New measurement of double-β decays of ^{100}Mo to excited states of ^{100}Ru with the CUPID-Mo experiment

C. Augier,[1] A. S. Barabash,[2] F. Bellini,[3,4] G. Benato,[5] M. Beretta,[6] L. Bergé,[7] J. Billard,[1] Yu. A. Borovlev,[8] L. Cardani,[4] N. Casali,[4] A. Cazes,[1] M. Chapellier,[7] D. Chiesa,[9,10] I. Dafinei,[4] F. A. Danevich,[11] M. De Jesus,[1] T. Dixon [7,12,*] L. Dumoulin,[7] K. Eitel,[13] F. Ferri,[12] B. K. Fujikawa,[14] J. Gascon,[1] L. Gironi,[9,10] A. Giuliani,[7] V. D. Grigorieva,[8] M. Gros,[12] D. L. Helis,[12,5] H. Z. Huang,[15] R. Huang,[6] L. Imbert,[7] J. Johnston,[16] A. Juillard,[1] H. Khalife,[7] M. Kleifges,[17] V. V. Kobychev,[11] Yu. G. Kolomensky,[6,14] S. I. Konovalov,[2] J. Kotila [18,19,20] P. Loaiza,[7] L. Ma,[15] E. P. Makarov,[8] P. de Marcillac,[7] R. Mariam,[7] L. Marini,[6,14,5] S. Marnieros,[7] X.-F. Navick,[12] C. Nones,[12] E. B. Norman,[6] E. Olivieri,[7] J. L. Ouellet,[16] L. Pagnanini,[21,5] L. Pattavina,[5,22] B. Paul,[12] M. Pavan,[9,10] H. Peng,[23] G. Pessina,[10] S. Pirro,[5] D. V. Poda,[7] O. G. Polischuk,[11] S. Pozzi,[10] E. Previtali,[9,10] Th. Redon,[7] A. Rojas,[24] S. Rozov,[25] V. Sanglard,[1] J. A. Scarpaci,[7] B. Schmidt,[14,†] Y. Shen,[15] V. N. Shlegel,[8] V. Singh,[6] C. Tomei,[4] V. I. Tretyak,[11] V. I. Umatov,[2] L. Vagneron,[1] M. Velázquez,[26] B. Welliver,[6] L. Winslow,[16] M. Xue,[23] E. Yakushev,[25] M. Zarytskyy,[11] and A. S. Zolotarova[7]

(CUPID-Mo Collaboration)

[1]*Univ. Lyon, Université Lyon 1, CNRS/IN2P3, IP2I-Lyon, F-69622, Villeurbanne, France*
[2]*National Research Centre Kurchatov Institute, Institute of Theoretical and Experimental Physics, 117218 Moscow, Russia*
[3]*Dipartimento di Fisica, Sapienza Università di Roma, P.le Aldo Moro 2, I-00185, Rome, Italy*
[4]*INFN, Sezione di Roma, P.le Aldo Moro 2, I-00185, Rome, Italy*
[5]*INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi (AQ), Italy*
[6]*University of California, Berkeley, California 94720, USA*
[7]*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France*
[8]*Nikolaev Institute of Inorganic Chemistry, 630090 Novosibirsk, Russia*
[9]*Dipartimento di Fisica, Università di Milano-Bicocca, I-20126 Milano, Italy*
[10]*INFN, Sezione di Milano-Bicocca, I-20126 Milano, Italy*
[11]*Institute for Nuclear Research of NASU, 03028 Kyiv, Ukraine*
[12]*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[13]*Karlsruhe Institute of Technology, Institute for Astroparticle Physics, 76021 Karlsruhe, Germany*
[14]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[15]*Key Laboratory of Nuclear Physics and Ion-beam Application (MOE), Fudan University, Shanghai 200433, P.R. China*
[16]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[17]*Karlsruhe Institute of Technology, Institute for Data Processing and Electronics, 76021 Karlsruhe, Germany*
[18]*Department of Physics, University of Jyväskylä, P.O. Box 35, FI-40014, Jyväskylä, Finland*
[19]*Finnish Institute for Educational Research, University of Jyväskylä, P.O. Box 35, FI-40014. Jyväskylä, Finland*
[20]*Center for Theoretical Physics, Sloane Physics Laboratory, Yale University, New Haven, Connecticut 06520-8120, USA*
[21]*INFN, Gran Sasso Science Institute, I-67100 L'Aquila, Italy*
[22]*Physik Department, Technische Universität München, Garching D-85748, Germany*
[23]*Department of Modern Physics, University of Science and Technology of China, Hefei 230027, P.R. China*
[24]*LSM, Laboratoire Souterrain de Modane, 73500 Modane, France*
[25]*Laboratory of Nuclear Problems, JINR, 141980 Dubna, Moscow region, Russia*
[26]*Université Grenoble Alpes, CNRS, Grenoble INP, SIMAP, 38420 Saint Martin d'Héres, France*

The CUPID-Mo experiment, located at the Laboratoire Souterrain de Modane (France), was a demonstrator experiment for CUPID. It consisted of an array of 20 Li$_2$ ^{100}MoO$_4$ (LMO) calorimeters, each equipped with a Ge light detector for particle identification. In this work, we present the result of a search for two-neutrino and neutrinoless double-β decays of ^{100}Mo to the first 0^+ and 2^+ excited states of ^{100}Ru using the full CUPID-Mo exposure (2.71 kg yr of LMO). We measure the half-life of $2\nu\beta\beta$ decay to the 0_1^+ state as $T_{1/2}^{2\nu \to 0_1^+} = (7.5 \pm 0.8 \,(\text{stat.}) \, ^{+0.4}_{-0.3} \,(\text{syst.})) \times 10^{20}$ yr. The bolometric technique enables measurement of the electron energies as well as the γ rays from nuclear deexcitation and this allows us to set new limits on the two-neutrino decay to the 2_1^+ state of $T_{1/2}^{2\nu \to 2_1^+} > 4.4 \times 10^{21}$ yr (90% c.i.) and on the neutrinoless modes of $T_{1/2}^{0\nu \to 2_1^+} > 2.1 \times 10^{23}$ yr (90% c.i.),

---

*toby.dixon@universite-paris-saclay.fr
†Present address: Northwestern University, Evanston, Illinois 60208, USA.

