(n|M, B) \times \mathcal{G}(\phi|\Phi) \\ &= \frac{B^b e^{-B}}{b!} \times \frac{(M + \alpha B)^n e^{-(M+\alpha B)}}{n!} \times \frac{1}{\sqrt{2\pi t_s^2 \Delta_\phi^2}} \, e^{-(\Phi-\phi)^2/2\Delta_\phi^2} \,, \end{aligned} \tag{3.6}$$

with $\alpha = t_s/t_b$, a measurement time duration of $t_s = 17\,\mathrm{d}$, a background data taking time of $t_b = 17\,\mathrm{d}$, an initial flux of $\phi_0 = (160 \pm 4)$ n/s, and a half-life $T_{1/2} = 432.7\,\mathrm{yr}$. We denote the obtained event counts during the source and background measurements with n and b, respectively. Our measurement was performed approximately eight years after the flux measurement, i.e., the exponential decay law yields a flux of $\phi = \phi_0\, e^{-\log 2\ t/T_{1/2}}$, $t = (8 \pm 1)\,\mathrm{yr}$, where we add an uncertainty of one year as a conservative estimate. Following the prescription of Ref. [139], we

take the logarithm of the likelihood with the signal given as $M = \epsilon t_s \Phi$ to find

$$\begin{aligned} &\log \mathcal{L}(n, b; \phi | \epsilon t_s \Phi, B; \Phi) - \text{const.} \\ &= b \log B + n \log(\alpha B + \epsilon t_s \Phi) - (\alpha + 1) B - \epsilon t_s \Phi - \frac{(\Phi - \phi)^2}{2 \Delta_\phi^2} \ . \end{aligned} \tag{3.7}$$

Forcing the partial derivatives to vanish for the maximum in order to determine maximum likelihood estimators (MLE), we now obtain the set of equations

$$\left. \frac{\partial \log \mathcal{L}}{\partial B} \right|_{\text{MLE}} = \frac{b}{B_{\text{MLE}}} + \frac{n\alpha}{\alpha B_{\text{MLE}} + \epsilon_{\text{MLE}} t_s \Phi_{\text{MLE}}} - (\alpha + 1) \stackrel{!}{=} 0 \ , \tag{3.8}$$

$$\left. \frac{\partial \log \mathcal{L}}{\partial \Phi} \right|_{\text{MLE}} = \frac{n \epsilon_{\text{MLE}} t_s}{\alpha B_{\text{MLE}} + \epsilon_{\text{MLE}} t_s \Phi_{\text{MLE}}} - \epsilon_{\text{MLE}} t_s - \frac{(\Phi_{\text{MLE}} - \phi)}{\Delta_\phi^2} \stackrel{!}{=} 0 \ , \tag{3.9}$$

and

$$\left. \frac{\partial \log \mathcal{L}}{\partial \epsilon} \right|_{\text{MLE}} = \frac{n t_s \Phi}{\alpha B_{\text{MLE}} + \epsilon_{\text{MLE}} t_s \Phi_{\text{MLE}}} - t_s \Phi_{\text{MLE}} \stackrel{!}{=} 0 \ . \tag{3.10}$$

We obtain the corresponding solutions as $B_{\text{MLE}} = b$, $\Phi_{\text{MLE}} = \phi$, and the efficiency reads

$$\epsilon := \epsilon_{\text{MLE}} = \frac{n - \alpha b}{t_s \phi} \ . \tag{3.11}$$

In addition, we need to estimate the systematic uncertainty on the efficiency, which we do via error propagation, yielding

$$\Delta_\epsilon = \epsilon^2 \sqrt{\frac{\Delta_\phi^2}{\phi^2} + \Delta_t^2 \frac{(\log 2)^2}{T_{1/2}^2}} \ . \tag{3.12}$$

Having obtained a large number of $\mathcal{O}(10^5)$ counts, the statistical uncertainty on the efficiency measurement can be accurately approximated via its Gaussian standard deviation as

$$\sigma_\epsilon = \sqrt{\frac{n + \alpha^2 b}{t_s^2 \phi^2}} \ . \tag{3.13}$$

When analysing the neutron flux measurement, let us keep in mind that we need to consider additional systematic uncertainties, including the corrections due to the solid angle, the different moderation lengths, the counting range, and the different emission spectra. For this analysis, we then combine the statistical uncertainty with all systematic contributions. To estimate the neutron flux, we can now proceed as follows. We use the ^{228}Th half-life $T_{1/2}$ =1.9116 yr, and an activity scaling due to the exponential decay as $A = A_0 \, e^{-\log 2 \, t / T_{1/2}}$ with $t = (10.5 \pm 1)$ months between γ-ray and neutron data taking. The likelihood for a counting statistic for the neutron flux Ψ, now evaluated with two Gaussian uncertainties for efficiency and activity, reads

$$\begin{aligned} &\mathcal{L}(n, b; A', \epsilon_w | M, B; A, \epsilon) = \mathcal{P}(b|B) \times \mathcal{P}(n|M, B) \times \mathcal{G}(A'|A) \times \mathcal{G}(\epsilon_w|\epsilon) \\ &= \frac{B^b e^{-B}}{b!} \times \frac{(M + \alpha B)^n e^{-(M + \alpha B)}}{n!} \times \frac{1}{\sqrt{2\pi t_s^2 \Delta_{A'}^2}} \, e^{-(A - A')^2 / 2 \Delta_{A'}^2} \times \frac{1}{\sqrt{2\pi \Delta_{\epsilon_w}^2}} \, e^{-(\epsilon - \epsilon_w)^2 / 2 \Delta_{\epsilon_w}^2} \ , \end{aligned} \tag{3.14}$$

where $\alpha = t_s/t_b$, and $t_s = 58.6\,\mathrm{d}$ (79.9 d) for the 17 (15) source measurement, and t_b =188.778 d. The logarithm of this likelihood with a source strength $M = f\epsilon A t_s \Psi$ becomes

$$\begin{aligned} &\log\mathcal{L}(n,b;A',\epsilon_w|f\epsilon A t_s\Psi,B;A,\epsilon) - \text{const.} \\ &= b\log B + n\log(\alpha B + f\epsilon A t_s\Psi) - (\alpha+1)B - f\epsilon A t_s\Psi - \frac{(\epsilon-\epsilon_w)^2}{2\Delta^2_{\epsilon_w}} - \frac{(A-A')^2}{2\Delta^2_{A'}} \ . \end{aligned} \tag{3.15}$$

Forcing the partial derivatives again to vanish for maximum in order to determine maximum likelihood estimators, we obtain

$$\left.\frac{\partial\log\mathcal{L}}{\partial B}\right|_{\mathrm{MLE}} = \frac{b}{B_{\mathrm{MLE}}} + \frac{n\alpha}{\alpha B_{\mathrm{MLE}} + \epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}} - (\alpha+1) \overset{!}{=} 0 \ , \tag{3.16}$$

$$\left.\frac{\partial\log\mathcal{L}}{\partial A}\right|_{\mathrm{MLE}} = \frac{nf\epsilon_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}}{\alpha B_{\mathrm{MLE}} + f\epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}} - f\epsilon_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}} - \frac{(A_{\mathrm{MLE}} - A')}{\Delta^2_{A'}} \overset{!}{=} 0 \ , \tag{3.17}$$

$$\left.\frac{\partial\log\mathcal{L}}{\partial \epsilon}\right|_{\mathrm{MLE}} = \frac{nf A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}}{\alpha B_{\mathrm{MLE}} + f\epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}} - f A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}} - \frac{(\epsilon_{\mathrm{MLE}} - \epsilon_w)}{\Delta^2_{\epsilon_w}} \overset{!}{=} 0 \ , \tag{3.18}$$

and

$$\left.\frac{\partial\log\mathcal{L}}{\partial \Psi}\right|_{\mathrm{MLE}} = \frac{nf\epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s}{\alpha B_{\mathrm{MLE}} + f\epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s \Psi_{\mathrm{MLE}}} - f\epsilon_{\mathrm{MLE}} A_{\mathrm{MLE}} t_s \overset{!}{=} 0 \ . \tag{3.19}$$

The solutions of this set of equations are $B_{\mathrm{MLE}} = b$, $A_{\mathrm{MLE}} = A'$, $\epsilon_{\mathrm{MLE}} = \epsilon_w$, and

$$\Psi := \Psi_{\mathrm{MLE}} = \frac{n - \alpha b}{f\epsilon_w A' t_s} \ . \tag{3.20}$$

We count $n = 60$ (77) events for the 17 (15) source run, and $b = 73$ events during the background measurement. Correcting for the $1.5\,\sigma$ ROI search region, assuming all events to be caused by neutrons, we divide the observed event numbers by a fraction of expected events in the chosen interval of $f = 0.8664$. The effect of the different neutron energies due to different emission spectra has been estimated in Ref. [137]. It is taken into account in calculating the systematic uncertainty, which we evaluate in the following. The individual source activities $A_{0,i}$ were measured on February 09, 2021, as published in Ref. [1], leading to a total survival activity of

$$A' = \sum_{i=1}^{17(15)} A_{0,i} e^{-\log 2 \ t/T_{1/2}} \ , \tag{3.21}$$

with initial uncertainty

$$\Delta_{A_0} = \sqrt{\sum_{i=1}^{17(15)} \Delta^2_{A_{0,i}}} \ , \tag{3.22}$$

and total uncertainty

$$\Delta_{A'} = \sqrt{\Delta^2_{A_0} e^{-2\log 2t/T_{1/2}} + \Delta^2_t \frac{(\log 2)^2}{T^2_{1/2}} A'^2} \ . \tag{3.23}$$

We define the source-activity and solid-angle weighted efficiency as

$$\epsilon_w = \epsilon \sum_{i=1}^{17(15)} \frac{A_{0,i}}{A_0} \frac{\Omega_i}{\Omega_\epsilon} \ , \tag{3.24}$$

where Ω denotes the solid angle of a source thread. Here we use the equal time-scaling of the individual and total activity, i.e., the exponential factor cancels out, reducing the uncertainty. The overall systematic uncertainty on the weighted efficiency then becomes

$$\Delta_{\epsilon_w} = \sqrt{\Delta_\epsilon^2 \frac{\epsilon_w^2}{\epsilon^2} + \epsilon^2 \frac{\Omega^2}{\Omega_\epsilon^2} \frac{1}{A_0'^2} \times \left(\Delta_{A_0}^2 + \sum_{i=1}^{17(15)} \Delta_{A_{0,i}}^2 \left(1 - \frac{A_{0,i}}{A_0}\right)^2 \right) + \epsilon^2 (\Omega \longleftrightarrow A_0)} \ , \tag{3.25}$$

where the $(\Omega \longleftrightarrow A_0)$ refers to the central term in the radicand with all Ω_i and $A_{0,i}$ exchanged due to the symmetry in Eq. (3.24). We find the effect induced by the solid angle uncertainty subdominant at around the 1% level only. We can now determine the overall systematic uncertainty by taking into account the correlation between A' and ϵ_w, plus an additional uncorrelated 12.1% uncertainty due to different neutron emission spectra, as

$$\Delta_\Psi = \Psi \sqrt{\left(\frac{\Delta_{A'}}{A'} + \frac{\Delta_{\epsilon_w}}{\epsilon_w} \right)^2 + 0.121^2} \ . \tag{3.26}$$

Note that even though the event numbers are not very high, we can here approximate the Poisson distribution via a Gaussian distribution, as the impact on the statistical uncertainty is marginal compared to the magnitude of the dominating systematic effect. We can then calculate the statistical uncertainty on the neutron flux simply as

$$\sigma_\Psi = \sqrt{\frac{n + \alpha^2 b}{f^2 \epsilon_w^2 A'^2 t_s^2}} \ , \tag{3.27}$$

which yields a symmetric interval. We used these formulae when analysing the neutron flux measurement above and in Ref. [1]. In Tabs. 3.4 and 3.5, we summarise all contributions to the uncertainty on the neutron detection efficiency, and on the neutron flux of the ^{228}Th sources, respectively.

Source of uncertainty	Magnitude $\times 10^{-4}$	Relative size [%]
^{241}Am^{9}Be source flux, with decay	0.11	2.5
moderation length	0.5	11
fit range	0.27	6
fit model	0.41	9
total, systematic	0.71	16
statistical	0.03	0.7
total, stat. + syst.	0.71	16

Table 3.4: List of contributions to the uncertainty on the neutron detection efficiency of the LiI(Eu) detector, together with their absolute size and relative to the best-fit value of the efficiency of $\epsilon = 4.55 \times 10^{-4}$. The total systematic and overall uncertainties are shown as well.

3.5 Simulations of source-induced background

We simulated the calibration procedure of L-200 to determine the impact of the calibrations on the expected background rate. Neutrons were emitted at the initially planned calibration positions inside the L-200 detector array, followed by dedicated simulations of the decay chains

Source of uncertainty	Magnitude $\times 10^{-4}$ [n / (kBq s)]	Relative size [%]
emission spectra [137]	0.48 / 0.60 / 0.52	12 / 12 / 12
^{228}Th sources activity, with decay	0.12 / 0.30 / -	3 / 6 / -
solid angle & activity weighting of efficiency	0.64 / 0.79 / -	16 / 16 /-
total, systematic	0.90 / 1.24 / 0.93	23 / 25 / 22
statistical	0.88 / 1.02 / 0.69	22 / 21 / 16
total, stat. + syst.	1.26 / 1.61 / 1.16	31 / 33 / 27

Table 3.5: List of contributions to the uncertainty on the neutron flux emitted by the L-200 calibration sources as measured with the LiI(Eu) detector, in combination with their magnitude and relative size when normalised to the best-fit flux results of $\Psi = 4.00$ / 4.96 / 4.30 $\times$ 10^{-4} n / (kBq s) for the measurements with 17 / 15 sources / global estimate. We also indicate the overall systematic and total uncertainties.

of the potentially dangerous produced isotopes, comprising locations within the Ge detectors and in external materials. As mentioned above, the most important contributor is the ^{77}Ge ground-state beta decay to ^{77}As, with a Q-value of 2.7 MeV, and a half-life of $T_{1/2} = 11.3$ h. The metastable state, usually considered one of the most critical isotope decays in LEGEND if caused by cosmogenic neutron activation [54], is not relevant in our case. The source removal time after finishing a calibration is approximately 14 min. With a half-life of $T_{1/2} = 52.9$ s, ^{77m}Ge isotopes decay fast enough to induce a suppression of their background contribution by more than five orders of magnitude by the end of a calibration. Other potential contributors comprise ^{41}Ar with a beta decay Q-value of 2.49 MeV and $T_{1/2} = 1.8$ h, produced upon neutron capture on the most abundant isotope in the cryostat, ^{40}Ar. As the branching ratio of the respective decay mode is 0.78% [49], and the emitted electrons need to propagate through the liquid and penetrate the detector beyond the dead layer without losing relevant energy, the expected contribution is still subdominant, as we will prove below. Potentially, ^{76}Ga could be produced by highly-energetic neutrons with a production threshold of 6.2 MeV, leading not only to a β-transition with a Q-value of 6.9 MeV [49], i.e., beyond 2 MeV, but also causing a γ-transition within 1 FWHM of $Q_{\beta\beta}$ [141]. Two important points prevent a relevant production. The neutron energy spectrum of the sources, peaking at around 3 MeV, is already highly suppressed above 6 MeV [123], meaning only a few ^{76}Ga isotopes may be produced at all. Given their $T_{1/2}$ of 33 s [49], these would then decay, yielding an overall suppression during the source removal duration of around eight orders of magnitude.

We investigated the production rate of the relevant isotopes during the simulations to confirm the agnostic expectations quantitatively. We ran all simulations in MaGe [142], a GEANT4 [143, 144, 145] -based simulation framework specifically tailored to the GERDA, MJD, and LEGEND experiments. The software simulates particle interactions inside the relevant geometrical configurations and the corresponding particle propagation and energy deposition. Due to a change in the array configuration from the planned 14 to 12 strings due to space considerations for the LAr veto barrels, the simulated geometry does not exactly correspond to the currently operated hardware layout. However, the overall dimensions are rather similar. We will see below that the induced background level is sufficiently low to neglect the uncertainty induced by the geometrical differences. We generated 10^7 neutrons initiated from the calibration positions foreseen earlier. Given the 14-fold symmetry in the assumed configuration (Fig. 3.23), we limit ourselves to source locations one and two, effectively doubling the statistics due to the symmetry. Within the MaGe reference system, these positions are parametrised as $S1 = (150, 0, z)$ mm, and $S2 = (0, 150, z)$ mm, where $z \in \{425, 325, 225, 105, 40, -60, -160, -280\} - 400$. The offset

at the end shifts the positions to the centre of the detector array in the vertical axis.

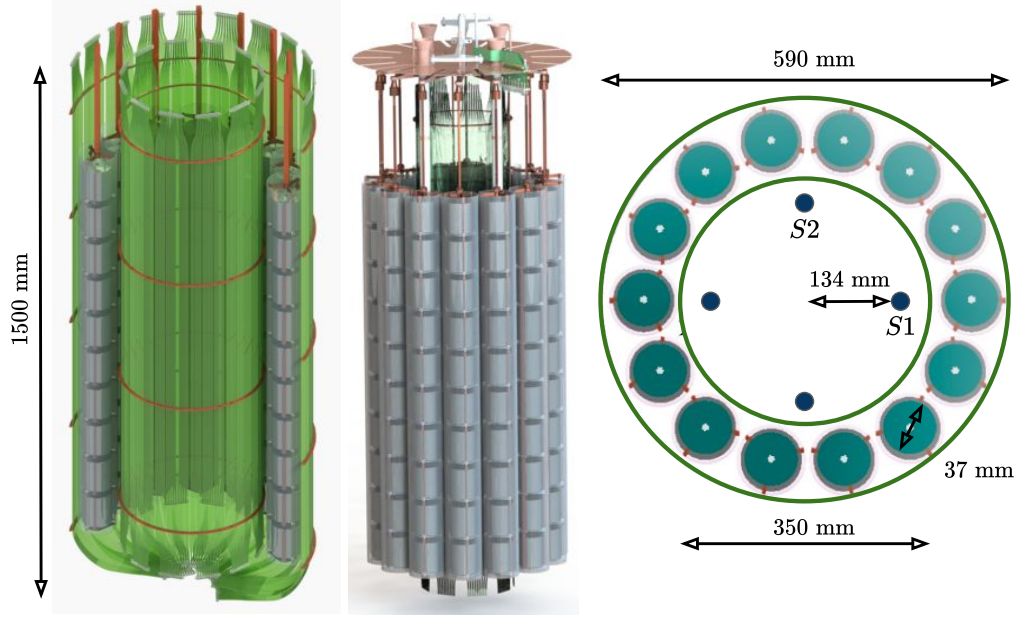

Figure 3.23: Preliminary geometry implementation of LEGEND-200 in the GEANT4-based MaGe simulation framework. Left: detector array with LAr veto fibres. Middle: x-z-view of the array. Right: x-y-view of the detector array with calibration source positions. The sketches are adapted from Ref. [54], P. Krause, R. Massarcyk, C. Ransom, and other collaboration members.

We determined the energy distribution of the neutrons via a cubic spline interpolation of the data of the neutron flux spectrum simulated for GERDA Phase I [123]. We then investigated the locations where simulated neutron captures were observed, identifying the detectors, the LAr veto setup, and the cryostat walls as the main locations, with captures in the entire LAr uniformly distributed. An overview of the positions of the simulated neutron capture reactions is given in Fig. 3.24.

Isotope	Number of produced isotopes	Q-value [MeV]	$T_{1/2}$	Suppression factor	Relevant fraction
any	8817752	not unique	not unique	not unique	9.764×10^{-2}
^{77}Ge	$21811\times\sim0.5$	2.7	11.3 h	0.986	1.075×10^{-3}
^{77m}Ge	$21811\times\sim0.5$	2.7	52.9 s	1.659×10^{-5}	1.810×10^{-8}
^{41}Ar	177512	2.49	109.34 min	0.915	1.750×10^{-2}
^{76}Ga	0	6.9	32.6 s	1.751×10^{-8}	$\leq 1.751\times10^{-15}$

Table 3.6: Numbers of different isotopes of interest as produced via neutron capture in a simulation of 10^7 neutrons sampled from the calibration positions. The suppression factors and relevant surviving isotope fraction due to the exponential decay during source removal assumed to last 14 min, corresponding to approximately 8400 mm moving distance, are listed as well.

We also counted the number of relevant isotopes produced via neutron capture, as listed in Tab. 3.6. Here we assumed an equal production rate for the ground state and the metastable state of ^{77}Ge, which is, in approximation, sufficiently accurate for our needs for the following reason. Taking into account that approximately 10 to 20% more ^{77m}Ge is produced [146, 147],

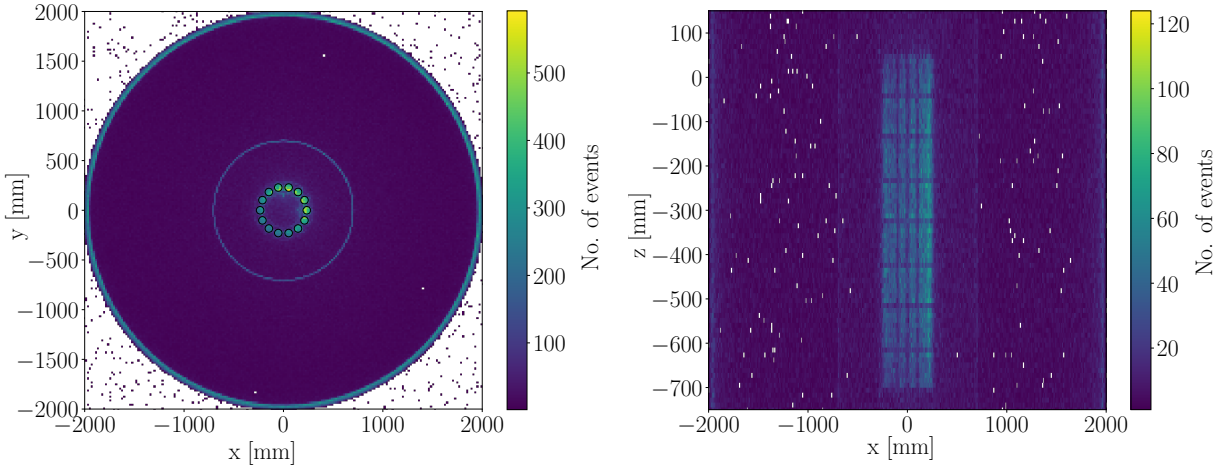

Figure 3.24: Locations of neutron capture reactions in the cryostat caused by neutrons emitted from the sources as simulated with a L-200-like GEANT4 geometry. The small, inner circles indicate the position of the detectors, and the centre and outer rings are the LAr veto system and the cryostat walls, respectively. Left: x-y-view. Right: x-z-view.

the 19% of ^{77m}Ge nuclei that decay via its ground-state result in a roughly equal abundance. In Tab. 3.6, we also estimate the fraction of relevant isotopes surviving the source removal time, i.e., that may potentially lead to a background event during physics data taking. From the survival fraction, we readily deduce the subsequent cases still to be simulated are most importantly ^{77}Ge, and in addition, ^{41}Ar, but ^{77m}Ge and ^{76}Ga are irrelevant due to the time suppression. We then simulated 10^7 ^{77}Ge full decay chains originating inside the detector strings.
We again effectively double the simulated events due to the 14-fold symmetry by using a uniform sampling among all channels in strings 1 to 7. We subsequently analysed the resulting smeared energy spectra before and after background reduction cuts to estimate the expected event rates in the region-of-interest for the $0\nu\beta\beta$ decay search. We emphasise the usage of the word *smeared*, meaning we took the finite detector energy resolution into account by adding to each detected energy value a Gaussian random variable with a mean equal to zero, and a standard deviation equal to the energy resolution of the measured GERDA IC detector resolution curved evaluated at the observed energy. We also applied the detector anti-coincidence cut, and a simplified pulse shape discrimination cut based on the spatial extent of the simulated hits in MaGe. This simplified approach mitigates the need for dedicated waveform simulations, still estimating the fraction of single versus multi-site events expected to be distinguished by the full pulse shape analysis. More details on this simplified method are given at the end of this section. The ^{77}Ge beta decay hit event distribution is shown in Fig. 3.25. The corresponding energy histogram, summed over all detectors channels, is plotted in Fig. 3.26, where the effects of the background reduction cuts are also indicated.

We list the counted simulated hits in different energy ranges, particularly in the ROI, here defined as 1950 to 2150 keV, in Tab. 3.7.

Note the reduction of two orders of magnitude upon application of the PSD cuts. This strong suppression can only be achieved due to the accompanying γ particles emitted in ^{77}As de-excitation processes following the ^{77}Ge beta decays, which causes a dominant multi-site event character for these decays. In Fig. 3.27, we illustrate the decay scheme of the $^{77(m)}$Ge isotope, which indicates the presence of the mentioned accompanying γ-ray emissions in the case of the

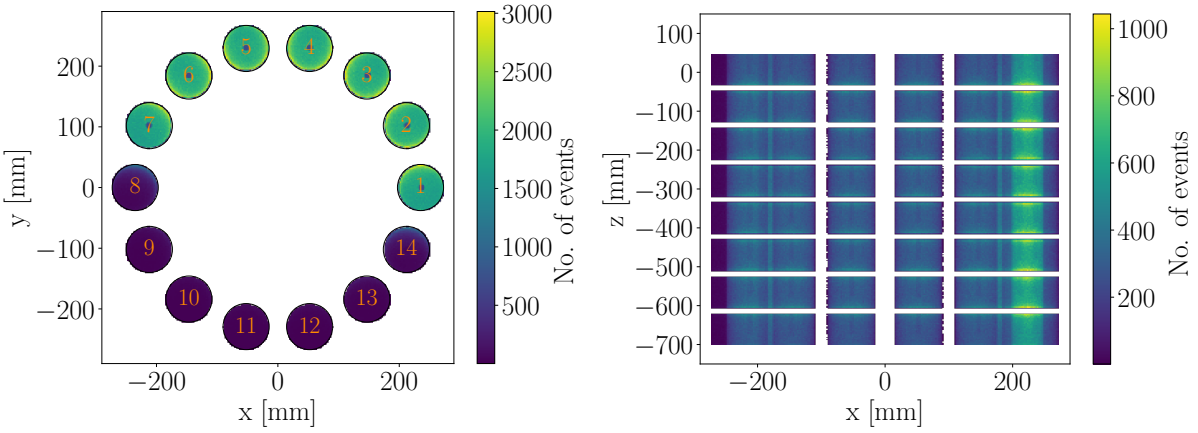

Figure 3.25: Simulated energy depositions from ^{77}Ge isotope decays inside the HPGe detectors when sampled uniformly among strings 1 to 7. Left: x-y-view. Right: x-z-view.

Cut level	Full spectrum	$\geq$ 200 keV	ROI 1950-2150 keV	ROI event fraction
before	14939830	12227300	434626	4.346×10^{-2}
M1/AC	6049439	5937062	346822	3.468×10^{-2}
PSD	3424357	1536459	3312	3.312×10^{-4}
M1/AC+PSD	505820	420063	2168	2.168×10^{-4}

Table 3.7: Number of energy depositions in given energy ranges as observed in detectors due to ^{77}Ge decay, as well as the relevant ROI fraction.

ground-state beta decay.

Using the measured neutron flux Ψ from Sec. 3.4, a nominal ^{76}Ge enrichment fraction of $f_{^{76}\mathrm{Ge}} = 0.9$, the relevant surviving isotope fraction s, the ROI width $w = 200\,\mathrm{keV}$, and the ROI event fraction f_{ROI} before / after cuts, we can now convert the neutron-induced background contribution into a full background index for L-200. Under the assumption of a weekly calibration time fraction of around $t_{\mathrm{cal}} = 2\,\mathrm{h}\ /\ (7 \times 24\ \mathrm{h})$ like in GERDA, of a nominal operated detector array mass of $m_0 = 200\,\mathrm{kg}$, and a nominal source activity of $A_0 = 16 \times 5\,\mathrm{kBq}$ compensating for potential source replacements instead of aggressively considering the activity decrease over time, we estimate an induced background before / after analysis cuts of

$$BI = \frac{\Psi\ A_0\ f_{\mathrm{ROI}}\ f_{^{76}\mathrm{Ge}}\ s\ t_{\mathrm{cal}}}{m\ w} = 1.4 \times 10^{-5}\,/\,6.7 \times 10^{-8}\ \mathrm{counts}\,/\,(\mathrm{keV\,kg\,yr})\ . \tag{3.28}$$

This value is more than three orders of magnitude below the L-200 background budget of $2 \times 10^{-4}\,\mathrm{counts}\,/\,(\mathrm{keV\,kg\,yr})$, leaving only the impact of ^{41}Ar as a potentially dangerous contributor. We simulate this case as follows. We initiate 10^7 full isotope decay chains uniformly distributed among a cylinder inside the cryogenic liquid with the following dimensions, radius $r = 750\,\mathrm{mm}$, height $h = 1500\,\mathrm{mm}$, centre vector $\vec{c} = (0, 0, 70)$ mm, and shift vector $\vec{s} = -(0, 0, 400)\,\mathrm{mm}$. These dimensions and the shift ensure that the detector array is fully enclosed by the cylinder, with an appropriate margin to cover propagation effects inside the LAr. A further expansion of the cylinder to the entire cryostat only reduces the chance for especially the electrons to reach the array, meaning it would require an extremely high number of simulations to achieve relevant statistics. Hence, we instead overestimate the potential contribution,

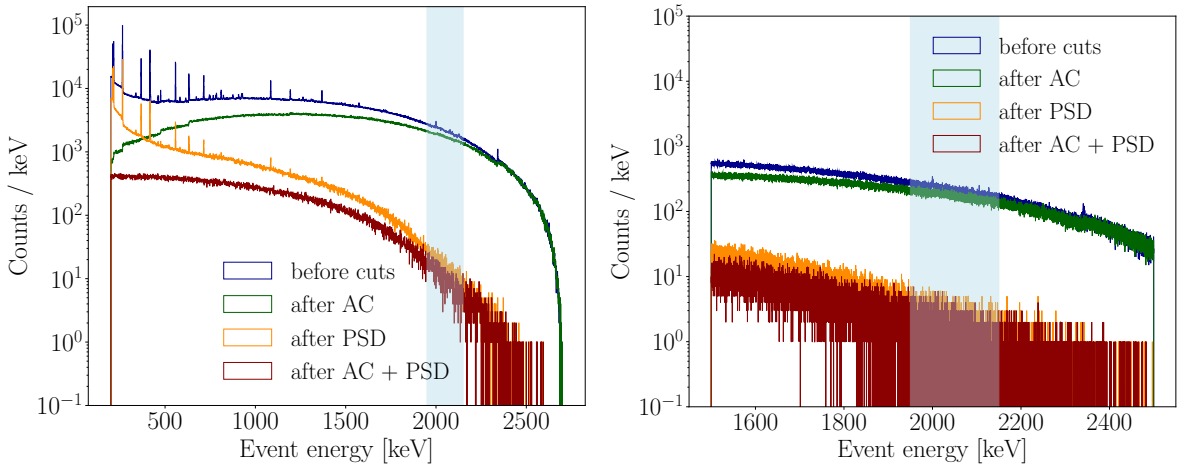

Figure 3.26: Simulated ^{77}Ge ground state decay spectra in the HPGe detectors after neutron capture on ^{76}Ge, before (blue) and after various background reduction cuts (green, yellow, red). The abbreviations refer to detector anti-coincidence (AC) and pulse-shape discrimination (PSD). The blue shaded region indicates the region-of-interest for the $0\nu\beta\beta$ search. Left: full spectrum. Right: zoom into the ROI for the $0\nu\beta\beta$ analysis.

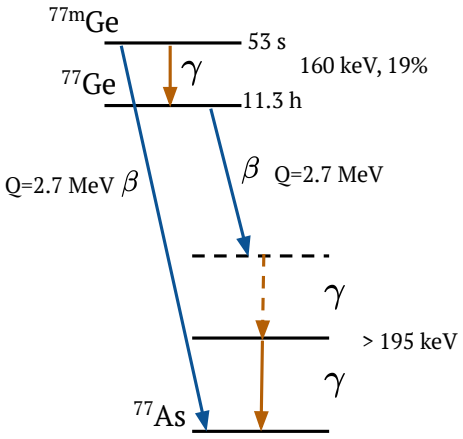

Figure 3.27: Simplified decay scheme of the metastable state and the ground-state of ^{77}Ge to ^{77}As. Whereas ^{77m}Ge either de-excitates to its ground-state by γ-ray emission or directly undergoes a beta decay to the ground-state of ^{77}As, the beta decay of ^{77}Ge always transits through an excited state of ^{77}As, and is hence always accompanied by one or more γ particle emissions during the subsequent de-excitation.

yielding a conservative result. As in the case of the ^{77}Ge study, we investigate the simulated energy depositions in the detector array, see Fig. 3.28.

We show the simulated histograms in Fig. 3.29, and the counted hits in Tab. 3.8.

Here we again apply energy smearing and evaluate the count numbers before / after the background reduction cuts, namely after AC and PSD cuts. After the latter, no events remain, which confirms that we can safely neglect the contributions caused by this isotope. Applying the same conversion scheme as above with a pseudo value of one hit, we obtain a more quantitative confirmation of a background index upper bound of less than 6×10^{-10} counts / (keV kg yr), which is fully irrelevant for the sensitivity of L-200.

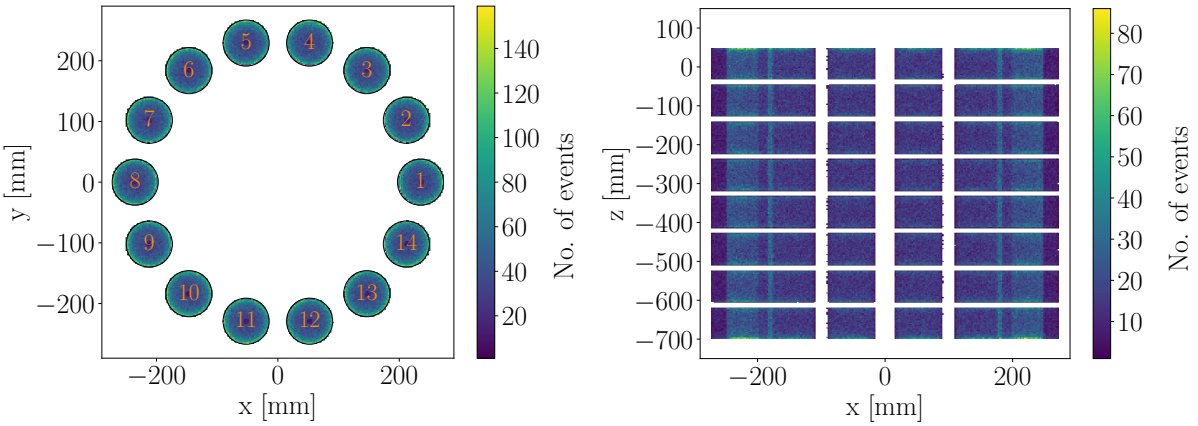

Figure 3.28: Simulated energy depositions from ^{41}Ar isotope decays around the detector array, distributed within a cylinder inside the LAr with radius $r = 750\,\text{mm}$, height $h = 2 \cdot 750\,\text{mm}$, and centre position $\vec{c} = (0, 0, 70 - 400)^T\,\text{mm}$. Left: x-y-view. Right: x-z-view.

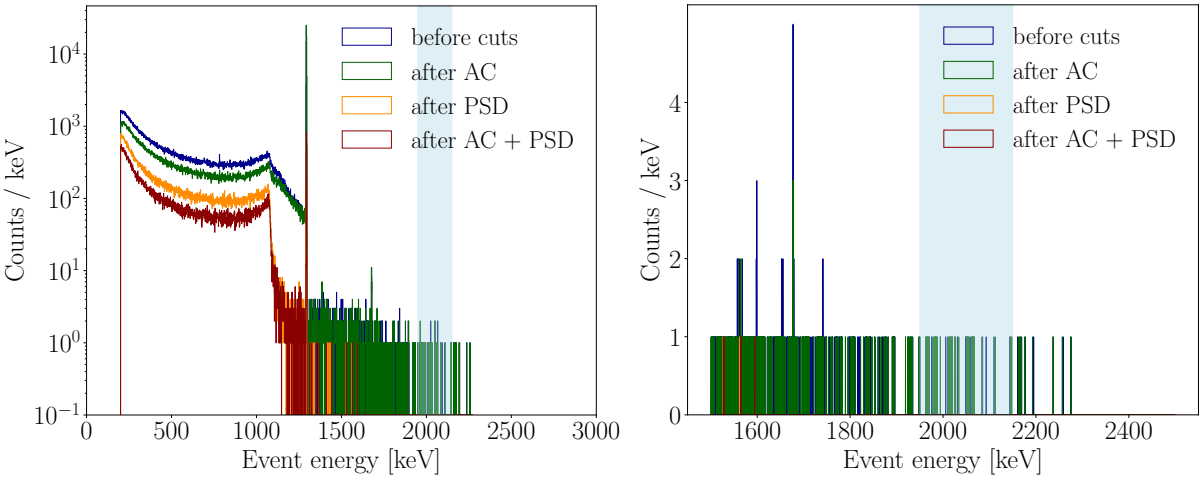

Figure 3.29: Simulated ^{41}Ar beta decay spectrum. The plotted individual functions, and the used abbreviations, follow those in Fig. 3.26. Left: full spectrum. Right: zoom into the region-of-interest for the $0\nu\beta\beta$ analysis.

Cut level	Full spectrum	$\geq 200\,\text{keV}$	ROI 1950-2150 keV	ROI event fraction
before	852736	539093	37	3.7×10^{-6}
M1/AC	605412	396310	34	3.4×10^{-6}
PSD	380545	154899	0	$\leq 10^{-7}$
M1/AC+PSD	247707	98277	0	$\leq 10^{-7}$

Table 3.8: Number of energy depositions in given energy ranges as observed in the detectors due to ^{41}Ar decay, as well as ROI fraction.

