AN IMPROVED BACKGROUND MODEL AND TWO-NEUTRINO DOUBLE-BETA DECAY MEASUREMENT FOR THE MAJORANA DEMONSTRATOR

Anna Leigh Reine

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Physics and Astronomy in the College of Arts and Sciences.

Chapel Hill
2023

Approved by:

John F. Wilkerson

Reyco Henning

Matthew P. Green

David C. Radford

Jonathan Engel

Joaquín Drut

ABSTRACT

Anna Leigh Reine: An Improved Background Model and Two-Neutrino Double-Beta Decay Measurement for the MAJORANA DEMONSTRATOR
(Under the direction of John F. Wilkerson)

Neutrinoless double-beta decay ($0\nu\beta\beta$) is a hypothesized nuclear process that would provide direct evidence of physics beyond the Standard Model in the form of lepton number violation. Its detection would prove that neutrinos are their own antiparticles and play a role in addressing questions of how neutrinos acquire their mass and why the universe exhibits a matter-antimatter asymmetry. Since $0\nu\beta\beta$ would have an extremely long half-life, one of the key challenges for any $0\nu\beta\beta$ experiment is mitigating radioactive backgrounds that could obscure a signal.

The MAJORANA DEMONSTRATOR was a germanium-based $0\nu\beta\beta$ decay search that operated at the Sanford Underground Research Facility and set a 90% CL lower limit of 8.3×10^{25} yr on the $0\nu\beta\beta$ half-life in ^{76}Ge. Through the use of a multi-layer passive shield, radiopure materials, and analysis-based background rejection techniques, the DEMONSTRATOR achieved a background index of $6.23^{+0.55}_{-0.52} \times 10^{-3}$ c/(keV kg yr) in its low-background configuration. This is one of the lowest achieved backgrounds in a $0\nu\beta\beta$ experiment, but it is in excess of the assay-based projection by over a factor of five.

Frequentist fits to the DEMONSTRATOR's data allow its energy spectrum to be decomposed into contributions from different decay chains and experimental components. This work details the development and results of the DEMONSTRATOR's data-driven background model, including its conclusions about the source of the background excess. The resulting model also enabled a precision measurement of the half-life of $0\nu\beta\beta$'s Standard Model counterpart, two-neutrino double-beta decay, one of the rarest processes ever detected. A number of statistical and systematic contributions to the two-neutrino double-beta decay half-life uncertainty are discussed and quantified. The findings described in this work have informed the design of the next-generation tonne-scale experiment, LEGEND-1000, and its precursor, LEGEND-200. LEGEND-1000 aims to achieve sensitivities to $0\nu\beta\beta$ half-lives beyond 10^{28} years, covering the entire inverted ordering parameter space.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, the Particle Astrophysics and Nuclear Physics Programs of the National Science Foundation, and the Sanford Underground Research Facility.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

$0\nu\beta\beta$	neutrinoless double beta decay
$2\nu\beta\beta$	two neutrino double beta decay
BEGe	Broad Energy Germanium
BI	Background Index
BSM	Beyond the Standard Model
DCR	Delayed Charge Recovery
FCCD	Full Charge Collection Depth
FWHM	Full Width Half Maximum
GAT	Germanium Analysis Toolkit
HPGe	High Purity Germanium
ICPC	Inverted Coaxial Point Contact
LAr	Liquid Argon
LEGEND	Large Enriched Germanium Experiment for Neutrinoless Double-Beta Decay
LMFE	Low Mass Front End
LQ	Late Charge
M1	Module 1
M2	Module 2
NLL	Negative log likelihood
PDF	Probability Distribution Function
PPC	P-type Point Contact
PSD	Pulse Shape Discrimination
SM	Standard Model
UGEFCu	Underground Electroformed Copper

LIST OF SYMBOLS

$m_{\beta\beta}$ Effective Majorana mass

$Q_{\beta\beta}$ Double beta decay Q-value

CHAPTER 1: Introduction to Neutrinoless Double-Beta Decay

The Standard Model (SM) of particle physics is a remarkably successful model for describing the fundamental particles that make up the universe and the mechanisms by which they interact. The model successfully unifies our understanding of electromagnetism, the strong force that governs the bonds between atomic nucleons, and the weak force that mediates radioactive decays. Its predictive power was most recently demonstrated by the 2012 discovery of the Higgs boson at the Large Hadron Collider, which confirmed the Standard Model theory for how mass is generated in at least some particles [12, 13].

However, the Standard Model is known to be incomplete. For example, it does not include dark matter or dark energy, which together account for approximately 95% of the energy in the universe. One of the most promising places to search for evidence of physics beyond the Standard Model is in the properties of neutral, weakly-interacting particles called neutrinos. Although neutrinos are included in the Standard Model, the discovery that they have a non-zero mass challenged Standard Model expectations and raised a number of questions that have yet to be resolved. These questions, including whether neutrinos are their own antiparticle and how they acquire their mass, can best be addressed experimentally through searching for a hypothesized rare process known as neutrinoless double-beta decay. Neutrinoless double-beta decay is predicted by many theories beyond the Standard Model, and its discovery would violate an accidental symmetry of the Standard Model. Its discovery could also play a part in addressing the fundamental question of how the universe came to be dominated by matter.

To understand the theory behind neutrinoless double beta decay and the motivation for searching for it, it is useful to begin with a brief introduction to neutrinos and the experimental evidence for their mass.

Section 1.1: Neutrino Properties in the Standard Model

The neutrino was first posited in 1930 by Wolfgang Pauli to resolve a troubling discrepancy in nuclear physics: the fact that beta decay appeared to violate the principle of conservation of energy [14]. Beta decay is the process by which a neutron is converted into a proton, accompanied by the emission of an electron, also known as a β^- particle. Beta decays occur spontaneously in free neutrons and within atomic nuclei, governed by characteristic half-lives. The difference in mass between the parent isotope and the more tightly

bound daughter determines the decay Q-value, the amount of energy released during the decay. Conservation of energy, momentum, and angular momentum suggested that beta decays of a particular isotope should always emit electrons with the same amount of kinetic energy. However, experimental tests instead found that measured electrons took on a range of energies between 0 and the decay Q-value, meaning that some fraction of the energy expected to be released in the decay was missing [15]. Pauli proposed a solution wherein the missing energy was carried away by undetected particles. It was not until 1956 that these particles, which came to be known as antineutrinos, were finally directly observed [16, 17].

The reason (anti)neutrinos were so challenging to discover is that they interact with other particles via the weak force, leading to low interaction cross-sections. (Anti)neutrinos are neutral particles, meaning that they do not participate in electromagnetic interactions. They also do not carry color charge, meaning that they cannot interact via the strong force, and at the time when the Standard Model was formulated, there was no evidence that neutrinos interacted gravitationally. Because the probability of a neutrino interacting with another particle is so low, they are very difficult to shield against or to detect.

Neutrinos can participate in both neutral current and charged current weak interactions. Charged current interactions, mediated by W-bosons, reveal another notable property of neutrinos, flavor. The three neutrino flavors correspond to the three types of charged leptons: electrons, muons, and taus. A charged current interaction involving a charged lepton always involves a neutrino or antineutrino of the same flavor. For example, antineutrinos emitted in β^- decay are always electron-type because their creation is accompanied by the creation of an electron. In the original formulation of the Standard Model, lepton flavor was posited to be a conserved quantity, meaning that a neutrino created in a charged current interaction with a particular charged lepton could only interact with that same flavor of charged lepton in later interactions.

The discovery of flavor oscillation in neutrinos conflicts with the neutrino properties posited by the Standard Model. In doing so, it raised a number of compelling questions about neutrinos and physics beyond the Standard Model.

Section 1.2: Neutrino Oscillation and Massive Neutrinos

Neutrino flavor oscillation refers to the phenomenon in which a neutrino created with a known flavor will, if detected later in time, have some probability of having changed to a different flavor. This process violates the principle of conservation of lepton flavor originally posited by the Standard Model. For flavor oscillation to occur, each neutrino flavor state must be a superposition of mass eigenstates, meaning that a neutrino

cannot simultaneously have a definite flavor and a definite mass. When neutrinos undergo charged current interactions, they do so in flavor eigenstates, but neutrinos propagate through space in mass eigenstates. Each neutrino flavor eigenstate (ν_e, ν_μ, and ν_τ) can be expressed in terms of the mass eigenstates (ν_1, ν_2, and ν_3):

$$|\nu_\alpha\rangle = \sum_{j=1}^{3} U^*_{\alpha j}|\nu_j\rangle, \tag{1.1}$$

where U is a 3x3 matrix know as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix [18]. The PMNS matrix can be parameterized by three mixing angles (θ_{12}, θ_{23}, and θ_{13}) and a CP-violating phases (δ_{CP}). If neutrinos are their own antiparticles, a possibility that will be discussed in Sec. 1.3, the PMNS matrix also depends on two additional Majorana phases (φ_1 and φ_2).

To see how this superposition leads to detectable flavor oscillation, we must first consider how the mass eigenstates evolve with time. Since these states are, by definition, eigenstates of the free space Hamilitonian, their time evolution can be described through a plane wave solution

$$|\nu_j(t)\rangle = e^{-iE_j t}|\nu_j(0)\rangle, \tag{1.2}$$

written here in natural units where $c = \hbar = 1$. Because neutrinos can be treated relativistically, this equation can be approximated in terms of length travelled, L, and neutrino energy, E, as

$$|\nu_j(t)\rangle = e^{-\frac{iL}{2E}(p+m_j^2)}|\nu_j(0)\rangle. \tag{1.3}$$

The probability of a neutrino changing flavor when traveling through vacuum can then be found by:

$$P_{\nu_\alpha\to\nu_\beta} = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 = |\sum_j\sum_k U_{\beta k}U^*_{\alpha j}\langle\nu_j|\nu_k(t)\rangle|^2. \tag{1.4}$$

Substituting in Eq. 1.3 and making use of the orthogonality of the mass eigenstates to eliminate terms where $j \neq k$, we get

$$P_{\nu_\alpha\to\nu_\beta} = |\sum_j U_{\beta j}U^*_{\alpha j}e^{-\frac{iL}{2E}(p+m_j^2)}|^2. \tag{1.5}$$

When evaluating the modulus squared, we once again end up with a double summation:

$$P_{\nu_\alpha \to \nu_\beta} = \sum_j \sum_k U^*_{\beta j} U_{\alpha j} U_{\beta k} U^*_{\alpha k} e^{\frac{iL}{2E}(p+m_j^2)} e^{\frac{-iL}{2E}(p+m_k^2)} \tag{1.6}$$

For terms where $j = k$, the exponentials cancel out entirely, but the remaining terms each involve an exponential of the form $e^{\frac{iL}{2E}(m_j^2 - m_k^2)}$. For convenience, we define the notation $\Delta m^2_{jk} \equiv m_j^2 - m_k^2$.

The full expression can be written in a simplified form as

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta} = \delta_{\alpha\beta} &- 4 \sum_{j>k}^{n} \mathrm{Re}(U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k}) \sin^2\left(\Delta m^2_{jk} \frac{L}{4E}\right) \\ &+ 2 \sum_{j>k}^{n} \mathrm{Im}(U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k}) \sin\left(\Delta m^2_{jk} \frac{L}{2E}\right). \end{aligned} \tag{1.7}$$

Eq. 1.7 fully describes the probability of flavor change as a function of energy and length for neutrinos travelling through vacuum. In practice, this equation must be modified for neutrinos traveling through dense matter, such as neutrinos escaping the Sun. Once the effect of interactions is medium is properly accounted for, the findings of oscillation experiments can be used to constrain the oscillation parameters. Since the probability of flavor change in either vacuum or matter involves components of the PMNS matrix, it depends on the mixing angles and CP-violating phase, but the Majorana phases can be shown to cancel out when evaluating this expression. In addition to these 4 parameters from the PMNS matrix, the differences in squared masses play a crucial role in the probability of flavor change. If neutrinos were massless, all the Δm^2_{jk} terms would go to 0, and Eq. 1.7 would simplify to the no-oscillation expression $P_{\nu_\alpha \to \nu_\beta} = \delta_{\alpha\beta}$. Because flavor oscillation is not possible for massless neutrinos, the discovery of neutrino oscillation contradicted the Standard Model prediction that neutrinos are massless. One of the first indications of neutrino oscillation came from Ray Davis's Homestake experiment, which used radiochemical methods to detect solar neutrinos [19]. Solar neutrinos originate from fusion processes in the Sun and are produced in the electron flavor state. The Homestake experiment, which was only sensitive to electron-type neutrinos, found approximately one-third of the neutrino flux predicted by stellar modeling. Flavor oscillation offered a possible explanation of this deficit of electron-type neutrinos. This interpretation was later confirmed by the Super-Kamiokande and Sudbury Neutrino Observatory (SNO) experiments [20–22]. Super-Kamiokande also studied muon neutrinos produced in the Earth's atmosphere, including those originating on the opposite side of the Earth that were required to travel through the planet to reach the experiment. The difference in the flux of

downward- and upward-going muon neutrinos clearly indicated a distance-dependence in the probability of flavor change. SNO was a solar neutrino experiment, like the Homestake experiment, but it made independent measurements of neutral current and charge current interactions. Measuring these two channels enabled a comparison of the ν_e flux to the total neutrino flux, providing direct evidence that a fraction of the solar neutrinos changed flavor.

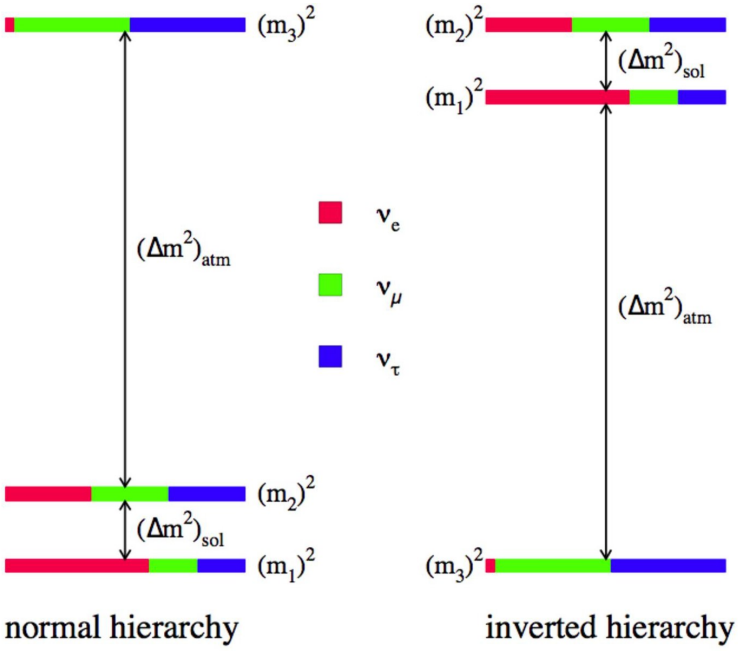


Figure 1.1: Representation of the two possible neutrino mass orderings, the normal ordering and the inverted ordering. Figure from [1]

Neutrino oscillation experiments are now at a precision measurement stage, where many of the oscillation parameters have been highly constrained. The squared sines of the mixing angles have been measured with ~3-4% precision, and the magnitudes of the Δm^2_{jk} are known with ~1-2% precision [2]. The sign of Δm^2_{12} is also well constrained, but the sign of $\Delta m^2_{13}/\Delta m^2_{23}$ and the CP phase are still largely unresolved [18]. The uncertainty in the sign of $\Delta m^2_{13}/\Delta m^2_{23}$ indicates that two possible orderings of the mass eigenstates are allowed. As shown in Fig. 1.1, the normal ordering refers to the case where the smaller mass separation is between the two lightest mass states, contrasted to the inverted ordering where the two heaviest states are closer together. Global fits of oscillation experiments currently favor the normal ordering scenario, but the question cannot be conclusively resolved without data from the next generation of oscillation experiments [18]. The Majorana phases cannot be determined from oscillation experiments at all, since they are not present in the final neutrino oscillation equation.

The absolute scale of the neutrino mass is also not measurable from neutrino oscillation experiments. However, other neutrino experiments indicate that the absolute mass scale must be at least six orders of magnitude below the next lightest Standard Model particle, the electron. Direct kinematic neutrino mass experiments are sensitive to m_β, which is the incoherent sum of the neutrino mass eigenstates, $m_\beta = \sum_k m_k |U_{ek}^2|$. The current best limit of 0.8 eV comes from the KATRIN experiment, which is kinematically searching for neutrino mass based on the shape of the tritium beta decay spectrum near its endpoint [23]. Cosmological data can be used to probe the sum of neutrino mass states, Σ, using observables such as the quantity of large-scale structure formation in the universe. The current bound from Planck found Σ to be less than 0.26 eV, but the combined results from many sources sets a more aggressive limit of $\Sigma < 0.09$ eV [18, 24]. The question of why neutrinos are so light relative to other Standard Model particles guides many of the theories of how neutrinos acquire their mass, to be discussed more in Sec. 1.4.

Section 1.3: Dirac vs. Majorana Neutrinos

The finding that neutrinos have mass introduced a fundamental question about the relationship between neutrinos and antineutrinos. For charged leptons, electric charge provides a quantum number by which leptons and antileptons can be clearly distinguished. Electrons and positrons, for example, must be distinct particles, since electric charge is a conserved quantity, preventing electrons from being converted into positrons or vice versa. For neutral leptons, the distinction between particles and antiparticles is less clear. In the case of a massless neutral particle, as neutrinos were originally thought to be, helicity could be used as a distinguishing quantity. A massless neutrino would be definable by its negative helicity, or the fact that its spin and momentum are oppositely directed, in contrast to a positive helicity antineutrino, where spin and momentum are aligned. However, the finding that neutrinos have mass prevents helicity from being an inherent property of a particle. This can be seen by the fact that for a massive particle, it is always theoretically possible to Lorentz boost into a reference frame where the particle is traveling in the opposite direction. Because the neutrino's spin is independent of reference frame, the neutrino's helicity is reference frame-dependent.

Based on this example, we can see that neutrinos must have a right-handed component, but the smallness of the neutrino mass means that this right-handed component is very highly suppressed. It is an open question whether this right-handed neutrino state is distinct from the right-handed antineutrino, which has been experimentally observed. Similarly, antineutrinos must have a very small admixture of left-handedness,

which could be either equivalent to or distinct from the left-handed neutrino state. This leaves us with two possibilities. If neutrinos are what is known as Dirac particles, they are totally distinct from their antiparticles. In this case, 4 degrees of freedom are required to fully describe all neutrino states (ν_L, ν_R, $\bar{\nu}_L$, $\bar{\nu}_R$). Conversely, neutrinos could be Majorana particles, meaning that they are their own antiparticles. Majorana neutrinos can be described with only 2 degrees of freedom.

If neutrinos are Dirac, some quantum number must exist that distinguishes the neutrino and antineutrino states with the same handedness. The natural choice for this quantity would be lepton number, which differentiates leptons from antileptons in the Standard Model. However, unlike in the original formulation of the Standard Model with massless neutrinos, if neutrinos are Dirac, lepton number must be imposed as a conserved quantity. The relationship between lepton number and the nature of the neutrino's mass will be discussed more in Sec. 1.5.

Section 1.4: Neutrino Mass Mechanism

The question of whether neutrinos are their own antiparticles is deeply tied to another question raised by the discovery of neutrino oscillation: the question of what underlying mechanism generates the non-zero neutrino mass. Since neutrinos were treated as massless in the formulation of the Standard Model, the model must be extended to incorporate a mechanism by which a mass can be acquired. There are two approaches by which this can be accomplished.

One method of generating neutrino masses is to postulate that neutrinos acquire their masses through the same basic mechanism as other fermions, the Higgs mechanism. However, when a particle interacts with the Higgs field, it changes handedness, so to generate a neutrino mass via the Higgs mechanism, the left-handed neutrino fields would be required to have at least one right-handed counterpart. Since only left-handed neutrinos have been observed in weak interactions, a minimal extension of the Standard Model where neutrinos acquire mass via the Higgs mechanism would require the presence of right-handed sterile neutrinos, which only interact gravitationally. Once such sterile neutrinos are introduced, it is possible to generate neutrino mass by adding a Yukawa coupling term, known as the Dirac mass term, to the Standard Model Lagrangian. In order to explain the neutrino's lightness, however, the strength of that coupling would need to be very small relative to the couplings for quarks and charged leptons.

In the case of Majorana neutrinos, the neutrino mass is not acquired through the same mechanism as other leptons but through either coupling with a new Higgs boson field or through the introduction of new

physics at a high energy scale [25]. The latter case is particularly interesting, because it implies the Standard Model is an effective field theory that breaks down above some cutoff energy scale, Λ. While this new physics could take many forms, below the cutoff it would manifest as an effective operator added to the Standard Model Lagrangian. The lowest dimensional operator that can be added to the SM Lagrangian while respecting gauge invariance is the dimension 5 Weinberg operator, which involves the lepton and SM Higgs doublets [26, 27]. At energies below electroweak symmetry breaking, this operator leads to a neutrino mass with a magnitude proportional to $\frac{v^2}{\Lambda}$, where v is the vacuum expectation value of the SM Higgs field. Having the neutrino mass be the consequence of two different symmetry-breaking scales could explain why its mass is qualitatively different from other Standard Model particles. One reason that the Majorana mechanism is compelling is that some high energy theories that incorporate the Weinberg operator can help to explain why the neutrino mass is so small. In high-energy seesaw mechanisms, for example, the masses of the light neutrinos could be inversely proportional to the mass of sterile heavy neutrinos that have not yet been observed. Discovering that neutrinos are their own antiparticles would not provide a complete answer for how neutrinos acquire their mass, but it would play an important part in addressing the question.

Section 1.5: Lepton Number and $B - L$ Violation

Another key consequence of discovering neutrinos to be Majorana particles would be the violation of a thus far observed symmetry of the Standard Model. The principle of conservation of lepton number, L, states that no process should change the balance between leptons and antileptons. All leptons are given a lepton number of 1, while antileptons have a lepton number of -1, so processes that create a lepton must simultaneously destroy another lepton or produce an antilepton to leave the total lepton number unchanged. A corresponding conservation law also exists for baryon number, B. Within the framework of the Standard Model, lepton number and baryon number are both anomalous symmetries, meaning that they can be violated by quantum fluctuations [28]. However, the combined quantity $B - L$ (baryon number minus lepton number) is a non-anomalous symmetry of the Standard Model, since anomalies in the baryon and lepton sector cancel out.

Majorana neutrinos violate both L and $B - L$. A Majorana neutrino has a small probability of behaving like a SM neutrino in one interaction and interacting like a SM antineutrino at a later time, meaning it violates lepton number by two units while leaving baryon number unaffected.

B, L, and $B - L$ are all associated with global symmetries, meaning that the Standard Model Lagrangian

is invariant under a constant transformation, unlike the gauge symmetries that define the Standard Model and give rise to particle interactions, where the transformation is a function of spacetime. Within the framework of the Standard Model, these three global symmetries are also accidental, meaning that they emerge from the SM rather than being externally imposed. While there is currently no experimental evidence of $B - L$ violation, there is not a fundamental reason why new physics at energy scales not yet probed must conserve $B - L$. Many grand unified theories violate lepton number and sometimes $B - L$. Attaining experimental evidence of $B - L$ violation would be a clear indication of physics beyond the Standard Model, and the level of violation could help distinguish between beyond the Standard Model (BSM) theories.

Section 1.6: Matter-Antimatter Asymmetry

$B - L$ violation would be an interesting discovery in its own right, but it also has the potential to weigh in on one of the biggest open questions in cosmology: the origin of the universe's matter-antimatter asymmetry. In principle, at the Big Bang, baryons and anti-baryons should have been produced in equal quantities. However, in the early universe, a small asymmetry emerged. While the vast majority of baryons and anti-baryons annihilated into photons, roughly one out of every 2×10^9 baryons remained, leading to the matter-dominated universe that we inhabit [29]. In 1967, Andrei Sakharov specify three requirements necessary to generate such an asymmetry: baryon number violation, CP violation, and interactions out of thermal equilibrium [30]. While there exist processes within the Standard Model that do not conserve the sum of baryon and lepton number, B+L, these processes are incapable of generating a sufficient degree of asymmetry to explain the measured imbalance, implying that a solution beyond the Standard Model is required. A discovery of neturinoless double beta decay would help satisfy the Sakharov conditions in two ways. First, the two additional phases in the PMNS matrix if neutrinos are Majorana fermions serve as an additional source of CP violation. Second, lepton-number-violating processes may be connected to the required baryon number violation.

Although a discovery of lepton number violation does not automatically constitute an explanation of the baryon asymmetry, it does strengthen the argument for BSM processes that can induce greater matter-antimatter imbalances than those allowed in the Standard Model. In addition, there exist a large landscape of theoretical leptogenesis models, models where an asymmetry between leptons and antileptons is converted into a baryon sector asymmetry. Like baryons, leptons and anti-leptons are expected to have been created in equal proportion at the Big Bang, so explaining the baryon sector asymmetry through leptogenesis requires

BSM lepton number violating processes. Many BSM theories connect the explanation of neutrino mass generation with leptogenesis, such as the original leptogenesis mode where heavy right-handed Majorana neutrinos in the early universe experienced lepton number violating decays [2]. Theories of leptogenesis are difficult to test in practice, but the potential connection between Majorana neutrinos and leptogenesis provides another compelling reason to search for lepton number violation in the neutrino sector.

Section 1.7: Introduction to Double Beta Decay

Experimentally determining that neutrinos are Majorana particles would be a paradigm-shifting discovery and would tie to deep questions about the underlying symmetries of the Standard Model and possible new physics at high energy scales. However, because the neutrino mass is so close to zero, processes that can be used to distinguish between Majorana and Dirac neutrinos are highly suppressed, making it challenging to realize any experiment capable of drawing a definitive conclusion. The only viable process for determining whether neutrinos are their own antiparticles is neutrinoless double beta decay, $0\nu\beta\beta$, which is the subject of this work. To understand what neutrinoless double beta decay is and how experiments can be optimized to search for it, it is first necessary to understand its Standard Model counterpart, two-neutrino double beta decay.

1.7.1: Two-Neutrino Double-Beta Decay

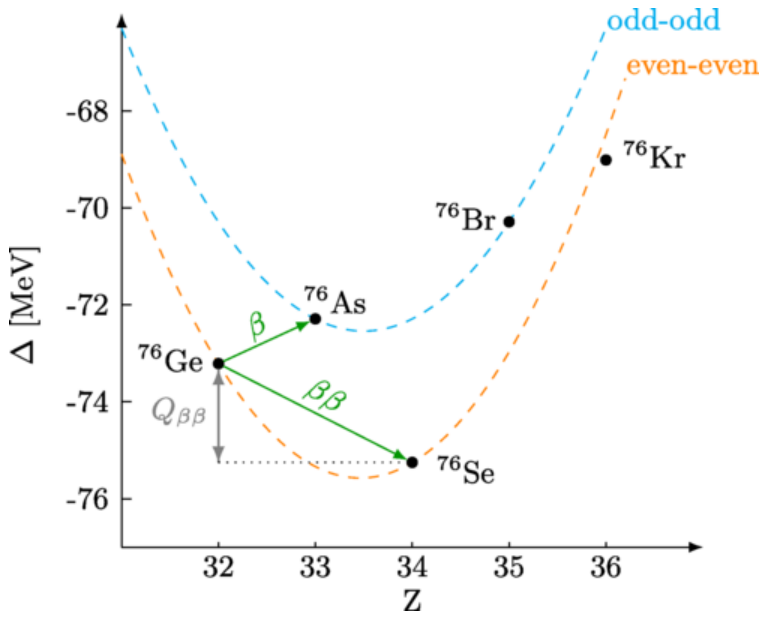


Figure 1.2: Isobar curves for a mass number of 76, where Δ represents the difference between the isotope mass and the mass of an equivalent number of unbound nucleons. A single beta decay from ^{76}Ge to ^{76}As is energetically forbidden, but two simultaneous decays to ^{76}Se is allowed. Figure from [2]

While beta decay is prevalent throughout nature, in a small number of isotopes, single beta decay is forbidden or highly suppressed. The most common example of this phenomenon occurs in isotopes where single beta decay is energetically disallowed because it would result in a less tightly bound nucleus. Even-even isotopes, or isotopes containing even numbers of both protons and neutrons, tend to be more tightly bound than their odd-odd counterparts with the same number of nucleons. Therefore, in some cases, such as the example of ^{76}Ge shown in Fig. 1.2, a single beta decay is forbidden, but two simultaneous beta decays can occur. This process of double-beta decay, first postulated by Maria Goeppert Mayer in 1935 [31] is a second order weak process, meaning it is highly suppressed relative to single-beta decay and has very long half-lives ($> 10^{18}$ yrs) [32]. The mode of double-beta decay allowed by the Standard Model is called two-neutrino double-beta decay ($2\nu\beta\beta$) because the change of atomic number by two units is accompanied by the emission of two antineutrinos in addition to two electrons ($(A, Z) \rightarrow (A, Z+2) + 2e^- + 2\bar{\nu}_e$). Since this process produces two leptons and two antileptons, it conserves lepton number. Despite its rarity, $2\nu\beta\beta$ has been experimentally observed in more than ten isotopes, making it one of the longest-lived decays ever detected [32].

1.7.2: Neutrinoless Double-Beta Decay

Neutrinoless double-beta decay is a version of double-beta decay not yet observed where no antineutrinos are emitted: $(A, Z) \rightarrow (A, Z+2) + 2e^-$. A number of possible mechanisms could contribute to $0\nu\beta\beta$, but any $0\nu\beta\beta$ signal requires physics beyond the Standard Model because the process violates lepton number (and $B - L$) conservation by two units.

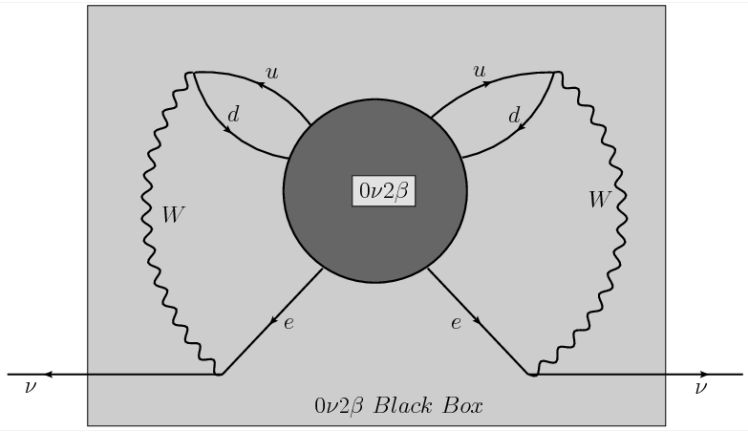


Figure 1.3: Black box diagram demonstrating that neutrinos must be Majorana particles if $0\nu\beta\beta$ is observed. Figure from [3]

Without making any assumptions about the mechanism by which it occurred, discovering neutrinoless

double-beta decay would imply that neutrinos are Majorana fermions. The Schechter-Valle black box theorem, illustrated in Figure 1.3, shows that an antineutrino can be converted to a neutrino using only a combination of Standard Model vertices and the unspecified $0\nu\beta\beta$ interaction [4]. While this theorem does not suggest that the illustrated process significantly contributes to the Majorana mass, it does indicate that, if $0\nu\beta\beta$ occurs, neutrinos cannot be Dirac particles.

In general, the $0\nu\beta\beta$ rate can be written as a sum of the rate due to each mechanism:

$$[T^{0\nu}_{\frac{1}{2}}]^{-1} = \sum_i G_i g_i^4 |M_i|^2 f_i(\Lambda) + \text{interaction terms}, \tag{1.8}$$

where f_i is the physical parameter corresponding to the $0\nu\beta\beta$ mechanism, written as a function of Λ, the scale of the lepton-violating process [2]. G_i is a phase space factor that takes into account the kinematics of the interaction and the number of possible final states, g_i is the hadronic matrix element containing information about the weak vertices involving the nucleons, and $M_i^{0\nu}$ is the nuclear matrix element, which contains more detailed information about nuclear structure. The most commonly considered mechanism, known as light neutrino exchange, introduces only a minimal extension to the Standard Model and is the mechanism associated with the dimension 5 operator.

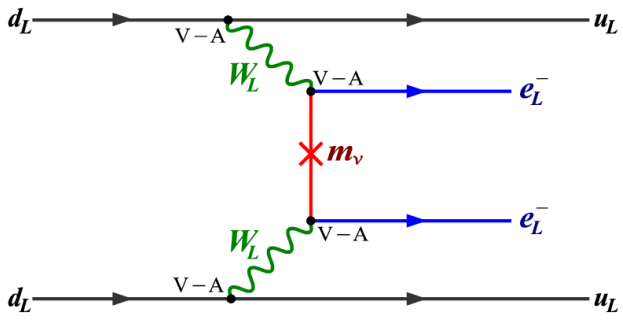


Figure 1.4: Feynman diagram for $0\nu\beta\beta$ via light neutrino exchange. Figure from [4]

Light neutrino exchange involves a right-handed Majorana neutrino being emitted from one weak vertex, changing helicity, and then being reabsorbed at the other vertex, as shown in Figure 1.4. If this mechanism dominates, the half life equation can be approximated as

$$[T^{0\nu}_{\frac{1}{2}}]^{-1} = G_{01} g_A^4 |M_{light}|^2 \frac{m^2_{\beta\beta}}{m_e^2}, \tag{1.9}$$

where the phase space factor, G_{01}, and nuclear matrix element, M_{light}, are specific to light neutrino exchange, m_e is the mass of the electron, and $m_{\beta\beta}$ is the effective Majorana neutrino mass.[1] $m_{\beta\beta}$ is the sum of the neutrino mass eigenstates weighted by the elements of the PMNS mixing matrix relating those eigenstates to the electron neutrino: $m_{\beta\beta} = \left| \sum_k m_k U_{ek}^2 \right|$. Because $m_{\beta\beta}$ is a coherent sum, it explicitly depends on the unknown Majorana phases that cannot be measured by oscillation experiments. Under the assumption of light neutrino exchange, a measurement of the neutrinoless double-beta decay half life can be used to determine the absolute mass scale of the neutrino, which also cannot be obtained from oscillation experiments .

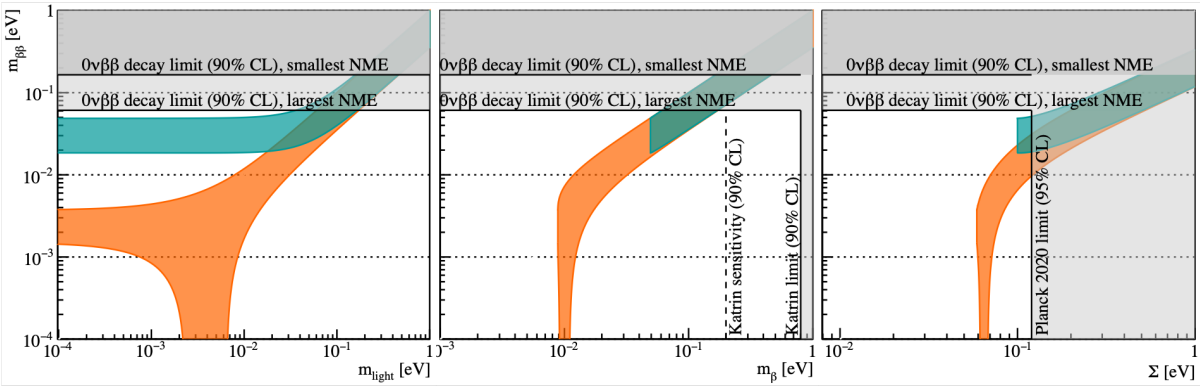


Figure 1.5: $m_{\beta\beta}$ parameter space as a function of the lightest neutrino mass (left), the effective kinematic electron neutrino mass (center), and the sum of the neutrino masses (right). The blue band indicates the inverted ordering allowed parameter space, while the orange band is the parameter space for the normal ordering. Figure from [2]

Based on the mixing angles and Δm_{jk}^2 values measured by neutrino oscillation experiments, the range of parameter space that $m_{\beta\beta}$ may occupy can be plotted as a function of the lightest neutrino mass. The $m_{\beta\beta}$ distribution can also be plotted against parameters that other kinds of neutrino experiments are sensitive to: the effective electron neutrino mass, m_β, or the sum of the neutrino mass states, Σ. These type of plots are shown in Figure 1.5, along with the current best limits on $m_{\beta\beta}$, m_β, and Σ. These different parameters relating to the absolute neutrino mass scale allow complementary measurements from different kinds of experiments to more fully elucidate the mass of neutrinos.

Figure 1.5 makes clear that the possible range of values that $m_{\beta\beta}$ can take is dependent on the neutrino mass ordering. In the case of the inverted ordering, where the heavier two mass states have a smaller separation, $m_{\beta\beta}$ must have a value on the order of or above tens of meV. In the case of the inverted ordering,

[1]The light neutrino exchange nuclear matrix element contains both short and long range components. While the long-range axial-vector coupling, g_A is factored out, the short range part is proportional to a different coupling [2].

experiments with a large enough half life sensitivity would be guaranteed to either discover or rule out $0\nu\beta\beta$ via light neutrino exchange.

It should be noted that the maximum available $m_{\beta\beta}$ parameter space is typically plotted on logarithmic scales, as in Fig. 1.5. This choice calls attention to the low mass region of the normal ordering parameter space, which is beyond experimental reach. However, if no mechanism that drives $m_{\beta\beta}$ or m_{light} towards 0 is invoked, this region if disfavored by naturalness arguments, since it comprises only a small portion of the available parameter space. The most probable range of $m_{\beta\beta}$ values based on this argument can be found by performing a Bayesian analysis with a flat prior on the Majorana phases. This was done in [5], as shown in Fig. 1.6. Under this set of assumptions, an experimental program capable of probing values of $m_{\beta\beta}$ at the 10-20 meV level would have a ~50% probability of discovering $0\nu\beta\beta$ in the normal ordering scenario, in addition to having a $> 95\%$ discovery probability for the inverted ordering. Although half-life goals for the next generation of $0\nu\beta\beta$ experiments are typically framed in terms of the inverted ordering, these experiments have a reasonable chance of discovering $0\nu\beta\beta$ even if oscillation experiments rule out the inverted ordering.

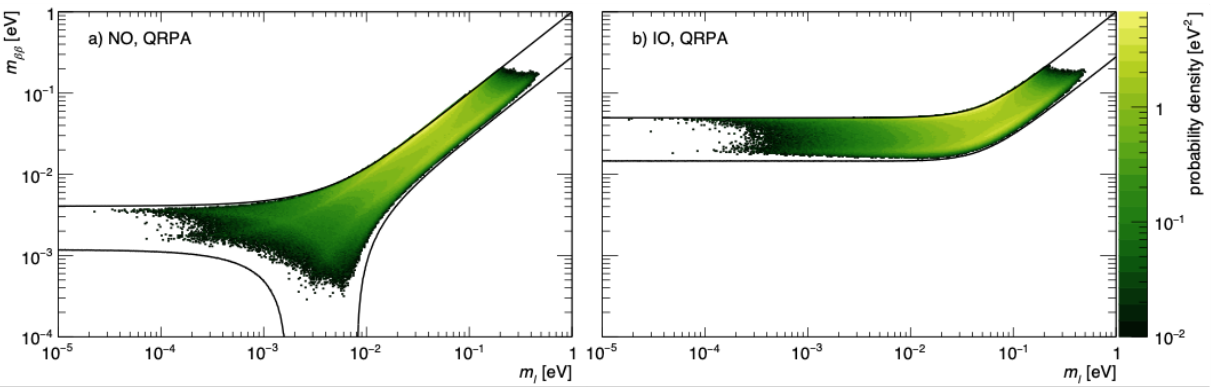


Figure 1.6: Marginalized posterior distribution for $m_{\beta\beta}$ and the lightest neutrino mass based on a Bayesian analysis where the prior incorporates the assumption that no external mechanism drives $m_{\beta\beta}$ or m_l to 0. Figure from [5].

While $0\nu\beta\beta$ could be generated by numerous lepton-number-violating mechanisms, light neutrino exchange is theoretically well-motivated and provides well-quantified goals for the next generation of neutrinoless double-beta decay experiments, especially if neutrinos are found to follow the inverted mass ordering. Different nuclear matrix elements for different $\beta\beta$ decay isotopes lead to different half-life requirements to cover the inverted ordering parameter space, but in general half-life sensitivities on the order of 10^{28} years are necessary. Searching for such a rare decay poses experimental challenges, but significant progress has

been made towards developing techniques to maximize the of sensitivity $0\nu\beta\beta$ experiments.

Section 1.8: Experimental Requirements of Double-Beta Decay Searches

1.8.1: Experimental Signature

For most mechanisms, including light neutrino exchange, neutrinoless double-beta decay would manifest as a peak at the decay Q-value. In the Standard Model process of two-neutrino $\beta\beta$-decay, the energy released in the decay is distributed between the electrons and antineutrinos, just as in single β-decay. Because neutrinos are weakly interacting particles, the two antineutrinos produced in $2\nu\beta\beta$ decay escape the experiment undetected, while the two electrons deposit their energy in the detector volume. The measured sum electron energy takes on a range of values between zero and the decay Q-value, $Q_{\beta\beta}$. In contrast, in neutrinoless double beta decay, no neutrinos are emitted to carry energy out of the detector, so the energy of the two electrons always sums to $Q_{\beta\beta}$. Because neutrinoless $\beta\beta$-decay would be highly suppressed relative to two-neutrino $\beta\beta$-decay, the number of counts in the $0\nu\beta\beta$ peak would be orders of magnitude below the the number of counts making up the broad $2\nu\beta\beta$ spectrum.

1.8.2: Maximizing Sensitivity to $0\nu\beta\beta$

Since $0\nu\beta\beta$ searches aim to detect a peak from a rare decay at a known energy, the main experimental challenge is to maximize the number of $0\nu\beta\beta$ signal events while minimizing the number of background events near $Q_{\beta\beta}$. The number of signal events is the product of the number of $0\nu\beta\beta$ decays occurring in the detector array, N_d, and the efficiency of detecting a given event, ϵ. Because the half-lives associated with double-beta decay are very long, the decay rate can be considered constant over the runtime of the experiment, meaning the number of $0\nu\beta\beta$ decays is directly proportional to the experiment's runtime, t, and to the number of atoms of the $\beta\beta$ decaying isotope. The number of detected signal counts, N_s, can therefore be written as:

$$N_s = \ln(2)\frac{N_A}{W}\left(\frac{a\epsilon Mt}{T^{0\nu}_{1/2}}\right), \tag{1.10}$$

where N_A is Avogadro's number, W is the molar mass of the $\beta\beta$-decay isotope, a is the isotopic abundance, and M is the total mass of all detectors [33]. This equation suggests a number of ways experiments can be designed to increase the number of signal events observed. To maximize the detection efficiency ϵ, $0\nu\beta\beta$ decay experiments typically utilize technologies where the $\beta\beta$-decay source is internal to the detector. For

example, a fraction of the atoms making up germanium detectors are ^{76}Ge, meaning that the probability of a $\beta\beta$-decay depositing all of its decay energy within the detector is high. The isotopic abundance a represents the fraction of the total detector mass that the $\beta\beta$-decay isotope constitutes. In many cases, the natural abundance of the double-beta decaying isotope is small, but isotopic enrichment enables substantial increases in a. A large exposure, $E = Mt$, is also necessary to increase the probability of observing a rare event. Since the half-life of neutrinoless double-beta decay is many orders of magnitude larger than the age of the universe, the only way to have a high probability of even a single decay occurring in the detector is to observe a large number of nuclei capable of undergoing double-beta decay for a long interval of time. Once an experiment has an isotopic abundance close to 1, increasing the total detector mass is the only way to significantly increase the amount of $\beta\beta$-decaying isotope.

Given the very high half-life limits for neutrinoless double-beta decay, a $0\nu\beta\beta$ signal can easily be obscured by even a small number of background counts. The level of backgrounds present at energies near $Q_{\beta\beta}$ can be expressed through a specific background rate, B, which has units of counts per unit mass per unit time per unit energy. Like signal counts, background counts are typically proportional to exposure, so the large detector masses necessary to maximize signal counts exacerbate the difficulty of minimizing background counts. The principal sources of backgrounds to $0\nu\beta\beta$ are discussed in the next section. A competitive $0\nu\beta\beta$ experiment must be designed to mitigate backgrounds as much as possible, typically through a combination of hardware and analysis-based background reduction techniques. Using a $\beta\beta$-decay isotope with a high Q-value can also facilitate a low background rate at $Q_{\beta\beta}$, since some common background sources contribute fewer counts or do not contribute at all at high energies. Finally, an experiment's energy resolution, ΔE, plays an important role in controlling the impact of backgrounds. The size of the energy interval necessary to contain the $0\nu\beta\beta$ signal peak is set by the energy resolution, and only background counts in this energy interval negatively impact the experiment's sensitivity. For the same background rate, an experiment with a better (lower) energy resolution will observe fewer background counts in the signal region.

The effects of the expected number of signal counts and the expected number of background counts at $Q_{\beta\beta}$ can be combined to determine an experiment's probability of detecting $0\nu\beta\beta$ for a given half-life [33].

The $0\nu\beta\beta$ half-life discovery sensitivity is governed by the following relationship:

$$S^{0\nu} \propto \begin{cases} aM\epsilon t & \text{Background-free or quasi-background-free} \\ a\epsilon\sqrt{\frac{Mt}{B\Delta E}} & \text{Background rate } B \end{cases} . \tag{1.11}$$

In this expression, discovery sensitivity always scales linearly with the isotopic abundance of the $\beta\beta$-decay isotope and with detection efficiency. However, the relationship between discovery sensitivity and exposure depends on whether the experiment can be considered quasi-background-free. A quasi-background-free experiment is one where B is sufficiently low that less than one background count is expected in the signal region over the experiment's full exposure. While discovery sensitivity typically scales as the square root of exposure, in the quasi-background-free scenario, sensitivity scales approximately linearly with exposure. As a consequence, operating in the quasi-background-free regime allows an experiment to explore a much larger range of half-lives for the same exposure. However, since the number of background events at $Q_{\beta\beta}$ increases as a function of exposure, a lower background index is required to fulfill the quasi-background-free condition for larger experiments. Experiments seeking to increase sensitivity to longer $0\nu\beta\beta$ half-lives while maintaining a reasonable runtime must increase detector mass while decreasing the background index.

1.8.3: Types of of Backgrounds in $0\nu\beta\beta$ Experiments

Backgrounds to $0\nu\beta\beta$ searches predominately originate from three sources: two-neutrino $\beta\beta$ decay, naturally-occurring radioactive decay chains, and cosmic rays. These sources produce multiple types of ionizing radiation, which can each be mitigated in different ways.

All candidate neutrinoless $\beta\beta$ decay isotopes also undergo two-neutrino double-beta decay. Two-neutrino $\beta\beta$ decay occurs with very long half-lives ($> 10^{18}$ years), but the decay rates are still orders of magnitude higher than the upper limits on the decay rate of neutrinoless $\beta\beta$ decay. Since $0\nu\beta\beta$ decay experiments are designed to maximize the number of signal events, the number of two-neutrino $\beta\beta$ decay events occurring in the detector is irreducible. However, two-neutrino $\beta\beta$ decay only poses a background for $0\nu\beta\beta$ if it contributes a significant number of counts to the signal region around $Q_{\beta\beta}$. Although the two-neutrino $\beta\beta$ decay spectrum spans from 0 keV to just below the decay Q-value, most of the decay phase space is near the middle of the spectrum, and only a very small fraction of the events occur near $Q_{\beta\beta}$. The magnitude of the background contribution in the signal region from events at the upper end of the two-neutrino $\beta\beta$ decay spectrum is directly related to an experiment's energy resolution. As energy resolution improves, the

width of the signal region decreases. Because the two neutrino spectrum drops off rapidly near $Q_{\beta\beta}$, small changes in the width of the signal region can have a large impact on the amount of background contributed by two-neutrino $\beta\beta$ decay. When energy resolution is sufficient, this background becomes negligible.

The dominant class of backgrounds at $Q_{\beta\beta}$ is generally from naturally-occurring radioactive decay chains. These decay chains are the result of primordial nuclides, long-lived isotopes created prior to the Earth's formation that are present to varying degrees in all materials. These primordial nuclides decay into radioactive progeny and then proceed through a series of additional radioactive decays with half-lives ranging over many orders of magnitude. The decays occur via the emission of alpha or beta particles. In some cases, there is a probability of decaying into excited progeny, which subsequently de-excite by emitting gamma rays. Alphas, betas, and gammas can all lead to backgrounds in germanium detectors, although alphas and betas have much shorter path lengths in materials and can only contribute when they originate near detectors.

The three major naturally-occurring radioactive decay chains are the ^{232}Th chain, the ^{238}U chain, and the ^{40}K chain. Events from all three of these chains are found in all $0\nu\beta\beta$ experiments and must be taken into account when modeling backgrounds, but not all decay chains produce sufficiently high-energy backgrounds to impact $0\nu\beta\beta$ searches. For example, the highest energy decay product in the ^{40}K chain is a 1460 keV gamma, which is lower than the Q-values of all commonly used $\beta\beta$-decay isotopes. The ^{232}Th and ^{238}U decay chains, which do contribute backgrounds near the Q-values of most $\beta\beta$ decay isotopes including ^{76}Ge, are shown in Fig. 1.7. One example of a decay product that can act as a background to $0\nu\beta\beta$ searches is a 2615 keV gamma in the ^{232}Th decay chain, which always accompanies the decay of ^{208}Tl into ^{208}Pb. If this gamma Compton scatters, it can deposit energy near the Q-value of $\beta\beta$ decay isotopes like ^{76}Ge and ^{136}Xe.

The other major source of backgrounds in $0\nu\beta\beta$ experiments is related to cosmic rays, high energy particles impinging on Earth from space. When cosmic rays interact with Earth's atmosphere, they produce showers of energetic particles including muons. Atmospheric muons can deposit energy directly in detectors, but they can also interact with materials to create secondary particles such as neutrons. Muons can be effectively shielded by going underground, after which the most significant class of backgrounds related to cosmic rays is caused by the delayed decay of cosmogenically-activated nuclei. When cosmic ray secondaries interact with nuclei in or near detectors, they can knock off nucleons to create new isotopes that are long-lived but unstable. When these isotopes eventually decay, they lead to additional backgrounds. Because some cosmogenic isotopes have half-lives of greater than a year, cosmogenic activation occurring before the start of an experiment can be a major source of backgrounds over the entire lifetime of the experiment. Since

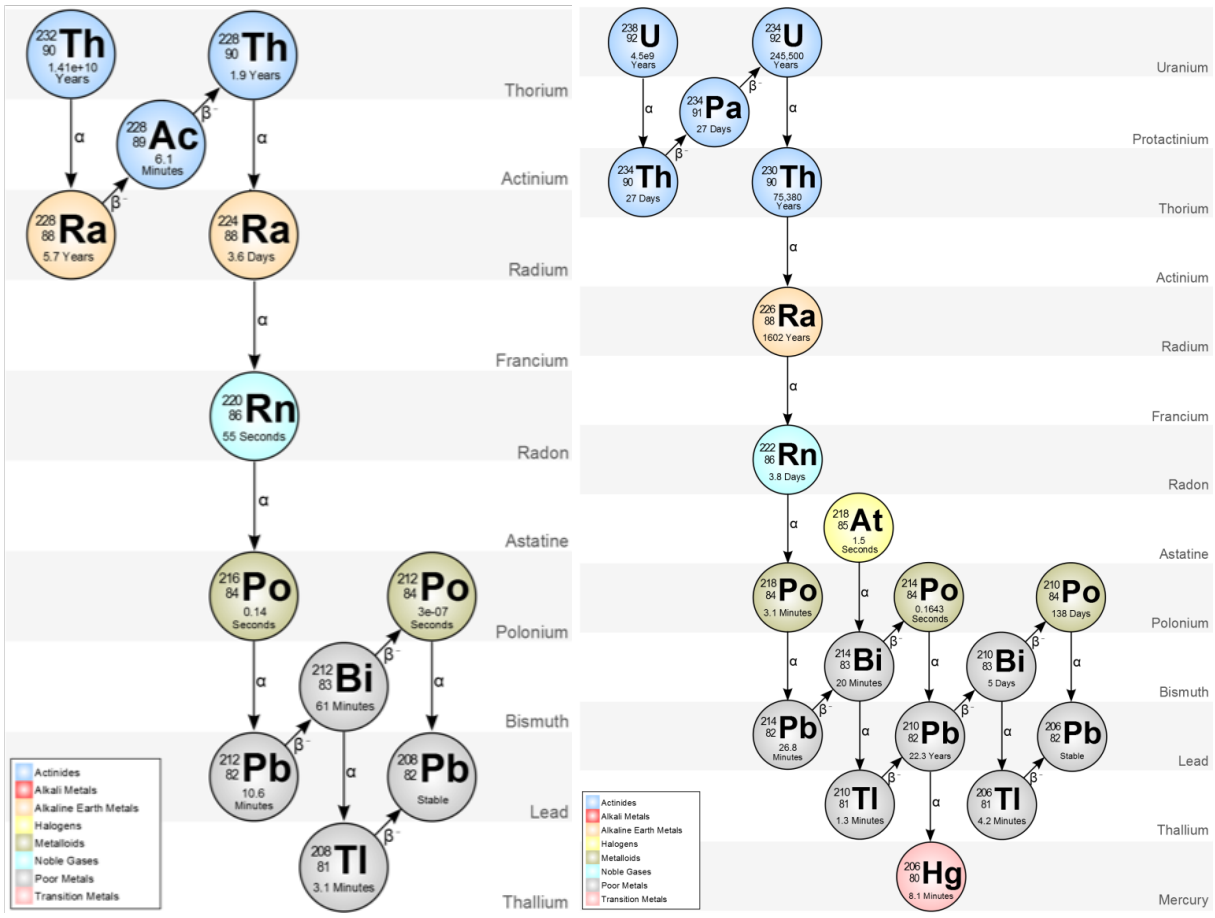


Figure 1.7: ^{232}Th and ^{238}U decay chains, which produce backgrounds with energies near the Q-values of most $\beta\beta$ decay isotopes. ^{208}Tl from the ^{2615}Th chain and ^{214}Bi from the ^{238}U chain produce high energy gammas that are of particular concern to $0\nu\beta\beta$ experiments. *Image credit: Wikipedia*

these backgrounds do not occur in coincidence with the large energy deposits associated with muons, they can be difficult to reject.

Section 1.9: Current Landscape of Neutrinoless Double-Beta Decay Experiments

Over the last decade, major progress has been made in improving $0\nu\beta\beta$ half-life limits and in developing, testing, and improving the technologies to be used in the next generation of experiments. Over the next one to two decades, $0\nu\beta\beta$ experiments aim to cover the full inverted mass ordering parameter space for the light neutrino exchange mechanism. This goal, which corresponds to probing half-lives on the order of 10^{28} years, would also cover a large portion of the normal ordering parameter space [5]. $0\nu\beta\beta$ decay searches have been conducted using a number of different detector technologies including liquid scintillators, bolometers, time-projection chambers, and semiconductor detectors. Additional technologies used in $0\nu\beta\beta$ are discussed in [2]. Each experimental approach has different advantages and drawbacks for meeting the requirements of

an ideal $0\nu\beta\beta$ search.

Liquid scintillator experiments use a monolithic detector design that makes them easily scalable to larger masses by allowing the inside of the detector to experience significant self-shielding. However, liquid scintillators cannot reach the same low energy resolution achievable with other detector technologies, leading to irreducible $2\nu\beta\beta$ backgrounds. The current world-leading limit on the half-life of $0\nu\beta\beta$ was set by the Kamland-Zen collaboration using a 1 kton volume of liquid scintillator surrounding an inner balloon containing Xe-loaded liquid scintillator [34]. Kamland-Zen's 90% confidence level upper limit of $T^{0\nu}_{1/2} > 2.3 \times 10^{26}$ yr for ^{136}Xe was achieved with 970 kg yrs of exposure. The SNO+ experiment also plans to search for neutrinoless double-beta decay using liquid scintillator technology but with a different double-beta decay isotope, ^{130}Te.

Time projection chambers (TPCs), as large monolithic detectors, have some similar advantages and challenges to liquid scintillator experiments. As detector mass scales, the inner regions of a TPC substantially benefit from self-shielding. However, the intrinsic energy resolution is modest, particularly for liquid phase TPCs. Unlike the liquid scintillator experiments described above, TPC-based $0\nu\beta\beta$ experiments also benefit from background reduction through particle discrimination. In a time projection chamber, both the ionization and scintillation signals are read out for each particle interaction in the detector. The ratio between these signal amplitudes differs between alpha particles and betas/gammas, allowing alpha backgrounds to be rejected. TPCs also facilitate event position reconstruction, which allows for backgrounds that deposit energy in multiple locations to be tagged. The current most competitive $0\nu\beta\beta$ half-life limit achieved using a TPC comes from EXO-200, a 200 kg single-phase liquid Xe TPC [35]. EXO-200 measured $T^{0\nu}_{1/2} > 3.5\times10^{25}$ yr for ^{136}Xe. The next generation nEXO experiment plans to improve on the techniques demonstrated by EXO to probe the inverted ordering parameter space using 5000 kg of enriched liquid Xe [36].

Cryogenic bolometer experiments take a different approach to optimizing $0\nu\beta\beta$ sensitivity, operating arrays of crystal calorimeters at mK temperatures. Each crystal is formed from compounds that include a $\beta\beta$-decaying isotope, and radiation interactions in a detector, like those caused by $\beta\beta$ decay, lead to a detectable increase in the crystal's temperature. Cryogenic bolometers can achieve much better energy resolution than large monolithic detectors, but the large number of cables and support structures necessary to support a large mass of small detectors can contribute to their background rates. The cryogenic bolometer approach to searching for $0\nu\beta\beta$ was demonstrated effectively by CUORE, the Cryogenic Underground Observatory for Rare Events. CUORE set a lower limit of 2.2×10^{25} years on the $0\nu\beta\beta$ decay half-life of ^{130}Te based on

more than a ton year of exposure [37]. To probe a large range of $0\nu\beta\beta$ half-lives, the next-generation CUPID experiment is building on the experience of CUORE but improving its background reduction capabilities [38][39]. CUPID, or CUORE Upgrade with Particle IDentification, will read out both heat and scintillation signals for each event. The relative magnitudes of the two signals allows alpha particles to be distinguished from betas and gammas, majorly reducing the dominant background observed in CUORE. CUPID will also use a different $\beta\beta$-decay isotope from CUORE, ^{100}Mo, which has a higher Q-value that reduces gamma backgrounds.

Finally, germanium experiments apply the well-established technology of semiconductor detectors to searching for $0\nu\beta\beta$ decay in ^{76}Ge. Although germanium detectors are also relatively small and do not benefit from self-shielding that scales with experiment size, the best energy resolution and the lowest background index of all $0\nu\beta\beta$ experiments were measured in germanium, by the MAJORANA and GERDA collaborations respectively [11][40]. In addition to their inherently good energy resolution, the point contact germanium detectors used by MAJORANA and GERDA allowed for analysis-based rejection of certain background populations, such as multi-site gammas and surface alphas and betas. The MAJORANA DEMONSTRATOR, which is the subject of this work, also mitigated backgrounds by encasing arrays of germanium detectors in a graded passive shield, while GERDA immersed detectors in liquid argon that functioned as an active veto system. Both experiments made careful use of radiopure near-detector materials to further reduce background levels. The best $0\nu\beta\beta$ half-life limit in ^{76}Ge to date, 1.8×10^{26} yr, was measured by GERDA with over 100 kg yr of exposure. The expertise developed by MAJORANA and GERDA has informed the design of the next-generation $0\nu\beta\beta$ decay experiment in ^{76}Ge, LEGEND [10]. LEGEND, or the Large Enriched Germanium Experiment for Neutrinoless $\beta\beta$ Decay, is taking a phased approach, with a 200 kg phase currently operating and a tonne-scale upgrade planned. The advantages of germanium detectors for $0\nu\beta\beta$ searches in general, the experimental setup of the DEMONSTRATOR, and the sensitivity goals of LEGEND are described in more detail in Chapter 2.

The neutrinoless double-beta decay community has developed a diverse set of techniques for maximizing sensitivity in rare-event searches. These technological advances have already enabled the exploration of theoretically-interesting parameter space, and the next generation of experiments is well-suited to either rule out an important subset of the parameter space or to make a discovery of $0\nu\beta\beta$ decay.

CHAPTER 2: The MAJORANA DEMONSTRATOR

The MAJORANA DEMONSTRATOR operated as a $0\nu\beta\beta$ experiment at the Sanford Underground Research Facility (SURF) in Lead, South Dakota [41] from 2015 until the removal of enriched detectors in 2021. The DEMONSTRATOR was designed with three primary goals:

- To search for $0\nu\beta\beta$ decay
- To demonstrate the feasibility of a ton-scale $0\nu\beta\beta$ experiment in germanium.
- To exploit its low backgrounds across a wide energy range to search for other rare processes, particularly physics beyond the standard model

This chapter serves as an overview of the DEMONSTRATOR's design and findings, with a focus on the strategies used to reduce backgrounds. Because the unique properties of germanium detectors were central to the DEMONSTRATOR's approach, it is useful to begin with an introduction to the general operation of germanium detectors and to the detector geometries developed to maximize suitability for rare-event searches.

Section 2.1: Overview of Germanium Detectors

Germanium detectors are a type of semiconductor diode detector that have historically been used in a wide range of radiation spectroscopy measurements [42]. When radiation interacts with a semiconductor, electrons receive sufficient energy to overcome the material's bandgap and enter the conduction band, where they can move freely through the detector. In the presence of an electric field, conduction electrons and the positively-charged holes they leave in the valence band are drifted in opposite directions. The drift of charge carriers through the detector creates a current that ultimately leads to a measurable ionization signal. Since the quantity of freed charge carriers is proportional to the energy deposited in the detector, the magnitude of the ionization signal reflects the energy of the interaction.

Because solid state germanium detectors have a low ionization energy (~3 eV to create a single electron-hole pair), a large number of charge carriers are produced for an event of a given energy in a germanium detector compared to in other detector technologies, such as liquid or gas TPCs. As a consequence, the

energy of the event can be determined with higher precision, since statistical fluctuations in the number of charge carriers are less significant. The intrinsically good energy resolution of germanium detectors is one of their main advantages in searches for rare events, particularly those that manifest as peaks in an energy spectrum.

2.1.1: P-type HPGe Detectors

All intrinsic semiconductors in practice contain residual impurities that dominate the material's properties. The amount of impurities present can be controlled and taken advantage of in a process known as doping. The DEMONSTRATOR used p-type detectors, in which the impurities in the detector bulk are dominated by acceptor impurities. An acceptor impurity is an atom containing one fewer valence electron than the surrounding semiconductor material, such as a boron atom in a germanium detector. This gap in the valence band functions like an extra hole, and the increased probability of electron-hole recombination due to the presence of more holes leads to a reduction in free electrons.

In p-type germanium detectors, much of the detector surface is covered by a contact with a very high concentration of donor impurities, impurities containing an extra valence electron that can easily be moved to the conduction band. This contact is known as the n+ contact. A heavily doped p+ contact is created on the opposite side of the detector. The boundary between the n+ contact and the p-type bulk forms a p-n junction. P-n junctions are beneficial for constructing detectors, because electrons from the highly doped n-type material diffuse into the p-type material and combine with the holes, while holes from the p-type material diffuse into the n+ contact, creating a depleted region where the quantity of free charge carriers is greatly reduced. The buildup of a net positive charge in the n+ region and a net negative charge in the p region creates an electric field that prevents further diffusion. When the detector is reverse biased by applying positive voltage to the n+ contact and grounding the p+ contact, the depletion region is extended to include the entire p-type detector bulk. The use of high purity germanium with impurity concentrations of around 10^{10} impurities/cm^3, known as HPGe, allows detectors to be fully depleted at reasonable voltages (3000-5000 V). When energy from ionizing radiation frees electron-hole pairs in the depletion region, the high electric fields in the bulk sweep holes towards the p+ contact and electrons towards the n+ contact.

2.1.2: P-type Point Contact and Inverted Coaxial Point Contact HPGe Detectors

The choice of detector geometry for the HPGe detectors used in the MAJORANA DEMONSTRATOR was based on optimizing their pulse shape discrimination (PSD) capabilities. Pulse shape discrimination refers

to the ability to distinguish $0\nu\beta\beta$-like and background-like events based on the shape of the waveform they produce. To understand how detector geometry influences PSD, it is necessary to first lay out the basics of signal formation in germanium detectors.

When electrons and holes undergo motion within a semiconductor detector, they induce charge on the detector contacts. The charge induced on the p+ electrode as a function of time is read out through an electronics chain and digitized in order to record the waveform for each event. The amount of charge induced, Q_{ind}, over a short time interval by a single moving charged particle is given by the Shockley-Ramo theorem:

$$Q_{ind} = q\Delta\phi_0, \tag{2.1}$$

where q is the charge of the moving particle and $\Delta\phi_0$ is the change in weighting potential between the particle's position at the beginning and end of the interval [42]. The weighting potential as a function of position in the detector is found by solving the Laplace equation with the boundary conditions that the potential goes to 1 at the readout (p+) contact and 0 at the other (n+) contact. Because the weighting potential is independent of the charge density in the detector and the voltages applied at the boundaries, it is only a function of the detector geometry.

The majority of detectors operated in the Demonstrator utilized a p-type point contact (PPC) geometry [43, 44]. This detector design is comprised of a cylindrical crystal where the n+ electrode makes up the top and side of the detector, while the p+ contact (or point contact) is limited to a small region on the bottom of the detector. Most of the detector's bottom is a passivated surface separating the p+ and n+ electrodes. The weighting potential map for a typical PPC is shown in Fig. 2.1. The small size of the p+ contact leads to a weighting potential that is relatively uniform and small throughout the majority of the detector bulk but rises sharply near the point contact. The Demonstrator, also included natural detectors with a slightly different geometry, referred to as Broad Energy Germanium detectors (BEGEs) [45]. The BEGe geometry is a variation on the point contact design where the n+ contact wraps around to cover a portion of the detector bottom and the p+ contact is larger. The weighting potential map and waveform shapes of BEGEs are very similar to those of PPCs because of the similarity in detector geometry.

According to the Shockley-Ramo theorem, large amounts of charge are induced during times when electrons or holes are traveling through regions of the detector where the weighting potential is rapidly

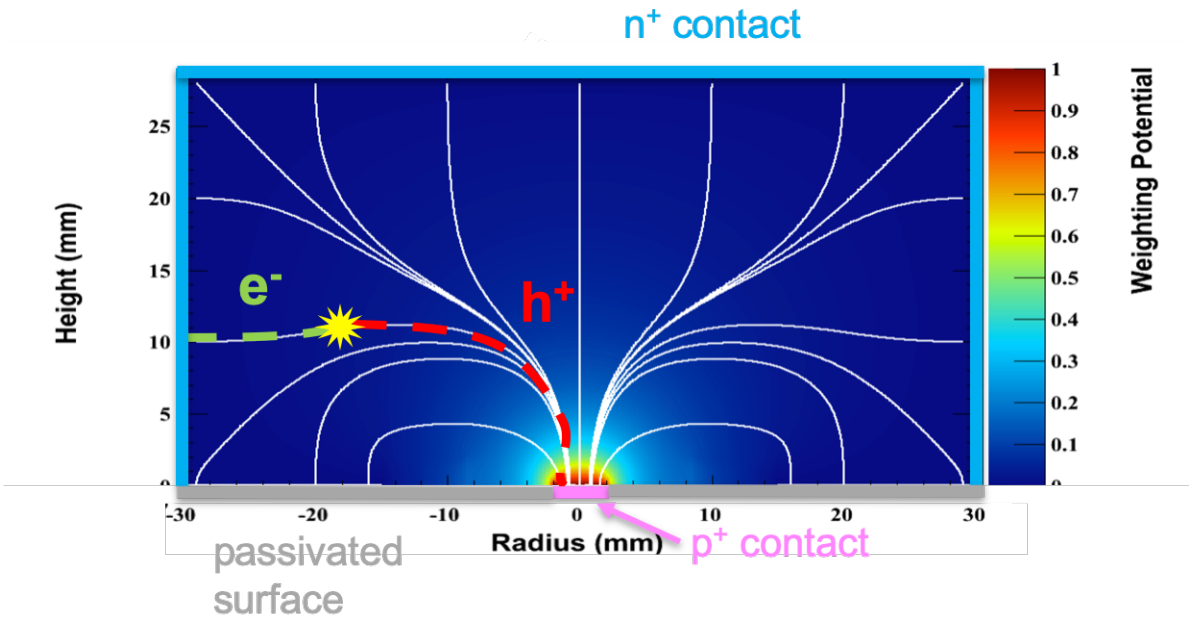


Figure 2.1: Weighting potential of a p-type point contact detector. White lines shown the drift paths of charged particles in the detector. The rapid increase in weighting potential near the p+ contact enables the good pulse shape discrimination capabilities of point contact germanium detectors.

changing. As a result, the vast majority of the signal from an energy deposit in a PPC is induced when the charge carriers are very close to the p+ contact. Since the waveform is a measure of induced charge, rather than charge that reaches the p+ contact, both electrons and holes contribute to the signal. However, except for energy deposits that occur very close to the p+ contact, the electrons typically only travel through the low weighting potential portion of the detector, meaning their contribution to the waveform is relatively small.

The highly concentrated weighting potential of PPCs leads to waveforms that consist of a gradual, shallow rise followed by a sharp increase when the holes reach the high weighting potential region of the detector. The length of time in which the induced charge is slowly increasing, known as the drift time, is much longer than the risetime that characterizes the rapid increase, and it depends on the position of the initial energy deposit in the detector. Interactions occurring far from the point contact can have drift times that are up to ~1μs longer than near point contact interactions [7]. When ionizing radiation deposits energy in two or more distinct locations in the detector, two signals with different drift times sum to produce the total waveform. The difference in drift times between the two components of the charge signal leads to a waveform with a kink in the rising edge. The distinct shape of these multi-site events allows for analysis-based rejection of some backgrounds to $0\nu\beta\beta$. More details about the cut parameter used to identify multi-site events are

provided in Sec. 2.5.2.

Near the end of the DEMONSTRATOR's runtime, four enriched detectors with a different geometry were added into the detector array. The Inverted Coaxial Point Contact (ICPC) detector geometry was tested in the DEMONSTRATOR's setup for eventual use in LEGEND-200. ICPCs were designed to circumvent a known shortcoming of traditional PPCs, the limitation on their maximum size [46]. PPCs exceeding ~1 kg exhibit pinch-off, meaning that an undepleted region remains in the middle of the detector even when the detector is fully biased. Increased detector size is advantageous in low-background experiments because the lower number of detectors necessary to reach the experiment's mass goal reduces the required number of cables, detector mounts, and front end electronics, decreasing the amount of near-detector material that can contribute to the background rate. Additionally, increasing detector size decreases detectors' surface-to-volume ratio, reducing the amount of surface backgrounds per unit mass. The ICPC design allows for larger detectors by using a semicoaxial shape with an n+ electrode that covers the inner cylindrical well in addition to the majority of the detector's outer surface. This design prevents pinch-off by ensuring that all locations in the detector bulk are relatively near the charged electrode even as the size of the detector grows. Similar to PPCs, ICPCs have a small p+ contact on their bottom surface, resulting in a similar sharp increase in weighting potential near the readout contact. Due to the similar structure of the weighting potential, ICPCs exhibit the same good PSD as PPCs despite their bigger size.

The PPC and ICPC geometries result in waveforms with long drift times and short risetimes, enabling discrimination between single-site and multi-site events. The ability to use waveforms not only to determine event energies but also to reduce backgrounds through pulse shape discrimination makes HPGe point contact detectors an ideal technology to use in searches for $0\nu\beta\beta$.