$T_{1/2}^{0\nu \to 0_1^+} > 1.2 \times 10^{23}$ yr (90% c.i.). Information on the electrons' spectral shape is obtained, which allows us to make the first comparison of the single and higher state dominance $2\nu\beta\beta$ decay models for the 0_1^+ excited state of ^{100}Ru.

I. INTRODUCTION

Two-neutrino double-β decay ($2\nu\beta\beta$) is an allowed standard model process which occurs in some even-even nuclei for which single-β decays are energetically forbidden or heavily disfavored due to large changes in angular momentum [1,2]. These decays have been observed to the 0^+ ground states ($0_{g.s.}^+$) in eleven nuclei and to the first zero-plus excited state (0_1^+) for two nuclei, ^{100}Mo and ^{150}Nd [1–3]. In addition, if neutrinos are Majorana particles then an additional decay mode becomes allowed: neutrinoless double-β decay ($0\nu\beta\beta$) [4–6]. Whereas $2\nu\beta\beta$ decay conserves lepton number, $0\nu\beta\beta$ decay would violate this symmetry [4,7,8] and provide a possible path to explain the predominance of matter over antimatter in the universe [4,8–11].

One of the most promising experimental techniques to search for $0\nu\beta\beta$ decay are cryogenic calorimeters. These detectors provide an excellent energy resolution, high detection efficiency, and are scalable to tonne scale arrays, such as the CUORE experiment (Cryogenic Underground Observatory for Rare Events) [12]. The CUPID experiment [13] (CUORE Upgrade with Particle IDentification) is an upgrade for CUORE which will use a scintillating bolometer technology to discriminate β and γ particles from α particles. CUPID-Mo [14] was a demonstrator experiment for CUPID aiming to validate the technology of lithium molybdate (LMO) scintillating calorimeters.

Multiple models exist to describe $0\nu\beta\beta$ decay [7,8,15–21], however the minimal extension to the standard model needed to explain $0\nu\beta\beta$ decay is the light Majorana neutrino exchange mechanism [22], in which $0\nu\beta\beta$ decay is mediated by a Majorana neutrino. In this framework, the decay rate of $0\nu\beta\beta$ decay is related to the effective Majorana neutrino mass $\langle m_{\beta\beta} \rangle$ (a weighted sum of the three neutrino masses) by

$$\left(T_{1/2}^{0\nu}\right)^{-1} = G_{0\nu} g_A^4 |M_{0\nu}|^2 \frac{\langle m_{\beta\beta} \rangle^2}{m_e^2}, \tag{1}$$

where $T_{1/2}^{0\nu}$ is the $0\nu\beta\beta$ decay half-life, $G_{0\nu}$ is the phase-space factor, $M_{0\nu}$ is the nuclear matrix element (NME), m_e is the electron mass, and g_A is the axial-vector coupling constant. The phase-space factor $G_{0\nu}$ can be computed accurately, but the NME is the result of complicated many-body nuclear physics calculations [23]. Several models with different approximations are used, but these only agree to within a factor of a few on the value of the NME (cf. Refs. [24–27]). At present the upper limit on the effective Majorana mass $\langle m_{\beta\beta} \rangle$ ranges from 60 to 600 meV [8,12,28–30]. Despite its low exposure, CUPID-Mo has set the most stringent limit on $0\nu\beta\beta$ decay in ^{100}Mo, with $T_{1/2}^{0\nu} > 1.8 \times 10^{24}$ yr or $\langle m_{\beta\beta} \rangle < (0.28$–$0.49)$ eV [31].

For $2\nu\beta\beta$ decay, the half-life is related to the NME as

$$\left(T_{1/2}^{2\nu}\right)^{-1} = G_{2\nu} |M_{2\nu}^{\text{eff}}|^2, \tag{2}$$

where $G_{2\nu}$ is the phase-space factor and $M_{2\nu}^{\text{eff}}$ is the dimensionless effective nuclear matrix element including the factor g_A^2, $M_{2\nu}^{\text{eff}} = g_A^2 M_{2\nu}$ [2]. Measurements of $T_{1/2}^{2\nu}$ for both ground states (g.s.) and excited states (e.s.) of a daughter nucleus provide experimental data which can be used to validate the methods to calculate the NME $M_{2\nu}$, and by extension, provide better confidence in the calculation of $M_{0\nu}$.

While the phase space, and therefore decay probability, is lower for decays to e.s. compared with transitions to the g.s., the presence of monochromatic γ's can lead to a very clear detection signature. In the case of a positive observation of $0\nu\beta\beta$ decay, measurements of the $0\nu\beta\beta$ decay to e.s. will be a useful tool to understand the decay mechanism [3]. Additionally, while the $0\nu\beta\beta$ decay to 0_1^+ state is expected to be suppressed relative to the decay to $0_{g.s.}^+$ state, it can still improve the overall sensitivity to $\langle m_{\beta\beta} \rangle$, especially if measurements of the ground-state decay are limited by backgrounds. Finally, a $2\nu\beta\beta$ decay to 2_1^+ e.s. is expected to be strongly suppressed by angular-momentum conservation. However, this mode is expected to be much more likely in a framework with bosonic neutrinos [32,33] and can be used to test this model.

We can exploit the specific combination of energy depositions associated with these events to measure $\beta\beta$ decays to e.s. As shown in Fig. 1, the two βs are accompanied by one or more γ particles. These have a much longer range (≈ 10 cm for 1 MeV γ [34]) compared with β^- (≈ 2 mm [35]) in LMO. A $\beta\beta$ e.s. event therefore has a high probability
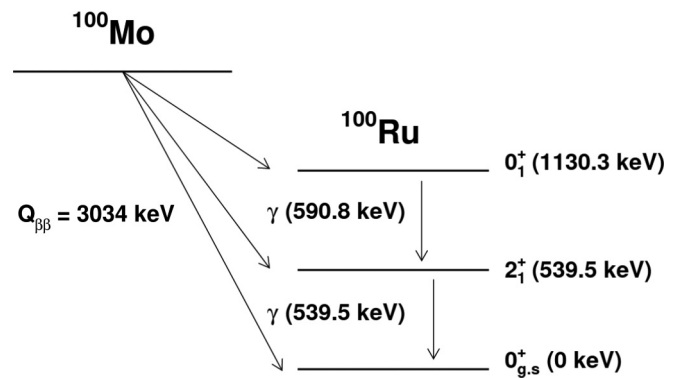