What we have not yet explained in detail, but have applied already, is the simplified PSD cut based on the spatial extent of simulated energy depositions inside the Ge detectors, mitigating the need to perform full waveform simulations with a subsequent event pulse shape discrimi-

nation study. Note that also in GERDA, a simplified method was applied frequently to obtain fast estimations after applying the PSD cuts. A simulated energy deposition was interpreted as single-site event if the corresponding spatial extent within a Ge detector does not exceed 2 mm. This value conservatively takes into account the approximate ∼1 mm mean free path of an electron of ∼2 MeV in Ge. For the complete $0\nu\beta\beta$ decay analysis, a more refined PSD cut investigation based on the actual pulse shapes had been applied [100]. We pursue the simplified approach in a slightly more refined manner as follows.
Similarly to the real PSD cut efficiency evaluation, based on the DEP of ^{208}Tl in the calibration data, we use simulated events in the DEP to determine the expected spatial extent of potential $0\nu\beta\beta$ signals. Recall that to generate the DEP, the two secondary γ particles escape the detector, and only a single site event deposition remains, i.e., a signal-like event is expected. Technically, we simulate 5×10^6 full ^{228}Th decay chain events each for the same set of 16 calibration positions as done for the neutron simulation, i.e., overall 80×10^7 decays. We again apply both the energy resolution smearing based on the measured IC resolution curve from GERDA, and the AC cut, keeping in mind that DEP events are of single-site type, meaning contained in one detector. A spectrum of such events is shown in Fig. 3.30, left.

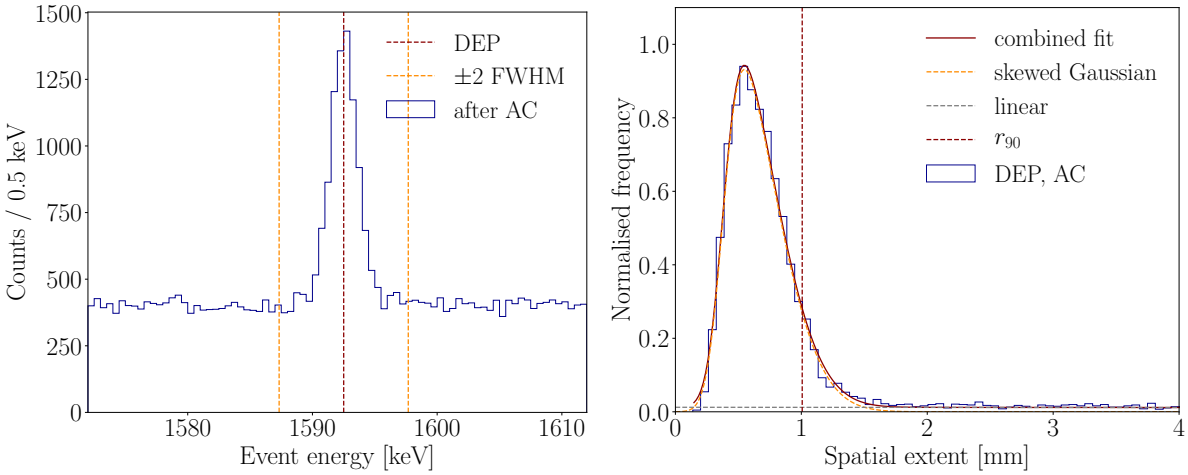

Figure 3.30: Left: simulated spectrum of DEP energy depositions summed over all detectors, after the application of the AC/M1 cut. The theoretical DEP energy is highlighted in red, and the selection window for accepted events in orange. Right: plot and fit of the spatial extent of the events falling into the selected event region from the left plot. The fit consists of a skewed Gaussian distribution to model the asymmetric signal peak shape, and a linear function for the background continuum. The r_{90} cut value of 1.0 mm corresponds to the 90% quantile of the fitted skewed Gaussian component.

Thirdly, we select energies within a window of ±2 FWHM around the DEP energy and histogram the spatial extent of the selected simulated events within the detector volume. Fourthly, we fit the obtained histogram with a skewed Gaussian to model the asymmetric peak shape of signal events, and a linear function to model the underlying continuum caused by Compton scattering events penetrating the region of interest. The histogram data, the combined fit, and the two individual components are shown in Fig. 3.30. Lastly, considering that the real PSD cuts are usually tuned to achieve a nominal signal acceptance of 90%, we interpret the 90% quantile of the signal, i.e., the skewed Gaussian, as an estimator to be used for the spatial extent cut replacing the original waveform based PSD cut. The resulting value is a spatial extent

r_{90} equal to 1.0 mm, which we have applied in the background evaluation for the isotope decay simulations outlined above. Even though the extent is smaller than what had been commonly used in GERDA, the result agrees well with what was found by other collaborators as well [121]. In addition, we investigate the impact of our spatial extent cut on the simulated data around the SEP and FEP, expecting a multi-site event character, and hence a wider distribution of spatial extents. We show a histogram of the extents in the three different peaks in Fig. 3.31.

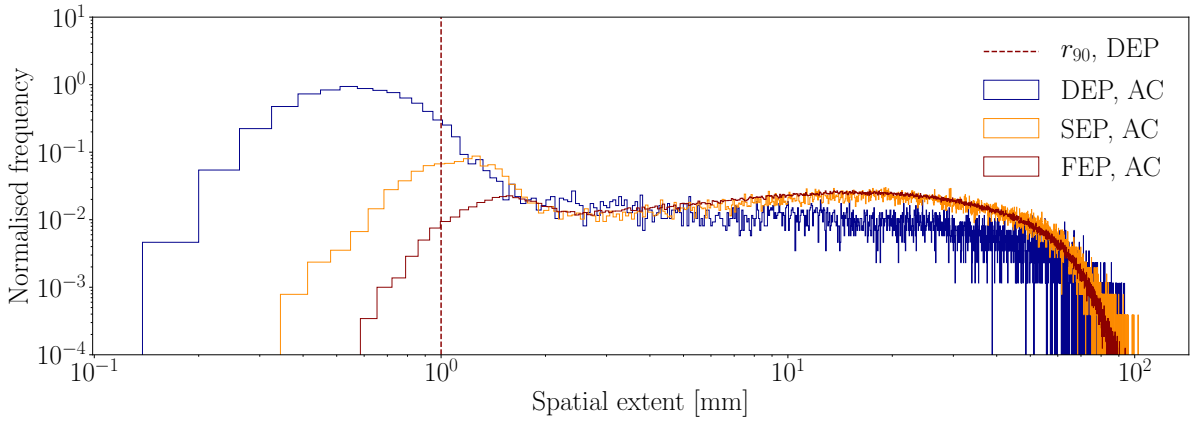

Figure 3.31: Histogram of the spatial extent of the simulated events in the DEP, SEP, FEP, after application of the AC cut. The selection windows are centred at the peak position, and span a width of ±2 FWHM of the DEP. The r_{90} cut value was determined from the DEP event study in Fig. 3.30.

Here we select the simulated events in the same manner as for the DEP study. This means we apply the AC cut, and constrain the energy interval to a window centred at the corresponding peak, with a width of ±2 FWHM of the DEP peak. We choose the width of the latter, as it has the smallest energy resolution among the three lines. From the plot we deduce a very strong reduction of the background-like multi-site SEP and FEP events when applying the spatial extent cut based on the r_{90} value determined above. We provide a quantitative overview of the survival fractions of simulated events in the DEP, SEP, and FEP in Tab. 3.9.

Peak	Events in peak region	Events after PSD	Survival fraction [%]
DEP	13871	6521	47.0
SEP	37444	650	1.7
FEP	468594	612	0.1

Table 3.9: Comparison of survival fractions in the DEP, SEP, FEP for the simplified pulse shape study based on the simulated spatial extent of events. The investigated spatial extents for each peak were selected in a window of ±2 FWHM of the DEP, centred at the mean values of 1592.5, 2103.5, 2164.5 keV for the DEP, SEP, FEP. The r_{90} cut value of 1.0 mm was determined in Fig. 3.30.

We emphasise that less than 2% of the background events survive the cut, whereas almost half of the simulated DEP events pass the criterion, implying a good enhancement of the signal-to-background ratio. The fraction of energy depositions in the DEP being rejected by the cut can be understood by the Compton continuum surrounding the single-site like DEP peak events. The presence of these events had already been compensated for in the estimation of the cut

value by fitting the skewed Gaussian peak profile simultaneously with a linear pedestal, as discussed in relation with Fig. 3.30. Hence, the performance of the background reduction cut in terms of the signal survival fraction is not altered relevantly. In comparison to real pulse shape discrimination studies performed by GERDA based on measured calibration data [100], we see qualitatively similar trends in terms of the survival of DEP events as opposed to the suppression of SEP and FEP events, however with strong differences in the actual numerical percentages. Given differently chosen analysis techniques and data selection windows between the simplified simulation approach and the actual PSD analysis, a direct comparison is expected to provide only a very rough idea.
Overall our simplified pulse shape studies put the background level estimations, in particular also after cut applications, on solid ground. In addition, the LAr veto cut has not been applied here, as we have not run any optical simulations. In the real experiment, the number of events accompanied by γ particles escaping a detector and causing scintillation signals in the LAr will thus only further decrease the background level. Considering all the cut uncertainties and the geometrical inaccuracies, let us underline that the actual numerical values are to be interpreted as order of magnitude estimates, though very promising ones. We suggest the collaboration to run detailed pulse shape simulations to obtain quantitatively more accurate background estimates in the future.
To conclude, our estimations prove that the background contribution in L-200 induced by calibrations is negligible, confirming the possibility of safely deploying the sources in the experiment. Given the strong suppression factors achieved via applying the cuts, a more refined evaluation, based on the actual L-200 geometry, more precise weight, activity, mass, and time duration estimates, does not seem to be required for the operations of L-200. However, a potential pitfall may appear. Let us remark that additional sources are currently produced for L-200 at LANL. These will be needed as the half-life of ^{228}Th is approximately 1.9 yr, which is shorter than the expected full run duration of five years. Later calibrations would take too long to provide sufficient statistics in all detector channels. Hence the new sources will be installed upon arrival and characterisation at LNGS, presumably in 2025. The neutron rate measurements with the LiI(Eu) detector should then be repeated with the new sources to ensure that a new production set does not cause higher neutron fluxes than expected and that our conclusions on the background level are not being altered.
Another important point to be made here is the case of L-1000. The 20 times lower background budget compared to L-200 and the different array configuration, potentially consisting of densely packed detector strings if built at LNGS, may result in a higher number and impact of neutron capture reactions. Therefore, we strongly recommend simulating the neutron capture processes in a realistic L-1000 geometry, guided by the procedure presented here for L-200. This will allow for an accurate evaluation of the expected background contribution. If an unacceptably high level of calibration-induced background counts is anticipated, there are a couple of possible solutions. One option is to extend the periods between calibrations from one to two weeks, which would effectively lower the ^{77}Ge background level by a factor of two, even without changing the calibration duration. Alternatively, the production process or the shielding method of the radioactive sources themselves could be improved to further reduce the neutron emission rate directly on the source side.

Chapter 4

Searches for signatures of new physics in GERDA

The GERDA and the LEGEND experiments offer a unique experimental configuration in the sense of combining the high efficiencies and unparalleled energy resolution of the HPGe semiconductor detectors with the active shielding benefits of the liquid scintillator. Achieving the ultra low background level as well as optimal efficiency and resolution performance implies potentially competitive sensitivities to BSM processes other than $0\nu\beta\beta$ decay. Here, we explore the sensitivity of GERDA to interactions caused by bosonic dark matter candidates, nucleon decays violating baryon number, and electron decays violating charge conservation. These physics channels are typically probed in large-scale experiments of multi-tonne exposures, clearly exceeding the reach of GERDA. However, as presented in this chapter, our results are among the strongest ones measured with semiconductor detectors. Based on our obtained sensitivities in GERDA, we will motivate the reach of LEGEND, potentially entering the regime of competitive constraints given the performance achievable in Ge detector-based experiments. We start with a brief introduction of the different physics models. Afterwards, we describe the full high-level data analysis for GERDA, covering the data selection including the signal and the background modelling of the data, the efficiency simulation, and the statistical methods. We then present the final GERDA results on these searches. We conclude with the projection of the sensitivity of L-1000 as deduced from the *measured* GERDA data.

4.1 New physics channels

4.1.1 Bosonic dark matter

One of the biggest puzzles in the physics of the Universe apart from the origin of the matter-antimatter asymmetry is the nature of the dark matter (DM), comprising around five times more energy density than baryonic, or ordinary, matter, as outlined in Chapter 1. Assuming the accuracy of the gravitational laws of Nature as described by General Relativity to hold, remarkable evidence for its existence on all astrophysical scales has been found. These range from, among many others, galactic rotational curve measurements, gravitational lensing and X-ray studies on the galaxy-cluster scale, up to investigations of the cosmic microwave background on the cosmological edge of the observable Universe. For a detailed discussion of the evidence for DM, we refer the reader to e.g. [148, 149]. The ideas to solve this riddle range over a similar scale, from tiny bosonic fields with masses from $\mathcal{O}(10^{-22}\,\mathrm{eV})$, via typical fermionic, weakly interacting massive particles (WIMPs) of masses around the GeV or TeV scale, all the way up to

macroscopic objects of astrophysical sizes such as black holes. Yet, no conclusive explanation for the nature of DM has been identified. Given this uncertain situation, many experimental efforts are ongoing, terrestrially in accelerator and fixed-target experiments, and even on satellites in Space [150]. Three main concepts are pursued. Collider experiments search for indications of missing energy upon the production of DM particles, which escape the detectors. In contrast, direct detection experiments seek to measure the momentum of matter particles when scattering with incoming DM particles. Indirect detection experiments instead look for excesses in the event rates of SM particle observations caused by the annihilation of DM particles. Despite a wide range of experimental approaches and programs, no convincing signature of any DM interaction with ordinary matter has been found in a laboratory [151]. With the excellent experimental performance, our experiments can contribute to the investigation of multiple potential solutions. One class of candidates to those interactions we are sensitive to are bosonic DM (bDM) particles with masses around the keV scale, sometimes referred to as superWIMPs. Considering this mass range, their coupling needs to be extremely small even compared to the weak scale in order to compose the entire DM in our galaxy, as e.g. pointed out in Ref. [152]. This explains why the authors thereof refer to these as superWIMPs. A low interaction rate though implies a large number density, which, as we will see below, makes it possible to probe at least a certain mass range relevant for DM in GERDA. These bDM particles are in fact a class of particles whose interactions, with SM fermions, can be described by different Lagrangians as listed below,

- scalar s, $\mathcal{L}_s = g_s s\ \overline{\psi}\psi$,
- pseudoscalar a, $\mathcal{L}_a = g_a a\ \overline{\psi} i\gamma_5 \psi$,
- vector V, $\mathcal{L}_V = g_V V_\mu\ \overline{\psi}\gamma^\mu\psi$,
- axialvector A, $\mathcal{L}_A = g_A A_\mu\ \overline{\psi}\gamma_5\gamma^\mu\psi$.

Here, ψ denotes a SM fermion field which the bDM particle of mass m_i couples to with coupling strengths g_i, where the index i denotes the corresponding particle class in terms of the behaviour under a parity transformation. All classes of particles may be observed because of their absorption inside a Ge detector when removing an electron from its atomic shell, exactly like for photons in a photoelectric effect [81]. In addition, they may scatter off an atomic shell electron, just like γ particles undergo Compton scattering, producing an electron-photon pair [82]. A sketch of the two channels is shown in Fig. 4.1.

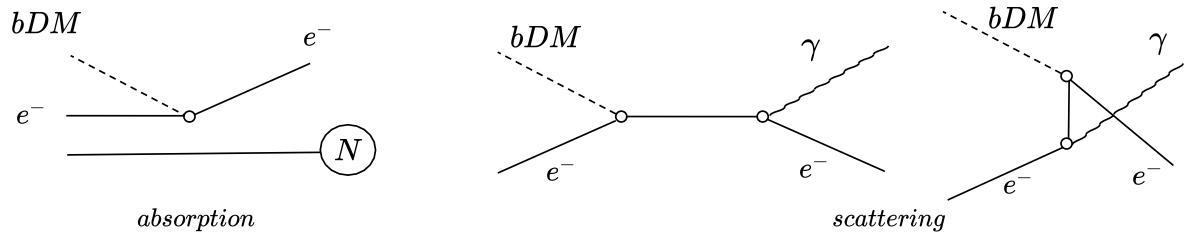

Figure 4.1: Left: sketch of the photoelectric-like absorption of a bDM particle by a Ge atom, represented by its nucleus N, causing the emission of an electron from its atomic shell. Right: sketch of the dark Compton scattering of a bDM particle with an atomic shell electron (s channel left, t channel right). Here, an electron and a photon are produced.

We refer to the latter channel as dark Compton scattering. The expected interaction rates for both processes for the case of pseudoscalar and vector particles were provided by DM phe-

nomenologists in Refs. [152, 153, 154]. We thus limit ourselves in this analysis to these two particle types, but in principle all results obtained could be reinterpreted to the cases of the scalar and/or axialvector candidates. We further limit the search to masses below $2m_e$, where m_e denotes the electron mass. Above this threshold, the decay into e^+e^- pairs prevents bDM to be stable over timescales of at least around the age of the Universe, a requirement to explain DM. The considered pseudoscalar candidates are sometimes referred to as axion-like particles (ALPs), particles that have similar properties to the axion, which is a pseudoscalar particle. The axion is the pseudo-Goldstone boson emerging from the spontaneuously broken Peccei-Quinn symmetry initially suggested to explain why the expected charge-parity violations in strong interactions are strongly suppressed [155], but is not necessarily further linked to our pseudoscalar bDM particle of interest. Vector bDM instead is sometimes referred to as dark photons, introduced analogously to SM photons as mediators of a dark U(1)' symmetry in hidden sector models. However, if dark photons are massive, they may either compose the entire DM, without the need for any additional fermionic candidate, or contribute to the dark matter density [156]. As dark photons mix with SM photons [157], they are potentially detectable. Due to couplings or mixing to SM photons, both our candidates are not fully stable over the age of the Universe [152]. In fact, it has been pointed out in Ref. [158] that the sensitivity of current experiments in the field is not sufficient to reach the regime where pseudoscalar particles can compose the entire DM for all masses below $2m_e$. We will come back to this argument when interpreting our results in Sec. 4.5. For the moment, we neglect the needed lifetime and focus on a generic search for any potential mass values below $2m_e$. The DM candidate needs to be dominantly cold, that is, non-relativistic, in order to explain the formation of structures such as galaxies and galaxy clusters [159, 160]. Taking into account a DM velocity[1] of today of $v \sim 10^{-3}$, we can approximate the incoming energy as the mass of the bDM, $E_{\mathrm{bDM}} = m_{\mathrm{bDM}}$. The approximate energy-mass equality implies a scan over our energy spectrum essentially poses a scan for bDM signals over certain masses. Using the formulae derived in Refs. [152, 153, 154], we expect the following interaction rates of pseudoscalar bDM of a given mass m_a when interacting via both absorption and dark Compton scattering in the energy spectrum of our Ge detectors,

$$R_a = \rho_{\mathrm{DM}} \frac{3 g_{ae}^2 m_a}{4 e^2 m_e^2} \sigma_{\mathrm{PE}}(m_a) \times \epsilon_{e^-} + \frac{n_e \rho_{\mathrm{DM}} e^2 g_{ae}^2 m_a (m_a + 2m_e)^2}{\rho_{\mathrm{t}}\ 16\pi\ m_e^2 (m_a + m_e)^4} \times \epsilon_{e^-\gamma}\ . \tag{4.1}$$

For the interactions of a vector bDM particle of mass m_V, we expect a rate of

$$R_V = \frac{\rho_{\mathrm{DM}}}{m_V} \frac{\alpha'}{\alpha} \sigma_{\mathrm{PE}}(m_V) \times \epsilon_{e^-} + \frac{n_e \rho_{\mathrm{DM}} e^4 \frac{\alpha'}{\alpha} (m_V + 2m_e)(m_V^2 + 2m_e m_V + 2m_e^2)}{\rho_{\mathrm{t}}\ m_V\ 24\pi\ m_e^2 (m_V + m_e)^3} \times \epsilon_{e^-\gamma}\ . \tag{4.2}$$

Here, ρ_{DM} and ρ_{t} denote the density of the DM in our galaxy and of the target, respectively, α and α' are the SM U(1) and the dark U'(1) fine structure constants, e and n_e are the electric charge and the number of electrons in the target, and σ_{PE} denotes the photoeletric cross section. The coupling of pseudoscalar bDM to electrons g_{ae} corresponds to g_a in the list of Lagragians when limiting ourselves to electrons, and the kinetic mixing strength α'/α is related to g_V in the vector bDM case via $\alpha'/\alpha = e^2 g_V^2$. In both equations, the first summand denotes the expected absorption rate, and the latter one the scattering contribution. Hence, $\epsilon_{e^-,e^-\gamma}$ denotes the efficiency of detecting either an e^-, for the absorption case, or an $e^-\gamma$-pair, for the scattering case, respectively, at a given energy. The electron energy, which is simply $E_e = m_{\mathrm{bDM}}$ for absorption, is given as

$$E_e = \frac{m_{\mathrm{bDM}}^2}{2\,(m_e + m_{\mathrm{bDM}})}\ , \tag{4.3}$$

[1] As mentioned in Chapter 1, in this thesis we always use natural units, i.e. the speed of light is set to unity.

in the scattering case, whereas the photon energy reads

$$E_\gamma = m_{\mathrm{bDM}} - E_e = \sqrt{E_e^2 + 2m_e E_e}\ . \tag{4.4}$$

We remark that the functional dependencies of the rates on the bDM mass for the several particle candidates and processes can be explained by the different cross sections. Naively, we would expect the rates to scale inversely with m_{bDM}, as the flux of the incoming bDM particles that can interact is proportional to the number density. As the overall rate is not only proportional to the flux, but also depends on the interaction cross section, this simple inverse proportionality is modified if the latter is not independent of the bDM mass. For the pseudoscalar bDM, the absorption cross section scales quadratically with the bDM mass [152], hence we see an overall linearity in m_{bDM} (see Eq. 4.1). In the case of the scattering, the scaling of the rate is linear as well for $m_a \ll m_e$, but approaches an inverse proportionality in the opposite regime $m_a \gg m_e$ [153]. For the vector bDM, the inverse dependency of the rate with the mass is retained in the absorption case, given a mass-independent cross section [152]. Similarly, in the asymptotic mass regimes $m_V \ll m_e$ and $m_V \gg m_e$ also the rate induced by the scattering process scales as $1/m_{\mathrm{bDM}}$ [153]. Given the more complicated dependencies of the cross sections on the bDM mass derived by calculating the Feynman diagrams, as done in Refs. [152, 153, 154], the simple functional behaviour does not hold for any bDM mass. We show a plot of the relative contribution of absorption and dark Compton scattering for fixed bDM mass, for both the cross sections and rates, in Fig. 4.2, adopting the formulae from [154].

To take into account a realistic experimental environment, we indicate additionally the rates for different efficiencies $\epsilon_{e^-\gamma}$, where the efficiency of the pair is assumed to describe the full combined energy measurement. For simplicity of the comparison, we set ϵ_{e^-} to unity. Looking at the plots, we see that the relative strength strongly depends on both the relative efficiency and the energy range. Whereas the absorption process dominates in the lower keV range, the scattering process becomes more and more important approaching the MeV regime. A former study in GERDA [161] already analysed the absorption process using Bayesian techniques, with a null result over the entire analysed mass range of 60 keV to 1 MeV. Similarly, several other experiments have investigated this channel. For an overview of recent experiments, we refer readers to Ref. [162]. Leading results in different mass ranges were obtained by EDELWEISS [163], MARONA DEMONSTRATOR (MJD) [164], and XENONnT [165], all with null results. Let us emphasise that no experiment has included the dark Compton scattering in the interpretation of their results, as motivated in Refs. [153, 154]. This means that our current study poses the very first combined interpretation performed by an experiment. Our findings above let us expect relevant sensitivity improvements towards higher energies.

Let us now briefly discuss the expected signal shapes to identify our analysis strategy. When a bDM particle is absorbed by the germanium atom, an electron is ejected. With a mean free path of around one mm in a semiconductor, the electron gets fully absorbed in the majority of the cases. Since mostly all of the induced charge would be collected, a full energy deposition corresponding to the incoming energy of the DM particle is expected. In other words, we will need to look for a peak-like signature in the energy spectrum located at m_{bDM}. In the case of the scattering, the final-state electron would also most likely be absorbed immediately. In contrast, the γ, having a mean free path of around 0.1 to 2 cm depending on the precise energy in the relevant range [153], can either fully escape the source detector or be partially or fully absorbed within. As the dimensions of the Ge detectors are around 3 to 10 cm, it is not obvious to estimate which scenario dominates. Hence, dedicated simulations were run by our colleagues from University of Padua (UniPD). These revealed that the highest efficiency is obtained for the combined detection of the electron and the γ, meaning again for the detection of the full incoming

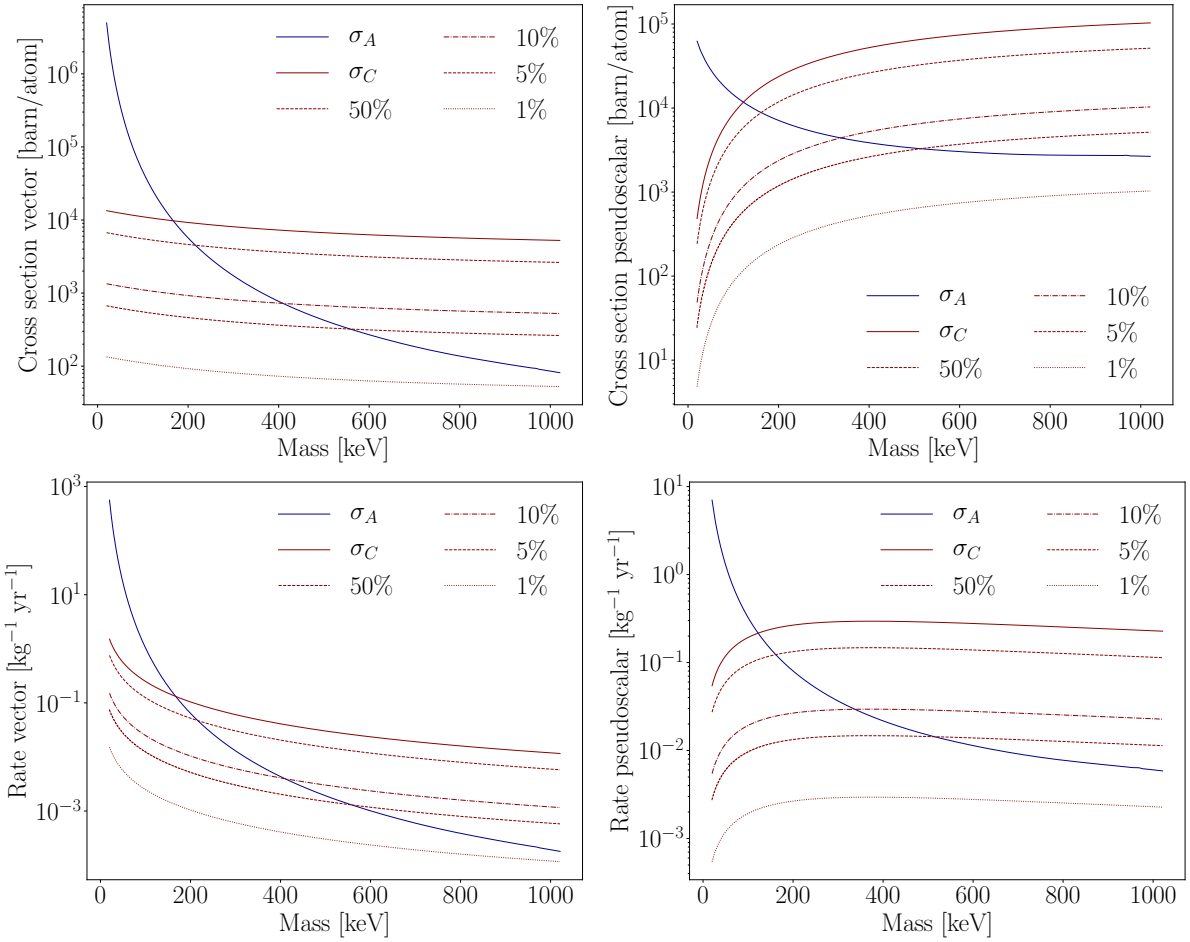

Figure 4.2: Top, left: plot of the crosssection of the photo-electric like absorption σ_A (blue), and dark Compton-like scattering σ_C (red) processes for vector bDM. Top, right: the same plot for pseudoscalar bDM. For illustration purposes, the couplings g_a and g_V are set equal to unity in both figures. Bottom, left: expected rates induced by both scattering and absorption processes in a Ge-based experiment for vector bDM. Bottom, right: the same plot for pseudoscalar bDM. For illustration purposes, we set $g_a = g_V = 10^{-12}$, corresponding to a very rough approximation of the sensitivity of GERDA, as we will see in Sec. 4.5. Realistic experimental scenarios are indicated by downscaling the detection efficiency for the Compton channel with respect to the absorption process (dashed, dotted). The former requires tagging of both β and γ particles, whereas in the latter case only the electron needs to be detected. The figure is adapted from Ref. [154].

energy, exactly as in the absorption case. We will discuss more details on these simulations in Sec. 4.3. With this in mind, we will apply the same technical analysis to search for an induced full energy peak, which can then be interpreted as the combined rate of both interactions, for either bDM candidate. As a last remark, regarding the methods, our analysis poses the first bDM search in GERDA that applies a full Frequentist statistical treatment.

4.1.2 Nucleon decay

Having addressed two major unsolved shortcomings of the SM, namely no viable mechanism to explain the matter-antimatter asymmetry of the Universe, see Chapter 1, and the nature of DM, as outlined in the previous section, we can now turn to a process potentially providing input on both simultaneously. An alternative solution to the DM puzzle via dark or a mirror neutrons has been suggested in, e.g. Refs. [166, 167]. Such mirror neutrons are further inspired by an experimental anomaly between measurements of the neutron lifetime with neutron beam and with bottle experiments. The former method relies on the counting of protons produced in the decay of neutrons in-flight, the latter on the exponential law, which the decay of ultra-cold neutrons trapped in a bottle follows [168]. The discrepancy between these measurements has been standing on the $\sim 4\sigma$ level for years, and a viable explanation in terms of a systematic effect has not yet been found [169, 170]. In the simplest model explaining the discrepancy, a new fermionic mirror particle χ mixes with its SM counterpart according to the Lagrangian

$$\mathcal{L}_{\mathrm{n}} = \theta \ (\overline{n}\chi + n\overline{\chi}) \ . \tag{4.5}$$

This mixing could induce invisible decays of the SM neutron. Extensions with two mirror neutrons χ and $\tilde{\chi}$ are proposed in Ref. [166] as well. Example processes of a neutron disappearing and producing dark neutrons, potentially accompanied by a γ particle, are sketched in Fig. 4.3.

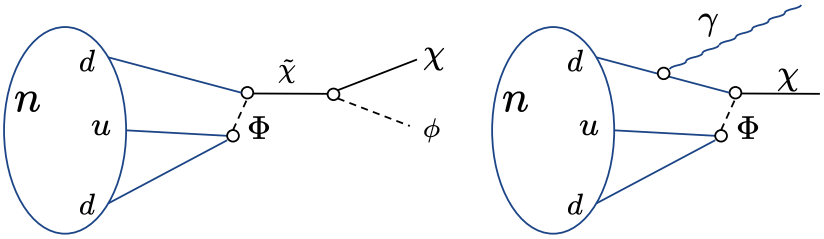

Figure 4.3: Left: illustration of an invisible neutron decay into dark sector particles. Right: illustration of a semi-visible neutron decay into dark sector particles. In both sketches, χ and $\tilde{\chi}$ denote dark fermions, i.e. one or two mirror (dark) neutrons, and Φ and ϕ are dark (mediator) bosons. The sketches follow the principles of the illustrations published in [171].

Adding to the standard beta decay of neutrons, such a mixing or oscillation may explain why the bottle experiments observe a reduced neutron lifetime. Let us remark that in Ref. [171], it was pointed out that phenomenological considerations in combination with existing experimental constraints imply a limited viability for these dark decays in heavy elements such as Ge as operated in GERDA. The dark neutron decays inside the latter must be suppressed to ensure the stability of SM protons and light nuclei such as ^{9}Be [172], assuming a coupling of the dark neutron to the proton, plus a similar behaviour of bound and free neutrons with respect to dark decays. However, a potential annihilation of a bound neutron with an incoming dark (anti-)neutron from the galactic bulk may still cause the disappearance of a neutron also from a heavy nucleus [172]. Moreover, let us underline that several BSM models predict single neutron decays independently of any stability and coupling conditions related to dark sector physics. An example of such a BSM decay channel is a neutron decaying into three neutrinos, which is e.g. predicted by models of extra dimensions [5], or two or three neutrons decaying simultaneously, as illustrated in Fig. 1.2. In these scenarios, the disappearance of one or more neutrons implies the violation of baryon number. This is of particular relevance, as we outlined in Chapter 1. Baryon number violation is one of the fundamental criteria as proposed by Sakharov to understand the

matter-antimatter asymmetry of the Universe [5, 173]. Similarly, also single or multiple proton decays would imply baryon number violation, and are hence of similar interest.
Any of these processes would open a large window of potential experimental signatures, as partially sketched in Ref. [171]. For instance, the disappearance of a nucleon, neutron or proton, out of a Ge diode enriched in ^{76}Ge, would cause an energy release at least as large as the minimal nuclear separation energy in this isotope, which is approximately 9.5 MeV [174]. This release may occur via a γ-cascade emission or induce further α or additional nucleon emissions. Searching for such a large energy deposition is very difficult, as the precise emission properties, more precisely the angular distribution and the type and the number of particles of the induced cascade, are not unique. Consequently, we cannot simply simulate these processes. However, if we restrict ourselves to the simplest case, a single neutron disappearing from the ^{76}Ge nucleus, we expect the production of ^{75}Ge. This isotope undergoes beta decays in several branches to ^{75}As, with a half-life of 82 min and a maximal Q-value of 1183 keV [175], as illustrated in Fig. 4.4.

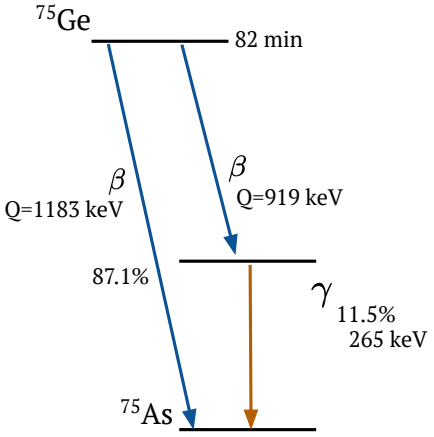

Figure 4.4: Simplified scheme of the decay of the isotope ^{75}Ge to ^{75}As, adapted from Ref. [175]. Whereas the direct beta decay of ^{75}Ge to the ground-state of ^{75}As is difficult to clearly identify in the detectors of GERDA due to the strong ^{39}Ar background, the beta decay of ^{75}Ge to an excited state of ^{75}As accompanied by γ-emissions cause potential coincident events in two detectors, strongly reducing the background. The dominant γ-transition occurs at an energy of approximately 265 keV, with a branching ratio of 11.5%. The branching ratios of all other γ-accompanied transitions is below percent-level, and hence neglected.

We could now look for single subsequent β particles in the HPGe detectors from this secondary decay process. However, the presence of ^{39}Ar in the natural argon of GERDA and of the $2\nu\beta\beta$ decay continuum, as detailed in Sec. 4.2, implies that this method is strongly background dominated. Instead, we look for a β particle with a maximal energy of $Q_\beta = 919$ keV in coincidence with a γ particle of 264.7 keV, where the latter is released during the ^{75}As de-excitation. The branching ratio of this process is 11.5%, which is the dominant γ-accompanied transition. Tagging the electron in its source detector and the photon of this known energy in a neighbouring detector of the detector array of GERDA, we obtain a distinctive, clear signal signature in the multiplicity 2 (M2) data. In addition, as opposed to the initial cascade, this signature can be readily simulated in the experimental configuration to understand the detection efficiency. Note that the half-life of 82 min is too long to make an additional delayed tagging of a potential immediate energy deposition, meaning directly after the neutron decay, feasible. The M2 coincidence method has been searched for in Ref. [176] using the Phase II pre-upgrade data only. Here, we extend the study to the full Phase II data set and further explore the impact

of a full Frequentist treatment, as opposed to Bayesian methods used in the aforementioned work. A notable side feature of this search has not been pointed out before. It is interesting to understand that the very same analysis channel can not only be applied to search for the decay of a single neutron but also for that of a single proton. If a proton decays, assuming a branch without the emission of accompanied nucleons, the produced ^{75}Ga isotope undergoes β decay to ^{75}Ge with a half-life of around 126 s, and a branching ratio of 100% [177]. Note again that this timing window is too long to apply an additional delayed tagging. Given that both neutron and proton decay are probed with the single coincident ∼265 keV M2 signal, we will refer to this particular analysis channel as nucleon decay. We illustrate the corresponding decay schemes for the two types of nucleon decays within our experimental context in Fig. 4.5, neglecting particles invisible to the detectors such as neutrinos or dark sector particles.