Section 2.2: Design of the DEMONSTRATOR

The DEMONSTRATOR consisted of approximately 44 kg of p-type point contact HPGe detectors, 30 kg of which were enriched in the double beta decay isotope, ^{76}Ge, and 14 kg of which were comprised of natural germanium. The enrichment process resulted in detectors with a 87.4% isotopic abundance of ^{76}Ge [47]. This high enrichment fraction marks another advantages of germanium detectors for $0\nu\beta\beta$ searches, since high isotopic abundance increases the number of atoms capable of undergoing $\beta\beta$-decay. Both enriched and natural germanium detectors were divided between two copper cryostats. The modular approach of deploying multiple cryostats was chosen to show the design's scalability to higher masses, setting the stage

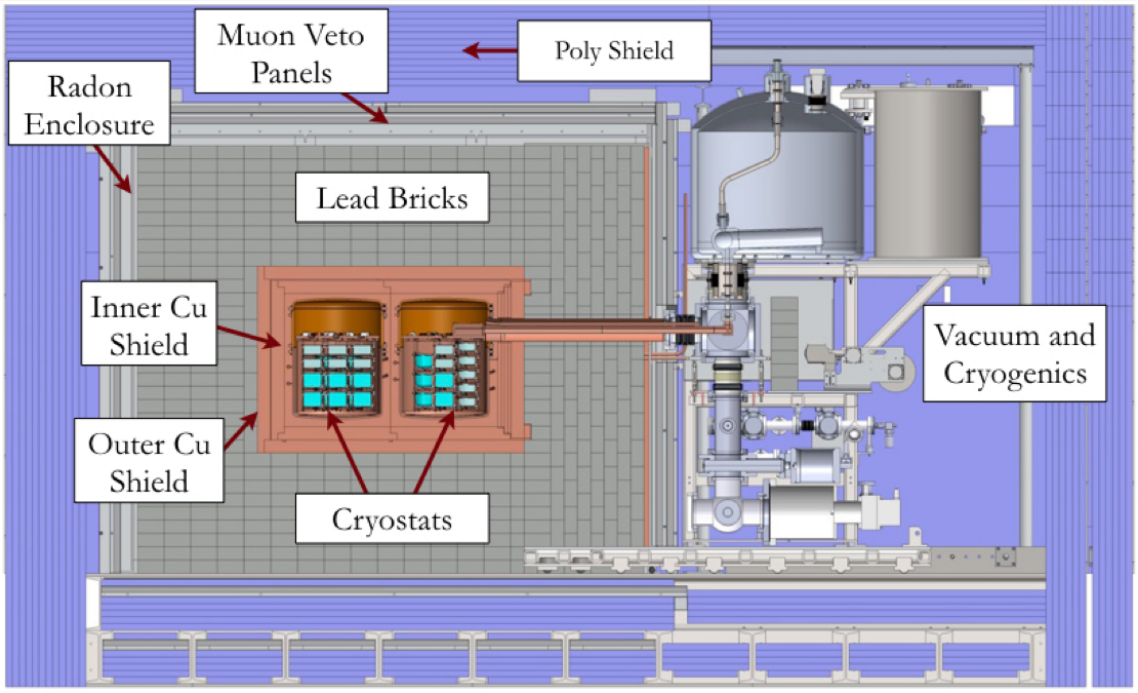


Figure 2.2: MAJORANA DEMONSTRATOR shield design [6]

for future experiments.

Each germanium detector was housed in a detector unit that contained all the hardware necessary to support the detector, apply voltage to its n+ contact, and read out and amplify the signal induced at the p+ contact. Up to five detector units were then stacked to form a detector string. Each cryostat contained seven such strings. The two cryostats were surrounded by multiple layers of shielding, as illustrated in Figure 2.2. A horizontal crossarm, through which high voltage and signal cables were routed, extended from each cryostat through the shield, as shown for the rightmost module in Fig. 2.2.

Germanium detectors require cryogenic temperatures to operate, so the detector arrays were kept under vacuum and cooled to a temperature of approximately 77 K. Each cryostat was cooled using liquid nitrogen via a two-phase thermosyphon [48]. Each module's crossarm contained a thermosyphon tube, which acted as a closed nitrogen loop. Outside the DEMONSTRATOR's shield, nitrogen within the closed system was cooled via heat exchange with an isolated source of liquid nitrogen. At the cryostat end of the crossarm, the thermospyhon cavity made thermal contact with the cold plate from which the detector strings hung. The heat load of the operating detector array caused nitrogen within the thermospyhon cavity to evaporate, cooling the array in the process.

Section 2.3: Hardware-Based Background Mitigation Strategies

Since low background rates are an essential component of increasing the sensitivity of rare-event searches, the experimental design of the DEMONSTRATOR was almost entirely driven by the challenge of minimizing the amount of background radiation that could reach the germanium detectors. This was achieved through a combination of techniques.

2.3.1: Underground Lab and Active Muon Veto

Locating the MAJORANA DEMONSTRATOR in an underground laboratory acted as one essential form of background reduction, providing a 4300 meters water equivalent rock overburden that shielded the experiment from cosmic ray muons. Atmospheric muons that did reach the experiment had a high probability of being detected by plastic scintillator panels located above, below, and to the sides of the detector array. If scintillation light from at least two panels was detected in coincidence, all events occurring in the germanium array in the preceding 20 ms and the following second were tagged. This approach effectively vetoed backgrounds due to muons and to secondary particles produced when the muons interacted with passive shielding. The active veto did not mitigate long-lived backgrounds due to cosmogenic activation, but the production rate for cosmogenics during the course of the experiment was low as a result of the rock overburden.

2.3.2: Graded Passive Shield

With the exception of the muon veto system, which served as an active shield, the DEMONSTRATOR mainly made use of passive shielding to reduce environmental backgrounds. The outermost shielding layer was polyethylene, which acted as a neutron moderator. The polyethylene consisted of a 25 cm layer of pure polyethylene to thermalize fast neutrons through scattering interactions, followed by a 5 cm layer of borated polyethylene, chosen for boron's high cross section for thermalized neutrons. The neutron shield surrounded the muon veto system, enclosed a radon seal. The area inside the radon enclosure was purged with liquid nitrogen boil-off to provide a radon-free environment. Directly inside the radon enclosure was located the thickest layer of shielding, 45 cm of lead. Because of its high atomic number, lead effectively attenuates gammas, including the gammas produced by neutron capture in the polyethylene shield. The lead shield surrounded a copper shield consisting of two 5-cm layers. The outer layer of copper was commercially acquired, while the inside layer was produced from highly radiopure underground electroformed copper (UGEFCu). The two cryostats, also constructed form UGEFCu, were located immediately inside the inner copper shield.

2.3.3: Radiopure Material Selection and Minimization of Surface Exposure

All of the materials used to construct the DEMONSTRATOR were carefully selected for their radiopurity. Particular care was taken to minimize the radioactivity of the detectors themselves and of parts located close to the detectors, such as detector holders, cabling, and near-detector electronics. The levels of radioactive contamination in all candidate materials and finished components were assessed through a detailed radioassay program [49]. Assay measurements included a combination of non-destructive and destructive techniques. The non-destructive methods, such as gamma-ray counting, could be uniformly applied to components prior to their use in the DEMONSTRATOR but were generally not sensitive enough to meet the DEMONSTRATOR's background goals. Destructive measurements, like mass spectroscopy, did meet the DEMONSTRATOR's stringent sensitivity requirements, but could only be applied to small representative samples. The combination of multiple assay techniques aimed to achieve the most reliable possible activity estimations for each part of the experiment.

The enrichment, zone refining, and crystal pulling processes involved in producing enriched PPCs intrinsically caused the detectors themselves to have very low levels of ^{232}Th and ^{238}U. The DEMONSTRATOR observed no evidence of internal ^{232}Th and ^{238}U in its enriched germanium detectors. However, two other radioactive isotopes that can lead to backgrounds at $Q_{\beta\beta}$, ^{68}Ge and ^{60}Co, are produced in germanium detectors through cosmogenic activation. To mitigate this background source, the MAJORANA collaboration carefully controlled and tracked the amount of time enriched detectors spent above ground. When detectors had to be transported between locations, air travel was avoided to prevent the increased cosmogenic activation that occurs at high altitudes.

The key material used for near-detector structural components and for the innermost layer of the DEMONSTRATOR's shielding was underground electroformed copper. Electroforming is the process of dissolving a metal, such as copper, in a chemical bath and electroplating it on a mandrel. The reconstituted metal contains a highly reduced concentration of impurities relative to the original material. For the DEMONSTRATOR's underground electroformed copper, nuggets of commercially-available high purity copper were used as the stock material for electroforming, resulting in a final product with $< 0.1\mu$Bq/kg of ^{232}Th and ^{238}U [49]. The copper was then machined underground to shape it into components suitable for use in the final experiment, such as detector mounts. Performing the electroforming and machining of copper parts in an underground cleanroom limited the introduction of cosmogenically-produced isotopes, such as ^{60}Co, during and after the slow electroforming process. This procedure resulted in a structural material that not only had good

mechanical and thermal properties but was also highly radiopure.

For components of the experiment that could not be constructed from UGEFCu, care was taken to select the most radiopure materials possible, such as using a clean plastic, pure polytetrafluoroethylene (PTFE), for insulating parts. Efforts were also made to minimize the mass of material present. For example, Axon' manufactured low-mass coaxial cables for use in the DEMONSTRATOR. The Axon' cables cables used to read out signals in the germanium detectors had an outer diameter of only 0.4 mm, and even the high voltage cables were limited to a 1.2 mm outer diameter, much smaller that typical commercial cables. Detector strings were constructed in a nitrogen-purged glove box within the underground cleanroom to limit the introduction of contamination at this stage.

Section 2.4: Energy Resolution

As was discussed in Sec. 1.8.2, good energy resolution also plays a role in limiting the number of background counts present in the $0\nu\beta\beta$ peak region. The DEMONSTRATOR's energy resolution of 2.52 ± 0.08 keV (0.12%) at the Q-value was the best of any $0\nu\beta\beta$ experiment to date [11]. Multiple aspects of the DEMONSTRATOR's hardware and analysis chain contributed to its ability to measure each event's energy with high precision. Key among these was the fact that germanium detectors have inherently good energy resolution due to the large number of charge carriers per keV of interaction energy, as described in Sec. 2.1.

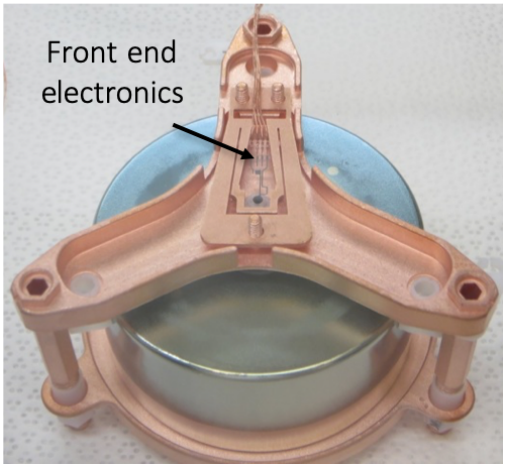


Figure 2.3: Photograph of a detector unit (upside down from its orientation in the DEMONSTRATOR) showing the position of the front end electronics (LMFE). The detector mount is made from UGEFCu.

The design of the electronics chain also played a role in the achieved energy resolution, since high electronic noise can lead to the degradation of energy resolution. The electronics chain by which the DEMONSTRATOR's signals were read out was divided into two stages, one located inside the cryostat and

one outside the DEMONSTRATOR's shielding [50]. The low-mass front end boards (LMFEs) that made up the first stage of the DEMONSTRATOR's electronics chain were located approximately 1 cm away from each detector's point contact, as shown in Fig. 2.3. This near-detector location minimized the length of cable required between the signal's generation and the first stage of amplification, improving the signal-to-noise ratio. The DEMONSTRATOR's LMFEs represented a major technological achievement, because the stringent radiopurity requirements for near-detector components imposed heavy constraints on the design of electronics immediately adjacent to the detectors. As with other components located inside the DEMONSTRATOR's shielding, the materials used in the LMFEs were carefully chosen and their masses limited to avoid introducing ^{232}Th and ^{238}U adjacent to the detectors.

Once a signal passed through the entire electronics chain, it was converted from an analog to digital pulse using GRETINA digitizers [51]. These 14-bit digitizers have a sampling rate of 100 megasamples per second. Each detector was connected to separate high and low gain digiziter channels. The low gain channels had dynamic ranges of up to ~10 MeV, while the high gain channels, which were used for the $0\nu\beta\beta$ analysis when possible, recorded higher resolution waveforms for events below ~3 MeV. The fast sampling rate and high resolution allowed for high precision energy measurements. They also preserved detailed information about the rising edge of waveforms that enabled the good pulse shape discrimination, to be discussed in the next section. A digitizer nonlinearity correction was applied to each waveform prior to extracting its energy [52].

Finally, refined signal processing techniques for energy estimation contributed to the achieved energy resolution of the DEMONSTRATOR. In principle, an event's energy can be determined simply by finding the waveform's maximum amplitude. However, in practice, effects like charge-trapping reduce the efficacy of this approach. When charge carriers are traveling through a germanium detector, they can become trapped in impurities in the detector bulk and only released on time scales longer than the waveform's data collection window, meaning they do not contribute to the measured signal. Because the magnitude of this effect is proportional to the path length that the charge carriers travel through the detector, it results in a drift-time-dependent energy degradation. To correct for this, a fixed-time pickoff approach was used [53]. First, a modified pole-zero correction was applied to each waveform. This pole-zero correction took into account both the decay constant due to the electronic response function and also the decay constant due to charge trapping. Then, each pole-zero corrected waveform was passed through a trapezoidal filter. The uncalibrated energy was determined not by the maximum of the filtered waveform but by its value a fixed time after

the start of the waveform's rise, effectively removing the drift time dependence. The application of the charge-trapping correction improved the DEMONSTRATOR's energy resolution by ~30% at $Q_{\beta\beta}$ [53].

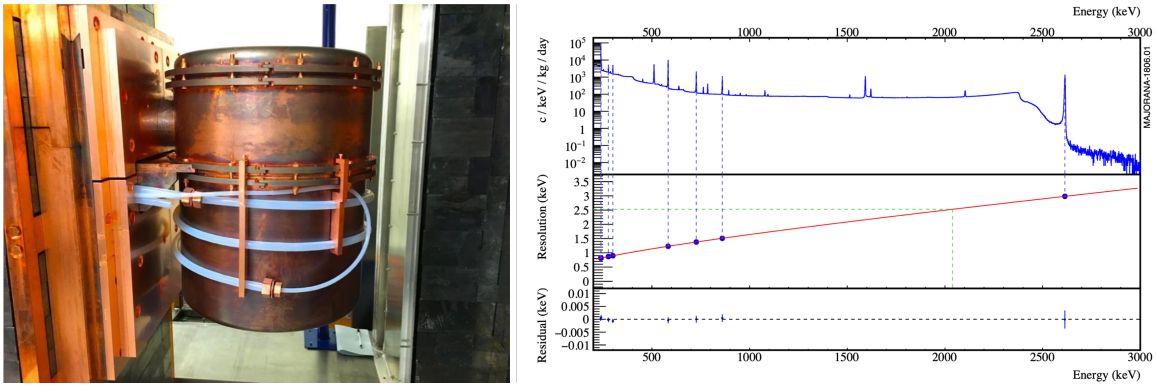


Figure 2.4: Left: Photo of the translucent calibration tube wrapped around one of the DEMONSTRATOR's cryostats. Right: Plot of the DEMONSTRATOR's energy resolution as a function of energy, as determined from calibration data (right)

The final step in the estimation of each waveform's energy was the application of calibration parameters, determined through ^{228}Th calibration runs. Each of the DEMONSTRATOR's copper cryostats was encircled by a helical calibration track with an insertable ^{228}Th line source, shown in Fig. 2.4. Every week, ~1-hr long calibration runs were taken with each module, and every few months, longer (~18-hr) runs were performed. In general, the two modules were calibrated separately, although early in data-taking concurrent calibrations of the modules were performed. Calibration parameters were determined by simultaneous fits of eight prominent gamma peaks between 238 and 2615 keV. Weekly calibrations allowed the time stability of each channel's gain, along with other parameters relevant to pulse shape discrimination, to be carefully monitored.

Section 2.5: Analysis-Based Background Mitigation Strategies

Many of the same qualities that enabled the DEMONSTRATOR's good energy resolution also contributed to its ability to efficiently reject backgrounds through pulse shape discrimination (PSD). The hardware-oriented background reduction techniques of shielding and radiopure material selection were supplemented with analysis-based tools, many of which depended on good pulse shape discrimination.

2.5.1: Data Quality Cuts

To be included in the main $0\nu\beta\beta$ analysis, runs and detectors were first required to first meet data quality standards. For example, runs where disruptive work to the shield was taking place were not included in the $0\nu\beta\beta$ analysis because of the additional noise they might include. Similarly, a detector that could not be properly calibrated would be excluded from the analysis. The analysis framework was flexible enough to

allow detectors, or even single digitization channels, to be excluded for some run ranges but included for others. This might be done if a detector had stable energy calibration parameters over a large number of weekly calibrations but then began to exhibit gain instability. For many of the more complex analysis cuts described later in this section, time instability in cut parameters led to some detectors being excluded from the $0\nu\beta\beta$ analysis for certain run ranges.

Data cleaning cuts were then used to eliminate pathological populations, such as pile-up waveforms and waveforms where the maximum voltage exceeded the digitizer's range. Events occurring in a module that was undergoing a liquid nitrogen fill were also removed because of the vibrational noise that occurs during a fill. Pulsers, which were used to monitor the experiment's dead time, were also removed from the data stream at this stage, as were events cut by the active muon veto described in Sec. 2.3.1.

2.5.2: Multiplicity and Multi-site Cuts

Since the path length of betas in germanium is smaller than the position resolution of germanium detectors, neutrinoless double beta decay is considered a single-site event. Monte Carlo simulations confirm that 90% of $0\nu\beta\beta$ decays in the Demonstrator would be single-site, taking into account X-ray excitations and bremsstrahlung [7]. In contrast, many backgrounds to $0\nu\beta\beta$ can deposit energy in multiple locations. For example, at the MeV-scale energies relevant to the Demonstrator, gammas have a high probability of Compton scattering, which can result in multiple interactions with active material. A multiplicity cut removes all events with energy deposits in multiple detectors, taking advantage of the Demonstrator's closely-packed detector arrays. The natural germanium detectors distributed throughout the two modules mainly served to increase the efficacy of this cut by increasing the amount of active material in the experiment. In many of the cases where detectors were rejected from the main analysis for data quality reasons, they were still fit for use in determining the multiplicity. These detectors were referred to as veto-only detectors for their role in rejecting multi-detector events.

The good pulse shape discrimination of point contact HPGe detectors allowed the identification of multiple-location events to be extended to include events where all interaction locations were within the same detector. As described in Sec. 2.1.1, the highly concentrated weighting potential of p-type point contact detectors leads to a charge signal characterized by long drift times followed by fast rise times. Because signals originating in different parts of the detector have significantly different drift times, energy deposits in different parts of the detector during the same event can be resolved. The multi-site cut, known as

AvsE, is based on the amplitude of an event's current pulse (A) compared to its energy (E) [7]. In a multi-site event, the fast rise from each energy deposit causes a distinguishable peak in the current signal. Since events of the same energy have the same total integrated current, the maximum current value at a given energy is smaller for multi-site events, where that current is divided between multiple peaks, as is shown in Fig. 2.5

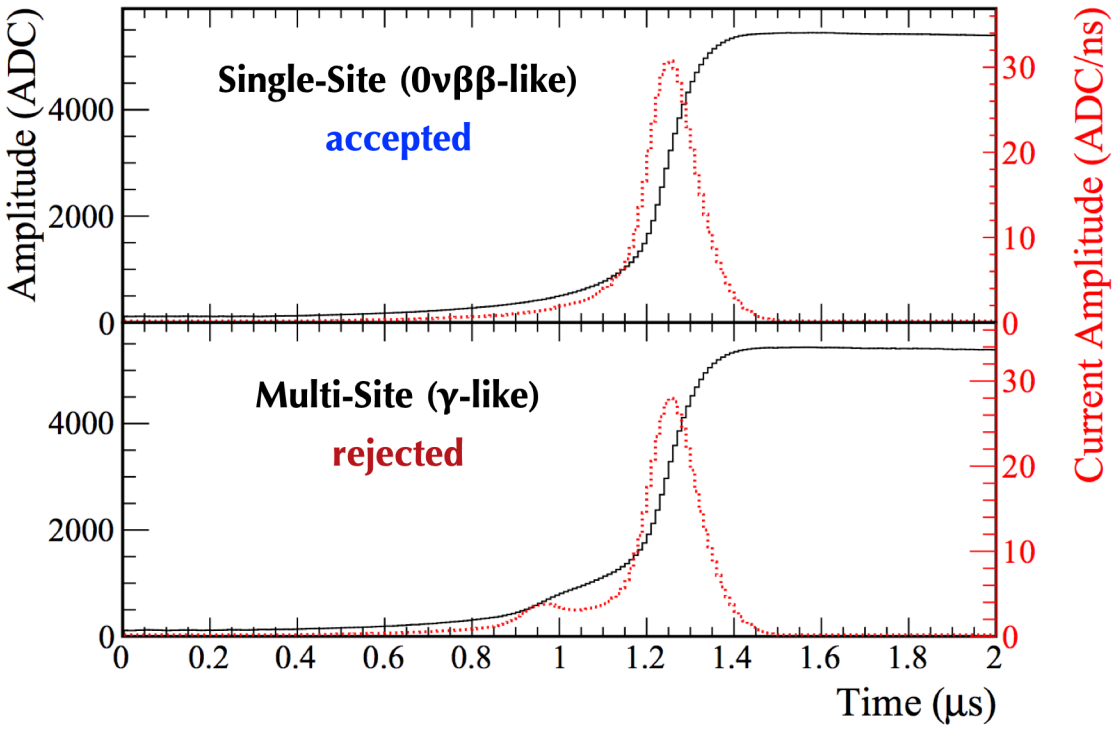


Figure 2.5: Comparison of the current and charge waveforms for a single-site and multi-site event [7]. For events of the same energy, multi-size events have a smaller maximum current amplitude, leading to a lower AvsE value.

Because ^{228}Th calibration data provided relatively pure populations of single site and multi-site events, calibrations were used for tuning and evaluating the performance of the AvsE parameter. Gamma rays with sufficient energy to interact with a germanium detector via pair production, such as the 2615 keV gamma from ^{208}Tl, result in spectral peaks that can be used to isolate single and multi-site events. Pair production occurs when a gamma ray is converted to an electron positron pair in the presence of an atomic nucleus [42]. The electron and positron quickly lose kinetic energy through scattering with the detector material, and then the positron annihilates with an electron in the germanium, producing two gammas at the electron rest mass of 511 keV. The double escape peak (DEP), which is located at 1592 keV for a 2615 keV gamma, occurs when both gammas escape the detector without interacting. The DEP predominately consists of single site events, because the distance that the positron travels in the detector before annihilating is small relative to

the detector's position resolution. The single escape peak (SEP) at 2103 keV peak is a result of one of the 511 keV gammas escaping the detector while the other deposits all of its energy via Compton scattering and the photoelectric effect. The single escape peak is dominated by multi-site events, since the initial pair production and the 511 keV gamma interaction occur at different locations within the detector.

Using dedicated ~12-18 hour ^{228}Th long calibration runs, the AvsE cut was set to accept 90% of events in the 1592 keV DEP. Since ^{56}Co has multiple DEPs near the $\beta\beta$-decay Q-value, a one-time ^{56}Co calibration was used to properly account for the changing width of the AvsE distribution as a function of energy to ensure a ~90% single-site acceptance at $Q_{\beta\beta}$. The AvsE cut parameter was also corrected to remove drift time dependence. After all corrections are applied, ~6% of events in the 2103 keV SEP survive a low AvsE cut, demonstrating how effectively the cut rejects multi-site gamma interactions [7]. The 2039 keV $\beta\beta$-decay Q-value is in the Compton continuum region below the 2615 keV peak, which consists of a combination of single and multi-site events. In this region, more than 50% of events in calibration data are cut by a low AvsE cut, showing that the AvsE cut can significantly reduce gamma backgrounds, although a population of single-site gammas that are indistinguishable from $\beta\beta$ decay events remains.

2.5.3: Delayed Charge Recovery Surface Alpha Cut

The DEMONSTRATOR's multiplicity and multi-site cuts effectively mitigate gamma-related backgrounds from radioactive decay chains. However, these decay chains also create alpha particles that can lead to backgrounds in the detector array. Since alphas are produced at fixed energies and have a short path length in germanium (tens of μm), alphas originating inside the detector bulk manifest as monoenergetic peaks, which do not pose a background to $0\nu\beta\beta$. Alphas from sources outside the detectors can be shielded by very small amounts of material. Even the ~1 mm lithiated n+ layer that covers most of the detector surface is too thick for alphas to penetrate. However, the ~0.1 μm passivated surface that covers the bottom surface of the detector and separates the n+ layer from the p+ point contact is susceptible to alpha backgrounds. As alphas pass through the passivated surface, a significant portion of their total charge becomes trapped and is re-released on a long time scale. This process results in energy-degraded alphas that no longer manifest as a distinct peak. 5.3 MeV ^{210}Po alphas from the ^{238}U decay chain lead to energy-degraded surface events that constitute an important background source for the DEMONSTRATOR in the $0\nu\beta\beta$ region of interest.

The slow release of charge trapped on the detector surface creates a signature in the waveform tail that can be used to distinguish surface alphas from bulk events. During data processing, a correction is applied

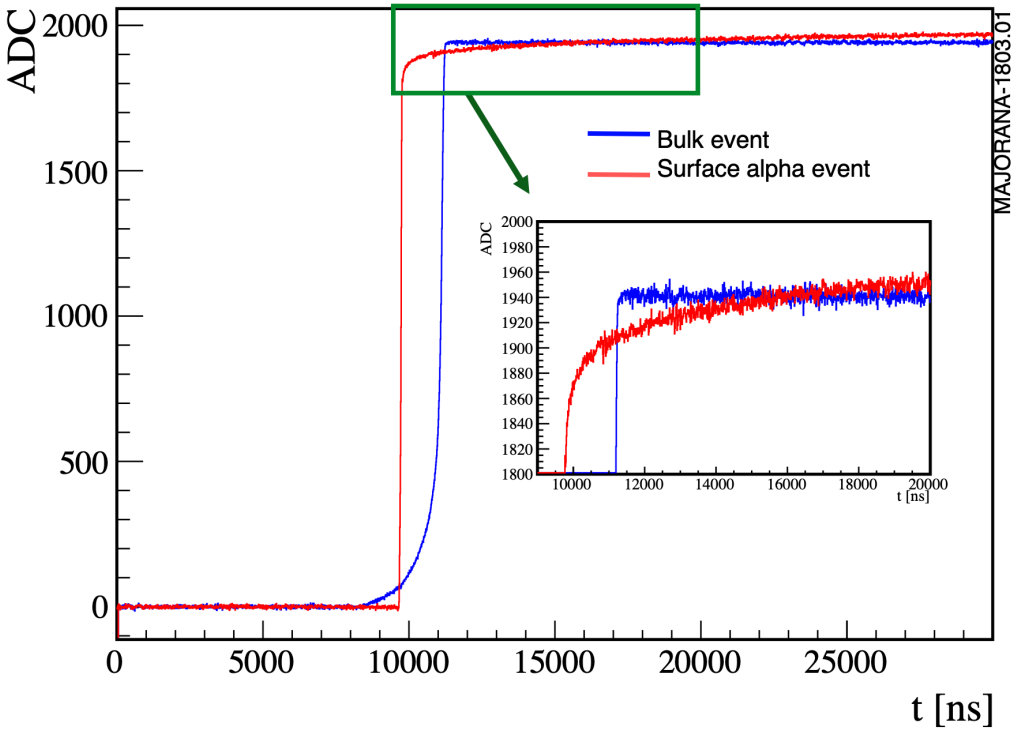


Figure 2.6: Comparison of waveforms for a typical bulk event (blue) and a surface alpha (red) after correcting for the decay constant of the detector's electronics chain. The positive slope of the surface alpha waveform's tail is due to the slow release of charge trapped on the detector surface [8].

to each waveform's tail to correct for the known decay constant due to the detector's electronics chain. Once this correction is applied, bulk events display flat tails following the fast rise in which the bulk charge is collected, as is shown in Fig. 2.6. In contrast, the tails of surface event waveforms have a positive slope, since trapped charges continue to be slowly released and collected over the entire digitization window of the waveform. The delayed charge recovery (DCR) parameter is a measure of waveform tail slope used to tag surface events [8].

The DCR parameter was tuned using ^{228}Th calibration data. The alpha population was negligible during calibration runs because the detector event rate was dominated by radiation from the calibration source, which was sufficiently shielded from the passivated surface to not contribute alphas. The DCR cut value was set to keep 99% of calibration events in the region surrounding the double beta decay Q-value from 2028 to 2050 keV, which contains bulk events from the 2615 kev peak's Compton continuum [11]. The cut was adjusted to remove drift time dependence and account for energy and time instabilities.

2.5.4: High AvsE Cut

Prior to the DEMONSTRATOR's final $0\nu\beta\beta$ data release, improvements in the stability of the AvsE parameter allowed a high AvsE cut to be added to the standard pulse shape analysis cuts. The high AvsE cut acts as a surface event cut that complements the DCR cut by targeting surface events close to or at the point contact, where the DCR cut loses discriminatory power [11]. Events occurring near the point contact exhibit atypically fast rise times. Since the current waveform is a measure of the slope of the charge waveform, waveforms with fast rise times have high AvsE values.

The impact of the high AvsE cut on detection efficiency was evaluated with ^{228}Th calibration data following the same procedure as was used for the DCR cut. At $Q_{\beta\beta}$, the survival fraction of bulk events was found to be 98% [11].

2.5.5: Late Charge Cut

An additional analysis cut was developed prior to the publication of the DEMONSTRATOR's final $0\nu\beta\beta$ decay half-life limit in order to improve background rejection [11]. The late charge (LQ) cut was initially introduced to target a population of multi-site events that the AvsE cut did not effectively tag [9]. Since events occurring near the point contact have atypically high AvsE values, multi-site events where the majority of the total energy is deposited near the point contact do not always exhibit the low AvsE values typically characteristic of multi-site events. To flag this category of waveforms, the LQ parameter calculates the area above the waveform during the final portion of its rise. In order to measure how much of the total charge is still uncollected late in the waveform's rise, the integral above the waveform is taken starting at the time point when the waveform reaches 80% of its maximum amplitude. The shaded area in the inset of Fig. 2.7 demonstrate the integrated LQ regions for a typical bulk waveform and a waveform with an atypically high LQ.

While the LQ cut was originally designed to supplement the AvsE cut in tagging multi-site waveforms, it was found to be particularly useful for identifying events with a slow component. Slow pulse events occur when energy is deposited in the detector's transition region, at the boundary between the n+ layer and the detector bulk. Unlike the n+ layer, the transition layer is not fully dead, but it is not subject to the high electric fields experienced by the majority of the detector. When energy is deposited in the transition layer, a fraction of the charge carriers slowly drift into the detector bulk and are subsequently read out at the point contact, while the remainder are unable to escape the transition region. The resulting events are

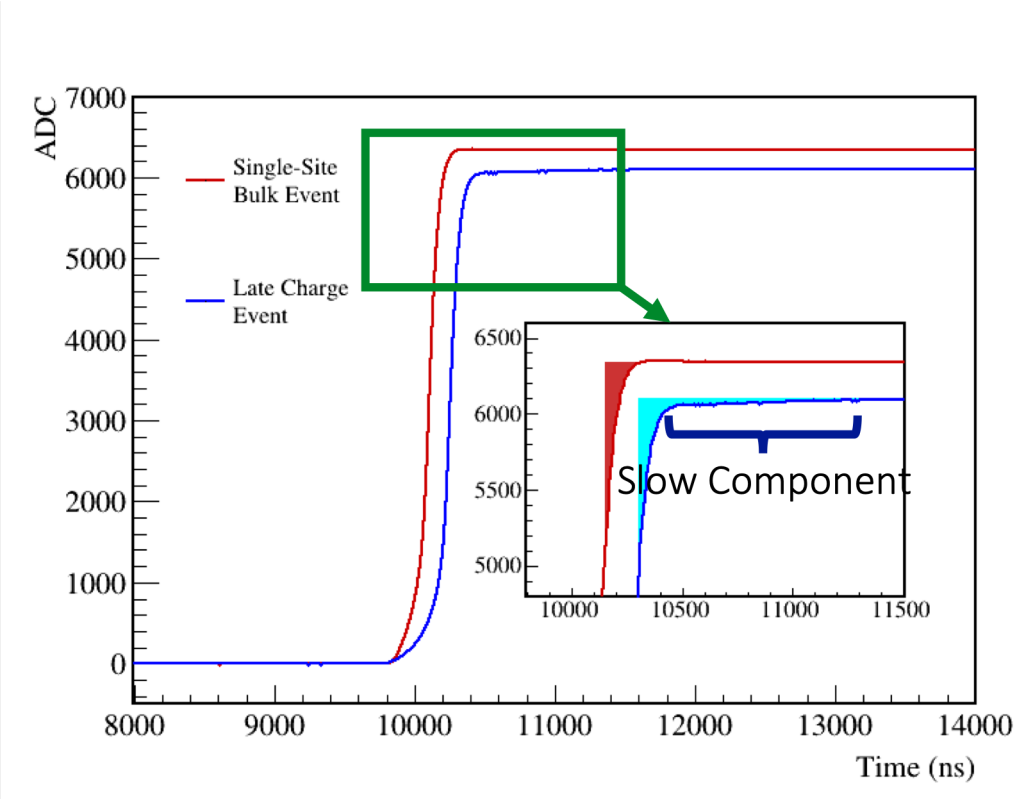


Figure 2.7: Comparison of the LQ calculation for a single-site bulk event (red) and an event with a slow component to its rise (blue). The shaded regions in the inset show the integrals computed to calculate LQ. For the waveform with a slow component, a larger amount of the total charge remains uncollected until the final 20% of the waveform's rise, resulting in a larger LQ value [9].

energy degraded and display a characteristic slow rise. If an event deposits some energy in the bulk and some in the transition region, it will have both a fast and slow component to its rise, resulting in a high LQ relative to bulk events. These events with slow components can be due to gamma-induced multi-site events where one energy deposit occurs in the transition layer, as well as to beta particles with sufficient energy to move beyond the fully dead detector region. This class of slow component events was more common in the DEMONSTRATOR than the near point-contact multi-site events that the LQ parameter was originally designed to target and was also a source of backgrounds in the $0\nu\beta\beta$ region of interest. Therefore, the LQ cut is primarily considered a surface event cut that complements the DCR cut.

For every detector and every weekly calibration, the LQ parameter was tuned and a correction was applied to remove dependence on drift time. The final $0\nu\beta\beta$ analysis used a 5σ cut on the drift-time-corrected LQ parameter. To determine LQ component of the detection efficiency at $Q_{\beta\beta}$, which was found to be 99.3%, the ^{228}Th DEP survival fraction was measured and an energy-dependent correction based on higher energy ^{56}Co calibration DEPs was applied [11]. The cut efficiency at lower energies was not well studied, since the

LQ parameter was targeted at rejecting backgrounds to $0\nu\beta\beta$.

Section 2.6: Blinding Scheme

Given that a high level of tuning of analysis parameters was required to optimize sensitivity, it was crucial to ensure that results were not inadvertently biased [54]. To this end, for the majority of its runtime, the DEMONSTRATOR employed a statistical blinding scheme. During periods where the blinding scheme was utilized, only 25% of physics runs were left open, while the other 75% of the data was made inaccessible to analysts. This was accomplished by alternating between 31-hour periods of open data-taking and 93-hour periods of blinded data-taking. These lengths were chosen to prevent the open cycles from lining up with any possible daily or weekly cycles in the lab environment. All calibration runs were left open in order to calibrate and monitor the stability of energy and PSD parameters. Once calibration data had been used to tune all necessary parameters, the $0\nu\beta\beta$ analysis was performed on open runs, and any necessary changes to analysis routines or parameters were implemented. A phased approach was then taken to unblinding, allowing data far from $Q_{\beta\beta}$ to be unblinded first in order to check for major anomalies before the energy region relevant to the $0\nu\beta\beta$ analysis was unblinded. During this process, some event populations, such as multi-detector events, remained blinded for other analyses. The DEMONSTRATOR's blinding strategy prevented decisions about the background reduction techniques employed from being influenced by their ultimate performance on the full low-background dataset.

Section 2.7: Timeline of the DEMONSTRATOR

MAJORANA ran with multiple different experimental configurations over the lifetime of the experiment. The DEMONSTRATOR's neutrinoless double-beta decay data was broken up into datasets, numbered DS0-DS8, based on these changes in experimental configuration. Two datasets, DS5 and DS6, were further divided into subdatasets based on smaller changes. Background counts, signal counts, exposures, and efficiencies were separated by subdataset during the calculation of the final $0\nu\beta\beta$ half-life limit in order to account for possible differences in background rates.

For approximately the first year of data-taking, only a single cryostat, Module 1, was deployed. Between June and October 2015, Module 1 ran with a partial shield, because the inner copper shield had not yet been installed. Data taken during this period of time, DS0, was considered commissioning data and had higher backgrounds. From December 2015 to July 2016, data were taken with Module 1 only but with the full shield installed. This period was divided into DS1 and DS2 based on a change to the data acquisition

system in May 2016 to test multisampling, which allowed longer waveforms to be acquired without loss of information from the waveform rising edge.

After DS2, Module 2 was deployed. Late August and September of 2016 was the first time data was collected with both modules in the shield. Unlike in later datasets, during this period, the two modules had independent data acquisition systems. DS3 and DS4 are comprised of data taken over this time span in Module 1 and Module 2 respectively. In October 2016, the data acquisition systems for the two modules were combined, marking the start of DS5. DS5 lasted until May 2017. However, it was later subdivided, mainly to take into account differences in electronic noise between different portions of DS5. DS5a was a high noise subdataset that consequently had worse energy resolution and pulse shape discrimination than other datasets. In January 2017, the DEMONSTRATOR's grounding scheme was improved, reducing noise levels and leading to the beginning of DS5b. While the change in grounding was being evaluated, no blinding was implemented. In March 2017, when sufficient data had been taken to confirm the improvements in electronics noise, the blinding scheme outlined in Sec. 2.6 was instated, marking the beginning of DS5c. In May 2017, the multisampling technique tested in DS2 was fully implemented. This began DS6, which lasted until November 2019. While DS6 was also subdivided, the differences between DS6a, b, and c were relatively minor and do not merit discussion here.

November 2019 marked the beginning of a period of more major change, undertaken in order to upgrade the DEMONSTRATOR's cables and connectors. Up to this point, problems such as high voltage breakdowns and failure of signal connections had led to a number of detectors not being operational. Some detectors had been biased down entirely, while others were operated below their ideal voltage and were only usable for veto purposes. These problems reflected the experimental challenges associated with operating a low-background experiment, where stringent radiopurity requirements had necessitated custom-built cables and connectors. These custom parts were designed to have lower masses and to use materials with lower radioactivities than commercially-available alternatives. However, the custom-designed parts also proved to be less robust and easier to damage during installation. As a part of the DEMONSTRATOR's goal of showing scalability, it was important to acquire evidence that these problems could be fixed, increasing the percentage of usable detector mass. The cable and connector upgrade of Module 2 did this. Following the upgrade, all Module 2 detectors were operational for the remainder of the experiment. The upgrade was also used as an opportunity to install 4 enriched ICPC detectors into the DEMONSTRATOR. To make space for the ICPCs, five enriched p-type point contact detectors were removed for testing in LAr in advance of LEGEND-200. Since cables and connectors

were only upgraded in Module 2, Module 1 continued acquiring usable $0\nu\beta\beta$ exposure while the upgrade was in progress. DS7, which spanned November 2019 to August 2020, represents another period of M1-only data, taken while Module 2 was outside the shield to be upgraded. DS8 consists of data from both modules after M2 had been reinstalled in the shield following the upgrade.

March 2021 marked the end of $0\nu\beta\beta$ data-taking. At this time, all enriched detectors were removed from the DEMONSTRATOR for installation in LEGEND-200. The remaining natural detectors were consolidated into a single cryostat, Module 2, to allow for further background studies. The data taken with natural detectors during this period is known as DS9. Finally, beginning in 2022, the module of natural detectors was rebuilt with tantalum disks installed between germanium detectors in order to pursue a new physics goal, the detection of the decay of the long-lived isomer ^{180m}Ta [55]. Data taken during this period, known as DS10, is not discussed in this work.

Section 2.8: $0\nu\beta\beta$ Half Life Lower Limit

The MAJORANA DEMONSTRATOR's final $0\nu\beta\beta$ half-life analysis was based on all enriched data from DS0-8, including both open and unblinded run ranges. The total enriched exposure for this period was 71.1 kg yrs, with an active exposure of 64.5 kg yrs [11]. The active exposure different from the total exposure because it accounted for the effect of dead time, including detector-dependent sources of dead time like the retrigger dead time following each digitized event. The active exposure calculation also excluded the dead region of each detector when determining the active mass. Both the fully dead n+ layer and the transition layer, where only partial charge collection occurs, were excluded from the active mass, since analysis cuts allow transition layer events to be rejected during analysis, leading to an active volume fraction of $92.0^{+1.3}_{-1.7}\%$ in the DEMONSTRATOR's PPCs [11]. Neither the total exposure nor the active exposure incorporated the isotopic abundance, the containment efficiency, or the efficiencies of analysis cuts, which were included separately in the half-life limit.

In 64.5 kg yrs of data, the MAJORANA DEMONSTRATOR found no evidence of $0\nu\beta\beta$ [11]. The combined enriched background index for all datasets, which will be discussed at length in Chapter 3, was found to be $16.6^{+1.4}_{-1.3} \times 10^{-3}$ cts/(FWHM kg yr), or, equivalently, $6.59^{+0.56}_{-0.53} \times 10^{-3}$ cts/(keV kg yr). Using an unbinned, extended profile likelihood analysis, the DEMONSTRATOR was able to set a 90% C.L. lower limit of 8.3×10^{25} years on the $0\nu\beta\beta$ half-life of ^{76}Ge. These results were consistent with the experiment's median sensitivity of 8.1×10^{25} years. Several alternative methods, including a Feldman-Cousins approach and Bayesian analyses

with two different priors, also found comparable results, ranging between 6.4×10^{25} and 1.1×10^{26} years. Assuming light neutrino exchange, the half-life limit from the primary analysis corresponds to an upper limit on the effective majorana neutrino mass, $m_{\beta\beta}$, of between 113 and 269 meV, depending on the choice of nuclear matrix elements [11].

Section 2.9: Looking Forward Towards LEGEND

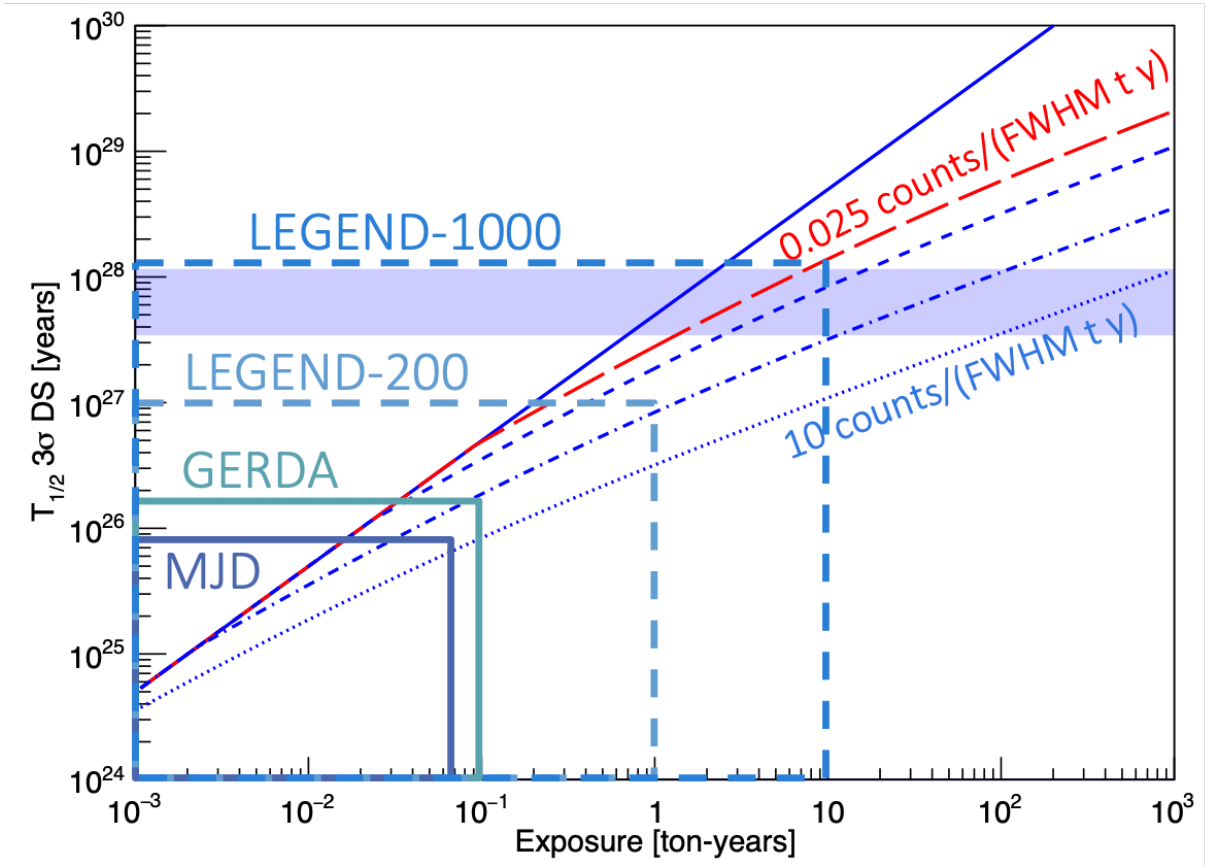


Figure 2.8: Half-life discovery sensitivity of ^{76}Ge $0\nu\beta\beta$ experiments as a function of exposure and background index. The blue diagonal line represents a completely background-free experiment. The lines for MAJORANA (MJD) and GERDA are based on their final exposures and measured background indices, while the LEGEND-200 and LEGEND-1000 lines indicate design goals. The blue shaded bar shows the range of half-lives associated with the inverted ordering region.

Given that one of the primary goals of the DEMONSTRATOR was to set the stage for a ton-scale experiment, its results should also be evaluated in the context of LEGEND. The final phase of LEGEND, LEGEND-1000, aims to attain a 3σ half-life discovery sensitivity of 1.3×10^{28} yr, covering the entire inverted ordering parameter space [10]. To do so, it will depend on the expertise developed by both MAJORANA and GERDA.

The first stage of LEGEND, LEGEND-200 began taking physics data in 2023 at Laboratori Nazionali del Gran Sasso (LNGS) in Italy, utilizing much of the infrastructure originally developed by GERDA. GERDA

very effectively proved that a germanium array could operate stably in liquid argon and that a liquid argon veto is a powerful tool for reducing backgrounds. GERDA's background index of 5.2×10^{-4} cnts/(keV kg yr) at $Q_{\beta\beta}$ is world-leading among $0\nu\beta\beta$ experiments [40]. However, LEGEND-200 and LEGEND-1000 have even more aggressive background goals of 2×10^{-4} and 1×10^{-5} cts/(keV kg yr) respectively [10]. These background levels correspond to remaining quasi-background-free for the intended exposure of each experiment with a 2.5 keV FWHM, as shown in Fig. 2.8.

For LEGEND-200 to attain a background index approximately a factor of 2.5 lower than GERDA's measured background, the background reduction techniques of GERDA are being supplemented by complementary techniques developed by MAJORANA, along with additional improvements such as increases to the efficiency of the liquid argon active veto system. In particular, some of the designs and low-background materials used for near-detector parts in the DEMONSTRATOR were repurposed for LEGEND-200. For example, LEGEND-200 is utilizing a front end electronics design similar to the design of the DEMONSTRATOR's LMFEs but modified for use in liquid argon. Locating this first stage preamplifier very close to the detector is important for achieving the best possible energy resolution, but it necessitates stringent radiopurity requirements. LEGEND-200 also builds on the DEMONSTRATOR's technique of machining structural components from underground electroformed copper. The excellent radiopurity of underground electroformed copper is particularly necessary for near-detector components because a liquid argon active veto is most efficient for components far from the detectors, where the probability of ionizing radiation interacting in the LAr is higher. Work towards evaluating the performance of these aspects of MAJORANA has helped shape LEGEND-200's approach to reducing backgrounds. For LEGEND-1000 to ultimately attain background levels that are a factor of 50 lower that GERDA's, GERDA and MAJORANA's strategies will be supplemented by additional measures, such as the replacement of all low mass point contact detectors with higher mass ICPCs and the replacement of the liquid argon in the active shield with underground-sourced argon. A clear path to achieving LEGEND-1000's background goals is more fully outlined in the experiment's pre-conceptual design report [10].

Fig. 2.8 demonstrates how important low backgrounds are to discovery sensitivity, but the benefit of a quasi-background-free spectrum to making an unambiguous detection can be more powerfully illustrated by Fig. 2.9. This figure shows a simulated spectrum comparing LEGEND-1000's expected background rate to a $0\nu\beta\beta$ signal with a 10^{28} yr half-life. Even with a small number of signal counts, a quasi-background-free experiment allows for a signal that is convincingly distinguishable from background. The ability to achieve

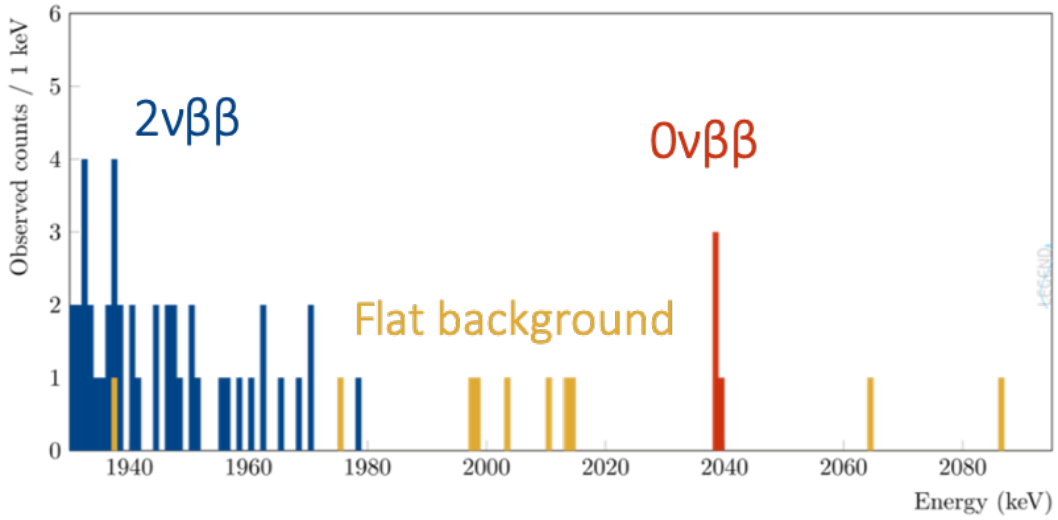


Figure 2.9: Simulated LEGEND-1000 post-cut spectrum after ten years of data taking, assuming a $0\nu\beta\beta$ half-life of 10^{28} years. Counts due to $2\nu\beta\beta$ are separated from other $0\nu\beta\beta$ backgrounds to demonstrate that the contribution from $2\nu\beta\beta$ to the background rate at $Q_{\beta\beta}$ is negligible given the energy resolution of germanium detectors. Figure adapted from [10]

a quasi-background-free ton-scale $0\nu\beta\beta$ experiment is one of the distinctive advantages of germanium experiments. The centrality of low backgrounds to the germanium $0\nu\beta\beta$ program adds to the importance of understanding the backgrounds of past experiments such as the DEMONSTRATOR.

CHAPTER 3: Study of Observed Backgrounds

Section 3.1: Overview of Observed Backgrounds

The enriched energy spectrum from the MAJORANA DEMONSTRATOR's final $0\nu\beta\beta$ analysis, shown in Fig. 3.1, effectively illustrates the backgrounds observed by the DEMONSTRATOR both before and after analysis cuts. The dominant visible feature is the $2\nu\beta\beta$ spectrum. The $2\nu\beta\beta$ rate in a $0\nu\beta\beta$ experiment cannot be reduced, but the DEMONSTRATOR's superb energy resolution ensured that it had a negligible contribution to the background level at $Q_{\beta\beta}$. The fact that $2\nu\beta\beta$ accounted for nearly all post-cut counts over a large energy region indicates how successful the DEMONSTRATOR was at achieving low backgrounds.

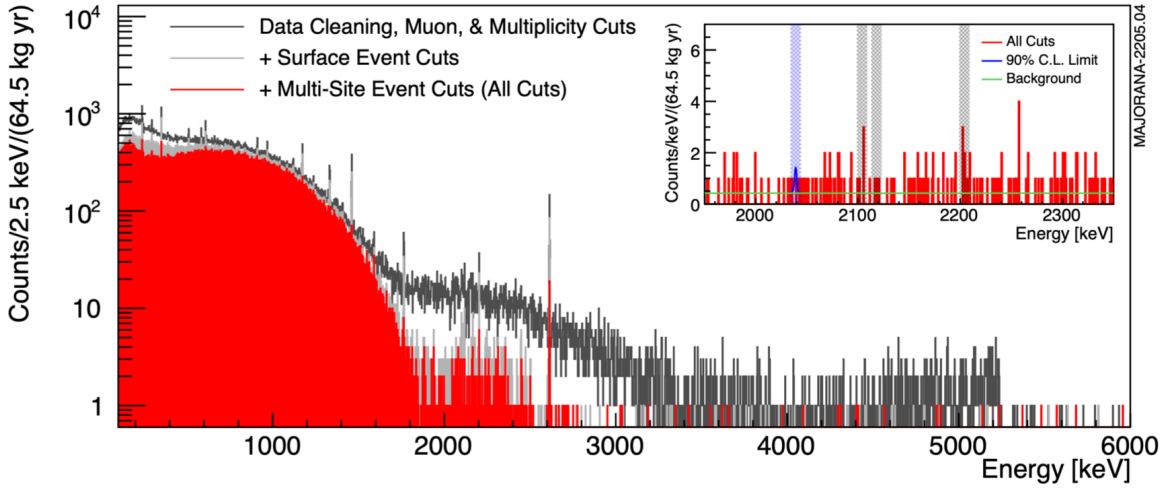


Figure 3.1: Energy spectrum in enriched detectors using final enriched active exposure of 64.5 kg-yrs [11]. The impact of surface event and multi-site cuts is shown.

The background rate of interest to the $0\nu\beta\beta$ search was the measured background after cuts in a 10 keV signal region around the $\beta\beta$ decay Q-value of 2039 keV. Since the DEMONSTRATOR's background spectrum near $Q_{\beta\beta}$ was relatively flat after analysis cuts, the rate at $Q_{\beta\beta}$ could be estimated by measuring the rate in a 360 keV background estimation window (BEW), shown in Fig. 3.2. The background estimation window included the region between 1950 and 2350 keV, excluding the 10 keV signal region and three other 10 keV regions at the locations of known gamma peaks. This approach enabled the DEMONSTRATOR's $0\nu\beta\beta$ half-life

limit to be independent of background modeling efforts, relying only on the assumption of a flat background.

In the background estimation window, the dominant background prior to cuts was energy-degraded surface alphas. The majority of events near $Q_{\beta\beta}$ were removed by surface alpha cuts, as can be seen by comparing the dark and light gray spectra in Fig. 3.1. The surface event cuts applied in this plot include the DCR, LQ, and high AvsE cuts discussed in Sections 2.5.3-2.5.5. The efficacy of these surface event cuts at removing alphas can be seen by examining the region above the 2615 keV ^{232}Th gamma peak in Fig. 3.1. Above 2615 keV, the DEMONSTRATOR's event rate after data cleaning and muon cuts was almost exclusively due to energy-degraded alphas. The very small number of counts above 2615 keV after cuts suggests that the surface event cuts were highly effective at removing surface alpha backgrounds. Because the surface event cuts were highly efficient, the remaining events in the background estimation window following cuts were mostly due to gammas. The multi-site cut effectively reduces gamma backgrounds, as shown by the sizable reduction in gamma peaks in the spectrum after cuts. However, single-site gamma events cannot be distinguished from signal events, so a significant fraction of gamma backgrounds could not be removed by analysis cuts.

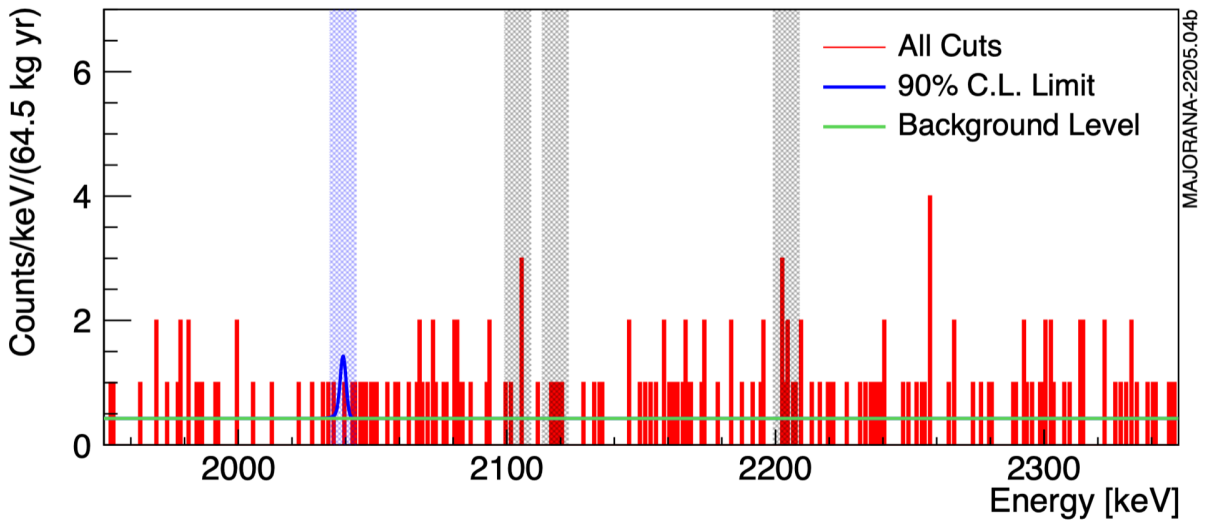


Figure 3.2: Event distribution in the 360 keV background estimation window after all analysis cuts [11]. The 10 keV regions around $Q_{\beta\beta}$ (blue) and three known gamma peaks (gray) were omitted when calculating the background index.

The DEMONSTRATOR's background index in the 360 keV background estimation window for enriched detectors was measured to be $16.6^{+1.4}_{-1.3} \times 10^{-3}$ cts/(FWHM kg yr), which is equivalent to $6.59^{+0.56}_{-0.53} \times 10^{-3}$ cts/(keV kg yr). In the low background configuration that excludes 1.2 kg yrs of commissioning data

taken prior to the full shield installation, the background index was $15.7^{+1.4}_{-1.3} \times 10^{-3}$ cts/(FWHM kg yr), or $6.23^{+0.55}_{-0.52} \times 10^{-3}$ c/(keV kg yr). This background index is one of the lowest of all backgrounds achieved in $0\nu\beta\beta$ experiments. However, it is significantly in excess of the assay-based background projection of $(2.9 \pm 0.1) \times 10^{-3}$ cts/(FWHM kg yr), or $(1.16 \pm 0.04) \times 10^{-3}$ cts/(keV kg yr) [56].

The assay-based background projection was based on a well-validated assay model, which incorporated assays performed prior to the construction of the experiment [49] as well as supplementary assay measurements taken following the discovery of the background excess. For each component, assay measurements were combined with efficiencies for detecting events in the background estimation window after pulse shape analysis cuts, found using simulations. The simulation framework used is described in more detail in Section 3.4. Simulations were updated following the beginning of data-taking in order to reflect changes between the DEMONSTRATOR's original design and the as-built experiment. A Monte-Carlo based error propagation technique was used to properly account for assay uncertainties and statistical uncertainties in the simulations [56]. The discrepancy between the assay-based background index and the DEMONSTRATOR's measurement motivates a data-driven approach to locating the background excess.

Multiple ^{232}Th peaks are more prominent in the data than was predicted by the assay-based background model. This conclusion, along with the shape of the post-cut spectrum surrounding $Q_{\beta\beta}$, indicates that the excess in the background estimation window can be attributed to Compton continuum events from the ^{232}Th decay chain. Attempting to determine the experimental component where the excess originated requires a more detailed analysis, which will follow in Sec. 3.3 and throughout the remainder of this work.

Section 3.2: Goals of Background Modeling

The presence of an unexplained background excess in the MAJORANA DEMONSTRATOR's data offers one but not the only compelling motivation for the development of a detailed background model. This section describes the three primary goals that the background model seeks to address and the significance of each.

3.2.1: Locating Background Excess

Determining what component or components are responsible for the excess ^{232}Th background described in the previous section matters not just for MAJORANA but also for LEGEND. Given that one of MAJORANA's principal functions was to demonstrate the feasibility of scaling up to a ton-scale experiment, assessing the reason MAJORANA did not meet its background budget has been of high importance. As was discussed in Sec. 2.9, LEGEND-200 seeks to achieve a lower background index than GERDA partially by adapting several

of the designs and materials that MAJORANA used for its near-detector components, such as underground electroformed copper and low mass front end electronics. This strategy means that it was imperative to understand whether the DEMONSTRATOR's background predominately originated in near-detector components. A background source originating in a component external to the cryostat or above the coldplate, such as the lead shield or crossarm, would not pose a problem for LEGEND, since LEGEND's design does not utilize these classes of components. Disentangling how much individual far components contribute to the background index would be of mostly academic interest, but determining whether near-detector components contributed significant to the overall background rate is relevant to the $0\nu\beta\beta$ community.

The background modeling procedures described in this work build on previous efforts by M. Buuck [57] and T. Gilliss [58] using a portion of the DEMONSTRATOR's exposure. That work came to the preliminary conclusion that the ^{232}Th background excess in data was not dominated by near-detector components. The follow-up efforts which will be described below seek to validate that finding and place more quantitative limits on how much radioactivity can be present in near-detector components by using refined techniques and taking advantage of the DEMONSTRATOR's full exposure.

3.2.2: Precision $2\nu\beta\beta$ Measurement

While data-driven background modeling of the MAJORANA DEMONSTRATOR plays an important role in setting the stage for future experiments, it can also be directly used in rare event searches or the measurements of physical processes occurring in the DEMONSTRATOR itself. In particular, a precision measurement of the two-neutrino double-beta decay half-life in ^{76}Ge relies heavily on background modeling. Since the two-neutrino double-beta decay signal is distributed across a wide-range of energies, a comprehensive background model with well-quantified uncertainties is necessary to determining what fraction of the total event rate can be attributed to $2\nu\beta\beta$.

This measurement has previously been conducted by other experiments, most recently GERDA [59]. A measurement using the DEMONSTRATOR's full exposure would provide a cross check on the GERDA $2\nu\beta\beta$ half-life using an experimental setup with different backgrounds. Below ~565 keV, GERDA's background spectrum was dominated by beta decays of the cosmogenic isotope ^{39}Ar. Since the DEMONSTRATOR did not utilize a liquid argon shield, $2\nu\beta\beta$ was the largest contributor to its event rate over a greater energy range.

3.2.3: Quantitative Background Model for Additional Analyses

A quantitative background model is also useful in searches for new physics using the DEMONSTRATOR's data. One example of this is a search for distortions in the $2\nu\beta\beta$ spectral shape relative to the theoretical prediction. Various beyond the Standard Model processes could result in this type of distortion. A few include $0\nu\beta\beta$ with the emission of one or more additional particles known as Majorons [60], $2\nu\beta\beta$ modified by the presence of Lorenz violation [61], and $2\nu\beta\beta$ with the emission of sterile neutrinos [62, 63]. The DEMONSTRATOR's data could be used to place limits on these Standard Model-violating processes by comparing the performance of fits with different simulated $2\nu\beta\beta$ spectra generated from different theoretical models, as has been done in other double-β decay experiments, such as GERDA [64]. A search for shape distortions in the $2\nu\beta\beta$ spectrum would also benefit from the DEMONSTRATOR's reduced backgrounds at energies below 500 keV, enabled by the lack of ^{39}Ar. Performing this type of analysis is beyond the scope of this work, but any such analysis of the DEMONSTRATOR's data will rely on the computational tools developed for background model fitting described in this work.

Section 3.3: Spatial Non-Uniformity of Backgrounds

The first step towards meeting the goals of background modeling was to perform a more detailed study of the DEMONSTRATOR's background spectrum, in particular how it varies between detectors. Although assay results and procedures for parts handling assumed that background rates would be fairly uniform across the two modules and their components, this was not born out by the data. The DEMONSTRATOR's measured background in the $0\nu\beta\beta$ background estimation window was significantly higher in Module 1 than in Module 2. When analyzed with the DEMONSTRATOR's full exposure, the Module 1 background index for enriched detectors after analysis cuts was $(7.38 \pm 0.71) \times 10^{-3}$ cnts/(keV kg yr), compared to a background index of $3.33^{+0.75}_{-0.67} \times 10^{-3}$ cnts/(keV kg yr) in Module 2.

To analyze the spatial distribution of backgrounds further, background rates were compared between individual detectors. In this work, detectors are referred to based on their location within the experiment. The first part of the detector name specifies which cryostat the detector was in, Module 1 (C1) or Module 2 (C2). The next part of the name gives the string position (P1-P7). P1 indicated the center position, and the rest of the string labels proceed counterclockwise when viewed from above, meaning P2 and P7 are next to one another. Finally, the end of the name indicates the detector's position within the string (D1-D5), with 1 being closest to the top and 5 nearest the bottom. C1P2D1, for example, indicates the top detector in string

2 of Module 1.

While most sources of background in Module 2 appear to have been relatively uniformly distributed among detectors, the distribution of Module 1 backgrounds points to a localized hot region. In particular, two natural detectors located adjacent to the Module 1 crossarm, C1P2D1 and C1P7D1, displayed much higher rates than other detectors. Since these detectors were not enriched, their high background rates did not directly contribute to the higher M1 background index quoted above. However, the enriched detectors located closest to these natural detectors also had elevated rates in the background estimation window, though the difference was much less pronounced. The three enriched detectors with the highest background indexes were C1P1D2, C1P2D2, and C1P7D2, the three second-row detectors located closest to the high rate natural detectors. If these three detectors were removed, the DEMONSTRATOR's background index for the full exposure would be reduced from 6.59×10^{-3} cnts/(keV kg yr) to 4.91×10^{-3} cnts/(keV kg yr) [9]. Nevertheless, excluding these detectors from the DEMONSTRATOR's $0\nu\beta\beta$ analysis would only lead to a small improvement in sensitivity (estimated at ~2%) [9].

The high backgrounds in the natural detectors C1P2D1 and C1P7D1 are evident from their high integrated count rate over a large energy range (between 100 and 3000 keV) as well as from their increased strength in prominent ^{232}Th decay chain peaks. The detector distribution of events in the 2615 keV ^{208}Tl peak is shown in Fig. 3.3. The 238 keV gamma from ^{212}Pb is also of particular interest. Because this gamma has a relatively low energy, it can be strongly attenuated by even a small amount of shielding material. Fig. 3.4a shows that C1P7D1 had a much more prominent 238 keV peak than the other Module 1 natural detectors, which were located in string 4, the string furthest away from the crossarm. While C1P2D1 also showed some hints of an elevated 238 keV peak rate, this evidence is weaker due to the detector's lower statistics, as shown in Fig. 3.4b. This detector was biased down early in DS5, so it had much less exposure than C1P7D1. Although C1P2D1's 238 keV peak alone is not statistically significant enough to conclusively demonstrate that the detector's rate excess was dominated by ^{232}Th decay chain events, Fig. 3.5b shows additional evidence for this conclusion from other peaks in the ^{232}Th decay chain. In particular, the ^{228}Ac peaks at 911 and 969 keV are visibly in excess of background, although the statistics in each peak are still low. At the level of statistics present in C1P2D1, the presence of an overall rate excess due to ^{232}Th is clear, but the relative intensities of the ^{232}Th peaks are not.

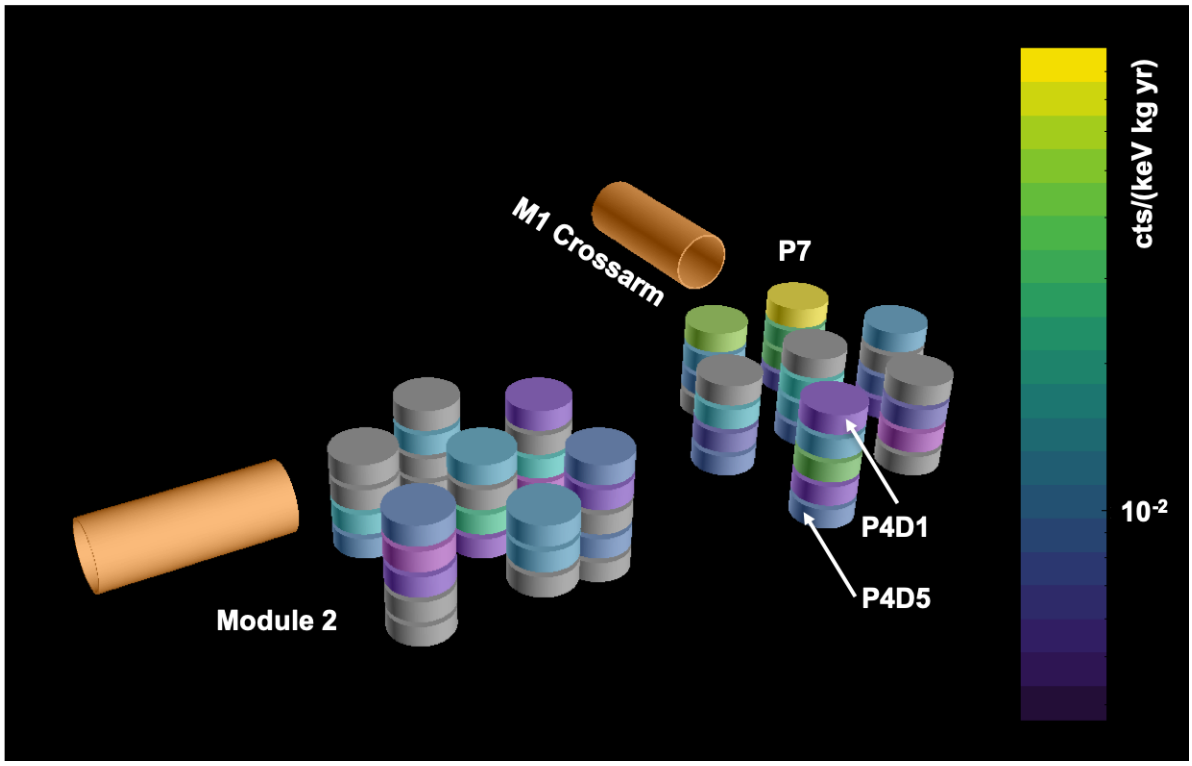
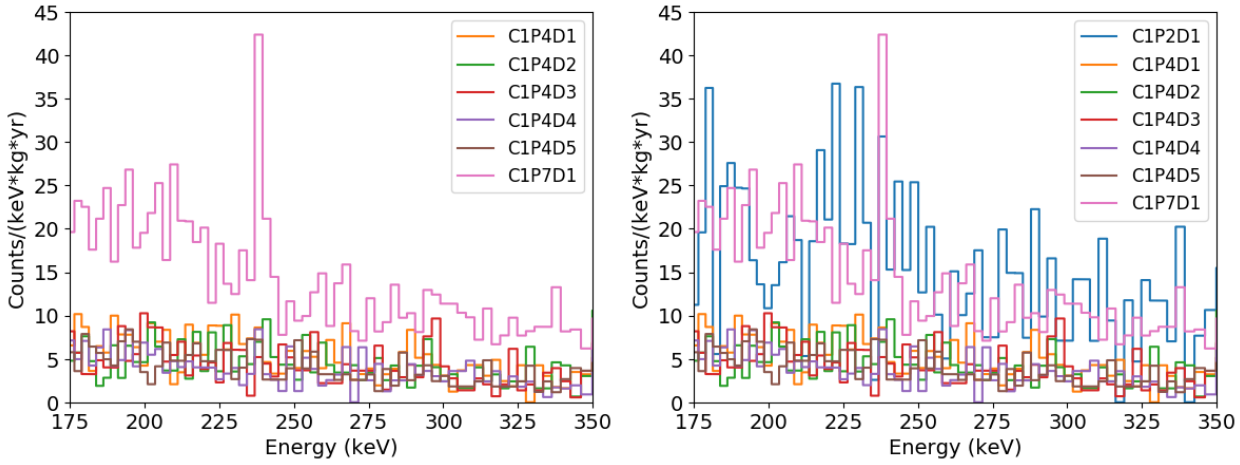


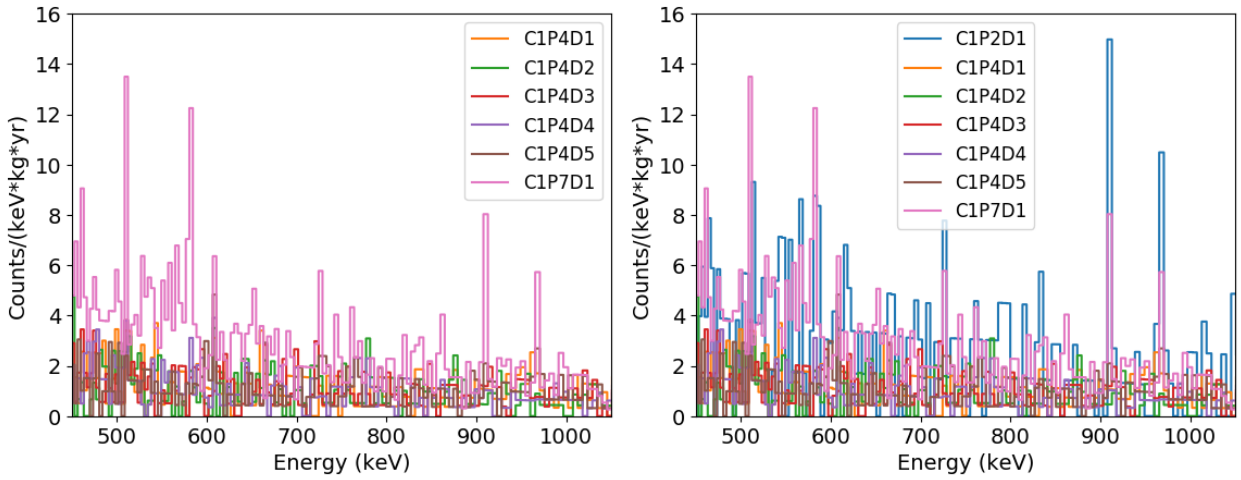
Figure 3.3: Heatmap of events in the 2615 keV ^{232}Th peak in open datasets prior to the upgrade (DS1-6c). Gray detectors are detectors that were biased down or veto-only for all included datasets, leading them to be excluded from this analysis. An excess of 2615 keV events is observed in detectors near the M1 crossarm, in particular C1P7D1. The placement of the other M1 natural detectors in string 4 is also indicated.