FIG. 1. Transition scheme for the decay of ^{100}Mo to ^{100}Ru showing the double-β decays to the 0_1^+ (1130.3 keV) and 2_1^+ (539.5 keV) states. The decay to the 0_1^+ state proceeds via the 2_1^+ state with $\approx 100\%$ probability, so both a 539.5 and a 590.8 keV γ are emitted in every decay, while only the 539.5 keV γ is emitted in the double-β decay to the 2_1^+ state.
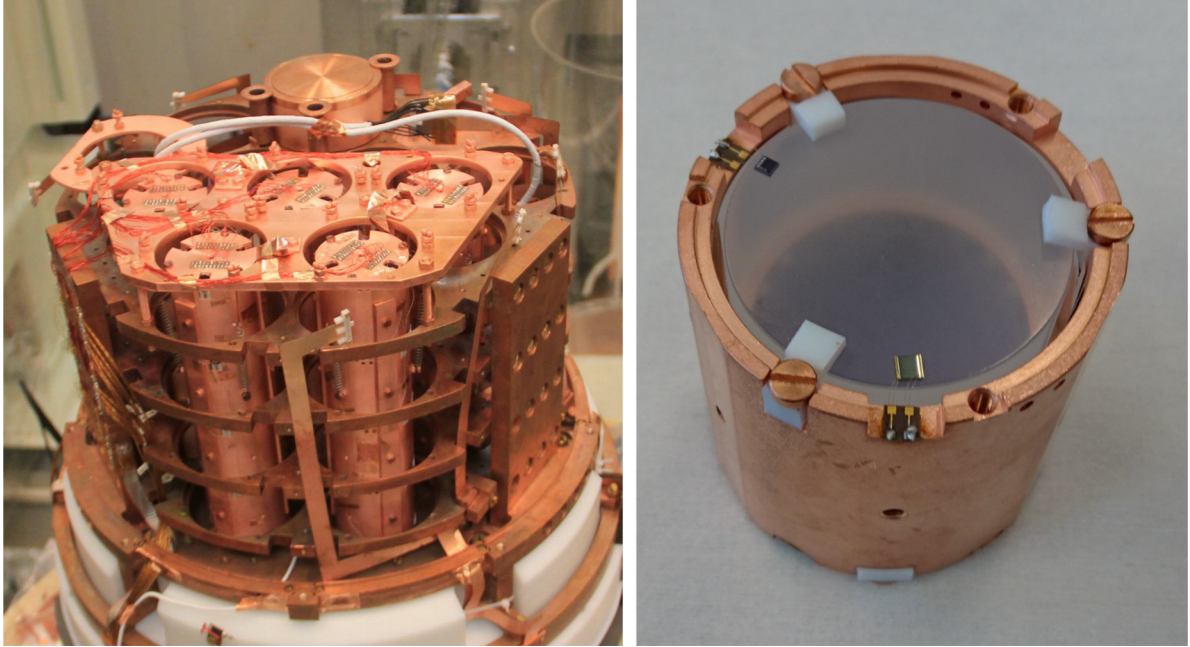
FIG. 2. CUPID-Mo experimental geometry showing (left) the CUPID-Mo towers mounted in the EDELWEISS cryostat and single detector module (right) with LMO crystal, PTFE clamps, NTD-Ge thermistor, Si heater, and Cu holder. The Ge LD is seen through the LMO crystal.

to induce a multidetector signature with the γ's escaping the detector where the decay took place and being measured in neighboring detectors. The additional information provided by these multidetector events allows us to significantly reduce the background rate while the γ peaks provide a clean signature to robustly measure the decay rate [36]. Since decays to higher e.s. are generally disfavored by the lower phase space due to lower Q value (and also by angular-momentum conservation for 2^+ decays), we focus on the lowest energy e.s.: the first zero plus (0_1^+) and two plus (2_1^+) e.s. of ^{100}Ru at 1130.3 keV and 539.5 keV (see Fig. 1).

The $2\nu\beta\beta$ decay of ^{100}Mo to the 0_1^+ e.s. has been measured in several experiments [37–42], mostly using high-purity germanium (HPGe) counters (cf. Refs. [2,3,43] for a recent review) with an external Mo (enriched) source and also NEMO-3 using a tracking detector and external sources. For the HPGe measurements, given that the source and detector did not coincide, only γ's were detected. The $2\nu\beta\beta$ decay to the 2_1^+ e.s. has not been observed and only lower limits on the half-life are available [39,42]. Placing constraints on this decay is complicated by the overlap with the decays to the 0_1^+ state since both decays include a 539.5 keV γ emission. By using a setup where the source is embedded in the detector, such as cryogenic calorimeters, we are able to measure the electron energies as well as the γ's. This provides additional information to separate the $2\nu\beta\beta$ decays to the 2_1^+ and 0_1^+ e.s., and to validate that the decay is indeed $2\nu\beta\beta \to 0_1^+$ and not a background process. This type of detector also allows us to effectively distinguish the neutrinoless and two-neutrino decay modes, which is not possible with HPGe detectors using an external $\beta\beta$ source.

In this work, we present an analysis of $2\nu\beta\beta/0\nu\beta\beta \to 0_1^+/2_1^+$ e.s. which exploits the measurement of both β and γ energies. In Sec. II we describe the CUPID-Mo experi-

ment; we introduce the search technique and GEANT4 Monte Carlo simulations in Sec. III. In Sec. IV, an overview of the experimental data and treatment of multisite events used for this analysis are given. We describe our Bayesian analysis framework in Sec. V, our treatment of systematic uncertainties in Sec. VI, and the obtained results in Sec. VII. We conclude in Sec. VIII by comparing our measurement of $2\nu\beta\beta \to 0_1^+$ state to theoretical calculations and we discuss the prospects for further investigations of the decay mechanism as well as sensitivity to models beyond the light Majorana neutrino exchange.

II. EXPERIMENT

CUPID-Mo consisted of an array of twenty ≈ 200 g LMO detectors (Fig. 2) enriched to $\approx 97\%$ in ^{100}Mo, with a total mass of 4.16 kg of LMO [14] and 2.26 kg of ^{100}Mo. These detectors were operated as cryogenic calorimeters at ≈ 20 mK, in the EDELWEISS [44] cryostat at the Laboratoire Souterrain de Modane (LSM), France. This technique allows for a very high detection efficiency ($\approx 75\%$ for $0\nu\beta\beta$ decay to g.s.) due to the detector containing the source, and excellent energy resolution ($\sigma \approx 0.1\%$) at 3 MeV [14]. In addition, CUPID-Mo employed a dual readout with twenty cryogenic light detectors (LDs) consisting of Ge wafers also operated as calorimeters. These allow for discrimination between (i) αs and (ii) βs and γ's since LMO is a scintillating crystal and the amount of light produced for β and γ is about five times higher compared with that produced by α particles of the same energy [14,45].