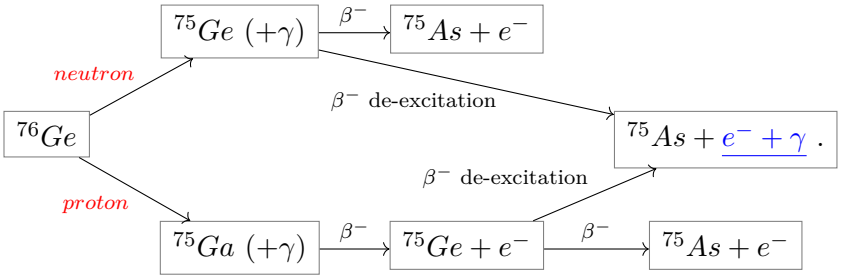

Figure 4.5: Flow chart of the decay schemes of single neutron and proton decays of ^{76}Ge, indicating the subsequent isotope transitions. In this analysis, we only consider the $\beta\gamma$-pair from the subsequent ^{75}As de-excitation, highlighted in blue.

The landscape of searches for such baryon number violating invisible single nucleon decays is dominated by large-scale experiments. To list some examples, the KamLAND experiment obtained a neutron lifetime limit of 5.8×10^{29} yr [178], the SNO [179], and SNO+ experiments [180] neutron (proton) lifetime lower limits of $1.9(2.1) \times 10^{29}$ yr and $2.5(3.6) \times 10^{29}$ yr, Super-Kamiokande a proton lifetime lower limit of 3.6×10^{33} yr [181], and Borexino a neutron (proton) lifetime lower limit of $1.8(1.1) \times 10^{25}$ yr [182], all at 90% CL. Of course, all these experiments have orders of magnitude larger exposures than GERDA. Correspondingly, their sensitivity is clearly beyond our reach. The motivation to perform this analysis is thus rather to establish a stability condition for the individual nucleons within the isotope ^{76}Ge, in a manner independent of the final state particles. This isotope-specific property has not yet been analysed in the literature. Yet, it might be relevant as the behaviour of bound nucleons may not be comparable between different nuclei. Similar ^{76}Ge-based studies conducted by MJD [183] and GERDA [184][2], only consider di- and/or tri-nucleon decay channels, but not single nucleon decays.

4.1.3 Electron decay

Historically, perhaps *the* most essential and fundamental building block of the SM of particle physics is the conservation of the electric charge. The well-known associated SM U(1) gauge symmetry, after electroweak symmetry breaking, protects, among others, the stability of an electron. Given the importance of this cornerstone of physics, it needs to be tested and confirmed experimentally with the highest precision possible. Indeed, several laboratory experiments focused on neutrino and DM searches also probed the electron stability, e.g. the Borexino experiment

[2]Paper accepted by EPJC for publication.

achieved a world-leading lifetime lower constraint of 6.6×10^{28} yr at 90% CL [185]. In addition, Ge-detector-based experiments with their excellent resolution properties have conducted such an analysis. For example, in Ref. [186] a lower lifetime limit of 9.4×10^{25} yr at 90% CL had been stated by Klapdor-Kleingrothaus and others, using data from the Heidelberg-Moscow experiment. Also, MJD [187] and the EDELWEISS-III experiment [163] published electron decay searches, both obtaining a lower limit of 1.2×10^{24} yr at 90% CL. For HPGe detector experiments, two search methods come into mind. The disappearance of an electron via the invisible channel $e^- \to 3\nu_e$ would only be detectable via the released binding energy of the electron from its respective atomic shell E_b, as the final state neutrinos would escape the detector. A semi-visible channel would be the two-body decay $e^- \to \nu_e\gamma$, with an expected γ particle energy of $m_e/2 \approx 255.5\,\mathrm{keV}$. Here, we can safely approximate the neutrino mass as vanishing and assume the decay of a free electron. In contrast, a bound electron would cause a γ particle energy of $E_\gamma = (m_e - E_\mathrm{b})/2$. For the GERDA experiment, the former channel is not accessible, as in the low energy region around the maximal binding energy in Ge of $\sim 11.1\,\mathrm{keV}$ [188] too many events appear in the spectrum. The origin of these events has not been clarified. Assumptions are that these might be caused by slow pulses from ^{39}Ar decays, and/or noise from the readout electronics components. The latter process though would fall nicely into our bDM energy range of interest, and we thus include this electron decay channel to our searches. When probing the signature γ-line, we will need to be aware of both the release of the relevant atomic binding energies, as well as the Doppler broadening as an effect of the different electron binding energies, as outlined in Ref. [186].
Assuming an electron decays inside a HPGe detector, and the photon remains contained, i.e. its total energy will be measured, we will see a signal at an energy of $E_\mathrm{t} = E_\gamma + E_\mathrm{b}$. The binding energy would be released via X-rays or Auger electrons. Given the low energy of $\mathcal{O}(10^{-3})$ to $\mathcal{O}(10^{1})\,\mathrm{keV}$, these would be contained in the detector as well, hence the shift upwards from $m_e/2$. In the case of a decay outside of a Ge detector, we could only measure E_γ. Let us remark that we only focus on the stability of the Ge atoms directly, i.e. we do not take into account external contributors such as materials from off-detectors, the detector holders, the nylon shrouds, the cryogenic liquid, or the contribution of the dead layers. Given the low efficiency of these materials to cause full energy peaks inside the active detector volume, especially after applying GERDA's background reduction cuts, and in particular the LAr veto cut, as described in Sec. 2.2, these simplifications are not expected to alter our results relevantly. Additional contributions may only improve upon these, but will not worsen the constraints, making our analysis conservative. To determine the expected mean energy measured in a detector, we calculate the weighted sum of the γ particle energy plus the binding energy contributions over all eleven different atomic shells i, where the weights are given as the individual electron occupancy n_i. We obtain a total energy estimate of approximately

$$E_\mathrm{t} = \frac{\sum_{i=1}^{11} n_i E_{\mathrm{t},i}}{32} = 256.03\,\mathrm{keV}\ , \tag{4.6}$$

i.e. a small shift from the expected energy of $m_{e^-}/2$ for an electron decaying freely. Here, the individual binding energies for all atomic shells are taken from Ref. [188]. Regarding the broadening of the signal peak, we model the full line I as a weighted sum of Doppler broadened Gaussian functions centred at the individual $E_{\mathrm{t,i}}$,

$$I(E) = \sum_{i=1}^{11} \frac{n_i}{\sqrt{2\pi\sigma_i^2}} e^{-\left(\frac{E-E_{\mathrm{t},i}}{\sqrt{2}\sigma_i}\right)^2}\ . \tag{4.7}$$

Again the weights are the electron occupancy per atomic shell n_i. Note that the energy contributions of the crystal structure of the Ge detectors to the binding energies are typically $\mathcal{O}(1\,\mathrm{eV})$, originating from interatomic forces. As the broadening of the Doppler-broadened peak is strongly dominated by the atomic shells with binding energies on the order of magnitude of one to ten keV, the contributions from the crystal bonding are expected to be of minor relevance. If we assume the electrons in the atom are in thermal equilibrium, we can apply the virial theorem, $E_{\mathrm{kin.}} = -\frac{1}{2}\, E_{\mathrm{pot.}}$, cf. [186], to estimate the widths of the individual line contributions as

$$\sigma_i = E_{\mathrm{t},i}\sqrt{\frac{k_{\mathrm{B}}T_i}{m_e}} \approx 0.0442\; E_{\mathrm{t},i}\sqrt{\frac{E_{\mathrm{b},i}}{\mathrm{keV}}}\;. \tag{4.8}$$

Here k_{B} denotes Boltzman's constant, and T the electron temperature, and we again use natural units. Note that in [186] a marginally larger constant of proportionality of 0.0447 had been stated, and the value above was found by us. With this analytical model, we can then determine the full signal peak shape as expected to be seen in the realistic experimental scenario as the convolution of the Doppler broadened line with the detector resolution broadening curve. In Sec. 4.2, we will explain how we model the latter contribution. In Fig. 4.6, we sketch the Doppler broadening of each atomic shell according to Eq. (4.8) and the full theoretical line shape according to Eq. (4.7). We also indicate the resolution curve modelled as a weighted Gaussian mixture distribution with the weights given as the exposures of different detector sets and phases (Sec. 4.2) and the final signal shape as obtained by convolving the Doppler broadened and the resolution broadened peaks.

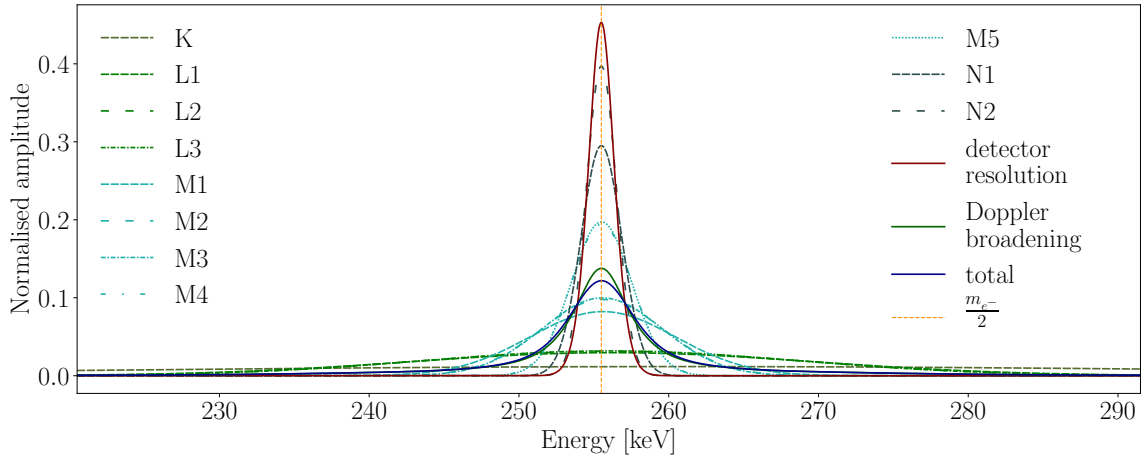

Figure 4.6: Doppler-broadened line shape for the semi-visible decay of an electron. The Gaussian contributions from the different atomic shells, also depicted individually as K, L1-L3, M1-M5, N1-N2, are taken into account. The broadening caused by the limited energy resolution as modelled with a Gaussian mixture model is shown for comparison (dark red), cf. Sec. 4.4. The full line shape (dark blue) is given as the convolution of the mixture model with the Gaussian functions of each atomic shell.

Having discussed the Doppler broadened line shape of the semi-visible electron decay signature, we conclude with a brief explanation of why these effects are negligible for the nucleon or the bDM analyses, apart from the broadening resulting from the detector resolution. In the case of the γ-ray signature from the ^{75}As de-excitation upon the ^{75}Ge decay, the intrinsic broadening is, according to the Heisenberg uncertainty principle, fully negligible, as already explained in

Sec. 2.1. There, we mentioned that excited nuclear states do not induce a broadening beyond around 0.1 meV. Also, the contributions of the nuclear recoil and the Doppler shift caused by the normal thermal motion of the nucleus are negligible, as the ^{75}As nucleus is contained inside a Ge crystal, and no resonance effects are induced. This means that the broadening will be fully suppressed by the crystal mass. In the case of a bDM particle interacting in our HPGe detectors, the situation is less straight-forward. Intuitively, we would expect a Doppler broadening induced by the relative motion between Earth and the galactic stream of order $v \sim 10^{-3}$, in natural units. Looking for bDM masses of up to $\sim 1\,\text{MeV}$, this seems to induce a measurable line broadening up to the keV level, i.e. comparable to the scale of the HPGe detector resolution. However, we also have to consider the thermal motion of the bDM, i.e. Eq. (4.8) comes into play. As in the case for the atomic shells described above, we can apply the virial theorem, because the DM particles in our galaxy are, forming a gravitationally bound object and not interacting otherwise, assumed to be on average in an equilibrium state. Let us remark that gravity is a conservative force, a necessary condition for the virial theorem to hold, cf. [189]. These considerations imply $k_\text{B}T \sim mv^2$, up to constants of order unity. Hence, the radicand in Eq. (4.8) scales with v^2. In other words, an additional suppression factor linear in the velocity appears, and the overall peak will be consequently broadened by maximally $\sim \mathcal{O}(v^2) \times 1\,\text{MeV} = \mathcal{O}(1)\,\text{eV}$. These back-of-the-envelope estimates confirm that we can neglect any broadening effect other than the detector resolution when searching for the bDM or single nucleon decay signatures.

4.2 Data modelling

All of the signal channels introduced above lead to full energy depositions within the HPGe detectors. Limiting ourselves to these, and neglecting other event topologies such as partial energy depositions and/or hits in surrounding materials, we search for monoenergetic excesses peaking on top of a continuous background spectrum. Other event topologies are constrained by efficiency considerations discussed below. In Fig. 4.7, we provide a sketch of the different probed signal signatures for all the new physics as expected to be seen in the Ge detectors.

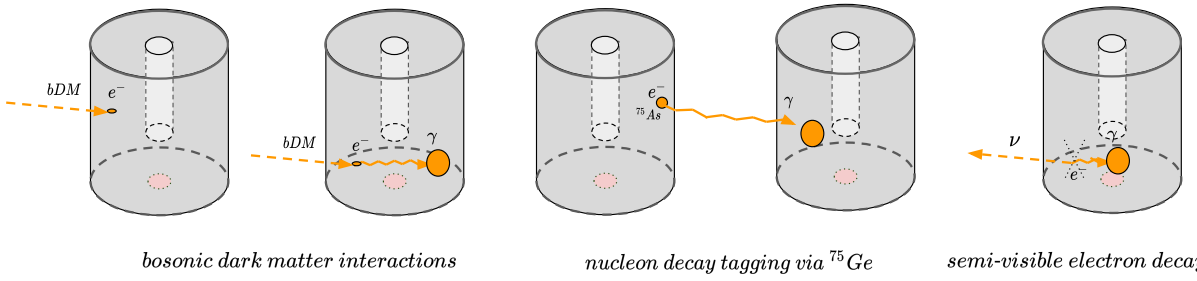

Figure 4.7: Sketch of the signal signatures sought-after in these exotic physics analyses. Searched for interaction patterns of incoming bDM particles cause an full energy depositions by absorption, or Compton scattering (left). A single nucleon decaying out of the ^{76}Ge nucleus is followed by the production of ^{75}Ge, which can subsequently cause a β event in the source detector plus a coincident γ event in a neighbouring detector (centre). The semi-visible electron decay can be searched for by looking for the produced γ-ray (right). The illustrations are adapted from sketches produced by C. Scaffidi.

Looking for such monoenergetic peaks at a probed energy in a given data set, we need to take into account the line broadening resulting from the limited energy resolution of the

measurement. As done in Sec. 3.3, we model the peak shape as a Gaussian signal profile under the assumption of a symmetric line shape for full charge collections. For the γ-line induced by the semi-visible electron decay channel, the Doppler effect leads to an additional line broadening, which, as described above, dominates the peak width. Given that every data set considered in these GERDA analyses is non background free, we need to simultaneously fit our Gaussian signal distribution of whichever width to be determined, together with a distribution modelling the background continuum. Details of how we model the background depending on the specific data set are given in the following subsections.

4.2.1 Multiplicity-one data set

Let us remind ourselves that for the bDM analysis, we consider energy depositions with total energies below $2m_e$, and for the electron decay signal of approximately $m_e/2$. That is, compared to the $0\nu\beta\beta$ search, we consider low and medium energy territories only. To suppress unphysical or external background, we apply basic data quality cuts removing discharges, baseline, and test pulser events, and high-level analysis cuts, i.e. the muon veto, multiplicity equals unity (M1), and the LAr veto cuts, as discussed in Sec. 2.2. The GERDA collaboration also performed PSD analyses for the $0\nu\beta\beta$ search. As these were specifically designed and especially validated only for the MeV range, we did not apply any PSD cuts here. Given the count number of events of orders 10^3-10^5 in this low and medium energy regime of GERDA in the selected data set, we clearly expect a binned analysis to lead to a sufficient accuracy. We will put this assumption on a solid grounding further below. We need to use a fit function to model the background, as the GERDA background model after the LAr veto cut does not fully cover the energy range of interest, and cannot reproduce the observed ^{39}Ar-dominated spectral shape at lower energies [110]. We show a plot of the GERDA data in the energy range below around 1 MeV in Fig. 4.8.

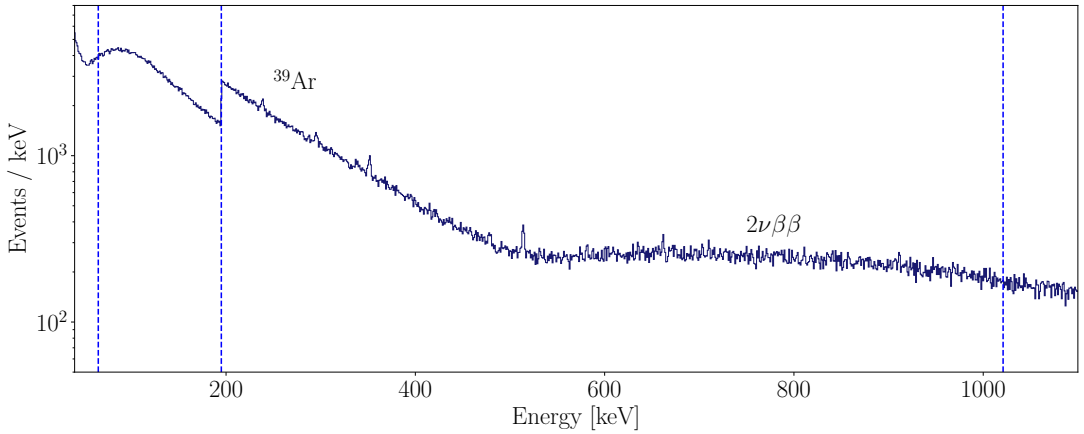

Figure 4.8: Plot of the GERDA Phase II and II+ low and medium energy data, applying data quality, muon veto, LAr veto, and M1 cuts. The dominating continuous background contributions, ^{39}Ar and $2\nu\beta\beta$ at lower and medium energies, respectively, are indicated. The vertical lines indicate the bDM search range, with a transition between the low and the medium data set at 195 keV.

We merge the data of Phase II and Phase II+ of all three enriched detector types, excluding the ones of natural Ge composition, as the spectrum of the latter ones more γ-ray background peaks are visible. After merging data sets, the background shape is a continuous function, whereas the signal shape is a mixture of the individual Gaussian distributions for each detector.

In the GERDA Phase II calibration paper [87], it is shown that the energy resolutions of the BEGe and ICPC detectors agree very well within their respective group. The combined resolution can hence be accurately described by a combined Gaussian model (see Fig. 5 therein). Only within the Coax detectors two diodes that have energy resolutions differing more strongly from the others have been observed. Given the scale of energy resolutions of $\mathcal{O}(1\,\mathrm{keV})$ and the minor impact of two detectors only, we apply a Gaussian mixture model with weights based on the exposure of each detector type, split by data-taking phase, instead of using a Gaussian mixture model consisting of individual components for each single detector channel. Note that we do not use a simplified single Gaussian distribution with exposure weighted variances to model all detector types and measurement phases at once. Here, exposure weighing refers to the exposures of the three individual detector types, as well as the two data-taking phases. We found that the difference between the simplified single Gaussian model and the accurate mixture distribution over the three detector types and two phases is rather small, but we obtain a slightly more peaked profile with the mixture model, implying a slightly increased sensitivity, as indicated in Fig. 4.9. We thus use the more complicated but more precise peak shape.

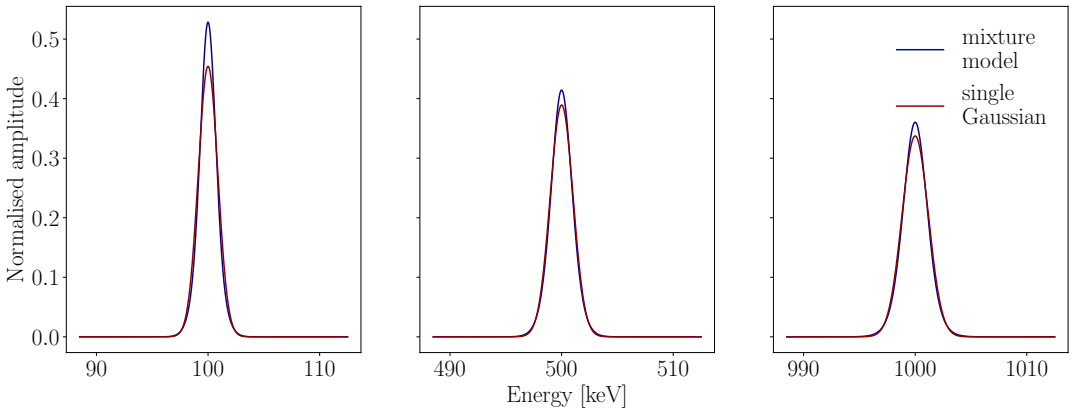

Figure 4.9: Comparison of the signal shape modelled with a Gaussian mixture model for the three detector types and the two phases, versus a simplified single Gaussian with an exposure-weighted variance, for energies of 100, 500, and 1000 keV. The widths of the peaks are very similar, but the appropriate mixture model is slightly more pronounced around the peak centre, implying a slightly improved sensitivity.

The systematic uncertainties on the energy scale are approximately 0.2 keV [87, 122], i.e. much smaller than the energy resolution of around 1 keV. We hence choose 1 keV bin width in correspondence with the approximate average energy resolutions of the detectors. We emphasise that the procedure followed here mitigates the parallel determination of signal-shape-related nuisance parameters, as we fix both the energy scale and the energy resolution for each detector type per data-taking phase as measured from calibration data [87]. Considering again the energy resolutions on the order of around 1 keV, it becomes evident that a window of 25 keV is clearly sufficient to contain all potential signal events, and in addition to constrain the background shape for the non-broadened peaks. Let us note that we tested that widening or reducing the window within a small range does not effect our later results on more than percent level. This holds as long as the window width remains within a reasonable range on the order of $\mathcal{O}(10\,\mathrm{keV})$ to run the combined fits of signal and background properly. Looking at the data shown in Fig. 4.8, this indicates that for small windows of interest, a linear distribution suffices to constrain the smooth $2\nu\beta\beta$ dominated background regime above approximately 500 keV.

However, in the region below the data taking threshold change at 195 keV, a second order polynomial fit is more advantageous for sufficient accuracy of the background estimation even in small windows because of the broad peak of the ^{39}Ar continuum. In the transition region of intermediate energies, a linear distribution has again been considered sufficient. These orders of the polynomials align with the previous bDM analysis of GERDA [161, 190], and we consider the technical details applied there reasonable. Note however that we symmetrise the search window to a centred 25 keV-wide region. For the electron decay channel, we choose a window ranging from $[196, 316]$ keV in order to take into account the Doppler broadening, centred at the expected Doppler shifted value of 256 keV.
In addition, let us point out that a fixed physics-motivated background model seems to be a more natural approach instead of a sliding window with individual polynomial fits for each signal model. The parameters of the polynomials should be correlated, and this correlation had been clearly neglected in the past. In contrast, here we develop a full background model for low energy M1 data after the LAr cut of GERDA, based on the actual dominating physics processes, the ^{39}Ar and the $2\nu\beta\beta$ components. This allows us to apply both approaches of a physically motivated model and the pure polynomial background fitting, and to determine all subsequent results in two independent ways. Given the physical origin, we choose the empirical background model as introduced below to evaluate our physical constraints. We study the systematic uncertainties induced by the background model via a comparison to the polynomial fitting approach in a sliding window. Let us summarise the technical details of the data modelling below.

Technical details of the data modelling

- Data: merged data set of Phase II and II+, from three enriched detector types
- Cuts: quality, muon veto, LAr veto, M1
- Histogram: integer-centred bin width of 1 keV
- Fit window: width of 25 keV, centred at assumed bDM mass, one search per keV;
  fit window from 196 to 316 keV for electron decay channel
- Background: empirical background model used for main analysis,
  polynomial background (first/second order above/below 195 keV) used for study of systematic uncertainty
- Signal: modelled as Gaussian mixture model; convolution of Gaussian mixture model and Doppler profile for electron decay

Empirical background model

In order to model the M1 data background continuum with a physically motivated distribution, we use an empirical fit model. It consists of two main contributions arising from the $2\nu\beta\beta$ and the ^{39}Ar background, which we parametrised by fitting directly to the data. The first parametrisation is a polynomial function of order 10, constraining the $2\nu\beta\beta$-dominated background regime beyond approximately 500 keV. Following the well-known, traditionally applied Primakoff-Rosen distribution [44], we force the polynomial to vanish in the origin and at the Q-value according to the phase space factor as introduced in Sec. 1.4. This leaves five parameters free, as opposed to only one free amplitude in the Primakoff-Rosen function. We will motivate below that more precise, recent calculations as e.g. applied in Ref. [191] are not needed to model our data. The fixed

polynomial function fitted directly to the data incorporates all experimental effects such as resolution broadening or detection efficiency automatically. The second distribution is a modified β distribution, which had been proposed in Refs. [192, 193] to provide more freedom when fitting skewed spectra with some potential distortions not allowing for direct β distribution fits. We applied the modified distribution here to model the ^{39}Ar-dominated low energy region between approximately 50 and 500 keV. We considered the modified β distribution a very reasonable choice as the original ^{39}Ar emission is a β decay. A modification of the original β shape of the emission spectrum could not be circumvented for the following reasons. As the emitted electrons propagate through the LAr volume, the majority of the electrons either do not reach the HPGe detectors because of the surrounding nylon shrouds or do not penetrate the detectors with depths deeper than the dead layer. Instead, the dominant signal of the ^{39}Ar background originates from bremsstrahlung emissions of β particles in the LAr. Hence, the spectral distortions needed to be modelled with a modified distribution. We tested different modifier distributions such as β, γ, χ, χ^2, skewed Gaussian, and exponential distributions, and found another β distribution as the modifier leading to the most accurate fits. Overall, we fitted ten free parameters for this distribution, two β shape parameters plus location and scale, for both β components, as well as one modification parameter and one amplitude parameter for the global scaling. We found the optimum parameters for both the $2\nu\beta\beta$ function and the ^{39}Ar contribution via a combined binned fit of the merged M1 data set after the LAr veto as defined above. We separately fitted the model to the low-energy data set in the region from 53 to 207 keV, where the data were taken during Phase II+, and to the medium energy range between 184 to 1033 keV, with data taken during Phase II and Phase II+. The lower end of this range was motivated by high noise conditions below around 50 keV in the GERDA data, which does not justify a search for a new physics signal. To constrain the background properly around the probed bDM models, the signal models are restricted to the range $[65, 195]$ keV and $[196, 1022)$ keV for the low and medium energy range, respectively. Using the Python3 minimiser SciPy optimise [194], the background model fits yielded a goodness-of-fit estimate of $\chi^2/dof \sim 1.51$ and 1.09 for the medium energy range and for the low energy interval, respectively. Here, dof denotes the degrees of freedom, which correspond to the number of bins subtracted by the number of free fit parameters. The resulting empirical background fit modelling the combined GERDA data is shown in Fig. 4.10. There we also indicate the residuals, which we evaluate as the difference between expected and observed counts over the square root of the expected counts, for both energy windows. The fit results demonstrate a very good agreement between data and model over extended ranges, except deviations at certain individual energy values. These are dominantly caused by expected, or at least known, γ-transitions. In Sec. 4.4, we will explain the actual signal search, which led us to the identification of these lines. In Appendix B, we will further discuss the investigation of potentially visible γ-lines in the data that may mimic a signal. Including all identified lines as detailed in Sec. 4.4 and Appendix B.2, an *a posteriori* model improves the χ^2/dof goodness-of-fit estimate to 1.06, with a p-value of 0.09 (before 0.000). Below we will also see that for the low energy part no γ-line is identified, and the p-value is always 0.23. We also applied additional goodness-of-fit criteria to validate the accuracy of our empirical model. A non-parametric test statistic based on the maximum distance between the expected and the observed cumulative distribution function (CDF) is the Kolmogorov-Smirnov (KS) test statistic [195]. We estimated the corresponding p-value, p_{KS}, by approximating the empirical CDF of the model and the data via the histograms directly. As this approach slightly differs from the direct, sample-based estimation of the empirical CDF, we did not rely on the known asymptotic test statistic distribution but simulated 10^5 model histograms to evaluate the expected distribution of the KS test. We find the resulting p-value to be 0.99 for the low, and 0.16 for the medium-energy data set. Again,

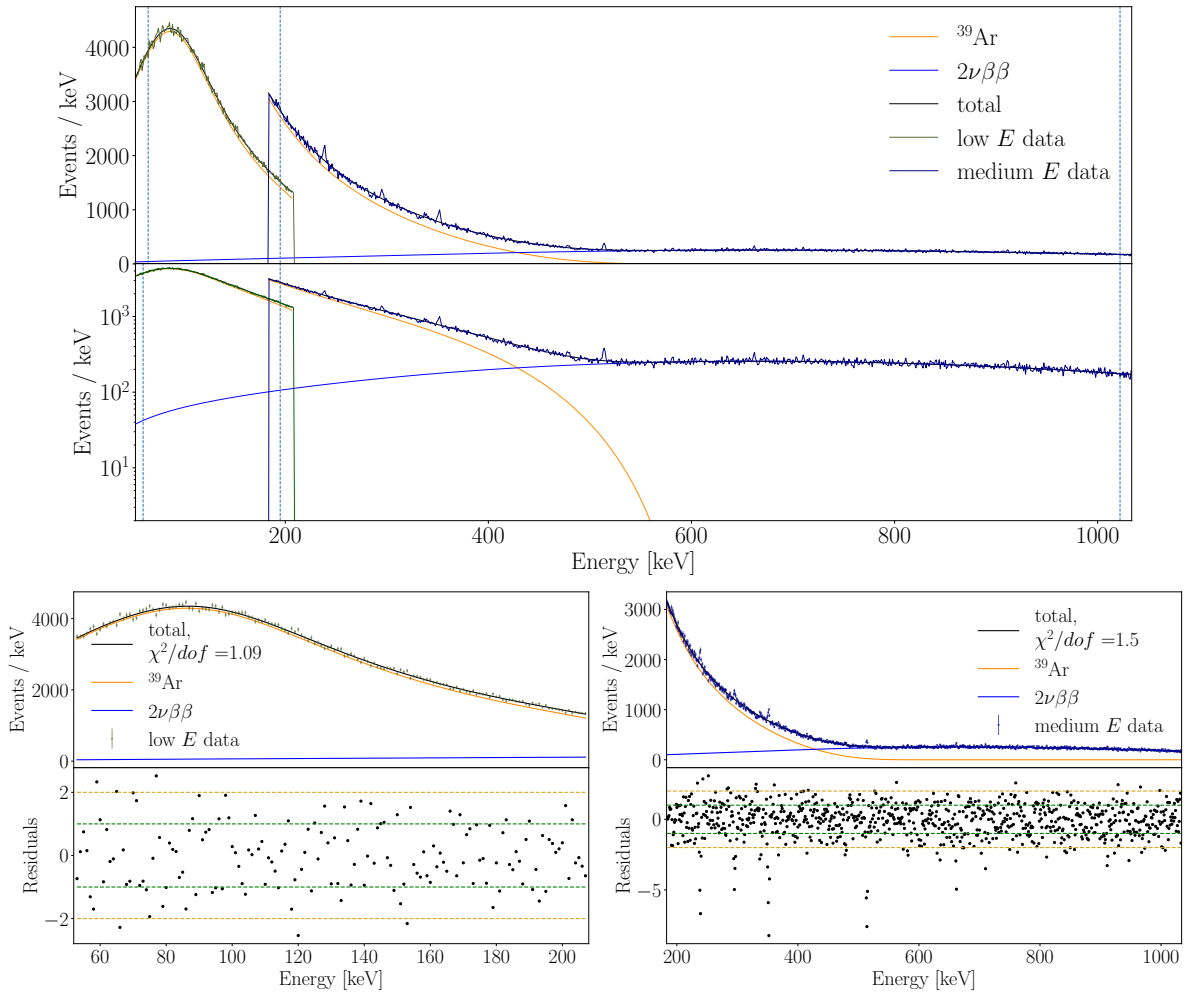

Figure 4.10: Top: plots of the empirical background fit model, in linear scale and logarithmic scale, compared to the low and the medium energy GERDA data after the M1 and LAr veto cut. The two fit components are also plotted individually. The vertical dashed, blue lines denote the lowest probed DM mass of 65 keV, the low-energy data set transition value of 195 keV, and 1021 keV as the highest potential integer DM mass below $2m_{e^-}$. Bottom left: low energy range of the fit model, including the residuals. Bottom right: medium energy range, including the residuals.

with the *a posteriori* model including the identified lines (Sec. 4.4) we obtained $p_{\mathrm{KS}} = 0.38$ for the medium energies. We also tested the normality of the distribution of the residuals with the help of the KS test. For the low energy spectrum fit residuals, we obtained $p_{\mathrm{KS}} = 0.70$, whereas for the residuals of the medium energy range fit, we got $p_{KS} = 0.005$, or 0.46 when excluding identified γ-lines. As the KS test is known to be not equally powerful to identify deviations from normality, we also apply a Shapiro-Wilk (SW) test [129]. Based on the ordering of a test sample and the deviations from the mean, this statistic had been found to be specifically suitable for normality testing [130]. Using the p-value estimate from SciPy [194], we obtain $p_{\mathrm{SW}} = 0.58$ for the low-energy data and $p_{\mathrm{SW}} = 0.000$ for the medium-energy data, rising to 0.12 after including the identified lines. Apart from the clear deviations at and around observed γ-line positions (Sec. 4.4), we conclude that the residuals largely fluctuate within the expected 1-2 σ ranges for

both energy ranges. To summarise, no broad deviations of the empirical model from the data have been identified. For convenience, we list the main properties of the empirical background model below.

Main properties of the empirical background model

- Components: ^{39}Ar distribution modelled as modified β distribution (ten free parameters), $2\nu\beta\beta$ spectrum described with a polynomial function of order ten (five free parameters)
- Energy ranges: 53 to 207 keV (for signal masses between 65 and 195 keV), 184 to 1033 keV (for signal masses between 196 and 1021 keV); fitted independently
- Goodness-of-fit investigations: χ^2 and Kolmogorov-Smirnov tests, plus Kolmogorov-Smirnov and Shapiro-Wilk tests for residuals; no significant deviation identified *a posteriori*

γ-line background

As mentioned above, several stronger peaks are present in the data, which were caused by γ radiation emitted within or near the detectors. It is crucial to understand that an induced γ-ray fully absorbed in a HPGe detector causes exactly the same event type as the signals we are searching for. Hence, these events cannot be distinguished from new physics signals, unless one has knowledge or an accurate estimate on the expected number of isotope transitions visible in the spectrum, i.e. on the signal strength. We do not have an accurate simulation available over the entire analysis range, and can thus not a priori estimate the γ-ray contamination accurately. In the former bDM study, γ-lines have been investigated and included in the combined signal plus continuous background fit if the branching ratio of a known γ-line is above 0.1% [190]. However, not all lines with this property are actually listed in Tab. II in Ref. [161]. Here, we changed this approach. We did not *a priori* include any potential γ-line in the background model but performed a generic search for any peak-like excess. If a line signal at a certain energy most likely caused by a known isotope transition appears to exceed an indication threshold, we explicitly stated the corresponding energy together with the potential origin, if known. We set this bound arbitrarily to the $3\,\sigma$ significance level, in correspondence with the typical threshold to claim an evidence in the field of high-energy physics. Subsequently, after obtaining the limits on potential new physics signals as described in Sec. 4.4, we deleted limits in the vicinity of identified γ-lines, where vicinity refers to the inner bin plus 3 bins on the right and 3 bins on the left of the known peak. This region corresponds to approximately a 2.5 FWHM resolution of the mixture distribution modelling the signal shape. Removing these results later on prevents us from interpreting excesses in all likelihood induced by known background contributions as an exotic physics signature. Let us remark that the official GERDA background model for the MeV range contains certain isotopes that undergo γ-rays with energy depositions falling into our analysis range [110]. Hence, our generic approach is expected to identify potential significant excesses because of the presence of such background contributions. We thus perform an independent, event-counting-based identification for the purpose of crosschecking the signals induced by γ-rays as well. We describe this method in Appendix B.2, yielding comparable results.

4.2.2 Multiplicity-two data set

To search for the single nucleon decay channel, limiting ourselves to the subsequent ^{75}Ge beta decay branch producing a coincident γ particle from the ^{75}As de-excitation, the data modelling becomes substantially easier. To select potential signal events, we require an M2 event with two different energy depositions. On the one hand, an event is interpreted as a β if the energy lies below $(Q + 2 \cdot \mathrm{FWHM}(Q))$, with $Q = 919\,\mathrm{keV}$ and a FWHM resolution of the mixture model of $\sim 3\,\mathrm{keV}$ to allow for some contingency beyond the spectral endpoint. On the other hand, the coincident γ particle is identified with an energy of $E_\gamma \pm 12.5\,\mathrm{keV}$, where $E_\gamma = 265\,\mathrm{keV}$. The window width of 25 keV is needed to appropriately constrain the background surrounding the potential signal at E_γ, which then provides an estimate of random coincidence events at the signal position. In addition, the value was chosen in correspondence with the bDM searches. If both energy values happen to lie within the γ particle energy region, the one closer to E_γ is interpreted as γ-signal, which causes a maximally conservative constraint on the number of γ-events. This happened in only four cases, without any relevant effect on our later results. To reduce the background rate, we applied the LAr veto, the muon veto, and the data quality cuts, but as before, no PSD cuts. With this selection, we modelled the continuum around the 265 keV line by a simple linear fit without the need for a dedicated background model. In Fig. 4.11, we show the events surviving the selection criteria for the nucleon decay search via ^{75}Ge isotope decays. Note that no relevant γ-ray transition is expected to lie in the selected window around E_γ and to survive the M2 cut at the same time.