(a) Excluding low statistics detector, C1P2D1

(b) Including C1P2D1

Figure 3.4: Comparison of the 238 keV ^{212}Pb peak from the ^{238}Th decay chain in data from Module 1 natural detectors in DS1-6c open.



(a) Excluding low statistics detector, C1P2D1

(b) Including C1P2D1

Figure 3.5: Comparison of multiple peaks from the ^{232}Th decay chain in data from Module 1 natural detectors in DS1-6c open.

The backgrounds excess observed in C1P2D1 and C1P7D1 was an effect of the detectors' position in the array, not something inherently anomalous about the two detectors. This point was illustrated clearly by comparing data taken in the standard configuration to data taken when the detectors were operated in a different position. In early 2021, all enriched detectors were removed from the DEMONSTRATOR for use in LEGEND-200, leading to the consolidation of the natural detectors in Module 2. The two detectors that were previously located next to the Module 1 crossarm were moved to the top of string 6 in Module 2, and the DEMONSTRATOR operated with only one cryostat in the shield. Fig. 3.6 compares detectors' integrated count rates between 100 and 3000 keV between these two configurations. In the figure, the two detectors originally located near the crossarm are boxed in maroon. In the original configuration, these detectors are clear outliers, displaying integrated count rates that are approximately a factor of two higher than the integrated count rates of other natural detector in either module. In contrast, following the rearrangement of detectors, these two detectors showed integrated count rates that were entirely in line with other detectors, pointing to a Module 1 background source in the general vicinity of the crossarm but not directly tied to either of the high rate detectors.

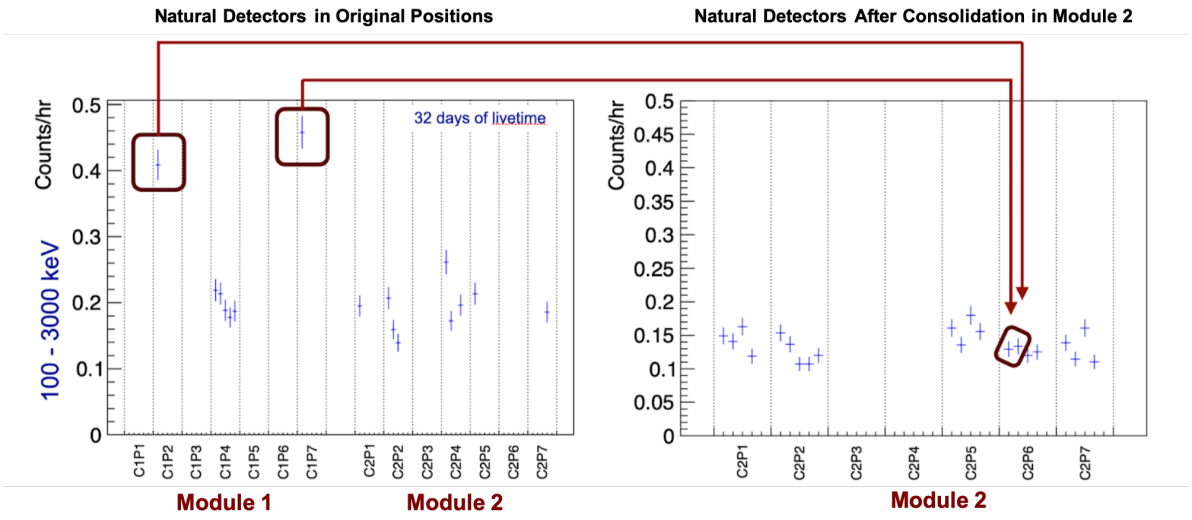


Figure 3.6: Comparison of the all natural detectors' integrated count rates from 100 to 3000 keV between two experimental configurations. Left: For the majority of data-taking, two Module 1 detectors located adjacent to the crossarm displayed anomalously high rates. These detectors are boxed in maroon. Right: After the removal of enriched detectors, the natural detectors were consolidated in Module 2. The new locations of the two previously hot detectors are again shown in maroon, demonstrating that in their new position, their rates were comparable to other natural detectors. Figure courtesy of V. Guiseppe

Characterizing the background excess more precisely and addressing the remaining background model goals requires comparing data to simulations. Chapter 4 will discuss the primary method of comprehensively performing such a comparison, spectral fitting. A number of simpler complementary methods incorporate information not fully utilized during spectral fitting and aid in interpreting fitting results. The remainder of this chapter will describe how the simulations are generated and will detail the results of comparisons performed between simulations and data.

Section 3.4: Simulations

The DEMONSTRATOR's background data spectrum consists of contributions from multiple different decay chains and cosmogenic isotopes, each of which can originate in many different components of the experiment. Each source results in a slightly different signature in the detector array, and simulations capture these variations. A simulated energy spectrum contains information about how many counts would be detected at each energy due to a single primary from a given decay chain in a specified part of the experiment. Spectral shape and overall detection efficiency both vary depending on where a background originates, making simulations a powerful tool for locating background sources. Drawing accurate conclusions from simulations requires that the experimental apparatus be modeled with high fidelity and that the details of detector response be carefully applied. The procedure for generating high quality simulations is described

in this section.

3.4.1: MaGe

Simulations of the Demonstrator were performed using the Geant4-based package MaGe, developed in collaboration between Majorana and GERDA [65–67]. Multiple versions of the experiment's as-built geometry were created in MaGe to account for major configuration changes in the experiment. A rendering of one of these configurations is shown in Fig. 3.7. When a simulation was performed, primaries were generally uniformly distributed throughout a component or group of components. After a primary is generated, it proceeds through its decay chain according to known branching ratios, producing ionizing radiation such as betas and de-excitation gammas. The particles produced during a decay are then transported through the simulated experimental space using Monte Carlo methods. At each step through a material, a particle has some probability of undergoing an interaction, such as absorption or scattering, determined by input physics lists containing information about interaction cross sections.

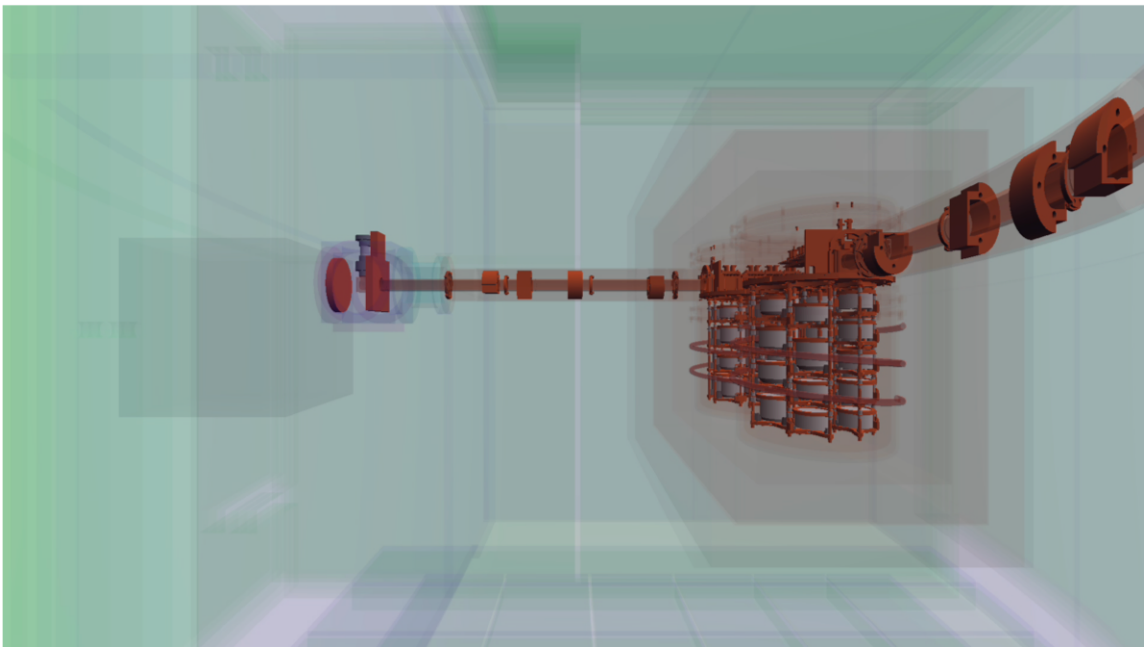


Figure 3.7: Rendering of the Demonstrator's simulated geometry in MaGe using raytracer. The graded shield and cryostat are made semi-transparent to reveal the detector arrays. Courtesy of M. Buuck and T. Caldwell.

If the simulated particle eventually reaches and interacts with one of the germanium detectors, the amount of energy deposited is recorded. A single primary can lead to multiple energy deposits in one or more detectors. Additional information about each event, such as the position of energy deposits within the detector, is also stored for use in event post-processing.

3.4.2: Component Groupings and Decay Chains

Approximately 4000 parts are included in the DEMONSTRATOR's simulated geometry. These parts are combined into ~27 groups of components that are considered together when generating primaries. In previous background modeling work performed by M. Buuck and T. Gilliss, components were grouped based on handling history, so that parts expected to have similar activity densities were considered together. The majority of components were simulated together across both cryostats, since it was expected that in most cases a class of parts, such as the front end electronics, would be identical between the two modules. The simulations used in the current work instead grouped components based on location. This allows components with very similar detection efficiency spectra to be grouped together. Module 1 components are now separated from Module 2 components, since the evidence from data suggests that an assumption of uniformity between the two modules is not justified.

Up to 9 decay chains or isotopes were simulated for each component group: ^{76}Ge ($2\nu\beta\beta$), ^{232}Th, ^{238}U, ^{40}K, ^{60}Co, ^{210}Pb, ^{68}Ge, ^{222}Rn, ^{57}Co, and ^{54}Mn. Some, like ^{232}Th, were simulated for every component group, while others, like ^{210}Pb, were only relevant for a small number of components. Decay chain were simulated in multiple segments, each of which is expected to be in secular equilibrium. In this work, these segments were later combined using the assumption that the full chain is in secular equilibrium.

3.4.3: Simulation Post-processing

After the simulations were completed, they were post-processed using the mage postproc package, where detector response effects were incorporated on a detector-specific basis. The detector response parameters used to apply these effects to the simulations were calculated from ^{228}Th calibration data and other supplementary sources, such as energy loss measurements performed by the vendor. More details about determining the values and uncertainties of post-processing parameters will be discussed in Chapter 5 in the context of determining the systematic uncertainty of the background model. A set of post-processed simulations was produced for each group of datasets that are considered together during background studies. Datasets were mainly grouped based on experimental configuration. The exception is that DS7 was post-processed separately from DS1-2, despite having the same configuration, because the datasets were significantly separated in time. Not grouping these datasets together during background studies allowed us to better take into account backgrounds that vary with time, such as cosmogenic isotopes with half-lives less than the runtime of the experiment. Post-processing parameters were determined separately for each dataset

group.

One of the principal effects introduced during post-processing is energy resolution. During MaGe simulations, the deposited energy in a detector is recorded with perfect accuracy, but this is not physically realistic. Simultaneous fits of multiple ^{228}Th peaks are performed to each detector's calibration energy spectrum to determine the detector's peak shape parameters, such as the full width half max as a function of energy. These peak shape parameters can then be used to randomly modify each simulated energy deposit to reflect the effect of instrumental response on the measured energy spectrum.

Another important effect included during this stage is the dead layer profile for each detector. Dead layer profile parameters characterize the fraction of the detector mass that is fully active and the impact of surface effects on the shape of the measured energy spectrum. Energy deposits that occur in the transition region, where the lithiated n+ surface meets the detector bulk, exhibit partial charge collection, meaning that only a portion of the electrons and holes freed by the interaction diffuse into the detector bulk to be read out [68]. Transition layer events are energy degraded, so the size and shape of the transition layer impacts the low energy tails of gamma peaks and the shape of the background spectrum at low energy. Although low energy transition events can be tagged due to their characteristic slow rise time, there is not a high efficiency analysis cut for removing transition layer events over a wide range of energies. To enable accurate comparisons between simulations and data, the simulations must be able to reproduce the impacts of transition events on the spectral shape. In addition, accurately accounting for the overall size of dead layer is important for correctly determining the $2\nu\beta\beta$ specific activity. The dead layer model applied to simulations assumes that the partial charge collection region of the detector consists of a portion where charge collection efficiency varies exponentially as a function of depth from the detector surface and a portion where it varies linearly [57, 68].

Tagging multi-detector and multi-site events is another process completed during simulation post-processing. To tag multi-detector events, energy deposits that occur in a short time window are grouped together into events. This procedure, which mimics the event building that occurs during the processing of the Demonstrator's data, determines the number of simultaneous detector hits in a event, known as the event's multiplicity. The procedure for tagging multi-site events in a single detector is more nuanced and requires tuning a parameter called the dt heuristic, which is intended to emulate the AvsE cut used on data. A clustering algorithm is applied to simulated energy deposits occurring in a detector. The dt heuristic is a measure of how far apart clusters are from one another. This directly impacts how much time separates

the read-out of the majority of the signal from each energy deposit. In data, there may be insufficient timing resolution to distinguish a waveform arising from two spatially close energy deposits from a single-site waveform. The dt heuristic is tuned to calibration data for each detector to determine what separation is needed for a hit to be considered multi-site. Single-site and multi-site events are currently not separated during background studies. If the dt heuristic method can be shown to match the performance of the AvsE cut over the full considered energy range, background modeling work might be improved in the future by utilizing this information.

Like the grouping of components during fitting, the simulations post-processing framework has been updated since the work of M. Buuck and T. Gilliss. The primary modification between magepostproc and older post-processing using the Germanium Analysis Toolkit (GAT) is the handling of detectors that are considered suitable for analysis for only part of a dataset. This can occur if a detector was biased down mid-dataset to address problems such as high voltage breakdowns. More commonly, it applies to detectors that were operating for entire datasets but were only used for veto purposes for some portion of that time. This can occur when a detector can not be properly calibrated for a period of time or if there are instabilities in one of the pulse shape parameters used for analysis cuts. Data from these detectors are rejected during these veto-only periods, and they are only used for the purpose of determining event multiplicity. To best align with data, simulations should also reflect the way that the configuration of good detectors can change throughout a dataset. In GAT post-processing, detectors simulated as good or bad for entire datasets. In magepostproc, a more nuanced system utilizes a list of all detector configurations that occur over the course of a dataset. For a given simulated event, a detector configuration is randomly chosen, with the fraction of the dataset spent in that configuration determining that configuration's probability of being drawn. If the event contains an energy deposit in a detector that is considered to be non-operational for the chosen detector configuration, the simulated hit is discarded. If a hit is in a veto-only detector, the hit energy is discarded, but the calculated multiplicity value includes the hit. Most detectors are considered good for the majority of data-taking, so this method of discarding energy deposits does not drastically reduce the efficiency of the simulations.

Another update to post-processing is a correction to reflect the efficiency of the DCR (delayed charge recovery) cut. Currently, the passivated surface effects that lead to energy-degraded surface alphas are not well-modeled. Therefore, simulations cannot reproduce the population of surface alphas that makes up the majority of the DEMONSTRATOR's pre-cut background near $Q_{\beta\beta}$. Prior to comparisons between data and

simulations, the DCR cut is applied to data in order to remove this population. The DCR cut is highly efficient and results in a low sacrifice of bulk physics events. However, based on a study of the cut efficiency performed by C. Haufe using ^{228}Th calibration data, the sacrifice does display some energy dependence, particularly for multi-site events. This effect must be accounted for during simulation post-processing. Using calibration data, the functional form of the DCR sacrifice as a function of energy was determined separately for single-site and multi-site events in each detector. To account for DCR sacrifice, during post-processing each event has a probability of being discarded based on its energy and on whether it consists of multiple energy deposits in a single detector.

3.4.4: Binned Spectra

Some comparisons between simulations and data directly make use of the detailed information about each event contained in the post-processed simulation files. Most analyses, however, only rely on histograms of the post-processed energy spectra, which can be directly compared to data binned in the same way. These histograms of the simulations are known as probability distribution functions, or pdfs.

When determining the binning for all pdfs, a variable binning scheme was used in order to best capture the features of the data. The position and width of bins containing each relevant peak were determined first. Based on the peak shape parameters for the entire DEMONSTRATOR, determined from calibration data, the function describing a peak at a particular energy was determined. The bin corresponding to that peak was chosen to contain 99% of expected events in the peak, with the 0.5% of expected counts to both the left and right of the peak being excluded. Since energy resolution increases with energy, high energy peaks require larger bin sizes than lower energy peaks. The binning used between peaks was optimized to increase approximately linearly with energy while still containing an integer number of bins between peaks. Unlike in previous background modeling work where the binning changed between datasets to account for differences in the peak shape parameters, identical binning is now used for all experimental configurations.

Once the locations of all bin edges were determined, histograms of the simulated data were produced using that binning scheme. Each set of post-processed simulations was divided into multiple spectra based on the detector in which energy was deposited and on event multiplicity. To compare spectra between data and simulations, a proper normalization needs to be applied to the post-processed events histograms. However, some analyses, like the spectral fits described in Chapter 4, utilize information about the unnormalized number of simulated counts in each bin to calculate the Poisson uncertainties of the simulations, so the

unnormalized histograms and the normalization factors must both be stored. Since different segments of decay chains were simulated independently, their unnormalized histograms and normalization factors were stored separately, rather than performing a weighted sum during pdf generation. The same was done in rare cases where components were simulated separately but combined into a single component group that was assumed to have a uniform activity density.

The normalization factor for each simulated spectrum is the product of a decay chain segment's branching ratio and a source component's mass divided by the number of primaries simulated. Dividing the counts spectrum by the number of simulated primaries removes dependence on how many primaries were generated, producing a histogram that represents the probability of an event being detected at a given energy for a single decay. For a couple of low efficiency component groups, simulations were generated in a specialized manner to increase the quantity of simulated statistics. In these cases, a correction factor was applied to the number of simulated primaries. For example, if simulated gammas were directed over a small solid angle, a correction was applied to determine the corresponding amount of isotropically generated gammas. The branching ratio piece of the normalization factor incorporates knowledge of how probable a particular decay chain segment is to occur for a single decay. In most cases, this probability is one, but it can be lower when a simulated segment represents only one possible branch of the decay chain. Finally, the source component mass factor in the normalization is used so that pdfs can be multiplied by an activity density rather than a full component activity. After each bin is divided by the bin width in keV, normalized pdfs have units of $\frac{\text{cnts}}{\frac{\text{decay}}{\text{kg}} * \text{keV}}$. Each pdf represents the detection efficiency for a particular background source as a function of energy. A normalized pdf can be multiplied by an activity density (in Bq/kg) and a runtime (in seconds) to produce a prediction for how many counts will be detected in each energy bin due to a particular background source.

Section 3.5: Comparing Simulations to Data

Once simulations were generated and post-processed, they could be directly compared to data to determine what combination of background sources best matches the DEMONSTRATOR's observations. Comparisons between simulations and data are useful for determining how much of the measured background is due to each decay chain but also for determining how much of the background due to a particular decay chain originates in each component group. A number of observables contain information about source location. For example, a background originating in a highly localized position will create a very different event distribution among detectors than a source distributed uniformly throughout a large component with a direct line of sight to

the entire array. While this can be qualitatively understood without simulations, simulations allow for numerical comparisons. Simulations also reflect the differences in the detected energy spectra between near-detector sources and far-detector sources. The ratio between low and high energy gamma peaks and the ratio between peaks and the Compton continuum both depend on the distance and amount of intervening material between the source and detector. One key reason these observables differ based on source location is the energy-dependent attenuation of gammas in materials. While different decay gammas are produced with a fixed relative ratio, low energy gammas are more strongly attenuated by shielding than higher energy ones. Figure 3.8 demonstrates this effect by showing the difference in spectral shape between simulations of ^{232}Th in a near-detector component (the Module 1 LMFEs) and a component outside the cryostat (the inner copper shield). Comparing observables between simulations and data is complicated by the knowledge that the data spectrum was created not just by a single source but by a linear superposition of multiple spectra. Spectral fits best address this complication, but in cases where a single source is dominant or where the size of the contributions due to most sources are well constrained, studies of individual observables can also be a powerful tool.

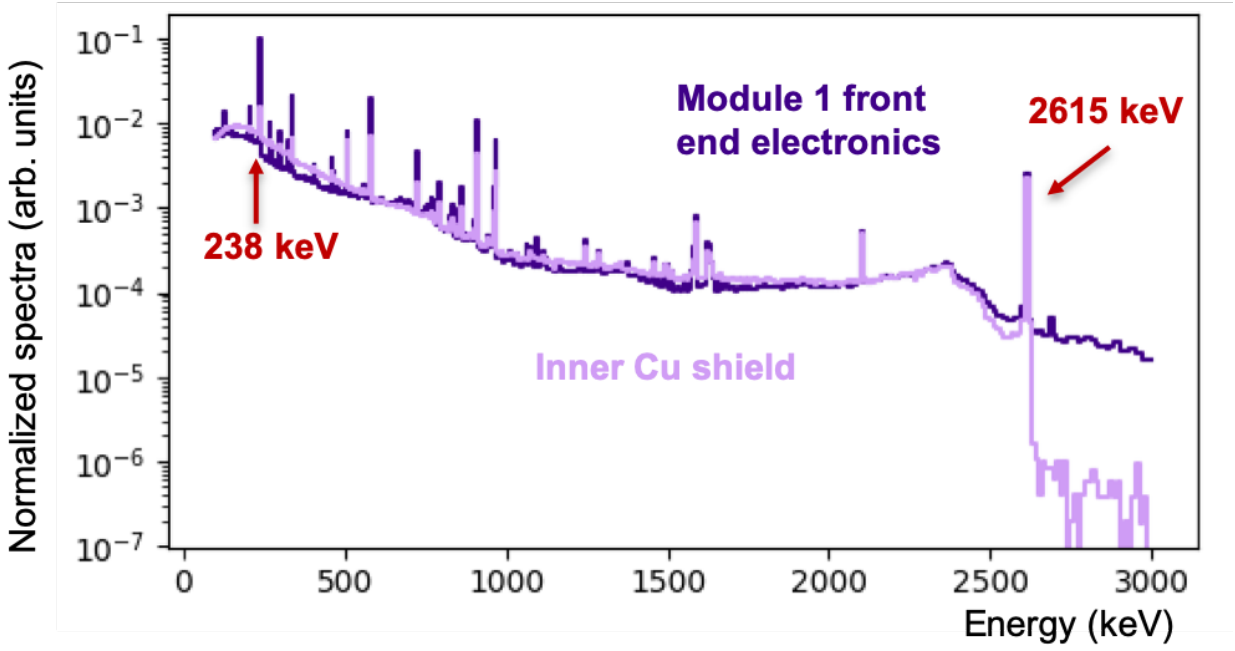
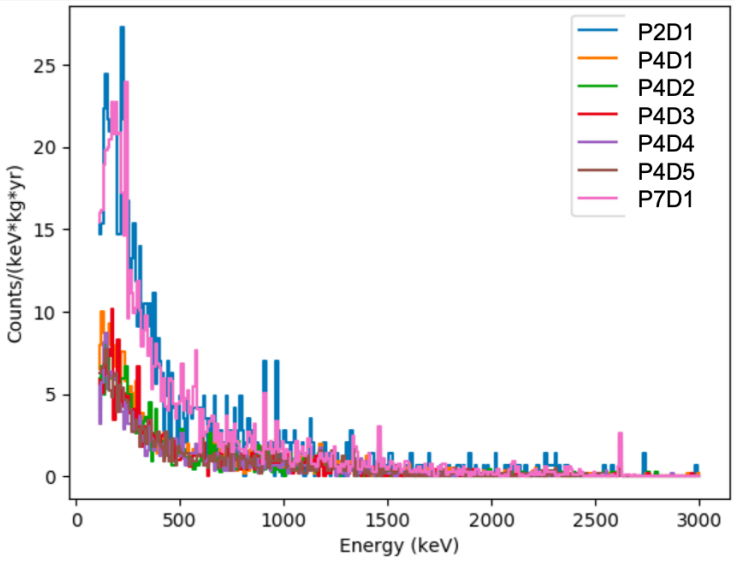


Figure 3.8: Comparison of simulated ^{232}Th spectral shape for the Module 1 LMFEs and the inner copper shield, which are examples of a near-detector background source and a source outside the cryostat. The relative heights of the 238 keV peak and the 2615 keV peak differ significantly between these two source locations because the material between the copper shield and the detectors attenuates low energy peaks more strongly than high energy ones. Spectra are normalized by dividing by the sum of all bin contents.

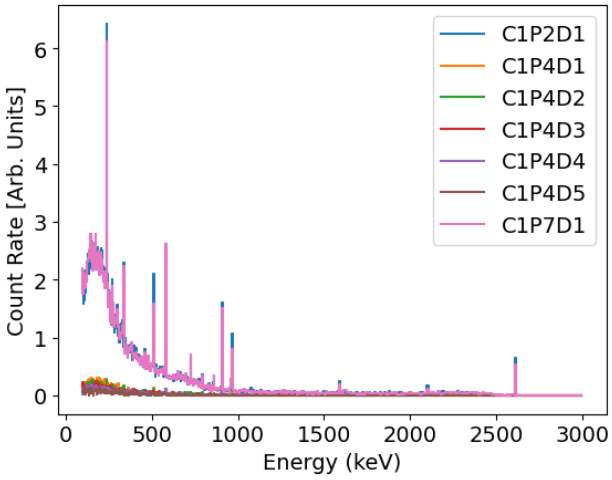
Before comparisons to simulations, cuts were applied to the DEMONSTRATOR's data to remove event

classes not included in the simulations or periods of time that did not contribute to the DEMONSTRATOR's calculated exposure. The applied cuts partially overlap with the analysis cuts described in Sec. 2.5, but not all cuts applied when calculating the $0\nu\beta\beta$ half-life limit were used when conducting background studies. In both cases, basic data quality cuts were applied to remove events that did not pass data cleaning criteria, events from bad or veto-only detectors, events tagged by the muon veto, and events that took place during liquid nitrogen fills. However, for background studies an AvsE cut was not applied to the data, since multi-site events contain useful information about the origins of the DEMONSTRATOR's backgrounds and because the corresponding heuristic for tagging multi-site events in simulations was not found to agree with the AvsE cut below 500 keV [57]. Like in the $0\nu\beta\beta$ analysis, events failing a DCR cut were removed prior to conducting background studies. As was discussed in Section 3.4.3, the DCR cut is used to remove surface events from energy degraded alphas, and it was applied prior to all background model work because there is not currently a pdf capable of describing these events. The late charge (LQ) cut, which also functioned principally as a surface event cut, was not applied prior to fitting. The LQ cut was optimized to cut events in the background estimation window, but its performance in other energy ranges was not well studied. The LQ distribution widens substantially at low energies, meaning that an energy-dependent cut would be necessary to avoid sacrificing a substantial population of bulk events. In addition, the LQ cut was used in the $0\nu\beta\beta$ analysis to reject multi-site events where the AvsE cut is not effective due to a large fraction of the total energy being deposited near the point contact. In order to match the simulated pdfs, which include both single-site and multi-site events, the LQ cut was not applied to data prior to background studies. Once the appropriate cuts were applied, data from the DEMONSTRATOR were stored as arrays of event energies for each detector and cut in each dataset. Binning that matches the binning of the simulated pdfs is applied when making spectral comparisons to inform the background model or performing spectral fits.

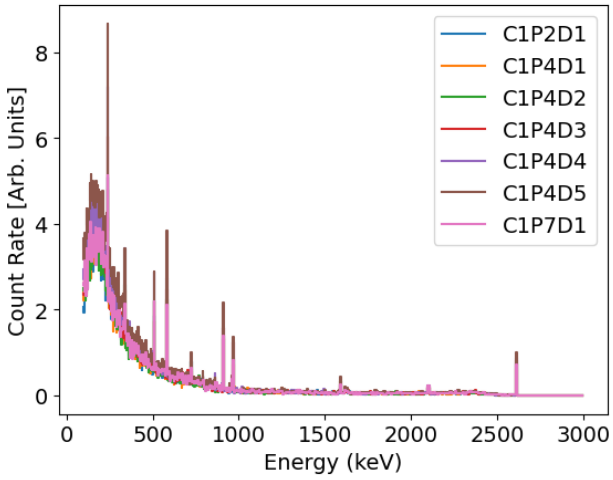
Section 3.6: Studies of Localized Module 1 Excess



(a) Background data (DS1-6c Open)



(b) Module 1 Crossarm Simulations



(c) Inner Cu Shield Simulations

Figure 3.9: Spectral comparison of the distribution of backgrounds across Module 1 natural detectors between data and selected simulations. The two high rate detectors, C1P2D1 and C1P7D1, are shown in blue and pink respectively. Other natural detectors, all of which were located in detector string 4, are shown in other colors.

As was described in Section 3.3, the DEMONSTRATOR's backgrounds were not uniformly distributed among all detectors. The conclusion of a localized background excess originating in the vicinity of the Module 1 crossarm is supported by qualitative comparisons with simulations. Simulations of a ^{232}Th background source in the portion of the Module 1 cryostat and thermospyhon copper located outside the shield, shown in Fig. 3.9b, display a rate excess in the same two natural detectors that exhibit an excess in data, as shown in Fig. 3.9a. In contrast, in simulations of an excess in the inner copper shield, plotted in

Fig. 3.9c, backgrounds are more uniformly distributed across the detectors.

3.6.1: Initial Candidate Sources of Excess

An initial set of potential candidate components in the crossarm and coldplate region were identified as possible sources of the excess that merited further examination. These candidates included a Vespel support structure in the crossarm and multiple electron-beam welds on the Module 1 crossarm and thermosyphon. The locations of these candidate components and some additional candidates considered later are shown in Fig. 3.10. Candidates were chosen primarily based on location near the high rate detectors, but attention was also paid to their material composition and handling history. For example, although electron-beam welding was not expected to introduce contaminants to copper components, the welded parts were considered a candidate when the excess was observed.

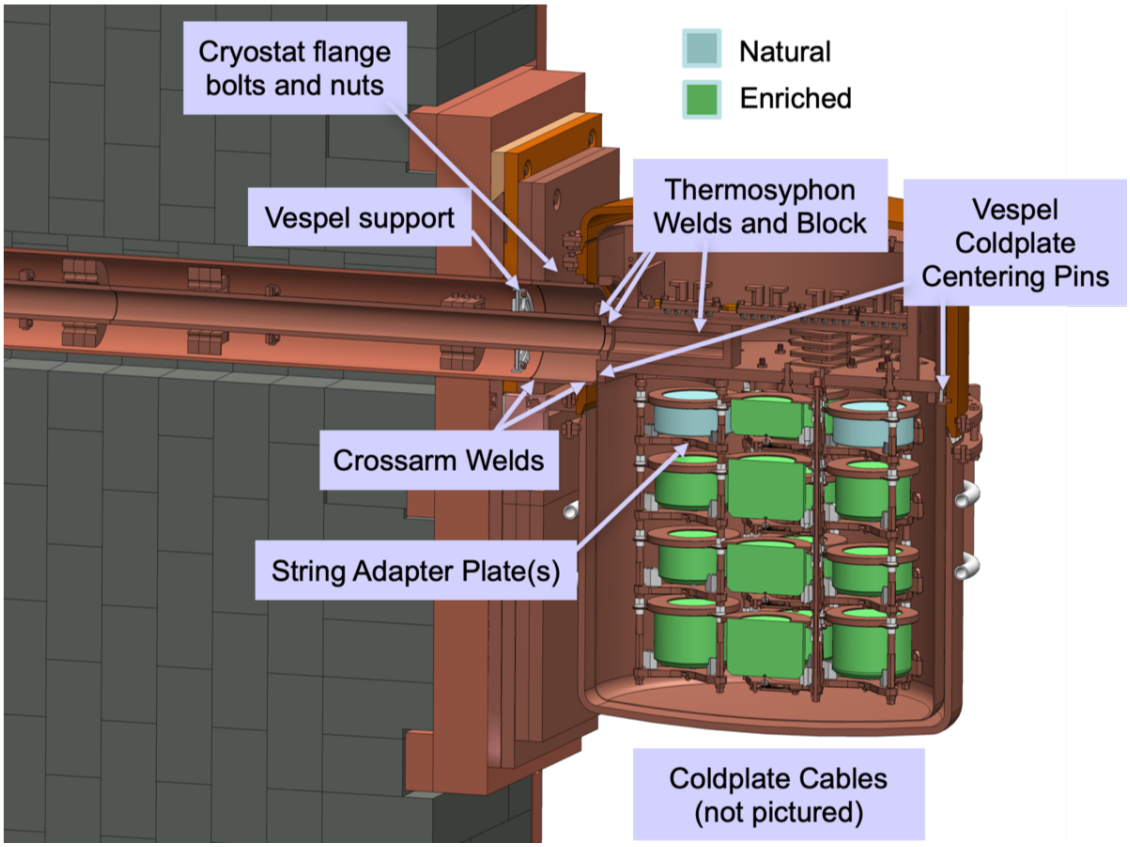


Figure 3.10: Locations of all initial candidate sources of the localized Module 1 excess ^{232}Th background, as well as some candidates that were identified and evaluated later. The natural detector (indicated in blue) in the upper left of this cross section was one of the two detectors that displayed elevated rates.

To evaluate candidate components, new simulations were generated isolating each component of interest. After each candidate was simulated, simulations were compared to data to determine how well each could

reproduce the spectral shape of the excess and to calculate the activity that would be required for it to explain the observed background. If a candidate could not be immediately ruled out based on substantial disagreements with data, the calculated activity was then used to evaluate the feasibility of assaying the component. An assay with sufficient sensitivity to detect the calculated activity could either confirm or rule out the component as the sole source of the background excess.

Multiple methods were used to estimate the approximate activity needed in each candidate to explain the ^{232}Th excess observed in data. Unless otherwise specified, these techniques held all components except the candidate at their assay-based activities, assuming that the excess over the assay model was entirely due to a single component. In the first method, the pdf for the component of interest was scaled so that the simulated model matched the 238 keV peak in C1P7D1. Background subtraction was performed prior to scaling. C1P7D1 was chosen because it was the near-crossarm detector with the best statistics. This detector's ^{232}Th background was expected to be less impacted by sub-dominant ^{232}Th sources than the module as a whole. Peak scaling to the C1P7D1 2615 keV peak was also performed. The 2615 keV peak in a single detector has a larger statistical uncertainty than the 238 keV peak, but it is not affected by continuum backgrounds. These two methods should predict similar activities if the simulated spectrum accurately describes the observed background, since different predictions reflect differences in the C1P7D1 238-to-2615 keV peak ratio between the model and data.

Finally, to make an activity prediction based on rates in more than one ^{232}Th peak, spectral fits were performed floating only the candidate component while other pdfs were fixed to their assay-based values. The general spectral fitting algorithm is described in detail in Chapter 4, but these simple preliminary fits were different in one notable way. Since fit complexity is reduced when only floating a single component, for these fits the fit was able to be subdivided by detector to take advantage of detector-specific information about the highly localized background source. One complication to drawing conclusions from fits floating a single component is the challenge of correctly accounting for uncertainties in the assay-based model. In particular, for multiple components, the assay results for ^{40}K activity were upper limits. If the pdfs for these components were fixed at their upper limits, the assay based model significantly overpredicted the ^{40}K gamma peak at 1460 keV, while the peak was severely underpredicted if these components were fixed to zero. If only one pdf was permitted to float during fitting, the fitted activity for that pdf could be biased by the chosen ^{40}K activity, since spectral fits utilize data from the continuum as well as the peaks. To account for this, fits where a few ^{40}K pdfs were floated, in addition to floating ^{238}U and ^{232}Th from the candidate

component, were also performed. Other ^{40}K components with only upper limits from assay were fixed to zero activity.

These methods provided a rough estimate of the sensitivities required to assay each Module 1 weld and the crossarm Vespel support structure. Required sensitivities ranged between ~1 and ~6 mBq, depending on the component. For a single component, the activity estimates from the methods discussed above could differ by up to a factor of ~3.5. For reference, these studies were repeated on simulations of ^{232}Th located in the two LMFEs adjacent to the high rate detectors, C1P2D1 and C1P7D1. For a ^{232}Th source in these two LMFEs, the prediction from scaling to the C1P7D1 2615 keV peak (~0.29 mBq) was ~14.5 times higher than the prediction based on scaling to the C1P7D1 238 peak (~0.02 mBq), indicating that the 238-to-2615 keV peak ratio for the simulated LMFEs is strongly in tension with the ratio in data. The lower 238-to-2615 keV peak ratio in data indicates that the LMFEs have a more direct line of sight to C1P7D1 than the true source of the excess. In general, this is a strong indication that source locations in the immediate proximity of one or both of the high rate detectors exhibit significant tension with important features of the observed data.

These studies showed that an excess in any of the initial candidate components could account for some of the discrepancies between data and the assay-based model, but no simulated pdf could fully reproduce all unexplained features of the data. Figure 3.11 is a heatmap plot comparing the performance of multiple components considered in the C1P7D1 238 keV peak scaling study. For a number of prominent ^{232}Th peaks, pulls were calculated for Module 1 and Module 2 detectors using the expression $\frac{c_{data}-c_{model}}{\sqrt{c_{data}}}$. c_{model}, or counts in the model, includes the combined contributions of the scaled pdf and of all other pdfs held at assay-based values during the scaling. The pull in the bin below each peak was subtracted from the pull in the peak bin to minimize the effect of continuum backgrounds not well accounted for in the assay-based model. Based on the consistent underprediction of Module 2 ^{232}Th peaks, as shown by the preponderance of red squares on the right half of the plot, the surveyed candidate sources for the localized Module 1 rate excess did not resolve the more modest ^{232}Th excess over assay values observed in Module 2. Since Module 2 data does not show evidence of significant variation in the detected ^{232}Th between detectors, this points to the DEMONSTRATOR's total ^{232}Th excess being split between a highly localized source near the Module 1 crossarm and a more uniform source that was underestimated in the assay campaign. The general overprediction of the Module 1 238 keV peak in the first column of Figure 3.11 suggests that models that explain the 238 keV peak well in C1P7D1 often result in too much strength in the 238 keV peak in

other Module 1 detectors.[1] Some components, especially the LMFEs near the hot detectors, overpredict the Module 1 238 keV peak (column 1) while underpredicting the Module 1 2615 keV peak (column 5), indicating that these components are not well shielded enough from Module 1 detectors to match the ratio between low and high energy peaks in data. Even candidates that do a reasonably good job of describing the Module 1 ^{232}Th peaks, such as the Vespel support structure (FrontSpider) shown in row 5, do not match all features of the data. Figure 3.12, which shows the C1P7D1 spectrum based on scaling the Vespel support structure pdf to match the excess in the C1P7D1 238 keV peak, demonstrates that the spectral shape below the 238 keV peak in data cannot be accounted for by this source.

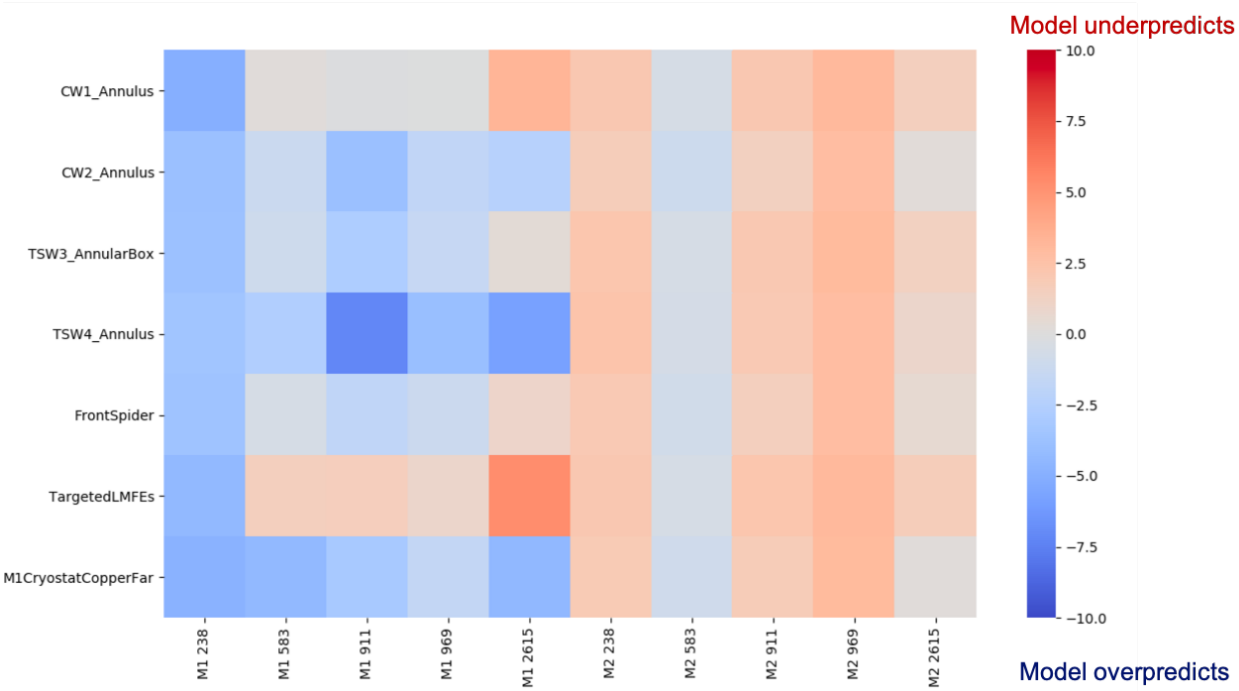


Figure 3.11: Results of a study scaling ^{232}Th activity in initial set of candidate components so that the 238 keV peak in the simulated model matches DS3-6 open data when all other components are held at assay-based activities. Background-subtracted pulls were calculated for a number of prominent ^{232}Th peaks and plotted as a heatmap with red designating that the model underpredicts the data and blue designating that the model overpredicts the data. "CW" stands for crossarm weld, "TSW" stands for thermosyphon weld, "Front Spider" indicates the Vespel support structure, and "TargetedLMFEs" indicates the LMFEs for C1P7D1 and C1P2D1.

[1]The Module 1 238 keV peak squares in this plot include the detector used for scaling (C1P7D1) as well as all other Module 1 detectors.

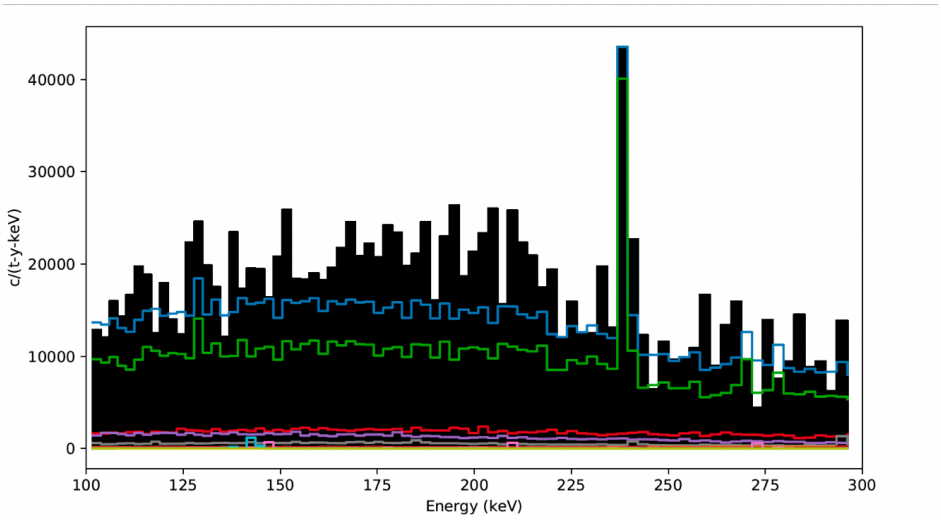


Figure 3.12: Comparison of the DS3-6 open low energy spectrum in C1P7D1 between data (black) and the model (blue) formed from scaling the Vespel support structure simulation to account for the excess over assay projections in the 238 keV peak. All other pdfs were fixed to their assay values. The summed ^{232}Th spectra from all components is shown in green, while the contributions from other decay chains are shown in other colors.

Based on the evaluation of the initial set of candidate components and their required sensitivities, an assay campaign was undertaken following the decommissioning of Module 1. The crossarm Vespel support structure was assayed using a gamma counting facility. However, the size constraints of assay facilities could not accommodate certain components, particularly the weld at the junction of the Module 1 crossarm and cryostat hoop. For this reason, Module 2 of the DEMONSTRATOR was used to perform in-situ assays of some components. After the removal of enriched detectors, Module 2 operated with only natural Ge detectors to collect data in support of the background model, making it available for temporary use as an assay system. The DEMONSTRATOR's ultra-low backgrounds and substantial detector mass enabled sensitive assay measurements in a reasonable time frame of approximately one month. To prepare the components to be assayed, the Module 1 crossarm and thermosyphon were cut to isolate the welds. The regions where cuts were performed were then cleaned with an etch solution, and the cut parts were placed inside the Module 2 shield, as shown in Fig. 3.13. Different parts were maximally separated within the shield so that the origin of any potential excess could be determined based on which detectors observed the highest rates.

Figure 3.13: In-situ assays of Module 1 candidate components using Module 2 detectors. Left: A cross sectional slice of the Module 1 crossarm centered around a weld was positioned inside the inner copper shield next to the Module 2 cryostat. Right: The Module 1 hoop and welded hoop/crossarm junction, shown after the remainder of the crossarm was cut away, was oriented so that the weld was adjacent to the Module 2 cryostat. A cut thermosyphon block containing two welds was placed in another corner inside the shield and is not pictured.

Based on 22 days of livetime, 95% CL upper limits were placed on each assayed component. For a given component, only the detectors in the closest detector string were used to set the limit. The results of the assays, compared to the estimated required sensitivities, are shown in Table 3.1. The measured activity upper limits all fell below the activities necessary to explain the excess observed in C1P7D1, ruling out the welded components as the main source of the elevated background rates in near-crossarm detectors. Assays performed at external facilities also ruled out a number of other candidates. No candidate assayed showed evidence of sufficient activity to cause the excess observed in Module 1 data.

Candidate	Lowest Required Activity (mBq)	Measured Activity (mBq)
Crossarm Weld 1	1.0	<0.52
Crossarm Weld 2	3.6	<1.3
Thermospyhon Welds	1.7	<0.72

Table 3.1: Results of the in-situ assay of Module 1 welded components. The quoted activity requirement for each component is chosen to be lowest prediction out of all estimates calculated in studies comparing simulations and data. For the two thermosyphon welds, which were assayed together but simulated separately, the lowest required sensitivity estimate for either weld is shown. For most components, the lowest estimate was based on fits floating only a few pdfs, including ^{238}U and ^{232}Th from the candidate and ^{40}K in other components.

3.6.2: Thermosyphon Cavity

Following the Module 2 assay and commercial assays that ruled out all initial candidate components, additional candidates were considered, such as the coldplate centering pins and cryostat flange bolts shown in Fig. 3.10. At this time, increased attention was also given to the possibility that the background source

might be a hot spot rather than a background uniformly distributed throughout a component. For example, a hot spot in the cables running along the crossarm and above the cold plate could cause an elevated rate in near-crossarm detectors, as is observed in data, while a uniform increase in cable activity would not agree well with the background distribution in data.

The possibility of a hot spot complicated the procedure for simulating excess candidates, because standard simulations assume decays are equally distributed throughout a component. Non-standard simulations were produced to test hot spots within some components, but selecting the best hot spot location and size in each case was challenging. For some components, multiple simulations of different possible hot spots were produced to study the variability in the resulting energy spectra and detector distributions.

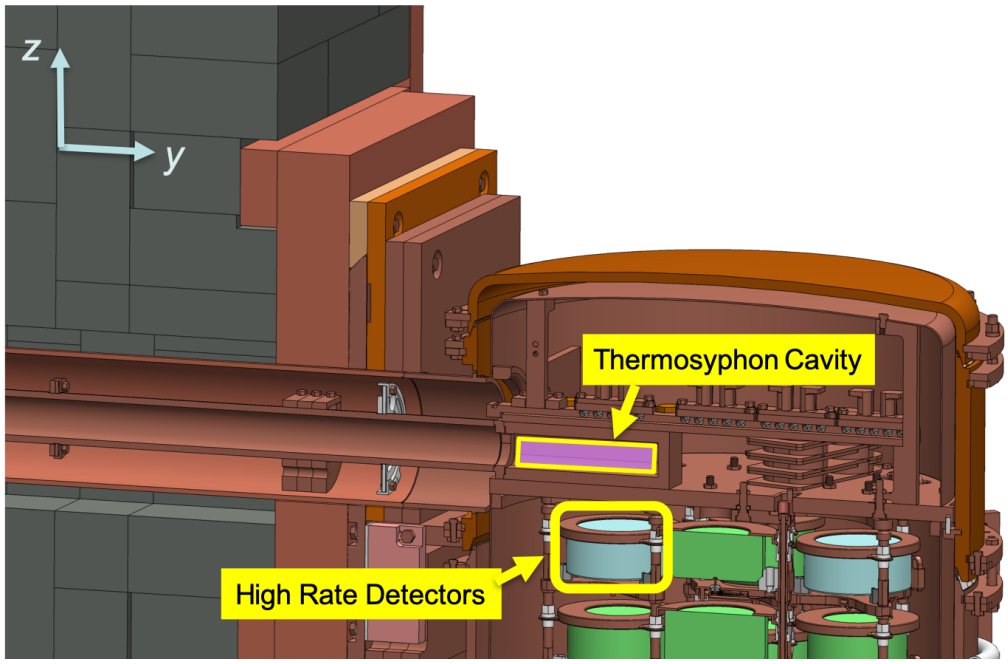


Figure 3.14: Location of the thermosyphon cavity region and its proximity to the high rate natural detectors. Simulations have shown that a hot spot in the cavity or a contamination on its bottom inner surface is able to reproduce many features of the localized ^{232}Th background excess observed in the DEMONSTRATOR's data.

For candidates not included in the initial analysis, C. Haufe performed studies similar to those described in the previous section [69]. These studies ruled out a number of candidate components and identified a region that performed particularly well, the M1 thermosyphon cavity. The location of the cavity region relative to the high rate detectors is illustrated in Fig. 3.14. Hot spots were simulated for multiple regions within the thermosyphon cavity. A uniform contamination on the bottom of the cavity's inner surface was also simulated. The bottom surface simulations and simulations of a hot spot in the bottom center of the

cavity both successfully reproduced a number of the most important observables in data. For convenience, in the following discussion these two source locations will generally be referred to together as a hot spot at the center of the cavity. A qualitative description of the major conclusions of thermosyphon cavity studies is included here, but full details can be found in [69]. In the following discussion, the success of this hot spot location at reproducing salient features of the observed data will be contrasted with other possible hot spot locations to support the conclusion that this source location is well-suited to explain the DEMONSTRATOR's findings. When coordinate directions are referenced, they are defined by the coordinate axis in Fig. 3.14, with x being the direction into or out of the page.

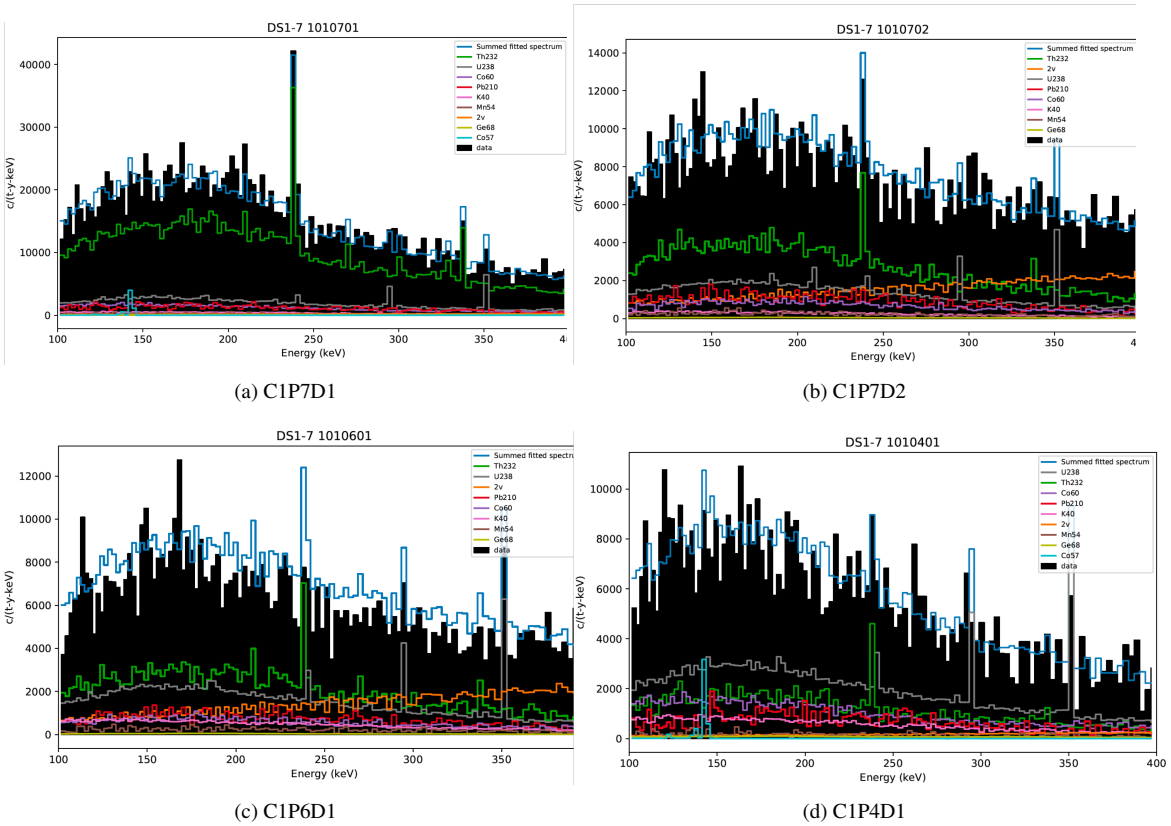
One relevant observable pertaining to the localized background source is the relative rates in the two high rate natural detectors. In datasets where both were active, these two detectors, C1P7D1 and C1P2D1, were observed to have similar continuum rates. The thermosyphon cavity is essentially equidistant from the two high rate detectors, and simulations confirm that backgrounds originating at the cavity's center produce similar event rates in C1P7D1 and C1P2D1. Any background source that is significantly displaced from the thermosyphon in the x-direction cannot match the relative rates of the two high rate detectors observed in the DEMONSTRATOR's data. Even simulations of a hot spot within the cavity but displaced in the x-direction to the cavity edge induce too much asymmetry to be consistent with the observed data.

Next, the ratio between a low energy (238 keV) and high energy (2615 keV) ^{238}Th peak in C1P7D1 agrees well between data and simulations of a hot spot at the center of the thermosyphon cavity. As above, an individual detector was used for this comparison rather than looking at the combined peak ratio over all Module 1 detectors because the C1P7D1 peak ratio was less significantly impacted by subdominant ^{232}Th sources. The success of the thermosyphon cavity simulations at replicating the C1P7D1 peak ratio in data indicates that the cavity has approximately the correct amount of shielding separating it from this detector. Conversely, hot spots simulated above or below the thermosyphon cavity are either too well-shielded or not well-shielded enough to reproduce the observed peak ratios.

Finally, a hot spot at the center of the thermosyphon cavity is largely effective at matching the relative event rates between different detectors. The comparatively low rates in several other detectors near the two high rate detectors, C1P7D1 and C1P2D1, can effectively constrain the region of the excess source. Although a number of nearby detectors were biased down and unable to provide information about the excess source location, two additional top row detectors were operational, one of which was immediately adjacent to C1P7D1. The event rate in this detector, C1P6D1, was modestly elevated compared to other enriched

detectors but much lower than the rate in the two detectors closest to the crossarm. The other top row detector, C1P4D1, was a natural detector that did not display a significant excess. The comparatively low rate in these top row detectors located further from the crossarm clearly indicates that the excess source was not uniformly distributed above the coldplate. The relative rates among top row detectors even limits where within the thermosyphon cavity a hot spot could be located. The best candidate hot spots in the center of the cavity slightly overpredict the rate in C1P6D1, but moving the hot spot deeper into the cryostat (farther from the crossarm) further exacerbates the problem. Sources from very far background sources with shine paths down the crossarm, such as the vacuum hardware outside the lead shield, also cannot faithfully represent this feature of the background distribution.

The detector located below C1P7D1, C1P7D2, was also operational and served as a useful tool for constraining the background region. For a source originating in the middle of the thermosyphon cavity, C1P7D2 is very well shielded by C1P7D1 itself, leading to very different background rates in the two detectors. Based on simulations, this difference is in fairly good agreement with the difference observed in data, both in terms of overall event rate and the magnitude of the 238 keV ^{232}Th peak. It is very difficult to identify other source locations within the DEMONSTRATOR where the shielding of C1P7D2 is so much larger than that of the detector immediately above it. Even a hot spot in the portion of the cavity closest to the crossarm can be ruled out, because C1P7D2 is not shielded by C1P7D1 from this portion of the cavity. In general, the relative rates between different detectors near the top of the Module 1 array constrain the y-position of the excess source.



(a) C1P7D1

(b) C1P7D2

(c) C1P6D1

(d) C1P4D1

Figure 3.15: Results of a preliminary spectral fit to DS1-7 open and blind data where a point source in the center of the M1 thermospyon cavity was the only ^{232}Th source allowed to float. Detectors were divided into groups during the fitting based on module and enrichment, but results are plotted separately for each detector. Note: Range of the y-axis varies significantly between plots to allow details of the lower rate detectors to not be obscured.

To supplement, C. Haufe's peak ratio studies, a preliminary fit to DS1-7 open and blind data using a pdf of a thermosyphon cavity hot spot as well as pdfs of all standard background sources (those included in the assay-based model) were performed as a part of this work. For this fit, all ^{232}Th sources except the thermosyphon cavity were held at their assay activity values to test the effect of placing the entire ^{232}Th excess in a single candidate. The pdfs in other decay chains were allowed to float. The results of these fits are shown in several key detectors in Fig. 3.15. Fig. 3.15a demonstrates that simulations of a thermosyphon cavity source very closely match the spectral shape of the data in C1P7D1, the high rate detector that was online for the entirety of data-taking. It is particularly useful to compare the overall performance in C1P7D1 between Fig. 3.15a and Fig. 3.12. Although there are differences in the exact datasets used and some of the details of the fits, it is evident from this comparison that the M1 thermosyphon cavity source matches the spectral shape below the 238 keV peak much more successfully than previously-tested components in the crossarm region. Fig. 3.15b shows that in the detector below C1P7D1, the thermosyphon cavity simulated

source leads to a model that is slightly high but still matches the data reasonably well. Note that the y-scales are different by more than a factor of two between C1P7D1 and C1P7D2, meaning that this candidate is successfully able to capture a large difference in the overall event rate between these two detectors. The two top row detectors that did not exhibit anomalously high rates, C1P6D1 and C1P4D1, are shown in Figs. 3.15c and 3.15d. The slight overprediction of the model in C1P6D1 is one of the weakness of the thermosyphon cavity hot spot source, but no other region of the experiment has been identified as able to reduce this rate without leading to larger tensions in other key observables. In general, the thermosyphon cavity source is successfully able to capture the comparatively low rates in these two top row detectors further from the crossarm.

A ^{232}Th hot spot in the M1 thermosyphon cavity accurately models many important features of the M1 localized background observed in the DEMONSTRATOR's data. However, neither it nor any other reasonable candidate source for this localized effect can fully explain all differences between the DEMONSTRATOR's data and the assay-based model. This is most evident by considering Module 2 ^{232}Th rates. Although the M2 background rate near $Q_{\beta\beta}$ is less than half as large as that in M1, it still is significantly in excess of the assay-based prediction. In general, a localized M1 source, scaled to account for the number of counts in the high rate detectors, does not contribute sufficient strength in M2 detectors to explain the background excess there. The results of full spectral fits in Chapter 6 quantify what percentage of the overall enriched detector background rate near $Q_{\beta\beta}$ can be explained by the localized M1 background source.

To test the hypothesis of a background source in the M1 thermosyphon cavity, the component was assayed using the GeMPI detector at LNGS [70]. The assay did not detect evidence of ^{232}Th contamination [71]. Based on two analyses, one assuming a uniform bulk contamination and one assuming an inner surface contamination, 90% upper limits were set on the ^{232}Th activity of the thermosyphon cavity. Both limits were lower than the required activity from peak-scaling studies and preliminary fits, meaning that, at the time of the assay, the thermosyphon cavity did not contain sufficient ^{232}Th to account for the excess in the DEMONSTRATOR's data. However, the thermosyphon cavity was particularly well-suited for a transient contaminant. During the DEMONSTRATOR's data-taking, the thermosyphon cavity was part of the experiment's nitrogen-based cooling system. The nitrogen that cycled through the thermosyphon was part of a closed system that never made direct contact with the LN powering the heat exchanger, which was periodically replaced. This closed loop, which included a dewar and the thermospyhon, limited the potential for a contaminant to be introduced during data-taking. A theory that ^{220}Rn from the ^{232}Th chain was circulating

with the nitrogen is also disfavored by the presence of prominent lines from ^{228}Ac in the high rate detectors, as shown in Fig. 3.16, since ^{228}Ac is above ^{220}Rn in the decay chain. A contaminant originating within the closed loop, however, could have been picked up by the cycling nitrogen and deposited on the bottom surface of the cavity. A background of this type could have evaporated away when the experiment was opened following the end of data-taking, meaning it would no longer be present at the time of the assay. This possibility cannot be directly tested. For this reason, the negative assay result does not rule out a point source in the thermosyphon cavity as a candidate for the DEMONSTRATOR's ^{232}Th excess. Because simulations of a hot spot in the M1 thermosyphon cavity point to this source as the most capable of producing the DEMONSTRATOR's observed background distribution, it was chosen for inclusion during spectral fits despite being disfavored by assays.

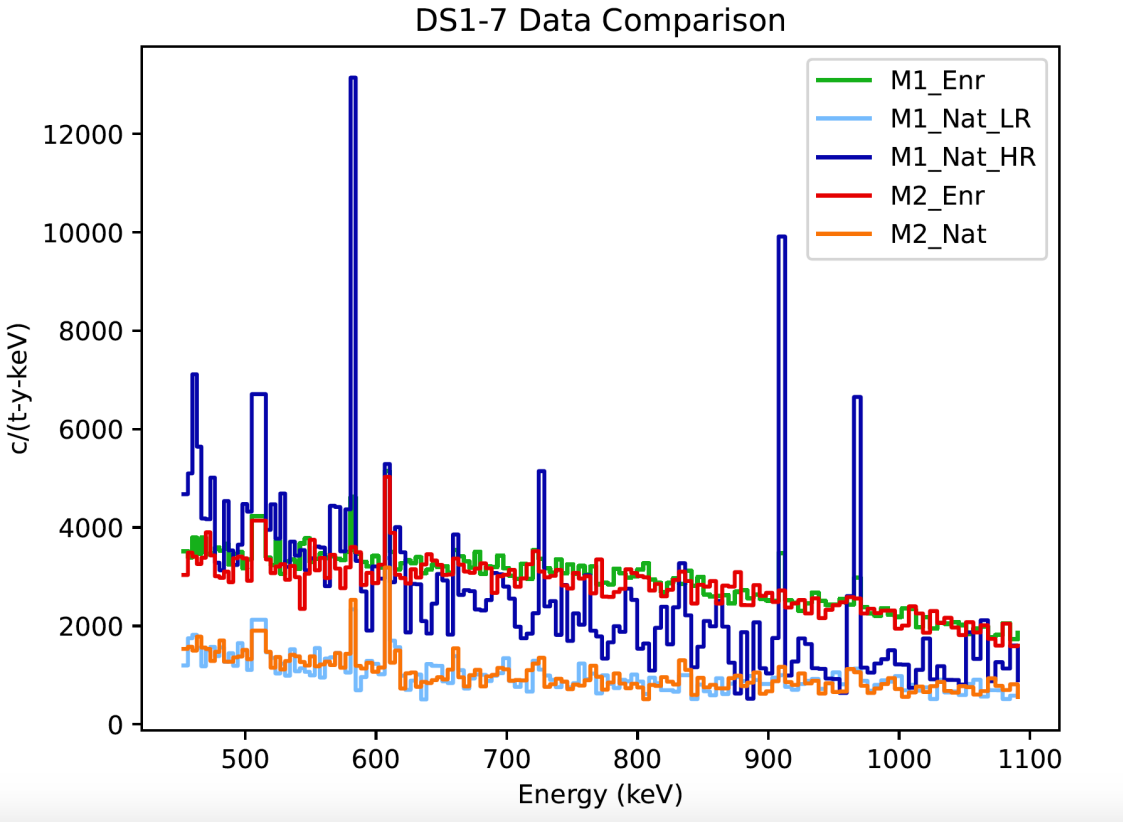


Figure 3.16: Comparison of gamma peaks from ^{228}Ac at 911 keV and 969 keV between the two high rate natural detectors (M1_Nat_HR) and other detector groups.

The background studies described throughout this chapter revealed several noteworthy features of the DEMONSTRATOR's data that provided insight into the dominant background sources, in particular the presence

of a localized ^{232}Th source near the M1 crossarm. These studies also led to the identification of a candidate M1 ^{232}Th hot spot source for inclusion in the final background model, setting the stage for full spectral fits to the DEMONSTRATOR's data.

CHAPTER 4: Spectral Fits

Although the studies described in Chapter 3 provide useful and interpretable information about the Demonstrator's backgrounds, each study only depended on a small subset of observables. In contrast, spectral fits incorporate a wide range of observables to build a quantitative model of the Demonstrator's background composition. These approaches complement each other well because supplemental studies help clarify which factors may be driving fitting results. In addition, some supplemental studies make use of information not directly accessible during fits. For example, during spectral fits, the spectra of multiple detectors are grouped together to ensure sufficient statistics in each bin, leading to a loss of information about the detector distribution of backgrounds. Particularly when used in conjunction with supplemental studies, spectral fits provide a powerful tool for making use of data across a wide range of energies to determine the contributions of many different background sources. This chapter describes the spectral fitting algorithm, the method for determining statistical uncertainties, and studies with simulated datasets that have been used to evaluate and refine the fitter.

Section 4.1: Fitting Technique

Spectral fits aim to determine the activity density by which each probability distribution function should be weighted such that the weighted sum over all pdfs best agrees with data. The frequentist technique for finding the optimal set of activity densities is to maximize the likelihood function, or equivalently to minimize the negative log likelihood.

4.1.1: Barlow-Beeston Likelihood

For a fixed exposure, the number of detected events in an energy bin due to a radioactively decaying source is governed by Poisson statistics. The Poisson negative log likelihood (NLL) takes the form

$$-\ln(\mathcal{L}) = -\sum_{i=1}^{n} d_i \ln(f_i) + f_i + \ln(d_i!), \tag{4.1}$$

where d_i represents the number of counts in the data in bin i, f_i represents the number of counts in the model in bin i, and n is the total number of bins.

A variation on the Poisson likelihood function, known as the Barlow-Beeston likelihood, is used in the frequentist fitting suite [72]. In addition to properly accounting for statistical fluctuations in data, the Barlow-Beeston method also takes into account the impact of statistical fluctuations in the Monte Carlo simulations used to create pdfs. This is implemented by assuming that the number of counts simulated for a background source, j, in bin i, a_{ji}, is Poisson distributed around the unknown expected number of counts, A_{ji}. The full negative log likelihood function is given by summing these two sets of Poisson NLLs:

$$-\ln(\mathcal{L}) = -\sum_{i=1}^{n} d_i \ln(f_i) + f_i + \ln(d_i!) - \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji} \ln(A_{ji}) + A_{ji} + \ln(a_{ji}!). \tag{4.2}$$

This expression depends on the activity density associated with each pdf, p_j, because the number of counts predicted by the model for each bin is given by $f_i = \sum_{j=1}^{m} p_j w_j A_{ji}$, where the weight of a pdf, w_j, is the product of runtime and source mass divided by the number of simulated primaries, and m is the total number of pdfs.

Barlow and Beeston show that the values of the A_{ji} that will maximize the likelihood can be found by solving a transcendental equation for each bin. The optimal set of A_{ji} can be written as:

$$A_{ji} = \frac{a_{ji}}{1 + p_j w_j t_i}, \tag{4.3}$$

where the t_i are found by solving

$$\frac{d_i}{1 - t_i} = \sum_j \frac{p_j w_j a_{ji}}{1 + p_j w_j t_i} \tag{4.4}$$

Section 5 of the Barlow-Beeston paper details a special treatment for bins where the pdf with the largest strength has zero counts. For non-zero values of a_{ji}, it can be shown that the optimal value of A_{ji} is also non-zero. However, if a_{ji} is zero for a bin with a nonzero number of counts in data ($d_i \neq 0$), the optimal A_{ji} may or may not be 0. A nonzero A_{ji} in this situation represents the case where a simulation has zero counts in a bin only due to insufficient simulation statistics. It can be shown that if d_i is nonzero and the pdf k with the largest strength has a bin content of 0, the optimal A_{ki} is given by

$$A_{ki} = \frac{d_i}{1 + p_k w_k} - \sum_{j \neq k} \frac{p_j w_j a_{ji}}{p_k w_k - p_j w_j}, \tag{4.5}$$

as long as this expression leads to a positive value. If the A_{ki} given by Eq. 4.5 is negative, $A_{ki} = 0$. The other A_{ji} are given by the usual calculation, Eq. 4.3. To summarize, if there are counts in data in a particular bin and the model does not have enough counts from pdfs with non-zero bin contents to easily account for it, a non-zero A_{ki} may be optimal for one of the pdfs. The pdf that is most heavily weighted can best account for the extra counts in data without incurring too high of a penalty from the last two terms of Eq. 4.2. However, this treatment is flawed in the context of the Demonstrator's spectral fits, because it fails to take into account that some sources, such as $2\nu\beta\beta$, cannot lead to counts in bins above a known energy. For these pdfs, a non-zero A_{ji} does not make physical sense for high energy bins. The statistics of the Demonstrator's simulations are sufficiently high that a large fraction of zero-count bins in pdfs can be explained by sources that are known to not contribute at high energies. For this reason, the special treatment of zero-count bins is neglected in this analysis, and A_{ji} is set to 0 for all cases where a_{ji} is zero.