Each LMO crystal is assembled into an independent detector module with a Ge LD, NTD-Ge [46] thermistors on both LMO and LD, Si heater, copper holder, reflective foil (3M Vikuiti™, which guides the light to the LDs), and PTFE

clamps (shown in Fig. 2). While this design allows for a modular pre-assembly and is compatible with the operation of multiple payloads in the EDELWEISS cryogenic system, it results in a relatively high copper/LMO mass ratio of ≈1.5. The individual detector modules are arranged in an array of five towers (see Fig. 2) so that each LMO detector faces two LDs (apart from those on the top floor which only have a lower light detector).

CUPID-Mo took data between 2019 and 2020, collecting a total exposure of 2.71 kg yr of LMO. The data taking was organized into nine datasets, periods of stable operations lasting around a month. As in Ref. [47], three short periods of data taking which could not be calibrated to the same accuracy were discarded.

III. SEARCH TECHNIQUE AND SIMULATIONS

In $\beta\beta$ decays to e.s. the electrons are accompanied by γ's so they often reconstruct in multisite or coincident events where multiple detectors are triggered simultaneously. The particular set of energies obtained allows us to make a pure selection of e.s. decay events with very low background. We define the multiplicity $\mathcal{M}$ as the number of crystals subject to simultaneous energy depositions above the analysis energy threshold, chosen as 40 keV, well above the detector trigger thresholds and within a time window of ± 10 ms well above the detector time resolution [31].

For an event which is a time coincidence between several LMO detectors, our minimal experimental observable information is the energy deposited in each detector. We define energies $E_1, E_2, E_3, \ldots, E_{\mathcal{M}}$ as the energy deposited in each detector, sorted so that

$$E_1 > E_2 > E_3 > \cdots > E_{\mathcal{M}}. \tag{3}$$

In principle, the most sensitive approach exploiting all the information contained in the data would be a fit directly to this multidimensional energy spectrum. However, this analysis has significant complexity due to the difficulty of quantifying the background shape in this high-dimensional space.

An alternative approach that has been used by the CUORE-0, CUORE, and CUPID-0 experiments [48–50] is to select categories of signal signatures consistent with being from an e.s. decay. For each category, only the energy variable containing the γ (or two electrons for $0\nu\beta\beta$ decay) peak is fit (the peak energy). The other energy variables are referred to as "projected-out" energies. This technique has the advantage that the decay rate can be extracted as a set of one-dimensional fits to a peak over a background, which is a simple and robust technique. We use this approach in our analysis; however, we modify this technique by dividing up categories based on their projected-out energy variables. Since the decays to different e.s. significantly overlap, this allows us to better discriminate between them. Furthermore, it avoids integrating together regions with different background levels.

A. Monte Carlo simulations

We use dedicated Monte Carlo (MC) simulations to identify the most promising categories of events to include and

optimize the associated energy cuts. The objective is to identify categories which possess a high signal-to-background ratio to define the energy variable to project onto, to select the energy boundaries for each signal category, and to assess the containment efficiency therein. A GEANT4 [51] model has been created which implements the CUPID-Mo detector into the existing EDELWEISS MC software package [52]. Special care is taken to implement the detector structure as accurately as possible, in particular accounting for the individual dimensions of each LMO crystal, and the passive holder material close to the detectors. We simulate both $2\nu\beta\beta$ and $0\nu\beta\beta$ decays to the 0_1^+ and 2_1^+ e.s. This accounts for the angular correlation between 539.5 and 590.8 keV γ's from the 0_1^+ ($0_1^+ - 2_1^+ - 0_{g.s}^+$) cascade as [53]

$$\rho(\theta) = 1 - 3\cos^2(\theta) + 4\cos^4(\theta). \tag{4}$$

Here $\rho(\theta)$ is the probability for γ's to be emitted at angle θ from each other. Particles are then propagated through the experimental geometry using the Livermore low-energy physics list [54], applicable down to very low energy (250 eV). For the $2\nu\beta\beta$ decay to 0_1^+ e.s. simulations are performed using both the single state dominance (SSD) and higher state dominance (HSD) models [55]. Currently, no experimental data favor one model over another for the 0_1^+ state; however, the SSD model is strongly favored for the ground-state decay [56,57] and so we use it as our default. The difference to the HSD model is then treated as a systematic uncertainty. We combine the MC simulations with inputs from experimental data to reproduce the detector response of each individual detector, accounting for

 (i) the energy threshold (set at 40 keV) [31];
 (ii) the multiplicity of events;
 (iii) the energy resolution of each detector-dataset pair (energy dependent);
 (iv) the rejected periods of detector operations due to periods of cryogenic instability (i.e., due to environmental disturbances);
 (v) scintillation light, light-detector energy resolution, and cuts based on the LDs (see Sec. IV).

In our MC simulations, we observe that most excited states events are in $\mathcal{M}_1$–$\mathcal{M}_2$ with a small contribution of $\mathcal{M}_3$ (Fig. 3). The large fraction of $\mathcal{M}_1$ events is caused by the significant volume of passive material (Cu housing) and a relatively small modular detector array. We exclude $\mathcal{M}_1$ $2\nu\beta\beta$ decay events as the continuous energy spectrum of the βs does not produce a peak to perform a fit. For $0\nu\beta\beta$ decays, there are monoenergetic peaks in $\mathcal{M}_1$, albeit with a high background. We therefore focus on $\mathcal{M}_2$ events which have the highest sensitivity, and we also include some categories of $\mathcal{M}_3$ events.