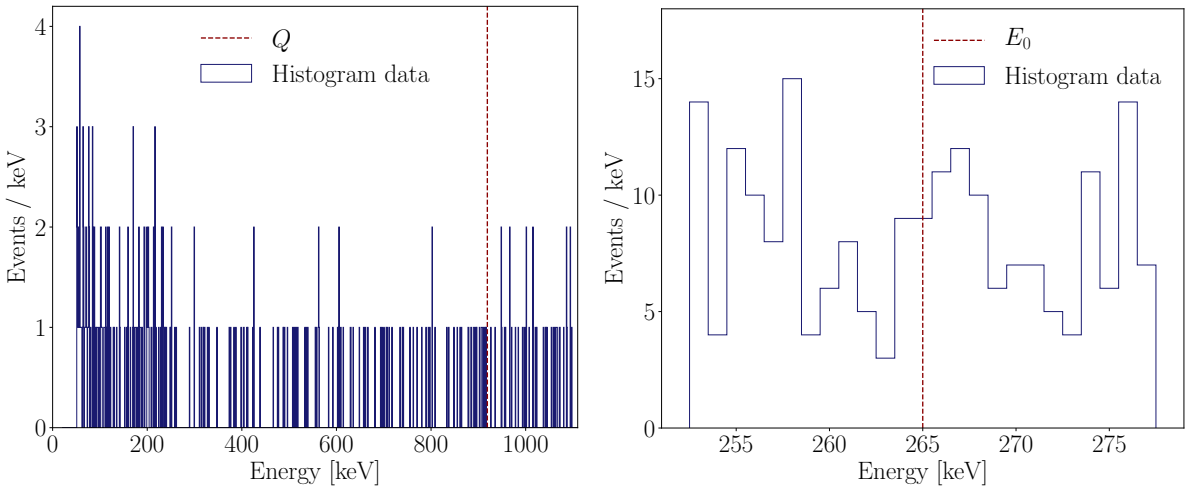

Figure 4.11: Left: β-component of candidate ^{75}Ge events in the M2 data set. Only energy depositions below the Q-value of ^{75}Ge pass the selection criterion. Right: potential coincident γ signal events surrounding the expected γ energy of $E_0 \sim 265\,\mathrm{keV}$.

4.3 Detection efficiencies

In order to understand how efficiently we can tag any of the signal events described above, our collaborators from UniPD simulated 10^7 signal events per analysis in MaGe [142], the code framework also used for the simulations discussed in Sec. 3.5. All experimental detector array configurations with which the real data were collected in Phase II and II+ were taken into account as separately modelled simulation geometries. The simulations were performed per

detector type and per phase, with all simulated events thrown as uniformly distributed events among all detectors, independently of their on or off status, and independently of dead or active material. After the simulation of the particle propagations, the downscaling to the active material, the detector-off phases, and non-physics data times was performed. Furthermore, the multiplicity cut and LAr veto cut were directly incorporated into MaGe. To determine the detection efficiency for the bDM absorption channel, 10^7 β particles were simulated each for integer values below 2.5 MeV in 10 keV steps. For the scattering channel, both 10^7 β and γ particles each were simulated parallelly from the same vertex, with the respective energies distributed according to Eqs. (4.3) and (4.4). For the electron decay channel, 10^7 uniformly distributed γ particles with the theoretical mean energy of $\sim$ 256 keV were simulated. For nucleon decay search, complete ^{75}Ge isotope decays including the subsequent ^{75}Ar de-excitation were run. Provided with the final detector-type specific distribution templates for all cases and energies, we fitted the efficiencies for each type and phase individually with a polynomial function. In the next step, we evaluated the final efficiency for the entire data set as an exposure weighted average at any integer energy. We show plots of the energy dependence for the different detector types and phases of the bDM absorption and scattering efficiencies in Fig. 4.12.

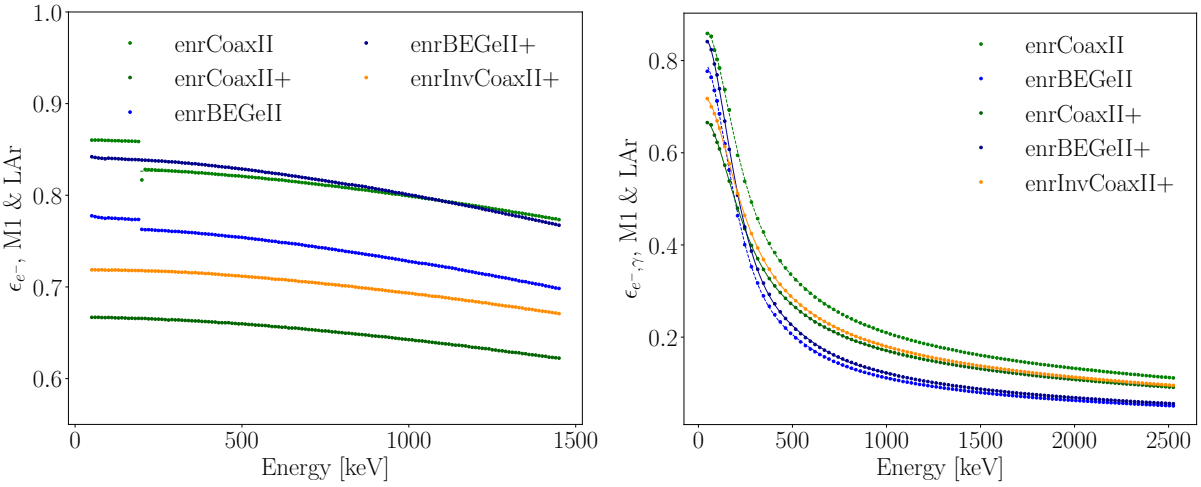

Figure 4.12: Simulated efficiencies of observing electrons (left), or β-γ-pairs (right), in the M1 data of GERDA after applying the LAr veto cut. Per data point, 10^7 events were simulated by our collaborators from UniPD. We evaluate the final efficiencies via weighing the results from a polynomial fit per detector type and phase with the corresponding exposures.

Next, we applied all cuts, assuming total signal survival for the data quality cuts and the muon veto, which were estimated in the past to be around 99.9% [47]. Taking into account the induced LAr veto dead time, this leads to reduction of, exposure weighted, approximately 2.1% [47]. The final exposure-weighted efficiencies after applying these cuts drop from 75.5% to 72.8% and from 74.4% to 14.9% for the absorption and the scattering channels, respectively, over the analysed bDM mass regime $[65, 1021]$ keV. Furthermore, we evaluate the detection efficiency for the semi-visible electron decay signature to be $\epsilon_e = 37.9\%$. The detection efficiency for the coincident M2 hits for the entire data set is $\epsilon_n = 0.185\%$, considering the energy ranges for the β-γ particle pair in correspondence with the M2 data selection (see Sec. 4.2). The overall observed expected energy spectrum of coincident M2 hits is pictured in Fig. 4.13. For both ϵ_e and ϵ_n, all relevant cuts and the LAr-induced dead-time were taken into account.
All the simulated efficiencies are prone to a certain systematic uncertainty. Potential marginal

inaccuracies of the simulated geometry of the experimental configuration are neglected, as no relevant deviations between data and simulations of GERDA had been found in the past [110]. We also ignore the uncertainty induced by the finite number of simulated primaries, given that $N_{\text{tot}} = 10^7$ events were simulated per individual model, implying an approximate contribution at a level of $\sqrt{N_{\text{tot}}}/N_{\text{tot}} = 0.03\%$. The detector-related parameters remain as relevant factors. For the active detector volume, calculating the exposure weighted average of the values stated in Ref. [47], we estimate a 4.0% effect. For the nucleon decay search, we additionally consider the uncertainty on the ^{76}Ge enrichment fraction also taken from there, inducing a 2.2% systematic uncertainty, evaluated as an exposure-weighted average. The enrichment level is quasi irrelevant for the other processes.
Let us emphasise that these efficiency simulations motivate us to restrict the analyses to the full energy peaks. Additional signal features dominantly cause Compton continua or other broad patterns induced by incomplete charge collections, and thus do not provide a distinct signal feature, which mitigates the chance for any convincing discovery. An exception is the peak induced by the dark Compton scattering when only the electron is fully absorbed, but the γ particle escapes its detector of the bDM interaction. This signature would offer a double peak feature, as motivated in Ref. [153]. However, this second peak would be located at the lower energy range even for the maximal bDM mass probed (Eq. (4.3)) and would thus be hidden under the strong ^{39}Ar continuum. For these reasons we do not include any other signal pattern in the main analyses. We remark though that an alternative method for the nucleon decay, suffering from the lowest efficiency, is motivated in Appendix B.1. Based on a continuous β signal in the M1 data with an increased efficiency but also suffering from a higher background rate, there we obtain a very similar result as here in the main search.

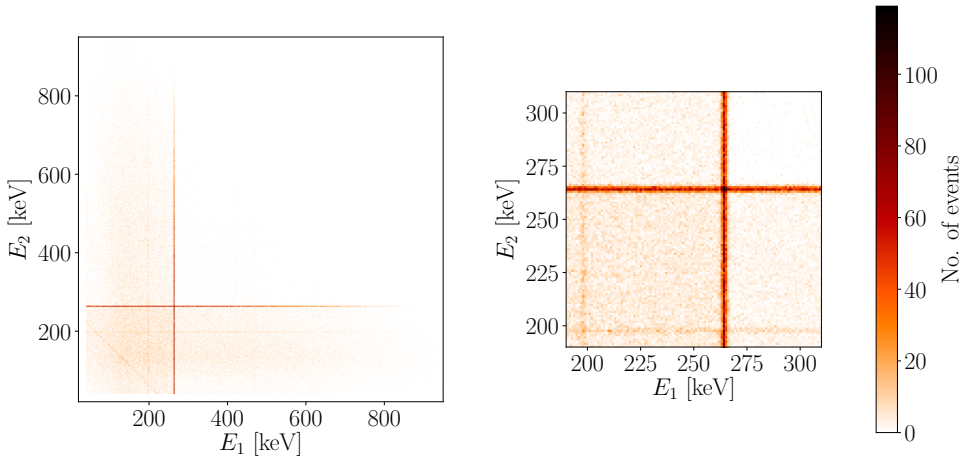

Figure 4.13: Simulated energy hits in the coincident M2 data in GERDA when throwing 10^7 ^{75}Ge isotope decays and zooming into the region around the γ-line at 265 keV caused by the de-excitation of the ^{75}As daughter isotopes produced from β decay. The simulation was run by our collaborators at UniPD.

We close this section with an overview tabulating the simulated efficiencies for the different analyses in Tab. 4.1.

Search	Channel	Energy [keV]	Efficiency [%]
bDM, absorption	β, M1	65-1021	75.5-72.8
bDM, scattering	$\beta + \gamma$, M1	65-1021	74.4-14.9
Electron decay	γ, M1	256	37.9
Nucleon decay	$\beta + \gamma$, M2	continuous (β)+ 265 (γ)	0.185

Table 4.1: List of simulated signal detection efficiencies for the different new physics analyses. The numerical values are obtained as exposure-weighted averages of the results of an interpolation between integer values in 10 keV steps, as simulated by our colleagues from UniPD. For each detector type, phase, and simulated energy value, 10^7 primaries were sampled. For the nucleon decay study, additional simulations were run. Quality and multiplicity cuts (M1/M2), as well as the dead-times induced by the muon and LAr veto cuts, are considered.

4.4 Statistical method

In both the former bDM study and the former neutron decay study, no signals had been found; thus, Bayesian credible intervals were determined [161, 176]. In this work, instead, we describe the extraction of a confidence interval on the signal strength in the Frequentist statistical interpretation based on the data from phases II and II+. In addition to these analyses, we also searched for the proton and the electron decay. We applied the likelihood-ratio test statistic q_0, following the nomenclature from Ref. [196], with the index indicating the background-only hypothesis H_0, i.e. a vanishing signal amplitude. Note that according to the Neyman-Pearson lemma [197], the likelihood-ratio test poses the most powerful test statistic, and is thus highly suitable to identify an incompatibility of the data with H_0, or in other words, highly sensitive to detect a new physics signature. We rely on the evaluated significance of a hypothesised signal peak at each potential bDM mass or at the γ-line energies induced by the probed single particle decay channels to statistically interpret how improbable deviations from the background are. Given the number of events in the data, we assumed the asymptotic $\chi^2_{1/2}(1)$ distribution to hold [196] when performing the hypothesis testing. Here the 1 in brackets denotes the number of degrees of freedom, and the index 1/2 indicates a δ-distribution with a weight 1/2 located at zero, under the assumption of no signal presence, i.e. noise under-fluctuations in half of the cases. Note that under the asymptotic distribution, assuming the background-only hypothesis H_0 to hold, the p-value, i.e. the difference between unity and the CDF of the test statistic indicating agreement between assumption and hypothesis test result, is simply given as [196]

$$p_0 = 1 - \Phi(\sqrt{q_0}) \ . \tag{4.9}$$

Here Φ denotes the CDF of the standard normal distribution with zero mean and unit variance. This result follows from a direct relation between the CDFs of χ^2 and normally distributed random variables. A full derivation based on the theorems of Wilks and Wald [198, 199] is outlined in e.g. Ref. [196]. As generally the p-value can be translated in a significance S via the well-known relation $S = \Phi^{-1}(1-p_0)$, we can simply estimate the significance via the square root of the test statistic result as measured for the best-fit value of the peak amplitude. To define our relevant evidence and discovery threshold *a priori*, we need to take into account the severe constraints from theoretical considerations, other direct detection experiments, and/or indirect constraints as discussed above and in Sec. 4.1. Furthermore, the large number of probed signal models in the bDM search of almost order 1000 let us expect a strong look-elsewhere effect, i.e. expected strong noise fluctuations resulting from the large number of trials performed but not necessarily caused by any physical effect. With these considerations in mind, we require a

local $5\,\sigma$ significance threshold for an evidence of a potential observation of an excess over the background. We show below that this evidence threshold corresponds to a global significance of approximately $4\,\sigma$ for the bDM analysis, prone to the strong look-elsewhere effect. For the particle decay channels, with known expected peak energies, local and global significance coincide, implying that we can immediately define a $4\,\sigma$ excess as signal evidence. In addition, we define a less stringent indication threshold of $3\,\sigma$ and provide a list of all peaks exceeding this threshold below. Furthermore, we do not define a discovery threshold, given the lack of a precise background γ-line strength prediction from a dedicated and well established GERDA background simulation model after the LAr veto cut in our entire analysis range.
If evidence for a signal is found, we will derive a central interval at 68% CL, corresponding to an approximate $1\,\sigma$ interval, on the signal strength parameter. In case no signal is observed, we will instead determine an upper limit, where we define the term upper limit as the confidence interval bound at 90% CL. For this purpose, we apply the test statistic $\tilde{t}_\mu$ from Ref. [196], where the index denotes the tested signal strength. As motivated therein, we simplified the interval evaluation by approximating the distribution of the test statistic by its known asymptotic form, a $\chi^2(1)$ distribution, in the case of a positive best-fit signal amplitude. In this case, the upper limit is the interval boundary of a central (two-sided) confidence interval. Fixing the CL to 90%, i.e. the significance level α to 0.1, this implies that we need to set the threshold of the likelihood profile to 2.71. This value corresponds to the solution of $\tilde{t}_\mu$ for the p-value under the signal hypothesis fixed to α via [196]

$$0.1 = \alpha \stackrel{!}{=} p_\mu = 2\left(1 - \Phi\left(\sqrt{\tilde{t}_\mu}\right)\right) \ . \tag{4.10}$$

We performed the profiling numerically. As we show later, the asymptotic approximation yields better than ~ 3 and $\sim 11\%$ level accuracy for the bDM search and for the decay searches, respectively, as determined via Monte Carlo simulations at selected energies. In the case of background under-fluctuations, we need to prevent obtaining non-physical negative limits. In these cases, we thus assume a non-central $\chi^2(1, \Lambda)$ distribution as outlined in Ref. [196]. The non-centrality parameter Λ was evaluated via the artificial Asimov data set, a data set with suppressed fluctuations [196]. This provides us with an estimation of the standard deviation of μ around its best-fit value via $\sigma_A^2 = \mu^2/\Lambda$. Here, we generated the Asimov data with signal strength 0, posing the physically relevant signal strength parameter closest to a negative best fit value. Additionally, we also used these data to evaluate the sensitivity for each model, independently of the best-fit value. We provide more details on the Asimov data set further below. The resulting upper limit for negative best-fit amplitudes naturally transits to a one-sided interval, as the test statistic $\tilde{t}_\mu$ is asymptotically equivalent to the procedure of unified intervals suggested by Feldman and Cousins [200] in the absence of nuisance parameters [196]. Moreover, relying on this test statistic prevents us from performing any result-dependent flip-flopping between central and one-sided intervals, which would violate the correct coverage, i.e. the stated confidence level, as pointed out in Ref. [200]. Let us remark that having fixed the parameters of the signal mixture model as deduced from calibration data, we do not need to further constrain any nuisance parameters. Following the prescription from Ref. [196], we determine the desired upper limit by numerically solving

$$0.1 = \alpha \stackrel{!}{=} p_\mu = 2 - \Phi\left(\sqrt{\tilde{t}_\mu}\right) - \Phi\left(\frac{\tilde{t}_\mu + \mu^2/\sigma^2}{2\mu/\sigma}\right) \ . \tag{4.11}$$

For the numerical minimisation, we made use of the asymptotic Wald approximation for $\tilde{t}_\mu$ and approximated the standard deviation σ obtained from the Asimov data as σ_A [196]. We obtained

the solution by iterating over μ until the difference between left and right hand side of Eq. (4.11) vanished. We now turn to the description of, firstly, the fitting procedure, secondly, the accuracy check of the asymptotic distribution of the test statistic, and thirdly, the investigation of the significance for the different physics models.

4.4.1 Fitting procedure

As outlined in Sec. 4.2, we applied a binned fit to determine the likelihood-ratio test statistic. At each signal position, we fitted the background-only model yielding the likelihood of the H_0 hypothesis, and a combined binned fit of a probed signal at fixed energies plus the surrounding background to obtain the likelihood of the alternative model H_1. The signal peak positions probed are in the interval $[65, 1022)$ keV, in 1 keV steps, for the bDM search, and at 256 keV for the electron decay channel. We use a bin width of 1 keV and centre each bin in our fit windows at the integer value of unit keV. As mentioned above, for the electron decay, we need to widen the window to 120 keV because of the Doppler broadening. The fit was run by minimising the negative logarithmic likelihood function for the bin count data as expected from the probed fit model with respect to the observed bin count data, which we implemented again with SciPy optimise [194]. When relying on the empirical fixed background model as introduced in Sec. 4.2, the only free parameter is the amplitude of the hypothesised signal peak. In comparison, in the sliding polynomial background, applied to evaluate the systematic uncertainty induced by the background modelling procedure, we additionally fit two parameters for the linear model above probed masses of 195 keV and three parameters for the quadratic model for probed masses below or equal to this value. As outlined previously, the signal is modelled as a Gaussian mixture distribution with weights being equal to the exposure of the data sets for each detector type and data-taking period. This means that we have five individual Gaussian terms summed together, three detector types and two detector phases, where no IC detectors were operated during Phase II. Between low and medium energy data set we adjust the corresponding Phase II exposures. Fig. 4.14 shows example fits for the two hypothesised bDM masses 65 keV, left, and 905 keV, right. In both fits, the empirical background model was used. At the bottom, we also indicate the likelihood profiling used to evaluate the upper limits.

In Fig. 4.11, right, we show the identified candidate ^{75}Ge events in the M2 data. On the left of Fig. 4.15, we illustrate the corresponding fitting procedure, with the best-fit hypothesised Gaussian mixture model indicated in dark red, on top of the linear fit drawn in dashed blue, and the 90% CL upper limit in dark green. Note that for this particular likelihood minimisation, we used the Python3 front-ends iminuit and probfit [201] to call the FORTRAN minimiser algorithm MINUIT2 [202]. The evaluation of the limit is described in the following subsection.

For the electron decay analysis, we fitted the introduced Doppler-broadened analytic line shape on top of the background fit model. In the widened search window, even the potential K-shell electron decays with the strongest broadening would be contained (Sec. 4.1.3). The fit of the total line shape, including the mixture model taking into account the different energy resolutions for the different detector types, on top of the empirical background model, is plotted in Fig. 4.15, right.

4.4.2 Test statistic distribution

With the help of the obtained likelihood-ratio test statistic, we estimate the significance of the probed signal models as the square root of the likelihood-ratio, if the best-fit amplitude is positive. Otherwise, when observing an under-fluctuation, the significance of an excess is set to zero, as we only focus on new physics channels inducing energy depositions. If we do not find

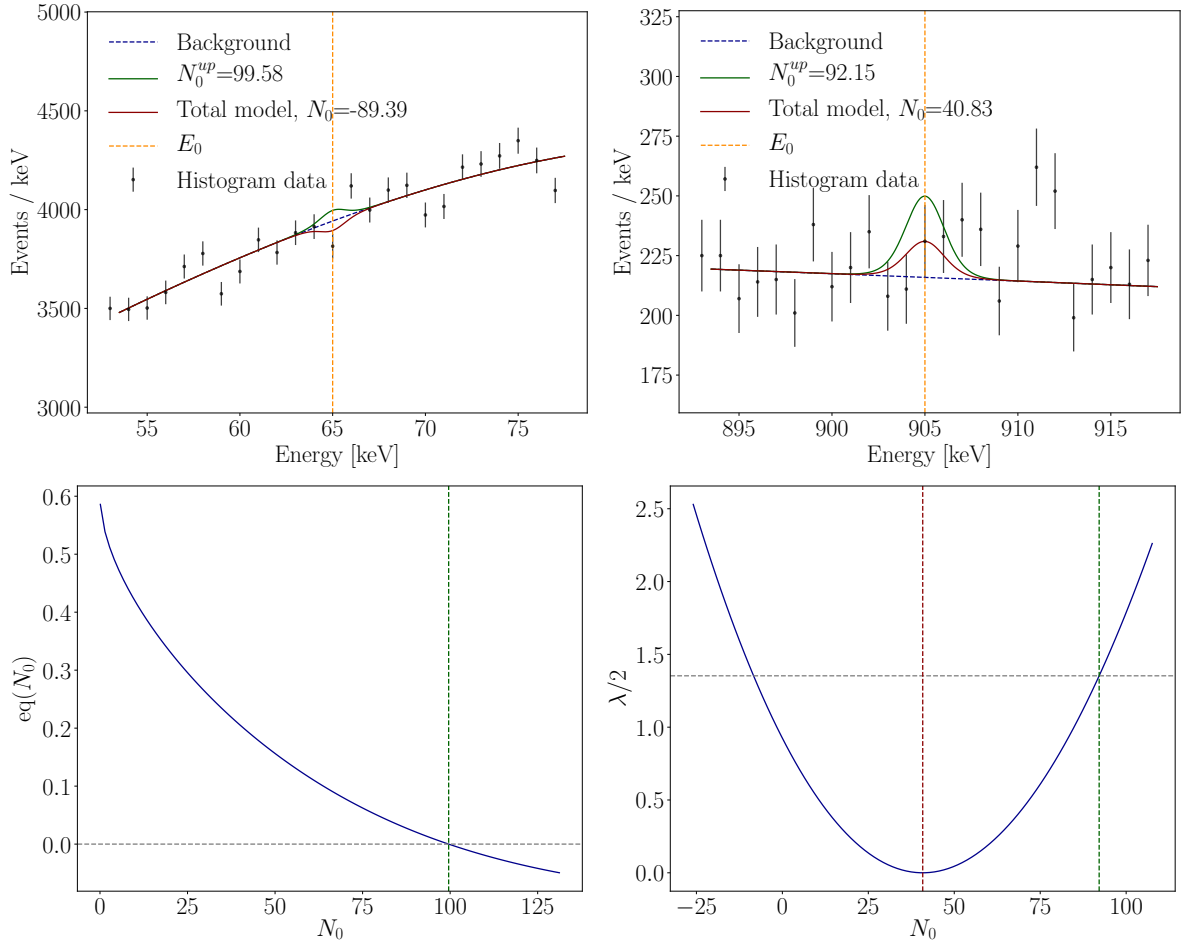

Figure 4.14: Top: examples of bDM model fits to determine the best-fit signal count strength N_0 and the corresponding upper limit. Left: for $E_0 = 65\,\text{keV}$. Right: for $E_0 = 905\,\text{keV}$. A maximum likelihood based combined fitting approach with a Gaussian mixture model signal shape on top of the background model has been applied. In the right plot a peak in the sidebands appears, most likely caused by the known ^{228}Ac transition at 911 keV. Bottom: profiles of the likelihood function over different signal strengths N_0, for the same energies as above. Left: numerically solving the difference between the two sides of Eq. (4.11), abbreviated as $eq(N_0)$, which is required because of a negative best-fit value. Right: numerically solving Eq. (4.10), where λ denotes the likelihood ratio.

an evident excess, we consequently extract upper limits for all models, using the asymptotic distribution of the test statistics for both significance test and confidence interval evaluation. In the second part of this subsection, we justify the asymptotic behaviour of these test statistics.

Firstly, we estimated the expected counts in a signal peak as the integrated count number of the empirical background model from Sec. 4.2 in a 1 FWHM wide window centred at the peak position. Fig. 4.16 shows that the expected signal counts range from several hundred to several thousand. This indicates that the expected count values can be well-approximated by a Gaussian distribution, as for sufficiently high statistics the underlying Poisson distribution converges to the normal distribution. These count numbers motivate us to run a binned analysis as mentioned in Sec. 4.2 and to approximate the test statistic distributions via their asymptotic

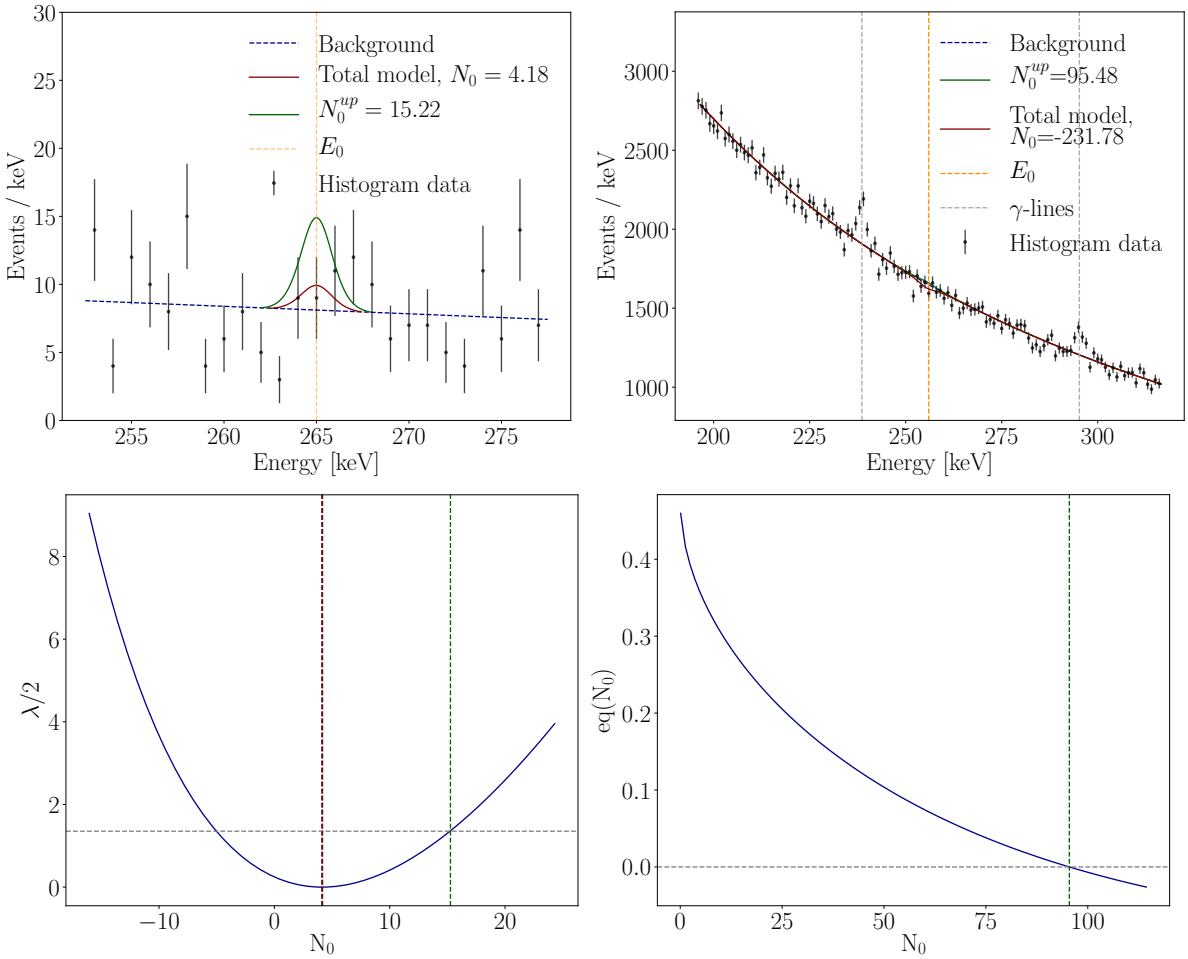

Figure 4.15: Top, left: combined likelihood fit of the probed, coincident signal γ-line at 265 keV. The best-fit amplitude, the 90% CL upper limit, and the linear background fit, are shown in dark red, dark green, and dark blue, respectively. Top, right: fit of the broadened line model for the electron semi-visible decay mode, in combination with the empirical background model, as well as the upper limit as deduced from the asymptotic behaviour. Bottom: corresponding likelihood profiles for the fits above to evaluate the upper limits via solving Eqs. (4.10) (left) and (4.11) (right), using the same notation as used in Fig. 4.14.

form. Secondly, we additionally crosschecked the accuracy of the asymptotic relations with the help of individual Monte Carlo simulations at energies $E \in \{100, 150, ..., 950, 1000\}$. For the electron and nucleon decay channels, we ran an additional Monte Carlo simulation each. Especially the nucleon decay signal window contains only 207 events, in contrast to the event counts between around 10^3 and 10^5 for the bDM search and the electron decay search. We show below that assuming the asymptotic form to hold, the test statistics lead to a conservative limit in all cases, with the most conservative estimate obtained for the M2 search. In order to test the distribution of the likelihood ratio test statistic under the background-only hypothesis (H_0), we ran 10^6 Monte Carlo iterations for each energy listed above. For each iteration, we first created event counts drawn from a Poisson distribution in all individual bins. Next, we fitted the signal amplitude by minimising the negative logarithmic likelihood in each search window, relying on the minimisation algorithm from SciPy optimise. Afterwards, we plotted the histogram of the

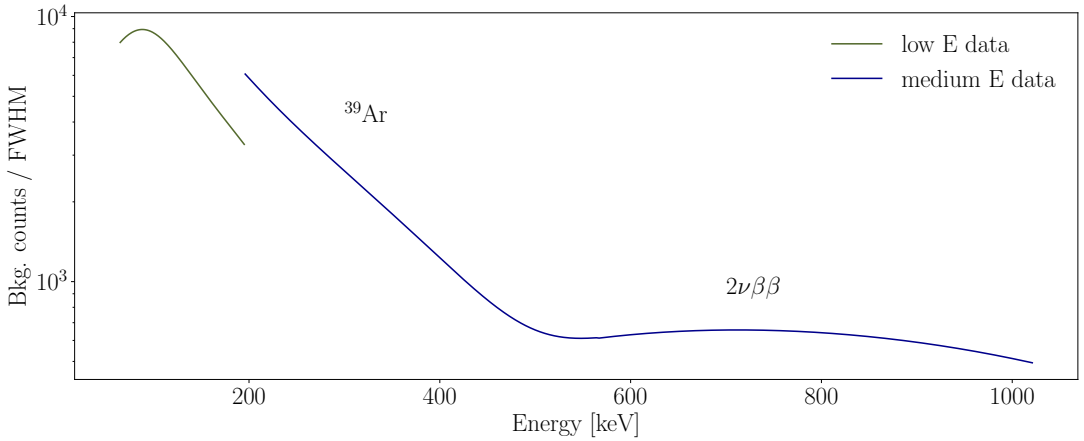

Figure 4.16: Plot of the expected counts in the signal window as a function of energy, for the low- and the medium-energy data set. We calculated the expectation value as the integral of the empirical background model fit (Fig. 4.10) in a 1 FWHM-wide window centred at the probed signal peak position. The obtained counts suggest that the underlying Poisson distribution for the bin counts can be well-approximated with a Gaussian shape, which implies that the asymptotic approximations for the distributions of the test statistics are accurate [196].

best-fit signal amplitudes and the test statistic, i.e. the square root of the likelihood ratio for these amplitudes.

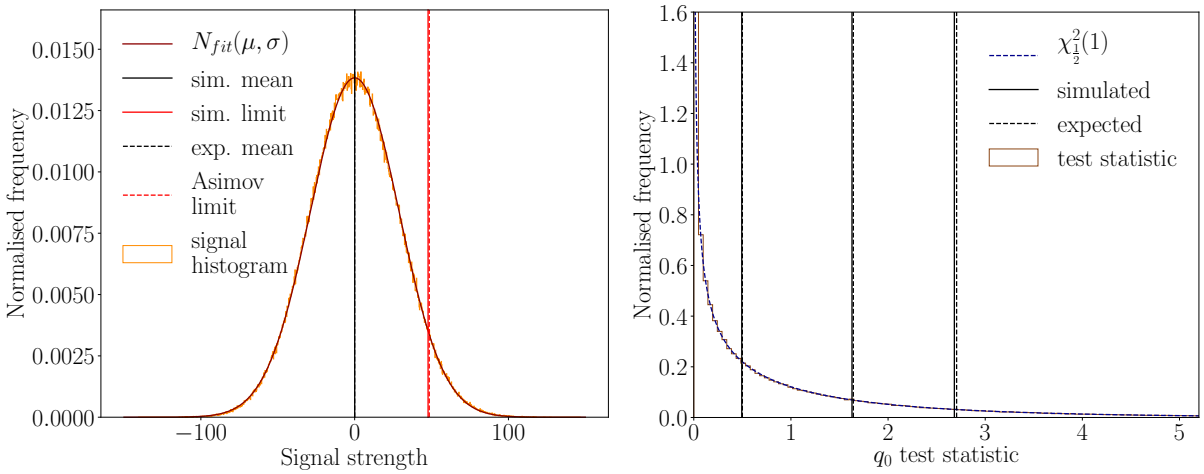

Figure 4.17: Left: histogram and best-fit Normal distribution of the signal strengths as fit on 10^6 Monte Carlo simulations of the fit window for $E_0 = 1000\,\text{keV}$. Right: histogram of the likelihood-ratio test statistic for each simulated data set for $E_0 = 500\,\text{keV}$ and plot of the $\chi^2_{1/2}(1)$ distribution. In both figures, the means and the two-sided 90% quantiles (in the right plot also one-sided) as deduced from the simulation, and as expected from Asimov data or the theoretical asymptotic distribution, are shown in solid and dashed lines, respectively, indicating perfect agreement. Note that the event count numbers in the region above approximately 500 keV are up to more than an order of magnitude lower than the observed event numbers around the maximum of the observed ^{39}Ar spectrum (Fig. 4.10). This implies that at lower energies, the asymptotic form will lead to even better accuracy than what is revealed here.

Example histograms are shown in Fig. 4.17. There, on the left hand side, we can see that even at a high energy such as 1000 keV, the simulated best fit signal amplitude histogram is perfectly well described by a Gaussian. Note that at these energies, we measure the lowest relevant count numbers for the M1 data set. The event numbers in the bins around this energy value are sufficiently high to accurately approximate the Poisson-distributed count numbers with a Gaussian-distributed variable. At lower energies, the approximation is even better given the higher event numbers. The accurate asymptotic form is also confirmed by the 90% quantile as expected from the Asimov data, which we detail below. It almost perfectly matches the corresponding quantile of the simulated histogram, being only slightly shifted to higher values (Fig. 4.17). The right hand side, showing both the histogram of the simulated test-statistic values for an intermediate energy of 500 keV and the probability density function of the $\chi^2_{1/2}(1)$ distribution, indicates that approximating the true critical threshold by the quantile of the asymptotic distribution yields accurate results on the limits on a level of a few percent, more precisely, on $< 3\%$ maximum deviation for the tested energies.

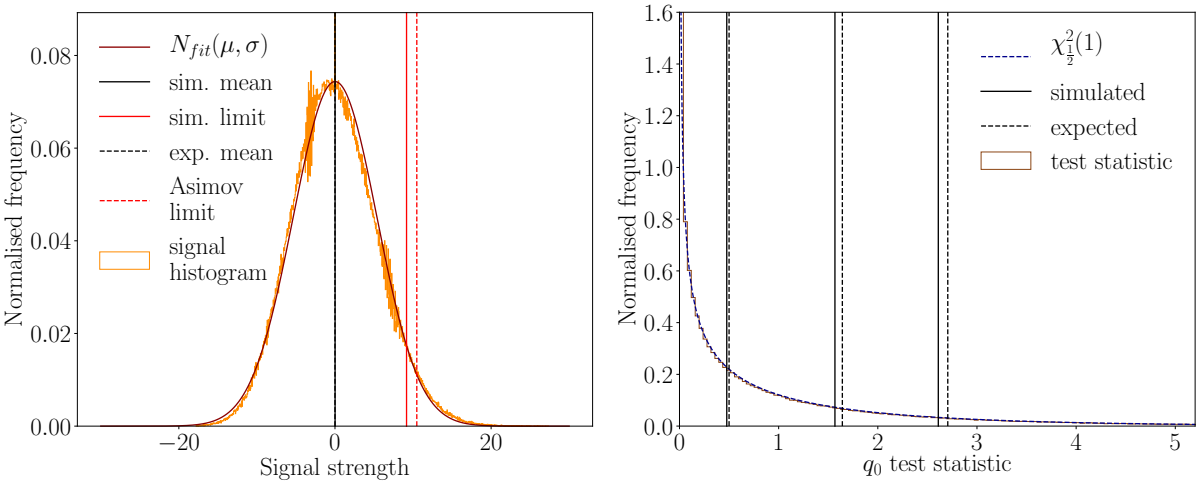

Figure 4.18: The same histograms and plots as those shown in Fig. 4.17 for the 265 keV line searched for in the M2 data set, again based on 10^6 Monte Carlo iterations. The small offset of the expected results to the right of the quantiles as deduced from the simulations implies a small overcoverage if applying the asymptotic properties of the test statistic, see text for details.