When minimizing the negative log likelihood, it can be helpful to make use of its analytic gradient. Each term of the gradient can be calculated by taking the partial derivative of the NLL with respect to a p_j:

$$\frac{\partial(-\ln(L))}{\partial p_j} = \sum_{i=1}^{n} \left(1 - \frac{d_i}{f_i}\right) w_j A_{ji}. \tag{4.6}$$

At each step in the minimization with respect to the p_j, the optimal A_{ji} previously solved for can be plugged into this expression. Details of how the analytic gradient is used during fitting, along with other details of the likelihood implementation in the frequentist fitting suite, will be described in Section 4.1.2

4.1.2: Implementation

The Barlow-Beeston likelihood provides a method for comparing a model to the data in each energy bin, but in the case of the Demonstrator, it is useful to subdivide both model and data based on more than just energy. The background composition of the Demonstrator's data differs between different groups of detectors and different periods of the Demonstrator's data-taking, meaning that useful information is lost if the fit is performed using only a single combined energy spectrum. Therefore, the Demonstrator's spectral fits are actually simultaneous fits to multiple energy spectra, and the background model can be thought of as a collection of related submodels.

The model is first subdivided into dataset submodels based on experimental configuration. An experimental configuration can span multiple datasets as long as no major changes to the Demonstrator's hardware occurred between them. For example, DS1 and DS2 were considered to have the same experimental

configuration, since both comprised only a single module and both had the same shielding configuration. The experimental configurations included in spectral fits were DS0, DS1-2, DS3-6, and DS7.

Each dataset submodel was then further divided based on detector group and multiplicity into detector submodels. In general, detectors were grouped together if they were located in the same module and had the same enrichment status. The natural detector group was further subdivided to separate the two high rate detectors identified in Sec. 3.3 into their own group. In principle, each detector could be separated into its own submodel, but this was found to worsen the performance of the spectral fitting algorithm because it led to a large number of low statistics bins. The spectrum for each detector was broken down into a multiplicity one spectrum and a high multiplicity spectrum depending on whether energy was simultaneously deposited in any other detector. For high multiplicity events, the spectrum consisted of single-detector energies, not the sum of the energy deposited in all detectors. Taking into account that not every detector group was present for every experimental configuration, 26 spectra were simultaneously fit.

The different submodels are linked by the activity densities associated with the different background sources in the model. A single set of activity density parameters is shared between all detector models for the same experimental configuration. In most cases, activities are also shared between datasets, with a transformation applied to account for backgrounds decaying away over time with a known half-life. The parameters floated during spectral fits can be interpreted as the activity densities at the beginning of the commissioning dataset, DS0. To transform this quantity into an average activity for a particular dataset submodel, the start time and end time of the dataset range relative to the beginning of DS0 are used to find the expectation value of the time t for that dataset, given the exponential decay of that isotope. If the half-life for an isotope is long enough that the expectation value calculation is unfeasible due to machine precision, the midpoint of the dataset is used instead. The exponential decay equation is then evaluated at this time to determine what fraction of the activity remains. This method is most significant for isotopes with a short half-life, such as ^{68}Ge and ^{57}Co, but it is applied even for longer-lived isotopes.

In rare cases, an exception is made and the activity density for a decay chain in a component group is independent between two experimental configurations. This occurs for the Module 1 gaskets, where a different material was used during commissioning than for the rest of data-taking. In the case of $2\nu\beta\beta$ decay, an additional constraint is applied requiring that the pdfs for decays originating in Module 1 enriched detectors and decays originating in Module 2 enriched detectors share the same activity density. The same is true for Module 1 and Module 2 natural detectors. An option also exists to constrain enriched and natural

detectors to have the same $2\nu\beta\beta$ decay half-life, meaning their activity densities are constrained to differ only by the ratio of their enrichment fractions. This option was not used in standard fits because the dead layer thicknesses were not as well constrained for natural detectors, as discussed in Chapter 5.

When a spectral fit is performed, the Barlow-Beeston negative log likelihood function is minimized using the migrad minimizer in iminuit [73]. The iminuit package is the python interface for the C++ minuit minimization package, and the migrad algorithm is the recommended minimization technique for most functions, which makes use of a variable-metric method [74] [75]. By default, the activity densities of all pdfs are floated, but a subset of the activity densities can also be held constant during fitting. One-sided limits are placed on all floated parameter values during the fit in order to ensure that no activity densities are assigned values less than 0. The 'strategy' argument in iminuit, which controls how frequently the Hessian is explicitly calculated during the fit, is set to 2, meaning that the Hessian is calculated at every step in the minimization. This approach means that the minimization process is slower, but it removes reliance on an approximation of the Hessian that can become distorted during minimization.

When the NLL is calculated during the fitting process, it is sequentially calculated for each detector group model in each dataset group and then summed. The analytic gradient of the NLL is simultaneously computed and passed to the migrad minimizer. The negative log likelihood and gradient calculation is implemented using cython, a python compiler designed to allow performance comparable to C [76]. In order to speed up the code, OpenMP is used to parallelize the parts of the calculation that can be performed independently on each bin [77]. The NLL is then calculated using Eq. 4.2. The constant factorial terms, which can in principle be omitted, are included in order to keep the magnitude of the NLL small, decreasing issues with machine precision in the fitter. In order to solve the transcendental equation for each bin needed to optimize the A_{ji}, Eq. 4.4, Newton's method is used, as suggested by the Barlow-Beeston paper [72]. If a solution within the tolerance is not found after 50 iterations of Newton's method, a bisection method is used. If no solution is found after 50 iterations of bisection, the code proceeds, but the failure is printed out to the log file.

Before performing spectral fits, all parameters were set to randomized initial values. To accomplish this, the activity densities of all pdfs were first adjusted such that the total number of integrated counts in the data was equally divided between all pdfs. In fits where some parameters were fixed, the counts resulting from the fixed pdfs were subtracted away from the total integrated count calculation before dividing counts among the remaining pdfs. In order to introduce a degree of randomness while roughly preserving the total number of counts in the model, each activity density was then multiplied by a random number between 0 and 2. This

method allowed the randomization to be carried out in terms of model counts rather than activities densities, preventing low efficiency pdfs from always contributing near zero counts at the start of the fit.

Section 4.2: Calculating statistical uncertainty

To draw conclusions from the results of background model fits, it is necessary to accurately determine the uncertainty of each fitted parameter value. This section details the procedure for ascertaining the statistical portion of each parameter uncertainty. Chapter 5 will describe how the procedure is extended to incorporate the contribution of systematic effects on the overall uncertainties.

Error evaluation in fits for the DEMONSTRATOR is complicated because the large number of floated parameters makes some error estimation techniques time and resource intensive. In addition, many parameters fit to values at or near the lower limit of zero activity. Also, in the DEMONSTRATOR's highly dimensional background model fits, the negative log likelihood is not parabolic around the minimum in all parameters. This limits the utility of methods that rely on the assumption of parabolic behavior, such as minuit's hesse algorithm, which determines parameter uncertainties using the diagonal terms of a covariance matrix approximated by inverting the Hessian.

Much of the background model development for the DEMONSTRATOR was performed using a non-parametric bootstrap to evaluate statistical uncertainty. The non-parametric bootstrap, also known as the resampling bootstrap, is a widely applicable technique for error estimation in which the data are randomly sampled with replacement to form bootstrap datasets. Each bootstrap dataset contains the same number of samples as the original dataset. This procedure approximates repeating the experiment many times by using the data as a proxy for the underlying distribution. Each bootstrap dataset is then independently fit. For a given parameter, the distribution of fitted values from all bootstrap datasets is used to calculate relevant statistics, such as confidence intervals.

Estimating errors through bootstrapping circumvented the requirement of a negative log likelihood function that behaves parabolically near the minimum, and it allowed the uncertainty estimates for all parameters to be performed simultaneously. It also provided a natural mechanism of combining the uncertainties of multiple component groups in a way that properly accounted for parameter correlations. However, in some cases, pdfs that fit to zero activity displayed unphysically small or non-existent uncertainties when using the bootstrap method of calculating error bars. This phenomenon occurred because the bootstrap technique does not achieve valid coverage for parameters at the boundary of the parameter space [78]. To understand why

this is, consider the case of a component that has zero activity in a particular decay chain. Due to statistical fluctuations, if the experiment were repeated many times, in some fraction of the trials the best fit activity density for the parameter would fall below zero for a fit performed without constraints. In these cases, if bootstrap datasets are generated from the data, their unconstrained best fit values would be centered around the sub-zero best fit value of the original data, rather than around zero. Since parameters are typically constrained to be non-negative during fits, any parameter that would fit to a negative value in an unconstrained fit instead fits to zero. When this occurs for a large portion of the bootstrap datasets, the standard deviation of the fitted activity distribution is artificially small, leading to a severe underestimation of the true parameter uncertainty. This issue is not limited to components with zero activity. Any pdf that leads to a small number of energy deposits in the detector array over the lifetime of the experiment has a chance of producing a negative best fit value, leading to similar problems. The difficulty is further exacerbated if missing components in the model or other mismodeling tends to drive some parameters to negative values larger than would be expected from statistical fluctuations. While remedies for the undercoverage of bootstrapping near parameter boundaries have been proposed, these alternatives tended to require significant added complexity and computation time.

Ultimately, the profile likelihood technique was chosen instead to determine the frequentist confidence interval for each floated parameter. This technique scans the negative log likelihood over a single parameter in the region surrounding the minimum. At each step in the scan, the scanned parameter is held fixed while the negative log likelihood is minimized with respect to all other parameters. Based on Wilks' theorem, in the large sample limit, a likelihood ratio Λ can be related to a chi-squared distribution by the relationship $-2 * \log(\Lambda) \simeq \chi^2(k)$, where k represents the number of degrees of freedom [79]. This relationship can be applied to each point in the likelihood scan to estimate its p-value. Using log rules to convert -$\log(\Lambda)$ to the difference between two NLLs, the NLL at the global minimum is subtracted from the NLL at each scanned point as a measure of how much the goodness of fit is penalized by varying the scanned parameter. Because the NLL is minimized over all other parameters at each point in the scan, those parameters can be considered a function of the scanned parameter, meaning the fit only has one degree of freedom. A one-sigma confidence interval can be found by identifying the points where the profiled NLL is greater than the minimum NLL by $\frac{1}{2}$, and in general, an N-sigma confidence inteval corresponds to a change in NLL by $\frac{N^2}{2}$.

Wilks theorem was derived under the assumption that all parameters of interest are far from parameter boundaries, which does not hold for the many floated parameters in the DEMONSTRATOR's background model

fits that fit to zero activity. However, Rolke, et. al., found that, under certain assumptions, the profile likelihood method of error approximation can achieve accurate coverage even near parameter boundaries [80]. The DEMONSTRATOR's background model uses what Rolke terms the bounded likelihood method. In this method, when a fitted parameter is on the boundary of the allowed parameter space, the change in likelihood used to determine the confidence interval is calculated relative to the likelihood at the boundary, rather than at the value the parameter would take in an unconstrained fit. This is notably different from the non-parametric bootstrap technique described above, where the uncertainty estimation cannot properly take into account the presence of the boundary.

The minos algorithm in iminuit implements the profile likelihood method. The iminuit package also includes a method called mnprofile, which traces out the shape of the NLL curve with respect to a particular parameter while profiling over the parameter at each step. After a spectral fit is performed on the DEMONSTRATOR's data, each parameter is profiled in a separate job, allowing the uncertainty estimation process to complete in a reasonable timeframe. For each parameter, minos is first used to determine the interval corresponding to a 3σ uncertainty, and then mnprofile is used to plot the profiled NLL curve over the 3σ interval. Smaller confidence intervals, such as the 1σ uncertainty, can then be extracted from the profiled curve.

Section 4.3: Evaluating Fitter Performance with Simulated Datasets

Prior to performing fits on the DEMONSTRATOR's data, the fitting and uncertainty estimation algorithms were validated using simulated datasets. Simulated datasets are datasets drawn from a known distribution. In the case of the DEMONSTRATOR, this typically involves first constructing a model from the same set of ~110 pdfs used during fitting. Each pdf is weighted by a user-chosen activity density and by the component mass. The weighted sum over all pdfs forms a binned distribution from which data samples can be drawn. The set of all data samples drawn from the distribution, known as a simulated dataset, can then be fit using the same procedure applied to experimental data. Simulated datasets have several advantages for evaluating fitter performance. Foremost among these is the fact that, for a simulated dataset, the parameters floated during fitting have known true values, the activity densities used to generate the simulated dataset. This means that the fit's ability to extract the true parameter values can be assessed.

Another advantage of simulated datasets is that the number of samples drawn can be selected based on the topic being studied. High statistics simulated datasets are most useful for method validation, while

simulated datasets that approximate the number of counts in the DEMONSTRATOR's full exposure dataset reveal how well the fitter will realistically be able to perform on data with the DEMONSTRATOR's level of statistics. Although they do not take into account systematic sources of uncertainty, fits to simulated datasets with MAJORANA-like statistics can provide insight into whether the DEMONSTRATOR has a sufficient level of statistics to meet the goals of background modeling, such as a precision measurement of the $2\nu\beta\beta$ half-life and the determination of the ^{232}Th excess source location. They can also be used to check the implementation of the profile likelihood method of determining statistical uncertainties.

For the simulated datasets discussed in the remainder of this chapter, samples were drawn from a model where pdfs were fixed to their assay-based activity densities unless otherwise indicated. For component groups where the assay result for a particular decay chain was an upper limit, the corresponding pdf was set to an activity of zero. The pdf for Pb bremsstrahlung in the lead shield was set to a data-driven value, since this activity could be well-constrained by comparing the DEMONSTRATOR's commissioning data to data where the full shield was installed. The M1 thermosphyon cavity was assumed to have no surface or hot spot contamination.

To determine the number of samples to draw when forming a simulated dataset, the number of counts in the assay-based model is found for a particular runtime. Typically this is either the DEMONSTRATOR's DS0-7 runtime or some multiple of it. In this section, the term "MAJORANA-level statistics" will be used to signify the number of counts expected from the assay-based model based on the DS0-7 exposure. However, the assay-based model is known to have lower backgrounds than were present in the DEMONSTRATOR's data, especially for the ^{232}Th decay chain. While assay-based simulated datasets function as a reasonable initial approximation of the background composition in data, this approach is refined in Sec. 4.3.3 when individual components are scaled to higher activities to match the level of ^{232}Th excess observed in data.

For simplicity, the description of simulated datasets thus far has assumed that the model from which samples are drawn consists of a single binned energy spectrum, formed from the weighted sum of all simulated spectra. However, during fits to the DEMONSTRATOR's experimental data, both data and pdfs are subdivided into multiple binned spectra based on detector group, experimental configuration, and event multiplicity, as described in Sec. 4.1.2. Since fits to simulated datasets are designed to closely replicate fits to the DEMONSTRATOR's data, the simulated datasets are subdivided in the same way. The assay-based simulated dataset model is generated assuming that each submodel contains the same proportion of the overall exposure as it does in the DEMONSTRATOR's data. In the final simulated dataset, the fraction of events

in each submodel is designed to agree within statistical fluctuation with the fraction of events that submodel contains in the assay-based model.

4.3.1: High Statistics Studies

To test the performance of the fitting algorithm without being statistics-limited, fits were initially performed on simulated datasets containing many more events than the DEMONSTRATOR's data. To determine the number of samples to draw, the number of counts predicted by the assay-based model over the DEMONSTRATOR's DS0-7 exposure was multiplied by 1000. Based on this exposure and assay activities, $\sim 1.6 \times 10^8$ samples were generated over the full fitting energy range (between 100 and 2620 keV).

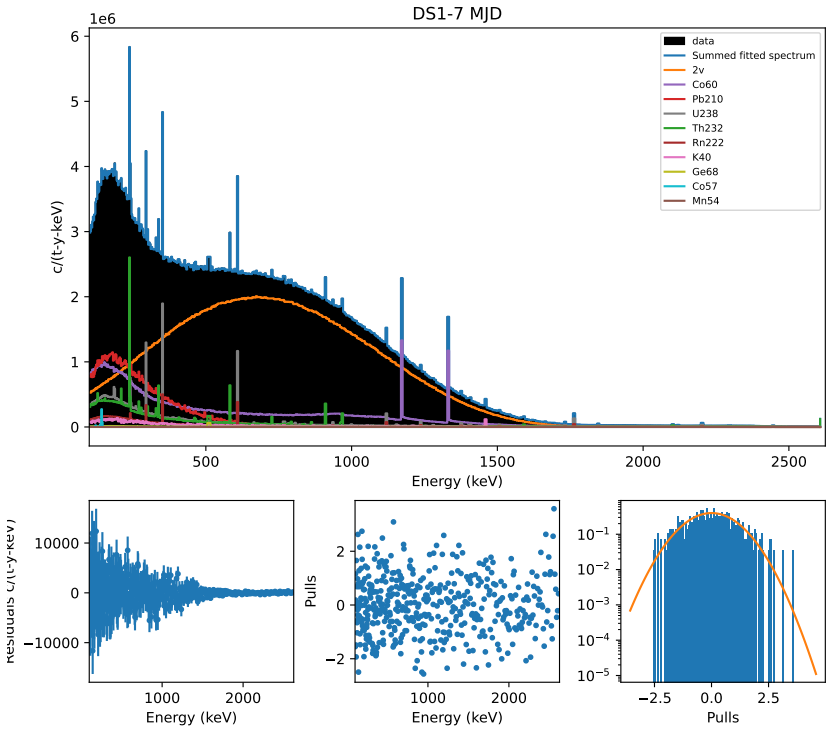


Figure 4.1: Comparison of a simulated dataset sampled from the assay-based model assuming 1000 times the DEMONSTRATOR's exposure (in black) and a fit to that simulated dataset (blue). The spectrum shown for each isotope is a sum over all pdfs for that isotope, each weighted by its fitted activity value. DS0 was simultaneously fitted but not included in the plotted results.

Fig. 4.1 show the results of a single fit to one simulated dataset generated in this fashion. The energy spectrum associated with the fitted model, shown in blue, demonstrates excellent agreement with the simulated dataset, shown in black. The fitted model spectrum is based on summing over all pdfs included in the fit, each weighted by its fitted activity density. The fit results for each decay chain are also plotted separately to show how much each contributes to the fitted model at different energies. They decay chain spectra,

plotted in color, are themselves combinations of many different pdfs representing different source locations. All plotted spectra, including the data spectrum, are also summed over many different submodels, meaning they combine many different detector groups and datasets that were separated during the fitting process. Submodels associated with the commissioning dataset, DS0, were included in the fit but are not incorporated in the plotted spectra.

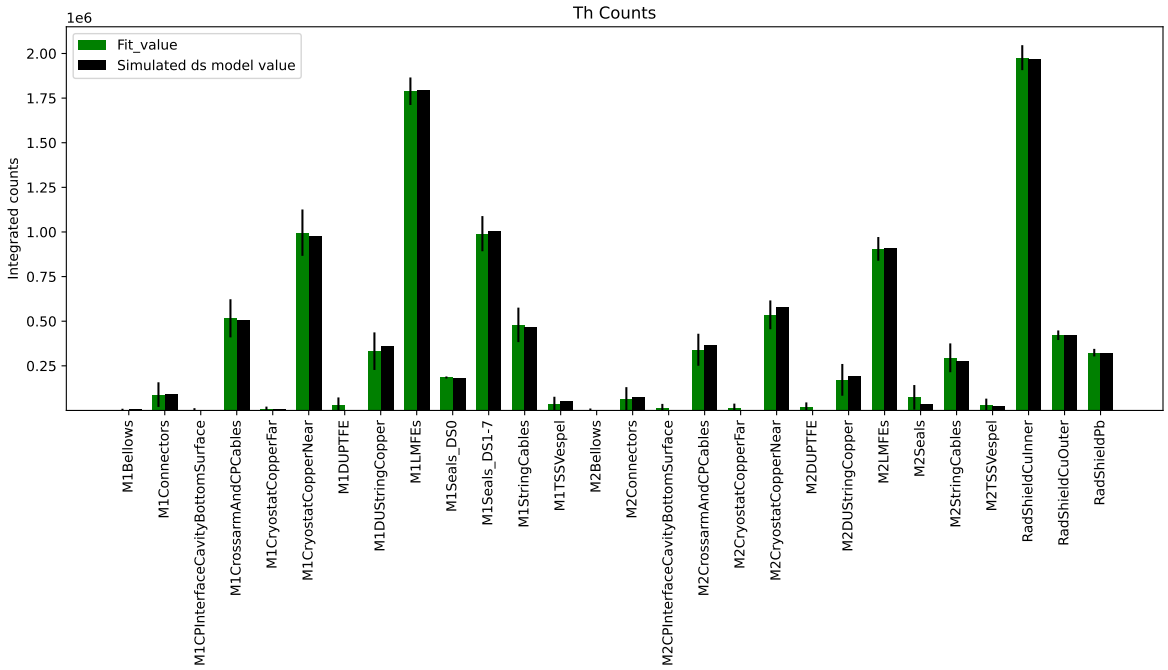


Figure 4.2: Comparison of expected vs. fit number of ^{232}Th counts in each component based on fits to 100 different assay-based simulated datasets, assuming 1000 times MAJORANA-level statistics.

While Fig. 4.1 clearly indicates the good spectral agreement between a simulated dataset and the model fitted to it, the real utility of simulated datasets is that they allow each fitted parameter to be compared to its "true" value, the value used to generate the simulated dataset. This could be done through a direct comparison of activity densities, but it is more illustrative to instead compare the total number of integrated counts fitted into each pdf. To understand why this is, consider the fact that, in most cases, a parameter fitting to a value an order of magnitude above the value used to generate the simulated dataset would be indicative of a problem. However, for a pdf that only contributed one count in the original model, a fitted activity density that is off by a factor of ten is still a very good result, since the pdf does not contribute significantly to the model in either case. Because pdfs associated with different experimental locations can have very different efficiencies, the pdfs that contribute significantly to the model are not immediately apparent through

comparing activity densities alone. To find the number of integrated counts for a pdf, the simulated spectrum (normalized by the number of primaries simulated) is multiplied by its fitted activity and the dataset runtime and is summed over all bins in the fitted energy range.

Bar charts act as a useful tool for visually comparing the integrated number of counts fitted to each pdf with the number of counts in that pdf in the assay-based model used to generate the simulated dataset. Instead of doing this comparison based on only a single fit to a single simulated dataset, 100 simulated datasets were drawn from the same underlying model, and each was fit once. Fig. 4.2 shows the result of this exercise for all ^{232}Th pdfs included in the fit. Each green bar indicates the average number of counts fitted into a pdf over all 100 trial fits. Its error bar corresponds to the standard deviation in the fitted number of counts over all trials. The number of integrated counts associated with each pdf in the simulated dataset model is shown in black. Comparing the green and black bars demonstrates that, on average, each floated parameter is fitting to approximately the correct value. There is no indication that certain parameters are systematically fitting to values that are too high or low. Based on the size of the error bar, at this level of statistics, the fitted parameter values do not display large variations between trials. Note that the central value plotted for each parameter is an average, meaning it performs better than an individual trial would. This leads to better agreement between the number of counts fitted into a component and its simulated dataset value than would be expected given the size of the error bars. In this plot, the error bars do not represent the expected uncertainty on the averaged value but the variation between trials. Similar plots were generated for other significant decay chains and cosmogenic isotopes, such as ^{238}U and ^{60}Co, and were found to show comparable levels of agreement.

These initial tests indicate that the fitter performs as designed and does not exhibit significant biases. Fits are shown to converge to the correct parameter values, indicating that the minimum NLL is being found successfully. At this level of statistics, the individual contributions of different pdfs can be clearly resolved, as shown by the small size of each error bar relative to the number of fitted counts for that component, meaning that the ^{232}Th background can be decomposed based on source location with high confidence. To study how the fitter performs in the case where statistics are more limited, as they are in the DEMONSTRATOR's data, these tests were repeated with lower statistics simulated datasets.

4.3.2: MAJORANA Statistics Studies

Simulated datasets with MAJORANA-level statistics were generated using the same procedure as was outlined for high statistics datasets but with a factor of one thousand fewer samples drawn. Before performing

fits to many different simulated datasets, an initial test was done by fitting the same simulated dataset one hundred times. The purpose of this test was to check whether the randomized start values for floated parameters impacted fitting results. Over all 100 trial fits, the best fit Barlow-Beeston NLL varied by less than 0.002, and the variation in fitted parameter values was also negligibly small. This study confirmed that fits to the same data successfully converge to the same minimum, regardless of the initial set of parameter values used. The best-fit NLL was also found to be lower than the NLL of the "true" model used to generate the simulated dataset, a good indicator that the minimum NLL is being successfully located. Fig. 4.3 shows an example spectral comparison between a MAJORANA-statistics simulated dataset and the model fitted to it. As in the high statistics case, the fitted model successfully reproduces the spectral shape of the simulated dataset, as evidenced by the lack of structure in the normalized residuals.

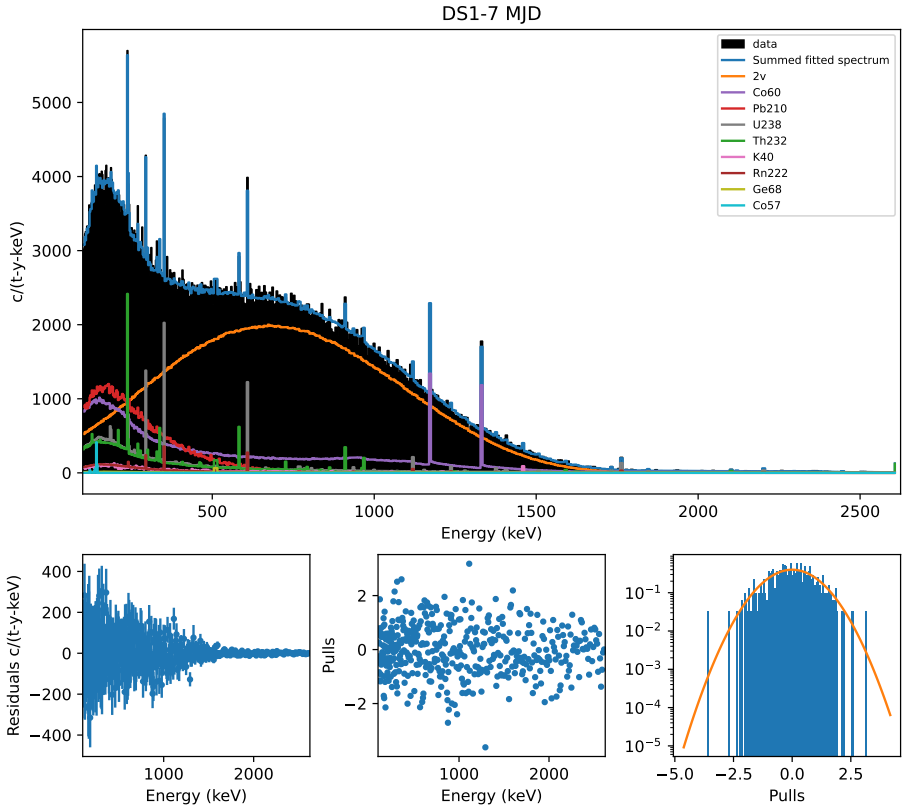


Figure 4.3: Comparison of a simulated dataset with MAJORANA-level statistics (in black) and a fit to that simulated dataset (blue). The spectrum shown for each isotope is a sum over all pdfs for that isotope, each weighted by its fitted activity value. DS0 was simultaneously fitted but not included in the plotted results.

To gain a fuller picture of fitter performance at this level of statistics, 100 MAJORANA-statistics datasets were drawn from the assay-based model, and each was fit once. The distribution of fitted activities over all

trials for each parameter demonstrates the expected amount of variation in the fitted value between repeated experiments. The distribution of $2\nu\beta\beta$ specific activities is shown in Fig. 4.4. The $2\nu\beta\beta$ distribution is fairly Gaussian and has a very small standard deviation of 2.1×10^{-7}, or 0.3% of the mean fitted activity density. This result is not indicative of the total uncertainty that can be expected for the DEMONSTRATOR's $2\nu\beta\beta$ result, since it reflects only statistical uncertainty. Unlike many others pdfs in the DEMONSTRATOR's background model, the $2\nu\beta\beta$ activity for enriched detectors contributes a large enough number of counts to the final model that its uncertainty was expected to be systematics-dominated. This study confirms the expectation that statistics is not a limiting factor in making a $2\nu\beta\beta$ half-life measurement with the DEMONSTRATOR's data.

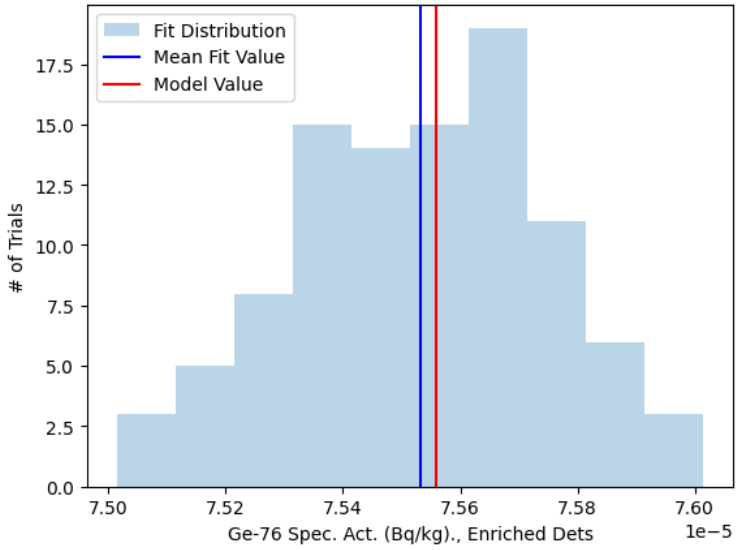


Figure 4.4: Distribution of fitted $2\nu\beta\beta$ specific activities from fitting 100 simulated datasets with MAJORANA-level statistics.

The overall ^{232}Th background rate is also relatively consistent between fits, demonstrated by the fact that the total number of fitted ^{232}Th counts has a standard deviation of only 5.5% over all 100 trials. However, the distribution of ^{232}Th source locations in the fitted model varies greatly between trials, as shown in Fig. 4.5. Although the average activity in each component is reasonably close to the model value, the large variation between trials, shown by the size of the error bars, indicates that this level of statistics is not sufficient to precisely distinguish the $^{232}\mathrm{Th}$ activity of each component or even to accurately determine which component groups are the dominant contributors to the $^{232}\mathrm{Th}$ background. It is notable that in the assay based model, no component contributes more than 2500 counts from the $^{232}\mathrm{Th}$ decay chain in the 100 keV-2620 keV energy region. This, in combination with the fact that the sum of all $^{232}\mathrm{Th}$ counts only account for $< 5\%$ of

the total counts in the assay model, makes resolving the composition of the ^{232}Th background challenging, particularly since some pdfs are highly correlated.

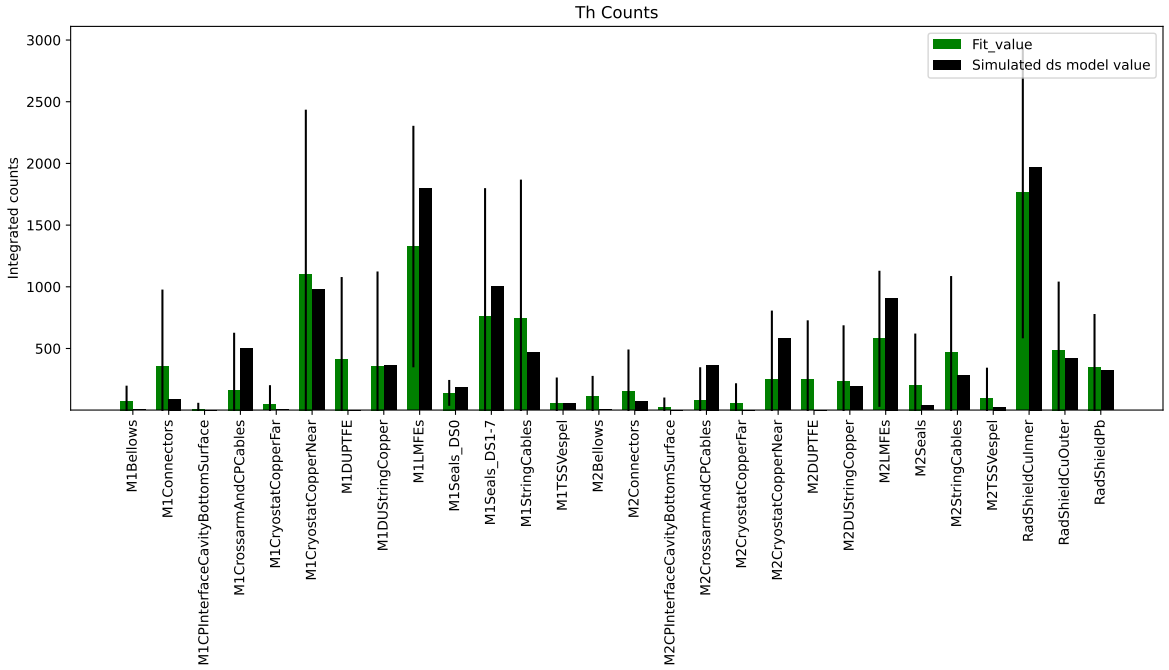


Figure 4.5: Comparison of expected vs. fit number of ^{232}Th counts in each component based on fits to 100 different assay-based simulated datasets at MAJORANA-level statistics.

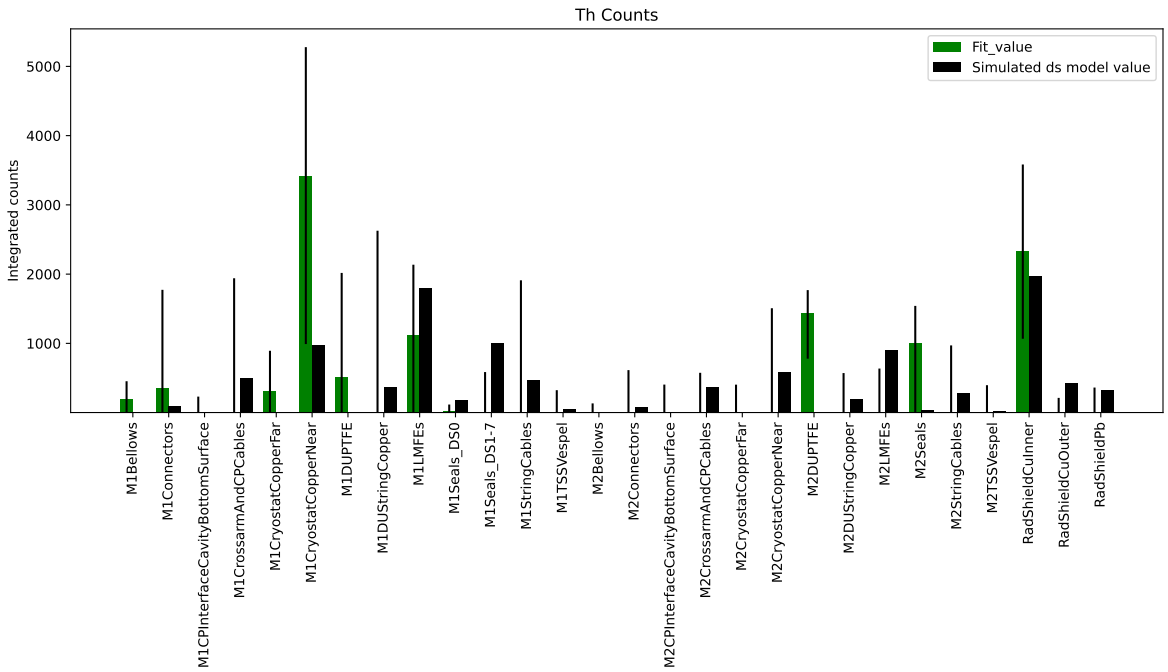


Figure 4.6: Comparison of expected vs. fit number of ^{232}Th counts from a single fit to an assay-based simulated dataset at MAJORANA-level statistics. The plotted error bars are based on the 1σ statistical uncertainties found by profiling the NLL.

Fits to simulated dataset also provide a good test case for evaluating the performance of the statistical uncertainty determination. Fig. 4.6 shows the composition of the ^{232}Th background in the fitted model for a single fit to a MAJORANA-level statistics simulated dataset. Because the central value of each bar is just based on a single fit, the agreement with the true parameter values is significantly worse than in Fig. 4.5 where the average over many trials was used. The error bars on the fitted values in Fig. 4.6 are the 1σ statistical errors found by performing a profile likelihood analysis for each floated activity density. In general, the sizes of the error bars are similar to those found in Fig. 4.5 by fitting many different simulated datasets. In addition, most but not all fitted activities are within 1σ of their true values. These two factors are a good indication that the statistical uncertainty evaluation was implemented correctly and produces reasonable results.

These studies validate the procedures for performing fits and calculating statistical uncertainty, but they also demonstrate the challenge of determining the composition of the ^{232}Th background at this level of statistics. However, the DEMONSTRATOR's data exhibits significantly higher ^{232}Th backgrounds than those predicted by the assay-based model, as discussed in Chapter 3. This means that the data contains more ^{232}Th counts than the assay-based simulated datasets, potentially allowing fits to data to have more discriminatory power. This difference could not be modeled perfectly, since the source location of the ^{232}Th excess in data

is not definitively known. Instead, multiple models assuming different excess sources were studied to test the fitter's ability to determine the source location of the DEMONSTRATOR's ^{232}Th excess.

4.3.3: ^{232}Th Excess Studies

The procedure for producing simulated datasets for the ^{232}Th excess studies was similar to the one used to generate MAJORANA-statistics simulated datasets, since most pdfs were still modeled at their assay-based activities. For the ^{232}Th excess studies, however, at least one pdf was scaled to a higher activity to make up the difference in the 2615 keV ^{232}Th peak height between data and the assay-based model. When the excess was shared between multiple components, the activity of each was adjusted so that the excess 2615 peak counts were split equally between them. This procedure was designed so that the size of the ^{232}Th contribution to each simulated dataset model would approximately match the ^{232}Th portion of the DEMONSTRATOR's data, regardless of where the excess was placed. For all source locations shown, fits were performed to 100 different simulated dataset drawn from the same underlying model, and the average and standard deviation over all 100 fits were used to evaluate how well the excess source could be identified.

For some excess source locations, spectral fits were able to clearly identify the component group responsible for the excess. For example, Fig. 4.7 shows the results of fits to simulated datasets where the lead shield was assumed to be entirely responsible for the difference in $^{232}\mathrm{Th}$ backgrounds between the assay model and data. Based on the size of the error bar on the fit number of RadShieldPb counts, it is evident that a uniformly distributed ^{232}Th excess in the lead shield can consistently be identified correctly. Fig. 4.8 shows a similar result for simulated datasets where the excess was assumed to be entirely located in the Module 1 thermosyphon cavity. Given that the M1 thermosyphon cavity was identified as the best candidate to explain the localized Module 1 background, as detailed in Chapter 3, this finding is an important result, indicating that the fitter would be able to identify an excess in this location with low uncertainty.

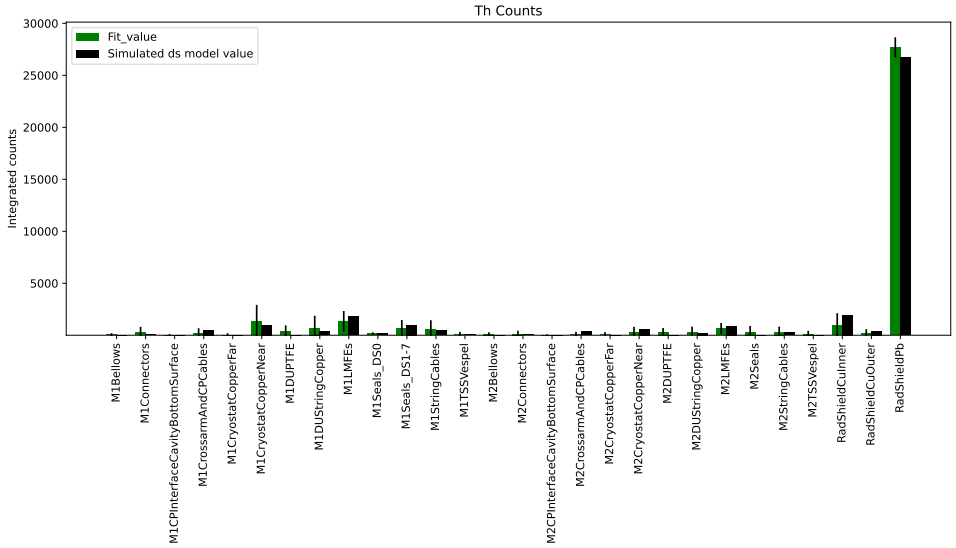


Figure 4.7: Comparison of expected vs. fit number of ^{232}Th counts in each component based on fits to 100 different simulated datasets, where the statistics are the full DS0-7 exposure and the activities of each pdf in the simulated dataset model were based on assay projections, except for the lead shield, which was given a large enough activity to account for the number of excess counts in the 2615 peak in data.

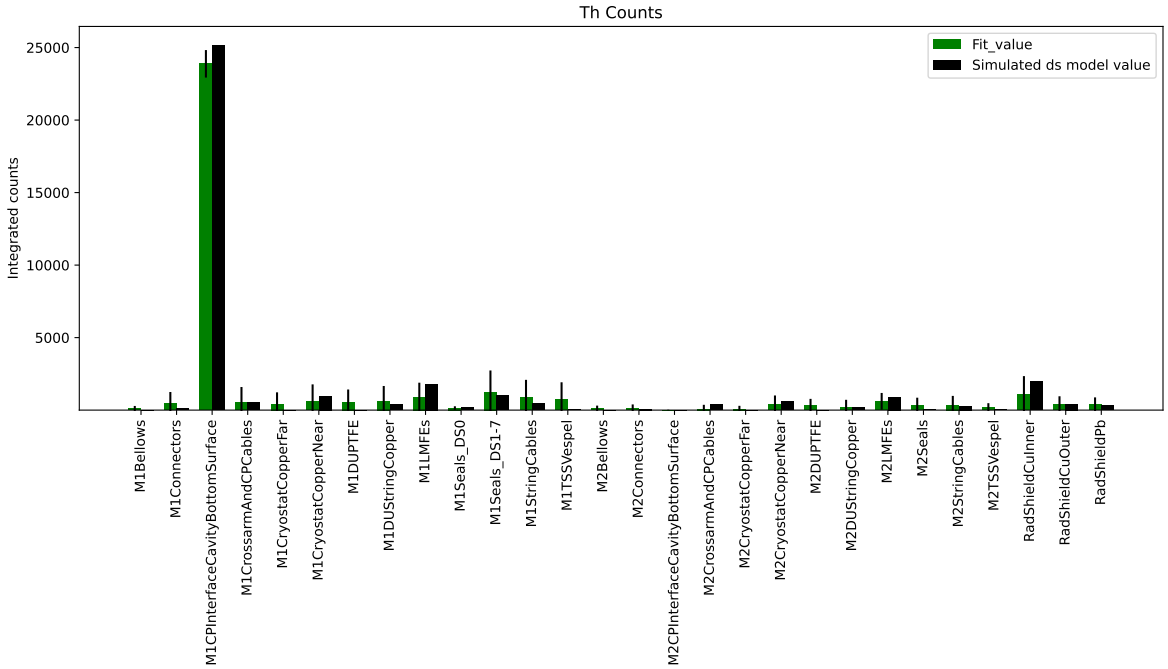


Figure 4.8: Comparison of expected vs. fit number of ^{232}Th counts in each component based on fits to 100 different simulated datasets, where the statistics are the full DS0-7 exposure and the activities of each pdf in the simulated dataset model were based on assay projections, except for the M1 thermosyphon cavity, which is given a large enough activity was account for the number of excess counts in the 2615 peak in data.

Not every source location of a ^{232}Th excess could be quite as definitively determined by the fitter.

Fig. 4.9 shows an example of one such case. When the ^{232}Th excess was modeled as originating in the crossarm and coldplate cables in both modules, the fit results varied significantly for different simulated datasets, as indicated by the large size of the error bars compared to the average number of counts fitted into the component. The average over all 100 trial fits also underestimated the total number of counts from the crossarm and coldplate cables while overestimating the number of counts originating from the connectors. This result can be understood by the fact that these two components are located in similar locations, leading to pdfs that are not identical but are highly correlated, making it difficult for the fitter to distinguish between them at this level of statistics.

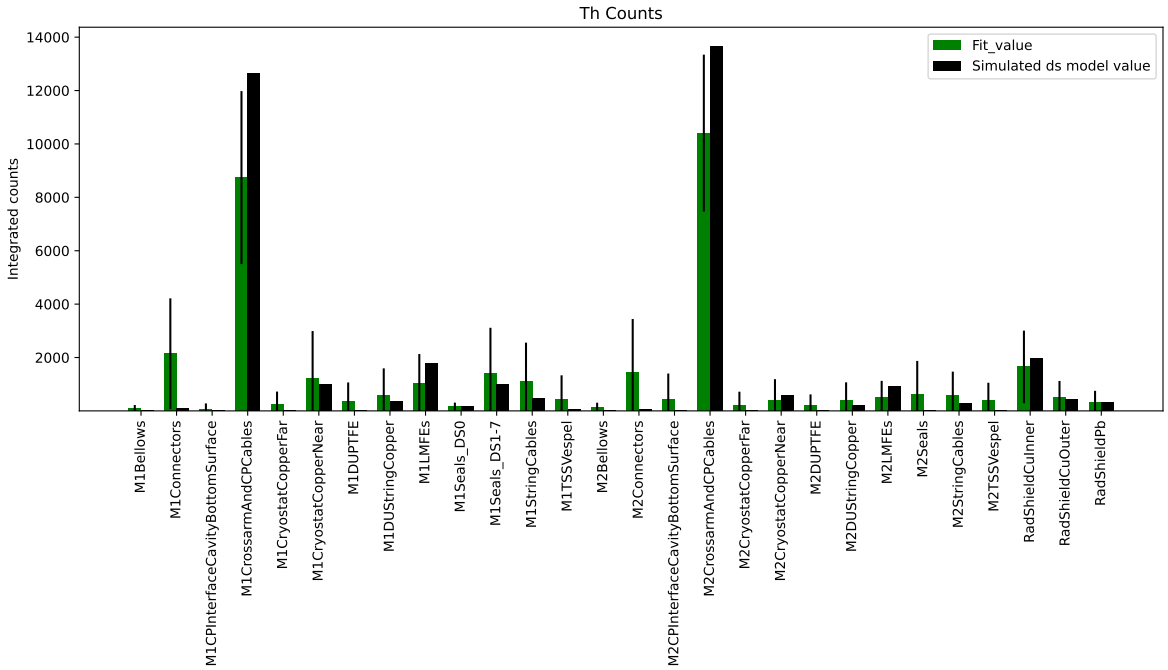


Figure 4.9: Comparison of expected vs. fit number of ^{232}Th counts in each component based on fits to 100 different simulated datasets, where the statistics are the full DS0-7 exposure and the activities of each pdf in the simulated dataset model were based on assay projections, except for the cables located in the crossarm and coldplate, which were given a large enough activity to account for the number of excess counts in the 2615 peak in data.

Although an excess in the coldplate and crossarm cables could not be pinpointed as definitively as some source locations, the excess counts were still attributed to the right overall region of the experiment. This can be visualized by regrouping the counts based on source region, as was done in Fig. 4.10. In this plot, the 'Near' groups include components inside the cryostat that are below the coldplate, while the 'Middle' groups consists of the cryostats, thermospyphons, and components above the coldplates. Fig. 4.10 was generated based on the same 100 fits included in Fig. 4.9. For each fit, the counts from components within the same

region were summed together. The distribution of summed counts for a group over all simulated datasets was then generated in order to calculate its mean and standard deviation, which manifest in the plot as the colored bar and its error bar. This technique provides a method for understanding the overall behavior of a group without having to assume that all components in the group have the same activity density.

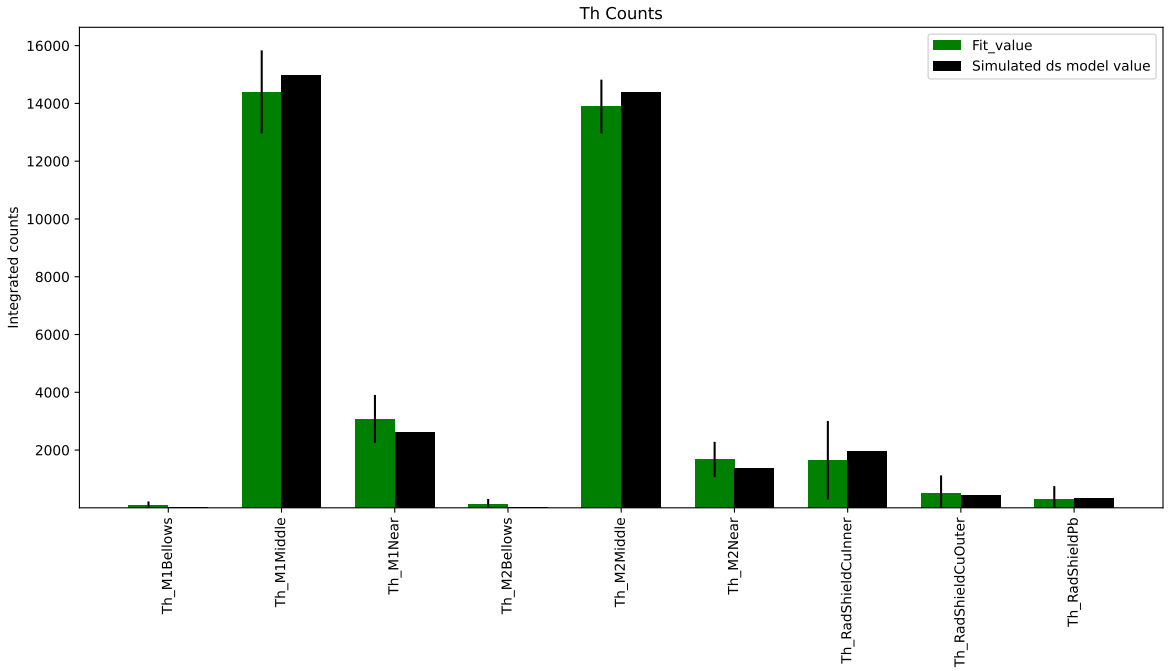


Figure 4.10: Regrouping of the results shown in Fig. 4.9 to emphasize the regions of the experiment where ^{232}Th backgrounds originated in fits to simulated datasets with an excess in the coldplate and crossarm cables. The Near groups included all components inside the cryostat and below the coldplate.

This procedure cannot be directly extended to data, since there is not a straightforward procedure for combining the profile-likelihood-based uncertainties for multiple components while properly taking into account the large correlations between them. Nevertheless, even without applying this post-facto grouping to fits of the DEMONSTRATOR's data, this study provides useful context for interpreting the results of those fits. In almost all cases of fitting simulated datasets with ^{232}Th excesses at the level observed in data, the general region of the excess could be definitively determined, even when the exact component could not be concluded. The rare exception to this finding occurred for the cryostat copper (CryostatCopperNear), which did not fit neatly into either group. Although the cryostat copper was grouped with middle components, it shares some similarities to the near-detector components based on its direct line of sight to the detectors. It is probably best matched by a linear combination of near and middle group components. Otherwise, there is not a large amount of overlap between counts attributed to the near and middle groups. This result demonstrates

that, if the ^{232}Th excess is largely dominated by a single component, spectral fits to the DEMONSTRATOR's data have a high probability of successfully locating the general region of the background source. In some cases, the exact component can be conclusively determined, but even when the uncertainty on the dominant component is large, fits can typically determine whether the excess originates in a near-detector component.

Overall, fits to simulated datasets effectively validated the procedures for performing spectral fits and calculating statistical uncertainties outlined in this chapter. They have also shown that the DEMONSTRATOR's DS0-7 exposure has sufficient statistics to determine the $2\nu\beta\beta$ half-life with reasonable statistical uncertainty and to draw reliable conclusions about the region from which the DEMONSTRATOR's dominant ^{232}Th source originates. However, the approaches discussed in this chapter do not take into account the systematic uncertainties associated with modeling the DEMONSTRATOR, particularly modeling the effects of detector response on spectral shapes and efficiencies. To determine the effect of uncertainties in the simulation post-processing on the results of spectral fits, a framework for evaluating systematic uncertainties is required.

CHAPTER 5: Framework for Quantifying Systematic Uncertainty

Drawing meaningful and reliable conclusions from background model fits to the MAJORANA DEMONSTRATOR's data requires a robust method for estimating the full uncertainty of each fitted parameter. While Sec. 4.2 details how the profile likelihood method is used to calculate the statistical component of the uncertainty, this method does not incorporate the systematic contribution. Many sources of systematic uncertainty exist when developing a precision model of experimental results, and not all sources can be quantified with the same procedure. However, multiple important sources of systematic uncertainty for the DEMONSTRATOR relate to simulation post-processing, and these uncertainties can be evaluated using a common framework.

As described in Sec. 3.4.3, simulation post-processing is the procedure by which details of the detector response, such as the energy resolution or the dead layer profile, are incorporated into the simulations. Each detector response effect is described by one or more parameters. The value of each parameter was typically determined from auxiliary data not included in background model fits, such as ^{228}Th calibration runs. However, each systematics parameter has a degree of uncertainty on its measured value, and the effect of this uncertainty must be propagated to the results of background model fits. Since systematics parameters affect both the efficiency and shape of each simulated spectrum, varying their values can affect fitting results in complex and unpredictable ways. To properly take these effects into account, pdfs must be generated using different values of the systematics parameters, and fits must be performed with these different sets of pdfs.

The general approach taken in the DEMONSTRATOR's systematics framework is to randomly sample the uncertainty distribution of each systematics parameter. For each random sample, a full set of pdfs is generated and used to fit the DEMONSTRATOR's data. The distribution of a parameter of interest over fits performed with different sets of systematics pdfs is then used to quantify the uncertainty on that parameter. This procedure can be applied to determine the systematics uncertainties on all activity densities floated during spectral fits, although in this work it is mainly discussed in the context of determining the systematic uncertainty associated with the $2\nu\beta\beta$ half-life.

Sections 5.1 and 5.2 briefly summarize the systematic parameters considered in this analysis and discuss the algorithm for sampling systematics parameters to generate new sets of pdfs. Section 5.3 gives more detail about how the results of repeated fits with different systematics pdfs are used to extract a combined

uncertainty for all systematics parameters. Finally, Section 5.4 discusses how to determine which systematics parameters have the greatest impact on the combined uncertainty.

Not all sources of systematic uncertainty can be evaluated using this framework. For example, each component of the DEMONSTRATOR's experimental apparatus had some design tolerance, meaning that its volume in the MAGE geometry could differ slightly from the volume of the manufactured part. This source of uncertainty must be evaluated by generating multiple simulations with different MAGE geometries, rather than simply applying different post-processing parameters to the same set of simulations. Since generating new simulations is much more computationally intensive than repeating post-processing, it is not feasible to quantify uncertainty contribution using a variation of the approach detailed in this chapter. Sources of uncertainty that do not fit into the post-processing framework will be discussed individually in the context of the $0\nu\beta\beta$ result in Chapter 6.

To illustrate how the systematics framework is applied, fits to calibration data are used as an example throughout this chapter. These fits utilized data from a single M1 long calibration run following data cleaning and surface alpha cuts. Calibration data has the advantage of possessing high event rates dominated by a single background source with a relatively well-known position. This factor was useful for framework validation and allowed studies to be performed with only a limited set of pdfs, enabling initial testing of the systematics framework to use limited computational resources. In addition to calibration line source pdfs for both modules, the fits included pdfs for radon in the nitrogen and for ^{232}Th in the copper of each cryostat. The M1 ^{232}Th cryostat copper pdf was particularly chosen as a source that was likely to have a similar spectral shape to the calibration line source pdf.

Section 5.1: Systematic Parameters

The values of a number of post-processing parameters are determined from external measurements with associated uncertainties. Some parameters that have been identified as potentially important to the results of background model fits are summarized briefly below.

5.1.1: Transition Dead Layer Profile

As was described in Sec. 3.4.3, events that deposit energy in the transition layer adjacent to a PPC's n+ contact exhibit partial charge collection. The fraction of the total charge that is detected varies as a function of the distance from the detector surface at which the interaction occurs. The functional relationship is modeled by an exponential region and a linear region, parameterized by a full charge collection depth and

two transition layer shape parameters. Each enriched detector's full charge collection depth was measured to within a 15% systematic uncertainty by the detector manufacturer, ORTEC, using a collimated ^{133}Ba source [81]. Since a detector's full charge collection depth can grow when it is stored at room temperature, this 15% uncertainty was combined in quadrature with a one-sided uncertainty reflecting the amount of time between each detector's ORTEC measurement and the cooling of the detector array. The full charge collection depth was not measured directly for natural detectors, so all natural detectors were modeled as having a 1 mm full charge collection depth. A 22.5% uncertainty in the natural detector FCCD was chosen to account for the level of variation observed between the FCCDs of enriched detectors. For the purposes of estimating systematic uncertainties, the FCCDs of enriched and natural detectors were varied separately, since their uncertainties were not expected to be highly correlated.

The values for the two shape parameters and their uncertainties were determined by fits to ^{228}Th calibration data performed by I. Guinn [81]. The low energy spectral shape and the tails of gamma peaks in the calibration data encode information about the dead layer profile, due to the energy degradation experienced by transition layer events. The two shape parameters are not independent, so their covariance matrix was also saved for use when varying the parameters.

5.1.2: Delayed Charge Recovery Survival Fraction

Since the DCR cut is applied to data prior to fitting to eliminate surface alphas, it is important to quantify how often non-surface events are incorrectly tagged and to adjust simulated spectra based on this efficiency. For single-site events, the bulk event survival fraction is very high (~99%) and relatively uniform as a function of energy, but for multi-site events, the survival probability displays some energy dependence [81, 82]. Therefore, the DCR survival fraction is parameterized by an overall single-site efficiency and by three Gaussian parameters capturing the additional energy-dependent rejection of multi-site events. The values and uncertainties of all four parameters were determined for each detector using ^{228}Th calibration data. The parameter uncertainties were treated as Gaussian and were quantified by studying the magnitude of effects like the time variation in parameter values over many calibration runs.

5.1.3: Energy Peak Shape and Non-linearity

Energy resolution in one of the most important effects incorporated during simulation post-processing, but energy-related systematics are not expected to have a large impact on pdfs due to the variable-width binning scheme, which was designed to almost fully contain each peak in a single bin. However, uncertainties

in energy parameters were still incorporated into the full systematics uncertainty through varying peak shape and energy non-linearity parameters. The peak shape function includes both a Gaussian component and a low energy tail. The energy-dependent sigma and tail parameters were both varied to explore the impact of uncertainties in the full width half max value. The energy non-linearity is described by two parameters, one of which is a phase factor that can take on any value between 0 and 2π [81].

Section 5.2: Varying Parameter Values

Since systematic parameters are based on measured values with corresponding uncertainties, each parameter can be treated as a distribution. In most cases, this distribution is assumed to be a Gaussian centered on the optimal parameter value. The standard deviation of the distribution is based on the uncertainty in the measurement used to determine the central value. Systematic parameters are essentially nuisance parameters in the DEMONSTRATOR's spectral fits, so in principle they could be allowed to float during fitting, with the parameter distributions from external measurements acting as penalty term to better constrain their values. However, since varying a systematics parameter requires regenerating a full set of pdfs, it is not computationally feasible to do so during each step in the minimization process. Instead, to explore the systematics parameter space, systematics parameters were sampled from their uncertainty distributions prior to fitting. This approach was chosen because the DEMONSTRATOR's low background dataset contains little information about the true values of systematics parameters and is unlikely to be able to constrain them accurately beyond their nominal uncertainties from auxiliary measurements.

For a given type of systematics parameter, such as the energy resolution, each detector has a different optimal value and uncertainty, but systematic errors are assumed to be correlated between detectors. To sample the detector-specific parameters in a correlated way, a sample is first drawn from a normal distribution with a mean of zero and a standard deviation of one. This sample, which will be referred to as the adjustment factor, is then multiplied by the detector-specific uncertainty and shifted by the detector-specific central value to transform the sample to a parameter value for each detector. In cases where systematics parameters are measured for each dataset, the central values and uncertainties used are dataset-specific as well as detector-specific. This technique is modified slightly for systematics described by multiple correlated parameters and parameters with non-Gaussian distributions, such as the correlated transition layer shape parameters and the asymmetric dead layer fccd parameter [81].

To take into account possible correlations between different systematics effects when determining the full

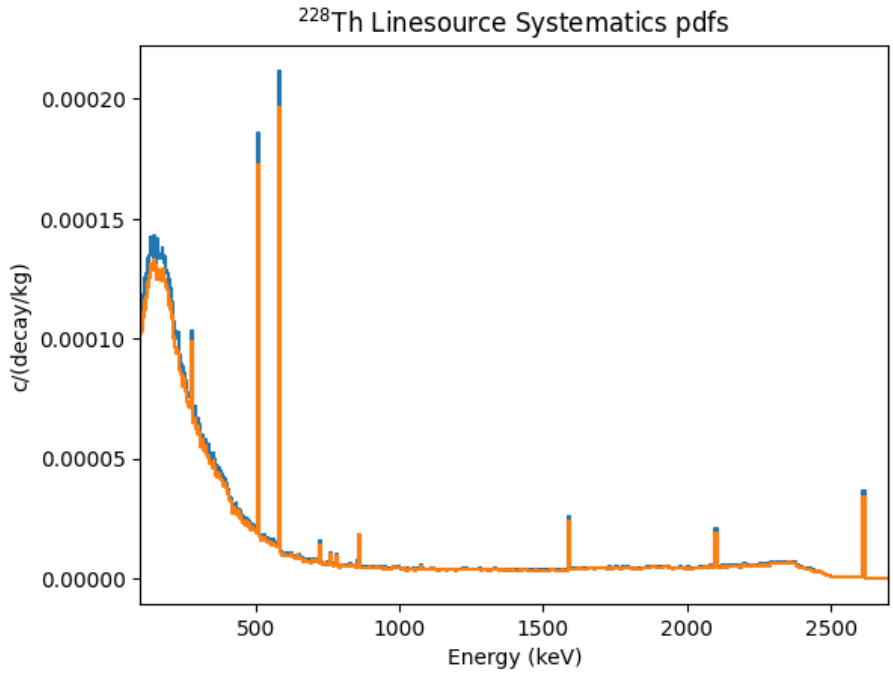


Figure 5.1: Comparison of two pdfs of the ^{228}Th M1 line source pdf generated with different systematics configurations. Multiple systematics effects are varied independently between the two pdfs, leading to a difference in efficiency and subtle differences in spectral shape

post-processing systematic uncertainty, the parameters describing these effects were varied simultaneously but independently. This means that to produce a single set of combined systematics pdfs, a separate adjustment factor was drawn for each systematics parameter and applied during post-processing. The full collection of adjustment factors used to generate a set of pdfs, known as the systematics configuration, was saved for use when analyzing fitting results.

Fig. 5.1 gives an example of two systematics pdfs produced with different sets of systematics parameters. The figure shows two different pdfs for the M1 ^{228}Th calibration line source, where the adjustment factors for each systematic parameter were sampled simultaneously but independently. The overall efficiency differs notably between the two pdfs, but there are also subtle differences in spectral shape that can impact fitting results.

Section 5.3: Combined Systematic Uncertainty Calculation

Once pdfs have been generated with many different combined systematics configurations, each set of pdfs is independently fit to the data. For any of the specific activities floated during fitting, the different best fit values found with different systematics pdfs form a distribution. The 1σ combined systematic uncertainty on a given activity density is found by taking the central 68% of that activity density's distribution. An

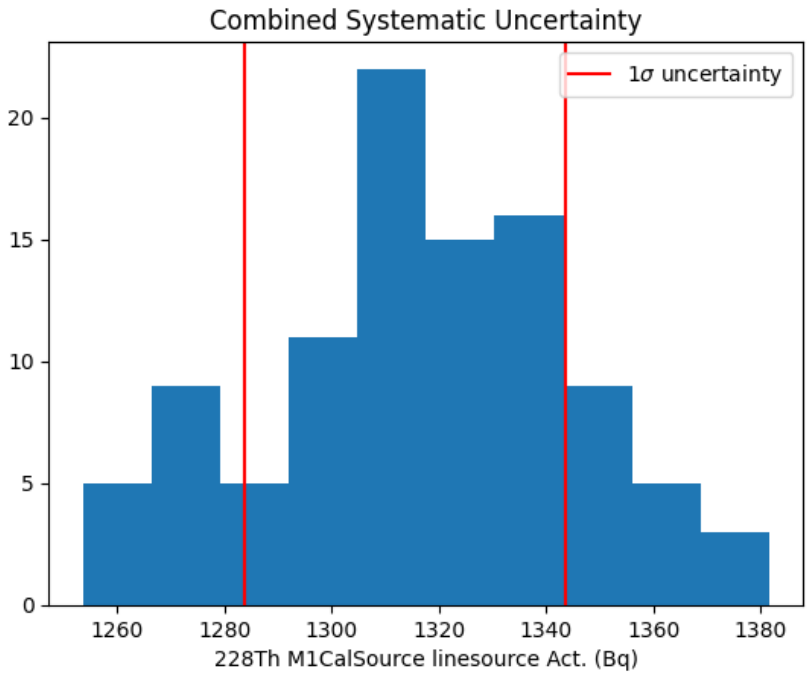


Figure 5.2: Distribution of fitted M1 ^{228}Th line source activities from fits to a Module 1 long calibration run. Each of the 100 fits was performed using a different set of systematics pdfs, where the systematic adjustment factors for different post-processing effects were independently varied.

example of this procedure is shown for the M1 ^{228}Th line source activity from fits to Module 1 calibration data in Fig. 5.2. Since all systematics parameter were varied when generating the systematics pdfs, the resulting uncertainty is a combined systematic uncertainty.

For fits to background data, an initial version of this study was performed with 100 systematics configurations, allowing the framework to be validated and initial conclusions to be drawn. Final uncertainty estimates will be based on repeating this study with 1000 systematics configurations, since the uncertainty on the systematic uncertainty goes as the square root of the number of configurations.

Section 5.4: Identifying Dominant Contributors to the Systematic Uncertainty

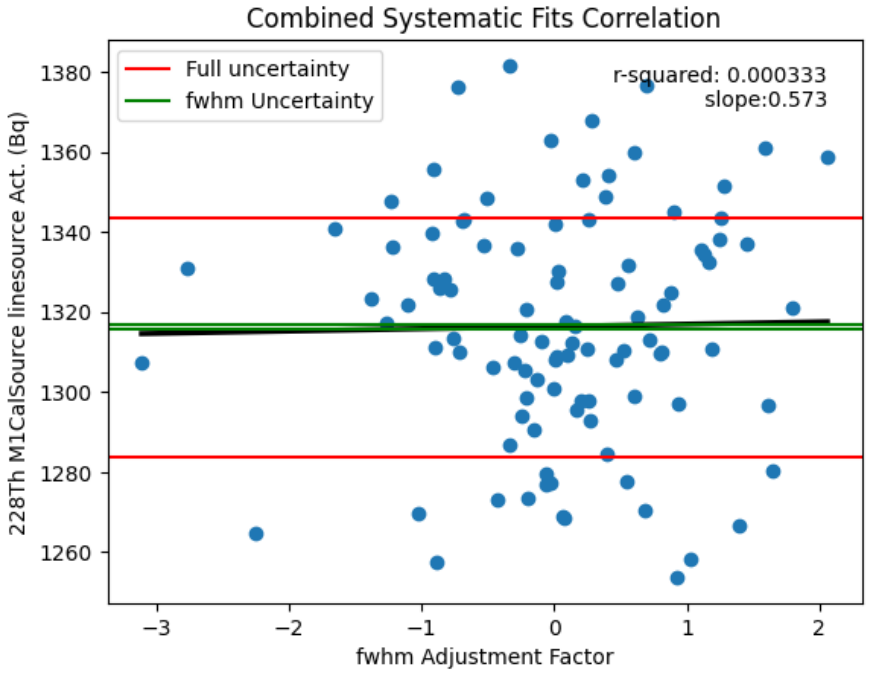
Although evaluating the combined uncertainty from all systematics parameters simultaneously allows correlations to be properly taken into account, it is also of interest to determine which parameters contribute most significantly to the overall uncertainty. In addition to the sets of combined systematics pdfs generated for the full uncertainty calculation, sets of pdfs varying only a single systematic were generated.[1] For each systematic effect considered, one hundred sets of pdfs were generated. In the future, fits performed with

[1]For systematics parameters that were varied in a correlated manner, such as the dead layer shape parameters, pdfs were generated varying both parameters.

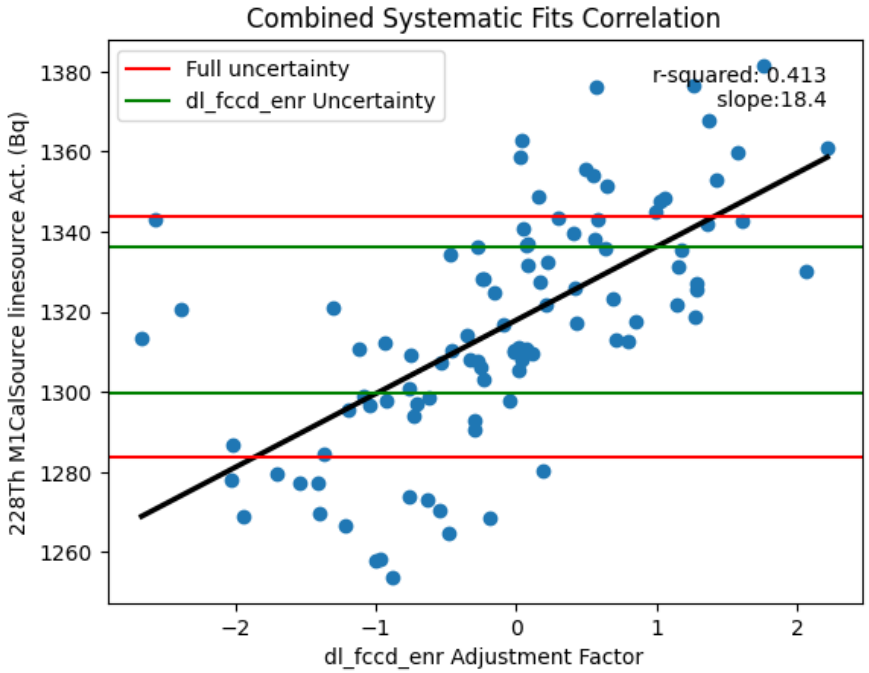
these pdfs will be used to quantify the contribution of each systematic effect. When only a single systematic is varied, the central 68% of fitted results can be used as an estimate of the uncertainty associated with that systematic.

However, the fits from Sec. 5.3 where all systematic parameters were varied can also be used to identify the dominant contributors to an activity's overall uncertainty. If a fitted parameter of interest is highly correlated with a systematic nuisance parameter, that systematic parameter contributes significantly to its total uncertainty. To determine the degree of correlation, a scatter plot is generated comparing the fitted activity to the systematic parameter's adjustment factor. For the example of fits to M1 calibration data, Fig. 5.3a plots the fitted ^{228}Th line source activity against the energy resolution (fwhm) adjustment factor. Each point on the plot is the result of a single fit with a different set of pdfs. The y-value of the point is based on the results of the fit, while its x-value is based on the systematics pdfs used to perform the fit. In this case, the adjustment factor on the x-axis shows how many standard deviations away from the optimal value the pdfs' energy resolution is. A linear fit to the plotted points is shown in black. A significantly non-zero slope is indicative of a systematic parameter with a large impact on the fit result. Fig. 5.3a displays no significant correlation, meaning that the results of fits to calibration data were not strongly impacted by the assumed energy resolution. This result was expected, since the variable binning scheme used in pdf generation is relatively insensitive to small changes in peak width. In contrast, Fig. 5.3b demonstrates that the fitted line source activity is highly correlated with the dead layer full charge collection depth. This result makes sense, because as the size of the dead layer activity increases, a higher source activity is required to explain the same number of counts in the data.