B. $2\nu\beta\beta$ decay categories

We first describe categories for $2\nu\beta\beta$ decay analysis. For the analysis of $\mathcal{M}_2$ events the energies can be visualized as two-dimensional histograms. We plot these two-dimensional energy spectra for both e.s. decays ($2\nu\beta\beta$ and $0\nu\beta\beta$ decay) to 0_1^+ and 2_1^+ e.s. in Fig. 4. When selecting categories of events to use in our analysis, we require at least one peak to perform
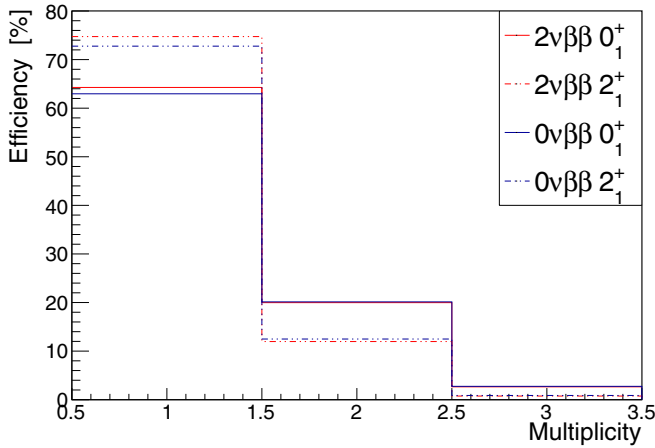
FIG. 3. Distribution of multiplicities (MC simulation) for each e.s. decay. Due to the small crystal array and significant amount of nonsensitive holder material, both decay modes to the $(0_1^+, 2_1^+)$ state are dominated by single crystals hits, (10%– 20%) of events cause energy deposits in two crystals, and $O(1\%)$ in three or more.

a fit. This selection is equivalent to looking for lines or points on the two-dimensional plane.

For $2\nu\beta\beta$ decays in MC simulation we observe a very clear signature of events where either E_1 or E_2 has a γ energy (either 539.5 or 590.8 keV) as shown in Fig. 4 (top). These events can be interpreted as the $\beta\beta$ being contained in one detector, giving rise to a continuous distribution, and a γ that is fully contained in a second LMO detector. In the case of the decay to the 0_1^+ state, the second γ either escapes the detector or is contained in the crystal with the $\beta\beta$s. We notice that the 2_1^+ and 0_1^+ decays (partially) overlap and the vertical and horizontal lines cover a wide range of projected out ($\beta\beta$) energies. This motivates us to divide the observed signal distribution into five slices, two for the horizontal and three for the vertical lines since the vertical lines cover a larger range in energy. The boundaries of this categories are defined to maximize the experimental sensitivity using MC simulations, as explained in Appendix A. This leads to the energy categories highlighted in Fig. 4 (top) and listed in Table I (index 0 to 4).

To identify $\mathcal{M}_3$ categories we note that in MC simulations generally the electrons carry the largest kinetic energy and
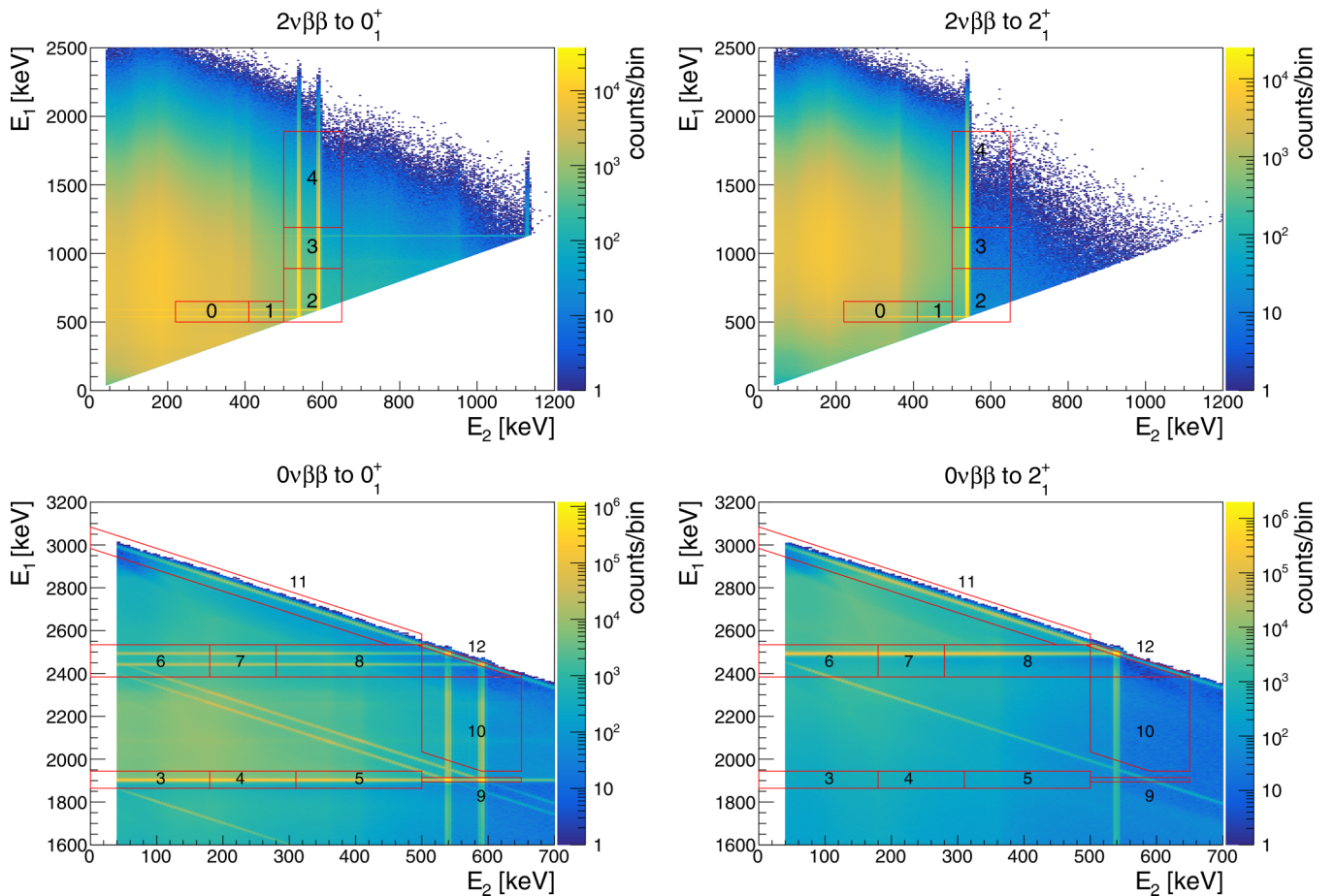

FIG. 4. Monte Carlo simulated two-dimensional $\mathcal{M}_2$ energy distributions for decays to e.s. (top left) $2\nu\beta\beta$ decay to 0_1^+, (top right) $2\nu\beta\beta$ decay to 2_1^+, (bottom left) $0\nu\beta\beta$ decay to 0_1^+, and (bottom right) $0\nu\beta\beta$ decay to 2_1^+. We show red boxes to highlight the $\mathcal{M}_2$ signal categories used in the analysis (see Tables I and II) and we label the associated category numbers.