The features can be more easily seen in Fig. 4.18, which shows the same type of histograms as before, but for the 265 keV γ-ray energy in the M2 data set caused by a potential nucleon decay. Again, the expected quantiles are clearly visible in the vicinity of the result as deduced from the simulations. In this search, the sensitivity as evaluated from the Asimov data is approximately 11% more conservative than the simulation. Given the intrinsic uncertainty of any numerical algorithm, we judge the deviations sufficiently small for our needs. We hence always follow the mathematically justified asymptotic behaviour. We underline again that the approximations are not only rather accurate but also always conservative, as indicated by the asymptotic threshold being even further right in the plot than the one obtained from the simulation. This implies that even in the case of non-perfect accuracy of the asymptotic approximation, it may lead to overcoverage but never to undercoverage, meaning never to over-aggressive results.
A point which has been mentioned several times by now, but not properly introduced and discussed, is the sensitivity of the experiment. We will now see that it is asymptotically related to the Asimov data set briefly mentioned before, cf. [196]. What we refer to as the sensitivity

of the experiment for a given signal model is more precisely the median expected sensitivity for setting a two-sided limit at 90% CL in the case of the absence of the probed signal. From the simulated best-fit amplitudes, we readily get this sensitivity simply as the upper *95%* quantile (*keep in mind that we desire a two-sided interval at 90% CL*), as the simulations were run under the background-only assumption. A faster and easier approach omitting the need for simulations is the Asimov procedure explained in Ref. [196]. The basic idea is to interpret an artificial background-only data set with fully suppressed fluctuations as representative for the expected data to be obtained over an ensemble of many experiments. If sufficient events are available, the Poisson distribution can be accurately approximated as a Gaussian, and the average bin count over many iterations would simply be the expected value. Consequently, this artificial Asimov data set can be used to determine the expected upper limit directly, without any random-number generation. Furthermore, even the expected fluctuation of the upper limit around its median can be estimated, which is also explained in Ref. [196]. We followed their prescription to estimate the median sensitivity and its expected $1\,\sigma$ fluctuation band, respectively. In Fig. 4.19, we compare the sensitivities as extracted from the Asimov data to the interpolation between the sensitivity values as extracted from the simulated data sets and the fits of the simulated signal amplitude histograms.

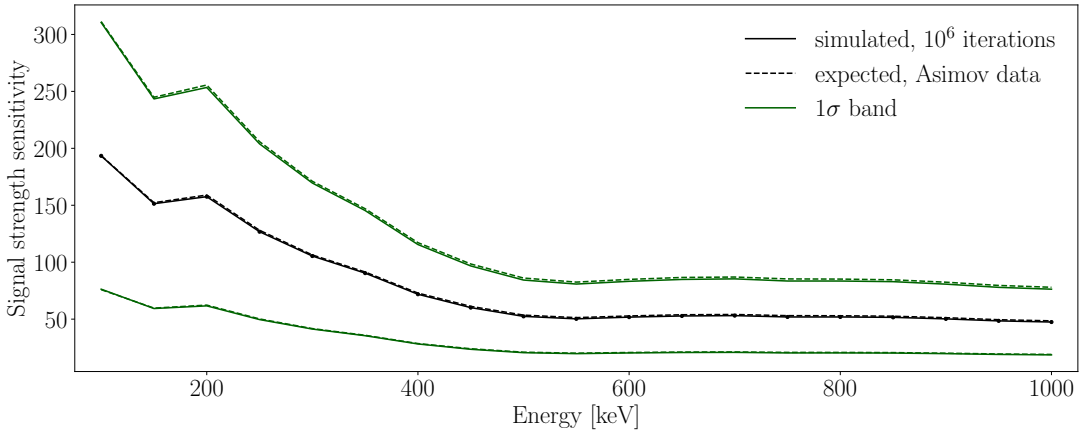

Figure 4.19: Comparison of the median sensitivity, including its $1\,\sigma$ fluctuations, as determined as an interpolation of the sensitivities obtained from 10^6 Monte Carlo simulations on a grid spanned by 19 uniformly spaced energies (solid lines), and as determined from background-hypothesis-only Asimov data sets with fully suppressed statistical noise at each energy in the range from 100 to 1000 keV (dashed). The lines almost perfectly overlap, verifying the accuracy of the Asimov sensitivity estimates used at each single energy value.

The lines overlap almost perfectly, so we justify that the Asimov approximation is accurate enough to both provide an estimate of the experimental sensitivity and to approximate the expected standard deviation of μ via σ_A for establishing an upper limit in the case of an under-fluctuation.
To investigate the systematic uncertainty induced by the background model, we also estimated the sensitivity via the Asimov approach for a sliding polynomial background fitting procedure within each 25 keV signal window instead of the empirical fixed background model. We used a second (first) order polynomial to constrain the background continuum for the low- (medium-) energy data set, as done in the former study [161] and motivated in Sec. 4.2. The sensitivities for all bDM candidates masses as obtained with the two different background modelling methods have a mean deviation of 0.2%, with a standard deviation of 0.7%, and the maximum deviation

at an individual energy value is 2.1%. These results indicate an excellent accuracy of the sensitivities independent of the background modelling approach. As for other sources of systematic uncertainty, an additional contribution comes from the chosen bin width. When using a 2 keV window, the sensitivities estimated with the sliding polynomial window approach worsened by $\sim 44\%$ compared to the 1 keV version. For a bin width of 0.5 keV, we got an improvement of 30% on average. Overall, this implies a mean uncertainty of around 7%, as our standard bin width is the closest to the energy resolution, which lies in between the two probed extreme values of 0.5 keV and 2 keV.

4.4.3 Significance and confidence intervals

After having discussed the fitting procedures as well as the experimental sensitivity, and convinced ourselves of the reliability of the asymptotic approach, we are now set to interpret the results of the performed fits of the different physics models. For the electron decay channel, we obtain a best-fit value for the amplitude of a Doppler-broadened line of -231.8 counts. The background fluctuates downwards strongly, meaning that the best-fit value lies outside the physically allowed range. Hence, the observed significance of the observed signal against the H_0 hypothesis is exactly 0. As mentioned before, for a known signal position, the global significance equals the local significance, which is obviously zero as well. The obtained confidence interval at 90% CL of 95.5 counts reaches below the sensitivity of the experiment of 268.5. Because of the underfluctuation, we here apply the non-central χ^2-distribution-based method assuming a physically allowed signal strength of zero. For the nucleon disappearance channel, we obtain a best-fit result of 4.2 events, with a, both locally and globally equally, significance of $0.7\,\sigma$. The upper limit determination yields 15.2 counts, assuming the conservative asymptotic distribution. Following the Monte Carlo estimated distribution, we obtain 15.1 events, which is in good agreement with the limit obtained from the asymptotic distribution. The corresponding sensitivities are 9.9 and 8.8 events for the asymptotic and the Monte Carlo procedure, respectively, meaning a maximum difference of $\sim 11\%$, which is the maximum discrepancy we see for any of the tested sensitivities. For the bDM study, we observe multiple $\geq 3\sigma$ excesses which are, in all likelihood, caused by known isotope transitions. We list all visible $3\,\sigma$ excesses together with their likely origin in Tab. 4.2.

Energy [keV]	Significance [σ]	Origin
237, 238, 239, 240	4.7, 7.9, 8.5, 5.3	^{212}Pb
294, 295, 296, 297	4.7, 6.7, 6.0, 3.5	^{214}Pb
338	3.0	^{228}Ac
350, 351, 352, 353	6.6, 9.9, 10.7, 7.1	^{214}Pb
478, 479	3.6, 3.5	^{228}Ac
512, 513, 514, 515, 516	4.8, 8.5, 10.2, 7.9, 3.6	^{85}Kr
581	3.1	^{208}Tl
661, 662, 663	4.7, 5.4, 3.4	^{137}Cs
710	3.3	unknown
911, 912	3.8, 3.1	^{228}Ac

Table 4.2: List of probed energy values where $\geq 3\,\sigma$ excesses were observed in the data, together with the corresponding local significance. The most likely origin from known γ-transitions expected to appear in the GERDA data are indicated as well.

We plot all local significance estimates versus the probed mass in Fig. 4.20, left, also indicating the identified γ-lines. There we also illustrate the distribution of the local significance, which is shown in the histogram on the right.

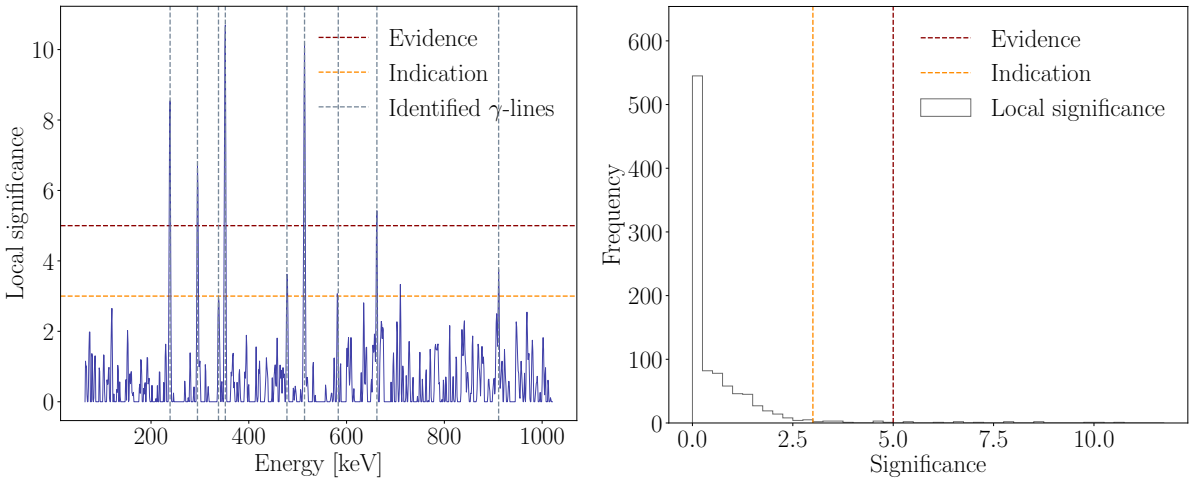

Figure 4.20: Left: plot of the local significance of all count strengths versus the tested energy value for the bDM analysis. The identified γ-lines have a local significance exceeding a $3\,\sigma$ indication threshold. At 710 keV, one excess is found that we cannot directly address to an expected γ-ray transition. Right: histogram of the corresponding local significance.

The only unknown line signal exceeding the indication threshold of $3\,\sigma$ appears at 710 keV, with a local significance of $3.3\,\sigma$. To our knowledge, this line cannot be reasonably attributed to an expected isotope transition. No excess beyond $3\,\sigma$ was found at this value in the former study [190]. A physical explanation could thus be related to a background source introduced upon the hardware upgrades performed before data taking phase II+, i.e. for data not taken into account previously. We checked all individual data sets around this energy value, and identified a small peak in any of the spectra independently of the operation phase, making a hardware work-related background source unlikely. Alternatively, the viability of an interpretation in terms of a statistical fluctuation instead is discussed below. As also this significance remains below our evidence threshold, we conclude that for all probed signal models, i.e. for both all bDM candidate masses and the two decay channels, the measured significance remains below $4\,\sigma$ also globally, apart from energies where γ-transitions are strongly expected.
As motivated before, the interpretation of the significance for the bDM models is different from the case of the particle disappearance searches. After removing the windows around known, identified lines, we have 894 trial experiments left. This large number of probed bDM models lets us expect a strong influence of the look-elsewhere effect, i.e. to see strong excesses from pure background fluctuation alone. To conservatively estimate how many upcrossings beyond a 3σ threshold we may approximately expect, we can, of course non-realistically, assume negligible correlations in the data. In such a simplified scenario, we would expect to see around 894 (number of experiments) $\times$ 0.3% (probability for $3\,\sigma$ occurrences per experiment) ~ 2.7 upcrossings, from pure noise fluctuations only. This simple estimation indicates that a few $3\,\sigma$ excesses are within our expectation. A conservative rescaling of the p-value by the number of trials neglecting correlations, known as Bonferroni correction [203], then clearly suggests a non-significant excess. The approximate global, Bonferroni-corrected significance is $S_{\mathrm{max,B}} = 0.17$, far from the indication or even the evidence threshold. To put this estimate on a more solid ground, we also

applied a data-driven estimate of the number of upcrossings to be expected from the likelihood ratio test statistic results directly, as described in Ref. [204]. Following their prescription, referred to as self-calibration, we artificially induced peaks of given local significance from 3 to $8\,\sigma$. We then compared these to the significance of the observed peaks beyond a $1\,\sigma$ threshold except for the strongest observed peak, in order to estimate the resulting global significance of a *salted* (meaning artificially induced) peak. Note that here we do not consider results located near the known γ-lines identified in the data.

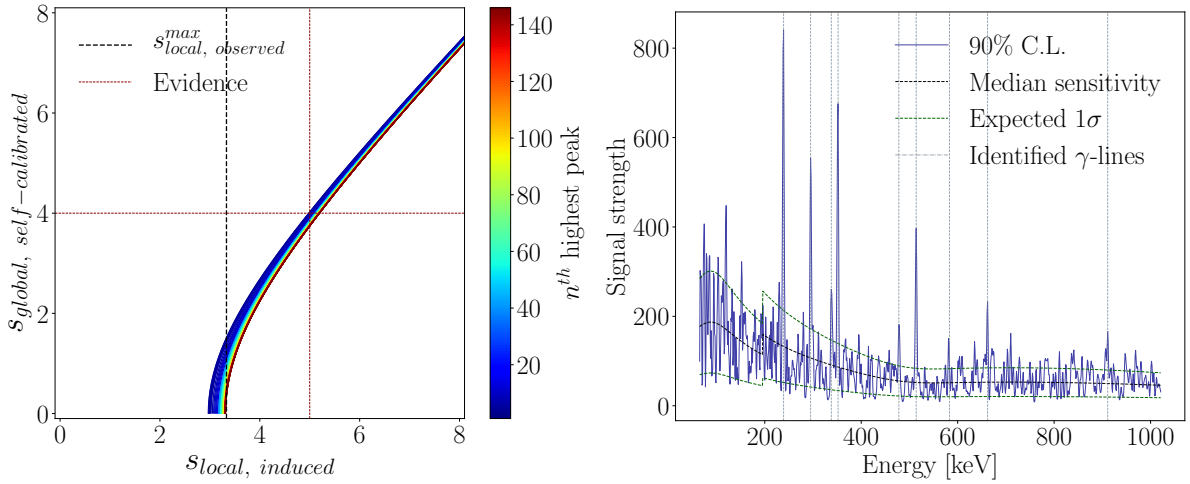

Figure 4.21: Left: plot of the global significance of artificial peaks induced in the data with a fixed local significance, and subsequently self-calibrated based on the n^{th} highest noise peak. Here, $n \in \{2, 3, ..., 146\}$, i.e. all peaks with an equal or higher than $1\,\sigma$ excess are considered, whereas the highest excess is blinded. See Ref. [204] for more details. Right: upper interval limits at 90% CL, shown in blue. Both the median sensitivity and its expected $1\,\sigma$ fluctuation band are shown as well. The positions of the identified γ-lines are indicated by the dashed, grey lines. Note that the limits are constrained to the positive regime.

Looking at Fig. 4.21, left, we readily realise that no globally relevant significance for the peak with the highest local significance, hence also not for any other peak, has been found. From the plot, we deduce that the excess has a global significance of roughly $1\,\sigma$ only, with an uncertainty of around $0.5\,\sigma$, i.e. way below our required evidence threshold. This statement holds independently of how many noise peaks we use for the comparison. To summarise, we conclude that we did not see any evidence for an excess beyond the expected noise fluctuations in the data apart from known isotope transitions. Thus, we always determined confidence intervals at 90% CL on the signal strength, one for each given input mass. Fig. 4.21, right, shows the obtained count limits as well as the median sensitivity and the expected 1σ range of all limits.

4.5 Results

4.5.1 Bosonic dark matter

We did not observe any globally significant monoenergetic excess above the continuous background potentially caused by bDM interactions in the GERDA detectors. We thus converted the evaluated count strength limits at 90% CL (Sec. 4.4) into an upper limit on the maximum physical interaction strength of the two considered bDM candidates, vectors and pseudoscalars.

We used the combined rate of photoelectric-like absorption and dark Compton scattering from Eqs. (4.2), (4.1) to obtain a limit on the total physical interaction strength. Let us repeat that the new feature compared to the former GERDA study [161] and to our knowledge, compared to any other direct bDM analysis performed by an experiment directly so far, is the inclusion of the Compton scattering process. Apart from this, we note that a recalculation of the constants of proportionality in the absorption case (Sec. 4.1.1) yielded more precise estimates of 4.683 instead of 4 for Eq. (4.2), and of 1.466 instead of 1.29 for Eq. (4.1), respectively, compared to the former study. The same numerical values have been published upon independent recalculation e.g. in Ref. [205]. Furthermore, we substituted $A_{\mathrm{tot}} = f_{76}A_{76} + f_{\mathrm{nat}}A_{\mathrm{nat}} \approx 75.5\,\mathrm{g/mol}$ into the calculation of the actual number of Ge atoms entering the target mass density, i.e. we considered the enrichment fraction of ^{76}Ge. As a third adjustment, when converting into physical limits, we used the experimental values for the photoelectric cross section of Ge from Ref. [206], instead of the phenomenological approximation used in Ref. [161].
The limits obtained upon conversion of the results from Sec. 4.4 are depicted in Fig. 4.22. There we also put the results into perspective with other leading direct detection experiments, indirect X-ray or γ-ray measurements, and the phenomenological constraint on the particle stability. In Tab. 4.3, we tabulate the detection principle and material used by each direct experiment, as well as the data exposure and the analysed bDM particle mass range, respectively.

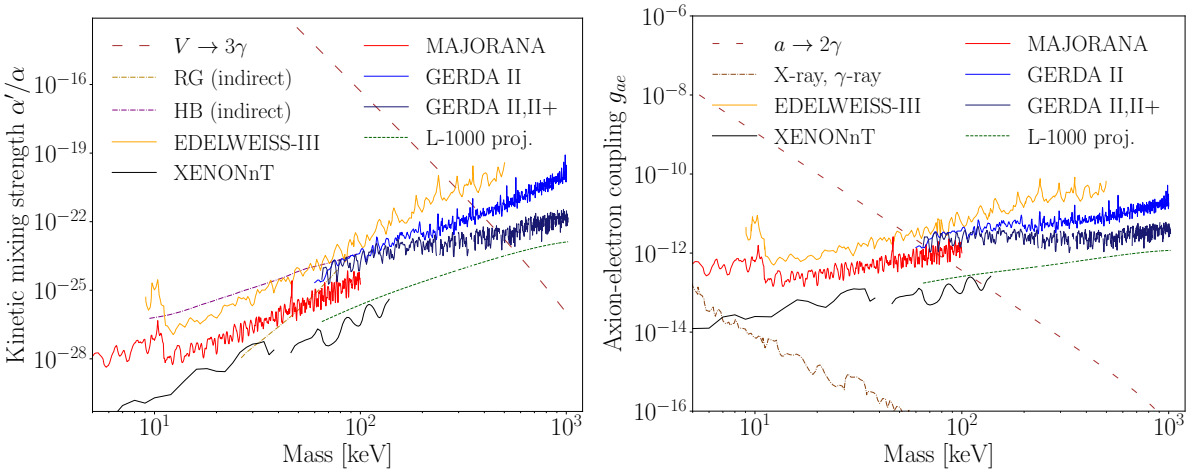

Figure 4.22: Left: constraints at 90% CL on the kinetic mixing strength of vector bDM. Right: constraints on the coupling strength of a pseudoscalar bDM to the electron. Both interaction channels, absorption and scattering, have been taken into account in both figures. The limits are set into perspective with results from other direct detection experiments, with the phenomenological stability condition on the scale of the age of the Universe (dotted, red) [158], and with indirect X-ray or γ-ray observations (dotted, red in the right plot) from astrophysical sources, red giants (RG, dash-dotted, yellow in left plot), and horizontal branch stars (HB, dash-dotted, violet in left plot), taken from Refs. [162, 158]. We also provide a projected sensitivity curve for L-1000 as evaluated via rescaling of our empirical fit model, assuming deployment of the HPGe detectors in an UGLAr shield.

Looking at Fig. 4.22, we can deduce that GERDA provides stringent direct bounds from slightly below 200 keV up to the MeV range, exceeding previous results by up to two orders of magnitude in the vector, and almost one order of magnitude in the pseudoscalar bDM case, respectively. However, the sensitivity increase relevantly drops at lower energies, dropping below the sensitivity of other experiments at the $\mathcal{O}(10\,\mathrm{keV})$ range. These phenomena can be

Experiment	Detector	Exposure	Mass range [keV]
GERDA II [161]	Ge diodes	14.6 kg yr	60-200
		58.9 kg yr	200-1000
GERDA II,II+ - this work	Ge diodes	67.7 kg yr	65-195
		116.7 kg yr	196-1021
Edelweiss-III [163]	Ge diodes	1149 kg d	0.8-500
MAJORANA DEMONSTRATOR [164]	Ge diodes	37.5 kg yr	1-100
XENONnT [165]	Xe-based dual-phase time projection chamber	1.16 t yr	1-39, 44-140

Table 4.3: Summary of the direct detection experiments from Fig. 4.22, listing the detection technology, the exposure, and the analysed mass ranges for the bDM search.

understood as follows. Limited by the orders of magnitude smaller exposure, GERDA is not competitive to the Xenon-based experiments. Also, the high ^{39}Ar background dominating the low energy part prevents us from exceeding the sensitivities of e.g. MJD at these lower energy values. Approaching the MeV range, though, the Compton process starts to dominate the total cross section with increasing energy, which triggers the strong improvement. Additionally, the exposure increase of approximately a factor of two compared to the previous study contributes to a marginal additional improvement of around $\sqrt{2} \approx 1.41$, implied by the assumption of Poisson fluctuations in the background. Compared to other principles, we see that the indirect constraints cannot be surpassed, and especially the pseudoscalar case is almost exclusively constrained by the lifetime condition to be stable over the age of the Universe. For the vector bDM, this constraint only sets in beyond approximately 500 keV. It is important to add that indirect and phenomenological constraints may suffer from severe systematic uncertainties or model-dependence, respectively. As an example, in Ref. [158], it had been pointed out that the stability condition may be circumvented with certain parameter fine-tuning in the ALP sector. Hence, all our direct measurement results are shown independently of how promising the phenomenological implications are. Physicists from the phenomenological sector have also reinterpreted absorption-only constraints for other purposes than DM models [207]. To provide input for such studies, we also indicate the individual effects of the absorption and the Compton processes separately in Fig. 4.23. There we also compare our constraints derived in the Frequentist statistical framework to two other GERDA analyses pursued in a Bayesian setting, the former study from Ref. [161] conducted at UZH, and a study by our colleagues from UniPD. The latter was pursued in collaboration with this work, and is also part of the prepared publication that is under collaboration-internal review (as of summer 2023).

4.5.2 Particle disappearances

We have neither identified any significant excess at the expected γ-line energies for the semi-visible electron decay implying charge non-conservation, nor for the ^{75}Ge decay following a single nucleon disappearance out of a ^{76}Ge nucleus indicating baryon-number violation. As shown in Sec. 1.5, we can derive limits on the investigated particle decay channels via conversion of the determined count strength limits into physical quantities by using the radioactive decay law. Given the order of magnitude of the lifetimes to be obtained beyond 10^{23} yr and higher, the expected event number for the rare processes described by the exponential decay law can be

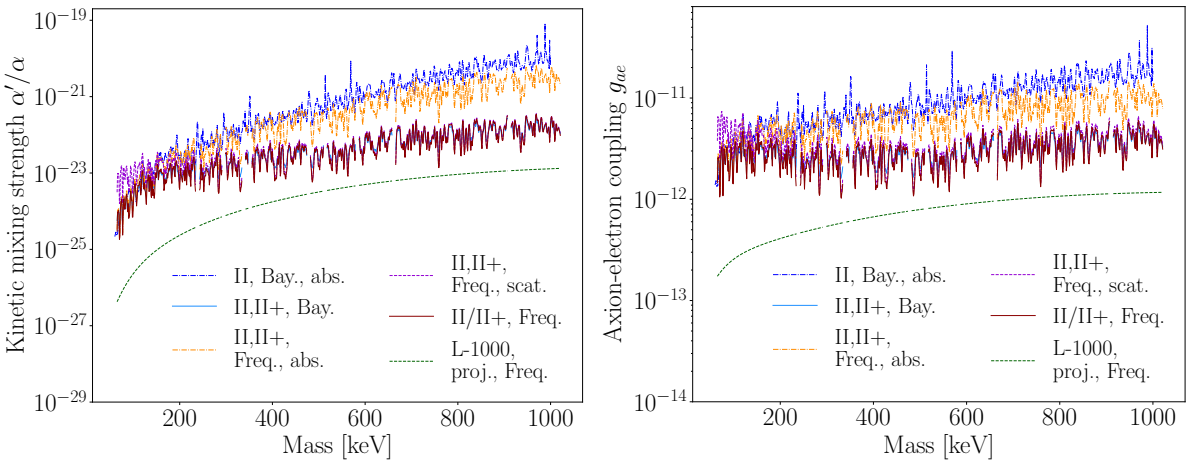

Figure 4.23: Left: overview of new and old Bayesian limits at 90% credible interval and Frequentist limits at 90% CL, on the kinetic mixing of vector bDM, plotted against the bDM mass. Right: comparison of the same limits for pseudoscalar bDM. For comparison, limits evaluated from the absorption or the scattering case only are illustrated, as well as a projection of the reach of L-1000 based on rescaling of our empirical fit model assuming UGLAr.

approximated very accurately by a simple first order Taylor expansion,

$$N = \epsilon n_{(76)} N_{(76),0} \left(1 - e^{-\frac{t}{\tau}}\right) \approx \epsilon n_{(76)} N_{(76),0} \frac{t}{\tau} = \epsilon n_{(76)} \frac{\lambda f_{(76)} N_A}{M_{(76)} \tau} \ . \tag{4.12}$$

We then derive the lower constraints on the particle lifetime τ based on the observed upper limit on the event number N_{up} as

$$\tau_{\text{low}} = \epsilon n_{(76)} \frac{\lambda f_{(76)} N_{\text{A}}}{M_{(76)} N_{\text{up}}} \ , \tag{4.13}$$

where ϵ denotes the corresponding detection efficiencies for the investigated channel, λ is the exposure, and N_{A} is Avogadros' constant. The index (76) indicates the channel dependency of the quantities nuclear mass M, enrichment fraction f, and number of source particles in each atom n. The index is relevant for the ^{75}Ge decay following a nucleon disappearance, but irrelevant for the electron decay channel analysed in all Ge isotopes. On the one hand, we evaluated the electron decay limit using $n = 32$, taking into account the contributions from all atomic shells in our signal model (Sec. 4.2). On the other hand, we stated a lifetime limit for the stability of the ^{76}Ge nucleus against a nucleon disappearance using unity $n_{(76)}$ instead of a lifetime of a neutron or a proton directly. This over-conservative interpretation may be relaxed in a manner similar to analyses conducted in e.g. Refs. [208, 209, 210, 211]. There effective number of nucleons inside the corresponding nucleus available for the decay have been estimated based on nuclear shell considerations. Applying a similar treatment, we would need to ensure that, when a nucleon disappears, no subsequent release of further nucleons or α particles occurs, since such emissions would prevent the ^{75}Ge production, either directly or via ^{75}Ga in the case of a disappearing proton, respectively. Assuming a nuclear shell model to accurately model a ^{76}Ge nucleus, we may expect only the less strongly bound outer shell nucleons fullfill the requirement. Taking 28 as a magic number of stable nuclei in the inner shell, we would then have $44 - 28 = 16$ effective neutrons and $32 - 28 = 4$ effective protons left. Hence, as less conservative estimates, our main lifetime limits on the single nucleon decays in ^{76}Ge may be multiplied by these factors. Let us remark that potentially even more nucleons could be

available, if bound less strongly than the energy released upon the nucleon decay. More precise considerations of the nucleon binding energies and their impact on the limits on the nucleon lifetimes were conducted by members of the collaboration working together with colleagues from the nuclear physics community, yielding more accurate factors of 16 effective neutrons and 14 effective protons. As these are simply constants of proportionality, the over-conservative estimates can be conveniently scaled linearly in future studies as soon as the results of their derivations will be published.
Regarding the actual observations, in Sec. 4.4, we obtained a marginal excess for the nucleon decay channel, but with a significance of $0.7\,\sigma$ only, induced by a signal strength of 4.2 counts. Having obtained a corresponding event count limit of 15.2 counts versus a sensitivity estimate of 9.9 counts, the conversion scheme described above yields a lifetime limit of 9.8×10^{22} yr, with a sensitivity of 1.5×10^{23} yr, as two-sided intervals at 90% CL. For the 256 keV Doppler- and resolution-broadened γ-line caused by a potential electron decay in Ge, we have not seen an excess at all but a strong under-fluctuation with an amplitude of -231.8 counts. The interval evaluation yields a limit of 95.5 counts versus a sensitivity of 268.5 counts. Thus, the conversion formula of Eq. (4.13) yields a lower lifetime limit of 1.2×10^{26} yr at 90% CL and a corresponding sensitivity of 4.2×10^{25} yr. We summarise the obtained constraints on the particle lifetimes in Tab. 4.4.

Particle	Signal count strength	Significance [σ]	Upper count limit / sensitivity	Lower lifetime limit / sensitivity [yr]
Electron	-231.8	0	95.5 / 268.5	$1.2{\times}10^{26}$ / $4.2{\times}10^{25}$
Neutron, proton	4.2	0.7	15.2 / 9.9	$9.8{\times}10^{22}$ / $1.5{\times}10^{23}$

Table 4.4: Overview of the determined lower lifetime limits at 90% CL for electrons and nucleons, together with the best-fit signal strength, the corresponding significance, and the upper count limit. Whereas a strong under-fluctuation in the vicinity of the expected electron decay energy leads to a limit severely exceeding the median experimental sensitivity, the presence of a small, non-significant peak at the nucleon decay energy causes a degradation of the limit with respect to the sensitivity, see Sec. 4.4.

The determined lower lifetime limit on the stability of a single nucleon inside the isotope ^{76}Ge is the only limit on these processes found in the literature. Comparing the measured GERDA result to sensitivities to the free single nucleon decay or the decay in any isotope, as obtained by, for example, the large-scale experiments mentioned in Sec. 4.1.2, however reveals that we remain several orders of magnitude below the competitive regime. This is expected, as we are constrained by our small exposure and higher background rate at low energies compared to the tonne-scale experiments. Our electron lifetime limit in Ge is similar to the result derived from the data of the Heidelberg-Moscow experiment by Klapdor-Kleingrothaus et al., and surpasses the limit by comparable experiments such as MJD or EDELWEISS. However, also for this channel, with the GERDA data, we do not reach the sensitivity of Borexino or other large scale experiments deploying other elements, as described in Sec. 4.1.3. Yet, let us emphasise that our results deduced from a small data set of $\mathcal{O}(100\,\mathrm{kg\,yr})$ exposure only will be relevantly strenghtened if similar analyses are pursued in L-1000, as we motivate in the following section.

4.6 Projection for Legend-1000

Having developed a background model for the data measured by GERDA, we scaled this model to explore the sensitivity reach of the future L-1000 experiment to the bDM interactions and the single particle decays.

4.6.1 Bosonic dark matter

In Figs. 4.22 and 4.23, we illustrate the sensitivity reach of L-1000. We project this line as follows. Firstly, we fit the empirical background model shown in Sec. 4.2 simultaneously to GERDA's medium-energy and low energy data below 145 keV, rescaled to the total Phase II and II+ exposure. The region between 145 and 183 keV was neglected here because of different data taking thresholds during runs just before the minor hardware upgrade discussed in Sec. 2.2. In this way, we obtained a single approximate background model for the full energy range. Including the identified lines in the fit model, this yielded a χ^2/dof goodness-of-fit estimate of 1.19 and a KS-test statistic value of 79.74, corresponding to a p-value p_{KS} of 0.69. The residuals were nicely distributed around zero, as indicated by KS and SW hypothesis test results of 0.026 and 1.00, corresponding to p-values of $p_{\mathrm{KS}} = 0.57$, and $p_{\mathrm{SW}} = 0.36$, when comparing the distribution of the residuals versus normality. This is illustrated in Fig. 4.24, where we also show a normal distribution fit on the residuals. The fit is centred at 0.04 with a standard deviation of 1.09, meaning that the fit result is very close to the expected standard normal distribution. Overall, the goodness-of-fit estimates again do not indicate any relevant broad, i.e. non-localised, discrepancy between the model and the data. Note that in correspondence with our analysis strategy for GERDA, the γ-lines are not part of the background model. Consequently, we did not derive any physics constraints in the seven 1 keV bins interval centred at the bin of the known line position.
Having built the model function for our prediction, we then assume the detector operation for L-1000 to take place in reentrance tubes filled with UGLAr, in contrast to the natural LAr deployed in the entire GERDA cryostat (Chapter 2). Thus, we reduced the amplitude of the ^{39}Ar component of our fitted background model by a factor 1400, in correspondence with an upper limit on the measured ^{39}Ar contamination of the UGLAr as published by the DarkSide collaboration [114]. As a next step, we scaled the amplitude of the fit to the nominal L-1000 exposure of 10 t yr. The L-1000 prediction model determined in this manner is plotted in Fig. 4.24.

Subsequently, we generated an Asimov data set based on the fit function. Finally, we followed the same procedure as outlined in Sec. 4.4 to determine the confidence interval at 90% CL at all energy values not falling into the deleted regions of identified γ-lines tabulated in Tab. 4.2, yielding our approximate sensitivity prediction as plotted in Figs. 4.22 and 4.23. Let us overview the assumptions made for the L-1000 sensitivity projections below.

Assumptions for the L-1000 sensitivity projections

- Nominal exposure: 10 t yr
- Efficiency: apply simulated values for GERDA Phase II+ IC detectors
- Energy resolution: use measured GERDA Phase II+ IC detector resolution curves
- Background contributions: after LAr veto cut, two components as for GERDA; ^{39}Ar and $2\nu\beta\beta$

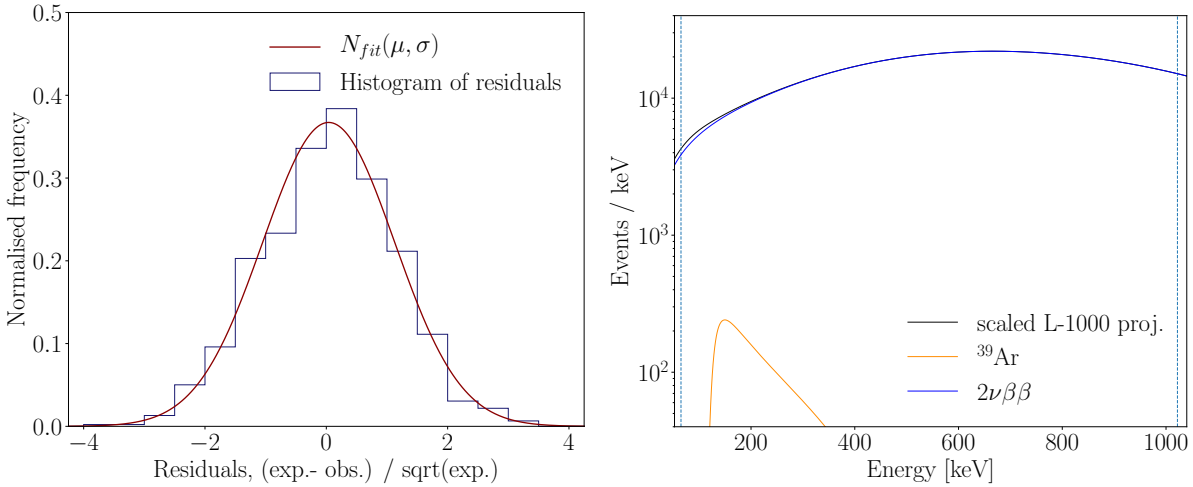

Figure 4.24: Left: histogram of the fit residuals for the empirical background model fit on the GERDA multiplicity-one data scaled to equal exposure, cf. Sec. 4.2. Note that the γ-lines identified in Sec. 4.4 have been included for this fit. Right: plot of the corresponding L-1000 prediction model obtained by rescaling the empirical background model for GERDA assuming that the detectors are directly operated in UGLAr. For simplicity, subdominant background contributions are fully neglected in the model. Here the γ-rays have been omitted, but we do not perform a search for a new physics signature in the vicinity of these.