To estimate the magnitude of each systematic parameter's contribution, the fitted correlation lines were evaluated at adjustment factors of ±1, plotted as horizontal green lines in Fig. 5.3. This gives an estimate of the uncertainty due to the parameter plotted on the x-axis. The magnitude of a single systematic parameter's contribution to the uncertainty can be compared to the combined uncertainty for all systematics parameters, shown in red. Both Fig. 5.3a and Fig. 5.3b are based on the same set of one hundred fits. In both cases, if the plot is projected onto the y-axis, we are left with the histogram from Fig. 5.2, where the combined uncertainty is based on the central 68% of points. These type of correlation plots allow the dominant contributors to the combined uncertainty to be identified and show what factors do not have a noticeable effect.



(a) Full width half maximum (fwhm)



(b) Dead layer full charge collection depth in enriched detectors (dl_fccd_enr)

Figure 5.3: Correlations plots showing how the fitted M1 calibration line source activity varies as a function of systematics adjustment factor in fits to M1 ^{228}Th calibration data. Both plots were generated from the same set of 100 fits where all systematic adjustment parameters were independently varied.

Section 5.5: Summary

The procedure described in this chapter allows uncertainties in detector response effects applied during simulation post-processing to be incorporated into fitting results. Using this method, multiple uncertainty

sources can be considered simultaneously to properly account for possible correlated effects. Correlation plots are then used to determine which effects most influence the combined result. In the next chapter, the techniques demonstrated on calibration data in this chapter are applied to the DEMONSTRATOR's low background data, particularly the $2\nu\beta\beta$ result.

CHAPTER 6: Fits to DS0-7 Background Data

The DEMONSTRATOR's data-driven background model is based on applying the fitting framework outlined in Chapter 4 to data taken between June 2015 and August 2020. This chapter details the set-up and results of spectral fits, with an emphasis on the conclusions that can be drawn from the fitted model and our level of confidence in them. Section 6.1 gives an overview of the data included in the final fits, including what cuts were applied prior to fitting and the energy range utilized. A high-level overview of the results is detailed in Section 6.2. Spectral comparisons give a visual representation of the level of agreement between data and the fitted model and show which decay chains contribute most significantly to the model over a range of energies. The agreement between model and data is then fully quantified in Section 6.3, where the negative log likelihood is compared between the fit to data and fits to simulated datasets to determine the goodness of fit. In Section 6.4, the fitted model's performance near $Q_{\beta\beta}$ is evaluated and the composition of observed backgrounds in this region is discussed.

The second half of the chapter focuses on the conclusions that can be drawn when the fitted model is subdivided not just based on decay chain but also on source location. Section 6.5 gives an overview of where in the experiment the predominant backgrounds originated. Section 6.6 examines the ^{232}Th excess in more detail and summarizes how the inclusion of a hot spot within the M1 thermosyphon chamber improves the model. Multiple systematic studies are then performed in Section 6.7 to determine how robust the fit is to changes in the fit setup, such as differences in the energy range used during fitting.

Finally, Section 6.8 reports the fitted $2\nu\beta\beta$ half-life and quantifies statistical and systematic uncertainty sources. The systematics uncertainty includes contributions from uncertainties in detector parameters such as the energy resolution and dead layer thickness, evaluated using the framework outlined in Chap. 5, as well as other sources of uncertainty, such as uncertainties in the simulated MAGE geometry. The impacts of the systematics checks discussed in Sec. 6.7 on the calculated $2\nu\beta\beta$ half-life are also disucssed.

Section 6.1: Fit Details

The DEMONSTRATOR's final background model fits include physics data taken prior to the M2 hardware upgrade (DS0-6) and data taken with Module 1 while Module 2 was being upgraded (DS7). DS5a, which

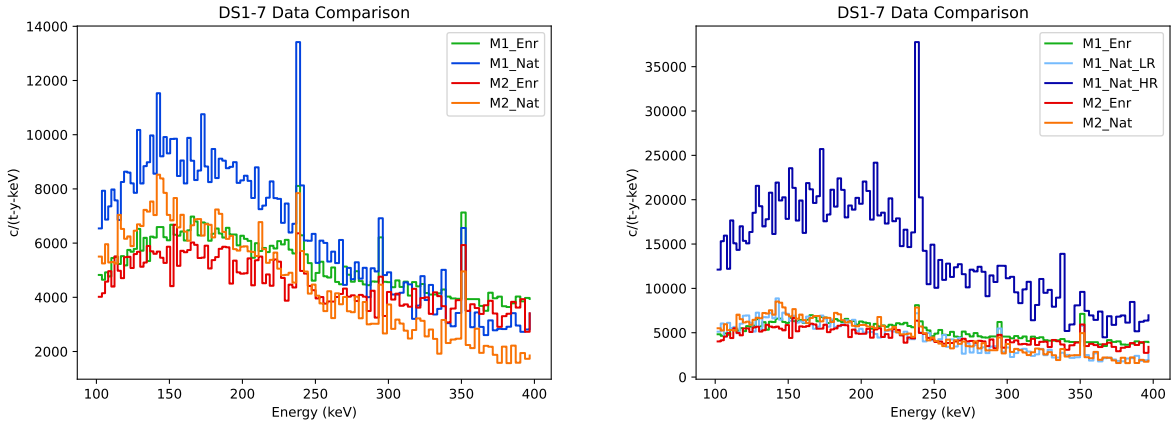


Figure 6.1: Comparison of the data spectra in the energy range between 100 and 400 keV if the two high rate detectors are included as part of Module 1 Natural detector group (left) or are separated into their own group (right). LR and HR refer to the low rate and high rate groups respectively.

consists of a few months of data with anomalously high electronic noise, was omitted because of its degraded energy resolution. Data taken following the Module 2 upgrade and after the removal of enriched detectors (DS8 and DS9) informed background studies but were not used for spectral fitting, since the replacement of multiple components during the upgrade would have increased the dimensionality of fits. As a consequence, none of the data taken with ICPC detectors are included in the background model. Data that were previously blinded for the main $0\nu\beta\beta$ result and other physics analyses are included in the fits described here. While fits were initially performed on open data before adding in unblinded data, the construction of the background model was not treated as a fully blind analysis.

The combination of open and unblinded data used for fitting corresponds to 51.9 kg-yrs of enriched exposure (39.3 kg-yrs in M1 and 12.6 kg-yrs in M2) and 22.4 kg-yrs of natural exposure (10.3 kg-yrs in M1 and 12.1 kg-yrs in M2). Based on the finding that the two natural Module 1 detectors near the crossarm displayed significantly hotter background rates than other Module 1 natural detectors, these detectors were separated into their own group during fitting, despite only comprising 2.07 kg-yrs exposure. Fig. 6.1 shows the impact of grouping the detectors in this way.

Prior to performing fits, data cleaning cuts and the DCR cut were applied to the data. During fitting, the hit spectra from multiplicity one and high multiplicity data were separated for each dataset and detector group. Single-site and multi-site events were not separated, since the dt heuristic is well-validated in the $0\nu\beta\beta$ region of interest but may not be well modeled over the entire fit range. Similarly, the LQ cut was not applied to data, because its efficiency is not fully quantified at low energies and because transition layer

events are not removed when generating pdfs from the simulations.

Standard background model fits employed a 100 keV threshold in order to minimize the impact of surface effects that are challenging to precisely model. Although the DEMONSTRATOR's low energy program developed a low energy analysis toolkit for rejecting unphysical and surface waveforms down to ~1 keV, fully incorporating the energy-dependent bulk waveform survival efficiency and its uncertainty would greatly increase the complexity of the background model. Analyses performed by the low energy group also excluded periods with high rates of low energy events, leading to the rejection of some exposure deemed acceptable for analyses that do not use low energy data. Spectral fits spanning both energy ranges would be required to either sacrifice a portion of the exposure across all energies or to split the spectrum arbitrarily into different energy ranges with different exposures.

The highest energy bin included during fitting is the bin containing the 2615 keV ^{232}Th peak. This high energy threshold was chosen because above the 2615 peak, the event rate in data is low and appears to be dominated by energy-degraded alphas. In enriched detectors, the energy spectrum between 2630 and 4500 keV is fairly flat, and approximately 60% of the waveforms that remain after a DCR cut are tagged as either LQ or high AvsE waveforms, indicating that the region is dominated by surface events. Since no pdf for energy-degraded alphas is included in the fit, including bins from a portion of this energy range during fitting would skew the results.

Section 6.2: Overview of Results

The data-driven background model produced by spectral fits reproduces the data well across a wide range of energies. Fig. 6.2 shows the spectral fit results for enriched detectors compared to data from the same detectors. The commissioning dataset, DS0, is excluded from this plot because its reduced shielding resulted in higher backgrounds, but DS0 data informed the plotted model through its inclusion as a submodel during simultaneous fits. The modeled spectrum for each decay chain in Fig. 6.2 is produced by summing all pdfs from that decay chain, weighted by their fitted specific activities. The summed fitted model in blue includes contributions from all decay chains, meaning it can be directly compared to the DEMONSTRATOR's data in black. Figs. 6.3-6.5 zoom in on the low, middle, and high energy scales of the fitted spectrum. Versions of Fig. 6.2 split by module and versions including natural rather than enriched detectors can be found in Appendix D.

Spectral comparisons allow for a number of broad conclusions to be drawn about fitting results. First,

some structure remains in the normalized residuals of these plots, particularly at low energy, indicating that the model does not perfectly describe all features of the data. While the deviations between data and model are not large, they do introduce an uncertainty into the fitted results that is discussed at length in future sections. Second, Fig. 6.5 demonstrates that at energies above the $2\nu\beta\beta$ spectrum, the fitted model in enriched detectors is dominated by ^{232}Th, as was expected based on the prominence of the 2615 peak in data. Finally, while the $2\nu\beta\beta$ decay spectrum is the dominant component of the model over a wide range of energies, at the lowest energies in the fitted range, multiple decay chains contribute significantly. One of these is ^{210}Pb, which includes pdfs for both bremsstrahlung in the lead shield and ^{210}Pb deposited on the surfaces of PTFE parts near the detectors. The contribution of ^{210}Pb in the lead shield is relatively well constrained in fits because of the differences between the low energy spectra in DS0 and later datasets. Since the inner copper shield was not yet installed during DS0, the detectors were less shielded from the lead than they were for the remainder of the experiment's runtime, causing large differences in background rate that simultaneous fits can exploit. However, the ^{210}Pb located near the detectors is more difficult to model and less well-constrained during fitting. Since this pdf plays an important role in the fits at low energies, systematic studies exploring its role in the model are discussed at length in Sec. 6.7.

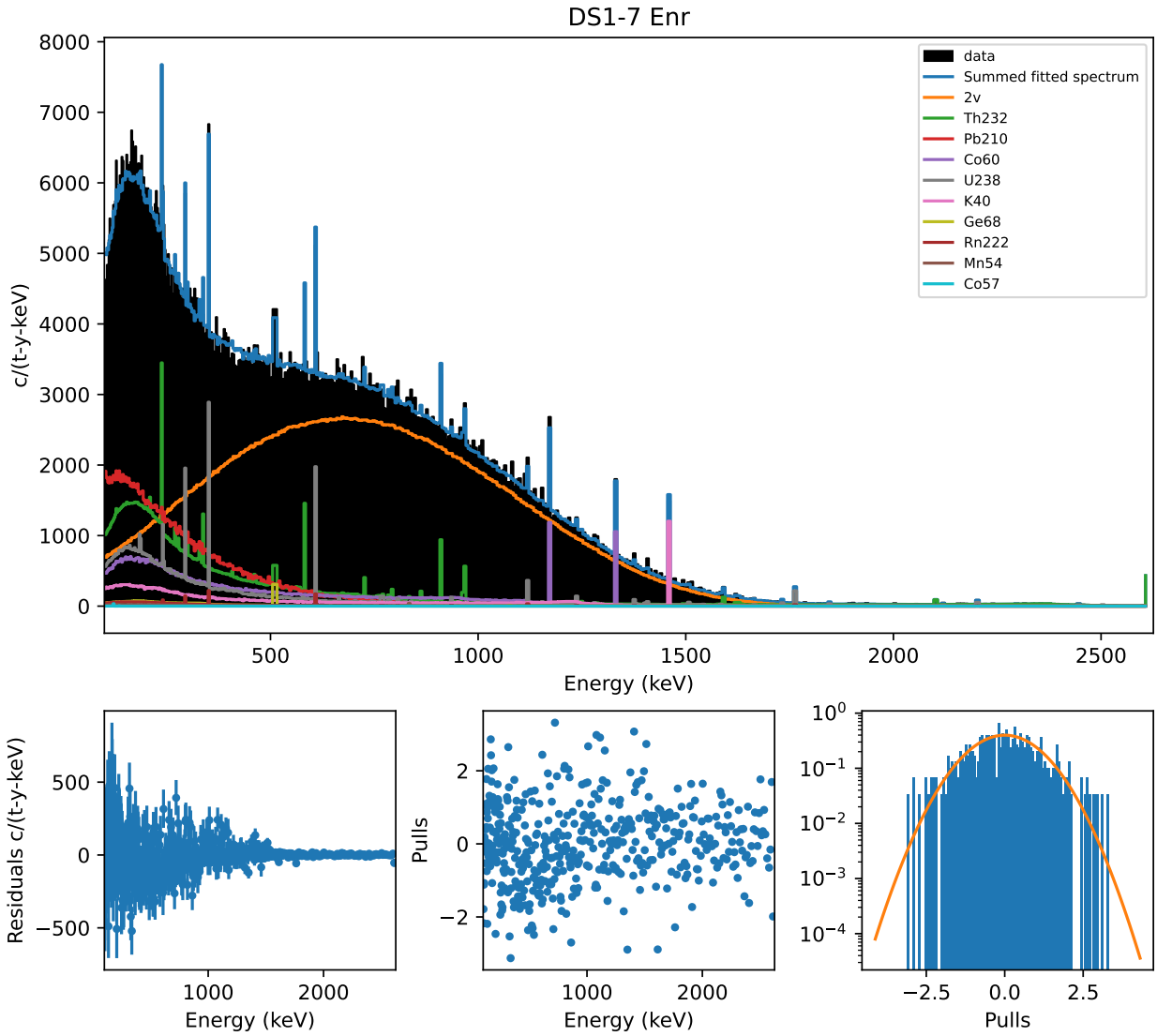


Figure 6.2: Result of spectral fits to DS0-7 data plotted over full fit energy range. Plotted for enriched detectors in both modules. Data from DS0 are excluded in the plot

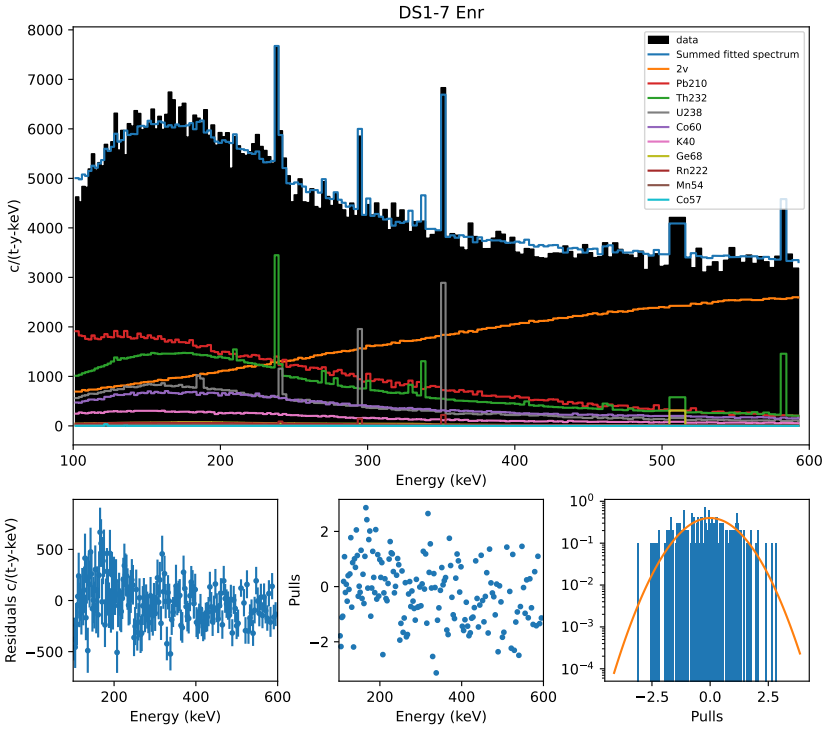


Figure 6.3: Zoom in of Fig. 6.2 in the range from 100 to 600 keV.

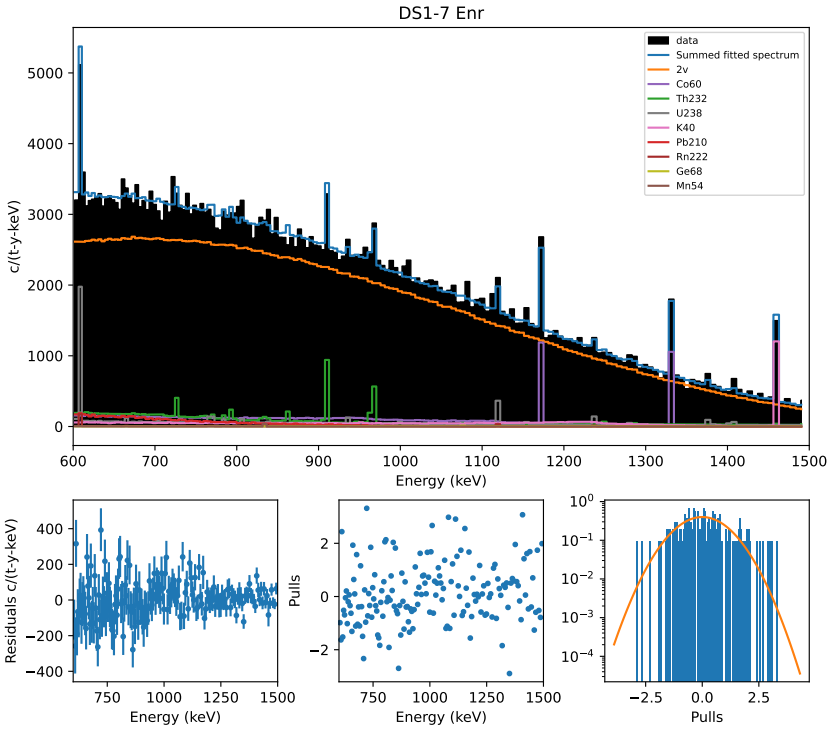


Figure 6.4: Zoom in of Fig. 6.2 in the range from 600 to 1500 keV.

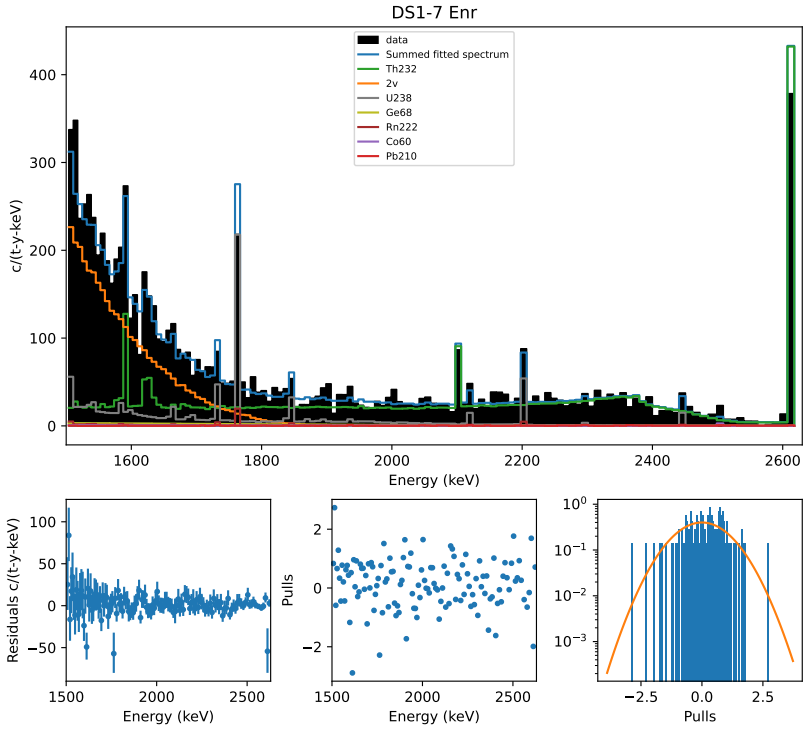


Figure 6.5: Zoom in of Fig. 6.2 in the range from 1500 to 2630 keV.

Since the full background model fit involves simultaneous fits between data across multiple detector groups, it can also be illustrative to directly compare fitting results between detector groups. Fig. 6.6 is one example of this type of plot that demonstrates the difference in ^{232}Th backgrounds between the two modules. The Module 1 and Module 2 enriched detector spectra are shown for data as well as for the ^{232}Th component of the fitted model. Module 1 data shows higher rates in both the continuum and in several prominent ^{232}Th peaks, and the model accurately reflects this difference.

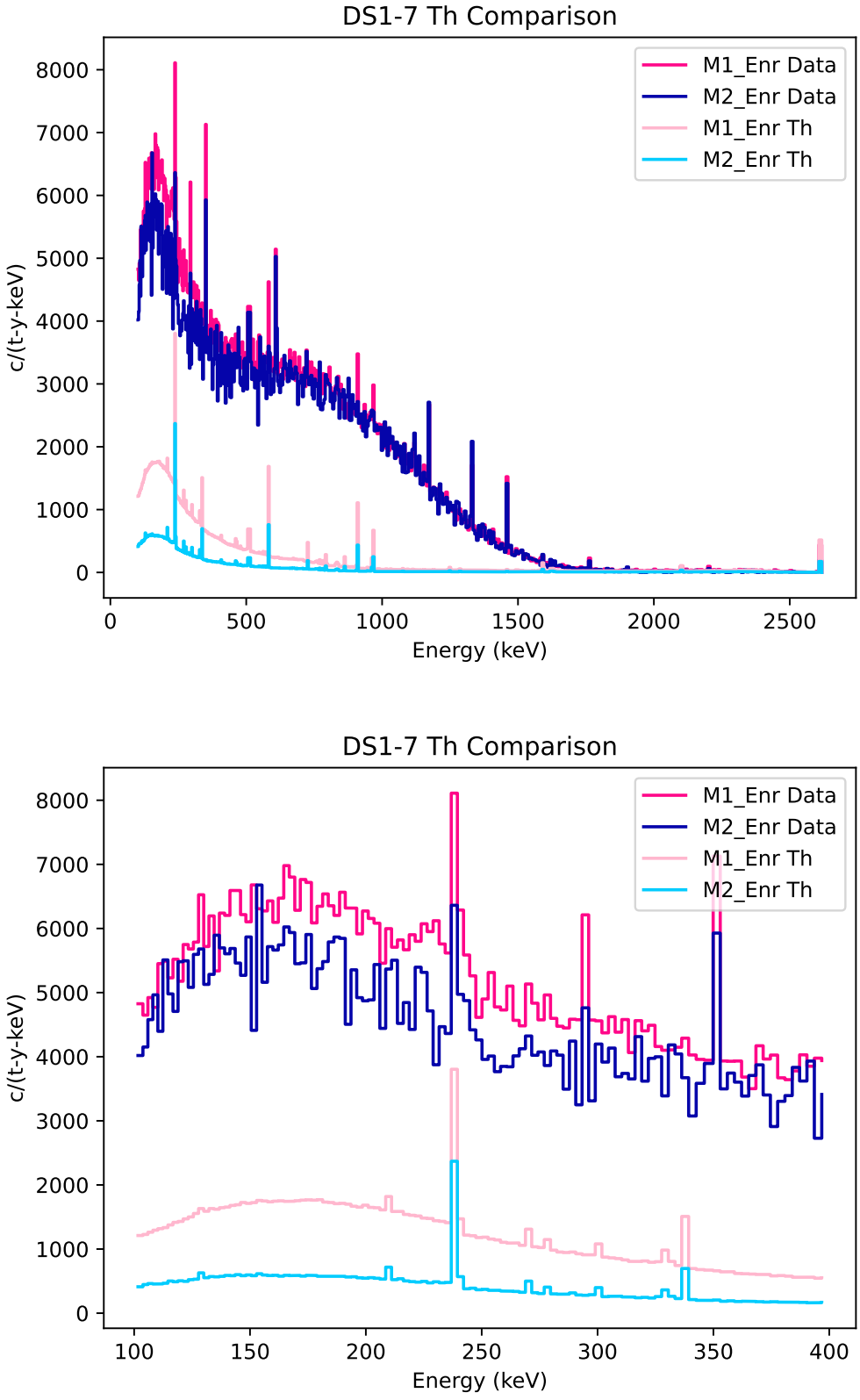


Figure 6.6: Comparison of data and the ^{232}Th portion of the background model between enriched detectors in Module 1 and Module 2. This comparison is shown for the full fit energy range (top) and for the low energy range below 400 keV (bottom).

Another feature of the fitted data is shown in Fig. 6.7, which compares enriched and natural detectors in Module 2. While the most prominent difference between these two data spectra is the lower $2\nu\beta\beta$ rate in natural detectors, this comparison also shows the higher rate of ^{60}Co backgrounds in natural detectors. The difference in cosmogenic-induced backgrounds between enriched and natural detectors was expected based on differences in surface exposure. The amount of time that enriched detectors spent without significant rock overburden was carefully minimized and tracked, while the surface exposure of natural detectors was less strictly controlled. Spectral fits show exactly how this difference influences the data across a wide range of energies. Module 2 detectors were chosen to demonstrate this effect because the data spectra are not shaped by highly localized backgrounds from other decay chains, such as the ^{232}Th hot spot observed in M1. Simultaneous fits of multiple spectra allow spectral fits to accurately capture features of the data that differ between detectors, enabling a better understanding of the background composition for different types of detectors. Appendix D includes versions of Figs. 6.6 and 6.7 that highlight other decay chains in the fitted model.

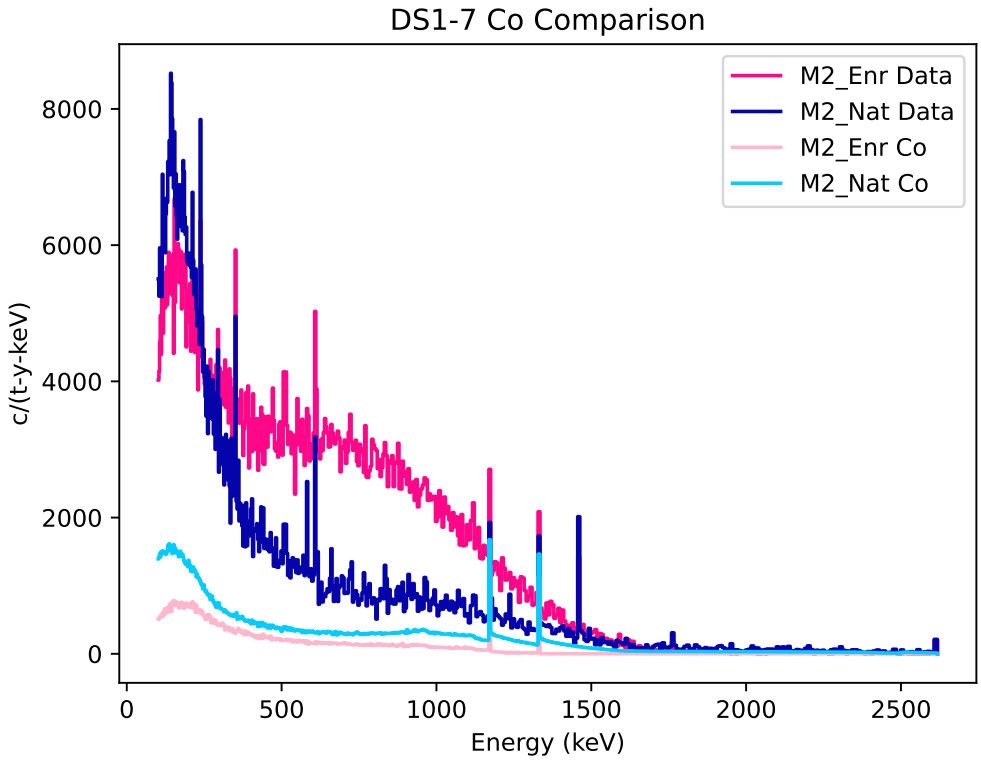


Figure 6.7: Comparison of data and the ^{60}Co portion of the background model between enriched and natural detectors in Module 2.

Section 6.3: Goodness of Fit

To fully evaluate the overall performance of fits to the DEMONSTRATOR's data, it is useful to supplement visual spectral comparisons with a quantitative figure of merit. The best-fit negative log likelihood can be used as a measure of the agreement between model and data if there exists some baseline NLL to which it can be compared. Since fits to simulated datasets represent an idealized setup with no systematic error, they can serve as a standard by which fits to data can be assessed. This study requires simulated datasets with a comparable level of statistics to the DEMONSTRATOR's data and also benefits from using a model where the spectral shape resembles that observed in data. This was achieved by sampling the best fit model obtained from fits to the DEMONSTRATOR's DS0-7 exposure. The number of samples drawn from the best-fit model was taken to be equal to the number of counts in the data between 100 and 2620 keV.

The best-fit negative log likelihood can vary between different simulated datasets drawn from the same underlying model, so this study required performing fits to many simulated datasets. The NLL from fitting the DEMONSTRATOR's data can then be compared to the distribution of best-fit simulated dataset NLLs. The fraction of simulated datasets with higher best fit NLLs than the data NLL can essentially serve as a p-value that captures the probability of obtaining the data NLL if the underlying model was completely comprehensive.

For the purpose of this study, one key change was made to the method of generating and fitting simulated datasets described in Sec. 4.3. In other simulated dataset studies, the weighted sum of pdfs formed a model from which samples were drawn, and the datasets were then fit using the same set of pdfs. However, this approach means that the simulated datasets and the pdfs used to fit them are impacted by the same statistical fluctuations, resulting in a lower negative log likelihood than if they were based on independent Monte Carlo simulations of the same background sources. This effect is particularly important because of the presence of some background sources with low efficiencies that nevertheless contribute a significant number of counts to the final model, such as Pb bremsstrahlung in the lead shield. Because simulations of inefficient background sources are very computationally intensive, the final pdfs have comparatively low statistics and are consequently subject to larger fluctuations. If such a source is a significant contributor to the background model, a large fraction of the total simulated counts will be drawn when generating the simulated dataset. To prevent this effect from erroneously impacting the goodness of fit calculation, the DEMONSTRATOR's simulations of each background source were divided in half, with half of the simulated statistics being used to generate simulated datasets, while the pdfs used in fitting were produced from the other half. For the

purpose of this study, the DEMONSTRATOR's data was re-fit with the same set of half-statistics pdfs used in fitting the simulated datasets.

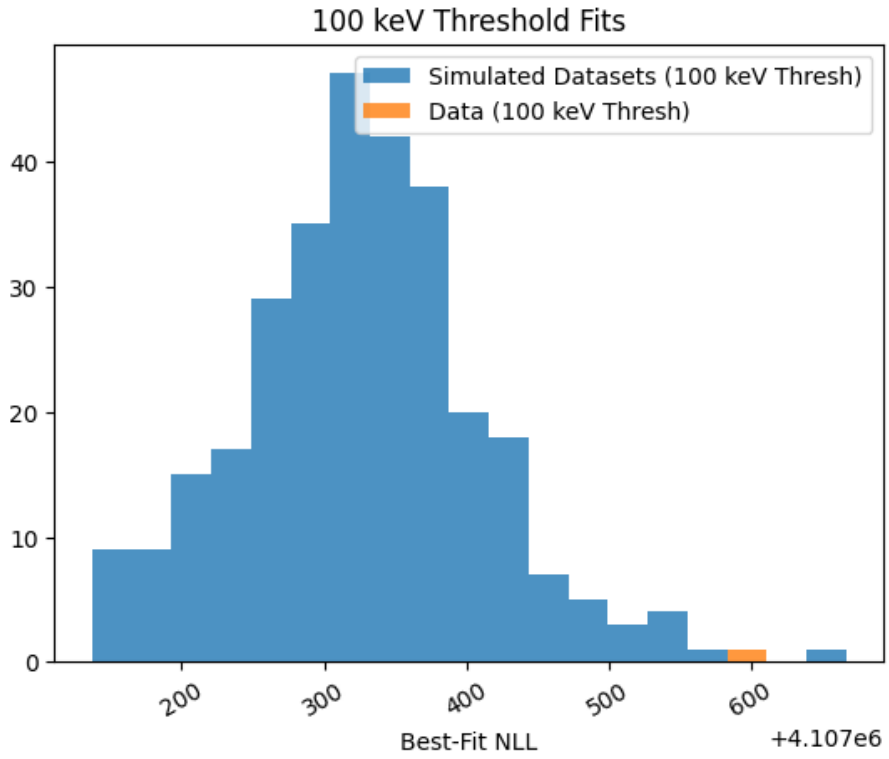


Figure 6.8: Distribution of the best fit negative log likelihoods from fits to 300 simulated datasets compared with the NLL found when fitting the DEMONSTRATOR's data. All simulated datasets contain the same number of samples as the the DEMONSTRATOR's data in the fitted energy range.

Fig. 6.8 shows the results of this exercise for the DEMONSTRATOR's standard fitting range between 100 and 2620 keV. Based on fits to 300 simulated datasets, the data has a higher NLL than 99.7% of simulated datasets, indicating that the fitted model displays more disagreement with the DEMONSTRATOR's energy spectrum than would typically be predicted for a perfect model. This is not unexpected, since all models rely on certain assumptions that cannot fully reproduce every detail of a real experiment. For example, except in the case of the M1 thermosyphon cavity, all background sources are assumed to be uniform within a given component group, which may not accurately represent the background distribution in data. The goodness of fits based on comparison to simulated datasets indicates that the DEMONSTRATOR's data-driven background model imperfectly reproduces the data, but it still is in good agreement with most features of the observed data, such as the background estimation window event rate.

Section 6.4: Results in the Background Estimation Window

A particularly relevant test of the performance of background model fits is the agreement between model and data in the $0\nu\beta\beta$ background estimation window. The model's ability to reproduce the data rate well in this low statistics region is a good indicator of both overall fit performance and the model's ability to locate the main backgrounds relevant to $0\nu\beta\beta$.

The background indices calculated here for both data and the fitted model are not directly comparable to the background index used in the $0\nu\beta\beta$ half-life limit, mainly due to the differences in applied cuts between the two analyses. The background indices calculated in this section use the same cuts applied during fitting, which include all data cleaning and muon veto cuts as well as a delayed charge recovery cut for removing surface alphas. In addition, only the multiplicity one spectrum is used when calculating the background index. In contrast, the background index for the $0\nu\beta\beta$ analysis includes AvsE and LQ cuts that supplement the cuts applied here. A more minor difference between the background index reported here and the one used for the $0\nu\beta\beta$ analysis is that the energy range of the background estimation window is defined slightly differently due to the applied binning scheme. Compared to to the 360 keV region described in Sec. 3.1, larger energy ranges are removed around each gamma peak in this analysis, leading to a ~330 keV energy window. Since the background is relatively flat in this region, the background index is not strongly impacted by small changes to the energy range used. The background index comparisons include data from DS1-7 (excluding DS5a). The DS0 spectrum was included in the fit used to define the model but not in the background index calculation due to its higher backgrounds.

The calculated background index for all enriched detectors in the fitted model is 17.7 cnts/(keV t yr), which can be compared to the background index with same cuts applied in data, 18.6 ± 1.1 cnts/(keV t yr).[1] The model rate is consistent with the data rate to within a 1σ Poisson statistical uncertainty, indicating that the model performs well in the energy region surrounding $Q_{\beta\beta}$. The difference between the background index presented here and the background index of 6.00 c/(keV t yr) quoted in Sec. 3.1 is almost entirely attributable to the ~65-70% reduction in backgrounds near $Q_{\beta\beta}$ from applying the remaining PSA cuts (low AvsE, high AvsE, and LQ) to data following the application of the DCR and multiplicity cuts.

The model is also able to accurately reproduce the difference in background index between the two modules. In the fitted model, the BI in M1 enriched detectors is 20.1 cnts/(keV t yr) and the BI in M2

[1]No uncertainty is quoted for the model rate because the uncertainties on each fitted activity density cannot be straightforwardly combined into a total model uncertainty.

enriched detectors is 10.3 cnts/(keV t yr), compared to the background indices in data of 21.2 ± 1.3 cnts/(keV t yr) for M1 and 10.8 ± 1.6 cnts/(keV t yr) for M2.

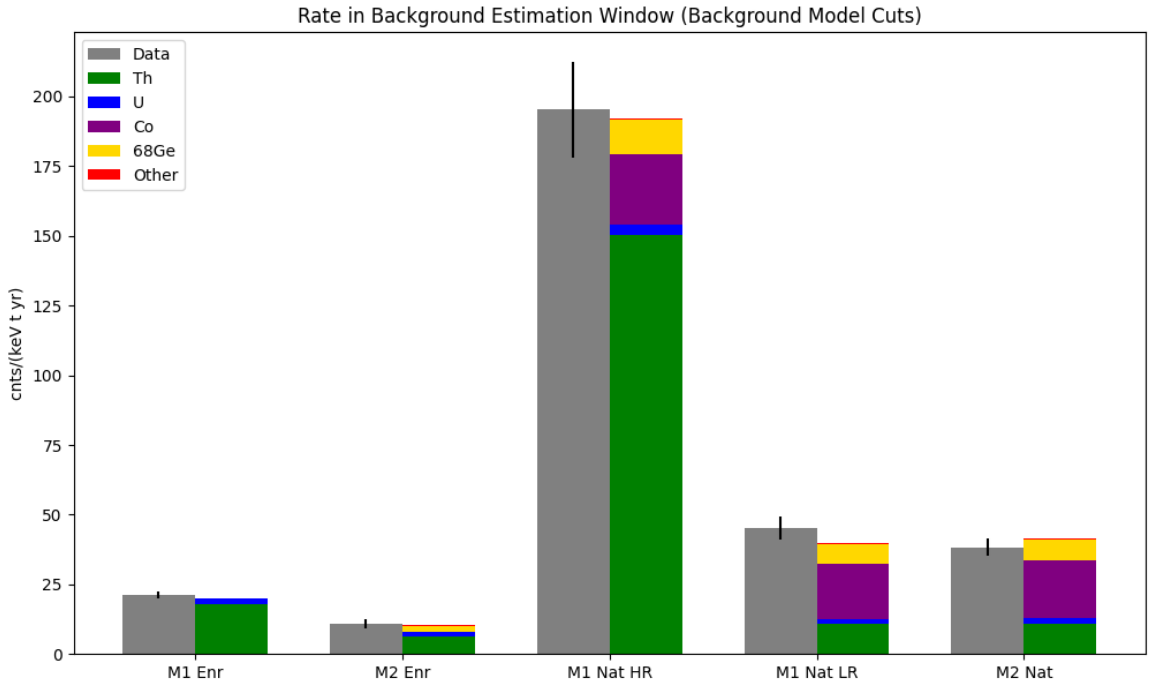


Figure 6.9: Comparison of the rate in the $0\nu\beta\beta$ background estimation window for each detector group between data and the fitted model. M1 Nat HR refers to the two high rate detectors near the M1 crossarm, while M1 Nat LR includes all other (lower rate) M1 natural detectors. In addition to the typical data cleaning and delayed charge recovery cuts applied to the data, a granularity cut is applied to both data and model. The model background index is broken down by decay chain/cosmogenic isotope.

In addition to being used as a test of fit performance, the modeled rate in the $0\nu\beta\beta$ background estimation window has further explanatory power when broken down by decay chain. Fig. 6.9 compares the background index of each detector group between data and the fitted model. The background index of the fitted model for each detector group is further subdivided based on the decay chain or cosmogenic isotope from which the backgrounds originated. For enriched detectors in both modules, the rate in the background estimation window is dominated by ^{232}Th. The higher rate in the background estimation window in Module 1 than in Module 2 can be entirely explained by the higher ^{232}Th in Module 1. The difference in background index between the two high rate Module 1 natural detectors near the crossarm and other natural detectors is also explainable based on ^{232}Th alone.

Fig. 6.9 also demonstrates that the difference in background index between enriched and natural detectors is caused by the increased presence of cosmogenic isotopes in natural detectors. Events due to the decay of ^{60}Co and ^{68}Ga (a progeny of ^{68}Ge) are a significant contributor to the rate in the background estimation

window for all natural detectors. This difference was expected due to the increased above-ground exposure natural detectors experienced prior to the start of the experiment, which led to greater cosmogenic activation. A difference in ^{68}Ge rates between enriched detectors in the two modules was also expected based on cosmogenic activation, although the magnitude of the difference in fitted results is larger than expected. Five of the fifteen enriched detectors in Module 2 were constructed from recycled material that spent a larger amount of time above ground, explaining the presence of a difference between modules. However, the magnitude of the fitted ^{68}Ge activity in M2 enriched detectors is a factor of 2-3 high compared to the level predicted by tracking the time enriched material spent above ground. This difference may be partly explained by the fact that ^{68}Ge is not well-constrained during fitting, given the energy range used. The only peak present above 100 keV in the ^{68}Ge decay chain is the 511 keV gamma peak, which has contributions from many different decay chains. Studies of low energy cosmogenic peaks not included in spectral fits are in good agreement with the predictions of surface exposure tracking, putting them in tension with the results of spectral fits. A systematic check where the M2 Enr ^{68}Ge was constrained to the prediction from exposure tracking will be discussed in Sec. 6.7.

Section 6.5: Source Locations of Observed Backgrounds

The data-driven background model can be further broken down to determine the locations in the DEMONSTRATOR where most backgrounds originated. Table 6.1 summarizes the ten floated pdfs that contributed the greatest number of counts to the fitted model. The number of counts comprises the integrated counts over the full fitted energy range and exposure (DS0-7) for both enriched and natural detectors. The full table, including all pdfs that contributed more than one count to the final model, is given in Appendix A.

parameter	Counts	% Counts
2v_M1EnrGe_bulk	9.3352e+04	32.244
2v_M2EnrGe_bulk	2.9800e+04	10.293
PbBrem_RadShieldAssembly_001_RadShieldPb_001_bulk	2.9527e+04	10.199
Th_M1CPInterfaceCavityBottomSurface_bulk	2.4370e+04	8.417
Co_RadShieldCuOuter_bulk	1.3782e+04	4.760

Continued on next page

parameter	Counts	% Counts
Pb_M1DUPTFE_Enr_surf	8.0937e+03	2.796
U_M1CryostatCopperNear_bulk	7.1090e+03	2.455
Co_M1NatGe_bulk	7.0307e+03	2.428
Co_M2NatGe_bulk	6.1380e+03	2.120
U_M2CryostatCopperNear_bulk	5.4699e+03	1.889

Table 6.1: Overview of the ten top contributors to the DEMONSTRATOR's data-driven background model. The number of integrated counts in the fitted energy range is shown, as is the percent of the total counts in the model that the pdf accounts for.

A table of the fitted activity densities of all pdfs included in spectral fits and their statistical uncertainties can be found in Appendix B.[2] For background sources that are decaying away with known half-lives, the reported activity densities correspond to their values at the beginning of DS0. The statistical uncertainty on the activity density for each pdf is determined by profiling the likelihood using the method described in Sec. 4.2. The profile curves for all pdfs included in fits to data are shown in Appendix C. The range that each curve is plotted over reflects a 3σ region around the best fit specific activity, while the 1σ statistical uncertainties quoted in Appendix B correspond to a ΔNLL of 0.5. For most profile curves, the negative log likelihood is relatively Gaussian over the floated parameter, but some distributions can be skewed, leading to asymmetric error bars.

The results of Appendix B can be better visualized by plots showing the total number of integrated counts attributed to each pdf in the fitted model. Given the large number of floated pdfs, each plot only includes pdfs from a single decay chain. Figs. 6.10 and 6.11 show this breakdown for the ^{232}Th and ^{238}U chains. Other decay chains or cosmogenic isotopes that contribute significantly to the model are included in Appendix D. These include ^{238}U, ^{60}Co, ^{40}K, $2\nu\beta\beta$ (^{76}Ge), ^{210}Pb, ^{222}Rn, and ^{68}Ge. For each plot, the composition of the fitted model, shown in color, is compared to that of the assay-based expectation, shown in black. The number of integrated counts reflects counts from all detectors, both enriched and natural, over the full fitted exposure and energy range. The error bar on the fitted number of integrated counts for each

[2]For the CPInterfaceCavityBottomSurface pdfs, the quoted numbers are activities, rather than activity densities, since the mass associated with the hot spot is not known.

pdf is based on converting the 1σ statistical uncertainty of the fitted activity density into detected counts. Systematic uncertainties are not included in the error bars for these plots.

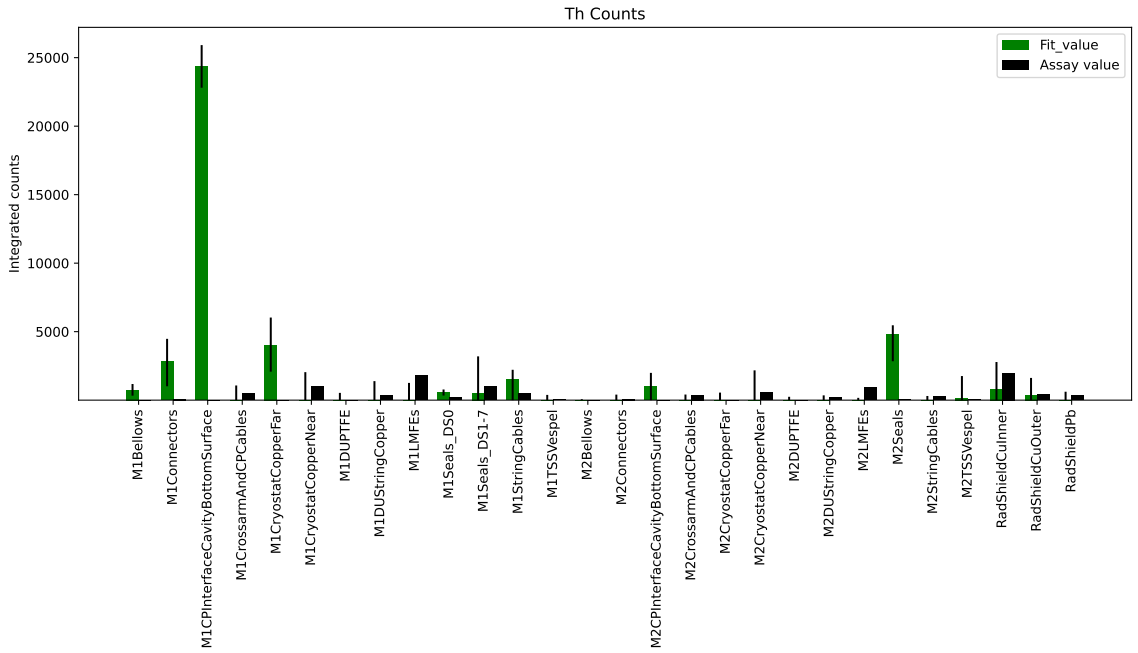


Figure 6.10: Distribution of ^{232}Th counts among different source locations in the fitted model (green) compared to the assay-based model (black). The number of integrated counts reflects the total number of counts that a pdf contributes to the model in the fit energy range (100-2620 keV) for the entirety of the fitted exposure in both enriched and natural detectors. Error bars on fitted count rates reflect only statistical uncertainties. Assay uncertainties are not shown, and assay values are plotted at zero for components where only upper limits were available.

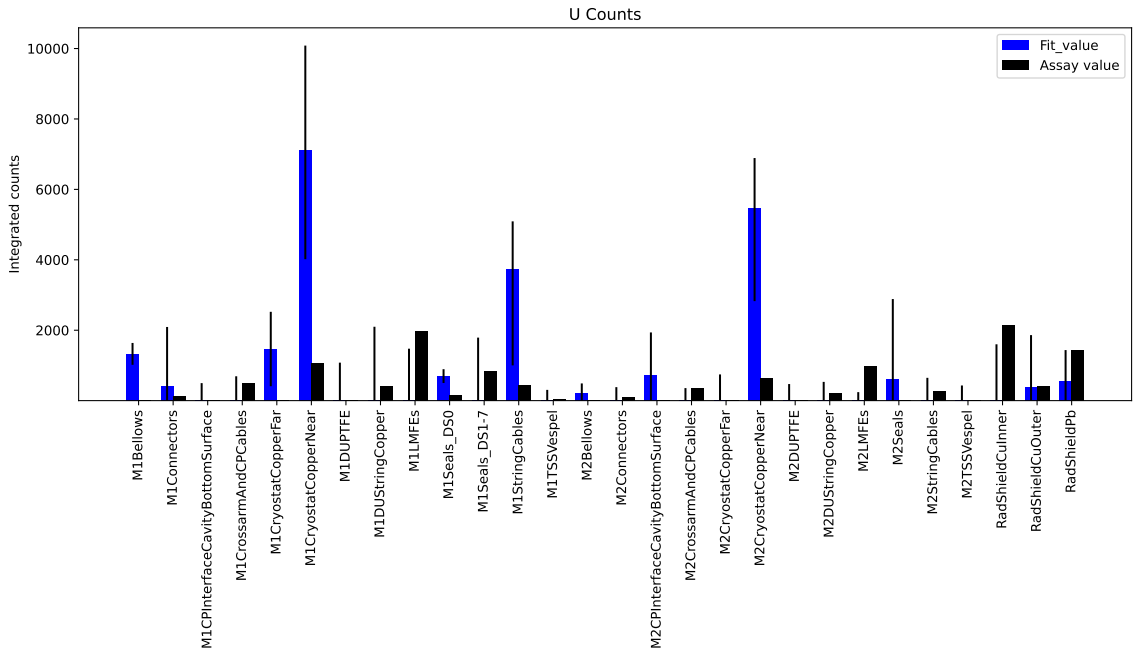


Figure 6.11: Distribution of ^{238}U counts among different source locations in the fitted model (blue) compared to the assay-based model (black). The number of integrated counts reflects the total number of counts that a pdf contributes to the model in the fit energy range (100-2620 keV) for the entirety of the fitted exposure in both enriched and natural detectors. Error bars on fitted count rates reflect only statistical uncertainties. Assay uncertainties are not shown, and assay values are plotted at zero for components where only upper limits were available.

Fig. 6.10 demonstrates that the fitted ^{232}Th is dominated by the contribution from a single component group, the bottom surface of the M1CPInterfaceCavity. This component, which represents a hot spot in the Module 1 thermosyphon cavity, was discussed in Chapter 3 as the component best able to match the detector distribution and peak ratios seen in data. This result is discussed in more detail in Sec. 6.6. Other major sources of ^{232}Th in the model cannot be as conclusively determined because of the size of the statistical error bars. It is worth noting that source locations near the detectors (including LMFEs, DUStringCopper, DUPTFE, and StringCables) all fit to ^{232}Th activities that are within 1σ of 0. This finding does not ensure that the total number of counts fitted into all near-detector components would be consistent with 0, given that there are significant correlations between pdfs from similar locations that influence the size of the error bars. Nevertheless, near-detector components do not appear to be a major ^{232}Th contributor to the model. ^{238}U components have similarly large uncertainties, indicating that the ^{238}U source location cannot be definitively determined, although the total amount of ^{238}U in the model is well constrained. Most ^{238}U in the fitted model is located in the Module 1 and Module 2 CryostatCopperNear groups, which predominately consist of the cryostat vessel and coldplate copper. The M1StringCables, which is a near-detector component group,

is the third-highest ^{238}U contributor. These results imply that the dominant ^{238}U sources may be located somewhat closer to the detector array than the dominant ^{232}Th sources, but most ^{238}U appears not to be located directly adjacent to the detectors.

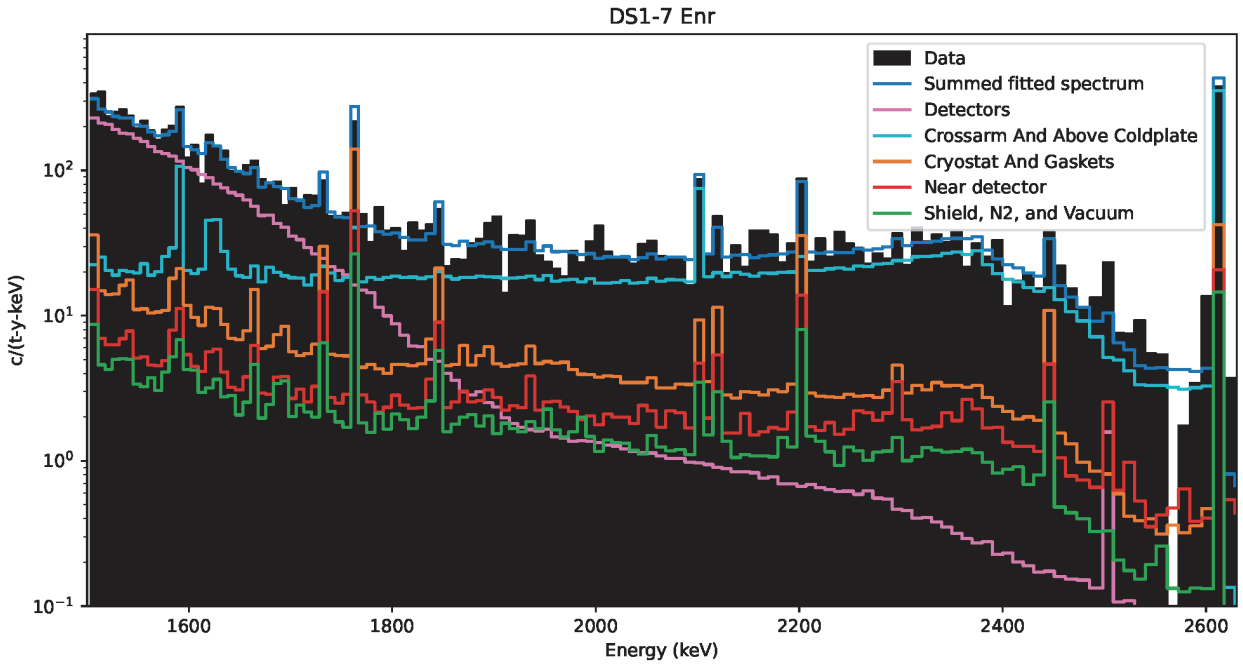


Figure 6.12: Result of spectral fits to DS0-7 data in the energy region surrounding $Q_{\beta\beta}$, plotted for enriched detectors in both modules. The fitted spectrum is subdivided based on source region, where multiple component groups from similar locations are grouped together, and each spectrum is summed over all decay chains. The M1 thermosyphon cavity hot spot discussed in Sec. 6.6 is part of the "Crossarm And Above Coldplate" region, which dominates the event rate near $Q_{\beta\beta}$. Data from DS0 are excluding in the plot.

The total effect of near-detector components on the best fit model is shown in Fig. 6.12, where the fitted spectrum near $Q_{\beta\beta}$ is subdivided based on source region rather than decay chain. Based on fits to DS0-7 data, most sources that led to counts in the $0\nu\beta\beta$ background estimation window originated in the region above the coldplate and along the crossarm, a region which includes the candidate hot spot in the M1 thermospyon cavity. In contrast, near-detector components from all decay chains made only a minor contribution to the final model, indicating that the DEMONSTRATOR's background index does not pose a problem for LEGEND-200.

Section 6.6: Fitting and the ^{232}Th Excess

Chapter 3 posited that a large portion of the ^{232}Th excess observed in data could be explained by a localized source in the region of the Module 1 crossarm. The detector distribution of backgrounds and the ratios of high and low energy ^{232}Th peaks in the natural detectors adjacent to the M1 crossarm suggested that

the localized background could best be explained by a source in or near the cavity of the M1 thermospyhon, located above the coldplate. Simulations of a point source in the center of the bottom surface and of a uniform source on the bottom surface of the cavity both matched the experimental data better than other simulated sources. Based on these studies, a simulated pdf of a contamination in the thermosyphon cavity was included in background model fits. The uniform bottom surface pdf was chosen for inclusion because it was easier to physically motivate, although either hot spot location was considered physically reasonable. Simulations of a ^{232}Th hot spot in the Module 2 thermosyphon cavity and of ^{238}U and ^{40}K hot spots in both modules' thermosphyon cavities were also included during fitting. While detector distributions of prominent peaks in data did not suggest that any of these additional pdfs would account for a significant portion of the observed backgrounds, floating the activities of these pdfs during fitting allowed this assumption to be tested. The results of spectral fits support the findings from supplementary background studies that a ^{232}Th hot spot in the region of the M1 thermospyhon cavity is responsible for a significant portion of the DEMONSTRATOR's backgrounds. Fig 6.10 demonstrates that the hot spot accounts for the majority of ^{238}Th counts over all detectors in fits of DS0-7 data. The activity fitted into the M1 thermosyphon cavity hot spot pdf is 1.73 mBq with a statistical uncertainty of ± 0.11 mBq. Varying the detector parameters applied to the simulations during post-processing results in a upper systematic error of $+0.14$ mBq and a lower systematic error of -0.07 mBq, based on fits with 100 sets of systematics pdfs. Future fits with 1000 sets of systematics pdfs will further refine these numbers. The dominant systematic contributing to this uncertainty is the full charge collection depth (fccd) of the dead layer in natural detectors, which is reasonable given that the two detectors closest to the hot spot are natural detectors.

Since the fitted ^{232}Th activity in the hot spot pdf is many standard deviations above 0, it is clear that this background source plays an important part in the model. To better understand how the presence of this source improves the model, it is instructive to compare fits including and excluding it. For the purpose of this comparison, the pdfs for other decay chains in the Module 1 thermosyphon hot spot and the Module 2 hot spot pdfs for all decay chains were still allowed to float, meaning that the fits only differed by a single degree of freedom. Fig. 6.13 demonstrates the results of this comparison in the two high rate M1 detectors. By comparing the residuals, it is clear that the fit omitting the ^{232}Th cavity hot spot, shown in the bottom plot, less accurately describes the data. Several ^{232}Th peaks are underpredicted in the high rate detector group when a hot spot pdf is not present, while a couple of the ^{238}U peaks are slightly overestimated. In other detector groups, differences between the two fits are less prominent. Another difference between the

two fits can be seen by contrasting the event rate in the $0\nu\beta\beta$ background estimation window for all M1 detector groups, shown in Fig. 6.14. Excluding the ^{232}Th cavity hot spot pdf from the fit decreases the modeled background index in the high rate detector group, worsening the agreement with these detectors' background index in data. While the changes from excluding the hot spot pdf are not dramatic, the fitted model is better able to describe multiple features of the data when this source in included in the fit. This conclusion is further validated by comparing the NLLs of the two fits, since the NLL is a measure of their relative goodness of fit. The NLL for the fit including a hot spot source in the M1 thermosyphon cavity has a Barlow Beeston NLL that is lower by 137 than the NLL of a fit where the hot spot pdf is constrained to have no activity. The Poisson NLL differs by 143 between the two fits. Given that the two fits only differ by one degree of freedom, the large difference in NLLs is a sign that the final fit models the data much better when a hot spot pdf is present.

Fig. 6.15 demonstrates another aspect of how fits to the DEMONSTRATOR's data change depending on whether the hot spot pdf is included by examining which source locations contribute large numbers of ^{232}Th counts to the fitted model. When the thermosyphon cavity hot spot is excluded from the fit, there is a large increase in the number of ^{232}Th counts attributed to the M1CryostatCopperFar and M1CrossarmAndCPCables component groups. The M1CryostatCopperFar group includes much of the M1 thermosyphon, while M1CrossarmAndCPCables consists of the portion of the signal and high voltage cables that lies on top the coldplate and along the crossarm. This suggests that, in the absence of a hot spot pdf, most of the ^{232}Th is still fit into the general region containing the suspected hot spot. The only other notable difference between the two fits is that more ^{238}U is fit into the M1 thermosyphon hot spot if the ^{232}Th in this location is constrained to be 0. This pdf helps account for the difference in continuum backgrounds between the high rate detector group and other M1 detectors, but its higher activity overestimates the uranium peaks in these two detectors.

From this comparison, it is evident that a large fraction of the observed ^{232}Th background is attributable to a source in the M1 thermosyphon region, regardless of whether a hot spot is considered. However, a bulk ^{232}Th source around the M1 thermosyphon is less successful at modeling the background excess than a hot spot, as evidenced by the difference in NLLs between fits with and without the hot spot pdf. No other component in the model can explain the Module 1 ^{232}Th excess as successfully as a hot spot in the M1 thermosyphon cavity. As described in Sec. 3.6, localized sources in multiple other locations near the high rate detectors were also simulated but were not included during spectral fits because they could not match the data as accurately as the thermosyphon cavity hot spot. Together, spectral fits and these supplementary

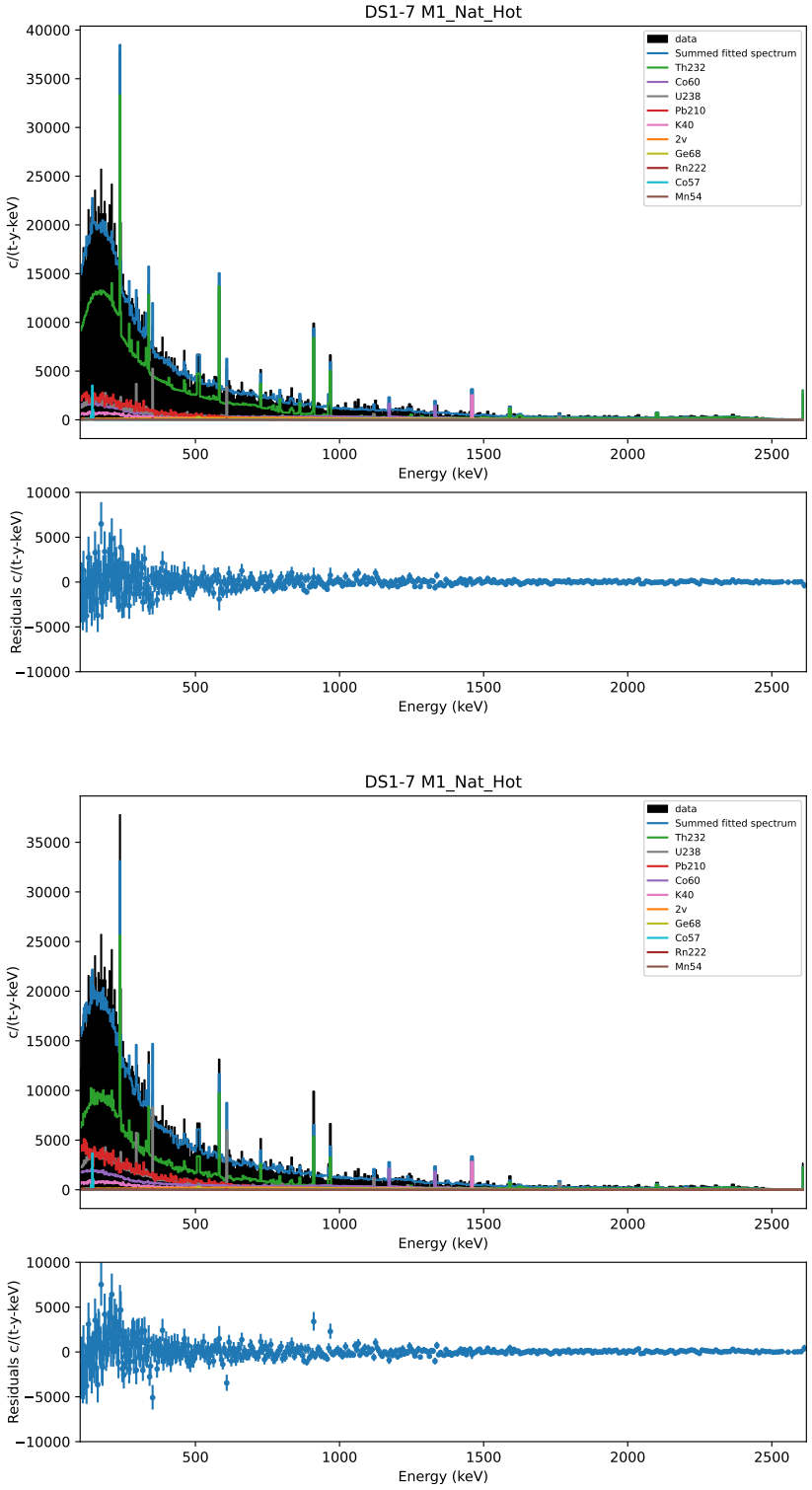


Figure 6.13: Results of full background model fits for the M1 high rate detector group compared between fits that included a pdf of a hot spot ^{232}Th source in the M1 thermosyphon cavity (top) and fits that do not include a hot spot pdf (bottom).

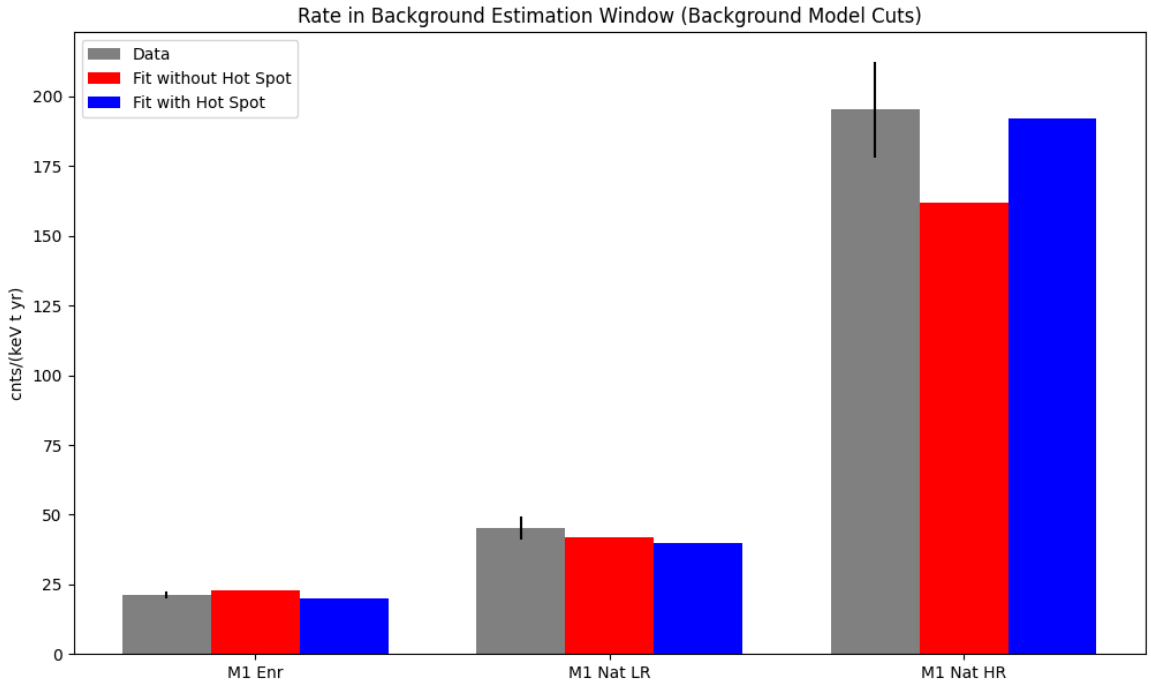


Figure 6.14: Comparison of the background rate in the $0\nu\beta\beta$ background estimation window for Module 1 detector groups in fits with and without a simulated M1 ^{232}Th hot spot. Module 1 natural detectors are split between low rate (LR) and high rate (HR) groups to separate the two natural detectors near the M1 crossarm from other other natural detectors. The background rate is calculated with the standard set of background model cuts and with a multiplicity cut.

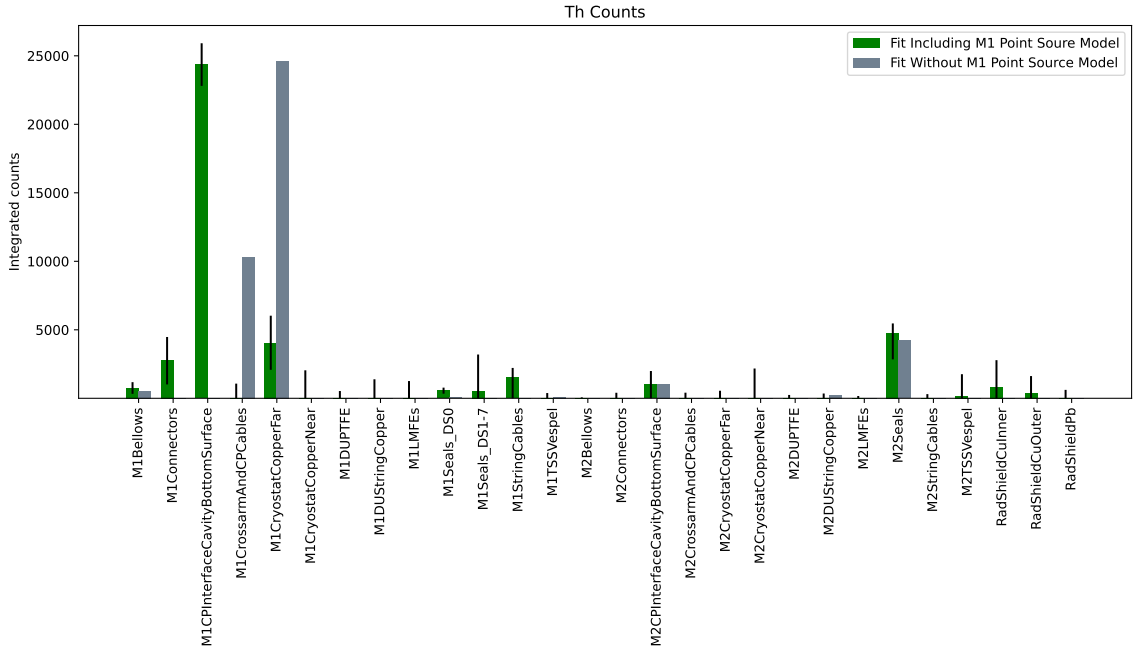


Figure 6.15: Composition of the fitted ^{232}Th spectrum based on the integrated counts fitted into each ^{232}Th pdf for the entire fit energy range. A fit including a hot spot source in the M1 thermosyphon cavity (M1CPInterfaceCavityBottomSurface) is compared to a fit where this pdf is constrained to have zero activity. This plot includes counts in both enriched and natural detectors in both modules. The error bars on the fit including the M1 hot spot source reflect only the statistical uncertainty.

background studies strongly point to a hot spot in the M1 thermosyphon cavity. However, the results of the 2023 assay of the suspected hot spot region in the thermosyphon cavity are inconsistent with the 1.73 mBq of ^{232}Th required by spectral fits. Assay results from the GeMPI detector at LNGS place a 90% upper limit of 0.6 mBq for ^{232}Th activity on the inner surface of the thermosphyon block [71]. This finding cannot rule out a source being present in this region during the DEMONSTRATOR's operation that was no longer present at the time of the assay, as was discussed in Chapter 3. If the assayed component was not responsible for the background excess, simulations of an excess in this region would likely still serve as a reasonable proxy for the true background source. Simulated spectra from a ^{232}Th contamination on the bottom surface of the thermosyphon cavity accurately model the spectral features observed in the two high rate natural detectors, where the hot spot accounts for the majority of events. Simulations of this source also do a reasonable job matching the detector distribution of M1 ^{232}Th backgrounds, as described in Section 3.6. Although the thermospyhon cavity region cannot be definitively confirmed as the hot spot location, it models the localized ^{232}Th excess observed in the DEMONSTRATOR well. The impact of modeling the hot spot in this location on the fitted $2\nu\beta\beta$ result is discussed further in Sec. 6.8.3.