TABLE I. Categories for $2\nu\beta\beta$ decay to 0_1^+, 2_1^+ e.s., these are given a name describing the multiplicity, peak energy variable, and then an index (if necessary to distinguish between categories). The table lists the energy cuts for the variable that is projected out, while the projected energy range for each category is 500–650 keV. For every category the γ's are either 539.5 or 590.8 keV.

Cat.	Multiplicity	Peak variable	Energy cuts	Interpretation
0	2	E_1	$E_2 \in [220, 410]$	γ in E_1, $\beta\beta$ in E_2
1	2	E_1	$E_2 \in [410, 500]$	γ in E_1, $\beta\beta$ in E_2
2	2	E_2	$E_1 \in [500, 890]$	γ in E_2, $\beta\beta$ in E_1
3	2	E_2	$E_1 \in [890, 1190]$	γ in E_2, $\beta\beta$ in E_1
4	2	E_2	$E_1 \in [1190, 1880]$	γ in E_2, $\beta\beta$ in E_1
5	3	E_2		γ in E_2, $\beta\beta$ in E_1, Compt. in E_3
6	3	$E_2 + E_3$	$E_2 < 500$	γ in $E_2 + E_3$ (shared Compt.), $\beta\beta$ in E_1

are thus likely to be contained in the E_1 detector due to the relatively low γ energies. We hence focus our search on events with γ's in the E_2 and E_3 variables. In particular, this leads to two categories, one where the γ energy is split between E_2, E_3 and another where one γ is fully contained in E_2 with a Compton scatter event in E_3.

Combined with the $\mathcal{M}_2$ categories this leads to a total of seven independent categories used for the $2\nu\beta\beta$ decay analysis.

C. $0\nu\beta\beta$ decay categories

Unlike for $2\nu\beta\beta$ decay, in $0\nu\beta\beta$ decay the electrons have a fixed total energy so we are also able to search for a peak from the summed electron energy. Therefore a different set of categories are necessary; however, the same strategy can be used. For $\mathcal{M}_1$ events we include three categories corresponding to the $\beta\beta$ and possibly γ's being contained in the detector where the decay took place. For $\mathcal{M}_2$ events, we show the two-dimensional distributions in Fig. 4 (bottom two figures). We observe a variety of features, vertical, horizontal diagonal lines and points. Each of these require a different set of energy cuts to contain the signal, as shown

in Table II. These cuts are carefully constructed to ensure that none of the categories overlap. Some categories feature a monoenergetic peak in both energy variables, for these we still choose one as the peak variable and use a range of ± 10 keV for the other, projected energy, much larger than the energy resolution. We employ a similar strategy to the $2\nu\beta\beta$ decay categories, dividing the horizontal lines, events with a $\beta\beta$ peak (1904, 2444, 2494 keV) in E_1 and a Compton scatter in E_2, into three categories each to account for the large changes in background level. This is not necessary for other categories due to the low background, or a small range in the projected-out energy. We neglect some possible categories which have very small containment efficiency (events with $E_2 > 600$ keV), that overlap significantly with $2\nu\beta\beta$ decay (events with $E_2 \approx 500$–650 keV, $E_1 < 1904$ keV) or with ^{214}Bi background, diagonal lines with energy summing to ≈ 2400 keV. This leads to 10 $\mathcal{M}_2$ $0\nu\beta\beta$ decay categories which are most easily interpreted geometrically (see the boxes in Fig. 4).

As in $2\nu\beta\beta$ decay analysis we also consider $\mathcal{M}_3$ events focusing on those where the electrons are contained in E_1. We choose four categories: For events with $E_1 \approx 1904$ keV, we

TABLE II. Categories for $0\nu\beta\beta$ to excited states, for each category we give an index, a name, the projected energy variable, and its energy window, the cuts used to define the category and a physical interpretation. Most of these cuts are more easily visualized in Fig. 4.

Cat.	Multiplicity	Peak variable	Range [keV]	Energy cuts [keV]	Interpretation
0	1	E_1	1864–1944		$\beta\beta$ in E_1, γ escape
1	1	E_1	2384–2534		$\beta\beta$ in E_1, γ escape
2	1	E_1	2984–3084		$\beta\beta$ and γ in E_1
3	2	E_1	1864–1944	$E_2 < 180$	$\beta\beta$ in E_1, Compt. γ in E_2
4	2	E_1	1864–1944	$E_2 \in [180, 310]$	$\beta\beta$ in E_1, Compt. γ in E_2
5	2	E_1	1864–1944	$E_2 \in [310, 500]$	$\beta\beta$ in E_1, Compt. γ in E_2
6	2	E_1	2384–2534	$E_2 < 180$	$\beta\beta$ in E_1, Compt. γ in E_2
7	2	E_1	2384–2534	$E_2 \in [180, 280]$	$\beta\beta$ in E_1, Compt. γ in E_2
8	2	E_1	2384–2534	$E_2 \in [280, 500]$	$\beta\beta$ in E_1, Compt. γ in E_2
9	2	E_2	500–650	$E_1 \in [1894, 2904]$	$\beta\beta$ in E_1, γ in E_2
10	2	E_2	500–650	$E_1 \in [1944, 2384], E_1 + E_2 > 2534$	$\beta\beta$ + Compt. E_1, γ in E_2
11	2	$E_1 + E_2$	2984–3084	$E_1 > 2534, E_2 < 500$	$\beta\beta$ + Compt. E_1, Compt. E_2
12	2	$E_1 - E_2$	1734–2034	$E_2 < 650, E_1 > 2384, E_1 + E_2 \in [3024, 3044]$	$\beta\beta$ in E_1, γ in E_2
13	3	$E_2 + E_3$	500–650	$E_1 \in [1894, 1914], E_2 < 500$	$\beta\beta$ E_1, Compt. E_2/E_3
14	3	E_2	500–650	$E_1 \in [1894, 1914]$	$\beta\beta$ in E_1, γ in E_2, Compt. E_3
15	3	E_1	1854–1954	$E_2 \notin [500, 650], E_2 + E_3 \notin [500, 650]$	$\beta\beta$ in E_1 and Compt. E_2/E_3
16	3	E_1	2384–2534		$\beta\beta$ in E_1

TABLE III. Table showing the γ-exclusion cuts used to remove events likely originating from a background γ ray.