- Background scaling: operation of HPGe detectors in UGLAr reduces ^{39}Ar beta decay rate by a factor of 1400; $2\nu\beta\beta$ decay rate as for GERDA; amplitudes exposure-rescaled with respect to GERDA

We emphasise that a simple background model scaling may not fully represent the power of L-1000 for these analyses. Neglecting other surely present but subdominant background contaminants may be considered aggressive. Nevertheless, our predicted sensitivity is expected to be conservative for the following reason. Without any relevant ^{39}Ar contamination, implying more than an order of magnitude lower background towards low energies compared to GERDA, L-1000 will be sensitive to the β particle-only peak in the Compton scattering case. The β event can be tagged both with or without a coincident scintillation or neighbouring detector event induced by the escaping γ particle, as we discussed towards the end of Sec. 4.3. Such an analysis does not only boost the overall detection efficiency, but additionally causes a double-peak signature consisting of one β particle -only and one combined β-γ pair-induced peak at known energies, meaning a very distinctive signature. On top of that, it is pointed out in Ref. [153] that the identification of a second monoenergetic peak at a correlated energy value strongly reduces the look-elsewhere effect, making a potential discovery claim more convincing. Lastly, the dominant remaining background, the $2\nu\beta\beta$ spectrum, causes single-site events. The combined β-γ pair Compton signal would be dominantly multi-site instead. This means that a dedicated PSD analysis in L-1000 extended to the low and medium energy regime may boost the sensitivity even further compared to the improvement we indicated in Figs. 4.22 and 4.23.

4.6.2 Particle disappearances

To estimate the sensitivity reach of L-1000 for the semi-visible decay of an electron, we reused the empirical, rescaled M1 data model used for the bDM analysis, precisely as done for GERDA.

Let us underline that the relevant search window, 196 to 316 keV, centred at the expected γ-line at $\sim m_e/2$, lies in a region that will no longer be dominated by the ^{39}Ar background but solely by the $2\nu\beta\beta$ spectrum depicted in Fig. 4.24, right. This obviously holds only for the L-1000 baseline design of operating the HPGe detectors directly in UGLAr being depleted in, among others, ^{39}Ar. Applying the Asimov approach as before, we evaluated a nominal sensitivity for the planned exposure of 10 t yr, which gives 1.5×10^{27} yr at 90% CL. This value clearly depends on the actual available low and medium energy exposure that will be accumulated in L-1000. Assuming Poisson statistics, we applied a simple square-root-scaling to estimate the median sensitivity for different exposure levels, as depicted in Fig. 4.25, left.

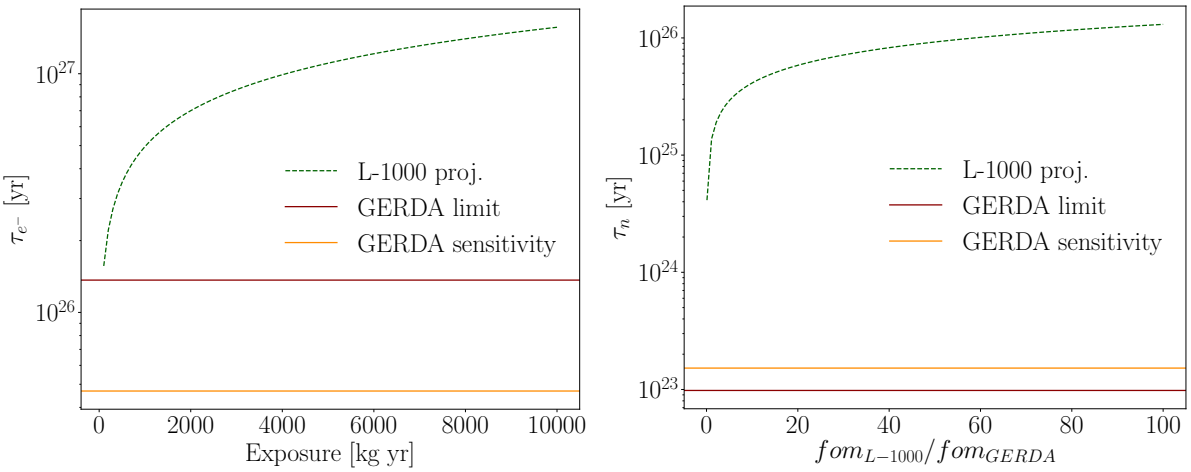

Figure 4.25: Left: scaling of the lifetime sensitivity of L-1000 to the semi-visible electron decay channel as a function of data exposure. The projected reach is based on a square-root-scaling of the 90% CL sensitivity expected from Poisson fluctuations in the background. The maximum sensitivity for the nominal exposure of 10 t yr was deduced from the rescaled empirical background model of GERDA assuming deployment of UGLAr. Right: dependence of the L-1000 sensitivity reach to the single nucleon decay as a function of the figure-of-merit (fom) improvement of L-1000 with respect to GERDA, assuming the over-conservative scenario of only one nucleon inside the ^{76}Ge nucleus decaying with the subsequent emission of γ particles without additional nucleons. Note that also here the exposure was set to 10 t yr, and the potential fom improvement is driven by an expected strong increase in the efficiency supposing that the subsequent ^{75}Ge decay can be directly searched for in M1 data sets. This method becomes available as the detector operation in UGLAr mitigates a relevant ^{39}Ar contribution.

This simple approach suggested that with approximately 4 t yr of exposure only, the sensitivity will already surpass the 10^{27} yr threshold. Again, the projection may be considered aggressive, as the empirical model deduced from the GERDA data neglected sub-dominant background contaminants. On the contrary, the mitigation of the ^{39}Ar β events opens up the possibility of a dedicated PSD analysis, potentially further improving the projected result. In the electron decay case, the expected γ-signal will share multi-site and single-site topologies (the energy is too low to expect quasi-only multi-site γ-events, which is the case for the γ particles with energies of $\mathcal{O}$(MeV)), whereas the $2\nu\beta\beta$ background will only induce single-site events, and may hence be at least partially suppressed given their distinct pulse shapes. Additionally, L-1000, if deploying low-noise front-end electronics with very low energy thresholds, may collect physics data down to energies of $\mathcal{O}$(keV) as done by MJD [164]. In this case, the invisible electron decay channel

$e^- \rightarrow 3\nu_e$ with a peak-like signature appearing at 11.1 keV may become accessible, opening an auxiliary search method.
For the single nucleon disappearance channel, the considerations towards L-1000 differ from the methods applied to GERDA's bDM and electron decay analyses using the M1 data. Let us repeat that the main ideas of searching for ^{75}Ge, the product of a single nucleon disappearance, via an M2 signal are the reduction of the potential background and a clearly detectable clean, i.e. symmetric, monoenergetic peak. Hence, as described earlier, we looked for such a ^{75}Ge β-decay in coincidence with a 265 keV γ-line signal from the de-excitation of the ^{75}As daughter. The downside of this coincident tagging is the low efficiency of around 0.19% only. The reason for this low value is that the chance for a γ particle to escape its source detector, to propagate through the LAr, to penetrate a neighbouring detector, and then to deposit its initial full energy there is very low. In fact, this process only works for γ particles being emitted close to the surface of the source detector. L-1000 however does not need to rely on this tagging technique with its limited efficiency. Without a relevant contribution of ^{39}Ar β-decays, meaning with a strongly reduced background level, L-1000 can perform the search for this channel directly in the M1 data as well. In this scenario, one looks for single-site β events from the ^{75}Ge decays within the source detector only. Assuming the GERDA geometry for a L-1000-like exposure, we expect an efficiency improvement by approximately a factor of 300. This value is based on additional simulations performed by our collaborators at UniPD. We applied these simulations to perform the corresponding M1-only search method for ^{75}Ge events in the GERDA data as an alternative search method. The analysis is added to this work as supplementary material in Appendix B.1. Even though the obtained limits and sensitivities are very similar to the M2-based search method, in the main branch, we focused on the latter because of its monoenergetic signal signature. Obviously, in the M1-based approach, the background level rises if one takes the entire β energy range as a potential signal interval. However, it is important to understand that if we interpret our lifetime conversion formula, Eq. (4.13), as a figure-of-merit (fom), Poisson statistics suggest that the background level rises only with the square root of the event count. The reason is that the count limit in the denominator can be, very roughly, approximated as the square root of all events in the signal region under the assumption of a Poisson distribution. The efficiency, however, enters the equation *linearly*. As a result, in the absence of ^{39}Ar contaminants, L-1000 can definitely employ this method to boost the sensitivity, as the background rise will not be as strong as in GERDA. As a second alternative option, to mitigate a large fraction of background events in M1 data, L-1000 may search for the coincident events in which the de-excitation γ particle escapes its source detector and loses its energy inside the LAr without reaching the neighbouring detector. The corresponding GERDA simulations suggest that with this method, the efficiency will still be improved by approximately a factor of 30, but not causing a background level as high as in the M1-only method. In addition to the different search methods, a PSD cut specifically tailored to the entire keV range, may allow L-1000 to separate signal from background events further. To apply PSD cuts one would need to be aware that depending on the search method, the signal topology changes, as β and γ particles cause mostly different pulse shape patterns (Sec. 2.3). Thus, from a quantitative point of view, it is unclear how precisely the extension of PSD analyses will impact the overall sensitivity to the single nucleon decay. We indicate the potential sensitivity reach achievable in L-1000 in Fig. 4.25, left, where we leave, admittedly slightly speculatively, the actual fom improvement compared to GERDA as a free parameter. As a rather qualitative statement, we can expect the combined effects of exposure increase and novel search techniques boosting the efficiency to lead to a two to three orders of magnitude higher sensitivity of L-1000 as compared to the GERDA result.
For all sensitivity projections for the different new physics searches, a more precise quantita-

tive assessment and the impact of potential keV-scale PSD analyses would require dedicated simulations. These should preferably be conducted with experimental geometries realistically modelling the future L-1000 experiment and rely on appropriate waveform analyses.

Chapter 5

Summary and Conclusions

Solving the mystery of the imbalance between the matter and the antimatter in our Universe is pressing. An observation of processes violating the conservation of lepton number could shed light on this mystery. One manifestation of lepton number violation would be a Majorana mass term generating the neutrino mass, which is known to differ from the value of zero assumed in the Standard Model given the measurements of neutrino oscillations. Neutrinoless double beta decay ($0\nu\beta\beta$) is the most promising experimental path to observe the creation of matter in the laboratory without antimatter and to consequently shed light on the question of the nature of the neutrinos and its mass.

The LEGEND experiment, a merger of the GERDA and the MAJORANA DEMONSTRATOR experiments, combines the experience of its predecessors, obtaining the lowest background level and the best energy resolution measured in the entire field of $0\nu\beta\beta$ decay searches. These parameters are crucial to perform a measurement in the quasi-free background regime, in which the experimental sensitivity scales quasi-linearly with the data exposure, as opposed to background-limited sensitivities scaling only with the square root of the exposure. By operating an array of high-purity Ge (HPGe) detectors enriched up to 90% in the target isotope ^{76}Ge inside an active liquid argon (LAr) shield, LEGEND will combine high signal detection efficiency with the lowest background levels achievable. We foresee a 3σ discovery sensitivity on the half-life of the decay exceeding 10^{28} yr, most likely covering the entire inverted neutrino mass ordering regime.

A key requirement for a successful experimental campaign is the regular accurate determination of the energy scale of the HPGe detectors and their measurement resolution in the vicinity of the Q-value of the decay $Q_{\beta\beta} = 2039.1\,\text{keV}$ via calibrations. In addition, the pulse shape analysis discrimination for background event suppression based on the event topology must be tailored and validated frequently. Illumination of the HPGe detectors with radioactive sources is thus regularly performed on an around weekly schedule between physics measurement periods. Based on the experience gained for GERDA, at the University of Zurich, we developed and produced five dedicated source insertion systems, featuring a movement control of a steel band in a cryogenic environment with a reliable position precision of a few mm. Four of these systems immerse four radioactive ^{228}Th sources per system into the cryostat of the first phase of the experiment, LEGEND-200, when a HPGe detector calibration is performed, and a fifth unit serves as a local backup for training purposes. I contributed to the hardware component testing, the assembly, the mounting, and the commissioning of the systems, as well as to the monitoring of the system performance and the corresponding maintenance work. All tasks provided valuable information for the development of the calibration systems for the future second stage of the experiment, LEGEND-1000.

I also investigated the illumination properties in terms of the γ-event homogeneity in the full

energy peak, the single-escape peak, and the double-escape peak of ^{208}Tl, an isotope present in the ^{228}Th decay chain, among the detectors in the array during the LEGEND-60 commissioning stage. I found acceptable, though not fully homogeneous event distributions, with particularly reduced event counts in the bottom-most detectors. This motivated an extension of the length of the nylon tubes guiding the sources next to the detector array, to improve the overall uniformity. Based on these commissioning data from several calibration runs, I also deduced overall good detector performances with full width at half maximum resolutions at $Q_{\beta\beta}$ below 3 keV, without any degradation for high-mass inverted coaxial detectors. These results are particularly important as the optimised surface-to-volume ratio of these detectors enables a lower background level.

Deploying the radioactive sources, emitting neutrons, may cause the activation of the detectors or the surrounding materials, producing isotopes potentially mimicking a signal. During the production process by our collaborators at Los Alamos National Laboratories, the radioactive source material was thus deployed on gold foils, featuring an (α, n) reaction threshold exceeding 9 MeV, i.e., higher than the maximal energy of any α particle in the decay chain of ^{228}Th. To determine the remaining neutron flux, we conducted a full neutron measurement campaign with a LiI(Eu) neutron counter underground at Laboratori Nazionali del Gran Sasso. We calibrated the detector with an ^{152}Eu γ-source, determined the thermal neutron detection efficiency with an ^{241}Am^{9}Be source, and measured the background level for more than half a year. In two separate, approximately two-month data taking periods, we then measured the emitted neutron strength of the radioactive sources. I carried out the related analysis, obtaining a global neutron flux, taking into account correlations in the data, of

$$\Psi = (4.30 \pm 0.69_{\text{stat}} \pm 0.93_{\text{syst}}) \times 10^{-4}\ \text{n} \,/\, (\text{kBq}\,\text{s})\ , \tag{5.1}$$

which is around one order of magnitude lower than the neutron flux of commercially available sources. By running dedicated subsequent simulations of the neutron emission rate during calibrations based on this result and on the subsequent decays of activated isotopes in and around the LEGEND-200 detector array, I found that the background level induced by the neutron emission of the sources is, after background suppression cuts, suppressed by around three orders of magnitude with respect to the background goal of LEGEND-200. We concluded that the custom-made radioactive sources feature a sufficiently low neutron emission rate for a safe operation in LEGEND-200. Considering a factor of 20 more ambitious background goal, plus vastly different configuration options for the detector array potentially causing higher neutron capture rates in the future LEGEND-1000 experiment, I strongly encourage the collaboration to perform similar neutron flux measurements, and to run the corresponding simulations in the relevant geometrical setups. These will be needed to ensure negligible background contributions also in the future, or in the case of a negative result, to address the source production process or the calibration procedures, if necessary.

LEGEND's predecessor GERDA developed the concept of the bare HPGe detector operation inside LAr. The experiment still holds the world-leading $0\nu\beta\beta$ half-life discovery sensitivity obtained in any isotope. With a sensitivity equaling the final measured limit of $T_{1/2} = 1.8 \times 10^{26}$ yr, it was the first semiconductor-based experiment to exceed the 10^{26} yr barrier, and to operate in the quasi-background-free regime. The low-background environment of GERDA, in combination with the excellent resolution of HPGe detectors, encouraged us to use the available high-quality physics data to search for monoenergetic signatures potentially caused by beyond Standard Model processes other than $0\nu\beta\beta$. These comprised potential dark matter particle interactions and particle decays violating baryon number or charge conservation.

In GERDA, we performed a dedicated search for imprints of bosonic keV-scale dark matter

(bDM) in the low energy regime of the data. As a novel feature, we did not only look for signals that might have been caused by the photoelectric-like absorption of bDM, but also by the dark Compton scattering, improving the sensitivity by up to two orders of magnitude towards masses around 1 MeV. For these analyses, I developed an empirical background model in the low-energy regime, for the first time in GERDA, taking into account the dominant physical background processes, namely the ^{39}Ar and the two-neutrino double beta decay ($2\nu\beta\beta$), and performed the analysis in the Frequentist statistical framework. I did not identify any *globally* relevant excess, which in all likelihood was not caused by expected γ-ray transitions. Let me underline that I considered the look-elsewhere effect, expecting to induce local excesses caused by pure noise fluctuations among the large number of trial experiments, which led us conclude that no evidence was found. We thus, also for the first time in GERDA, evaluated Frequentist intervals on the allowed coupling strength and kinetic mixing of pseudoscalar and vector bDM, respectively, for each mass model in the range 65 to 1021 keV individually. In the case of pseudoscalar bDM, our lower limits on the coupling strength to electrons vary from approximately 1 to 6 $\times 10^{-12}$. These results are not sensitive enough to probe the parameter space needed for the stability of these particles on a timescale of the age of the Universe. For vector bDM, however, we obtained constraints on the kinetic mixing strength in the range of approximately 1.8 $\times 10^{-25}$ to 3.5 $\times 10^{-22}$. These outcomes make GERDA the world-leading fixed-target experiment for vector bDM in the mass range between around 150 and 450 keV, which covers parameter space allowed by the required particle stability considerations.

Moreover, we also used the low energy data and the background model developed by myself to search for the charge conservation violating decay of an electron. We found no excess at the expected energy of $\sim$256 keV, and thus evaluated an interval of 1.2×10^{26} yr at 90% confidence level (CL) with a corresponding sensitivity of 4.2×10^{25} yr. As a third search for new physics, we studied the presence of the ^{75}Ge isotope. It could be produced by the disappearance of a single nucleon from the ^{76}Ge nucleus in a HPGe detector, either directly if a neutron disappears, or secondarily upon ^{75}Ga production if a proton decays. We searched for coincident two-detector events, triggered by the emitted β particle from the decay of the ^{75}Ge isotope, which would be accompanied by a ^{75}As de-excitation γ-ray with an energy of $\sim$265 keV in 11.5% of the cases. The latter could then be detected in a detector surrounding the source detector. Also for this channel, no significant signal was found, and we set an (over-)conservative interval on the neutron or proton disappearance in the ^{76}Ge nucleus of 9.8×10^{22} yr at 90% CL with a sensitivity of 1.5×10^{23} yr. Let me emphasise that these values can be relevantly improved upon precise calculations of the effective number of nucleons decaying with the subsequent emission of γ particles only, i.e., without any additional nucleon emission.

All the obtained constraints are among the most stringent measurements obtained with semiconductor detectors, but do not yet reach indirect bDM estimates or those of tonne-scale experiments, in particular in terms of the probed decay channel modes. However, from our *measured* results, I estimated the sensitivity reach of the future LEGEND-1000 experiment and commented on potential extensions of the analyses, which may boost the sensitivities further. Benefitting from a suppressed ^{39}Ar background by deploying the HPGe detectors directly inside underground-sourced argon, LEGEND will be able to close the gap to the orders of magnitude where the currently most-stringent results lie. Hence, our studies reveal excellent physics opportunities of HPGe-detector-based experiments ahead.

Fully making use of the future high-quality experimental data of LEGEND, an exciting time lies ahead for the collaboration. Probing the lepton-number violating $0\nu\beta\beta$ decay deep into unexplored territory and widening the physics program to any feasible dark matter or conservation-law violating search channel, LEGEND will provide input into the fundamental properties not

only of neutrinos, but into several phenomena going severely beyond our current understanding of particle physics and cosmology.

Appendix A

Supplementary material for the LEGEND-200 calibration hardware and the corresponding control software

A.1 Technical drawings of the LEGEND-200 calibration hardware

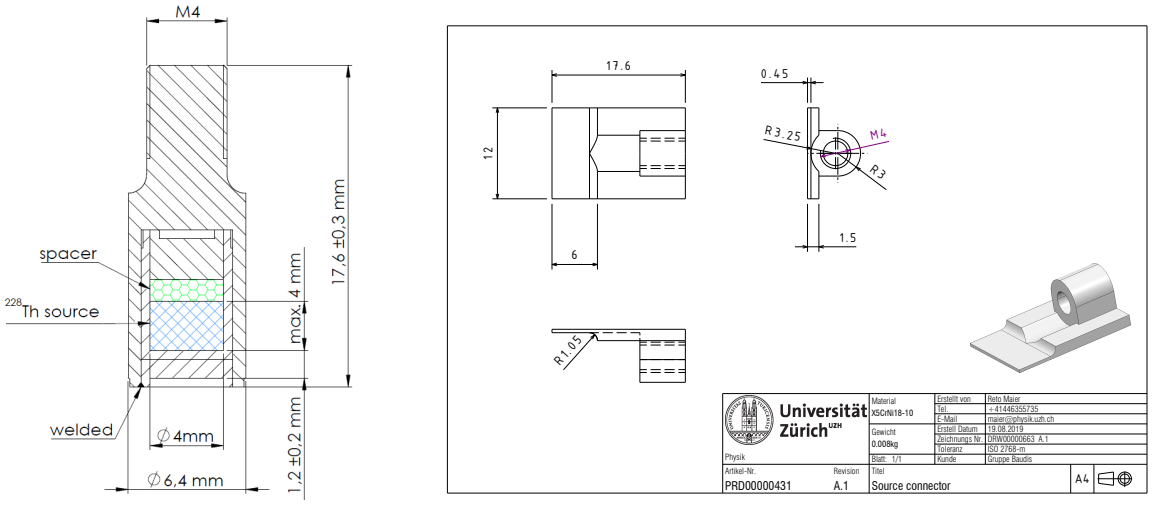

Figure A.1: Left: drawing of the radioactive calibration sources for L-200, as published in Ref. [1]. The sources were produced at LANL, following the UZH design. Drawing provided by UZH technician J. Franchi. Right: technical drawing of the source holder attached to the SIS stainless steel band via resistance spot-welding. The holders for the sources were produced at UZH directly. Drawing made by the head of the UZH workshop R. Maier.

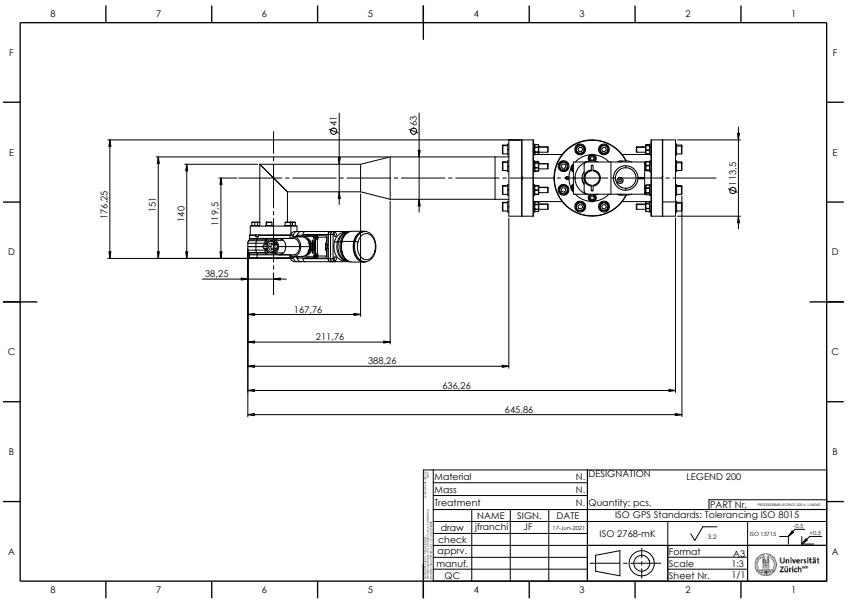

Figure A.2: Drawing of the L-200 SIS used to deploy the calibration sources designed and machined at UZH, with the exception of certain subcomponents, which were purchased commercially. The drawing was made by J. Franchi, and the overall SIS design was made by the UZH workshop.

A.2 Source insertion system control code

Here, we provide an overview of the working principles of the SIS control code developed to perform general hardware and dedicated positioning testings. Via serial communication, a set of control commands can be sent to the microcontroller controlling the power supply of the motor. The commands are encoded via specified byte arrays, which are transferred from a user's computer via a USB-to-RS422 serial converter. The microcontroller firmware was developed by the electronic workshop of UZH already for GERDA Phase I, and the commands have not been changed. For the brief summary and the code explanations given here, we follow the content of Ref. [124]. Active and passive commands are available, plus a set of protected commands and hidden commands only available for developers. Protected commands cause permanent changes on the microcontroller storage and are hence password protected. Among others, these comprise correction values for the calibration of absolute encoder positioning system or the effective band thickness parameter (Sec. 3.2). Also other physical properties such as the band length, the cryogenic temperature correction values, and the inner and outer diameter of the steel roll are protected. Hidden commands are callable only at the front panel directly. These allow to change the controller operation mode. For example, one can enter an error-debugging mode instead of the standard user mode. This alternative mode is used to investigate SIS malfunctions. It can also be employed for the calibration of the absolute encoder. Upon commissioning of the SIS and its positioning systems, the following passive and active commands are relevant for standard procedures like actual calibrations or testings.

- GET STATUS - passive
  The status of the control box and of all units connected to it is returned. The information comprises the running firmware revision, the supply voltage (standard 24 V), the operation mode, the motor currents, the front panel status, the internal option switch status, the

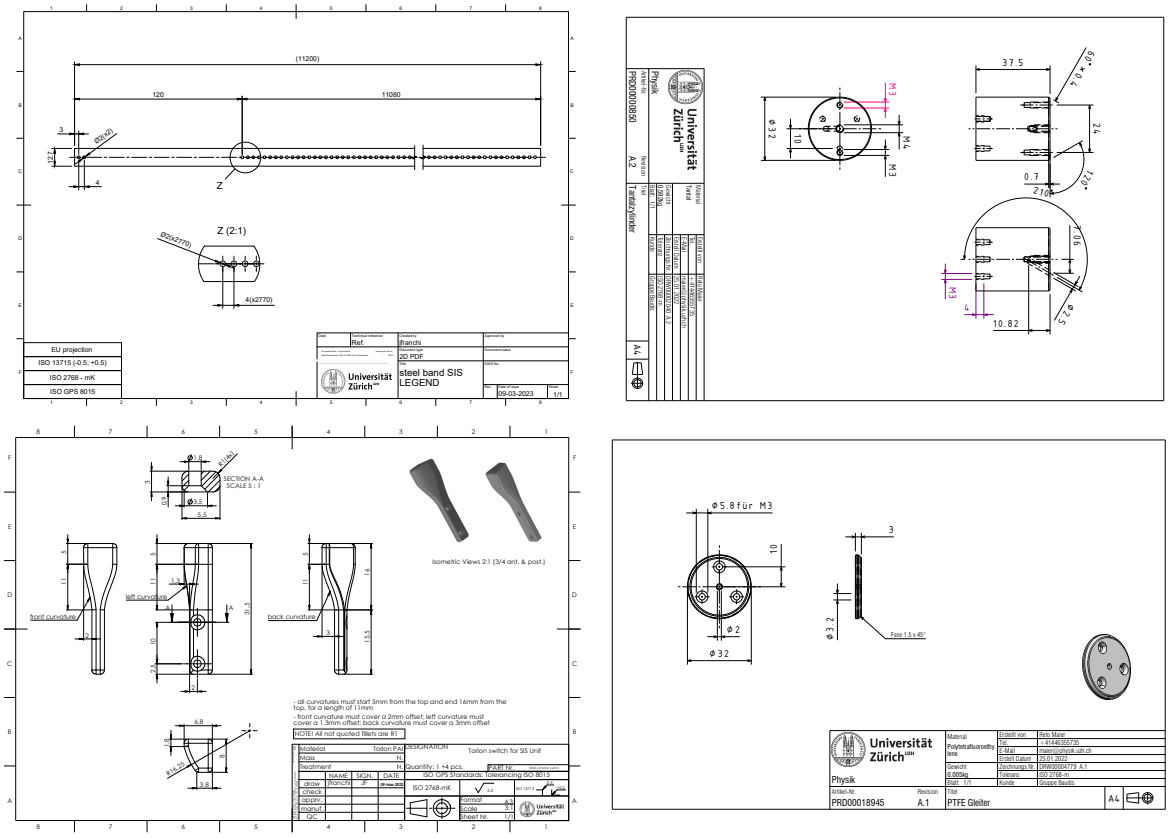

Figure A.3: Drawings of certain hardware components of the L-200 SIS as designed and produced in-house at UZH. On the top, the stainless steel band (left) and the Ta absorber (right), and at the bottom, the Torlon® pin for initialisation (left), and the PTFE sliding plate below the absorber are shown (right). The top right and bottom right drawings were made by the head of the workshop R. Maier, and the top left and bottom left ones by J. Franchi.

motor movement status, the initialisation status, the operation mode, and the error flags.

- GET POSITION - passive
  The positions of both position sensing systems of all initialised units connected to the control box are returned (in mm). For non-initialised units, the returned value of the incremental encoder is -9999 instead. The absolute encoder position is permanently stored, and thus available also before initialisation. Also, invalid positions are masked as -9999.

- INIT - active
  The desired SIS unit is initialised, meaning that it is moved up to the initialisation pin, hits the end-switch, and then moves down 4 mm to the parking position 0 mm defined via this procedure. The positions of both sensors are (re-)set to zero. This is the first active command to be performed upon system startup.

- STOP - active
  During movements, the specified unit is stopped by disconnecting the power to the corre-

sponding motor.

- GOTO POSITION - active
  The specified unit is moved to the desired position (in mm), where the position is defined via one least and one most significant byte. When the incremental encoder position aligns with the requested position, the motor power is stopped.

Note that sending active commands to the microcontroller via remote control is only accepted if remote control access button is activated, and some details as listed above are only valid in the standard operation mode. We developed a framework of several scripts written in Python3 code for specific small tests or for dedicated tasks. Among others, the individual scripts comprise tests of the basic communication functionalities, sending individual control commands only, performing long-term movement tests, simulating calibration procedures, investigating the measured positions, or calibrating the absolute encoder system. Also, analysis tools such as position fitting methods, or plotting functionalities are part of the code framework. The main feature of all scripts is the serial communication using the pySerial library[1]. We illustrate a typical code version used to monitor the positions iteratively over multiple deployment cycles in Fig. A.4.

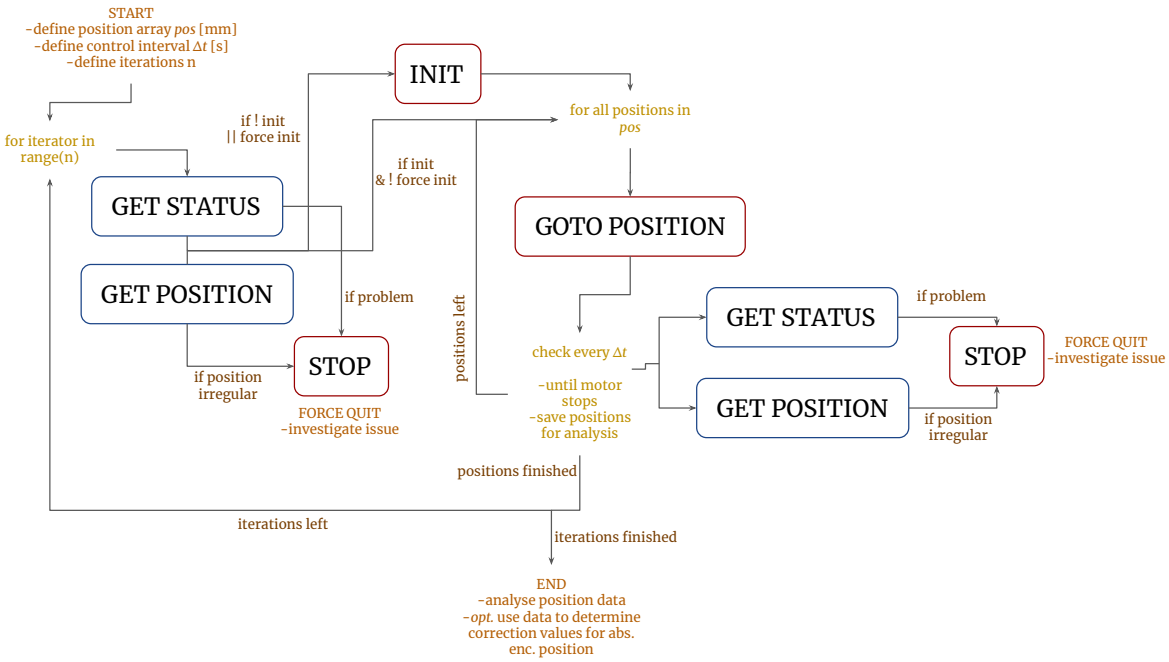

Figure A.4: Flow chart of the basic principle of the SIS control codes used to perform hardware testings, to set up the positioning sensors, and to investigate the positioning performance. The capitalised blocks represent the microcontroller internal commands as listed above.

Note that the Python scripts helped as guidance for the LEGEND collaboration members from UniPD when incorporating the UZH SIS control into the collaboration's slow control systems.

[1] https://pyserial.readthedocs.io/en/latest/pyserial.html

Parts of the code are available to the collaboration and the entire software to members of the local group at UZH.

Appendix B

Additional new physics and background analyses in GERDA

B.1 Alternative nucleon decay search method

The main ideas of searching for ^{75}Ge, the product of a single nucleon disappearance, via an M2 signal, are the reduction of the potential background and a clearly detectable clean signal shape. Hence, we looked for such a ^{75}Ge β-decay in coincidence with a 265 keV γ-line signal from the de-excitation of the ^{75}As daughter. The downside of this coincident tagging is the low efficiency of 0.185% (Sec. 4.3). Thus, a direct M1 signal search with a clearly increased efficiency, though in the presence of a higher background level and without a localised peak, may strengthen our limit for this nucleon decay channel. This approach, however, implies that we need to understand our M1 background extremely well. Otherwise, we would need to interpret every single event in the M1 data below the Q-value as a signal, and the limit would not improve. Given these considerations, we applied the published and validated GERDA background model for Phase II only [110]. In contrast to what is described there, we used the official model after the application of the LAr veto cut and for BEGe and Coax detectors only. Note that our empirical background model (Sec. 4.2) cannot be applied here. Our model was constructed by a fit on the data directly instead of a simulation. Hence, it does not yield an independent prediction for a continuous signal shape, which would have affected the fit result itself. Consequently, the empirical background model can only be applied for a search for a localised signal signature but not for a continuous β spectrum.
As the low energy region is highly dominated by ^{39}Ar (Sec. 4.2), we limited ourselves to the region above the Q-value of ^{39}Ar, 565 keV. In analogy to the M2 search, we did not precisely end at the Q-value of the ^{75}Ge β-only transition, but allowed for a contingency of an additional 2 FWHM, i.e., at 1190 keV. In this interval, we counted an expected event number based on the background simulation model of 76282, whereas in the corresponding real M1 data, we found 76252 energy depositions. Running a binned likelihood fit as indicated in Fig. B.1, we obtained a best-fit for a free signal strength of 673 counts with a significance of 1.45σ, where we allowed for an overall normalisation uncertainty of the simulated background model. The corresponding Frequentist limit is 1443 counts, with a sensitivity of 393 counts. In contrast, when we optimistically fully fixed the background simulation prediction, we obtained a best-fit value of 275 counts, with an upper limit of 584 counts. The sensitivity remains unchanged, as it was evaluated from the simulation model without any statistical fluctuations. The simulated efficiencies to detect M1 events after the LAr veto cut in the region from 565 to 1190 keV are 25.6% and 24.9% for Coax

and BEGe detectors, respectively. Using the same conversion formula as shown in Sec. 4.5, but now with the exposure levels of Phase II only, we obtained lifetimes of 0.7 (2.7) $\times 10^{23}$ yr for the Frequentist case. The limit (sensitivity) worsened (improved) by approximately a factor of 0.8 (1.8), or improved by a factor of 1.9 interpreting the simulation model optimistically as fully accurate. Given these overall very similar results, but considering the absence of a localised signal signature and consequently the correlations between the background and the signal shapes, we decided to interpret the peak search in the M2 data as the main method.

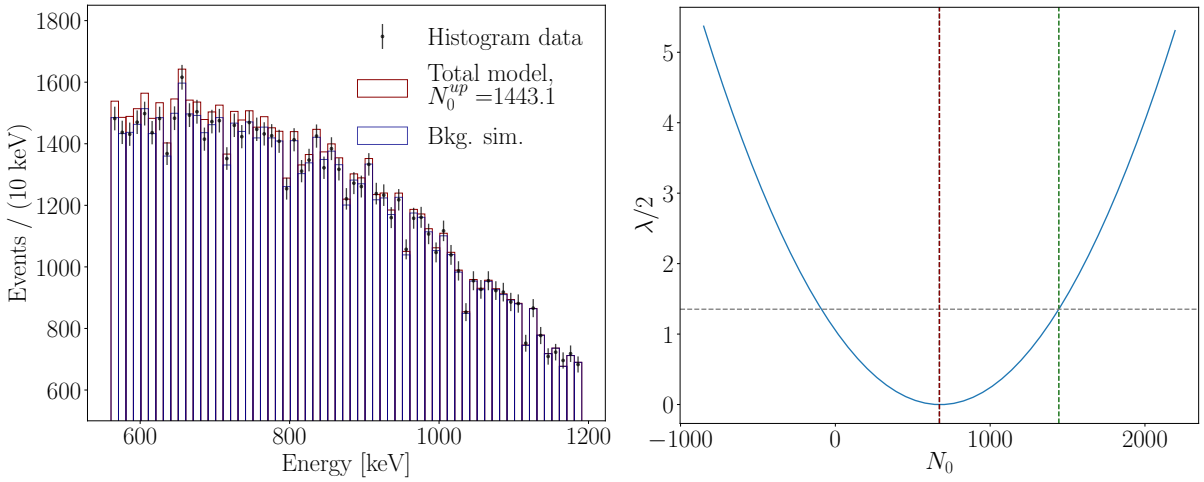

Figure B.1: Left: GERDA Phase II-only data between the endpoints of ^{39}Ar and the β-only transition of ^{75}Ge, after applying the muon, the LAr veto, and the M1 cuts. The simulated background model is taken from Ref. [110]. The full model includes the β decay spectrum of ^{75}Ge as expected to be seen in the data. Right: likelihood profile for the variation of the ^{75}Ge signal count strength N_0. The best-fit value and the corresponding two-sided interval limit at 90% CL are indicated in dark red and dark green, respectively. Note that during the profiling, we included an overall uncertainty on the number of background counts. If a perfect accuracy of the background simulation models is assumed, the count limit improves by more than a factor of 2.