Based on this model allowing for a hot spot in the Module 1 thermosyphon cavity, the results of spectral fits can be used to determine how much of the background rate near $Q_{\beta\beta}$ is explained by the hot spot, as shown in Fig. 6.16. In Module 1 enriched detectors, ^{232}Th in the hot spot pdf accounts for 12.2 cnts/(keV t yr) in the background estimation window with the standard background model cuts and a multiplicity cut. This rate accounts for 60.4% of the total M1 Enr BI and 67.6% of the ^{232}Th contribution to the M1 Enr BI. Since the M1 thermosyphon hot spot is relatively far from Module 2, it only contributes minimally to the M2 Enr BI (~0.6 cnts/(keV t yr), or 5.9% of the total M2 Enr BI). When the contribution from the Module 1 hot spot pdf is excluded, the ^{232}Th rate in the background estimation window is comparable in both modules: 5.8 cnts/(keV t yr) in M1 vs. 5.7 cnts/(keV t yr) in M2.

In the best fit model, a ^{232}Th hot spot in the thermospyhon cavity for Module 2 fits to a much lower activity than the Module 1 hot spot, only contributing 0.7 cnts/(keV t yr) in the background estimation window for M2 enriched detectors (6.5% of the total M2 Enr BI). Based on the 1σ statistical uncertainty determined from a profile likelihood technique, the activity in this pdf is consistent with zero. This supports the conclusion from detector distribution studies that there is no strong evidence for a ^{232}Th hot spot in Module 2. Similarly, the pdfs for other isotopes in the cavity hot spot do not significantly contribute to the fitted model.

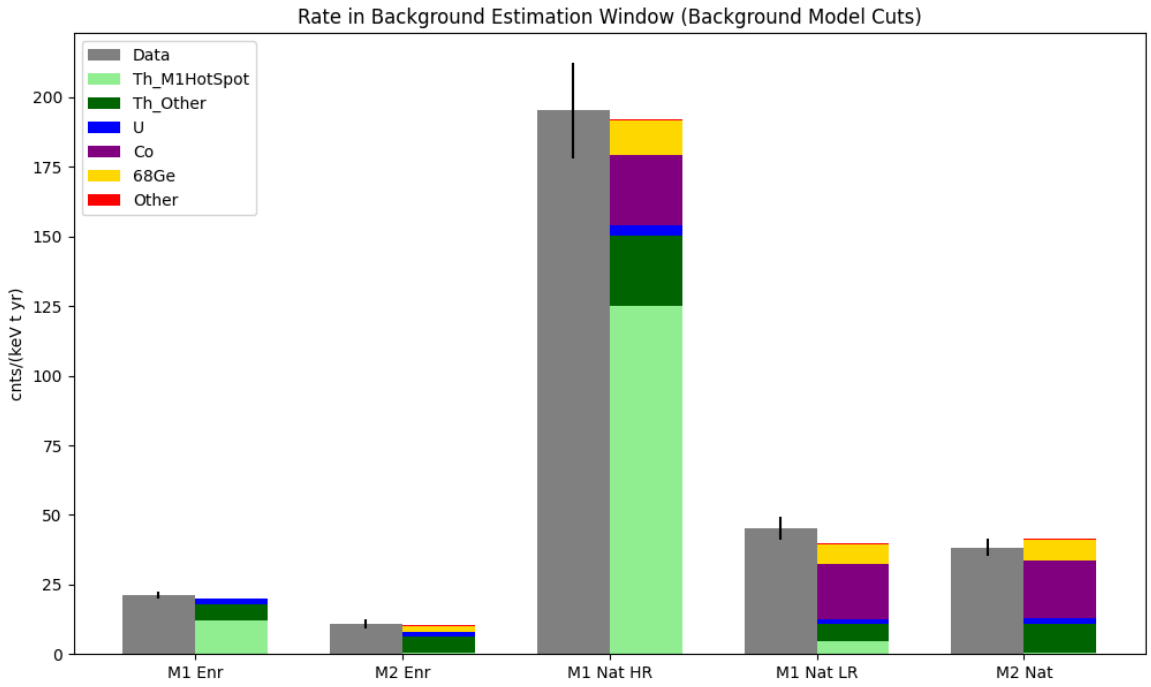


Figure 6.16: Adaptation of the background estimation window breakdown shown in Fig. 6.9 with the ^{232}Th contribution to the event rate divided between ^{232}Th originating in the Module 1 thermosyphon cavity hot spot and all other ^{232}Th source locations

Section 6.7: Additional Systematic Tests

The study described in the previous section evaluating the effect of excluding the M1 hot spot pdf from spectral fits is one example of a systematic test performed on the DEMONSTRATOR's data. Changing fit parameters, such as which pdfs are allowed to float, can be useful for checking the robustness of key results. Given that no model can perfectly capture all details of an experimental apparatus, these tests give an estimate of how inaccuracies in the model impact the conclusions that can be drawn from it. The effect of the systematic studies described in this section on the best-fit $2\nu\beta\beta$ half-life will be examined in Sec. 6.8.

One systematic parameter that can be examined is the low energy threshold of the fit. For the studies described in this section, energy thresholds ranging between 44 keV and 565 keV were tested. The impact of varying the fit range on the modeled background index and on the activity fitted into the M1 ^{232}Th hot spot pdf are shown in Tables 6.2 and 6.3. In a well-performing model, both of these parameters should not be dramatically impacted by the choice of fitted energy range.

Table 6.2 shows that varying the low energy threshold does not lead to large changes in the background index of the fitted model. Over all thresholds tested, the fitted M1 Enr BI only ranges between 0.4σ and 1.2σ below the background index for M1 enriched data, where σ is based on the data-derived statistical

	M1 Enr BI (cnts/(keV t yr))	M2 Enr BI (cnts/(keV t yr))
Data	21.2	10.8
44 keV Thresh Model	20.6	10.2
100 keV Thresh Model (Standard)	20.1	10.3
200 keV Thresh Model	20.1	10.6
300 keV Thresh Model	20.1	11.2
565 keV Thresh Model	19.7	11.2

Table 6.2: Comparison of the background indices for M1 and M2 enriched detectors between the data and models obtained from fits performed with different low energy thresholds. In addition to the typical data cleaning and delayed charge recovery cuts applied to the data, a granularity cut is applied to both data and model, but the AvsE and LQ cuts are not used.

Low Energy Threshold	Fitted ^{232}Th Hot Spot Activity (mBq)	% Change from Standard Fits
44	1.614	-6.8%
100 (Standard)	1.731	N/A
200	1.681	-2.9%
300	1.678	-3.1%
565	1.711	-1.2%

Table 6.3: Fitted activity for the hot spot pdf in the M1 thermsopyhon cavity as a function of low energy threshold applied used during fitting.

uncertainty of 1.3 cnts/(keV t yr). The fitted M2 Enr BI spans from 0.4σ below to 0.3σ above the background index for M2 Enr detectors in data, where σ = 1.6 cnts/(keV t yr)). Regardless of the energy threshold applied in the fit, the data in this region is well-described by the model.

The fitted activity in the hot spot displays some variation as the energy threshold is varied but continues to be the dominant ^{232}Th source in the model independent of energy threshold. The level of variation shown in Table 6.3 is reasonably consistent with the ~6.3% statistical uncertainty for this component found in Sec. 6.6, although fits with different thresholds cannot be considered statistically independent. Since the thermosyphon cavity hot spot accounts for most of the ^{232}Th counts in the model regardless of energy threshold, this study strengthens the conclusion from Sec. 6.6 that spectral fits support a hot spot as the dominant source of ^{232}Th in Module 1.

The lowest energy threshold tested for this study was chosen to be at ~44 keV because sources of ^{210}Pb with a direct line of sight to the detectors lead to a spectral peak at 46.5 keV. This peak is observed in the DEMONSTRATOR's data and is suspected to originate from a surface contamination of the PTFE components used in the DEMONSTRATOR's detector units (DUPTFE). This component group includes the bushing used to hold the contact pin that reads out the detector's p+ electrode, which has a line of sight to the detector passivated surface. A number of small support structures and nuts located near the detector are also included in the DUPTFE group. If radon plated out onto these surfaces during the construction or deployment of the

experiment, it could contribute 5.3 MeV alphas and 47 keV gammas throughout the experiment's lifetime. While other materials located near the detectors could also serve as the source of the observed ^{210}Pb gamma, a surface contamination in the PTFE should act as a reasonable proxy for any source internal to the cryostat. This is because in the standard fitting energy range between 100 and 2620 keV, the ^{210}Pb spectrum contains no prominent gamma peaks, but does contribute continuum backgrounds. It is consequently an important component to include in the background model in order to correctly fit other components, but it is difficult to constrain well with the features in the standard energy range.

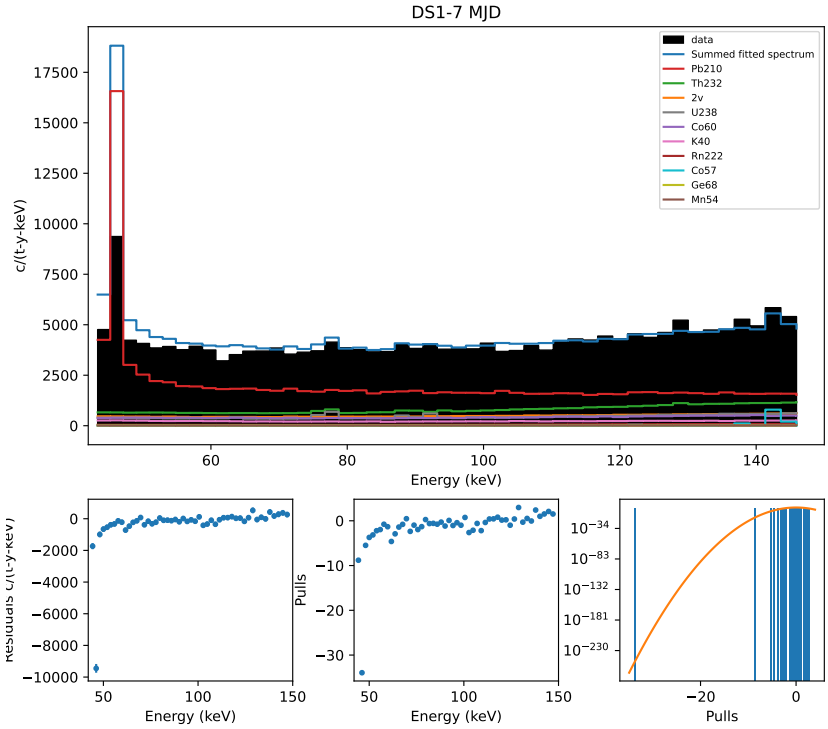


Figure 6.17: Results of background model fits with the standard 100 keV threshold projected down to lower energies for all detectors (enriched and natural). The ^{210}Pb DUPTFE spectrum is floated independently for different detector groups, and only multiplicity one data is included in the fit. The ^{210}Pb 47 keV peak displays some disagreement between model and data.

Although incorporating the spectrum below 100 keV into final fits is not feasible for the reasons discussed in Sec. 6.1, a fit with an energy threshold slightly above 40 keV can be used to evaluate how well standard fits do at estimating the ^{210}Pb DUPTFE contribution. Fig. 6.17 shows how the model from a fit performed with a 100 keV threshold performs when projected to lower energies. The strength of the 47 keV peak differs by approximately a factor of two between data and the model resulting from a 100 keV threshold fit. Fig. 6.17 combines enriched and natural detectors, but a similar level of disagreement is observed across

detector groups, even though the ^{210}Pb DUPTFE spectrum was allowed to float independently for each detector group. This choice was made to account for the fact that the strength of the 47 keV line is highly detector-dependent, suggesting different amounts of plate-out on different PTFE pieces. Another reason for detector dependence in the strength of the 47 keV peak may be differences in the passivated surface profile between detectors, particularly between enriched and natural detectors. These differences are challenging to fully account for in simulations, since sources are typically assumed to be uniform within a component group. Given these factors, the difference in 47 keV peak between model and data could indicate that ^{210}Pb DUPTFE is being overestimated in the model or could relate to surface effects and non-uniform activity distributions not being properly modeled.

One proposed method for taking into account information from the 47 keV peak in standard fits was to estimate the ^{210}Pb DUPTFE specific activity for each detector group based on low threshold fits and use it to introduce Gaussian penalty terms to the negative log likelihood during standard fits. This method was tested using Gaussian penalties with two different standard deviations: 10% of the mean and 50% of the mean. The 10% standard deviation was chosen to reflect a constraint that is unrealistically tight given our level of confidence in our ability to model this source of ^{210}Pb at low energies. The results of the fit with tightly constrained DUPTFE ^{210}Pb is shown in Fig. 6.18. Even in this extreme case, the 46.5 keV is significantly overestimated by the model, although the model performs reasonably well in the region above the peak. In tests of the looser constraint, where the standard deviation is 50% of the specific activity, the Gaussian penalty had only a minor impact on fit results. The fit DUPTFE ^{210}Pb activity only differed from unconstrained fits by 5% in M1 enriched detectors and 8% in M2 enriched detectors, which does not significantly improve agreement with the 47 keV peak. Since a penalty term with a realistic level of uncertainty did not largely impact the results of fitting, no penalty on the NLL was included when performing final fits.

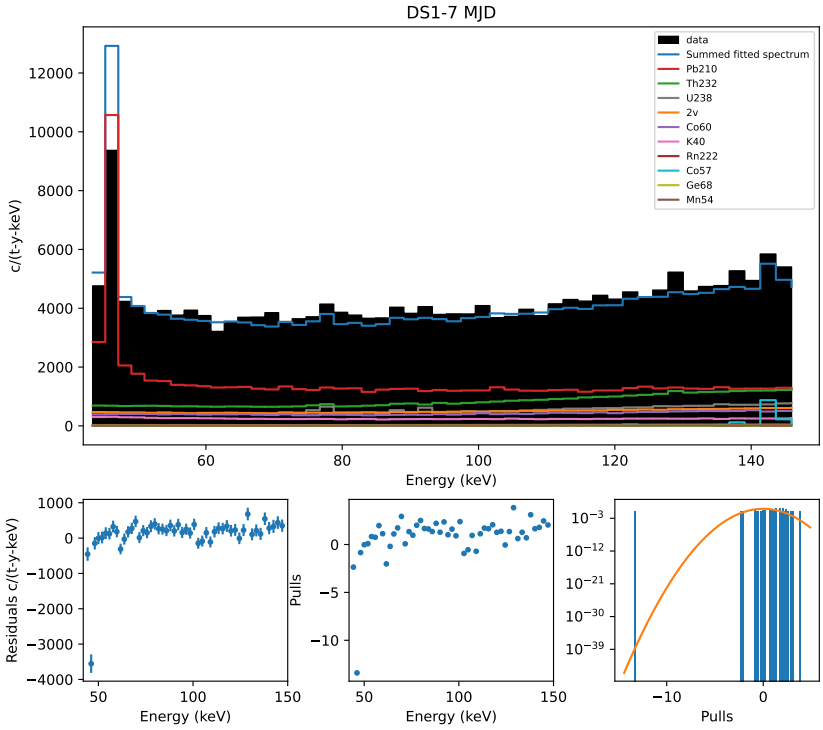


Figure 6.18: Results of background model fits with the standard 100 keV threshold projected down to lower energies in fits where tight (10% stan. dev.) Gaussian constraints are imposed on the ^{210}Pb DUPTFE activities based on the results of fits with a lower energy threshold. Multiplicity one data from all detector groups is included in the plot.

Since there are significant uncertainties in our ability to accurately model the unknown source or near-detector ^{210}Pb, it was still important to test the overall impact of this pdf on the fit. Since standard fits model a higher value for the 47 keV peak than is observed in data, this test was done by excluding all sources of near-detector ^{210}Pb from the fit to study how big an impact underestimating this background source could introduce. The fitted model from this study shows higher rates in the $Q_{\beta\beta}$ background estimation window than standard fits: 21.7 cnts/(keV t yr) for the M1 Enr BI and 11.7 cnts/(keV t yr) for the M2 Enr BI. These rates are still only 0.4σ above the data rate in enriched detectors in M1 and 0.6σ high in M2, indicating that fit still gives reasonable results at high energies when near-detector source of ^{210}Pb are excluded from the model. The fact that the fit background indices are systematically higher in this study compared to the standard model makes sense because counts that would otherwise be fit into near-detector ^{210}Pb must instead be placed into pdfs for other decay chains that impact the spectrum more at high energies. The fitted ^{232}Th hot spot activity in this fit is ~9% higher here than in standard fits, again indicating that the absence of near-detector ^{210}Pb does not change most of the important qualitative conclusions drawn from

the background model. This study will be discussed in more detail in reference to the fitted $2\nu\beta\beta$ half-life in Sec. 6.8.3.

A final additional systematic test was performed with the ^{68}Ge activity in M2 enriched detectors fixed to a lower value to better match the amount of cosmogenic activation predicted by tracking how long enriched material spent above ground. For this study, the specific activity was constrained to be 2.8 μBq/kg at the start of DS0, a little over a factor of two lower than its typical fitted value.[3] This value was a very rough estimate based on the fraction of the M2 enriched exposure that was due to detectors manufactured from recycled material. While this calculation should not be taken as a precise or finalized value, it allowed for a systematic test of how much a better constrained ^{68}Ge activity in M2 would change other fitting results. Since ^{68}Ge for M1 enriched detectors did not contribute to the fitted model, it was not included in this study.

Constraining the M2 ^{68}Ge activity does not strongly impact the major conclusions from spectral fits. In fits with constrained M2 ^{68}Ge, the amount of ^{232}Th in the M2 background estimation window increases slightly to make up for the counts previously fitted into ^{68}Ge. The overall rates in the background estimation window demonstrate modest improvements over standard fits: the M1 Enriched BI in the model changes from 5.2% lower than the data BI to 3.8% lower than data, and the M2 Enriched BI changes from 4.6% lower than the data BI to 1.9% lower than the data BI. The fitted activity of the ^{232}Th hot spot in the M1 thermospyhon cavity also only shows modest changes, increasing by 3.8%. In general, all systematic studies performed showed that the background index of the fitted model and the source responsible for the ^{232}Th excess were reasonably robust against changes in the model.

Section 6.8: $2\nu\beta\beta$ Half-Life Determination

Fits to DS0-7 data result in a preliminary $2\nu\beta\beta$ half-life measurement of $(2.085 \pm 0.033) \times 10^{21}$ years. The following section describes the procedure used to arrive at this number, sources of uncertainty considered, and systematic tests performed, some of which may be incorporated into the total uncertainty in the future.

6.8.1: Best Fit Half-Life

The fitted activity density for $2\nu\beta\beta$ in enriched detectors was converted into a $2\nu\beta\beta$ half-life using the equation:

[3] While the beginning of DS0 is perhaps not the most logical value to use for Module 2 detectors that were not brought online at that time, the fitting algorithm reports all activities relative to this date.

$$\tau_{1/2}^{2\nu} = \frac{\ln 2 \cdot N_A \cdot f_{76}}{M_{enr} \cdot p_{2\nu} \cdot (3600 \cdot 24 \cdot 365.25)}, \tag{6.1}$$

where $f_{76} = 0.874$ is the isotopic abundance of ^{76}Ge in the DEMONSTRATOR's enriched PPCs [47], N_A is Avogadro's number, $M_{enr} = 75.668 \times 10^{-3}$ kg is the molecular weight of enriched germanium [47], and $p_{2\nu}$ is the fitted $2\nu\beta\beta$ activity density in enriched detectors in units of Bq/kg. During standard background model fits, the $2\nu\beta\beta$ specific activity is constrained to be the same for enriched detectors in both modules. The $2\nu\beta\beta$ specific activity in natural detectors is similarly constrained to be the same between both modules, but it is floated entirely independently of the specific activity in enriched detectors, and the calculated $\tau_{1/2}^{2\nu}$ is based only on the enriched detector specific activity, 7.326×10^{-5} Bq/kg. Before taking the DEMONSTRATOR's uncertainty into account, the resulting half-life of 2.085×10^{21} years is in slight tension with the most recent published $2\nu\beta\beta$ half-life from GERDA, $(2.022 \pm 0.018_{stat} \pm 0.038_{sys}) \times 10^{21}$ years [59].

6.8.2: Half-Life Uncertainty

Multiple factors in the $\tau_{1/2}^{2\nu}$ equation have non-negligible uncertainties that must be propagated to the final half-life uncertainty. First, the activity density $p_{2\nu}$ is impacted by both statistical and systematic uncertainties. The statistical uncertainty, σ_{stat}, is found through the profile likelihood technique described in Sec. 4.2. Fig. 6.19 shows the profile likelihood curve for the $2\nu\beta\beta$ specific activity in enriched detectors. The curve is very Gaussian, leading to symmetric upper and lower errors. The 1σ uncertainty, corresponding to ΔNLL=0.5, is $\pm 0.029 \times 10^{-5}$ Bq/kg. This 0.4% statistical uncertainty is smaller than the 0.9% statistical uncertainty on the GERDA $2\nu\beta\beta$ half-life [59]. However, the GERDA result only uses a small subset of their total exposure, 11.8 kg yr, compared to the 51.9 kg-yrs from enriched detectors used in this analysis. The limited exposure included in the GERDA analysis was chosen because a subset of their detectors had dead layer measurements taken both before and after the experiment's runtime, allowing for a major reduction in systematic uncertainties. In addition, the statistics in GERDA's fit are reduced by the application of a 565 keV low energy threshold, since below this energy beta decays from ^{39}Ar dominate the energy spectrum. These two factors mean that the DEMONSTRATOR's fitted data includes a higher number of $2\nu\beta\beta$ counts, reducing the statistical uncertainty.

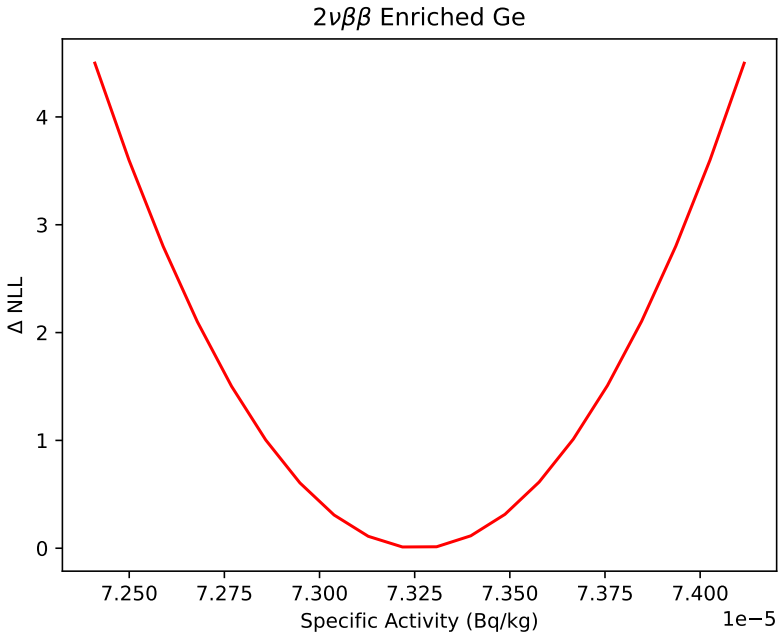


Figure 6.19: Profile likelihood curve for the $2\nu\beta\beta$ specific activity in enriched detectors in a standard fit. A 68% confidence interval corresponds to where ΔNLL crosses 0.5, which occurs between 7.297×10^{-5} and 7.355×10^{-5}

A number of systematic contributions to the uncertainty on the $2\nu\beta\beta$ specific activity are quantified through varying the parameters applied to simulations during post-processing. These include the dead layer shape and thickness, the energy-dependent efficiency of the delayed charge recovery cut, and the energy resolution and non-linearity. Many sets of pdfs were produced where all systematics parameters were randomly varied simultaneously but independently according to their uncertainties, following the procedure described in Section 5.1. After fits were performed to the DEMONSTRATOR's low-background data using each set of systematics pdfs, the post-processing systematic uncertainty σ_{post} was found by computing the 16% and 84% percentiles of fitted $2\nu\beta\beta$ activity densities. Using 100 sets of pdfs, the calculated confidence interval is $[7.233 \times 10^{-5}, 7.429 \times 10^{-5}]$, corresponding to a lower and upper uncertainty on the activity density of -0.093×10^{-5} and $+0.103 \times 10^{-5}$ respectively. Since there is not strong evidence of an asymmetry in the distribution, these quantities are averaged to give $\sigma_{post} = 0.098 \times 10^{-5}$. In the future, this uncertainty can be determined to higher precision by performing fits with 1000, rather than 100, sets of pdfs, and the choice to use a symmetric interval can be reevaluated. σ_{post} should be interpreted as a combined uncertainty due to all effects applied during post-processing.

The relative impact of different systematics parameters on the magnitude of σ_{post} was estimated by

analyzing correlation plots where the fitted $p_{2\nu}$ is plotted as a function of each systematic parameter value. As was described in detail in Chapter 5, each point in correlation scatter plots represents a single fit using a single set of systematics pdfs. The fitted activity density for $2\nu\beta\beta$ is plotted against the systematic adjustment factor applied to the systematic parameter of interest. (The adjustment factor was discussed in more depth in Chapter 5 but in most cases corresponds to how many sigma away from the optimal value the parameter is in a set of pdfs.) The values of the correlation plot's best fit line at adjustment factors of plus and minus 1 approximate the systematic parameter's contribution to the overall systematic uncertainty. Fig 6.20 shows the correlation plots corresponding to the dead layer full charge collection depth in enriched detectors and the overall DCR efficiency. These parameters are the dominant contributors to σ_{post}, and no other systematic parameters (including those governing the dead layer shape) display a convincing correlation with $p_{2\nu}$. In Fig 6.20, the percentiles lines (shown in red) show the total systematic uncertainty on $p_{2\nu}$, while the projection lines (shown in green) represent the plotted systematic parameter's contribution to the uncertainty based on the values of the correlation line at adjustment factors of ± 1.

When interpreting the size of σ_{post}, it is worthwhile to make a note regarding the dead layer full charge collection depth systematic. It is possible to make a correlation plot comparing the total integrated counts fitted into a pdf to the systematic adjustment factor rather than plotting the pdf's specific activity on the y-axis. This was done for the fccd in enriched detectors in Fig. 6.21. It is interesting to note that while $2\nu\beta\beta$ activity density is positively correlated with the size of the fccd, the number of fitted $2\nu\beta\beta$ counts is negatively correlated. This means that as the size of the dead layer increases, some counts are moved from the $2\nu\beta\beta$ pdf to other pdfs in the final fit. If the number of fitted counts in the $2\nu\beta\beta$ pdf was instead constant, the activity density would need to increase by more to account for the larger dead layer, which would have led to a larger systematic uncertainty.

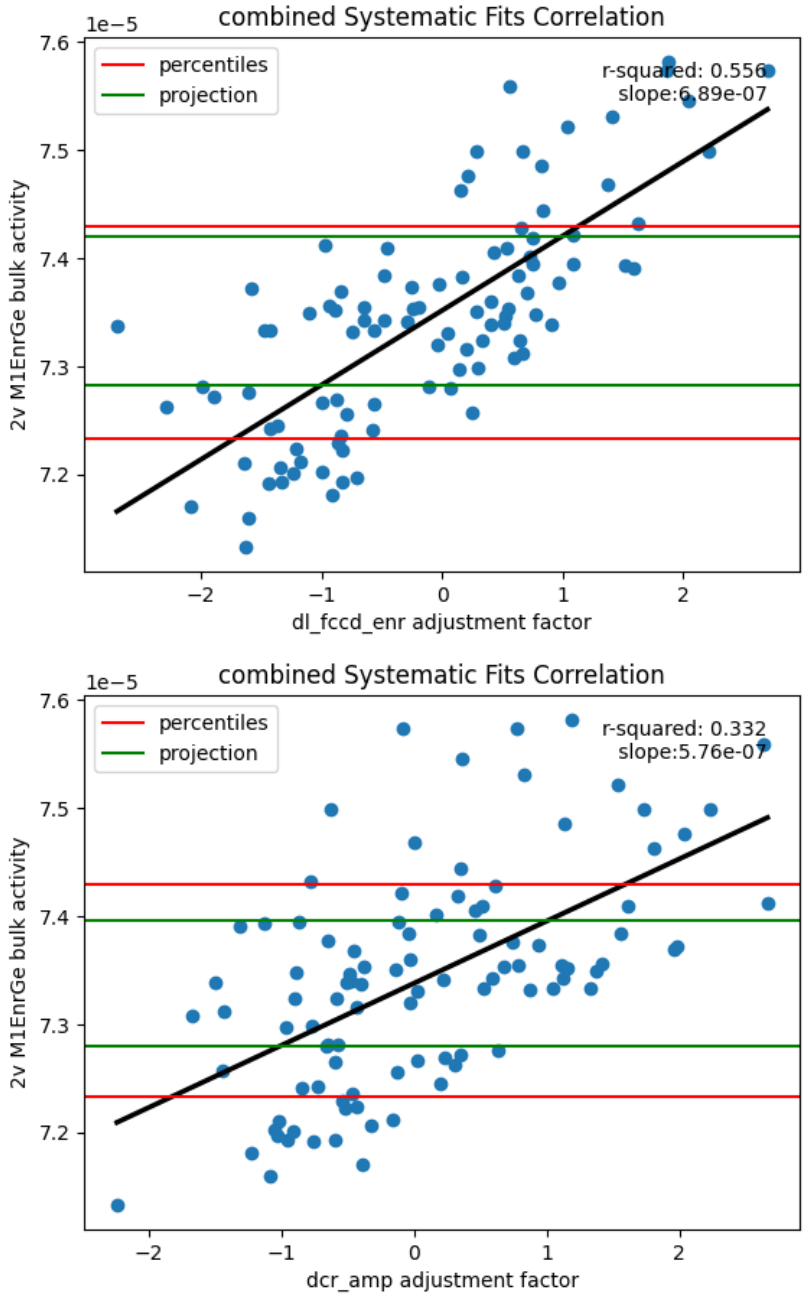


Figure 6.20: Correlation plots of the fitted $2\nu\beta\beta$ specific activity vs. systematic adjustment factors for the two dominant systematics effects varied during post-processing. The top plot shows how the $2\nu\beta\beta$ specific activity varies with the dead layer full charge collection depth in enriched detectors, while the bottom plot shows the variation as a function of the DCR cut efficiency. The projection lines show the portion of the uncertainty related to the particular systematic parameter, while the percentiles lines represent the full systematic uncertainty for all post-processing effects.

An additional source of systematic uncertainty on the fitted activity density is the uncertainty in the simulated MaGe geometry. Because this is not a systematic that can be applied during simulation post-processing, it cannot be included in the main systematics framework described in Chapter 5. While all

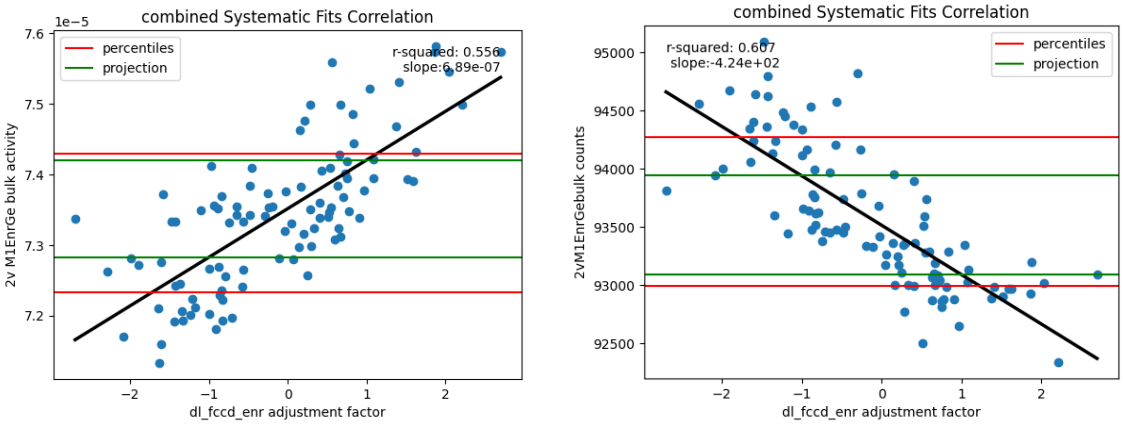


Figure 6.21: Comparison of correlation plots showing the fitted activity density (left) and the fitted number of model counts (right) associated with $2\nu\beta\beta$ in M1 enriched detectors as a function of dead layer full charge collection depth.

components in the MaGe model have some associated uncertainty because the geometry of every component of the Demonstrator is only known to within its design tolerance, nonuniformities in the cryostat thickness were deemed most likely to have a significant impact on fitting results. Varying the cryostat thickness should mainly impact what fraction of backgrounds from a given decay chain originate outside the cryostat, but it can also have an indirect effect on the fitted $2\nu\beta\beta$ activity. To quantify this uncertainty, first the magnitude of the cryostat thickness uncertainty was estimated by I. Guinn based on comparing the detector distribution of events in ^{228}Th calibrations between data and simulations with varying cryostat thickness. This method gave an estimated 1σ uncertainty of 1.37 mm [81]. While this value is large relative to the total cryostat thickness, which has a maximum of 6.6 mm in the simulated geometry, the estimate was intended to act as a proxy for all uncertainties in the simulated model rather than just accounting for the design tolerance of the cryostat.

To determine the impact of this level of cryostat thickness uncertainty on the fitted $2\nu\beta\beta$ half-life, ^{228}Th M1 calibration source simulations were generated with seven different cryostat thicknesses between -1 and +2.[4] For each cryostat thickness, a high statistics simulated dataset was generated from ^{228}Th line source and $2\nu\beta\beta$ pdfs. These simulated datasets were then fit using $2\nu\beta\beta$ pdfs and thorium pdfs from all components including the calibration line source, simulated using the standard cryostat thickness. The goal of this exercise was to determine how many counts in the fitted model were moved into or out of the $2\nu\beta\beta$ spectrum as a result of an incorrectly modeled cryostat thickness. The ^{228}Th line source pdf acts as a proxy

[4]This asymmetric range was used because of a problem generating simulations with thinner cryostat values that can be corrected in the future.

for all radiation sources originating outside the cryostat. By using simulated datasets that only included one background source external to the cryostat, this procedure limited the number of simulations that needed to be generated for each cryostat thickness, making this systematic computationally feasible. For each fit, the number of fitted $2\nu\beta\beta$ counts was compared to number of $2\nu\beta\beta$ counts in the model used to generate the simulated dataset. This difference was then expressed as a fraction of the number of the number of counts originating outside the cryostat in the simulated dataset model in order to allow it to easily be scaled to the DEMONSTRATOR's data. The percentage change in fitted $2\nu\beta\beta$ counts was found to vary relatively linearly with cryostat thickness, although decreases below the nominal cryostat thickness had a larger impact than increases above it. The size of the effect at ± 1.37 mm was estimated, using the steeper slope for decreasing thicknesses to form a conservative symmetric uncertainty. When scaled to the number of counts fitted to sources outside the cryostat in the DEMONSTRATOR's standard fits, the cryostat thickness uncertainty was found to change the total number of $2\nu\beta\beta$ counts, $c_{2\nu} = 126,633$ counts, by $\sigma_c = 610$ counts, leading to a 0.5% uncertainty to the fitted $2\nu\beta\beta$ activity. This number is still being refined but should give a reasonable estimate of the expected size of this effect.

In addition to the uncertainty on the fitted activity density, the $2\nu\beta\beta$ half-life uncertainty depends on the uncertainty of the ^{76}Ge isotopic abundance. Ref. [47] calculates an uncertainty of $\pm 0.5\%$ on the isotopic abundance of 87.4% for the DEMONSTRATOR's PPCs. Since the background model fits only include data taken prior to and during the Module 2 upgrade, the isotopic abundance in ICPCs was not used in calculating the half-life limit. Ref. [47] also calculates an uncertainty on the molecular weight of enriched detectors but it is small enough (0.01% of the molecular weight) that its effect on the half-life uncertainty is negligible.

While detector mass is not explicitly referenced in Eq. 6.1, each fitted activity density implicitly depends on the mass of the source component, which has an associated uncertainty. During fitting, pdfs are weighted by the source mass in order to allow the specific activity to act as the floated parameter. In the case of the $2\nu\beta\beta$ pdfs, the source component is enriched germanium detectors, which have a combined mass of $m_{enr} = 29.71$ kg (16.82 kg in M1 and 12.89 kg in M2). This mass is the sum of the physical masses of all enriched detectors rather than an active mass adjusted to subtract dead regions of the detectors. It also includes detectors that were non-operational throughout the runtime of the DEMONSTRATOR, since primaries in the simulations were generated in all enriched germanium regardless of whether the detector is biased. As described in Sec. 3.4.3, simulation post-processing corrects the simulations to account for the changing percentage of detectors that were non-operational or veto-only over different portions of data-taking. The uncertainty in detector mass

used for this analysis is consequently much smaller than the active mass uncertainty incorporated in the $0\nu\beta\beta$ half-life lower limit. In this analysis, uncertainties related to the size and shape of the dead layer are incorporated into σ_{post} through varying dead layer parameters in post-processing instead of being considered a part of the mass uncertainty. Following the procedure of the DEMONSTRATOR's double beta decay to excited states result, a 1 gram uncertainty was assumed for each of the 35 enriched detectors, leading to an uncertainty on the enriched mass of $\sigma_{mass} = 0.035$ kg.

Equation 6.2 shows how all these sources of uncertainty were added in quadrature to achieve a total uncertainty of 0.033×10^{21} years, or 1.6% of the best fit half-life:

$$\begin{aligned}
\sigma_{\tau_{2\nu}} &= \tau_{1/2}^{2\nu}\left\{\left(\frac{\sigma_{stat}}{p_{2\nu}}\right)^2 + \left(\frac{\sigma_{post}}{p_{2\nu}}\right)^2 + \left(\frac{\sigma_c}{c_{2\nu}}\right)^2 + \left(\frac{\sigma_{f_{76}}}{f_{76}}\right)^2 + \left(\frac{n_{det}\cdot\sigma_{mass}}{m_{enr}}\right)^2\right\}^{\frac{1}{2}} \\
&= 2.085\times10^{21}\text{ yr}\left\{\left(\frac{0.029}{7.326}\right)^2 + \left(\frac{0.098}{7.326}\right)^2 + \left(\frac{610}{126,633}\right)^2 + \left(\frac{0.005}{0.874}\right)^2 + \left(\frac{35\cdot0.001}{29.7054}\right)^2\right\}^{\frac{1}{2}} \\
&= 2.085\times10^{21}\text{ yr}\cdot 0.0159 \\
&= 0.033\times10^{21}\text{ yr}
\end{aligned} \tag{6.2}$$

6.8.3: Systematic checks

The systematic tests described in Sec. 6.7 were also used to test the robustness of the calculated $2\nu\beta\beta$ half-life. For example, the relationship between the fitted $2\nu\beta\beta$ activity density and the low energy threshold employed during the fit is shown in Fig 6.22. Compared to the study in Sec. 6.7, the range of energy thresholds tested was expanded to more fully explore the observed dependence. The lowest tested energy threshold of 44 keV was chosen to test the impact of including the 46.5 keV ^{210}Pb peak in the fitting range, while the highest considered threshold, 1170 keV, was chosen to be above the endpoint of the ^{210}Bi spectrum. A 565 keV threshold was also tested to approximately match GERDA's fitting theshhold.

For thresholds below approximately 400 keV, the fitted $2\nu\beta\beta$ activity density appears to exhibit a linear dependence on the low energy threshold used in fitting. Conversely, above ~400 keV, the energy threshold does not significantly alter the $2\nu\beta\beta$ result. The goodness of fit procedure described in Sec. 6.3 was repeated for the 565 keV threshold fit to evaluate the overall performance of a fit with a higher energy threshold. Based on a study with 300 simulated datasets, the fit to the DEMONSTRATOR's data was found to achieve a lower NLL than 1.3% of simulated datasets, a modest improvement over the 0.3% found for a 100 keV threshold

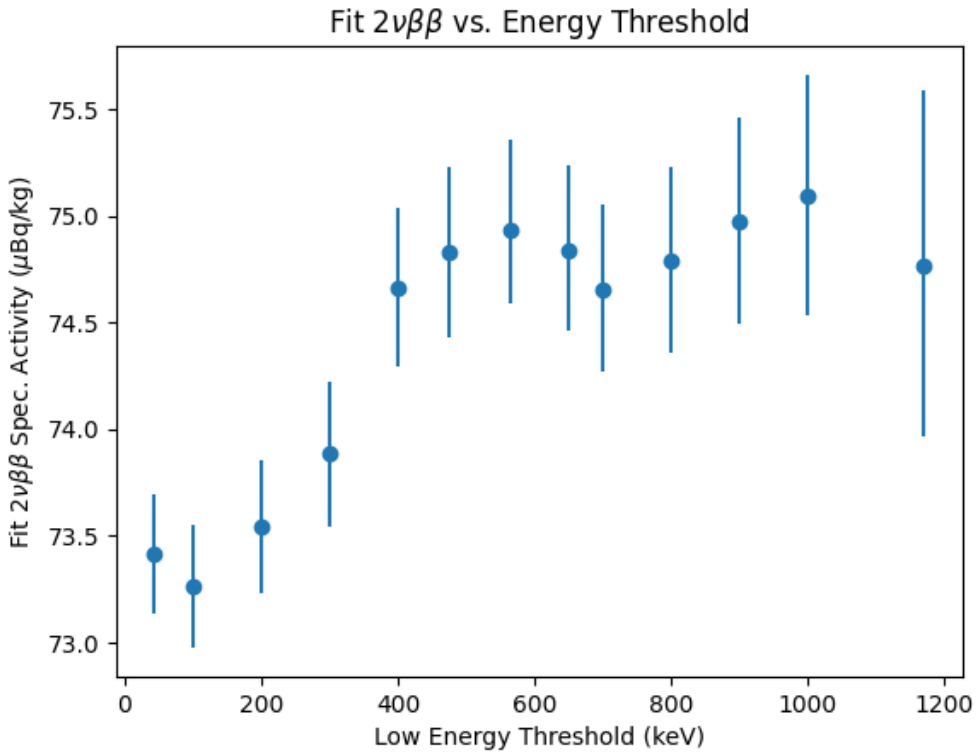


Figure 6.22: Dependence of the fitted $2\nu\beta\beta$ specific activity on the low-energy threshold used in the fit. The error bars on each data point reflect only its statistical uncertainty.

fit. In this study, only the fitted energy region is used in the NLL calculation, so the negative log likelihoods are not directly comparable between fits with different thresholds, but the p-values should provide a valid reference for comparison. The improved figure of merit and increased stability over changes in threshold for higher threshold fits may suggest that these results are more reliable, but this question is still being actively investigated.

To interpret the different $2\nu\beta\beta$ results for fits with different low energy thresholds, it is useful to investigate what factors may be driving the lower threshold fits to lower $2\nu\beta\beta$ activities than are favored by the high energy region of the spectrum. Fig. 6.23 compares the spectra from 100 to 600 keV between standard fits and fits with a 565 keV threshold. In fits with a 565 keV threshold, most of the plotted energy range was not utilized during fitting, so the fitted model was extrapolated to lower energies. In both cases, low energy ^{232}Th and ^{238}U peaks display reasonable agreement between model and data. However, in the fit with a 565 keV threshold, the model significantly underpredicts the data below 230 keV where the $2\nu\beta\beta$ spectrum no longer dominates. This is mainly a result of reduced activity in the near-detector ^{210}Pb pdfs for the 565 keV threshold fits, a result which is better demonstrated by Fig. 6.24. Fig. 6.24 shows the integrated count rate in

different fitted models above 565 keV divided by decay chain. Comparing the 100 keV threshold model (in blue) to the 565 keV threshold model (in red), most of the difference in the total number of $2\nu\beta\beta$ counts is compensated by ^{210}Pb. The ^{210}Pb pdf that accounts for the spectral strength at low energy in Fig. 6.23 also has a small contribution at higher energies, which pushes counts out of the $2\nu\beta\beta$ pdf in the fitted model. The third model shown in Fig. 6.24 is a 100 keV threshold fit where all near-detector ^{210}Pb sources were constrained to zero activity to test how the fit performs when ^{210}Pb was not allowed to make up as much of the strength at low energy. In this case, the fitted $2\nu\beta\beta$ value is greater than in standard 100 keV threshold fits, but not as large as in high threshold fits. Pdfs from other decay chains, mainly ^{238}U, were pushed to higher values to account for some of the missing strength at low energy. One possible explanation of the threshold dependence of the fitted $2\nu\beta\beta$ value is that the source of low energy strength in the data is not properly included in the model.

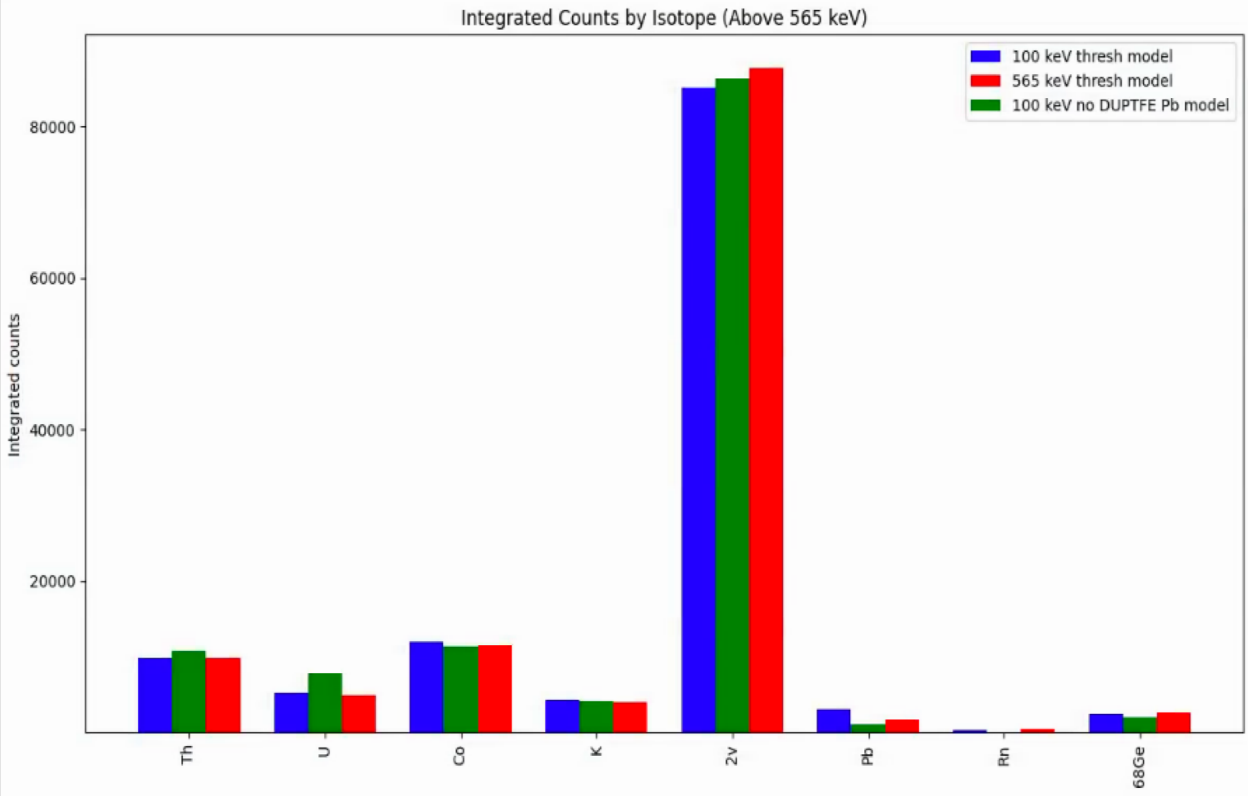


Figure 6.24: Comparison of the decay chain distribution of integrated counts above 565 keV for three different models. Two of these models were from fits with different thresholds (100 keV vs 565 keV), while the third model was from a fit with a 100 keV threshold where all potential source of near-detector ^{210}Pb were constrained to have no activity

The study omitting all sources of ^{210}Pb in detector unit PTFE conservatively estimates the effect of inaccuracies in modeling the source of the 46.5 keV peak has on the fit. Other systematic studies exploring

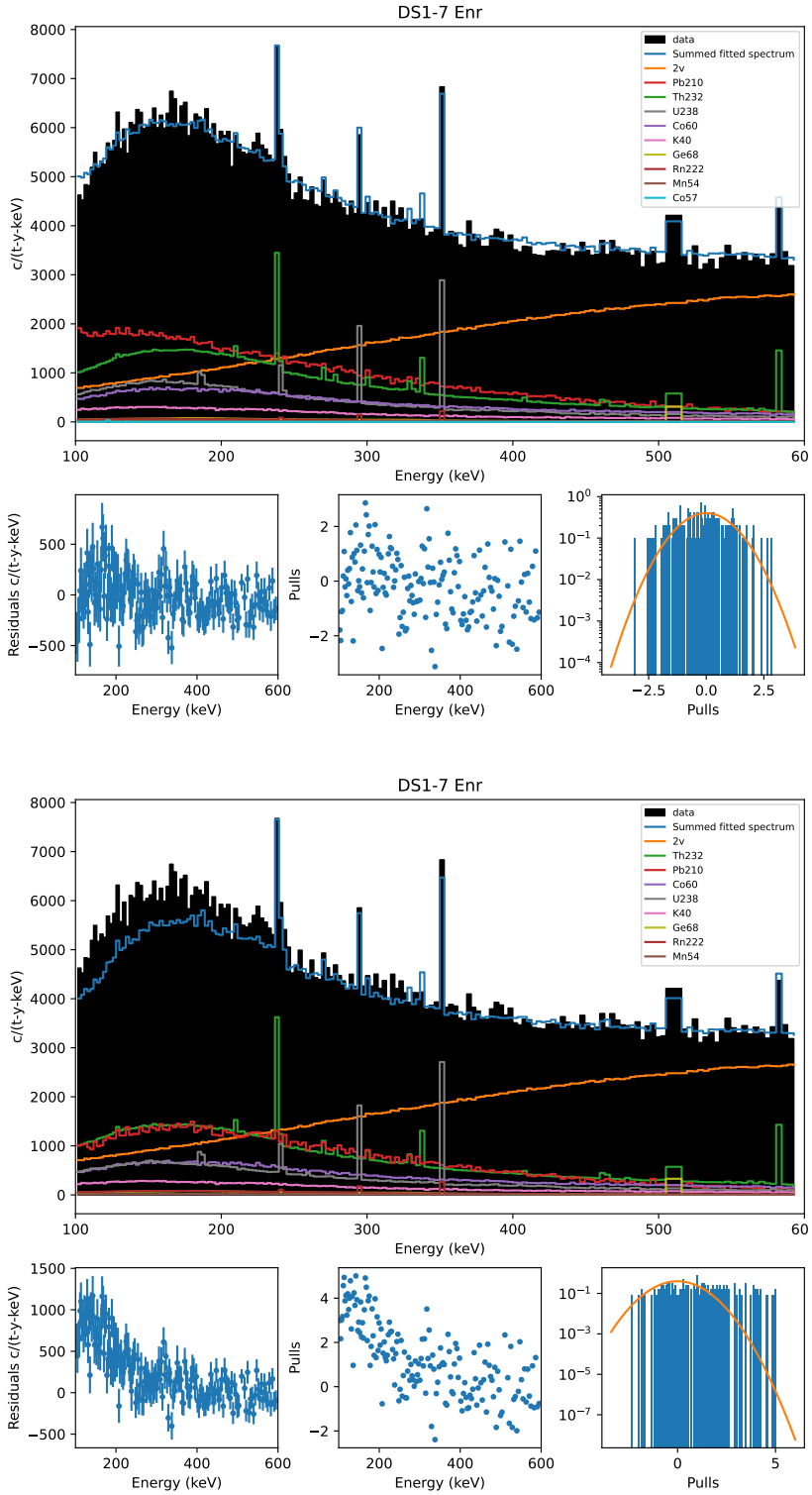


Figure 6.23: Comparison of the low energy spectrum resulting from a fit with the standard 100 keV threshold (top) and from projecting the model from a fit with a 565 keV threshold down to 100 keV. DS1-7 data in all enriched detectors is plotted together. DS0 data and data from natural detectors was simultaneously fitted but is not included in the plot.

Omitted/Fixed pdfs	Fitted $2\nu\beta\beta$ Spec. Act. ($\times 10^{-5}$ Bq/kg)	% Change from Standard Fits
None	7.326	N/A
Th M1 Interface Cavity	7.293	-0.45%
Pb DUPTFE (all)	7.376	+0.68%
^{68}Ge M2 Enr	7.322	-0.05%

Table 6.4: Variation in fitted $2\nu\beta\beta$ activity density due to changes in which pdfs are included (or allowed to float) in the fit. The percent difference from standard (100 keV threshold) fits is shown in the last column. In most cases, the indicated pdf(s) were fixed to zero activity during fitting, but in the final row, ^{68}Ge was fixed to a non-zero value based on expectations from tracking cosmogenic activation.

	Fitted $2\nu\beta\beta$ Spec. Act. ($\times 10^{-5}$ Bq/kg)	% Change from Standard Fits
Both modules (Standard)	7.326	N/A
Module 1	7.311	-0.20%
Module 2	7.371	+0.61%

Table 6.5: Results of systematic study where the $2\nu\beta\beta$ activity density in enriched detectors was floated independently between the two modules. The percent difference from standard fits (where the specific activity is constrained to be the same in both modules) is shown in the last column.

the impact of removing certain pdfs from the fit were also used to determine how sensitive the fitted $2\nu\beta\beta$ result is to uncertainties in modeling. For example, fits excluding the ^{232}Th hot pdf in the M1 thermosyphon cavity, the suspected source of the ^{232}Th excess over assay expectations, serve as a conservative estimate of the impact that mis-modeling the source of the ^{232}Th excess has on the fitted $2\nu\beta\beta$ half-life. Fits were also performed constraining the M2 ^{68}Ge activity in enriched detectors, since this background source may be overestimated in standard fits. All of these changes to the included pdf list only resulted in modest changes in the fitted $2\nu\beta\beta$ activity density, as shown in Table 6.4.

In addition to the systematics described in Sec. 6.7, fits were performed floating the $2\nu\beta\beta$ decay specific activity independently between the two modules. These fits included data from both modules, but the $2\nu\beta\beta$ specific activity was not constrained to be the same between M1 and M2 enriched detectors. The results of this study are shown in Table 6.5. The variation in fitted activity density between the two modules is less than 1%, even though Module 2 contains only 24% of the enriched exposure, supporting the validity of the constrained result.

As a final systematic check, a $2\nu\beta\beta$ half-life was derived from the fitted activity density in natural detectors of 6.372×10^{-6} Bq/kg. Using the ^{76}Ge abundance of 7.75% for natural germanium and the natural atomic weight of 72.630×10^{-3} kg/mol leads to a half-life of 2.22×10^{21}, which is about 6.5% higher than the half-life derived from enriched detectors. However, the smaller number of $2\nu\beta\beta$ decays in natural germanium leads to a much larger statistical uncertainty of 0.28×10^{-6} or 4.46%. The post-processing

systematic uncertainty in natural germanium does not appear to be symmetric around the best fit value from standard pdfs ($[6.173 \times 10^{-6}, 6.761 \times 10^{-6}]$). When translated into an uncertainty on the half-life, the statistical plus systematic uncertainty lower uncertainty is -7.5%, while the upper uncertainty is +5.4%, meaning that the result from natural detectors is consistent with the result from enriched detectors.

Based on the results of some of the systematic studies described in this section, the final estimate of the uncertainty on the DEMONSTRATOR's $2\nu\beta\beta$ half-life will almost certainly increase beyond the 1.6% uncertainty calculated in Sec. 6.8. Some or all of the studies varying the pdfs included during fitting, summarized in Table 6.4, may be incorporated as a modeling uncertainty. Based on the amount of variation in fitted $2\nu\beta\beta$ specific activity, this systematic would be a sub-dominant contribution that would not have a large impact on the total magnitude of the uncertainty. More importantly, the 2-3% variation in the fitted $2\nu\beta\beta$ half-life as a function of low energy threshold, shown in Fig. 6.22, indicates a source of uncertainty that needs to be properly incorporated. In this work, 100 keV threshold fits were considered the standard result, but further investigation is underway to determine which energy threshold should be used when reporting the best fit half-life and how the threshold dependence should contribute to the total uncertainty. It is worth noting that the higher fitted activity densities for fits with higher thresholds result in $2\nu\beta\beta$ half-lives that are in better agreement with the measured GERDA half-life. For example, the fitted $2\nu\beta\beta$ activity density from the 565 keV threshold fit corresponds to a half-life of 2.039×10^{21} years, which is well within uncertainty of the GERDA half-life of $(2.022 \pm 0.018_{stat} \pm 0.038_{sys}) \times 10^{21}$ years [59]. While this should not be taken as a reason to favor higher threshold fits, it is a good indication that incorporating the energy threshold dependence into the total uncertainty should lead to reasonably consistent results.

CHAPTER 7: Summary and Future Directions

The low background rate achieved by the DEMONSTRATOR made background model development a challenging enterprise. This was particularly true because much of the DEMONSTRATOR's background mitigation was achieved by passive means, limiting the amount of data available for analysis. Since the DEMONSTRATOR's $0\nu\beta\beta$ analysis did not rely on background modeling, this difficulty did not impact the $0\nu\beta\beta$ result. However, the development of a detailed background model was important for a number of the DEMONSTRATOR's other goals, such as informing the design of the next-generation LEGEND experiment. Despite the challenge posed by performing a high dimensional fit to a low statistics dataset, frequentist fits, in conjunction with other background studies, have been able to draw a number of important conclusions about the composition of the DEMONSTRATOR's background spectrum and its implications for LEGEND.

Section 7.1: Summary of Important Findings

The DEMONSTRATOR's background model sought to measure the half-life of $2\nu\beta\beta$ in ^{76}Ge and to explain many observed features of the DEMONSTRATOR's data, including:

- The discrepancy between the measured background index in enriched detectors, $6.23^{+0.55}_{-0.52} \times 10^{-3}$ c/(keV kg yr), and the assay-based projection of $(1.16 \pm 0.04) \times 10^{-3}$ cts/(keV kg yr)
- The difference between the Module 1 BI of $(7.38 \pm 0.71) \times 10^{-3}$ cnts/(keV kg yr) and the Module 2 BI of $3.33^{+0.75}_{-0.67} \times 10^{-3}$ cnts/(keV kg yr)
- The elevated rate in ^{232}Th peaks and in the continuum for two natural detectors near the Module 1 crossarm

These goals were successfully achieved by performing frequentist fits to 51.9 kg-yrs of enriched exposure and 22.4 kg-yrs of natural exposure from the DEMONSTRATOR, including both open and previously-blinded data from multiple experimental configurations. More than 100 simulated spectra were used in fitting, including spectra from up to 10 decay chains in ~30 component groups. In general, background sources were assumed to be uniform over the volume (or surface) of a given component group, but hot spot sources in the thermospyhon cavities for both modules were also included. Key conclusions from spectral fits and

supplementary background studies are summarized in Sec. 7.1.1-7.1.4.

7.1.1: Performance of the Fitted Model

- Although fits to the DEMONSTRATOR's data did not result in as good an NLL as most fits to simulated datasets, where the underlying model was known to be perfect, the fitted model accurately reproduced the spectral shape of the DEMONSTRATOR's data across a wide range of energies.
- The fitted model successfully accounts for differences in the $0\nu\beta\beta$ background estimation window event rate between different detector groups.

7.1.2: $2\nu\beta\beta$ Half-Life Result

- The half-life for two-neutrino double-beta decay in ^{76}Ge was measured to be 2.085×10^{21} years.
- A number of statistical and systematic contributions to the half-life uncertainty were identified and quantified, leading to a 1.6% total uncertainty. Dominant sources of uncertainty included the size of the dead layer full charge collection depth and the efficiency of the surface alpha cut.
- A dependence on the low energy threshold of the fit is still under investigation, which could impact the final half-life by up to $2-3\%$ and which will lead to an increased uncertainty.
- The fitted $2\nu\beta\beta$ was robust under other systematic studies.

7.1.3: Localized Background Source

- The difference in background index between enriched detectors in the DEMONSTRATOR's two modules can be explained by a localized ^{232}Th source on the bottom inner surface of the Module 1 thermosyphon cavity.
- No other considered ^{232}Th hot spot location was able to explain the peak ratios and spatial distribution of the DEMONSTRATOR's data with as much accuracy. In particular, near-detector hot spots below the cold plate displayed significant disagreement.
- Although assays of the thermosyphon cavity found no evidence of a hot spot, this finding cannot rule out a transient background source in the region that evaporated or was displaced upon opening the experiment.

- Spectral fits that did not include a localized ^{232}Th simulated source reproduced many aspects of the DEMONSTRATOR's data with reasonable fidelity but led to increased tension with the spectrum of the two high rate natural detectors.

- Excluding the hot spot ^{232}Th source from the fit only changed the fitted $2\nu\beta\beta$ half-life by approximately 0.5%, indicating that the choice of hot spot source location in the model has only a minor impact on the final $2\nu\beta\beta$ result.

7.1.4: Implications for LEGEND

- The dominant source of ^{232}Th events in the DEMONSTRATOR, the M1 thermosyphon cavity, does not have a counterpart in LEGEND and consequently does not pose a problem for LEGEND-200.

- Although the DEMONSTRATOR's background essay over assay projections cannot be entirely explained by the M1 thermosyphon cavity source, spectral fits suggest that the near-detector components, such as the LMFEs and detector unit electroformed copper, are minor contributors to the rate in the background estimation window. This finding takes into account backgrounds from all decay chains.

- This result suggests that similar near-detector components should be acceptable for use in LEGEND-200.

Section 7.2: Possible Improvements

The most important next steps in the background model development involve finalizing the uncertainty calculation for the fitted $2\nu\beta\beta$ half-life. First, additional tests will be performed using both data and data-driven simulated datasets, like those generated for the goodness of fit calculation, to further investigate the energy threshold dependence of the fitted $2\nu\beta\beta$ rate in data. These studies may aid in determining whether the reported $2\nu\beta\beta$ half-life should be based on fits with a 100 keV threshold or higher threshold fits. The difference between fits with different energy thresholds will be incorporated into the reported uncertainty, but further consideration is needed to decide how best to do so and what other systematic studies might influence the final uncertainty number. In addition, the post-processing portion of the systematic uncertainty will be updated based on performing 1000, rather than only 100, fits with different sets of systematics pdfs. This will improve the precision on this portion of the uncertainty.

Looking further into the future, there are additional directions that could be explored to improve the performance of spectral fits. A few of these possible directions are discussed in Sec. 7.2.1-7.2.3. Some of

these changes could potentially be incorporated into the DEMONSTRATOR's background model, but others would be more realistic to consider on the timescale of LEGEND-200.

7.2.1: Combined Component Group Contributions

One potential improvement that would not change the results of spectral fits but could improve our ability to draw conclusions from them relates to combining fit results from different component groups. For example, it could be useful to quantify with uncertainty how much of the total background index can be explained by near-detector ^{232}Th components. It is not difficult to calculate the count rate attributable to a particular region of the experiment in the best fit model, but finding the uncertainty on such a number is more involved. Since the fitted activities for components located in the same region tend to be highly correlated, their uncertainties cannot be assumed to be independent and straight-forwardly combined.

One way to calculate a combined uncertainty on the total number of counts from a group would be to do a profile likelihood analysis where the number of counts in the group was consider the parameter of interest. This would require a reparametrization of the fit during profiling. For pdfs considered part of the group, the fraction of the total group counts coming from each pdf would be floated instead of floating its activity density. Lagrange multipliers would be necessary to enforce the constraint that the floated fractions add up to 1. If this kind of analysis is deemed important to drawing conclusions about the background composition, this procedure could be implemented and tested.

7.2.2: Further Subdividing Spectra

Other areas of improvement would have the potential to improve the accuracy or precision of spectral fits, rather than simply changing how results are presented. One such area involves how the data are subdivided during fitting. Although fits to the DEMONSTRATOR's data benefit from subdividing the data into multiple spectra that are simultaneously fit, fit performance degrades if the data are subdivided too finely. This limitation led to detectors being combined into groups during fitting, preventing spectral fits from fully utilizing information about the detector distribution of backgrounds. One possible recourse to address this would be to increase the size of non-peak bins. The variable binning scheme employed in frequentist fits is advantageous because it prevents gamma peaks from being split between multiple bins. However, in the continuum, where the spectra are relatively featureless, a larger bin size would increase bin statistics, reducing statistical fluctuations. This could be especially beneficial for the high multiplicity spectra, which have lower count rates than their multiplicity one counterparts. This strategy could be tested with simulated

datasets closely resembling the data to determine whether changing the binning scheme and splitting the detector groups improves the fitter's ability to identify backgrounds with reduced uncertainty.

In addition to subdividing detector groups, it would also be advantageous if spectra could be divided to separate single-site and multi-site events. However, as discussed in Sec. 3.4.3, in this case the issue is not just a matter of bin statistics but also of emulating the behavior of the AvsE cut well across the full fitting energy range. If the dt heuristic used to approximate the multi-site cut in simulations could be modified to match the data AvsE cut down to energies of 100 keV in ^{228}Th calibration data, it would be worth considering incorporating this cut into spectral spectra simultaneously being fit. This change could be particularly beneficial because the turnover of the $2\nu\beta\beta$ spectrum is more clearly defined in single-site data, potentially improving the estimated $2\nu\beta\beta$ half-life.

7.2.3: Passivated Surface Effects and Alpha Model

Another possible improvement in the long term involves the handling of passivated surface events. As was discussed in Sec. 3.4.3, a DCR cut is performed prior to fitting the DEMONSTRATOR's data to remove surface alphas from the spectrum. The probability of this cut removing bulk events as a function of energy is accounted for during simulation post-processing. If a well-validated model of the surface alpha spectrum in PPCs could be developed, this cut would not need to be performed, and the alpha contribution could instead be allowed to float during fitting. Reliably modeling passivated surface effects is challenging, however. The energy degradation of surface events displays a dependence on their distance from the p+ contact, and the theoretical model of signal formation for surface events is still under investigation but may depend on a uniform surface charge density of unknown magnitude. Work is ongoing within the LEGEND collaboration to better model passivated surface effects, both through collimated scans with alpha and beta sources at test stands and through waveform simulations that take into account the evolution of the charge cloud distribution during signal formation.

Even if an alpha pdf is not added into spectral fits, improving the modeling of passivated surface effects could improve the fidelity of simulated spectra in a few cases. Most importantly, these effects influence the pdfs of ^{210}Pb surface contaminations on components with a direct line of sight to the passivated surface. However, improving passivated surface modeling would rely on ongoing work by other analysts to generate a map of charge trapping in the passivated surface region as a function of radius and surface depth. This map could be modified on a detector-specific basis to accommodate different detector sizes. These detector-

dependent charge trapping maps could be used in the post-processing of relevant simulations. Since charge trapping efficiency maps depend on the amount of surface charge assumed, the impact on fitting results from varying the surface charge density would need to be incorporated into the systematic uncertainty analysis.

Section 7.3: Conclusion

A discovery of neutrinoless double-beta decay would mark a huge step in the journey to understand the nature of neutrinos and its implications for new physics beyond the Standard Model. To push the sensitivity of $0\nu\beta\beta$ searches into previously unexplored parameter space, the next generation of experiments must expand in size while simultaneously reducing their background rates. Improving on the ultra-low backgrounds of previous $0\nu\beta\beta$ experiments like the DEMONSTRATOR can best be accomplished by understanding what those experiments did well and where they can be improved. Background modeling plays an important part in this continuous effort to learn from each generation of experiments how to better design the next.