Variable	Cut [keV]	Decay
E_1	1333 ± 5	^{60}Co
E_1	1173 ± 5	^{60}Co
$E_1 + E_2$	1333 ± 5	^{60}Co
$E_1 + E_2$	1173 ± 5	^{60}Co
$E_1 + E_2$	1461 ± 5	^{40}K
E_1	2615 ± 7	^{208}Tl
$E_1 + E_2$	2615 ± 7	^{208}Tl

split the data into $E_2 \in [500, 650]$ keV, $E_2 + E_3 \in [500, 650]$ keV or neither of these two. For events with $E_1 \approx 2400$ keV the background is expected to be low so we just include one category. This procedure leads to 17 experimental signatures, as shown in Table II.

D. Determination of signal shape and efficiency

We use MC simulations to determine the signal shapes and efficiencies for our Bayesian analysis. For each category we make the energy cuts as described in Tables I and II. We also apply a set of selection cuts to the data and MC simulations to remove events that likely arise from a known γ-ray background peak. The cuts are chosen to correspond to roughly $\pm \Delta E$ (FWHM) with the values shown in Table III. After the projection onto the peak energy variable we extract one-dimensional histograms for each decay mode (0_1^+ or 2_1^+) in each category from MC simulations.

From these histograms we obtain the containment efficiencies $\varepsilon_{\text{cont},r}^{0+}$, $\varepsilon_{\text{cont},r}^{2+}$, or the probability for a decay to 0_1^+ or 2_1^+ e.s. to reconstruct in category r. We also compute the total containment efficiency (summed over all categories) as 4.4(2.0)% for $2\nu\beta\beta$ to the $0_1^+ (2_1^+)$ e.s. and 46(57)%, respectively, for $0\nu\beta\beta$ decay.

We fit these histograms to phenomenological functions to extract analytic signal shapes to use in our analysis. These functions are described in more detail in Appendix B, the photopeak is approximated via three Gaussian's to account for the slight non-Gaussianity of the lineshape (due to each detector-dataset pair having a different resolution), this is discussed more in Sec. IV A. Each function is normalized to unity and then we model the MC simulated spectrum as

$$f(E) = N_{\text{tot}} \sum_{i=1}^{N} p_i f_i(E), \qquad (5)$$

where the sum runs over the normalized functions f_i, p_i is the probability for an event to be in a given spectral feature i of the signal distribution for the considered category, for example, in one of the gamma peaks or in the continuum, and N_{tot} is the total number of events. We show two example fits (both 0_1^+ signal) in Fig. 5, $2\nu\beta\beta$ category two and Fig. 6, $0\nu\beta\beta$ category two. We see that these models describe the MC simulations very well, in particular the slightly non-Gaussian photopeak shape. These functions encode most of our systematic uncertainties, in particular the resolution and peak position of Gaussian and normalization of the peaks which can then be treated as nuisance parameters of the final fit.
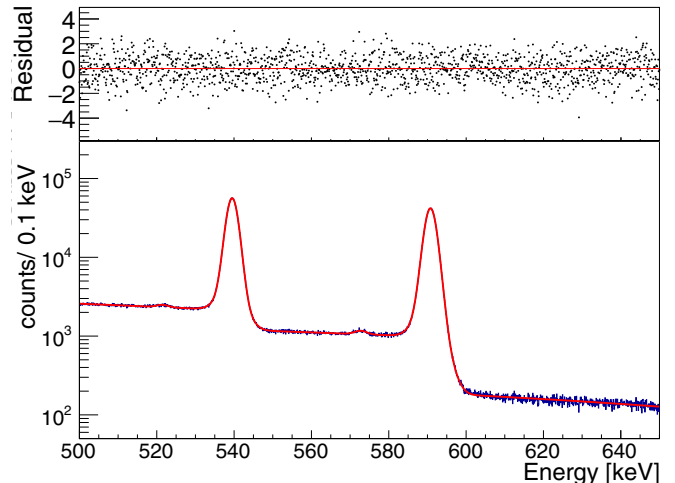

FIG. 5. Parametrization of $2\nu\beta\beta$ to 0_1^+ category two signal. The model consists of two photopeaks, with their respective ^{100}Mo x-ray escape peaks, two smeared step functions modeling Compton interactions and a linear background. The residual defined as (value − model)/error is shown in the upper figure.

IV. DATA ANALYSIS

For this search we use the full CUPID-Mo data corresponding to 2.71 kg yr exposure of LMO. We process our data using a C++ software, DIANA [58,59], developed by the Cuoricino, CUORE, and CUPID-0 Collaborations and further developed by CUPID-Mo. Most of the data processing steps are the same as described in Refs. [31,47]. An optimal trigger is used to identify physics events and an optimal filter which maximizes the signal-to-noise ratio is used to compute amplitudes (both LD and LMO detectors) [60]. Thermal gain stabilization and calibration are performed using data collected with a Th or U calibration source. For each LMO detector we associate "side-channels," as the LDs which face this crystal. For $\mathcal{M}_1$
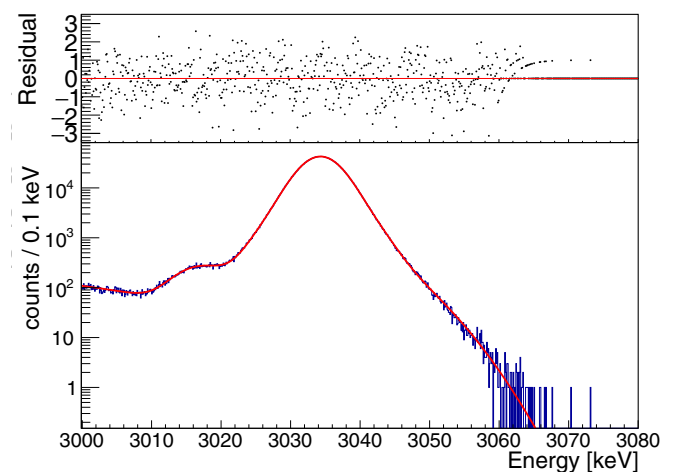

FIG. 6. Parametrization of $0\nu\beta\beta$ decay to 0_1^+ category two signal. The model contains a photopeak, x-ray escape peak, and linear smeared step functions. The residual defined as (value − model)/error is shown in the upper figure.

data (3 $0\nu\beta\beta$ decay categories) we use the same selection cuts as in Ref. [31]: a principal component analysis (PCA) based pulse shape discrimination (PSD) cut [61], a normalized light distance cut, delayed coincidence (DC) cuts to remove ^{222}Rn, ^{228}Th decay chain events and muon veto anticoincidence.