Additionally, we investigated a coincident tagging of an M1 event with a LAr veto scintillation signal as well. Nevertheless, the efficiency of this process is around 1% only, i.e., a factor of 25 lower. The event count in the real data in the interval defined above is approximately a factor of 3 lower, implying an approximate factor of $\sqrt{3} \sim 1.73$ improvement assuming Poisson statistics. Clearly, the background reduction relying on the coincident selection does not overcome the lower efficiency compared to the pure M1 signal search. Consequently, this method may only provide additional input but would not improve our former result based on the M2 signal.

B.2 Gamma-ray line investigation

In order to probe the presence of any potential γ-line in the data of GERDA causing a monoenergetic excess independently, we estimated their significance with the help of an auxiliary measurement. In contrast to the main method based on the global empirical background model and a peak fit, and also contrarily to the systematic crosscheck based on a combined polynomial plus peak fit (Sec. 4.2), we applied a local counting experiment in the signal region of potentially present γ-lines. Conducting an analysis independent of a signal fit provides a means of

crosschecking whether a 3σ excess as estimated via the full fitting frameworks can be reasonably explained by a known line and strenghtens such an interpretation. To determine the statistical significance of a known γ-peak in this second manner, we rely on the Frequentist prescription of Ref. [139], which is based on the likelihood-ratio test statistic as well. Hence, according to the Neyman-Pearson lemma [197], we do not lose background rejection power relevantly. We again fitted a polynomial function within a 25 keV wide window centred at a potential γ-ray energy to constrain the background continuum surrounding this energy value. In correspondence with the polynomial approach, we fitted a quadratic function to the search windows of the low-energy data set and a first order polynomial to the windows of the medium-energy data. Here, we used the same data as those selected in Sec. 4.2, to which we again applied the data quality, the M1, and the LAr veto cuts. The main difference here is that while fitting, we kept the inner five bins centred at the nearest integer of the γ-line energy fully blinded. If a second, potentially present, known γ-line fell into the fit window, we added an additional Gaussian mixture term to properly constrain the continuous background component in the side bands. We then determined the likelihood-ratio test statistic for the expected total count number, i.e., the summed counts of the five inner bins inside the blinded region, with versus without a potential excess in the event number. The signal and the background counts follow a Poisson distribution. We emphasise that in addition, we took into account the uncertainty on the background count number from the fit in the side bands. For this uncertainty, we assumed a Gaussian distribution propagated from the fit uncertainty into the signal region [139]. Given the sufficiently large number of counts in the GERDA data below the MeV range, we applied the test statistic q_0 to evaluate the significance, as was discussed in Sec. 4.4. Relying on the asymptotic $1/2\chi^2(1)$ distribution, this implies that the square root of the test statistic can be used as a sufficiently accurate estimate for the statistical significance [196]. As potential γ-ray transition positions are physically known, any look-elsewhere effect is absent, and the estimated significance can be directly interpreted as global. In Fig. B.2, we show two example γ-line investigations, one for the ^{212}Pb transition at $E_0 = 239$ keV, and one for the ^{208}Tl transition at $E_0 = 583$ keV.

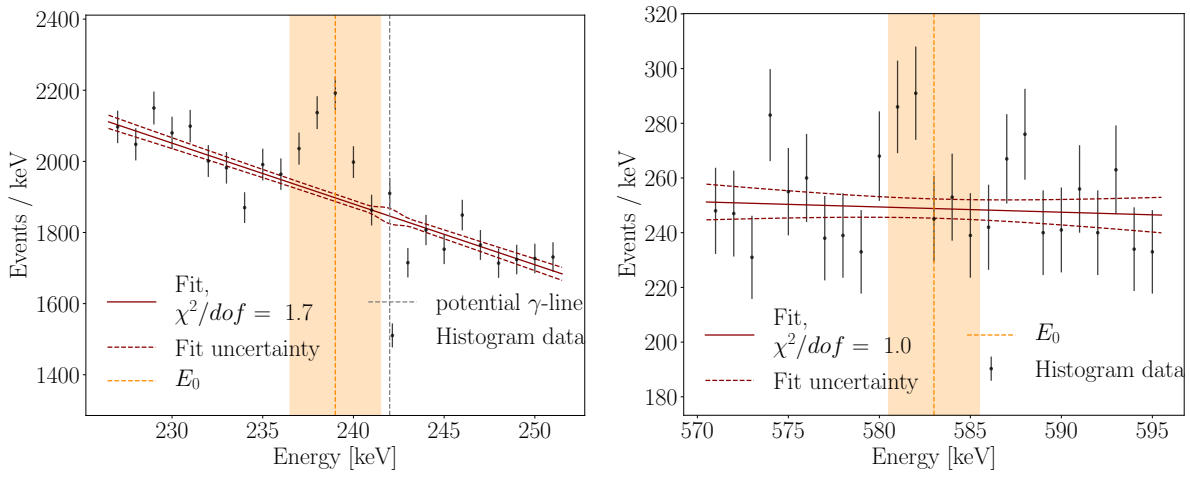

Figure B.2: Examples of a γ-line investigation in a method independently from the main signal fitting procedure. Left: for ^{212}Pb, $E_0 = 239$ keV. Right: for ^{208}Tl, $E_0 = 583$ keV. Here we use a maximum-likelihood-based counting approach with a Gaussian uncertainty on the polynomial fit in the side bands. The marked orange area indicates the signal region, which is blinded for the fit. In the plot on the left we include a Gaussian term to take into account the even only potentially present, known γ-ray at 242 keV from ^{214}Pb.

The left plot reveals a clear presence of ^{212}Pb in the setup, quantitatively confirmed by a significance of 7.4. This confirms the observation made in Sec. 4.4. In addition, in the right plot we see a mild indication of ^{208}Tl, with a significance of 1.9, which suggests that the peak identified at 581 keV (Sec. 4.4) may indeed be caused by this isotope. The obtained significance estimates of all probed transitions are listed in Tab. B.1.

Origin	Energy [keV]	Significance
^{212}Pb	239	7.37
^{214}Pb	242	0.00
^{214}Pb	295	7.17
^{228}Ac	338	3.42
^{214}Pb	352	9.46
^{108m}Ag	434	1.35
^{228}Ac	478	3.94
annihilation	511	0.80
^{85}Kr	514	8.70
^{208}Tl	583	1.94
^{214}Bi	609	1.39
^{108m}Ag	614	0.42
^{137}Cs	662	3.37
^{108m}Ag	723	0.00
^{212}Bi	727	1.45
^{214}Bi	768	0.42
^{228}Ac	795	0.00
^{208}Tl	861	0.00
^{228}Ac	911	1.29
^{214}Bi	934	0.00
^{228}Ac	969	1.86
^{234m}Pa	1001	2.00
^{214}Bi	1120	1.45
^{65}Zn	1125	3.38
^{60}Co	1173	2.41

Table B.1: γ-line investigation for the purpose of crosschecking results. The listed energy corresponds to the nearest integer value in keV. We estimated the significance with the likelihood-ratio test statistic by comparing the count numbers as observed in a blinded window of 5 keV with the expected count number deduced from a polynomial fit in the side bands. We took into account a Gaussian uncertainty on the background estimate propagated from the fit uncertainty when calculating the counting statistic. See Fig. B.2 for illustration and text for a detailed description.

The significance estimates of the excesses observed with the main fitting method described in Sec. 4.4 and tabulated in Tab. 4.2 agree well with the expected isotope transitions and the excess strengths obtained here (Tab. B.1). Among all, only the ^{228}Ac transition at 911 keV appears to be severely stronger in the potentially more sensitive full fitting method than in the counting method, likely caused by an under-fluctuation at the edge of the peak (Fig. 4.14, right, top). Nevertheless, with the alternative method, we find a, though mild, indication. Overall, we

confirm our previous conclusions that the only possible unidentified 3σ excess, which we cannot attribute to an expected isotope transition, is the line at 710 keV. As a side remark, let us note that we also list potential lines above the relevant bDM and electron decay search range. These values do not affect our main new physics search results as presented in Chapter 4, given that their energy exceeds $2m_e$. The analysis indicates a partially mild presence of the γ-rays at 1120, 1125, and 1173 keV from ^{214}Bi, ^{65}Zn, and ^{60}Co, respectively (Tab. B.1). Hence, our results suggest that these lines may impact the alternative nucleon decay search methods based on the M1 data set, as presented in Appendix B.1. Future analysts will thus need to be aware of the potential presence of these isotopes when conducting these searches in LEGEND.

B.3 Muon-induced ^{77m}Ge background in Gerda

Besides the searches for exotic new physics channels, the statistical analysis developed in Chapter 4 has a very useful side application, namely a straight-forward study of potential background contributions via their γ-transitions. The γ-line investigation presented in Appendix B.2 obviously poses a background study, but focused on the new physics searches. Here we apply our framework to a more important case, the study of the cosmogenic production of $^{77(m)}$Ge. It is produced upon captures of secondary neutrons, which are themselves produced by incoming cosmic muons, penetrating the enriched Ge detectors. As already briefly touched in Secs. 2.3, 3.4, and 3.5, decays of this isotope may pose the dominant background contamination for the L-1000 experiment if built at LNGS because of its lower overburden compared to SNOLAB. With a Q-value of 2.7 MeV [49], the β particles emitted from ^{77m}Ge decays are only rarely accompanied by de-excitation γ particles, as the transition to the ground state of ^{77}As has a branching ratio of approximately 80%. They thus potentially survive the background cuts, mimicking $0\nu\beta\beta$ signals [54]. Given this dangerous situation, the GERDA and the LEGEND collaborations run dedicated measurement and simulation campaigns to understand the potential impact very precisely. One option is to search for γ-emission signals following the neutron capture inside the Ge detectors. Another option is to investigate the available GERDA data to search for secondary transitions in the As daughter isotopes, caused by the ground-state ^{77}Ge decay. Of the ^{77m}Ge isotopes, $19 \pm 2\%$ decay to the ^{77}Ge ground state, wheres the remainder decays directly to ^{77}As. If the decay happens to populate the ^{77}Ge ground-state, a γ with an energy of 160 keV is emitted. This allows us to apply our peak search framework to a third option. We can look for these primary 160 keV γ particles from the ^{77m}Ge to ^{77}Ge transition, as done in Ref. [212] for the GERDA data before the hardware upgrade in 2018. To reduce the ^{39}Ar-dominated background at these low-energy events, we applied a delayed coincide cut, based on the Cherenkov veto triggered when an external muon enters the water tank. As the half-life of ^{77m}Ge is $T_{1/2} \approx 47 - 54$ s [49], 90% of the decays are expected to occur within 3 min considering the most conservative half-life of 54 s. This implies that we can define the fraction of potentially visible decays as $f_{3\text{min}} = 0.9$. We selected the data as follows. As in Sec. 4.2, we applied the quality cuts mitigating baseline and test pulser events, the M1, and the LAr cuts. We constrained the energy range to a 25 keV interval centred at 160 keV and only collected events that occurred withing 3 min after a muon veto trigger. Using the Phase II and II+ data, we found 6254 muon veto triggers overall. Applying the full selection, we counted 1225 events that occurred within the 3 min interval after one of these triggers. Note that we achieve almost a factor of 50 in background reduction thanks to the delayed coincide selection criterium. The obtained data set is shown in Fig. B.3, left. There we also provide a histogram of time stamps within the relevant half-life intervals, indicating a rather uniform distribution, which implies that no clear exponentially falling behaviour is seen.

We then applied our fitting method, using the mixture distribution to model the signal, as

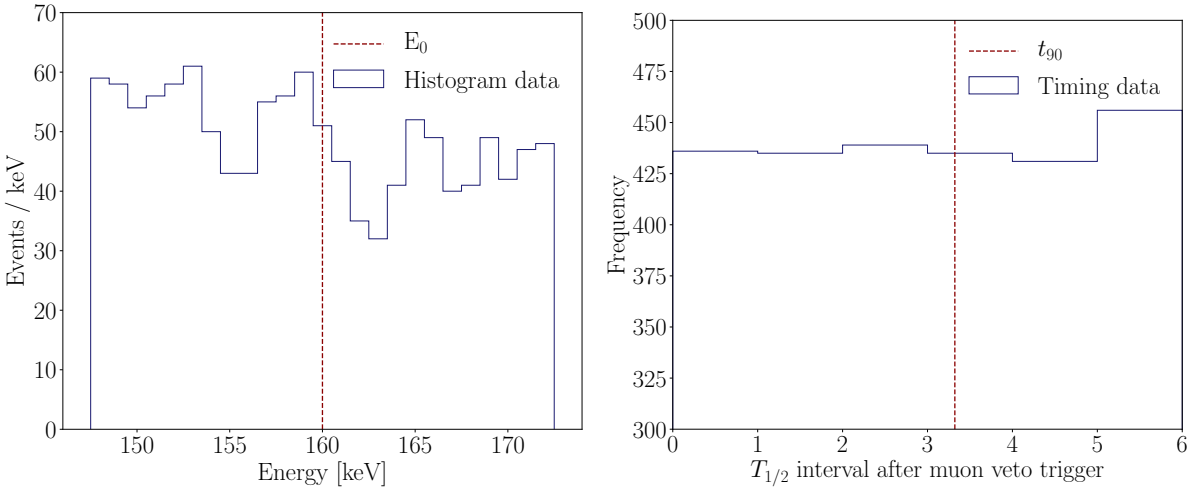

Figure B.3: Left: the selected events in the energy range of interest with a timestamp in a 3 min interval after a Cherenkov veto trigger, after applying the muon, the LAr, and the M1 cuts. Right: the corresponding time interval in units of the relevant $T_{1/2}$ of ^{77m}Ge, indicating a uniform time distribution, i.e., a domination of a continuous background over a potential ^{77m}Ge contribution, which would manifest as an exponential behaviour.

explained in Sec. 4.2. The background was modelled with a second-order polynomial, similar to the method described in Sec. 4.2 in the absence of a specific background model for this data set. The fit yielded a best-fit signal strength of 19.5 counts with a significance of 1.8. We conclude that we found only a very mild, but not a strong indication of the presence of ^{77m}Ge events in the GERDA data. The corresponding two-sided 90% CL upper interval edge, relying again on the profile likelihood method, is $N' = 45.05$ counts, as shown in Fig. B.4.

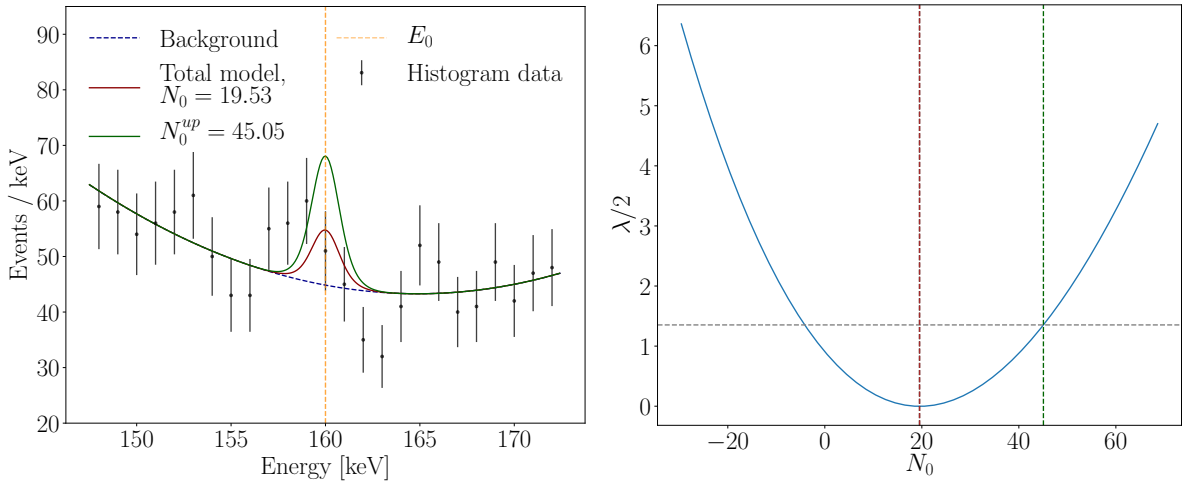

Figure B.4: Left: binned log-likelihood fit (second-order polynomial plus Gaussian mixture model) of the selected data after the delayed muon-veto coincidence cut as described in the text. Right: corresponding likelihood profile, where λ denotes the likelihood-ratio test statistic, when varying the signal strength parameter N_0 for the determination of the limit.

We estimated the sensitivity of our method to be $N_S = 20.8$ counts, using the Asimov data

introduced in Sec. 4.4. We determined the corresponding limit on the background contamination as follows. As mentioned before, we have a reduction of $f_{160} = 0.19$ from the branching ratio. Given an approximately equal cosmogenic production of metastable and ground-state isotopes if taking into account the metastable to ground-state transition [146, 147], we can further assume a reduction of $f_{77\mathrm{m}} \approx 0.5$. The efficiency can be simply taken from our exotic physics analysis. After weighing over the relevant low energy data sets as defined in Sec. 4.2, we obtained an efficiency of $\epsilon_{\mathrm{w}} = 60.0\%$. Also, the corresponding exposure of $\lambda_{\mathrm{low}} = 67.7\,\mathrm{kg\,yr}$ and the weighted ^{76}Ge enrichment fraction of $f_{76} \approx 0.88$ were determined there. Converting the signal count limit into a limit on the physical background rate, we obtain

$$R' \leq \frac{N'}{\epsilon_w \; f_{76} \; \lambda_{\mathrm{low}} \; f_{77\mathrm{m}} \; f_{160} \; f_{3\mathrm{min}}} \approx 14.9\,\mathrm{counts\,/\,(kg\,yr)} \; . \tag{B.1}$$

The analogue conversion of the sensitivity yields $R_S = 6.9\,\mathrm{counts\,/\,(kg\,yr)}$. More sophisticated analyses, e.g., based on the event pulse shapes of events assumed to be caused by β-γ-pairs during the mentioned ^{77}As de-excitation processes, provide approximately an order of magnitude higher sensitivities thanks to the lower background levels [213]. The same holds for pulse shape studies based on the primary neutron capture process. However, any event pulse shape study requires very detailed validation work, and hence time. Let us conclude that our fitting framework, of course with a limited sensitivity compared to dedicated pulse shape studies, can instead be readily applied to estimate relevant background contributions in very short time. In the future, an analyses as exemplarily outlined here may be helpful to provide quick but solid estimates on expected or measured background levels, and/or crosscheck the more dedicated studies.

List of abbreviations

$0\nu\beta\beta$, neutrinoless double β decay
$2\nu\beta\beta$, two-neutrino double β decay
AC, anti-coincidence, corresponds to M (multiplicity)
ADC, analog-to-digital converter
ALPs, axion-like particles
bDM, bosonic dark matter
BEGe, broad-energy germanium detector
BSM, beyond Standard Model
CDF, cumulative distribution function
CL, confidence level
Coax, coaxial-shaped detector
cts, counts
DAQ, data acquisition system
DEP, double escape peak
DM, dark matter
dof, degrees-of-freedom
FEP, full energy peak
fom, figure-of-merit
FWHM, full width at half maximum
GERDA, Germanium Detector Array
gof, goodness-of-fit
HPGe, high-purity germanium detector
IC/ICPC, inverted coaxial (point-contact) detector
INFN, Istituo Nazionale di Fisica Nucleare
KS, Kolmogorov-Smirnov test
L-60/L-200/L-1000, LEGEND-60/200/1000
LANL, Los Alamos National Laboratories
LAr, liquid argon
LED, light-emitting diode
LEGEND, Large Enriched Germanium Experiment for Neutrinoless $\beta\beta$ Decay
LNGS, Laboratori Nazionali del Gran Sasso
M1/M2, multiplicity one/two, corresponds to AC (anti-coincidence)
MaGe, MAJORANA/GERDA Monte Carlo simulation
MC, Monte Carlo simulation
MJD, MAJORANA DEMONSTRATOR
MLE, maximum-likelihood estimator
PE, polyethylen
PGT, Post-GERDA-Test

PMT, photo-multiplier tube
PPC, p-type point-contact detector
ppm, parts per million
PSD, pulse shape discrimination
PTFE, polytetrafluoroethylene
PVC, polyvinyl chloride
ROI, region-of-interest
RS-422, recommended standard 422
SEP, single escape peak
SiPM, silicon photo-multiplier
SIS, source insertion system
SM, Standard Model of particle physics
Surf, Sanford Underground Research Facility
SW, Shapiro-Wilk's test
TPB, tetraphenyl butadiene
UGLAr, underground-sourced liquid argon
USB, universal serial bus
UniPD, University of Padua
UZH, University of Zurich

List of Tables

List of Figures

Bibliography

[1] L. Baudis et al. "Calibration sources for the LEGEND-200 experiment". In: *JINST* 18.02 (2023), P02001. DOI: 10.1088/1748-0221/18/02/P02001. arXiv: 2211.05026 [physics.ins-det].

[2] R. L. Workman. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[3] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[4] Q. R. Ahmad et al. "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015.

[5] Julian Heeck and Volodymyr Takhistov. "Inclusive nucleon decay searches as a frontier of baryon number violation". In: *Phys. Rev. D* 101 (1 Jan. 2020), p. 015005. DOI: 10.1103/PhysRevD.101.015005. URL: https://link.aps.org/doi/10.1103/PhysRevD.101.015005.

[6] Wolfgang Pauli. "Pauli letter collection: letter to Lise Meitner". Typed copy. URL: https://cds.cern.ch/record/83282.

[7] J Chadwick. "Intensitätsverteilung im magnetischen Spectrum der β-Strahlen von radium B + C". In: *Verhandl. Dtsc. Phys. Ges.* 16 (1914), p. 383. URL: http://cds.cern.ch/record/262756.

[8] E. Fermi. *Tentativo di una Teoria Dei Raggi β*. 1934. DOI: 10.1007/BF02959820.

[9] C. L. Cowan et al. "Detection of the free neutrino: A Confirmation". In: *Science* 124 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103.

[10] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36.

[11] K. Kodama et al. "Observation of tau neutrino interactions". In: *Phys. Lett. B* 504 (2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035.

[12] C. S. Wu et al. "Experimental Test of Parity Conservation in β Decay". In: *Phys. Rev.* 105 (1957), pp. 1413–1414. DOI: 10.1103/PhysRev.105.1413.

[13] M. Goldhaber, L. Grodzins, and A. W. Sunyar. "Helicity of Neutrinos". In: *Phys. Rev.* 109 (1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015.

[14] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. "The Standard model of particle physics". In: *Rev. Mod. Phys.* 71 (1999), S96–S111. DOI: 10.1103/RevModPhys.71.S96. arXiv: hep-ph/9812285.

[15] Marco Drewes. "The Phenomenology of Right Handed Neutrinos". In: *Int. J. Mod. Phys. E* 22 (2013), p. 1330019. DOI: 10.1142/S0218301313300191. arXiv: 1303.6912 [hep-ph].

[16] M. Aker et al. "Direct neutrino-mass measurement with sub-electronvolt sensitivity". In: *Nature Phys.* 18.2 (2022), pp. 160–166. DOI: 10.1038/s41567-021-01463-1. arXiv: 2105.08533 [hep-ex].

[17] Ettore Majorana. "Teoria simmetrica dell elettrone e del positrone". In: *Nuovo Cim.* 14 (1937), pp. 171–184. DOI: 10.1007/BF02961314.

[18] Boris Kayser. "Are neutrinos their own antiparticles?" In: *J. Phys. Conf. Ser.* 173 (2009). Ed. by Frank Avignone, Richard Creswick, and Kuniharu Kubodera, p. 012013. DOI: 10.1088/1742-6596/173/1/012013. arXiv: 0903.0899 [hep-ph].

[19] Tsutomu Yanagida. "Horizontal Symmetry and Masses of Neutrinos". In: *Prog. Theor. Phys.* 64 (1980), p. 1103. DOI: 10.1143/PTP.64.1103.

[20] Rabindra N. Mohapatra and Goran Senjanovic. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (1980), p. 912. DOI: 10.1103/PhysRevLett.44.912.

[21] M. Fukugita and T. Yanagida. "Baryogenesis Without Grand Unification". In: *Phys. Lett. B* 174 (1986), pp. 45–47. DOI: 10.1016/0370-2693(86)91126-3.

[22] W. Buchmuller, R. D. Peccei, and T. Yanagida. "Leptogenesis as the origin of matter". In: *Ann. Rev. Nucl. Part. Sci.* 55 (2005), pp. 311–355. DOI: 10.1146/annurev.nucl.55.090704.151558. arXiv: hep-ph/0502169.

[23] Chee Sheng Fong, Enrico Nardi, and Antonio Riotto. "Leptogenesis in the Universe". In: *Adv. High Energy Phys.* 2012 (2012), p. 158303. DOI: 10.1155/2012/158303. arXiv: 1301.3062 [hep-ph].

[24] Chris Waltham. "Teaching neutrino oscillations". In: *American Journal of Physics* 72.6 (June 2004), pp. 742–752. DOI: 10.1119/1.1646132. URL: https://doi.org/10.1119%5C%2F1.1646132.

[25] P. Hernandez. "Neutrino Physics". In: *8th CERN–Latin-American School of High-Energy Physics*. 2016, pp. 85–142. DOI: 10.5170/CERN-2016-005.85. arXiv: 1708.01046 [hep-ph].

[26] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". In: *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247.

[27] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[28] P. F. de Salas et al. "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity". In: *Phys. Lett. B* 782 (2018), pp. 633–640. DOI: 10.1016/j.physletb.2018.06.019. arXiv: 1708.01186 [hep-ph].

[29] Ivan Esteban et al. "The fate of hints: updated global analysis of three-flavor neutrino oscillations". In: *JHEP* 09 (2020), p. 178. DOI: 10.1007/JHEP09(2020)178. arXiv: 2007.14792 [hep-ph].

[30] T. M. C. Abbott et al. "Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing". In: *Phys. Rev. D* 105.2 (2022), p. 023520. DOI: 10.1103/PhysRevD.105.023520. arXiv: 2105.13549 [astro-ph.CO].

[31] Lachlan Lancaster et al. “A tale of two modes: Neutrino free-streaming in the early universe”. In: *JCAP* 07 (2017), p. 033. DOI: 10.1088/1475-7516/2017/07/033. arXiv: 1704.06657 [astro-ph.CO].

[32] E. W. Otten and C. Weinheimer. “Neutrino mass limit from tritium beta decay”. In: *Rept. Prog. Phys.* 71 (2008), p. 086201. DOI: 10.1088/0034-4885/71/8/086201. arXiv: 0909.2104 [hep-ex].

[33] S. Abe et al. “Search for the Majorana Nature of Neutrinos in the Inverted Mass Ordering Region with KamLAND-Zen”. In: *Phys. Rev. Lett.* 130.5 (2023), p. 051801. DOI: 10.1103/PhysRevLett.130.051801. arXiv: 2203.02139 [hep-ex].

[34] M. Goeppert-Mayer. “Double beta-disintegration”. In: *Phys. Rev.* 48 (1935), pp. 512–516. DOI: 10.1103/PhysRev.48.512.

[35] V. I. Tretyak and Yuri G. Zdesenko. “Tables of double beta decay data: An update”. In: *Atom. Data Nucl. Data Tabl.* 80 (2002), pp. 83–116. DOI: 10.1006/adnd.2001.0873.

[36] M. G. Inghram and J. H. Reynolds. “Double beta-decay of Te-130”. In: *Phys. Rev.* 78 (1950), pp. 822–823. DOI: 10.1103/PhysRev.78.822.2.

[37] S. R. Elliott, A. A. Hahn, and M. K. Moe. “Direct Evidence for Two Neutrino Double Beta Decay in ^{82}Se”. In: *Phys. Rev. Lett.* 59 (1987), pp. 2020–2023. DOI: 10.1103/PhysRevLett.59.2020.

[38] Steven R. Elliott and Petr Vogel. “Double beta decay”. In: *Ann. Rev. Nucl. Part. Sci.* 52 (2002), pp. 115–151. DOI: 10.1146/annurev.nucl.52.050102.090641. arXiv: hep-ph/0202264.

[39] Matteo Agostini et al. “Toward the discovery of matter creation with neutrinoless double-beta decay”. In: (Feb. 2022). arXiv: 2202.01787 [hep-ex].

[40] J. Kotila and F. Iachello. “Phase space factors for double-β decay”. In: *Phys. Rev. C* 85 (2012), p. 034316. DOI: 10.1103/PhysRevC.85.034316. arXiv: 1209.5722 [nucl-th].

[41] Dusan Stefanik et al. “Reexamining the light neutrino exchange mechanism of the $0\nu\beta\beta$ decay with left- and right-handed leptonic and hadronic currents”. In: *Phys. Rev. C* 92.5 (2015), p. 055502. DOI: 10.1103/PhysRevC.92.055502. arXiv: 1506.07145 [hep-ph].

[42] Jonathan Engel and Javier Menéndez. “Status and Future of Nuclear Matrix Elements for Neutrinoless Double-Beta Decay: A Review”. In: *Rept. Prog. Phys.* 80.4 (2017), p. 046301. DOI: 10.1088/1361-6633/aa5bc5. arXiv: 1610.06548 [nucl-th].

[43] J. Schechter and J. W. F. Valle. “Neutrinoless Double beta Decay in SU(2) x U(1) Theories”. In: *Phys. Rev. D* 25 (1982), p. 2951. DOI: 10.1103/PhysRevD.25.2951.

[44] H. Primakoff and S. P. Rosen. “Double beta decay”. In: *Rept. Prog. Phys.* 22.1 (1959), pp. 121–166. DOI: 10.1088/0034-4885/22/1/305.

[45] N. Aghanim et al. “Planck 2018 results. VI. Cosmological parameters”. In: *Astron. Astrophys.* 641 (2020). [Erratum: Astron.Astrophys. 652, C4 (2021)], A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[46] F. T. Avignone, G. S. King, and Yu. G. Zdesenko. “Next generation double-beta decay experiments: Metrics for their evaluation”. In: *New J. Phys.* 7 (2005), p. 6. DOI: 10.1088/1367-2630/7/1/006.

[47] M. Agostini et al. “Final Results of GERDA on the Search for Neutrinoless Double-β Decay”. In: *Phys. Rev. Lett.* 125.25 (2020), p. 252502. DOI: 10.1103/PhysRevLett.125.252502. arXiv: 2009.06079 [nucl-ex].

[48] Arkadiusz Sitek and Anna Celler. "Limitations of Poisson statistics in describing radioactive decay". In: *Physica Medica* 31 (Oct. 2015). DOI: 10.1016/j.ejmp.2015.08.015.

[49] S. Y. F. Chu, L. P. Ekstrom, and R. B. Firestone. *WWW Table of Radioactive Isotopes*. http://nucleardata.nuclear.lu.se/nucleardata/toi/, database version 1999-02-28, accessed 23-January-2023.

[50] Werner Rodejohann. "Neutrino-less Double Beta Decay and Particle Physics". In: *Int. J. Mod. Phys. E* 20 (2011), pp. 1833–1930. DOI: 10.1142/S0218301311020186. arXiv: 1106.1334 [hep-ph].

[51] Michelle J. Dolinski, Alan W. P. Poon, and Werner Rodejohann. "Neutrinoless Double-Beta Decay: Status and Prospects". In: *Ann. Rev. Nucl. Part. Sci.* 69 (2019), pp. 219–251. DOI: 10.1146/annurev-nucl-101918-023407. arXiv: 1902.04097 [nucl-ex].

[52] M. Agostini et al. "Limits on uranium and thorium bulk content in GERDA Phase I detectors". In: *Astropart. Phys.* 91 (2017), pp. 15–21. DOI: 10.1016/j.astropartphys.2017.03.003. arXiv: 1611.06884 [physics.ins-det].

[53] I. J. Arnquist et al. "Final Result of the Majorana Demonstrator's Search for Neutrinoless Double-β Decay in Ge76". In: *Phys. Rev. Lett.* 130.6 (2023), p. 062501. DOI: 10.1103/PhysRevLett.130.062501. arXiv: 2207.07638 [nucl-ex].

[54] N. Abgrall et al. "The Large Enriched Germanium Experiment for Neutrinoless $\beta\beta$ Decay: LEGEND-1000 Preconceptual Design Report". In: (July 2021). arXiv: 2107.11462 [physics.ins-det].

[55] E. Aprile et al. "Double-Weak Decays of ^{124}Xe and ^{136}Xe in the XENON1T and XENONnT Experiments". In: *Phys. Rev. C* 106.2 (2022), p. 024328. DOI: 10.1103/PhysRevC.106.024328. arXiv: 2205.04158 [hep-ex].

[56] G. Anton et al. "Search for Neutrinoless Double-β Decay with the Complete EXO-200 Dataset". In: *Phys. Rev. Lett.* 123.16 (2019), p. 161802. DOI: 10.1103/PhysRevLett.123.161802. arXiv: 1906.02723 [hep-ex].

[57] V. Alvarez et al. "NEXT-100 Technical Design Report (TDR): Executive Summary". In: *JINST* 7 (2012), T06001. DOI: 10.1088/1748-0221/7/06/T06001. arXiv: 1202.0721 [physics.ins-det].

[58] J. B. Albert et al. "Sensitivity and Discovery Potential of nEXO to Neutrinoless Double Beta Decay". In: *Phys. Rev. C* 97.6 (2018), p. 065503. DOI: 10.1103/PhysRevC.97.065503. arXiv: 1710.05075 [nucl-ex].

[59] F. Agostini et al. "Sensitivity of the DARWIN observatory to the neutrinoless double beta decay of ^{136}Xe". In: *Eur. Phys. J. C* 80.9 (2020), p. 808. DOI: 10.1140/epjc/s10052-020-8196-z. arXiv: 2003.13407 [physics.ins-det].

[60] C. Adams et al. "Sensitivity of a tonne-scale NEXT detector for neutrinoless double beta decay searches". In: *JHEP* 2021.08 (2021), p. 164. DOI: 10.1007/JHEP08(2021)164. arXiv: 2005.06467 [physics.ins-det].

[61] Shaobo Wang. "The TPC detector of PandaX-III Neutrinoless Double Beta Decay experiment". In: *JINST* 15.03 (2020). Ed. by Marzio Nessi, p. C03052. DOI: 10.1088/1748-0221/15/03/C03052. arXiv: 2001.01356 [physics.ins-det].

[62] D. Q. Adams et al. "Search for Majorana neutrinos exploiting millikelvin cryogenics with CUORE". In: *Nature* 604.7904 (2022), pp. 53–58. DOI: 10.1038/s41586-022-04497-4. arXiv: 2104.06906 [nucl-ex].

[63] O. Azzolini et al. “Final Result on the Neutrinoless Double Beta Decay of ^{82}Se with CUPID-0”. In: *Phys. Rev. Lett.* 129.11 (2022), p. 111801. DOI: 10.1103/PhysRevLett.129.111801. arXiv: 2206.05130 [nucl-ex].

[64] E. Armengaud et al. “The CUPID-Mo experiment for neutrinoless double-beta decay: performance and prospects”. In: *Eur. Phys. J. C* 80.1 (2020), p. 44. DOI: 10.1140/epjc/s10052-019-7578-6. arXiv: 1909.02994 [physics.ins-det].

[65] V. Alenkov et al. “Technical Design Report for the AMoRE $0\nu\beta\beta$ Decay Search Experiment”. In: (Dec. 2015). arXiv: 1512.05957 [physics.ins-det].

[66] R. Nakamura et al. “Research and development toward KamLAND2-Zen”. In: *J. Phys. Conf. Ser.* 1468.1 (2020). Ed. by Masayuki Nakahata, p. 012256. DOI: 10.1088/1742-6596/1468/1/012256.

[67] Yoshiyuki Fukuda et al. “ZICOS - Neutrinoless Double Beta Decay experiment using Zr-96 with an organic liquid scintillator -”. In: *J. Phys. Conf. Ser.* 1468.1 (2020). Ed. by Masayuki Nakahata, p. 012139. DOI: 10.1088/1742-6596/1468/1/012139.

[68] Josephine Paton. “Neutrinoless Double Beta Decay in the SNO+ Experiment”. In: *Prospects in Neutrino Physics*. Mar. 2019. arXiv: 1904.01418 [hep-ex].

[69] S. Ajimura et al. “Low background measurement in CANDLES-III for studying the neutrino-less double beta decay of ^{48}Ca”. In: *Phys. Rev. D* 103.9 (2021), p. 092008. DOI: 10.1103/PhysRevD.103.092008. arXiv: 2008.09288 [hep-ex].

[70] Victor Tretyak. “Investigation of ^{100}Mo two-neutrino double beta decay in NEMO-3”. In: *AIP Conf. Proc.* 2165.1 (2019). Ed. by Osvaldo Civitarese, Ivan Stekl, and Jouni Suhonen, p. 020028. DOI: 10.1063/1.5130989.

[71] Andrea Jeremie. “The SuperNEMO demonstrator double beta experiment”. In: *Nucl. Instrum. Meth. A* 958 (2020). Ed. by Manfred Krammer et al., p. 162115. DOI: 10.1016/j.nima.2019.04.069.

[72] W. Pauli. “Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren”. In: *Zeitschrift fur Physik* 31.1 (Feb. 1925), pp. 765–783. DOI: 10.1007/BF02980631.

[73] S.M. Sze. *Physics of Semiconductor Devices*. Wiley-Interscience publication. Wiley, 1981. ISBN: 9780471056614. URL: https://books.google.ch/books?id=LCNTAAAAMAAJ.

[74] J. Singh. *Physics of Semiconductors and Their Heterostructures*. Electrical engineering series. McGraw-Hill, 1993. ISBN: 9780071128353. URL: https://books.google.ch/books?id=y2LoJQAACAAJ.