APPENDIX A: FITTED COUNTS

parameter	Counts	% Counts
2v_M1EnrGe_bulk	9.3352e+04	32.244
2v_M2EnrGe_bulk	2.9800e+04	10.293
PbBrem_RadShieldAssembly_001_RadShieldPb_001_bulk	2.9527e+04	10.199
Th_M1CPInterfaceCavityBottomSurface_bulk	2.4370e+04	8.417
Co_RadShieldCuOuter_bulk	1.3782e+04	4.760
Pb_M1DUPTFE_Enr_surf	8.0937e+03	2.796
U_M1CryostatCopperNear_bulk	7.1090e+03	2.455
Co_M1NatGe_bulk	7.0307e+03	2.428
Co_M2NatGe_bulk	6.1380e+03	2.120
U_M2CryostatCopperNear_bulk	5.4699e+03	1.889
Th_M2Seals_bulk	4.7843e+03	1.652
K_M1Connectors_bulk	4.5809e+03	1.582
Th_M1CryostatCopperFar_bulk	4.0318e+03	1.393
Pb_M2DUPTFE_Enr_surf	3.7272e+03	1.287
U_M1StringCables_bulk	3.7212e+03	1.285
Pb_M2DUPTFE_Nat_surf	2.9397e+03	1.015
Th_M1Connectors_bulk	2.8172e+03	0.973
2v_M2NatGe_bulk	2.4911e+03	0.860
68Ge_M1NatGe_bulk	2.4776e+03	0.856
K_M1StringCables_bulk	2.2661e+03	0.783
2v_M1NatGe_bulk	2.1128e+03	0.730
Pb_M1DUPTFE_Nat_surf	2.0296e+03	0.701
68Ge_M2NatGe_bulk	1.8066e+03	0.624
Co_M2CryostatCopperFar_bulk	1.7485e+03	0.604

Continued on next page

parameter	Counts	% Counts
K_M2StringCables_bulk	1.7061e+03	0.589
Rn_N2_bulk	1.6014e+03	0.553
Co_M1CryostatCopperNearWeldedParts_bulk	1.5885e+03	0.549
Th_M1StringCables_bulk	1.5176e+03	0.524
U_M1CryostatCopperFar_bulk	1.4687e+03	0.507
54Mn_RadShieldCuOuter_bulk	1.3683e+03	0.473
K_M2Connectors_bulk	1.3638e+03	0.471
U_M1Bellows_bulk	1.3253e+03	0.458
Th_M2CPInterfaceCavityBottomSurface_bulk	9.9853e+02	0.345
Th_RadShieldCuInner_bulk	8.0375e+02	0.278
Co_M1CryostatCopperFar_bulk	7.6076e+02	0.263
Th_M1Bellows_bulk	7.4958e+02	0.259
U_M2CPInterfaceCavityBottomSurface_bulk	7.3680e+02	0.254
U_M1Seals_DS0_bulk	6.9814e+02	0.241
U_M2Seals_bulk	6.0820e+02	0.210
Th_M1Seals_DS0_bulk	5.5723e+02	0.192
Pb_M1DUPTFE_Hot_surf	5.5566e+02	0.192
U_RadShieldAssembly_001_RadShieldPb_bulk	5.5290e+02	0.191
Th_M1Seals_DS12_DS345bc6abc_DS7_bulk	5.1340e+02	0.177
68Ge_M2EnrGe_bulk	4.4131e+02	0.152
Co_M1LMFEs_bulk	4.2177e+02	0.146
U_M1Connectors_bulk	4.0552e+02	0.140
K_M2CPInterfaceCavityBottomSurface_bulk	4.0234e+02	0.139
K_RadShieldAssembly_001_RadShieldPb_bulk	3.9987e+02	0.138
U_RadShieldCuOuter_bulk	3.9701e+02	0.137

Continued on next page

parameter	Counts	% Counts
Th_RadShieldCuOuter_bulk	3.7565e+02	0.130
K_M1Seals_DS0_bulk	2.7075e+02	0.094
U_M2Bellows_bulk	1.9886e+02	0.069
Th_M2ThermosyphonAndShieldVespel_bulk	1.6076e+02	0.056
57Co_M2NatGe_bulk	1.2025e+02	0.042
57Co_M1NatGe_bulk	1.2017e+02	0.042
K_M1Seals_DS12_DS345bc6abc_DS7_bulk	9.5009e+01	0.033
Th_M2CryostatCopperNear_bulk	1.6719e+01	0.006
68Ge_M1EnrGe_bulk	1.1744e+01	0.004

APPENDIX B: FITTED SPECIFIC ACTIVITY AND STATISTICAL UNCERTAINTY

parameter	Spec. Act. (μBq/kg)	Lower limit	Upper limit
2v_M1EnrGe_bulk	7.326e-05	7.298e-05	7.355e-05
2v_M1NatGe_bulk	6.372e-06	6.092e-06	6.653e-06
2v_M2EnrGe_bulk	7.326e-05	7.298e-05	7.355e-05
2v_M2NatGe_bulk	6.372e-06	6.092e-06	6.653e-06
54Mn_RadShieldCuOuter_bulk	3.279e-05	2.743e-05	3.846e-05
57Co_M1NatGe_bulk	2.945e-06	2.388e-06	3.525e-06
57Co_M2NatGe_bulk	3.218e-06	2.623e-06	3.838e-06
68Ge_M1EnrGe_bulk	4.219e-08	0.000e+00	6.247e-07
68Ge_M2EnrGe_bulk	6.313e-06	4.651e-06	7.953e-06
68Ge_M1NatGe_bulk	3.423e-05	3.164e-05	3.685e-05
68Ge_M2NatGe_bulk	3.070e-05	2.716e-05	3.427e-05
Co_M1CrossarmAndCPCables_bulk	5.516e-11	0.000e+00	1.191e-03
Co_M1CryostatCopperFar_bulk	1.107e-04	0.000e+00	2.441e-04
Co_M1CryostatCopperNearWeldedParts_bulk	1.271e-05	0.000e+00	2.205e-05
Co_M1EnrGe_bulk	3.927e-16	0.000e+00	1.111e-07
Co_M1LMFEs_bulk	3.430e-03	0.000e+00	6.813e-03
Co_M1NatGe_bulk	2.021e-05	1.925e-05	2.119e-05
Co_M1StringCables_bulk	2.315e-09	0.000e+00	2.156e-03
Co_M2CrossarmAndCPCables_bulk	8.797e-11	0.000e+00	5.408e-04
Co_M2CryostatCopperFar_bulk	5.252e-04	2.139e-04	6.452e-04
Co_M2CryostatCopperNearWeldedParts_bulk	1.817e-11	0.000e+00	1.244e-05
Co_M2EnrGe_bulk	6.130e-15	0.000e+00	1.535e-07
Co_M2LMFEs_bulk	2.185e-11	0.000e+00	1.537e-03

Continued on next page

parameter	Spec. Act. (μBq/kg)	Lower limit	Upper limit
Co_M2NatGe_bulk	1.752e-05	1.666e-05	1.839e-05
Co_M2StringCables_bulk	1.256e-11	0.000e+00	6.697e-04
Co_RadShieldCuOuter_bulk	3.953e-05	3.690e-05	4.221e-05
K_M1Connectors_bulk	1.759e-01	1.306e-01	2.144e-01
K_M1CPInterfaceCavityBottomSurface_bulk	3.036e-11	0.000e+00	2.064e-04
K_M1CrossarmAndCPCables_bulk	7.347e-11	0.000e+00	1.150e-02
K_M1CryostatCopperFar_bulk	3.387e-09	0.000e+00	6.898e-04
K_M1CryostatCopperNear_bulk	4.651e-10	0.000e+00	2.451e-05
K_M1DUPTFE_bulk	2.536e-12	0.000e+00	1.573e-04
K_M1DUStringCopper_bulk	1.559e-11	0.000e+00	1.920e-05
K_M1LMFEs_bulk	7.520e-09	0.000e+00	2.723e-02
K_M1Seals_DS0_bulk	2.390e-01	1.567e-01	3.394e-01
K_M1Seals_DS12_DS345bc6abc_DS7_bulk	1.502e-02	0.000e+00	2.923e-01
K_M1StringCables_bulk	3.795e-02	1.154e-02	5.026e-02
K_M1ThermosyphonAndShieldVespel_bulk	9.700e-10	0.000e+00	2.873e-02
K_M2Connectors_bulk	6.733e-02	9.416e-03	1.091e-01
K_M2CPInterfaceCavityBottomSurface_bulk	7.066e-04	0.000e+00	3.143e-03
K_M2CrossarmAndCPCables_bulk	6.438e-10	0.000e+00	1.541e-02
K_M2CryostatCopperFar_bulk	3.290e-13	0.000e+00	5.933e-04
K_M2CryostatCopperNear_bulk	1.695e-09	0.000e+00	3.118e-05
K_M2DUPTFE_bulk	5.099e-12	0.000e+00	2.108e-04
K_M2DUStringCopper_bulk	1.256e-11	0.000e+00	1.928e-05
K_M2LMFEs_bulk	2.413e-09	0.000e+00	2.497e-02
K_M2Seals_bulk	1.505e-09	0.000e+00	2.669e-02

Continued on next page

parameter	Spec. Act. (μBq/kg)	Lower limit	Upper limit
K_M2StringCables_bulk	4.809e-02	1.830e-02	6.178e-02
K_M2ThermosyphonAndShieldVespel_bulk	2.764e-09	0.000e+00	8.716e-02
K_RadShieldCuInner_bulk	3.961e-13	0.000e+00	3.385e-06
K_RadShieldCuOuter_bulk	1.731e-14	0.000e+00	5.671e-06
K_RadShieldAssembly_001_RadShieldPb_bulk	6.197e-05	0.000e+00	1.320e-04
Pb_M1DUPTFE_Enr_surf	4.699e-02	4.399e-02	5.009e-02
Pb_M1DUPTFE_Nat_surf	9.035e-02	8.147e-02	9.917e-02
Pb_M1DUPTFE_Hot_surf	9.383e-02	6.608e-02	1.217e-01
Pb_M2DUPTFE_Enr_surf	6.156e-02	5.762e-02	6.557e-02
Pb_M2DUPTFE_Nat_surf	8.183e-02	7.405e-02	8.964e-02
PbBrem_RadShieldPb_bulk	4.984e-01	4.886e-01	5.081e-01
Rn_N2_bulk	2.417e-03	0.000e+00	5.027e-03
Rn_N2_surf	3.855e-11	0.000e+00	4.288e-05
Th_M1Bellows_bulk	1.806e+00	7.877e-01	2.838e+00
Th_M1Connectors_bulk	6.521e-03	2.346e-03	1.036e-02
Th_M1CPInterfaceCavityBottomSurface_bulk	1.731e-03	1.620e-03	1.840e-03
Th_M1CrossarmAndCPCables_bulk	1.350e-10	0.000e+00	5.710e-04
Th_M1CryostatCopperFar_bulk	4.705e-04	2.428e-04	7.036e-04
Th_M1CryostatCopperNear_bulk	2.615e-12	0.000e+00	1.467e-06
Th_M1DUPTFE_bulk	9.434e-12	0.000e+00	3.124e-05
Th_M1DUStringCopper_bulk	9.618e-13	0.000e+00	2.689e-06
Th_M1LMFEs_bulk	7.204e-09	0.000e+00	5.300e-03
Th_M1Seals_DS0_bulk	2.486e-02	1.504e-02	3.472e-02
Th_M1Seals_DS12_DS345bc6abc_DS7_bulk	4.190e-03	0.000e+00	2.609e-02

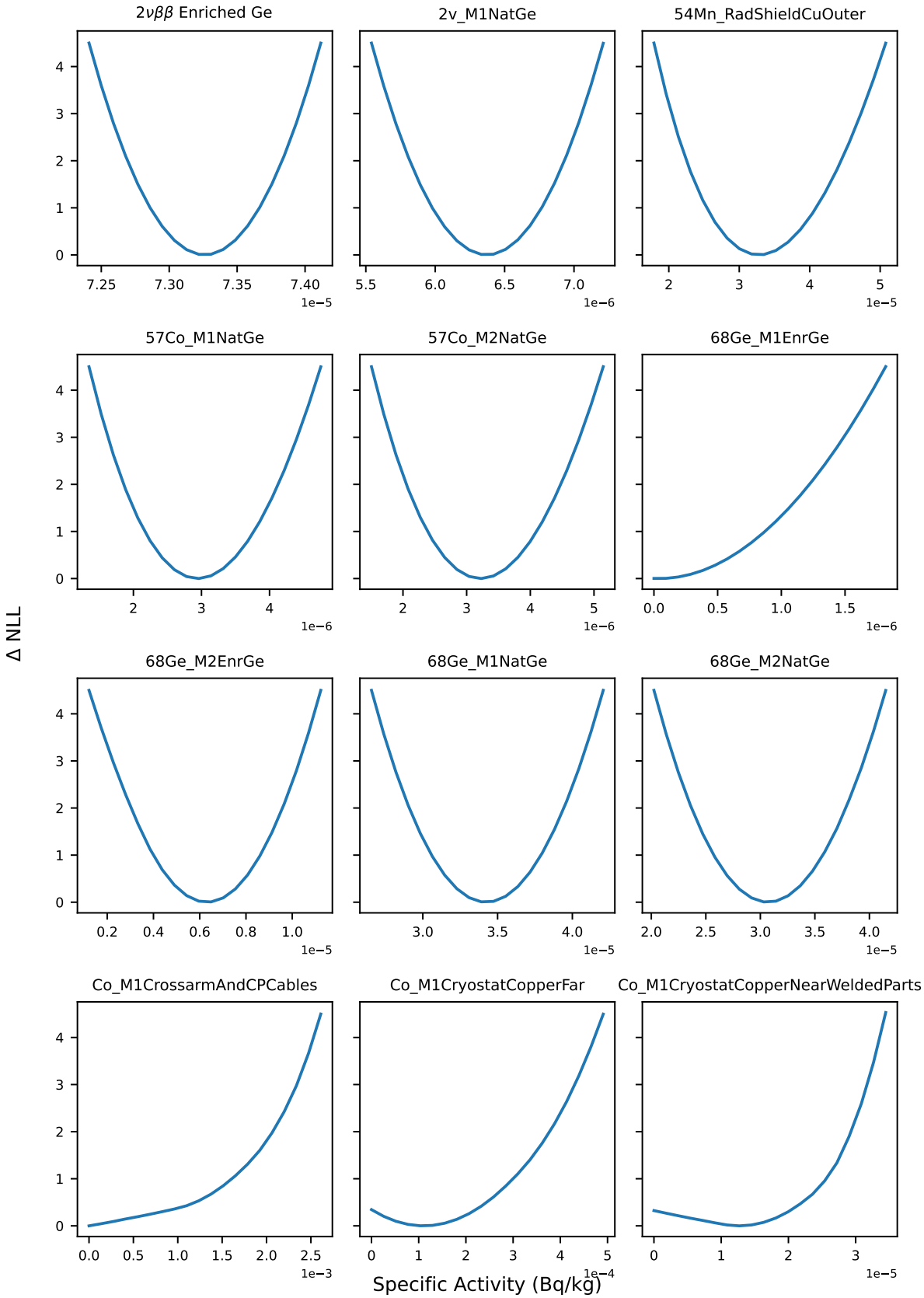
Continued on next page

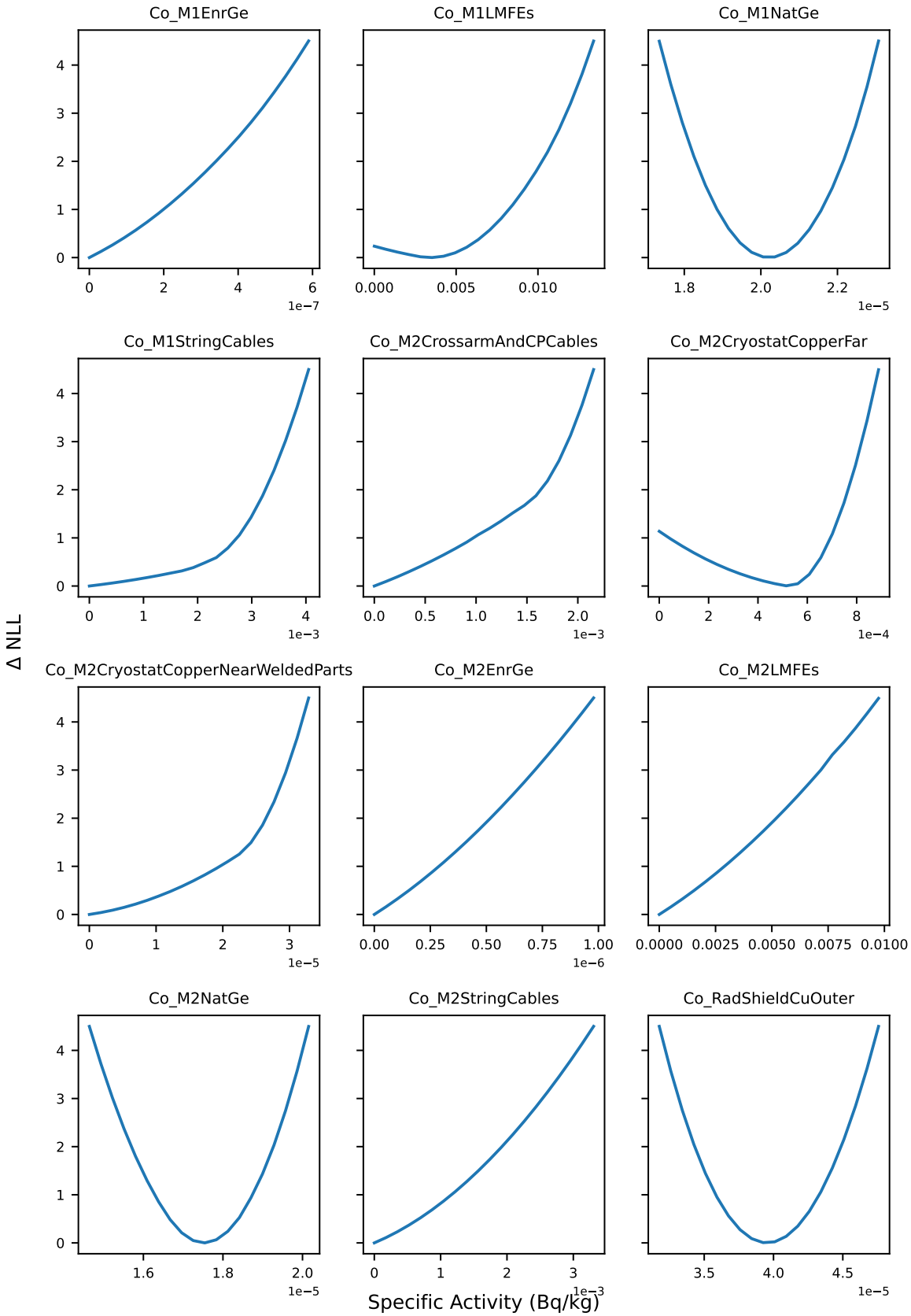
parameter	Spec. Act. (μBq/kg)	Lower limit	Upper limit
Th_M1StringCables_bulk	1.386e-03	0.000e+00	2.025e-03
Th_M1ThermosyphonAndShieldVespel_bulk	1.272e-10	0.000e+00	4.270e-03
Th_M2Bellows_bulk	3.424e-09	0.000e+00	2.895e-01
Th_M2Connectors_bulk	5.720e-10	0.000e+00	1.153e-03
Th_M2CPInterfaceCavityBottomSurface_bulk	1.150e-04	0.000e+00	2.293e-04
Th_M2CrossarmAndCPCables_bulk	1.794e-10	0.000e+00	3.018e-04
Th_M2CryostatCopperFar_bulk	3.115e-11	0.000e+00	1.301e-04
Th_M2CryostatCopperNear_bulk	2.024e-08	0.000e+00	2.634e-06
Th_M2DUPTFE_bulk	1.381e-11	0.000e+00	2.744e-05
Th_M2DUStringCopper_bulk	2.415e-12	0.000e+00	1.271e-06
Th_M2LMFEs_bulk	1.423e-10	0.000e+00	1.402e-03
Th_M2Seals_bulk	8.660e-03	5.143e-03	9.893e-03
Th_M2StringCables_bulk	1.175e-10	0.000e+00	4.678e-04
Th_M2ThermosyphonAndShieldVespel_bulk	3.504e-03	0.000e+00	3.834e-02
Th_RadShieldCuInner_bulk	2.863e-07	0.000e+00	9.914e-07
Th_RadShieldCuOuter_bulk	9.809e-07	0.000e+00	4.238e-06
Th_RadShieldAssembly_001_RadShieldPb_bulk	1.389e-13	0.000e+00	1.001e-05
U_M1Bellows_bulk	2.631e+00	2.022e+00	3.253e+00
U_M1Connectors_bulk	1.181e-03	0.000e+00	6.094e-03
U_M1CPInterfaceCavityBottomSurface_bulk	1.958e-11	0.000e+00	4.516e-05
U_M1CrossarmAndCPCables_bulk	3.458e-10	0.000e+00	4.655e-04
U_M1CryostatCopperFar_bulk	2.162e-04	6.031e-05	3.715e-04
U_M1CryostatCopperNear_bulk	6.433e-06	3.635e-06	9.125e-06
U_M1DUPTFE_bulk	2.044e-11	0.000e+00	7.997e-05

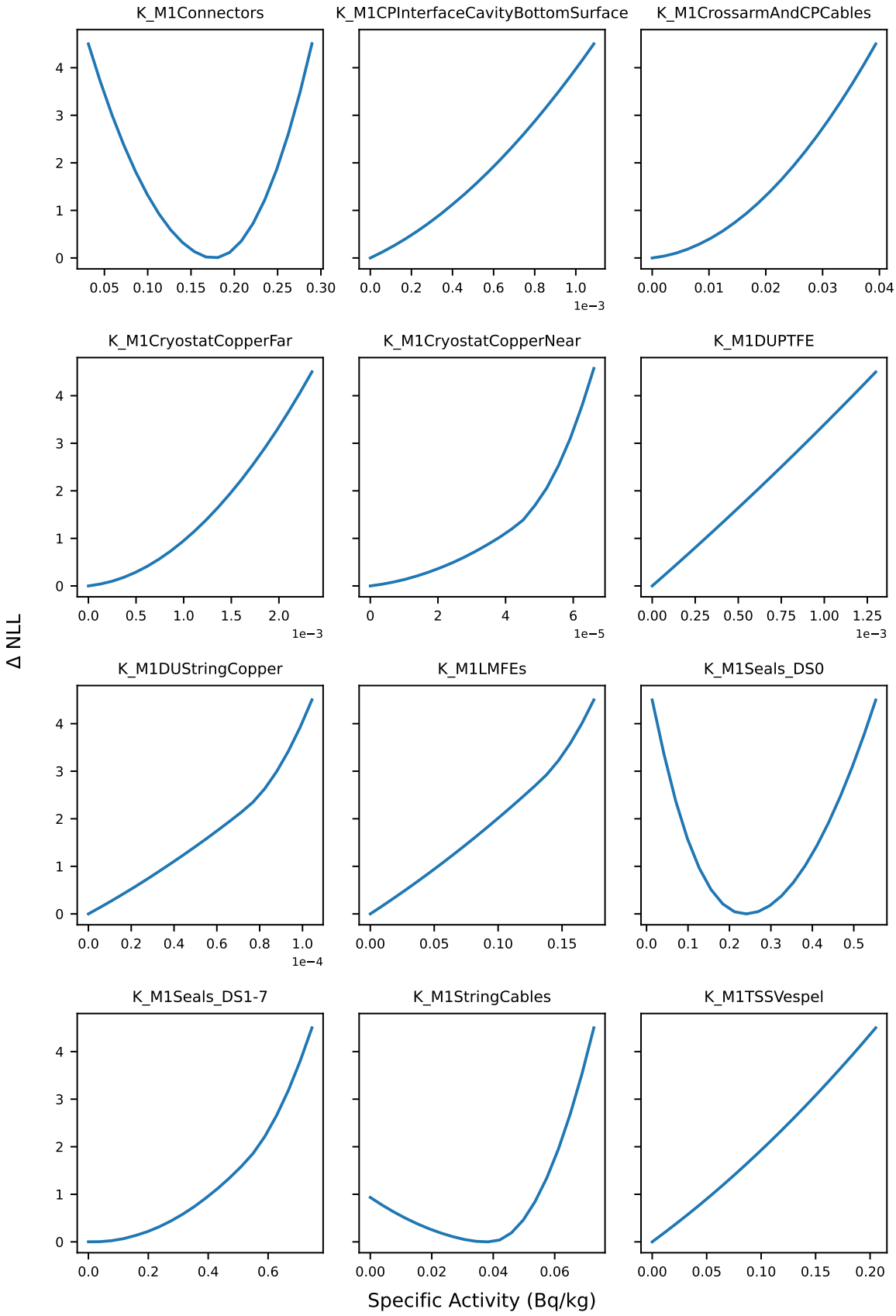
Continued on next page

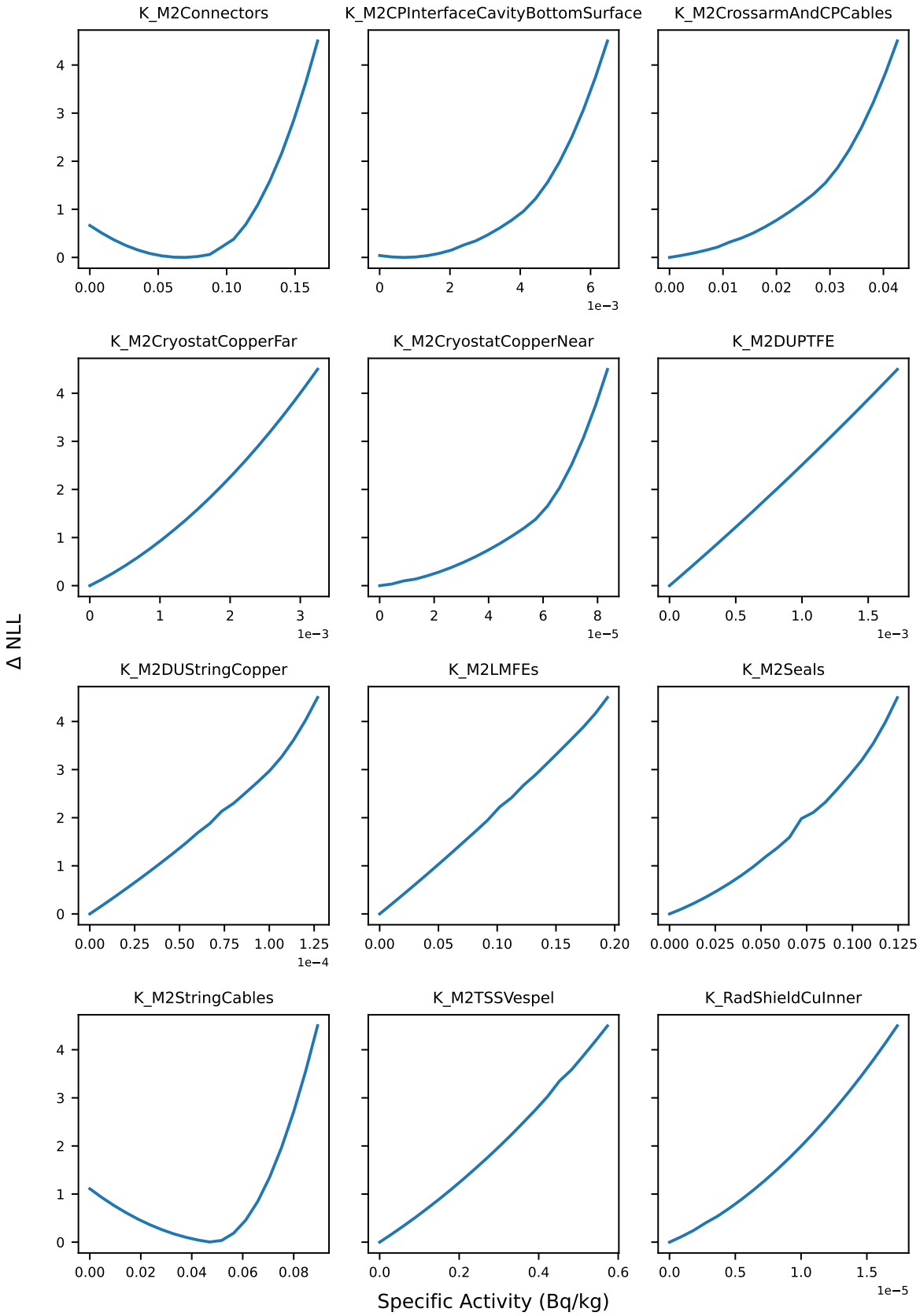
parameter	Spec. Act. (μBq/kg)	Lower limit	Upper limit
U_M1DUStringCopper_bulk	1.304e-10	0.000e+00	5.101e-06
U_M1LMFEs_bulk	8.661e-08	0.000e+00	7.922e-03
U_M1Seals_DS0_bulk	3.954e-02	2.838e-02	5.060e-02
U_M1Seals_DS12_DS345bc6abc_DS7_bulk	3.010e-08	0.000e+00	1.862e-02
U_M1StringCables_bulk	4.148e-03	1.117e-03	5.676e-03
U_M1ThermosyphonAndShieldVespel_bulk	4.034e-10	0.000e+00	4.343e-03
U_M2Bellows_bulk	6.954e-01	0.000e+00	1.711e+00
U_M2Connectors_bulk	4.464e-10	0.000e+00	1.380e-03
U_M2CPInterfaceCavityBottomSurface_bulk	1.080e-04	0.000e+00	2.840e-04
U_M2CrossarmAndCPCables_bulk	2.243e-10	0.000e+00	3.305e-04
U_M2CryostatCopperFar_bulk	3.532e-11	0.000e+00	2.219e-04
U_M2CryostatCopperNear_bulk	8.357e-06	4.318e-06	1.053e-05
U_M2DUPTFE_bulk	7.164e-10	0.000e+00	6.653e-05
U_M2DUStringCopper_bulk	2.403e-12	0.000e+00	2.440e-06
U_M2LMFEs_bulk	5.807e-09	0.000e+00	2.578e-03
U_M2Seals_bulk	1.403e-03	0.000e+00	6.657e-03
U_M2StringCables_bulk	6.187e-09	0.000e+00	1.217e-03
U_M2ThermosyphonAndShieldVespel_bulk	8.823e-09	0.000e+00	1.201e-02
U_RadShieldCuInner_bulk	2.170e-12	0.000e+00	7.229e-07
U_RadShieldCuOuter_bulk	1.399e-06	0.000e+00	6.559e-06
U_RadShieldAssembly_001_RadShieldPb_bulk	1.383e-05	0.000e+00	3.592e-05

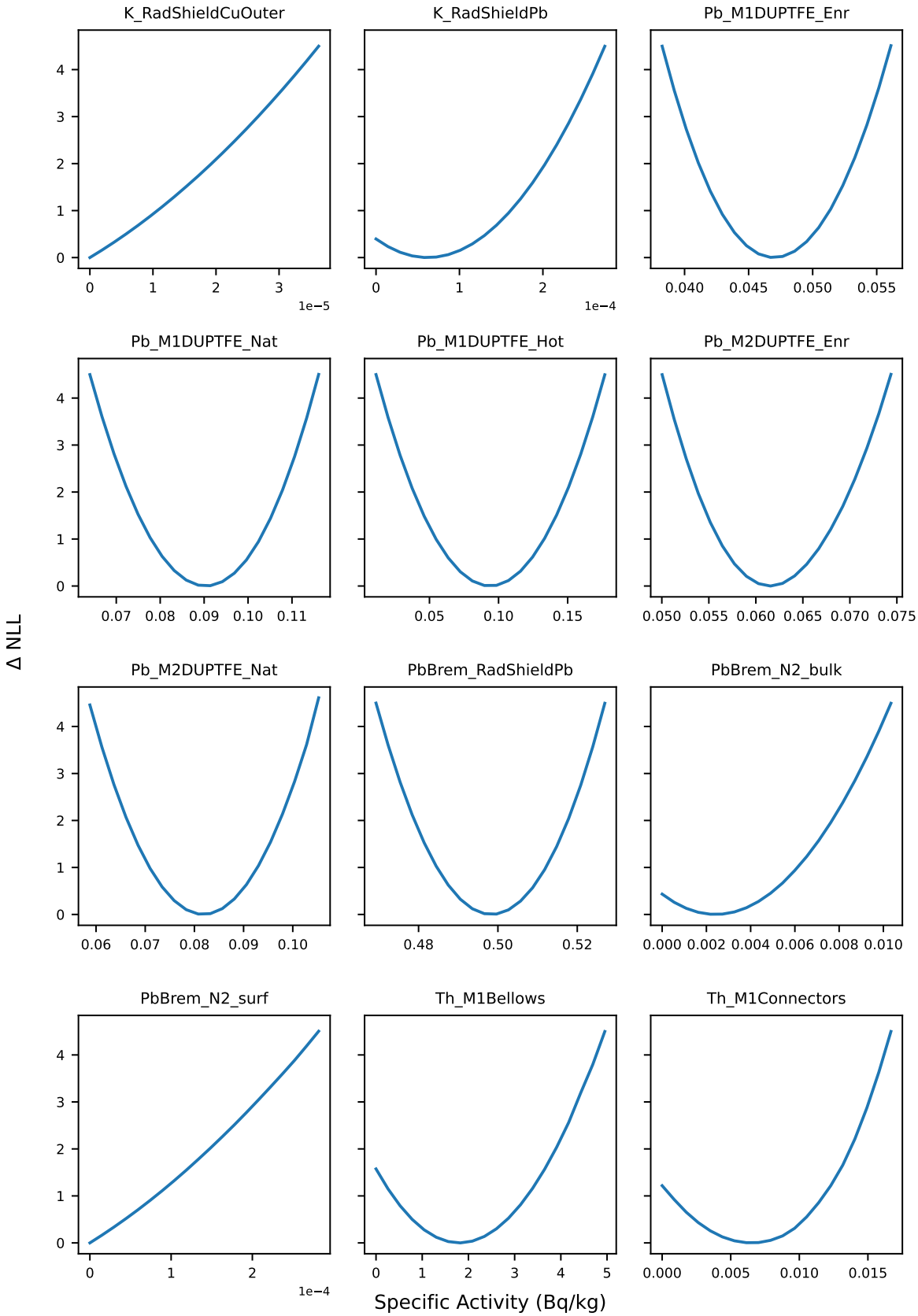
APPENDIX C: PROFILE LIKELIHOOD CURVES

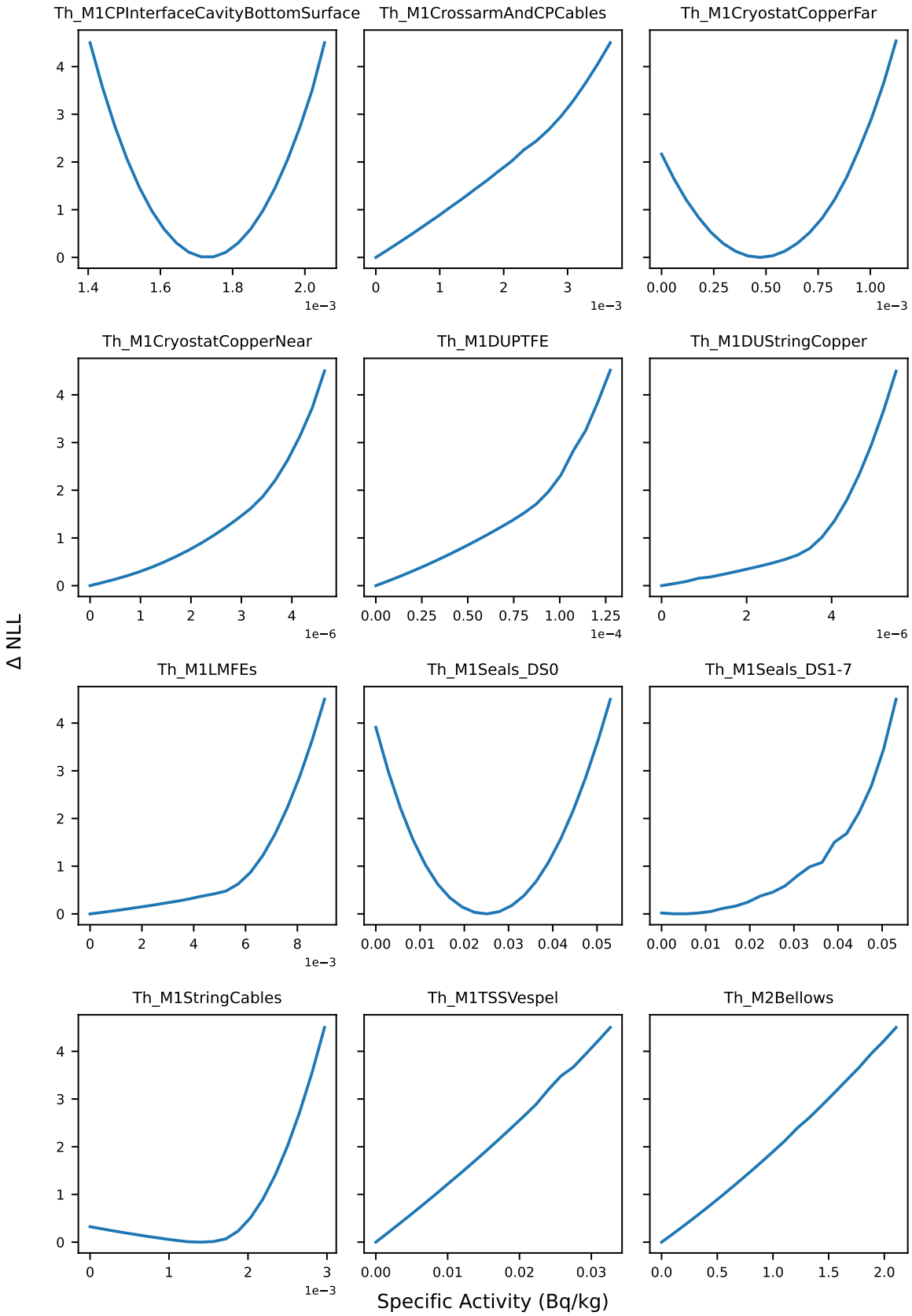

2νββ Enriched Ge
2v_M1NatGe
54Mn_RadShieldCuOuter
57Co_M1NatGe
57Co_M2NatGe
68Ge_M1EnrGe
68Ge_M2EnrGe
68Ge_M1NatGe
68Ge_M2NatGe
Co_M1CrossarmAndCPCables
Co_M1CryostatCopperFar
Co_M1CryostatCopperNearWeldedParts
Δ NLL
Specific Activity (Bq/kg)

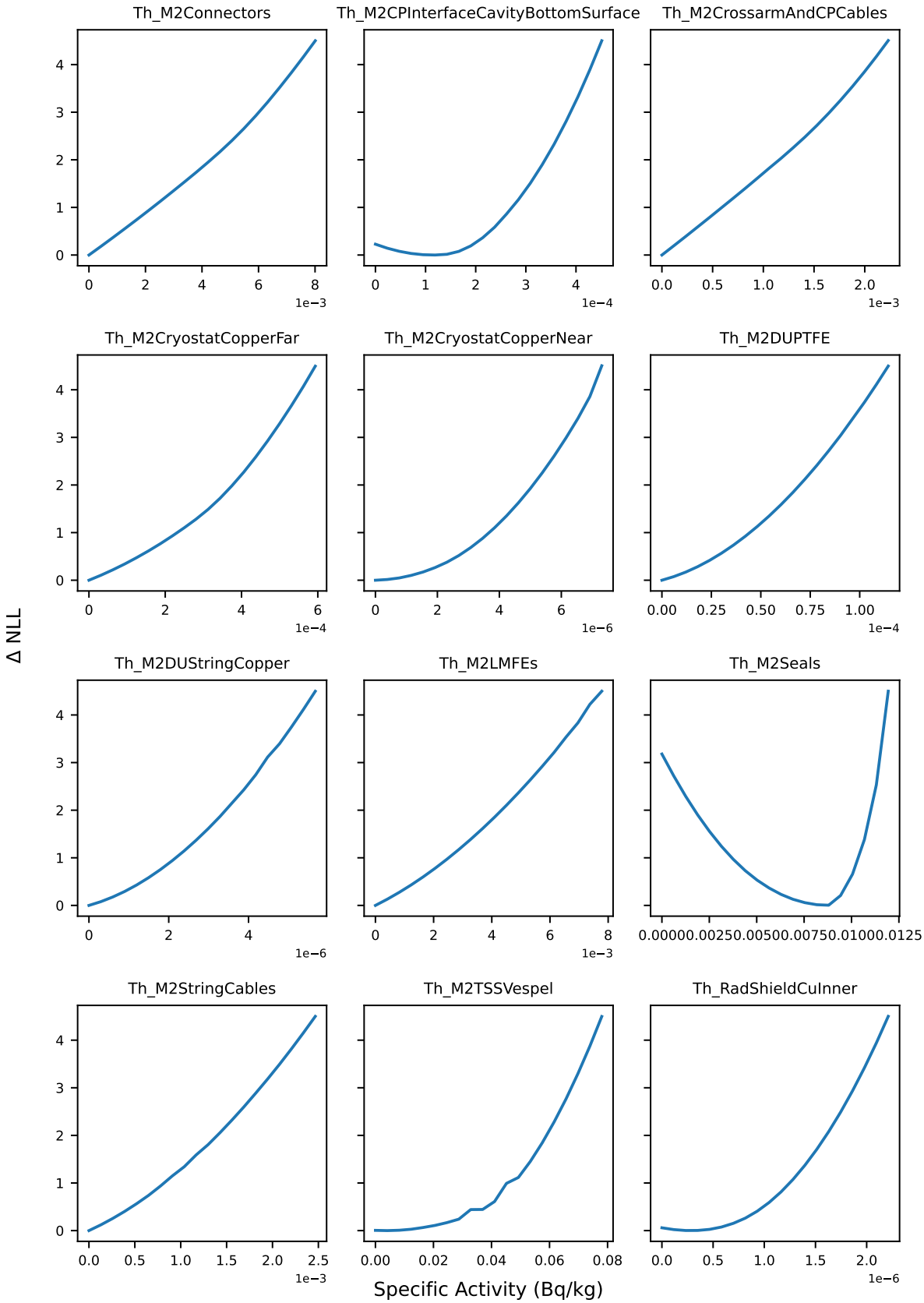

Co_M1EnrGe
Co_M1LMFEs
Co_M1NatGe
Co_M1StringCables
Co_M2CrossarmAndCPCables
Co_M2CryostatCopperFar
Co_M2CryostatCopperNearWeldedParts
Co_M2EnrGe
Co_M2LMFEs
Co_M2NatGe
Co_M2StringCables
Co_RadShieldCuOuter
Δ NLL
Specific Activity (Bq/kg)

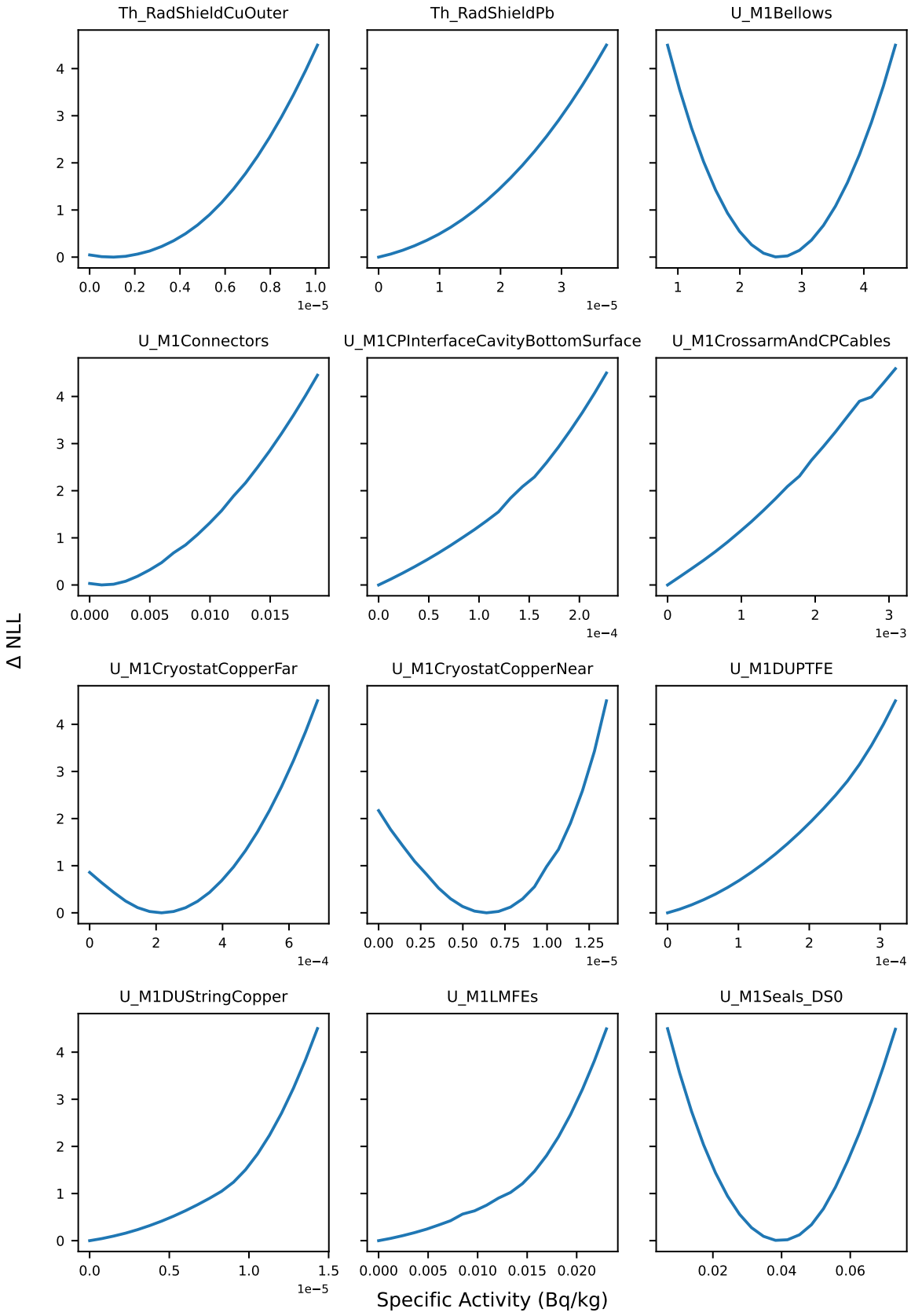

K_M1Connectors
K_M1CPInterfaceCavityBottomSurface
K_M1CrossarmAndCPCables
K_M1CryostatCopperFar
K_M1CryostatCopperNear
K_M1DUPTFE
K_M1DUStringCopper
K_M1LMFEs
K_M1Seals_DS0
K_M1Seals_DS1-7
K_M1StringCables
K_M1TSSVespel
Δ NLL
Specific Activity (Bq/kg)
1e−3
1e−5
1e−4

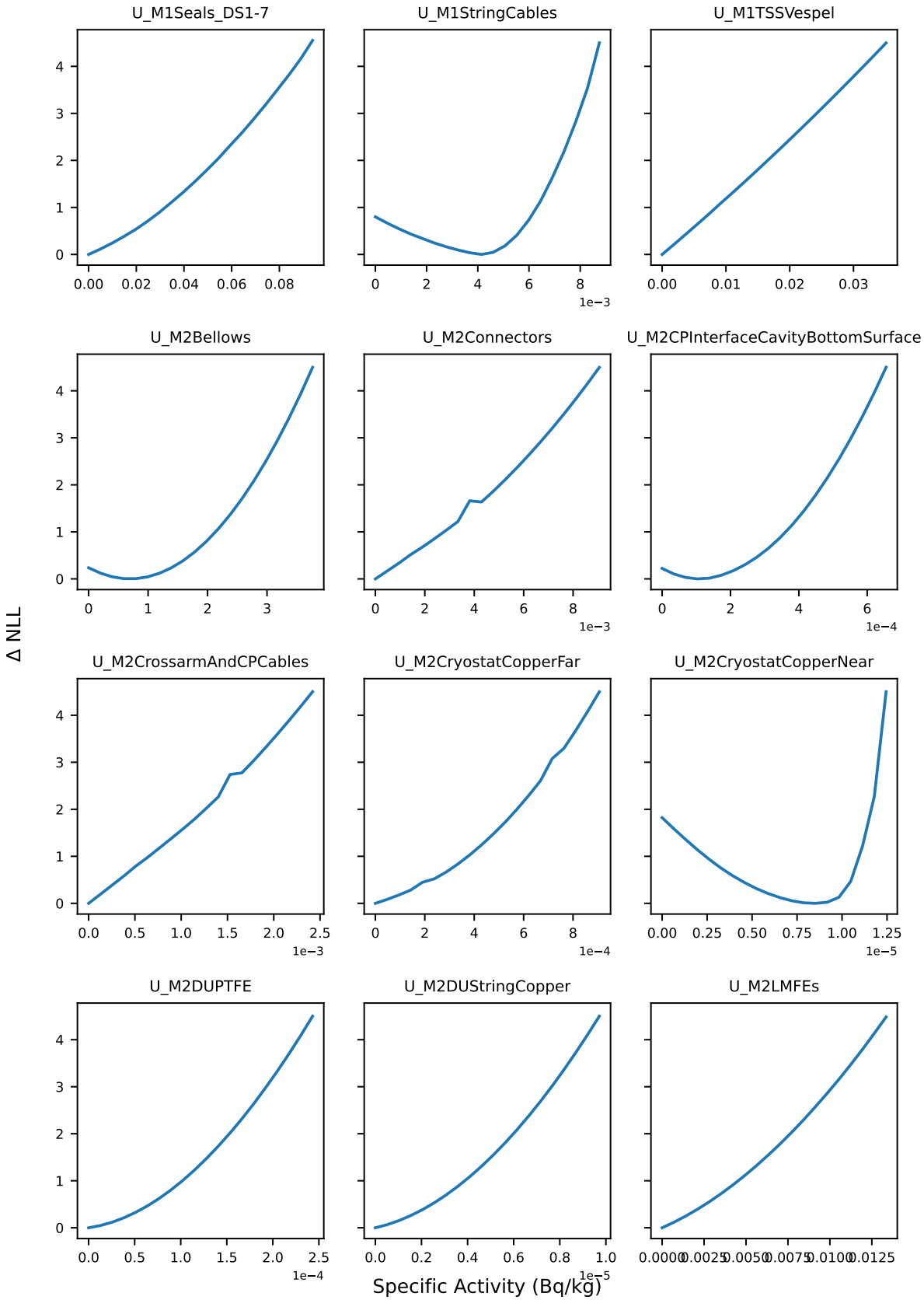

K_M2Connectors
K_M2CPInterfaceCavityBottomSurface
K_M2CrossarmAndCPCables
K_M2CryostatCopperFar
K_M2CryostatCopperNear
K_M2DUPTFE
K_M2DUStringCopper
K_M2LMFEs
K_M2Seals
K_M2StringCables
K_M2TSSVespel
K_RadShieldCuInner
Δ NLL
Specific Activity (Bq/kg)
1e−3
1e−5
1e−4

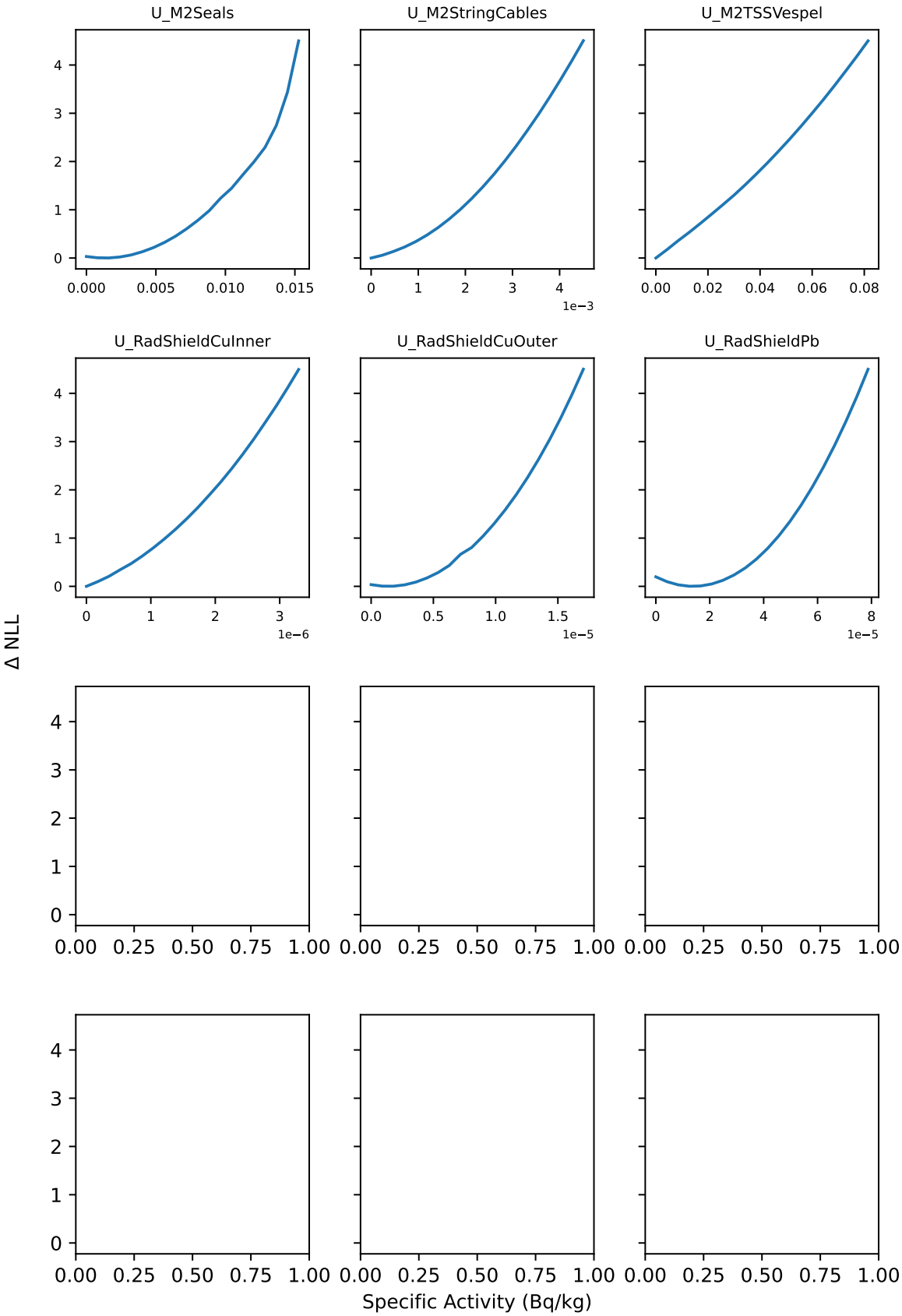

K_RadShieldCuOuter
K_RadShieldPb
Pb_M1DUPTFE_Enr
Pb_M1DUPTFE_Nat
Pb_M1DUPTFE_Hot
Pb_M2DUPTFE_Enr
Pb_M2DUPTFE_Nat
PbBrem_RadShieldPb
PbBrem_N2_bulk
PbBrem_N2_surf
Th_M1Bellows
Th_M1Connectors
Δ NLL
Specific Activity (Bq/kg)
1e−5
1e−4


Th_M1CPInterfaceCavityBottomSurface
Th_M1CrossarmAndCPCables
Th_M1CryostatCopperFar
Th_M1CryostatCopperNear
Th_M1DUPTFE
Th_M1DUStringCopper
Th_M1LMFEs
Th_M1Seals_DS0
Th_M1Seals_DS1-7
Th_M1StringCables
Th_M1TSSVespel
Th_M2Bellows
Δ NLL
Specific Activity (Bq/kg)

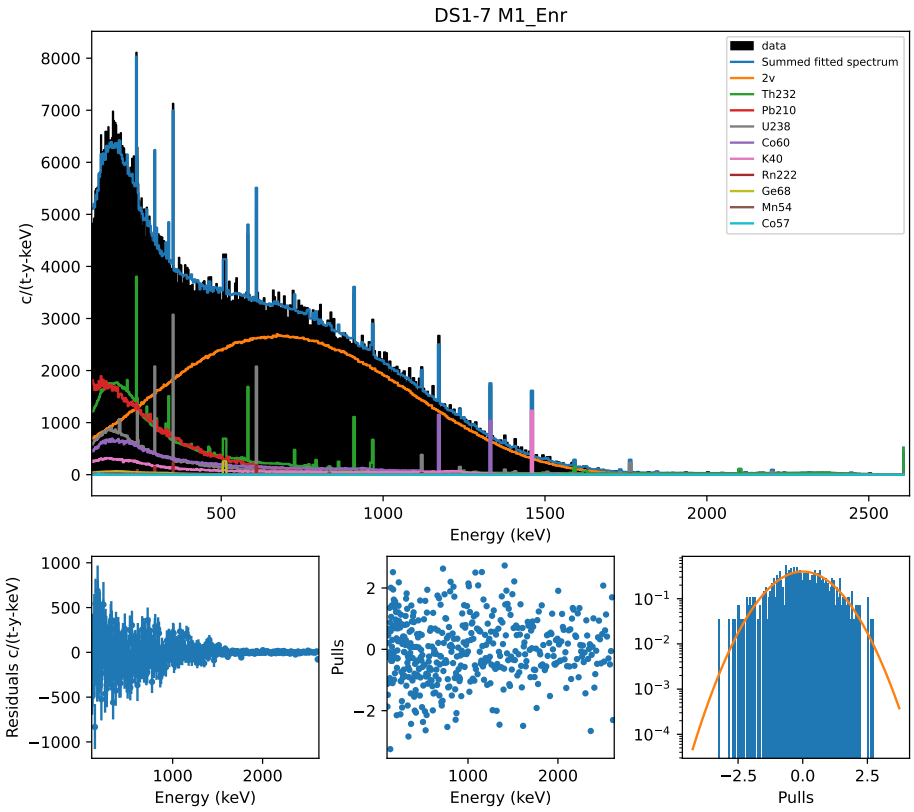

Th_M2Connectors
Th_M2CPInterfaceCavityBottomSurface
Th_M2CrossarmAndCPCables
Th_M2CryostatCopperFar
Th_M2CryostatCopperNear
Th_M2DUPTFE
Th_M2DUStringCopper
Th_M2LMFEs
Th_M2Seals
Th_M2StringCables
Th_M2TSSVespel
Th_RadShieldCuInner
Δ NLL
Specific Activity (Bq/kg)

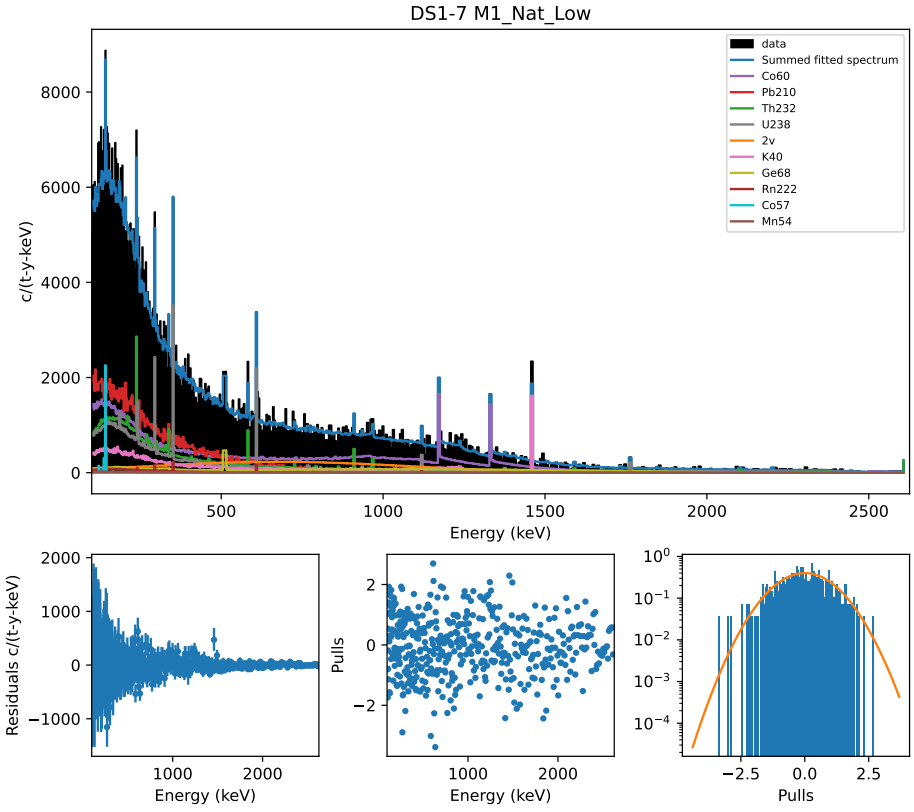

Th_RadShieldCuOuter
Th_RadShieldPb
U_M1Bellows
U_M1Connectors
U_M1CPInterfaceCavityBottomSurface
U_M1CrossarmAndCPCables
U_M1CryostatCopperFar
U_M1CryostatCopperNear
U_M1DUPTFE
U_M1DUStringCopper
U_M1LMFEs
U_M1Seals_DS0
Δ NLL
Specific Activity (Bq/kg)

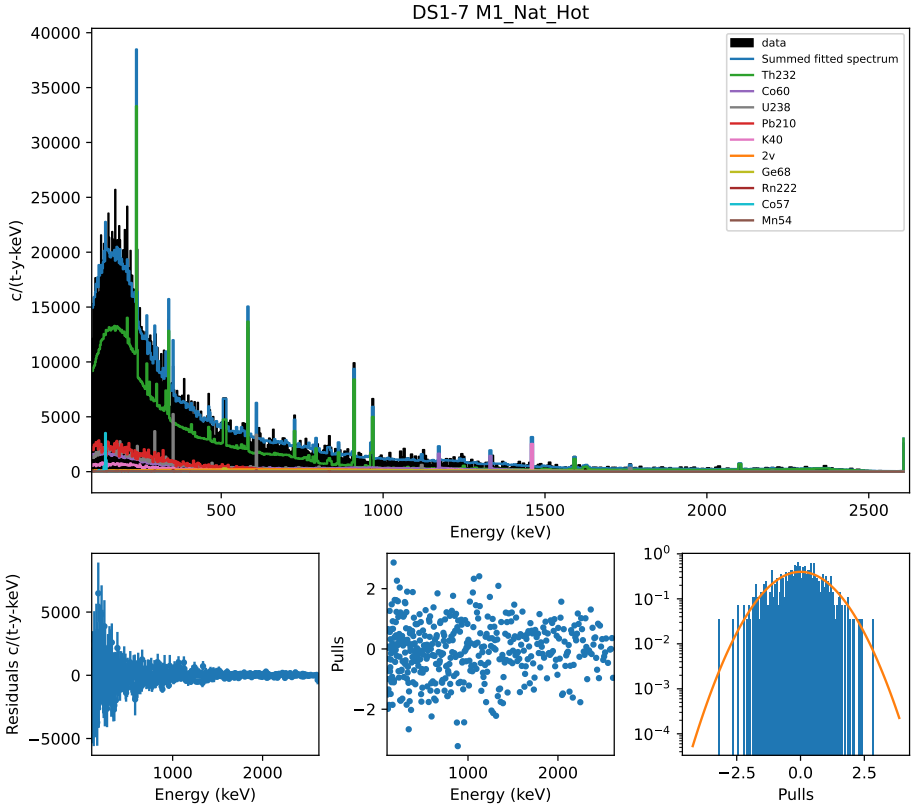

U_M1Seals_DS1-7
U_M1StringCables
U_M1TSSVespel
U_M2Bellows
U_M2Connectors
U_M2CPInterfaceCavityBottomSurface
U_M2CrossarmAndCPCables
U_M2CryostatCopperFar
U_M2CryostatCopperNear
U_M2DUPTFE
U_M2DUStringCopper
U_M2LMFEs
Δ NLL
Specific Activity (Bq/kg)
1e−3
1e−4
1e−5

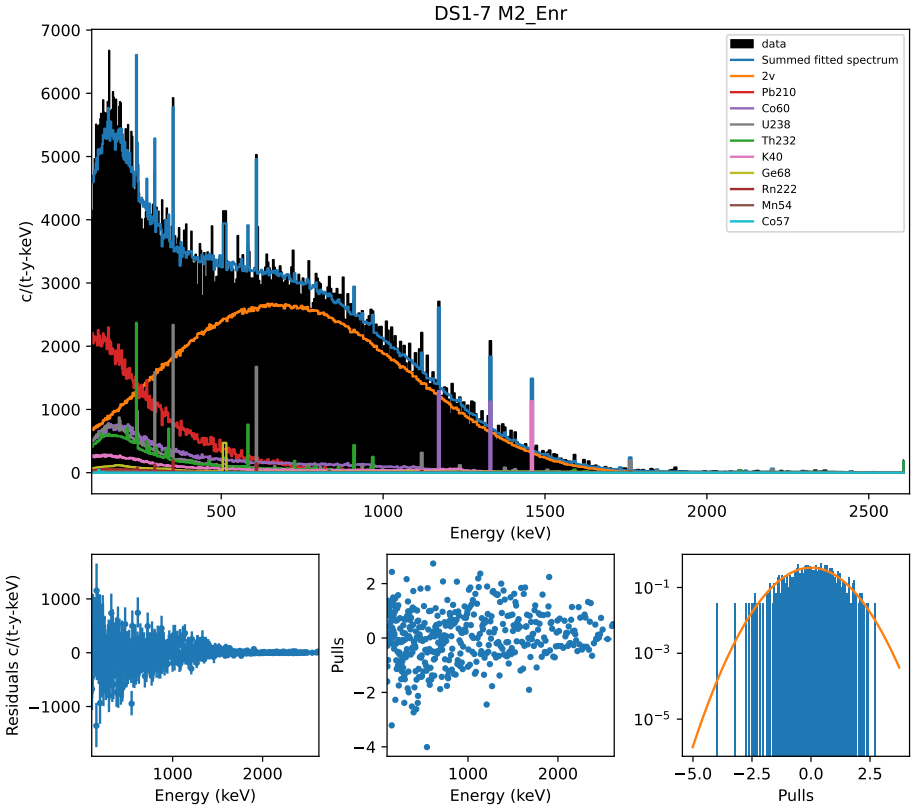

U_M2Seals
U_M2StringCables
U_M2TSSVespel
U_RadShieldCuInner
U_RadShieldCuOuter
U_RadShieldPb
Δ NLL
Specific Activity (Bq/kg)
1e−3
1e−6
1e−5
1e−5

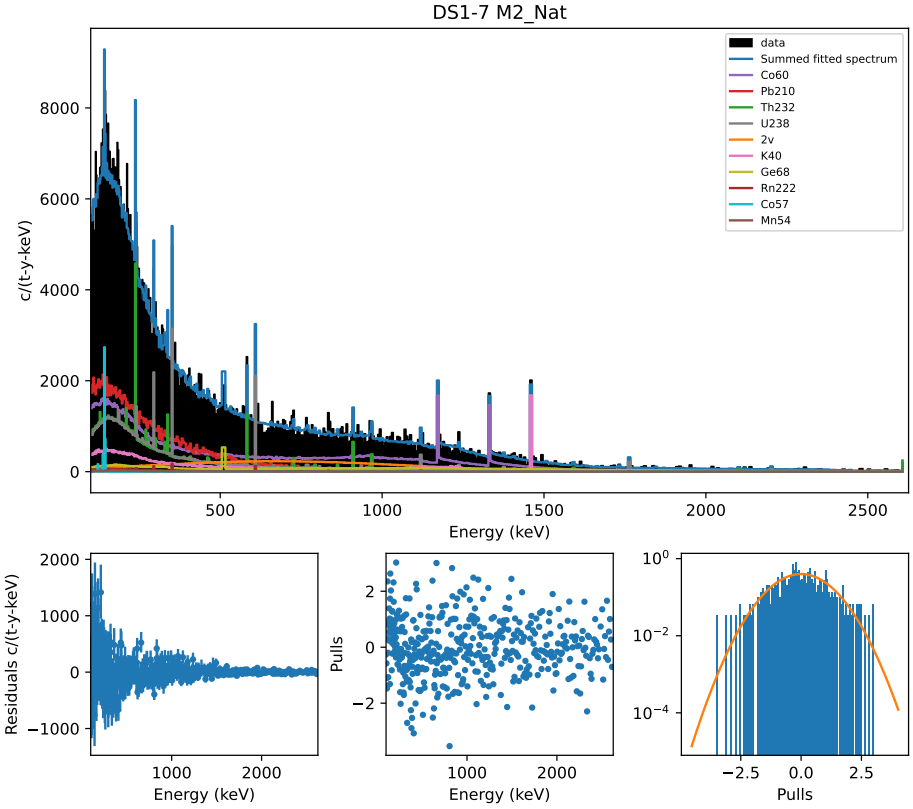
APPENDIX D: FIT RESULTS PLOTS

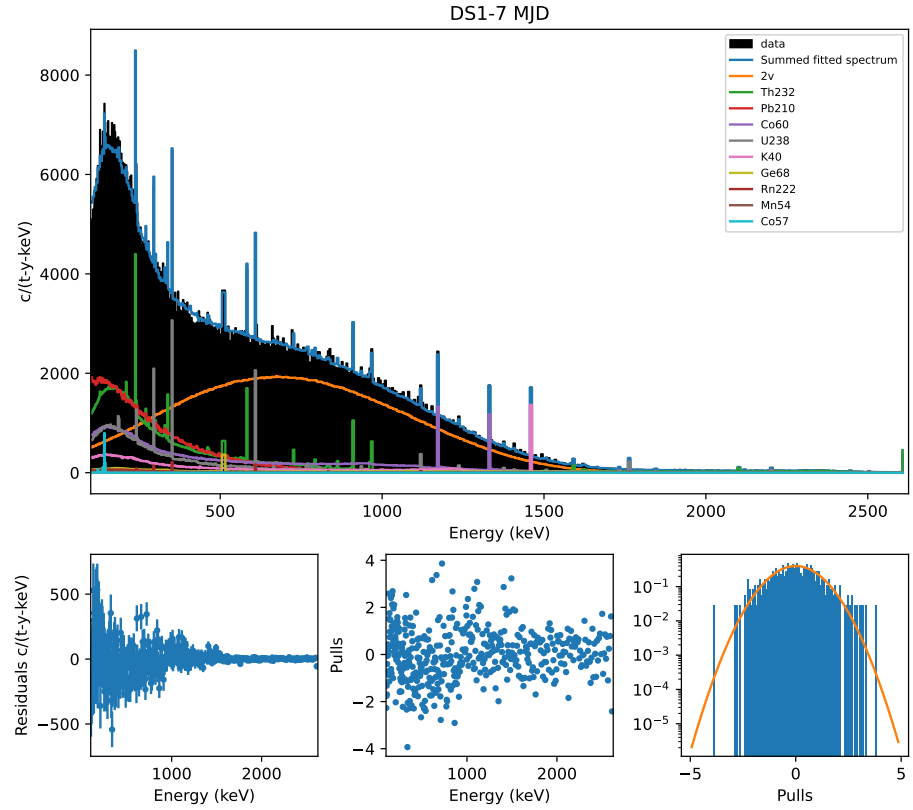

DS1-7 M1_Enr
data
Summed fitted spectrum
2v
Th232
Pb210
U238
Co60
K40
Rn222
Ge68
Mn54
Co57
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls
Energy (keV)
Energy (keV)
Pulls

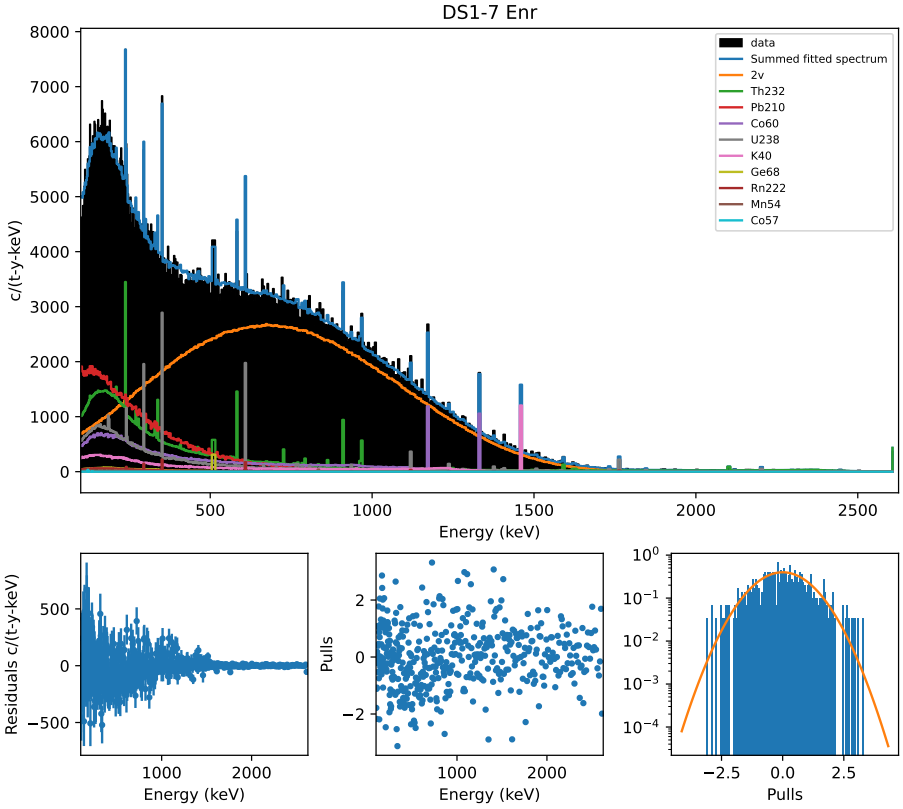

DS1-7 M1_Nat_Low
data
Summed fitted spectrum
Co60
Pb210
Th232
U238
2v
K40
Ge68
Rn222
Co57
Mn54
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls
Energy (keV)
Energy (keV)
Pulls

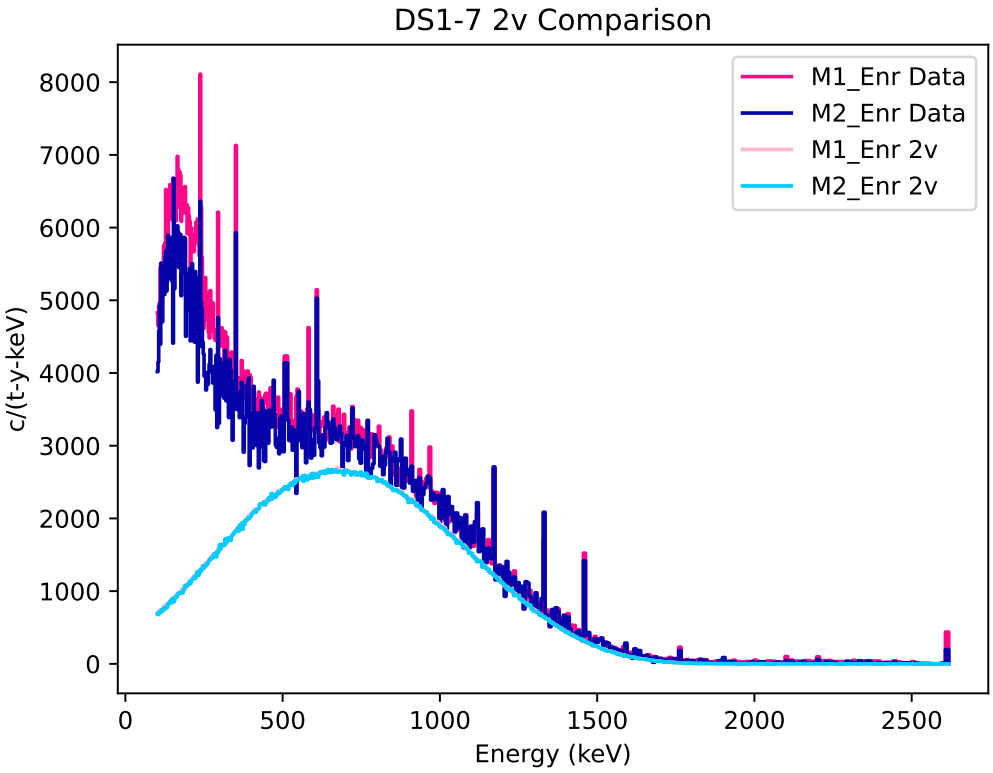

DS1-7 M1_Nat_Hot
data
Summed fitted spectrum
Th232
Co60
U238
Pb210
K40
2v
Ge68
Rn222
Co57
Mn54
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls

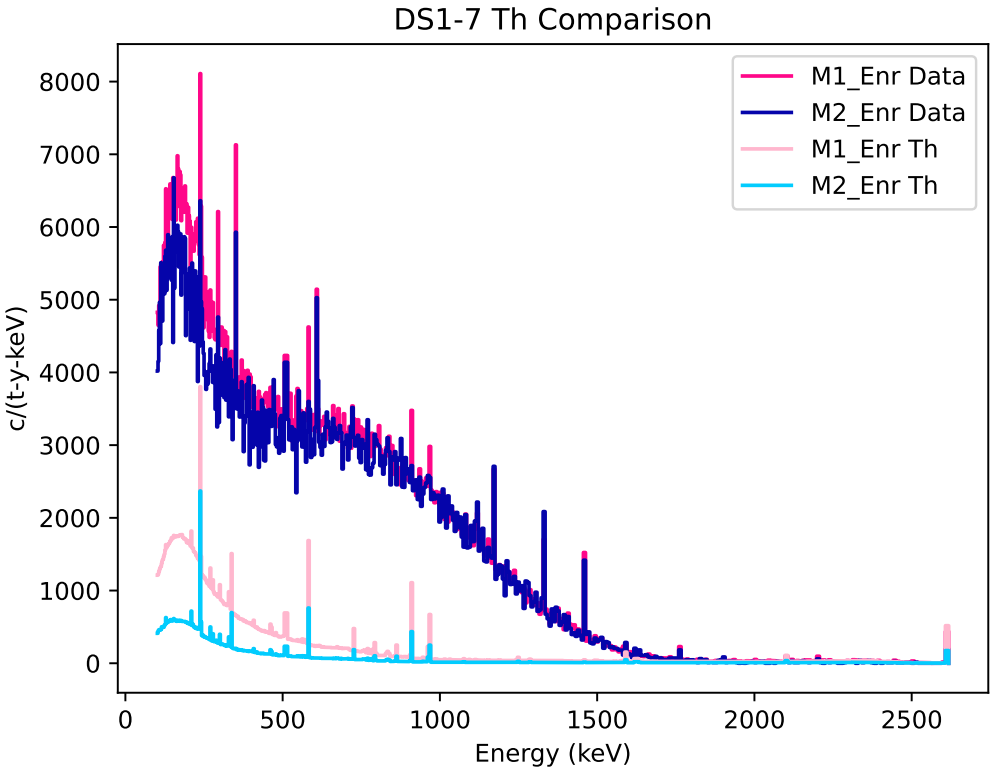

DS1-7 M2_Enr
data
Summed fitted spectrum
2v
Pb210
Co60
U238
Th232
K40
Ge68
Rn222
Mn54
Co57
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls

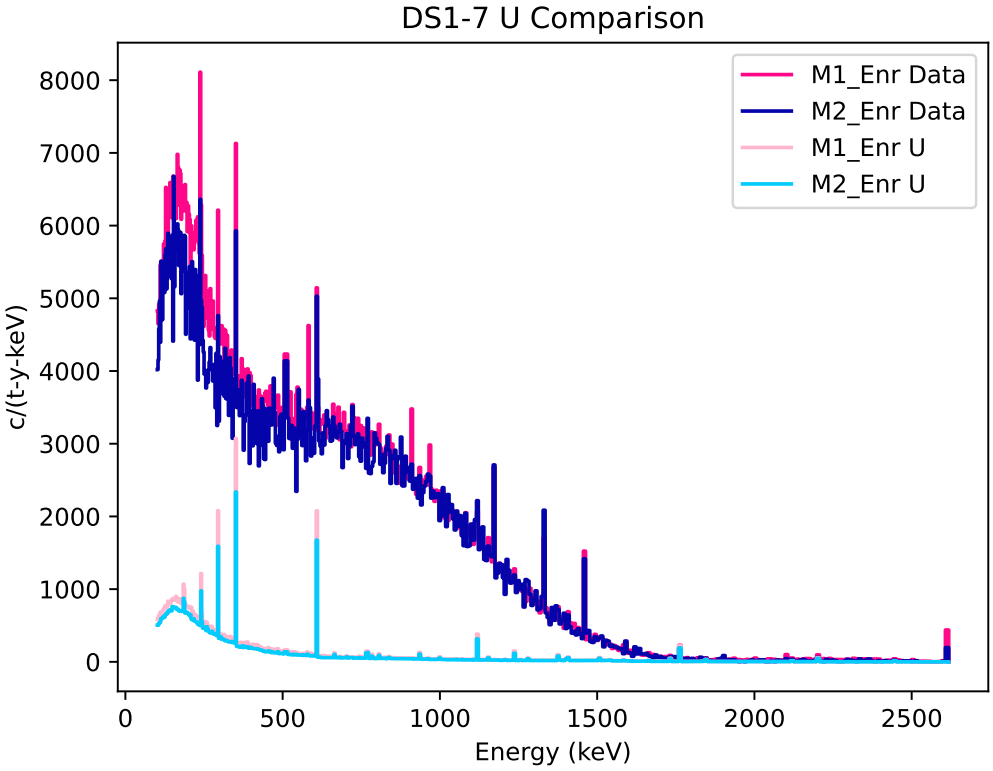

DS1-7 M2_Nat
data
Summed fitted spectrum
Co60
Pb210
Th232
U238
2v
K40
Ge68
Rn222
Co57
Mn54
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls
Energy (keV)
Energy (keV)
Pulls

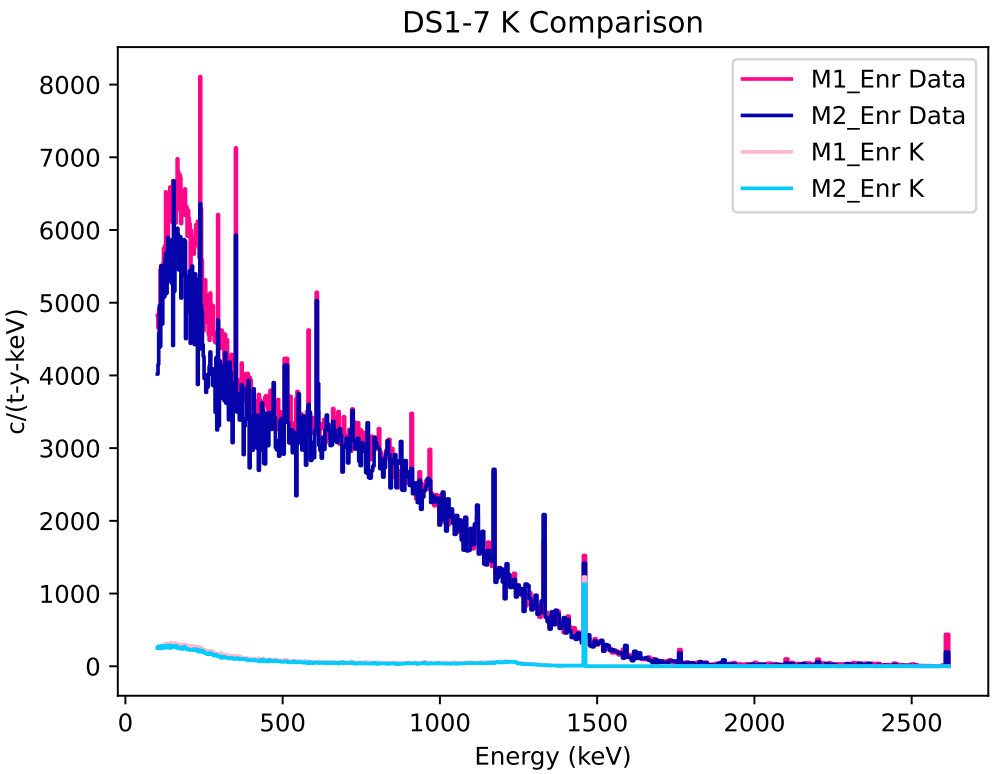

DS1-7 MJD
data
Summed fitted spectrum
2v
Th232
Pb210
Co60
U238
K40
Ge68
Rn222
Mn54
Co57
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Pulls
Energy (keV)
Energy (keV)
Pulls

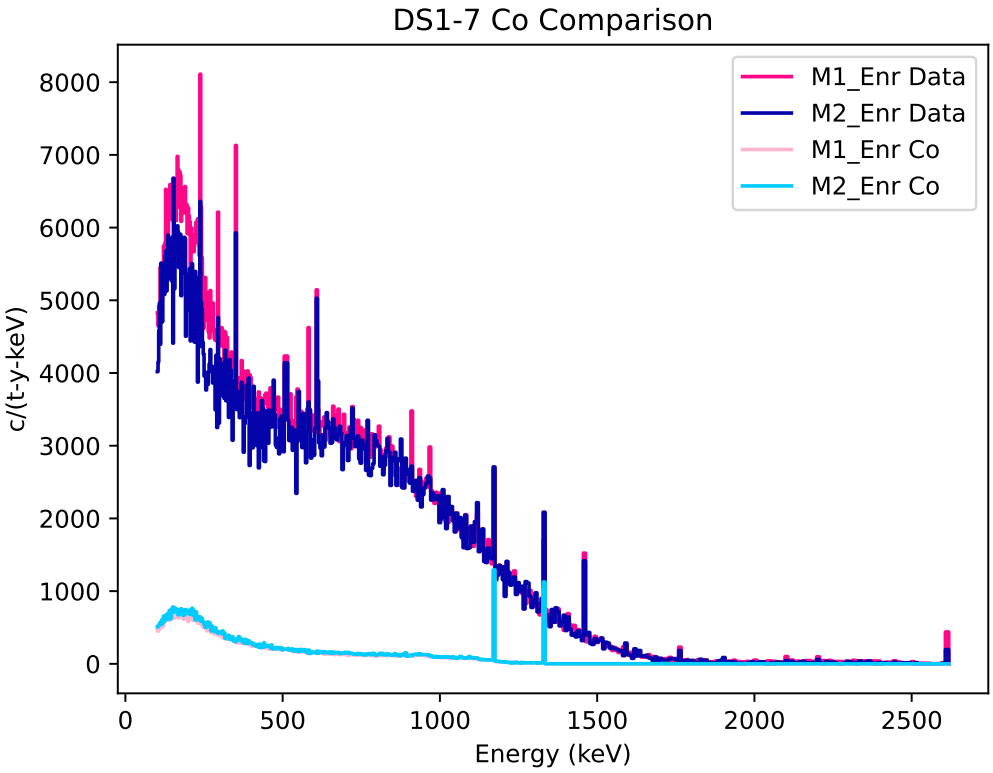

DS1-7 Enr
data
Summed fitted spectrum
2v
Th232
Pb210
Co60
U238
K40
Ge68
Rn222
Mn54
Co57
c/(t-y-keV)
Energy (keV)
Residuals c/(t-y-keV)
Energy (keV)
Pulls
Energy (keV)
Pulls

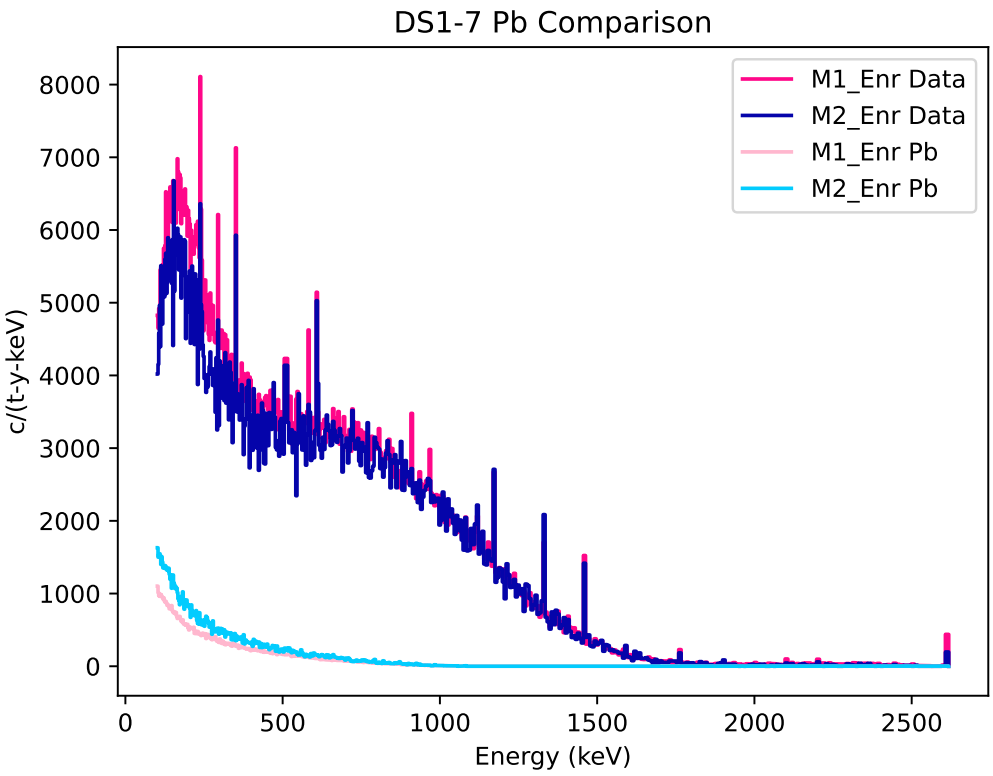

DS1-7 2v Comparison
M1_Enr Data
M2_Enr Data
M1_Enr 2v
M2_Enr 2v
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

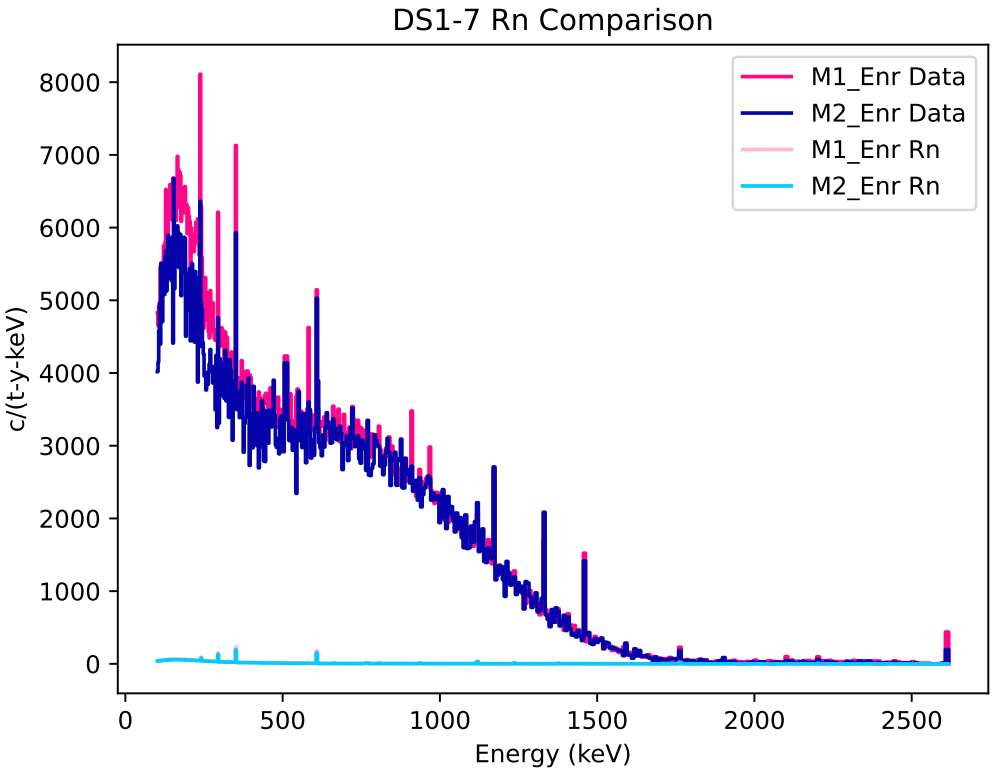

DS1-7 Th Comparison
M1_Enr Data
M2_Enr Data
M1_Enr Th
M2_Enr Th
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

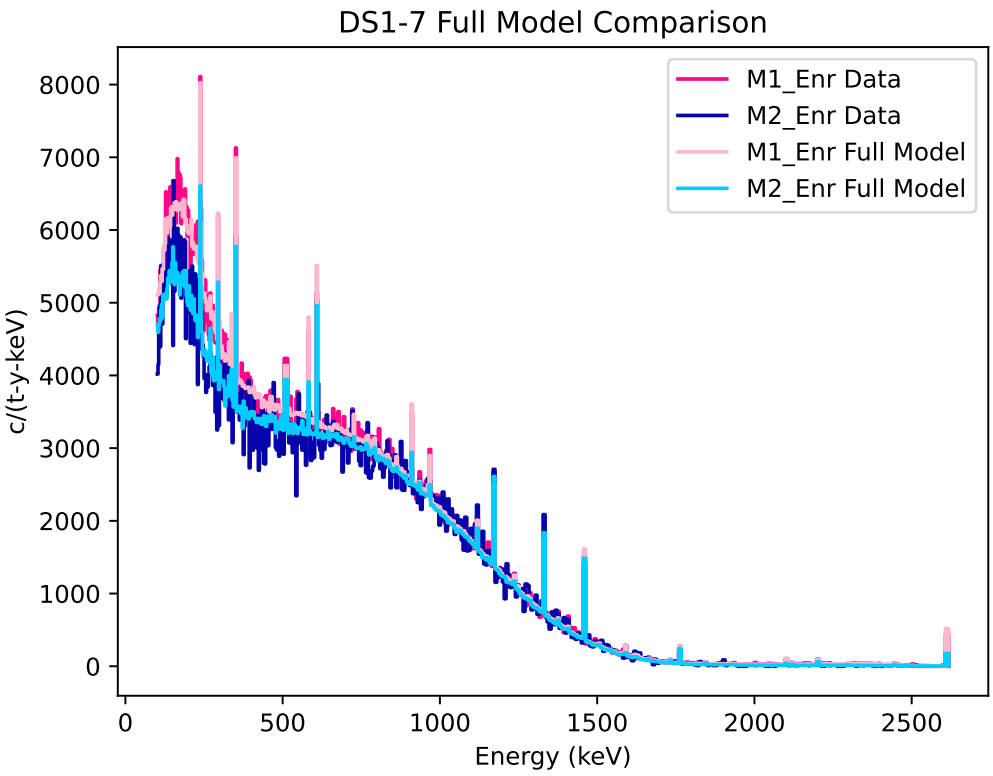

DS1-7 U Comparison
M1_Enr Data
M2_Enr Data
M1_Enr U
M2_Enr U
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

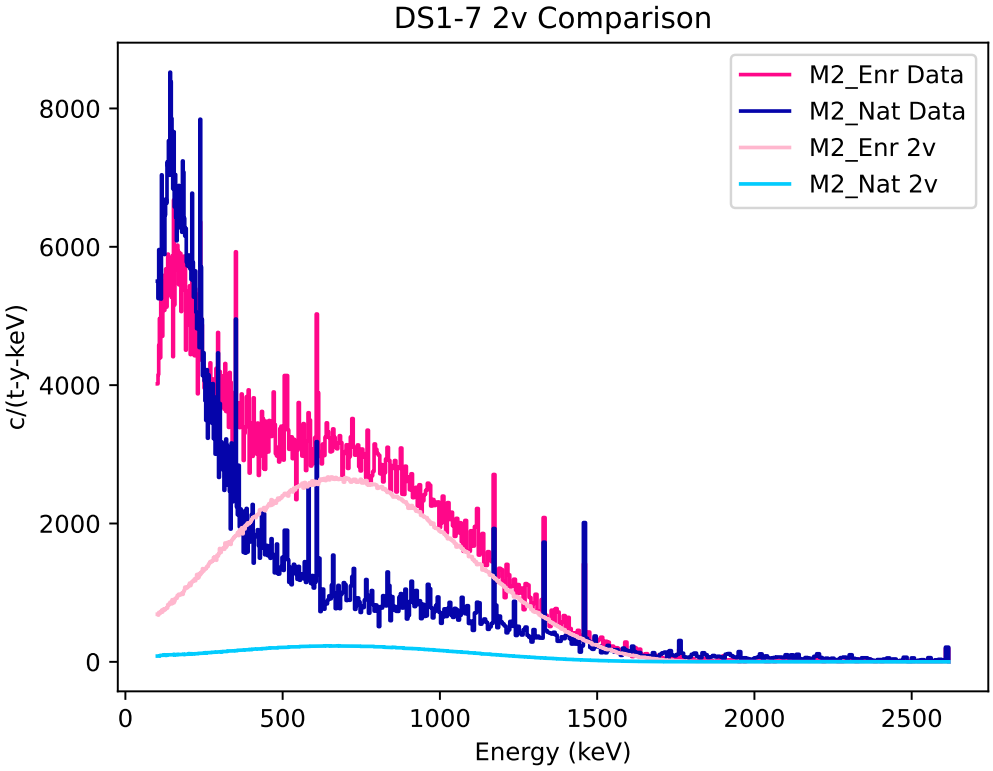

DS1-7 K Comparison
M1_Enr Data
M2_Enr Data
M1_Enr K
M2_Enr K
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

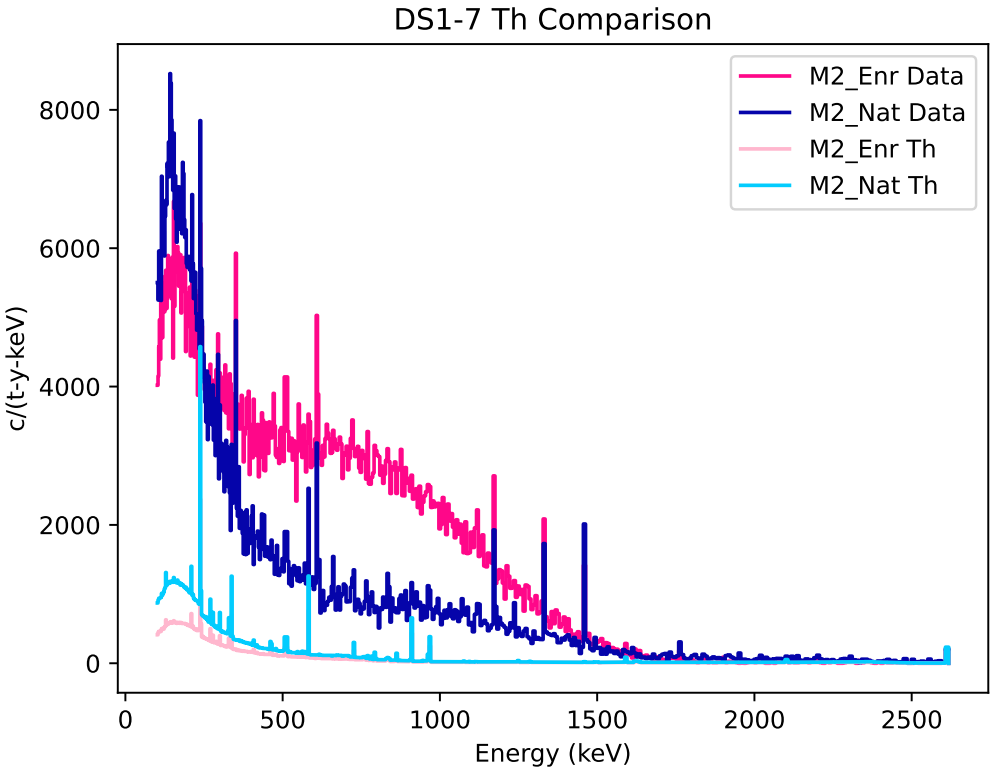

DS1-7 Co Comparison
M1_Enr Data
M2_Enr Data
M1_Enr Co
M2_Enr Co
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

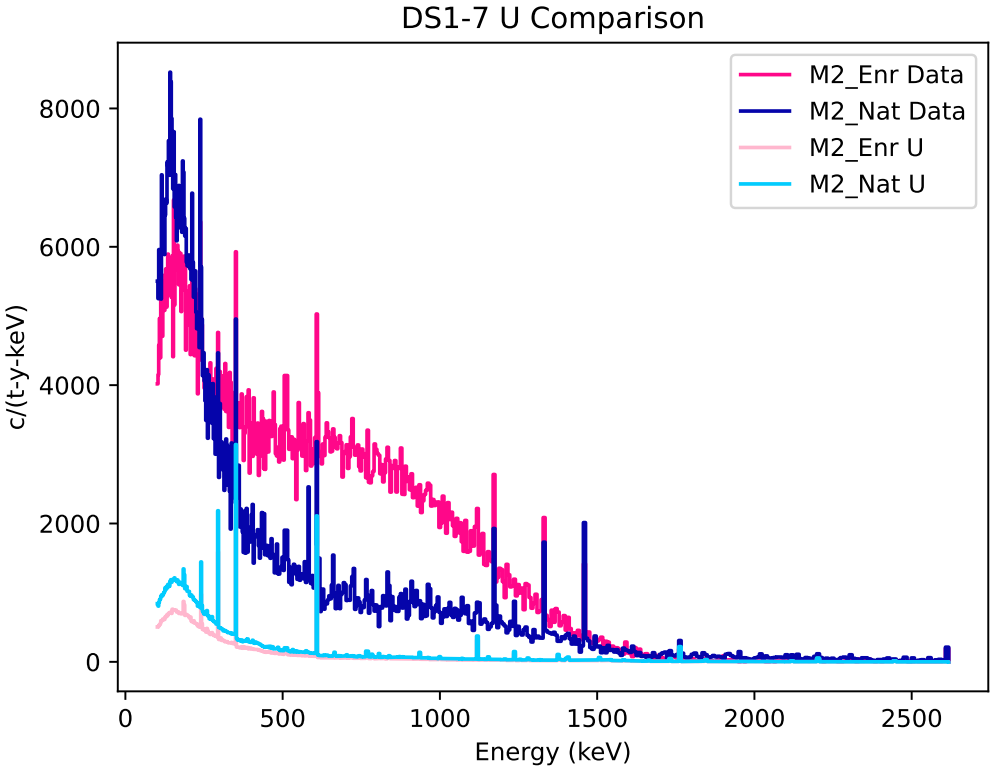

DS1-7 Pb Comparison
M1_Enr Data
M2_Enr Data
M1_Enr Pb
M2_Enr Pb
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

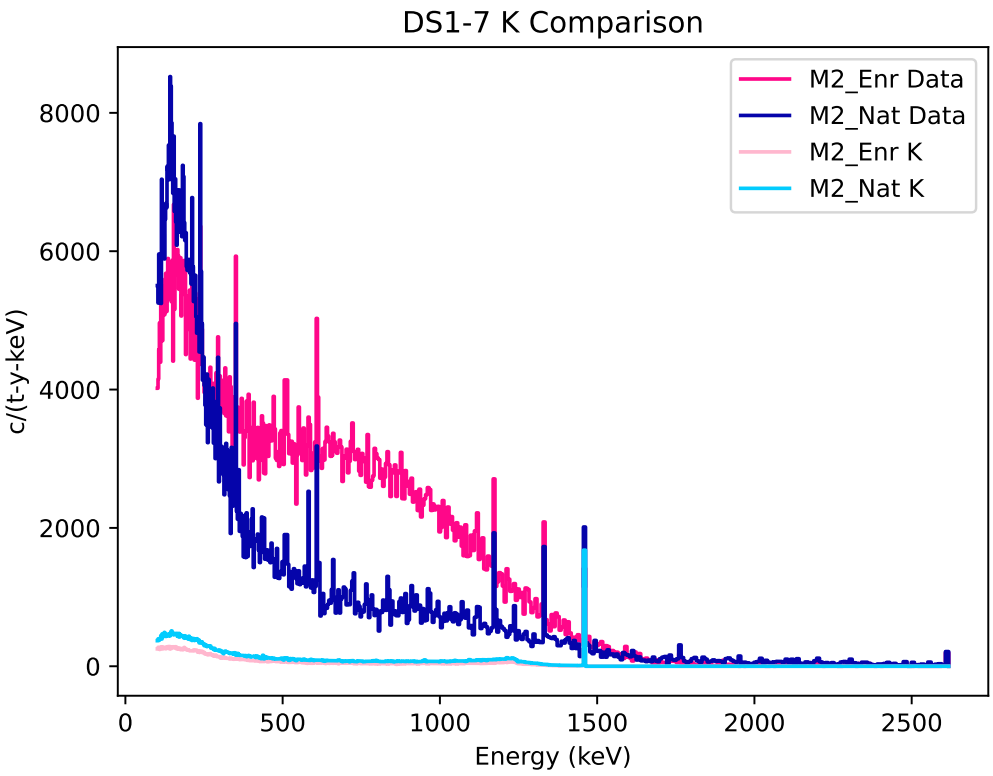

DS1-7 Rn Comparison
M1_Enr Data
M2_Enr Data
M1_Enr Rn
M2_Enr Rn
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

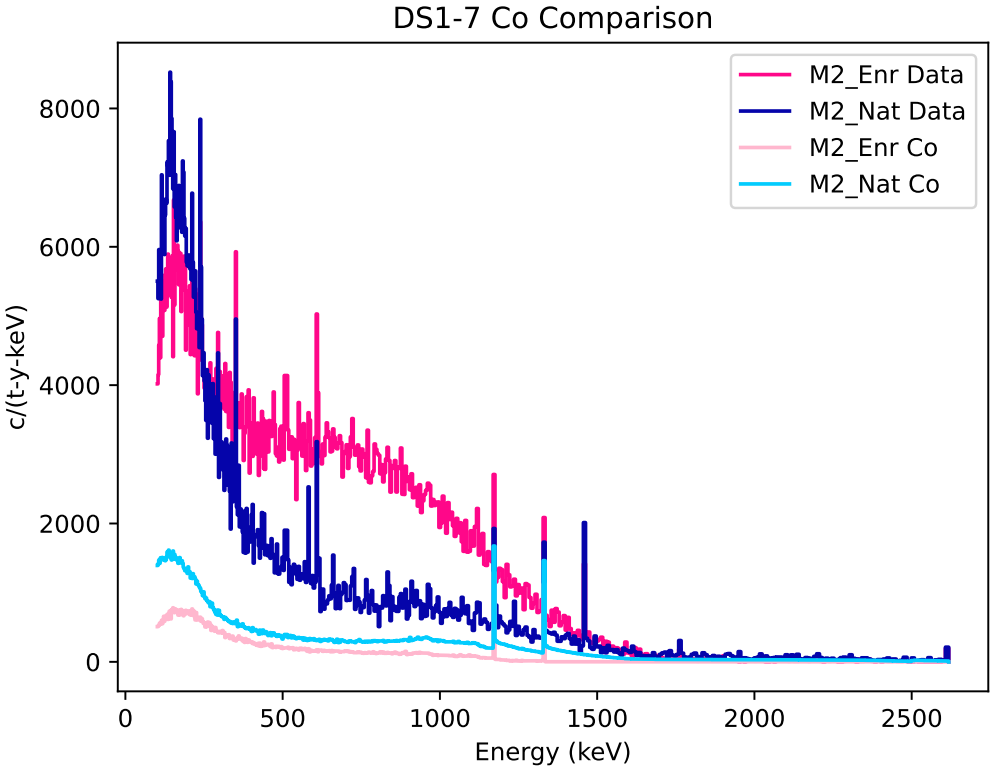

DS1-7 Full Model Comparison
M1_Enr Data
M2_Enr Data
M1_Enr Full Model
M2_Enr Full Model
c/(t-y-keV)
Energy (keV)
0
1000
2000
3000
4000
5000
6000
7000
8000
0
500
1000
1500
2000
2500

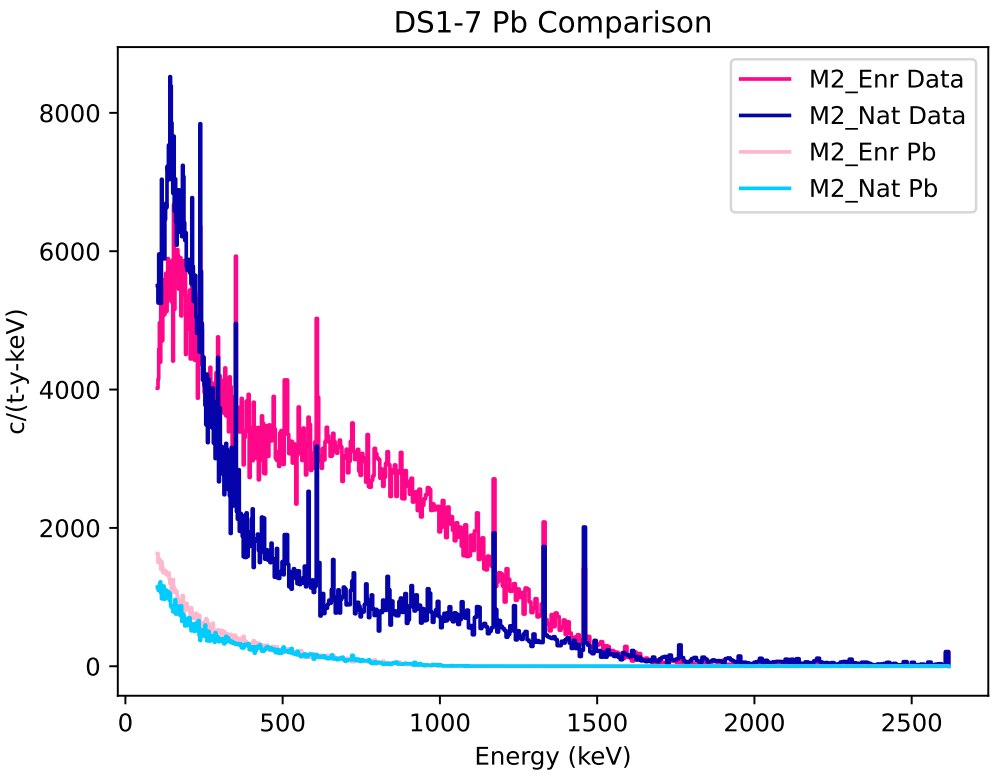

DS1-7 2v Comparison
M2_Enr Data
M2_Nat Data
M2_Enr 2v
M2_Nat 2v
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

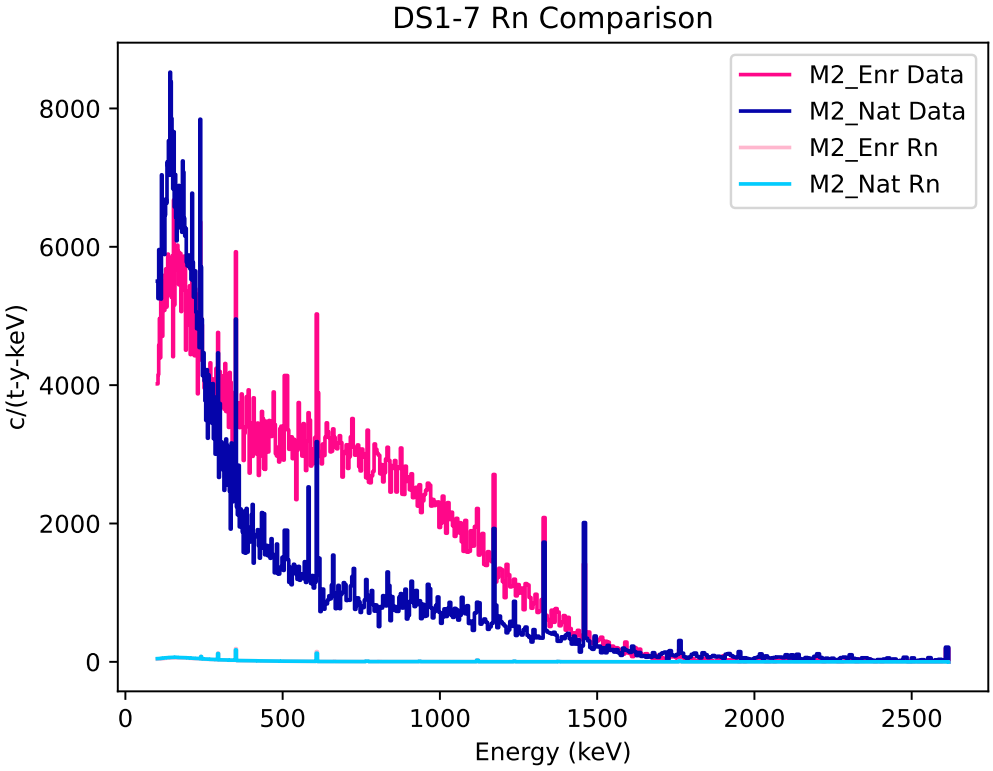

DS1-7 Th Comparison
M2_Enr Data
M2_Nat Data
M2_Enr Th
M2_Nat Th
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

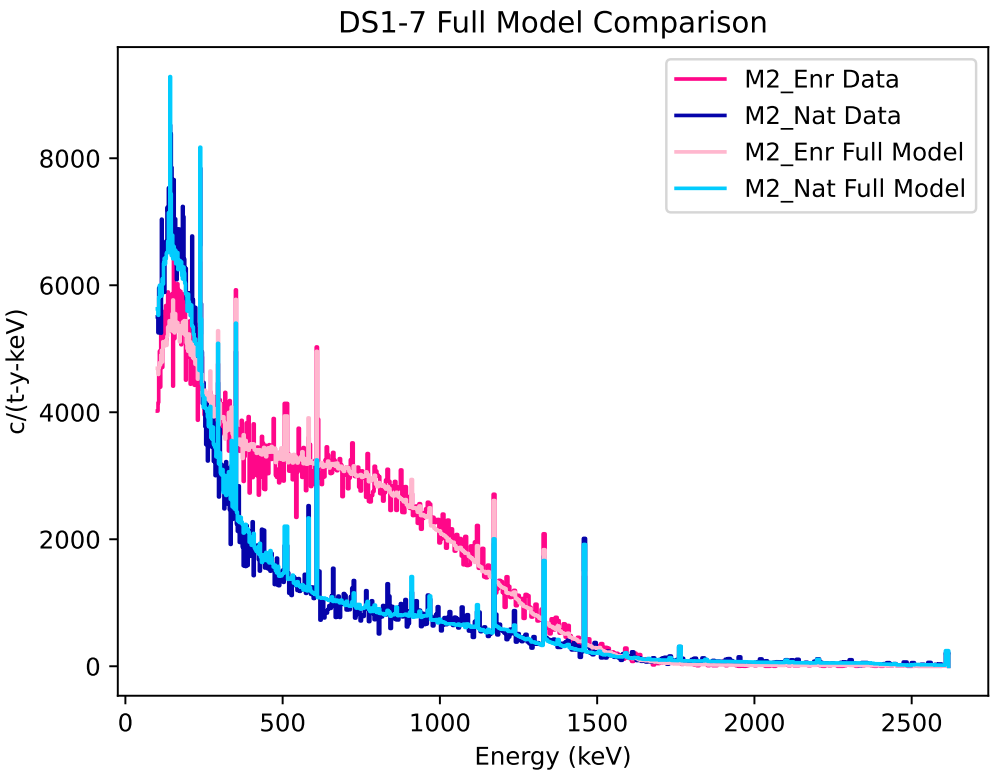

DS1-7 U Comparison
M2_Enr Data
M2_Nat Data
M2_Enr U
M2_Nat U
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

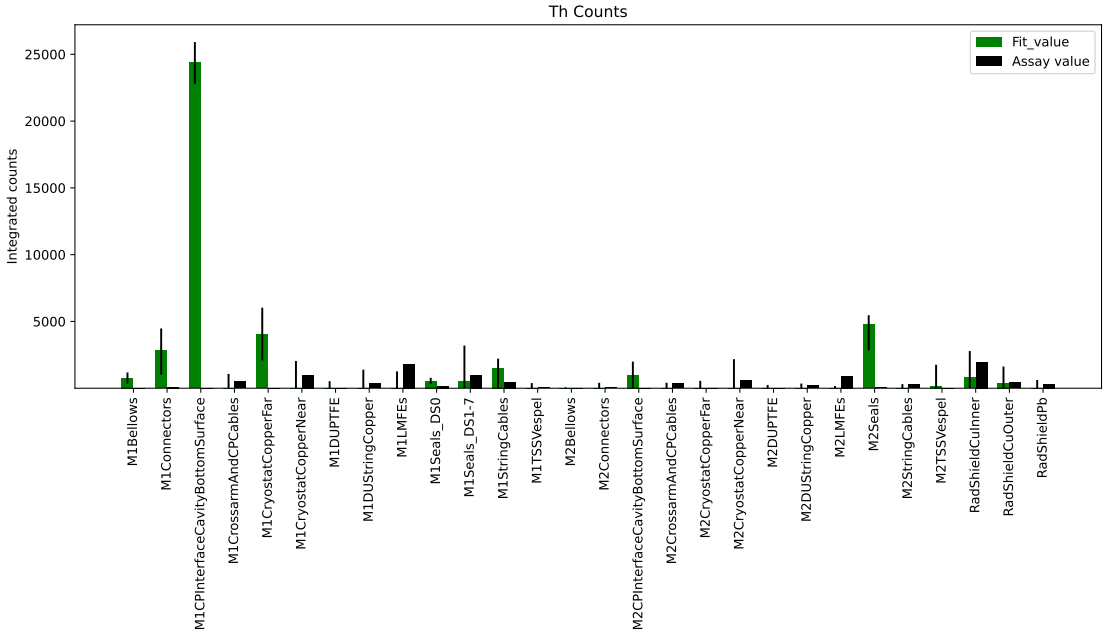

DS1-7 K Comparison
M2_Enr Data
M2_Nat Data
M2_Enr K
M2_Nat K
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

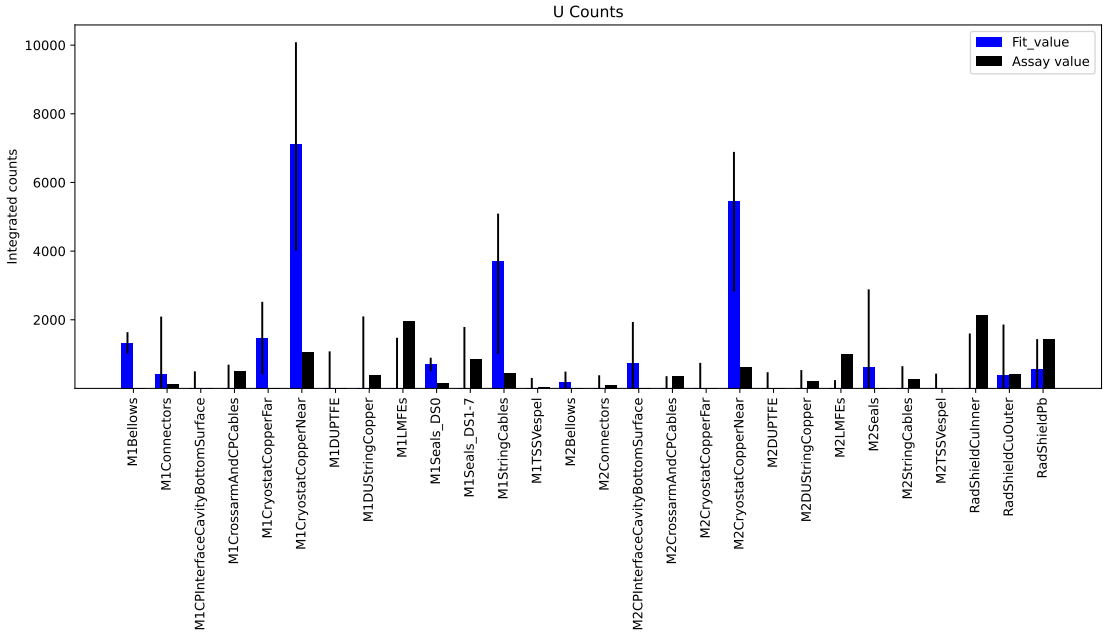

DS1-7 Co Comparison
M2_Enr Data
M2_Nat Data
M2_Enr Co
M2_Nat Co
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

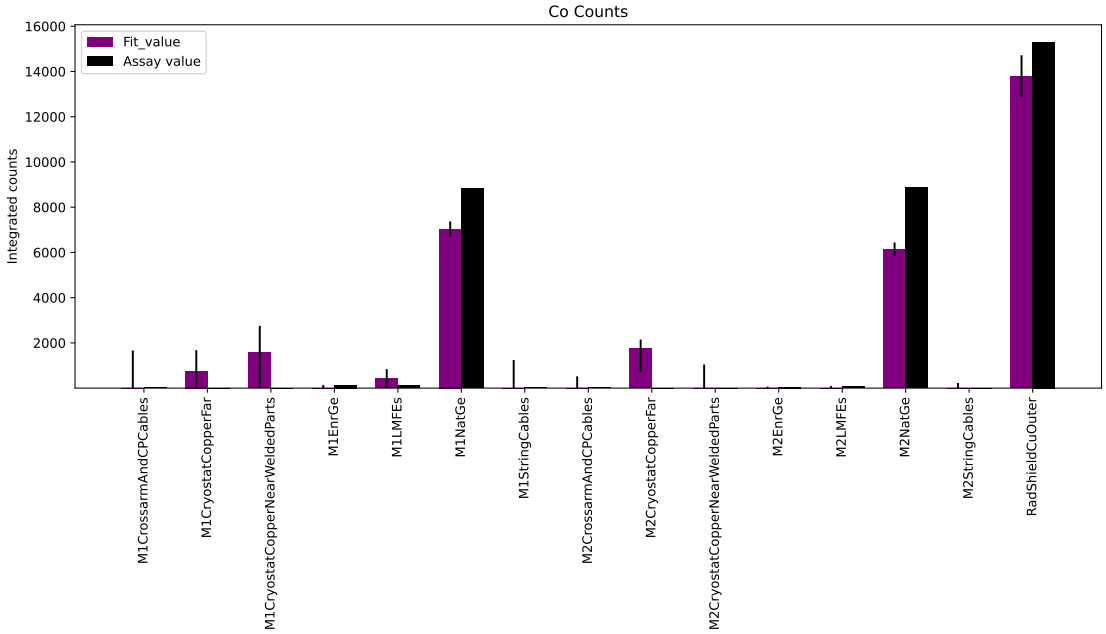

DS1-7 Pb Comparison
M2_Enr Data
M2_Nat Data
M2_Enr Pb
M2_Nat Pb
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

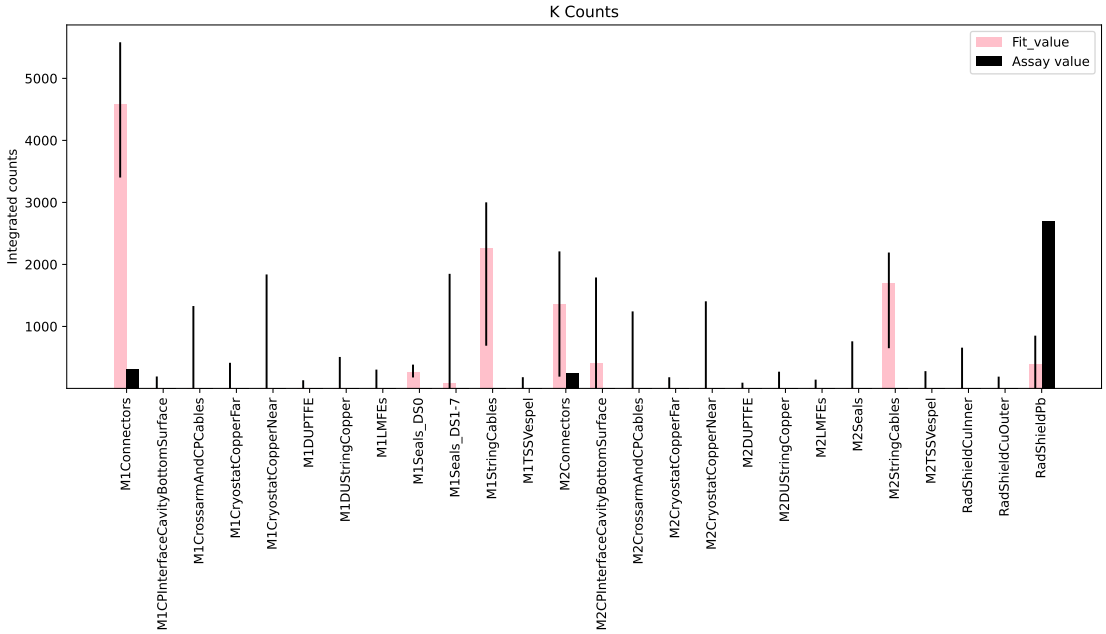

DS1-7 Rn Comparison
M2_Enr Data
M2_Nat Data
M2_Enr Rn
M2_Nat Rn
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

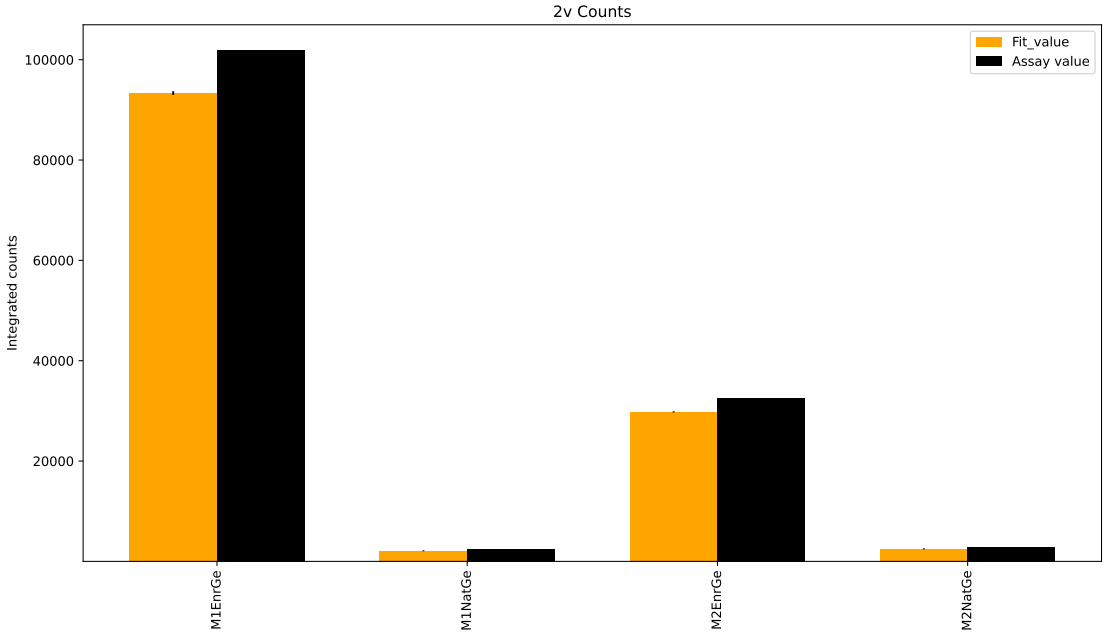

DS1-7 Full Model Comparison
M2_Enr Data
M2_Nat Data
M2_Enr Full Model
M2_Nat Full Model
c/(t-y-keV)
Energy (keV)
0
2000
4000
6000
8000
0
500
1000
1500
2000
2500

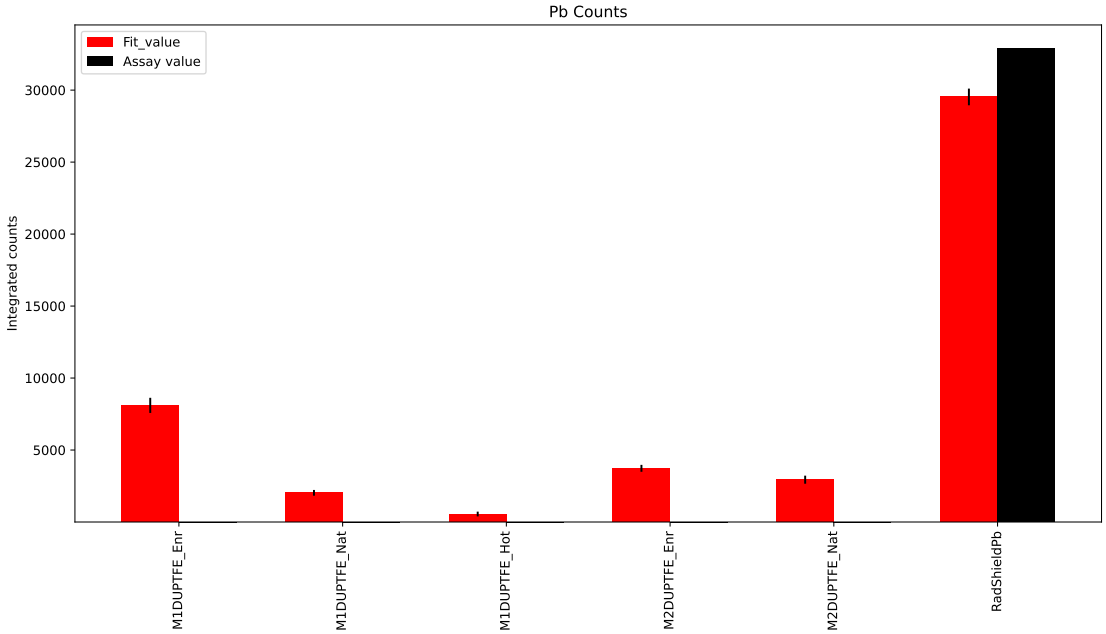

Th Counts
Fit_value
Assay value
Integrated counts
25000
20000
15000
10000
5000
M1Bellows
M1Connectors
M1CPInterfaceCavityBottomSurface
M1CrossarmAndCPCables
M1CryostatCopperFar
M1CryostatCopperNear
M1DUPTFE
M1DUStringCopper
M1LMFEs
M1Seals_DS0
M1Seals_DS1-7
M1StringCables
M1TSSVespel
M2Bellows
M2Connectors
M2CPInterfaceCavityBottomSurface
M2CrossarmAndCPCables
M2CryostatCopperFar
M2CryostatCopperNear
M2DUPTFE
M2DUStringCopper
M2LMFEs
M2Seals
M2StringCables
M2TSSVespel
RadShieldCuInner
RadShieldCuOuter
RadShieldPb

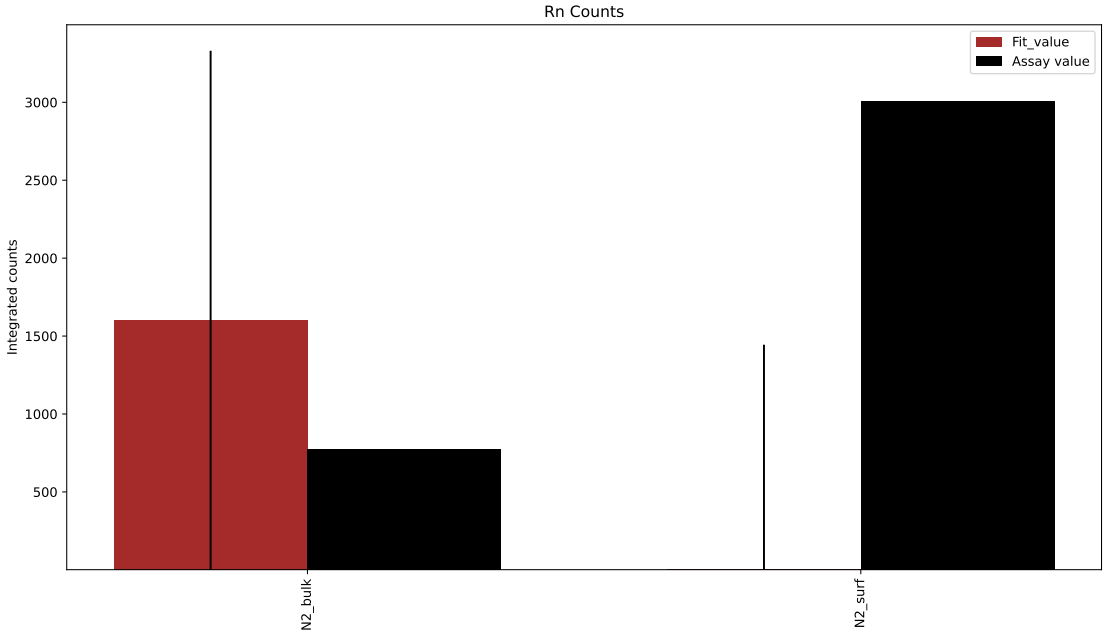

U Counts
Fit_value
Assay value
Integrated counts
10000
8000
6000
4000
2000
M1Bellows
M1Connectors
M1CPInterfaceCavityBottomSurface
M1CrossarmAndCPCables
M1CryostatCopperFar
M1CryostatCopperNear
M1DUPTFE
M1DUStringCopper
M1LMFEs
M1Seals_DS0
M1Seals_DS1-7
M1StringCables
M1TSSVespel
M2Bellows
M2Connectors
M2CPInterfaceCavityBottomSurface
M2CrossarmAndCPCables
M2CryostatCopperFar
M2CryostatCopperNear
M2DUPTFE
M2DUStringCopper
M2LMFEs
M2Seals
M2StringCables
M2TSSVespel
RadShieldCuInner
RadShieldCuOuter
RadShieldPb

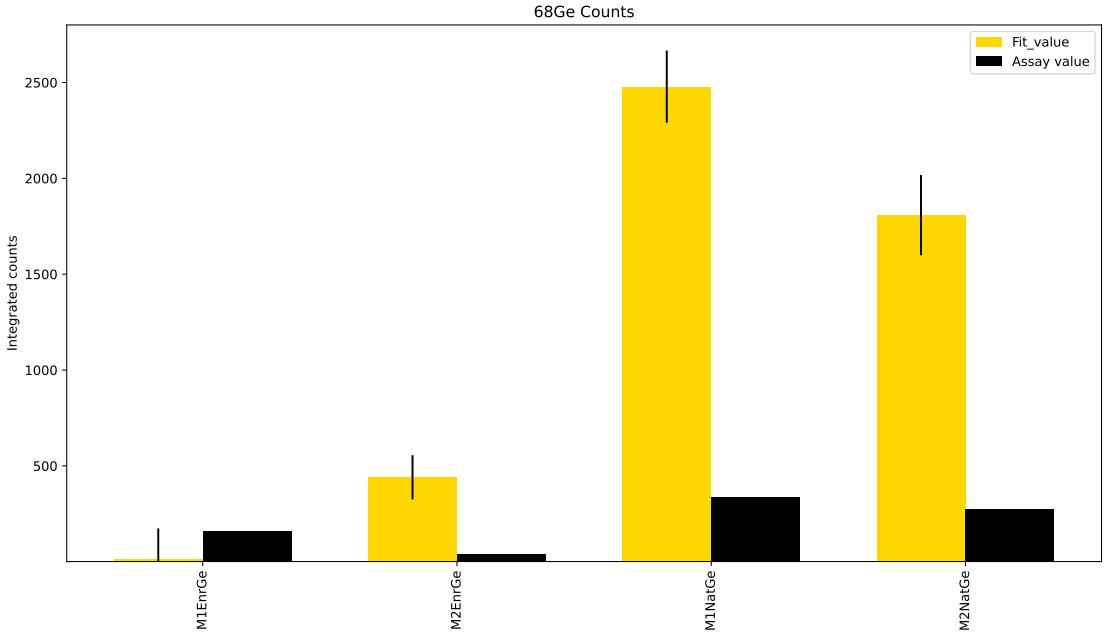

Co Counts
Fit_value
Assay value
Integrated counts
16000
14000
12000
10000
8000
6000
4000
2000
M1CrossarmAndCPCables
M1CryostatCopperFar
M1CryostatCopperNearWeldedParts
M1EnrGe
M1LMFEs
M1NatGe
M1StringCables
M2CrossarmAndCPCables
M2CryostatCopperFar
M2CryostatCopperNearWeldedParts
M2EnrGe
M2LMFEs
M2NatGe
M2StringCables
RadShieldCuOuter


K Counts
Fit_value
Assay value
Integrated counts
5000
4000
3000
2000
1000
M1Connectors
M1CPInterfaceCavityBottomSurface
M1CrossarmAndCPCables
M1CryostatCopperFar
M1CryostatCopperNear
M1DUPTFE
M1DUStringCopper
M1LMFEs
M1Seals_DS0
M1Seals_DS1-7
M1StringCables
M1TSSVespel
M2Connectors
M2CPInterfaceCavityBottomSurface
M2CrossarmAndCPCables
M2CryostatCopperFar
M2CryostatCopperNear
M2DUPTFE
M2DUStringCopper
M2LMFEs
M2Seals
M2StringCables
M2TSSVespel
RadShieldCuInner
RadShieldCuOuter
RadShieldPb


2v Counts
Fit_value
Assay value
Integrated counts
100000
80000
60000
40000
20000
M1EnrGe
M1NatGe
M2EnrGe
M2NatGe


Pb Counts
Fit_value
Assay value
Integrated counts
30000
25000
20000
15000
10000
5000
M1DUPTFE_Enr
M1DUPTFE_Nat
M1DUPTFE_Hot
M2DUPTFE_Enr
M2DUPTFE_Nat
RadShieldPb


Rn Counts
Fit_value
Assay value
Integrated counts
3000
2500
2000
1500
1000
500
N2_bulk
N2_surf


68Ge Counts
Fit_value
Assay value
Integrated counts
2500
2000
1500
1000
500
M1EnrGe
M2EnrGe
M1NatGe
M2NatGe

REFERENCES

[1] A. de Gouvêa, *On Determining the Neutrino Mass Hierarchy*, FNAL Theory Seminar (2006).

[2] M. Agostini, G. Benato, J. A. Detwiler, J. Menéndez, and F. Vissani, *Toward the discovery of matter creation with neutrinoless double-beta decay*, Rev. Mod. Phys. **95**, 025002 (2023).

[3] M. Hirsch, R. Srivastava, and J. W. F. Valle, *Can one ever prove that neutrinos are Dirac particle*, Phys. Lett. B **781**, 302 (2018).

[4] F. Deppisch, M. Hirsch, and H. Päs, *Neutrinoless double-beta decay and physics beyond the standard model*, J. Phys. G: Nucl. Part. Phys. **39**, 124007 (2012).

[5] M. Agostini, G. Benato, and J. Detwiler, *Discovery probability of next-generation neutrinoless double-β decay experiments*, Phys. Rev. D 96 **96**, 053001 (2017).

[6] N. Abgrall *et al.*, *The Majorana Demonstrator Neutrinoless Double-Beta Decay Experiment*, Adv. High Energy Phys. **2014**, 365432 (2014).

[7] S. Alvis *et al.*, *Multisite event discrimination for the Majorana Demonstrator*, Phys. Rev. C **99**, 065501 (2019).

[8] I. J. Arnquist *et al.*, *α-event Characterization and Rejection in Point-Contact HPGe Detectors*, Eur. Phys. J. C **82**, 226 (2022).

[9] I. Guinn *et al.*, *Updated Limit on $T^{0\nu}_{1/2}$ from DS0-8 of the Majorana Demonstrator*, Internal document (2022).

[10] N. Abgrall *et al.*, *LEGEND-1000 Preconceptual Design Report*, arXiv e-prints 2107.11462 (2021).

[11] I. J. Arnquist *et al.*, *Final Result of the Majorana Demonstrator's Search for Neutrinoless Double-β Decay in ^{76}Ge*, Phys. Rev. Lett. **130**, 062501 (2023), arXiv:2207.07638.

[12] A. Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012).

[13] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012).

[14] W. Pauli, *Dear radioactive ladies and gentlemen*, Phys. Today **31N9**, 27 (1978).

[15] J. Chadwick, *The Intensity Distribution in Magnetic Spectrum of β-Rays of Radium $B + C$*, Verhandl. Dtsch. Phys. **16** (1914).

[16] F. Reines and C. L. Cowan, *Detection of the free neutrino*, Phys. Rev. **93**, 830 (1953).

[17] C. L. Cowan *et al.*, *Detection of the free neutrino: A Confirmation*, Science **124**, 103 (1956).

[18] R. L. Workmanet *et al.*, *Review of Particle Physics*, Prog. Theor. Exp. Phys. **083C01** (2022).

[19] B. T. Cleveland *et al.*, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, Astrophys. J. **496**, 505 (1998).

[20] Y. Fukuda *et al.*, *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998).

[21] Q. R. Ahmad *et al.*, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89**, 011301 (2002).

[22] T. Kajita, *Nobel Lecture: Discovery of atmospheric neutrino oscillations*, Rev. Mod. Phys. **88**, 030501 (2016).

[23] M. Aker *et al.*, *Direct neutrino-mass measurement with sub-electronvolt sensitivity*, Nature Phys. **18**, 160 (2022).

[24] N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020).

[25] A. de Gouvêa, *Annual Review of Nuclear and Particle Science*, Annu. Rev. Nucl. Part. Sci. **66**, 197 (2016).

[26] S. Weinberg, *Baryon and lepton nonconserving processes*, Phys. Rev. Lett. **43**, 1566 (1979).

[27] S. Weinberg, *Varieties of baryon and lepton nonconservation*, Phys. Rev. D **22**, 1694 (1980).

[28] G. 't Hooft, *Symmetry breaking through bell-jackiw anomalies*, Phys. Rev. Lett. **37**, 8 (1976).

[29] P. F. Perez *et al.*, *Baryogenesis via Leptogenesis: Spontaneous B and L Violation*, Phys. Rev. D **104**, 055007 (2011).

[30] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).

[31] M. Goeppert-Mayer, *Double beta-disintegration*, Phys. Rev. **48**, 512 (1935).

[32] A. Barabash, *Precise Half-Life Values for Two-Neutrino Double-β Decay: 2020 Review*, Universe **6**, 159 (2020).

[33] M. Dolinski, A. Poon, and A. Rodejohann, *Neutrinoless Double-Beta Decay: Status and Prospects*, Annu. Rev. Nucl. Part. Sci. **69**, 219 (2019).

[34] S. Abe *et al.*, *Search for the Majorana Nature of Neutrinos in the Inverted Mass Ordering Region with KamLAND-Zen*, Phys. Rev. Lett. **130**, 051801 (2023).

[35] G. Anton *et al.*, *Search for Neutrinoless Double-Beta Decay with the Complete EXO-200 Dataset*, Phys. Rev. Lett. **123**, 161802 (2019).

[36] G. Adhikari, *NEXO: Neutrinoless double beta decay search beyond* 10^{28} *year half-life sensitivity*, J. Phys. G: Nucl. Part. Phys. **49**, 015104 (2022).

[37] C. Alduino *et al.*, *Search for Majorana neutrinos exploiting millikelvin cryogenics with CUORE*, Nature **604**, 53 (2022).

[38] W. Armstrong *et al.*, *CUPID Preconceptual Design Report*, arXiv e-prints 1907.09376 (2019).

[39] K. Alfonso *et al.*, *CUPID: The Next-Generation Neutrinoless Double Beta Decay Experiment*, Journal of Low Temperature Physics **211**, 5-6 (2022).

[40] M. Agostini *et al.*, *Final Results of GERDA on the Search for Neutrinoless Double-β Decay*, Phys. Rev. Let. **125**, 252502 (2020).

[41] J. Heise, *The Sanford Underground Research Facility at Homestake*, J. Phys. Conf. Ser. **606**, 012015 (2015).

[42] G. F. Knoll, *Radiation Detection and Measurement*, (John Wiley & Sons, Inc., 2000), 3rd ed.

[43] P. Luke *et al.*, *Low capacitance large volume shaped-field germanium detector*, IEEE Transactions on Nuclear Science **36**, 926 (1989).

[44] P. S. Barbeau *et al.*, *Large-mass ultralow noise germanium detectors: performance and applications in neutrino and astroparticle physics*, J. Cosmol. Astropart. Phys. **09**, 009 (2007).

[45] *Canberra Industries Inc. (now Mirion Technologies)*, `https://www.mirion.com/products/bege-broad-energy-germanium-detectors`, 800 Research Parkway Meriden, CT 06450.

[46] R. J. Cooper *et al.*, *A novel HPGe detector for gamma-ray tracking and imaging*, Nucl. Instrum. Methods Phys. Res. A **665**, 25 (2011).

[47] J. A. Detwiler, *Updated Isotopic Abundances and Molecular Weights of the Majorana Demonstrator Enriched Detectors*, Internal document (2022).

[48] E. Aguayo *et al.*, *The design of an ultra-low background thermosyphon for the Majorana Demonstrator*, Nucl. Instrum. Methods Phys. Res. A **709**, 17 (2013).

[49] N. Abgrall *et al.*, *The Majorana Demonstrator Radioassay Program*, Nucl. Instrum. Meth. A **828**, 22 (2016).

[50] N. Abgrall *et al.*, *The Majorana Demonstrator readout electronics system*, JINST **17**, T05003 (2022).

[51] J. Anderson *et al.*, *Data Acquisition and Trigger System of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA)*, IEEE Trans. Nucl. Sci. **56** (2009).

[52] N. Abgrall *et al.*, *ADC Nonlinearity Correction for the Majorana Demonstrator*, IEEE Trans. Nucl. Sci. **68** (2021).

[53] I. J. Arngquist *et al.*, *Charge trapping correction and energy performance of the Majorana Demonstrator*, Phys. Rev. C **107**, 045503 (2023).

[54] J. R. Klein and A. Roodman, *Blind analysis in nuclear and particle physics*, Annu. Rev. Nucl. Part. Sci. **55**, 141 (2005).

[55] I. J. Arngquist *et al.*, *Constraints on the decay of* ^{180m}Ta, Phys. Rev. Let. **131**, 152501 (2023).

[56] I. Arnquist *et al.*, *An assay-based background projection for the Majorana Demonstrator using Monte Carlo Uncertainty Propagation*, In preparation.

[57] M. Buuck, *A Radiogenic Background Model for the Majorana Demonstrator*, Ph.D. thesis, The University of Washington (2019).

[58] T. Gilliss, *Statistical Modeling and Markov Chain Monte Carlo Inference of Majorana Demonstrator Background Data*, Ph.D. thesis, The University of North Carolina at Chapel Hill, (2019).

[59] M. Agostini *et al.*, *Final Results of GERDA on the Two-Neutrino Double-β Decay Half-Life of* ^{76}Ge, arXiv e-prints 2308.09795 (2023).

[60] M. Hirsch *et al.*, *On the observability of majoron emitting double beta decays*, Phys. Lett. B **372** (1996).

[61] O. Nitescu, S. Ghinescu, and S. Stoica, *Lorentz violation effects in $2\nu\beta\beta$ decay*, J. Phys. G **47**, 055112 (2020).

[62] P. Bolton *et al.*, *Two-Neutrino Double Beta Decay with Sterile Neutrinos*, Phys. Rev. D **103**, 055019 (2021).

[63] M. Agostini *et al.*, *Search for Light Exotic Fermions in Double-Beta Decays*, Phys. Lett. B **815**, 136127 (2021).

[64] M. Agostini *et al.*, *Search for exotic physics in double-β decays with GERDA Phase II*, JCAP12 **012** (2022).

[65] M. Boswell *et al.*, *MaGe - a Geant4-based Monte Carlo Application Framework for Low-backgroiund Germanium Experiments*, IEEE Trans. Nucl. Sci. **58**, 1212 (2011).

[66] S. Agostinelli *et al.*, *Geant4–a simulation toolkit*, Nucl. Instrum. Meth. A **506**, 250 (2003).

[67] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53**, 1, 186 (2006).

[68] G. Giovanetti, *P-type point contact germanium detectors and their application in rare-event searches*, Ph.D. thesis, The University of North Carolina at Chapel Hill, (2015).

[69] C. R. Haufe, *A Study of Majorana Demonstrator Backgrounds with Bayesian Statistical Modeling*, Ph.D. thesis, The University of North Carolina at Chapel Hill, (2023).

[70] G. Heusser, M. Laubenstein, and H. Neder, *Low-level germanium gamma-ray spectrometry at the μBq/kg level and future developments towards higher sensitivity*, in P. P. and S.-C. J., eds., *Radionuclides in the Environment*, vol. 8, 495–510, (Elsevier, 2006).

[71] M. Laubenstein, Personal Communciation (2023).

[72] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77**, 219 (1993).

[73] H. Dembinski, P. Ongmongkolkul, *et al.*, *scikit-hep/iminuit* (2020).

[74] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10**, 343 (1975).

[75] R. Fletcher, *A new approach to variable metric algorithms*, The Computer Journal **13**, 317– (1970).

[76] S. Behnel *et al.*, *Cython: The best of both worlds*, Computing in Science & Engineering **13**, 2, 31 (2011).

[77] OpenMP Architecture Review Board, *OpenMP Application Program Interface Version 4.5* (2015).

[78] D. Andrews, *Inconsistency of the Bootstrap When a Parameter is on the Boundary of the Parameter Space*, Econometrica **68**, 399 (2000).

[79] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist **9**, 60 (1938).

[80] W. A. Rolke *et al.*, *Limits and Confidence Intervals in the Presence of Nuisance Parameters*, Nucl. Instrum. Meth. A **551**, 493 (2005).

[81] M. P. Green *et al.*, *A model of backgrounds in the Majorana Demonstrator*, Internal document (2023).

[82] A. Hostiuc and J. Gruszko, *Delayed Charge Recovery Tagging Update*, Internal document (2022).