[75] K.K. Ng. *Complete Guide to Semiconductor Devices*. McGraw-Hill series in electrical and computer engineering. McGraw-Hill, 1995. ISBN: 9780071135276. URL: https://books.google.ch/books?id=JJhfAAAACAAJ.

[76] Gilmore G.R. *Practical Gamma-ray Spectrometry – 2nd Edition*. John Wiley Sons, Ltd., 2008. ISBN: 978-0-470-86196-7.

[77] W. Shockley. “Currents to conductors induced by a moving point charge”. In: *J. Appl. Phys.* 9.10 (1938), pp. 635–636. DOI: 10.1063/1.1710367.

[78] Simon Ramo. “Currents induced by electron motion”. In: *Proc. Ire.* 27 (1939), pp. 584–585. DOI: 10.1109/JRPROC.1939.228757.

[79] M. Agostini et al. “Characterization of 30^{76}Ge enriched Broad Energy Ge detectors for GERDA Phase II”. In: *Eur. Phys. J. C* 79.11 (2019), p. 978. DOI: 10.1140/epjc/s10052-019-7353-8. arXiv: 1901.06590 [physics.ins-det].

[80] Glen Knoll. *Radiation Detection and Measurement (4th ed.)* Hoboken, NJ: John Wiley, 2010. ISBN: 978-0-470-13148-0.

[81] P. Lenard. "Ueber die lichtelektrische Wirkung". In: *Annalen der Physik* 313.5 (Jan. 1902), pp. 149–198. DOI: 10.1002/andp.19023130510.

[82] Arthur H. Compton. "A Quantum Theory of the Scattering of X-rays by Light Elements". In: *Phys. Rev.* 21 (5 May 1923), pp. 483–502. DOI: 10.1103/PhysRev.21.483. URL: https://link.aps.org/doi/10.1103/PhysRev.21.483.

[83] J.H. Hubbell. "Electron–positron pair production by photons: A historical overview". In: *Radiation Physics and Chemistry* 75.6 (2006). Pair Production, pp. 614–623. ISSN: 0969-806X. DOI: https://doi.org/10.1016/j.radphyschem.2005.10.008. URL: https://www.sciencedirect.com/science/article/pii/S0969806X0500263X.

[84] Giovanni Benato. "Data Reconstruction and Analysis for the GERDA Experiment". PhD thesis. University of Zurich, 2015.

[85] G. Benato et al. "Improvement of the GERDA Ge Detectors Energy Resolution by an Optimized Digital Signal Processing". In: *Phys. Procedia* 61 (2015). Ed. by Frank Avignone and Wick Haxton, pp. 673–682. DOI: 10.1016/j.phpro.2014.12.069.

[86] M. Agostini et al. "Improvement of the energy resolution via an optimized digital signal processing in GERDA Phase I". In: *Eur. Phys. J. C* 75.6 (2015), p. 255. DOI: 10.1140/epjc/s10052-015-3409-6. arXiv: 1502.04392 [physics.ins-det].

[87] M. Agostini et al. "Calibration of the Gerda experiment". In: *Eur. Phys. J. C* 81.8 (2021), p. 682. DOI: 10.1140/epjc/s10052-021-09403-2. ar Xiv: 2103.13777 (physics.ins-det).

[88] W. a Heisenberg. "Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Z. Phys.* 43 (1927), pp. 172–198. DOI: 10.1007/BF01397280.

[89] Rudolf L. Mössbauer. "Kernresonanzfluoreszenz von Gammastrahlung in ^{191}Ir". In: *Z. Phys.* 151.2 (1958), pp. 124–143. DOI: 10.1007/BF01344210.

[90] Hans R. Bilger. "Fano Factor in Germanium at 77°K". In: *Phys. Rev.* 163 (2 Nov. 1967), pp. 238–253. DOI: 10.1103/PhysRev.163.238. URL: https://link.aps.org/doi/10.1103/PhysRev.163.238.

[91] N. Abgrall et al. "The processing of enriched germanium for the Majorana Demonstrator and R&D for a next generation double-beta decay experiment". In: *Nucl. Instrum. Meth. A* 877 (2018), pp. 314–322. DOI: 10.1016/j.nima.2017.09.036. arXiv: 1707.06255 [physics.ins-det].

[92] M. Agostini et al. "Pulse shape discrimination for GERDA Phase I data". In: *Eur. Phys. J. C* 73.10 (2013), p. 2583. DOI: 10.1140/epjc/s10052-013-2583-7. arXiv: 1307.2610 [physics.ins-det].

[93] M. Agostini et al. "Background-free search for neutrinoless double-β decay of ^{76}Ge with GERDA". In: *Nature* 544 (2017), p. 47. DOI: 10.1038/nature21717. arXiv: 1703.00570 [nucl-ex].

[94] M. Agostini et al. "Upgrade for Phase II of the Gerda experiment". In: *Eur. Phys. J. C* 78.5 (2018), p. 388. DOI: 10.1140/epjc/s10052-018-5812-2. arXiv: 1711.01452 [physics.ins-det].

[95] K. T. Knöpfle and B. Schwingenheuer. "Design and performance of the GERDA low-background cryostat for operation in water". In: *JINST* 17.02 (2022), P02038. DOI: 10.1088/1748-0221/17/02/P02038. arXiv: 2202.03847 [physics.ins-det].

[96] G. Bellini et al. "Cosmic-muon flux and annual modulation in Borexino at 3800 m water-equivalent depth". In: *JCAP* 05 (2012), p. 015. DOI: 10.1088/1475-7516/2012/05/015. arXiv: 1202.6403 [hep-ex].

[97] 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* 73.3 (2013), p. 2330. DOI: 10.1140/epjc/s10052-013-2330-0. arXiv: 1212.4067 [physics.ins-det].

[98] M. Agostini et al. "Characterization of 30 ^{76}Ge enriched Broad Energy Ge detectors for GERDA Phase II". In: *Eur. Phys. J. C* 79.978 (2019). DOI: https://doi.org/10.1140/epjc/s10052-019-7353-8.

[99] M. Agostini et al. "Characterization of inverted coaxial ^{76}Ge detectors in GERDA for future double-β decay experiments". In: *Eur. Phys. J. C* 81.6 (2021), p. 505. DOI: 10.1140/epjc/s10052-021-09184-8. arXiv: 2103.15111 [physics.ins-det].

[100] M. Agostini et al. "Pulse shape analysis in Gerda Phase II". In: *Eur. Phys. J. C* 82.4 (2022), p. 284. DOI: 10.1140/epjc/s10052-022-10163-w. arXiv: 2202.13355 [physics.ins-det].

[101] M. Agostini et al. "Results on Neutrinoless Double-β Decay of ^{76}Ge from Phase I of the GERDA Experiment". In: *Phys. Rev. Lett.* 111.12 (2013), p. 122503. DOI: 10.1103/PhysRevLett.111.122503. arXiv: 1307.4720 [nucl-ex].

[102] Ettore Segreto. "Properties of liquid argon scintillation light emission". In: *Phys. Rev. D* 103 (4 Feb. 2021), p. 043001. DOI: 10.1103/PhysRevD.103.043001. URL: https://link.aps.org/doi/10.1103/PhysRevD.103.043001.

[103] Nuno Barros et al. "In-situ measurement of the scintillation light attenuation in liquid argon in the GERDA experiment". In: *Nucl. Instrum. Meth. A* 953 (2020), p. 163059. DOI: 10.1016/j.nima.2019.163059. arXiv: 1906.11589 [physics.ins-det].

[104] Laura Baudis et al. "Enhancement of Light Yield and Stability of Radio-Pure Tetraphenyl-Butadiene Based Coatings for VUV Light Detection in Cryogenic Environments". In: *JINST* 10.09 (2015), P09009. DOI: 10.1088/1748-0221/10/09/P09009. arXiv: 1503.05349 [physics.ins-det].

[105] N. Abgrall et al. "The Large Enriched Germanium Experiment for Neutrinoless Double Beta Decay (LEGEND)". In: *AIP Conf. Proc.* 1894.1 (2017). Ed. by Osvaldo Civitarese, Ivan Stekl, and Jouni Suhonen, p. 020027. DOI: 10.1063/1.5007652. arXiv: 1709.01980 [physics.ins-det].

[106] N. Abgrall et al. "The Majorana Demonstrator radioassay program". In: *Nucl. Instrum. Meth. A* 828 (2016), pp. 22–36. DOI: 10.1016/j.nima.2016.04.070. arXiv: 1601.03779 [physics.ins-det].

[107] N. Abgrall et al. "The Majorana Demonstrator readout electronics system". In: *JINST* 17.05 (2022), T05003. DOI: 10.1088/1748-0221/17/05/T05003. arXiv: 2111.09351 [physics.ins-det].

[108] Frank T. Avignone and Steven R. Elliott. "The Search for Double Beta Decay With Germanium Detectors: Past, Present, and Future". In: *Front. in Phys.* 7 (2019), p. 6. DOI: 10.3389/fphy.2019.00006. arXiv: 1901.02805 [nucl-ex].

[109] S. Schönert. *Legend-1000 Portfolio Review: The Legend-1000 experiment design*. Unpublished material for the US Department of Energy portfolio review on future ton-scale double-β decay experiments. July 2021.

[110] M. Agostini et al. "Modeling of GERDA Phase II data". In: *JHEP* 03 (2020), p. 139. DOI: 10.1007/JHEP03(2020)139. arXiv: 1909.02522 [nucl-ex].

[111] J Hubbell and Stephen Seltzer. *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest*. en. 1995.

[112] W. S. Lyon and J. S. Eldridge. "Radioactive Ge^{77} and Ge^{77m}". In: *Phys. Rev.* 107 (4 Aug. 1957), pp. 1056–1057. DOI: 10.1103/PhysRev.107.1056. URL: https://link.aps.org/doi/10.1103/PhysRev.107.1056.

[113] Christoph Wiesinger, Luciano Pandola, and Stefan Schönert. "Virtual depth by active background suppression: Revisiting the cosmic muon induced background of GERDA Phase II". In: *Eur. Phys. J. C* 78.7 (2018), p. 597. DOI: 10.1140/epjc/s10052-018-6079-3. arXiv: 1802.05040 [hep-ex].

[114] P. Agnes et al. "DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon". In: *Phys. Rev. D* 98.10 (2018), p. 102006. DOI: 10.1103/PhysRevD.98.102006. arXiv: 1802.07198 [astro-ph.CO].

[115] Kevin-Peter Gradwohl et al. "Hydrogen reduction of enriched germanium dioxide and zone-refining for the LEGEND experiment". In: *JINST* 15.12 (2020), P12010. DOI: 10.1088/1748-0221/15/12/P12010. arXiv: 2009.07585 [physics.ins-det].

[116] Dongming Mei and A. Hime. "Muon-induced background study for underground laboratories". In: *Phys. Rev. D* 73 (2006), p. 053004. DOI: 10.1103/PhysRevD.73.053004. arXiv: astro-ph/0512125.

[117] Aldo Ianni. "Considerations on Underground Laboratories". In: *J. Phys. Conf. Ser.* 1342.1 (2020). Ed. by Ken Clark et al., p. 012003. DOI: 10.1088/1742-6596/1342/1/012003.

[118] Y. Efremenko et al. "Use of poly(ethylene naphthalate) as a self-vetoing structural material". In: *JINST* 14.07 (2019), P07006. DOI: 10.1088/1748-0221/14/07/P07006. arXiv: 1901.03579 [physics.ins-det].

[119] D. E. Fields et al. "Understanding the enhancement of scintillation light in xenon-doped liquid argon". In: *Nucl. Instrum. Meth. A* 1046 (2023), p. 167707. DOI: 10.1016/j.nima.2022.167707. arXiv: 2009.10755 [physics.ins-det].

[120] I. J. Arnquist et al. "Interpretable boosted-decision-tree analysis for the Majorana Demonstrator". In: *Phys. Rev. C* 107.1 (2023), p. 014321. DOI: 10.1103/PhysRevC.107.014321. arXiv: 2207.10710 [physics.data-an].

[121] Tommaso Comellato, Matteo Agostini, and Stefan Schönert. "Charge-carrier collective motion in germanium detectors for $\beta\beta$-decay searches". In: *Eur. Phys. J. C* 81.1 (2021), p. 76. DOI: 10.1140/epjc/s10052-021-08889-0. arXiv: 2007.12910 [physics.ins-det].

[122] Chloe Ransom. "Energy Calibration for the Gerda and Legend-200 Experiments". PhD thesis. University of Zurich, 2021.

[123] L. Baudis et al. "Monte Carlo studies and optimization for the calibration system of the GERDA experiment". In: *Nucl. Instrum. Meth. A* 729 (2013), pp. 557–564. DOI: 10.1016/j.nima.2013.08.003. arXiv: 1303.6679 [physics.ins-det].

[124] Michał Tarka. "Studies of Neutron Flux Suppression from a γ-ray source and The GERDA Calibration System". PhD thesis. University of Zurich, 2012.

[125] Francis Froborg. "Calibration of Phase I of the GERDA Double Beta Decay Experiment". PhD thesis. University of Zurich, 2012.

[126] Manuel Walter. "Background Reduction Techniques for the GERDA Experiment". PhD thesis. University of Zurich, 2015.

[127] P. A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[128] I. J. Arnquist et al. "Charge Trapping and Energy Performance of the MAJORANA DEMONSTRATOR". In: (Aug. 2022). arXiv: 2208.03424 [physics.ins-det].

[129] S. S Shapiro and M. B. Wilk. "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52 (3-4 December 1965), pp. 591–611. URL: https://doi.org/10.1093/biomet/52.3-4.591.

[130] Nornadiah Mohd Razali and Bee Yap. "Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests". In: *J. Stat. Model. Analytics* 2 (Jan. 2011).

[131] Laura Baudis et al. "Production and characterization of ^{228}Th calibration sources with low neutron emission for GERDA". In: *JINST* 10.12 (2015), P12005. DOI: 10.1088/1748-0221/10/12/P12005. arXiv: 1508.05731 [physics.ins-det].

[132] Michael Miloradovic. "Characterisation of Inverted Coaxial Detectors and Calibration Source Production for the GERDA Experiment". PhD thesis. University of Zurich, 2020.

[133] Gabriela R. Araujo et al. "The upgraded low-background germanium counting facility Gator for high-sensitivity γ-ray spectrometry". In: *JINST* 17.08 (2022), P08010. DOI: 10.1088/1748-0221/17/08/P08010. arXiv: 2204.12478 [physics.ins-det].

[134] W.R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Springer, 1994. ISBN: 9783540572800. URL: https://books.google.ch/books?id=W7vHQgAACAAJ.

[135] E. Aprile et al. "Response of the XENON100 Dark Matter Detector to Nuclear Recoils". In: *Phys. Rev. D* 88 (2013), p. 012006. DOI: 10.1103/PhysRevD.88.012006. arXiv: 1304.1427 [astro-ph.IM].

[136] M.J. Prata. "Solid angle subtended by a cylindrical detector at a point source in terms of elliptic integrals". In: *Rad. Phys. Chem.* 67 (5 2003), pp. 599–603. ISSN: 0969-806X. DOI: 10.1016/s0969-806x(03)00144-0. URL: http://doi.org/10.1016/s0969-806x%2803%2900144-0.

[137] W. Maneschg et al. "Production and characterization of a custom-made ^{228}Th source with reduced neutron source strength for the Borexino experiment". In: *Nucl. Instrum. Meth. A* 680 (2012), pp. 161–167. DOI: 10.1016/j.nima.2012.04.019. arXiv: 1110.1217 [physics.ins-det].

[138] Jens Erler. "On the Combination Procedure of Correlated Errors". In: *Eur. Phys. J. C* 75.9 (2015), p. 453. DOI: 10.1140/epjc/s10052-015-3688-y. arXiv: 1507.08210 [physics.data-an].

[139] G. Vianello. "The Significance of an Excess in a Counting Experiment: Assessing the Impact of Systematic Uncertainties and the Case with a Gaussian Background". In: *Astrophys. J. Suppl.* 236.1 (2018), p. 17. DOI: 10.3847/1538-4365/aab780. arXiv: 1712.00118 [physics.data-an].

[140] Wolfgang A. Rolke, Angel M. Lopez, and Jan Conrad. "Limits and confidence intervals in the presence of nuisance parameters". In: *Nucl. Instrum. Meth. A* 551 (2005). Ed. by L. Lyons and Muge Karagoz, pp. 493–503. DOI: 10.1016/j.nima.2005.05.068. arXiv: physics/0403059.

[141] W. Tornow et al. "Fast-neutron-induced potential background near the Q value of neutrinoless double-β decay of ^{76}Ge". In: *Phys. Rev. C* 93.1 (2016), p. 014614. DOI: 10.1103/PhysRevC.93.014614.

[142] Melissa Boswell et al. “MaGe-a Geant4-based Monte Carlo Application Framework for Low-background Germanium Experiments”. In: *IEEE Trans. Nucl. Sci.* 58 (2011), pp. 1212–1220. DOI: 10.1109/TNS.2011.2144619. arXiv: 1011.3827 [nucl-ex].

[143] S. Agostinelli et al. “GEANT4–a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[144] John Allison et al. “Geant4 developments and applications”. In: *IEEE Trans. Nucl. Sci.* 53 (2006), p. 270. DOI: 10.1109/TNS.2006.869826.

[145] J. Allison et al. “Recent developments in Geant4”. In: *Nucl. Instrum. Meth. A* 835 (2016), pp. 186–225. DOI: 10.1016/j.nima.2016.06.125.

[146] G. Meierhofer et al. “Neutron Activation of 74Ge and 76Ge”. In: *EPJ Web of Conferences* 2 (Mar. 2010). DOI: 10.1051/epjconf/20100205002.

[147] J. Marganiec et al. “Neutron capture cross sections of Ge-74, Ge-76, and As-75 at 25 keV”. In: *Phys. Rev. C* 79 (2009), p. 065802. DOI: 10.1103/PhysRevC.79.065802.

[148] Gianfranco Bertone, Dan Hooper, and Joseph Silk. “Particle dark matter: Evidence, candidates and constraints”. In: *Phys. Rept.* 405 (2005), pp. 279–390. DOI: 10.1016/j.physrep.2004.08.031. arXiv: hep-ph/0404175.

[149] Vasiliki A. Mitsou. “Dark matter: experimental and observational status”. In: *The Fifteenth Marcel Grossmann Meeting* (June 2022). DOI: 10.1142/9789811258251_0180. URL: https://doi.org/10.1142%5C%2F9789811258251_0180.

[150] Dan Hooper and Edward A. Baltz. “Strategies for Determining the Nature of Dark Matter”. In: *Ann. Rev. Nucl. Part. Sci.* 58 (2008), pp. 293–314. DOI: 10.1146/annurev.nucl.58.110707.171217. arXiv: 0802.0702 [hep-ph].

[151] Lars Bergstrom. “Dark Matter Evidence, Particle Physics Candidates and Detection Methods”. In: *Annalen Phys.* 524 (2012), pp. 479–496. DOI: 10.1002/andp.201200116. arXiv: 1205.4882 [astro-ph.HE].

[152] Maxim Pospelov, Adam Ritz, and Mikhail B. Voloshin. “Bosonic super-WIMPs as keV-scale dark matter”. In: *Phys. Rev. D* 78 (2008), p. 115012. DOI: 10.1103/PhysRevD.78.115012. arXiv: 0807.3279 [hep-ph].

[153] Yonit Hochberg et al. “Impact of Dark Compton Scattering on Direct Dark Matter Absorption Searches”. In: *Phys. Rev. Lett.* 128 (19 May 2022), p. 191801. DOI: 10.1103/PhysRevLett.128.191801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.128.191801.

[154] Young Ju Ko and HyangKyu Park. “Remarks on bosonic super-WIMP search experiments”. In: *Phys. Rev. D* 104 (8 Oct. 2021), p. 083030. DOI: 10.1103/PhysRevD.104.083030. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.083030.

[155] R. D. Peccei and Helen R. Quinn. “CP Conservation in the Presence of Instantons”. In: *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440.

[156] Paola Arias et al. “WISPy Cold Dark Matter”. In: *JCAP* 06 (2012), p. 013. DOI: 10.1088/1475-7516/2012/06/013. arXiv: 1201.5902 [hep-ph].

[157] Bob Holdom. “Two U(1)’s and Epsilon Charge Shifts”. In: *Phys. Lett. B* 166 (1986), pp. 196–198. DOI: 10.1016/0370-2693(86)91377-8.

[158] Ricardo Z. Ferreira, M. C. David Marsh, and Eike Müller. “Do Direct Detection Experiments Constrain Axionlike Particles Coupled to Electrons?” In: *Phys. Rev. Lett.* 128 (22 June 2022), p. 221302. DOI: 10.1103/PhysRevLett.128.221302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.128.221302.

[159] J. Silk. *The Big Bang: Third Edition*. Henry Holt and Company, 2000. ISBN: 9780805072563. URL: https://books.google.ch/books?id=XLwe1lUmz5kC.

[160] Lerh Feng Low. "Constraints on the composite photon theory". In: *Mod. Phys. Lett. A* 31.36 (2016), p. 1675002. DOI: 10.1142/S021773231675002X.

[161] M. Agostini et al. "First Search for Bosonic Superweakly Interacting Massive Particles with Masses up to 1 MeV/c^2 with GERDA". In: *Phys. Rev. Lett.* 125.1 (2020). [Erratum: Phys.Rev.Lett. 129, 089901 (2022)], p. 011801. DOI: 10.1103/PhysRevLett.125.011801. arXiv: 2005.14184 [hep-ex].

[162] Ciaran O'Hare. *cajohare/AxionLimits: AxionLimits*. Version v1.0. July 2020. DOI: 10.5281/zenodo.3932430. URL: https://cajohare.github.io/AxionLimits/.

[163] E. Armengaud et al. "Searches for electron interactions induced by new physics in the EDELWEISS-III germanium bolometers". In: *Phys. Rev. D* 98 (8 Oct. 2018), p. 082004. DOI: 10.1103/PhysRevD.98.082004. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.082004.

[164] I. J. Arnquist et al. *Exotic dark matter search with the Majorana Demonstrator*. 2022. DOI: 10.48550/ARXIV.2206.10638. URL: https://arxiv.org/abs/2206.10638.

[165] E. Aprile et al. "Search for New Physics in Electronic Recoil Data from XENONnT". In: *Phys. Rev. Lett.* 129.16 (2022), p. 161805. DOI: 10.1103/PhysRevLett.129.161805. arXiv: 2207.11330 [hep-ex].

[166] Bartosz Fornal and Benjamın Grinstein. "Dark Matter Interpretation of the Neutron Decay Anomaly". In: *Phys. Rev. Lett.* 120 (19 May 2018), p. 191801. DOI: 10.1103/PhysRevLett.120.191801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.191801.

[167] H Ejiri and J D Vergados. "Neutron disappearance inside the nucleus". In: *Journal of Physics G: Nuclear and Particle Physics* 46.2 (Jan. 2019), p. 025104. DOI: 10.1088/1361-6471/aaf55b. URL: https://doi.org/10.1088/1361-6471/aaf55b.

[168] Stephan Paul. "The Puzzle of Neutron Lifetime". In: *Nucl. Instrum. Meth. A* 611 (2009). Ed. by Torsten Soldner et al., pp. 157–166. DOI: 10.1016/j.nima.2009.07.095. arXiv: 0902.0169 [hep-ex].

[169] A. P. Serebrov et al. "Neutron lifetime measurements with a large gravitational trap for ultracold neutrons". In: *Phys. Rev. C* 97 (5 May 2018), p. 055503. DOI: 10.1103/PhysRevC.97.055503. URL: https://link.aps.org/doi/10.1103/PhysRevC.97.055503.

[170] F. M. Gonzalez et al. "Improved neutron lifetime measurement with UCNτ". In: *Phys. Rev. Lett.* 127.16 (2021), p. 162501. DOI: 10.1103/PhysRevLett.127.162501. arXiv: 2106.10375 [nucl-ex].

[171] Bartosz Fornal and Benjamın Grinstein. "Neutron's dark secret". In: *Mod. Phys. Lett. A* 35.31 (2020), p. 2030019. DOI: 10.1142/S0217732320300190. arXiv: 2007.13931 [hep-ph].

[172] Bartosz Fornal and Benjamin Grinstein. "Dark Matter Interpretation of the Neutron Decay Anomaly". In: *Phys. Rev. Lett.* 120.19 (2018). [Erratum: Phys.Rev.Lett. 124, 219901 (2020)], p. 191801. DOI: 10.1103/PhysRevLett.120.191801. arXiv: 1801.01124 [hep-ph].

[173] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[174] S.I. Sukhoruchkin and Z.N. Soroko. *Atomic Mass and Nuclear Binding Energy for Ge-76 (Germanium): Datasheet from Landolt-Börnstein - Group I Elementary Particles, Nuclei and Atoms · Volume 22A: "Nuclei with Z = 1 - 54" in SpringerMaterials, https://doi.org/10.1007/978-3-540-69945-3_1462*. Ed. by H. Schopper. accessed 2022-07-19. DOI: 10.1007/978-3-540-69945-3{_}1462.

[175] A. W. Schardt and Joan P. Welker. "Energy Levels of As^{75}". In: *Phys. Rev.* 99 (3 Aug. 1955), pp. 810–824. DOI: 10.1103/PhysRev.99.810. URL: https://link.aps.org/doi/10.1103/PhysRev.99.810.

[176] Aline Schneuwly. "A Search for Neutron Disappearance in GERDA Experiment". Bachelor thesis. University of Zurich, 2021.

[177] Wolfram Research. *IsotopeData*. https://reference.wolfram.com/language/ref/IsotopeData.html, accessed: 19-July-2022. 2014.

[178] T. Araki et al. "Search for the Invisible Decay of Neutrons with KamLAND". In: *Phys. Rev. Lett.* 96 (10 Mar. 2006), p. 101802. DOI: 10.1103/PhysRevLett.96.101802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.101802.

[179] S. N. Ahmed et al. "Constraints on Nucleon Decay via Invisible Modes from the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 92 (10 Mar. 2004), p. 102004. DOI: 10.1103/PhysRevLett.92.102004. URL: https://link.aps.org/doi/10.1103/PhysRevLett.92.102004.

[180] M. Anderson et al. "Search for invisible modes of nucleon decay in water with the SNO+ detector". In: *Phys. Rev. D* 99 (3 Feb. 2019), p. 032008. DOI: 10.1103/PhysRevD.99.032008. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.032008.

[181] R. Matsumoto et al. "Search for proton decay via $p \to \mu^+ K^0$ in 0.37 megaton-years exposure of Super-Kamiokande". In: *Phys. Rev. D* 106.7 (2022), p. 072003. DOI: 10.1103/PhysRevD.106.072003. arXiv: 2208.13188 [hep-ex].

[182] H. O. Back et al. "New limits on nucleon decays into invisible channels with the BOREXINO counting test facility". In: *Phys. Lett. B* 563 (2003), pp. 23–34. DOI: 10.1016/S0370-2693(03)00636-1. arXiv: hep-ex/0302002.

[183] S. I. Alvis et al. "Search for trinucleon decay in the Majorana Demonstrator". In: *Phys. Rev. D* 99 (7 Apr. 2019), p. 072004. DOI: 10.1103/PhysRevD.99.072004. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.072004.

[184] M. Agostini et al. "Search for tri-nucleon decays of ^{76}Ge in GERDA". In: (July 2023). arXiv: 2307.16542 [nucl-ex].

[185] A. Vishneva et al. "Test of the electron stability with the Borexino detector". In: *J. Phys. Conf. Ser.* 888.1 (2017), p. 012193. DOI: 10.1088/1742-6596/888/1/012193.

[186] H. V. Klapdor-Kleingrothaus, I. V. Krivosheina, and I. V. Titkova. "A new experimental limit for the stability of the electron". In: *Phys. Lett. B* 644 (2007), pp. 109–118. DOI: 10.1016/j.physletb.2006.11.013.

[187] N. Abgrall et al. "New limits on Bosonic Dark Matter, Solar Axions, Pauli Exclusion Principle Violation, and Electron Decay from the Majorana Demonstrator". In: *Phys. Rev. Lett.* 118.16 (2017), p. 161801. DOI: 10.1103/PhysRevLett.118.161801. arXiv: 1612.00886 [nucl-ex].

[188] T.A. Carlson. *Photoelectron and Auger Spectroscopy*. Plenum Press New York, 1975.

[189] R. Clausius. "XVI. On a mechanical theorem applicable to heat". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 40.265 (1870), pp. 122–127. DOI: 10.1080/14786447008640370. eprint: https://doi.org/10.1080/14786447008640370. URL: https://doi.org/10.1080/14786447008640370.

[190] Rizalina Mingazheva. "Calibration, Background Study, and Search for New Physics with the GERDA Experiment". PhD thesis. University of Zurich, 2019.

[191] M. Agostini et al. "Search for exotic physics in double-β decays with GERDA Phase II". In: *JCAP* 12 (2022), p. 012. DOI: 10.1088/1475-7516/2022/12/012. arXiv: 2209.01671 [nucl-ex].

[192] P.O. Awodutire et al. "The modified beta transmuted family of distributions with applications using the exponential distribution". In: *PLoS ONE* 16.11 (2021). DOI: 10.1371/journal.pone.0258512.

[193] Teimouri Nadarajah S. and S.H. M. Shih. "Modified Beta Distributions". In: *Sankhya* B.76 (2014), pp. 19–48. DOI: 10.1007/s13571-013-0077-0.

[194] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[195] A. N. Kolmogorov. "Sulla Determinazione Empirica di Una Legge di Distribuzione". In: *Giornale dell'Istituto Italiano degli Attuari* 4 (1933), pp. 83–91.

[196] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[197] Jerzy Neyman and Egon Sharpe Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Phil. Trans. Roy. Soc. Lond. A* 231.694-706 (1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009.

[198] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[199] A. Wald. "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large". In: *Transactions of the American Mathematical Society* 54 (1943), pp. 426–482. DOI: 10.1090/S0002-9947-1943-0012401-3.

[200] Gary J. Feldman and Robert D. Cousins. "A Unified approach to the classical statistical analysis of small signals". In: *Phys. Rev. D* 57 (1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873. arXiv: physics/9711021.

[201] Hans Dembinski and Piti Ongmongkolkul et al. "scikit-hep/iminuit". In: (Dec. 2020). DOI: 10.5281/zenodo.3949207. URL: https://doi.org/10.5281/zenodo.3949207.

[202] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[203] C.E. Bonferroni. *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze. Seeber, 1936. URL: https://books.google.ch/books?id=3CY-HQAACAAJ.

[204] Adrian E. Bayer, Uros Seljak, and Jakob Robnik. "Self-calibrating the look-elsewhere effect: fast evaluation of the statistical significance using peak heights". In: *Mon. Not. Roy. Astron. Soc.* 508.1 (2021), pp. 1346–1357. DOI: 10.1093/mnras/stab2331. arXiv: 2108.06333 [astro-ph.IM].

[205] E. Aprile et al. “Excess electronic recoil events in XENON1T”. In: *Phys. Rev. D* 102.7 (2020), p. 072004. DOI: 10.1103/PhysRevD.102.072004. arXiv: 2006.09721 [hep-ex].

[206] XCOM, NIST [accessed: September 2020]. https://www.nist.gov/pml/xcom-photon-cross-sections-database.

[207] Rick S. Gupta, Joerg Jaeckel, and Michael Spannowsky. “Probing Poincaré Violation”. In: (Nov. 2022). arXiv: 2211.04490 [hep-ph].

[208] John C. Evans Jr. and Richard I. Steinberg. “Nucleon Stability: A Geochemical Test Independent of Decay Mode”. In: *Science* 197 (1977), pp. 989–991. DOI: 10.1126/science.197.4307.989.

[209] R. Hazama et al. “Limits on single- and multinucleon decays in I-127 by inclusive measurements of nuclear gamma and x rays”. In: *Phys. Rev. C* 49 (1994), pp. 2407–2412. DOI: 10.1103/PhysRevC.49.2407.

[210] R. Bernabei et al. “Search for the nucleon and di-nucleon decay into invisible channels”. In: *Phys. Lett. B* 493 (2000), pp. 12–18. DOI: 10.1016/S0370-2693(00)01112-6.

[211] R. Bernabei et al. “Search for rare processes with DAMA/LXe experiment at Gran Sasso”. In: *Eur. Phys. J. A* 27.S1 (2006). Ed. by Zsolt Fülöp, György Gyürky, and Endre Somorjai, pp. 35–41. DOI: 10.1140/epja/i2006-08-004-y.

[212] Laura Vanhoefer. “Limitations of Rare Event HPGe Experiments due to Muon-Induced Neutron Background”. PhD thesis. Technische Universität München, 2018.

[213] Christoph Wiesinger. “No neutrinos not found”. PhD thesis. Technische Universität München, 2020.

Acknowledgements

The day I presented my PhD thesis I felt truly honoured achieving this great milestone. I am extremely grateful for having received very kind compliments not only on the successful completion of the projects, but also on the work ethic and collegiality. What deserves even more emphasis though, is the simple fact that a PhD in physics cannot be accomplished by a person alone. In my case, actually more than a hundred people contributed, no matter how heavily, to the successful completion of these projects. I thus consider this last part among the most important ones of this entire thesis; the part where I can say thanks to all those contributing to the work, supporting me, or just sharing wonderful moments together during a mind-polishing chapter of life.

I would hence like to thank Laura, not only for giving me the opportunity to work on these intriguing projects aiming to reveal the very reason of the existence of the matter in our Universe, but also for guiding and supporting me through four long years, and for granting me the freedom of choosing which topics to pursue deeper. I would further like to thank the committee members and reviewers, Gino, Riccardo, Andreas, Nico, Michelle, Junting, Pin-Jung, for the excellent input during the PhD, on the thesis itself, and on the presentation and the defence. Especially the last three, as both direct supervisors and work colleagues, I would like to highlight and thank them also for the almost daily input, for all the interest, the detailed suggestions, the guidance, and also the patience with me and my work. Junting's physics-first approach served as an extremely valuable compass when trying to find the optimal approach to solve a problem, or when balancing between the urgency of the fulfilment of different tasks. Pin-Jung's scientific motivation when performing any task and experiment, as well as the care for details, have been very inspiring, and so have been the chats during relaxed walks around Irchel campus with Michelle. I would also like to further thank the former and current PhD students and Postdocs of the group at UZH; Adam, Fréd, Ale, Chiara, Rizalina, Christian, Diego, Jose, Marta, Chloe, Yanina, Kevin, for extremely detailed input on the work, the thesis, the presentation, or just generally for contributing to a motivating and challenging, but still relaxed and enjoyable work environment. Our discussions, sports sessions, or lunch breaks will always stay in my mind.

I express my gratitude to the fellow PhD students still balancing the ship through their own journey, who have made my journey such a pleasant time. My thoughts are with you during this process, Alex, Paloma, Mariana, you are all making impressive progress, and you will certainly succeed in your journeys. The fantastic work environment at UZH would have not been the same without extremely friendly, and also very knowledgeable people from the physics institute, the group technicians Jonathan and Francesco, the mechanics and electronics workshops, let me mention Dani, Reto, Chris, Bruno, Marcel explicitly, the computing staff, namely Roland and Paul, and the secretaries, in particular Regina, Monika, Anna, and Gaby. Thank you all for supporting me, and I hope I did not cause you too many overhours with all my requests...

Going beyond the institute, it is needless to say that without the support from the collaborations, the projects would have never been accomplished. Hence, I would like to thank the entire GERDA

collaboration for all the detailed input and the suggestions on the data analyses and the paper under review. And I would like to thank the entire LEGEND collaboration as well, for all the input on the hardware, the simulations, the measurements, the different analyses, and the related publications. Let me add here also the members of the laboratories at Gran Sasso and Berkeley, for the chance to experience onsite field work, and to gain new insights into different both life and work environments. Given the long list of people in those categories, I will just mention some names here, not in any particular order, representing even more people; namely Bernhard, Sofia, Gio, Ralph, Riccardo, Matthias, Konstantin, Matteo, Ivano, Valentina, Katharina, Elisabetta, Luigi, Anatoly, Tasso, Mike, Brady, Patrick, Laszlo, Stefan, Thomas, Simon, George, Natalia, Mariia, Valerio, Tommaso, Luciano, Julietta, Rushabh, Nabin, Paudel, Wenxin, Steve, Sam, Evelyn, Chiara, Matthew, Ian, Matthew, Jason, CJ, Alan, Marcos, Alexey, Lisa, Ann-Kathrin. You, dear reader, may forgive me if your name is not on this list explicitly, again, many more people contributed in one way or the other, directly or indirectly, to a wonderful PhD journey, and I also thank you, either for your contribution, or simply for the interest in this thesis.
You may have further noticed that I mentioned the Postdocs and current PhD students of the UZH group, but above I skipped some fellow colleagues who shared most of the PhD journey with me. I did so on purpose, having in mind their extraordinary value not only for work-related matters, but also for another decisive part of life, a fulfilling social environment. I can consider myself a lucky person, having worked with and gotten to know colleagues exceeding this description by far. A way better term for Gabriela, my direct, fellow PhD colleague, for Giovanni, and Ricardo would simply be friends, a simple, but deep word. Not only excellent junior physicists, but friends whom you enjoy every single moment with, in and outside the lab. Thank you for bringing enormous joy and wonderful memories to this journey.
Lastly, I would like to thank my families and friends back home and across our beautiful planet, for all the appreciation, for all the support, from near or far, from Rigiblick to South America, over many kilometres and across borders. To my family back home in the Palatinate, thank you for always being there, especially during the moments when it counts and is needed the most. During my PhD I have also found a second family, in Sicily. Thank you for welcoming and accepting me with such wide and open arms. And lastly, thanks to my tiny, little, own family in Zurich, built and strengthened during a global pandemic. Tiny, but with the greatest support there is. In certain cases words are not the right means to express gratitude adequately.