However, for coincidence events ($\mathcal{M} > 1$) it is beneficial to modify several of these steps. Random coincidences between noncausally related events (hereafter *accidentals*) could provide a large possible background. Similar to CUORE [58] and as described in Ref. [31] we adjust the time difference of events to account for the characteristic detector response based time offsets between detector pairs. This allows us to reduce the time window used to define coincidences to 10 ms (from 100 ms in previous analyses [47]) and thus reduce the accidental coincidence background.

To select a clean sample of higher multiplicity e.s. candidate events, we make use of the dual readout and implement a light-yield-based cut for $\mathcal{M} > 1$ events. Unlike the $\mathcal{M}_1$ analysis where this cut primarily rejects α backgrounds, it is designed here to tag and reject events where an energetic electron escapes a LMO crystal, punches through a LD, and is stopped in the adjacent crystal within a tower. In addition, this cut very efficiently tags events from a ^{60}Co contamination identified on one of the LDs where the β particle typically leaves significant excess energy in the LD. Developing a light cut to remove these events is less straightforward than for $\mathcal{M}_1$ data, since a coincidence between two vertically adjacent crystals is accompanied by scintillation light signal that can be absorbed in the common, intermediate LD. For each event we compute the expected scintillation light deposited on the LD based on the values observed in $\mathcal{M}_1$ events (mainly $2\nu\beta\beta$ decays to the g.s.). This accounts for these multiple contributions to the light yield. We then normalize the light energies using a conservative estimate of the LD resolution ($\sigma_{i,s}^{\mathrm{LD}}$) measured in high-energy background data ($\mathcal{M}_2$ total energy >1500 keV) as

$$n_{i,s} = \frac{E_{i,s}^{\mathrm{LD}} - E_{i,s}^{\mathrm{LD,expec}}}{\sigma_{i,s}^{\mathrm{LD}}}, \qquad (6)$$

where i refers to the pulse index (one for E_1 etc.), s is the side channel (either 0 or 1), $E_{i,s}^{\mathrm{LD}}$ is the measured LD energy, while $E_{i,s}^{\mathrm{LD,expec}}$ is the expected LD energy for γ- or β-like event and $\sigma_{i,s}^{\mathrm{LD}}$ is the predicted energy resolution of the LD. We then place a cut of $|n_{i,s}| < 10$ for each LD and side channel in an $\mathcal{M} > 1$ event. We place a cut of $-10 < n_{i,s} < 3$ for the LD with the ^{60}Co contamination. This contamination can also lead to background events in other LMO detectors; to remove this we make a global LD anticoincidence cut. We remove any LMO event (excluding those who directly face this LD) with a trigger on this LD with LD energy >5 keV within a ± 5 ms window.

We also adjust the PCA-based PSD cuts [31,47] to place a cut on the shape of each pulse in a higher multiplicity event. We require that each pulse has a normalized PCA reconstruction error of less than 23 median absolute deviations. This was optimized by comparing the efficiency of $\mathcal{M}_2$ events ε_s, obtained from events summing to the ^{40}K 1461 keV peak and estimating the background efficiency, ε_b, in a sideband

$E_2 \in$ (450–520) keV. This sideband, which is only used in optimizing the PCA cut, approximates the background for the dominant $2\nu\beta\beta$ decay categories (2, 3, and 4) while not including any signal peaks. We then maximize a figure of merit $\varepsilon_s/\sqrt{\varepsilon_b}$ which is proportional to the experimental sensitivity. For $\mathcal{M}_1$ data a cut of <9 on the PCA normalized reconstruction error is used (as in Ref. [31]).

Finally, we employ a data blinding by adding simulated MC events directly to the data files. The rate is chosen by sampling randomly from a uniform distribution with range, $\Gamma \in \Gamma_{\mathrm{measured}} \times (2, 10)$, where $\Gamma_{\mathrm{measured}}$ is the current leading limit or measurement [2] and (2,10), is a uniform distribution between 2 and 10 chosen to ensure the injected signal is significantly larger than any possible signal in the data. The injected rate is hidden during the analysis of the blinded data. The blinded data are used to optimize and test the Bayesian fitting routine and prevent biasing our results.

A. Energy resolution and linearity

We determine the response of our detector to a monochromatic energy deposit using both calibration sources and γ peaks from natural radioactivity using the same procedure described in detail in Ref. [31]. In particular, a simultaneous fit, over each detector in each dataset, of the 2615 keV peak in the calibration data is used to extract the resolution of each detector-dataset pair at 2615 keV $\sigma_{\mathrm{chan,dataset}}^{2615}$.

We model the lineshape of our peaks in $\mathcal{M}_1$ physics data as an exposure-weighted sum of Gaussians for each channel-dataset pair. The individual Gaussians have a common mean μ but a resolution which is a product of $\sigma_{\mathrm{chan,dataset}}^{2615}$ and a global scaling $R(E)$. We fit this lineshape model to our peaks to extract $R(E)$ and $\mu(E)$ for each peak in physics data and parametrize the energy dependence as $R(E) = (r_0^2 + r_1 E)^{1/2}$ and $\mu(E) = b_0 + b_1 E + b_2 E^2$. The resulting parametrization is used in the MC simulations to model the detector response. In particular, we account for the uncertainty on the parameters of these functions as systematic uncertainties (Sec. VI).

B. Cut efficiencies

To measure $2\nu\beta\beta$ and $0\nu\beta\beta$ decay rates it is necessary to determine the analysis efficiency, or probability that a signal event will pass all selection cuts. We employ the same strategy that was utilized in the $\mathcal{M}_1$ $0\nu\beta\beta$ analysis [31] in order to compute these. We evaluate the pileup efficiency, or probability that a signal event will not have another trigger in the same waveform using random triggers. Several other cuts (multiplicity selection, muon veto cut, and LD anticoincidence) induce dead times in one or more detectors. For these we evaluate the efficiencies using ^{210}Po Q-value peak events by counting the number of events passing the cuts with energy in (5407 ± 50) keV. These are a proxy for physical $\mathcal{M}_1$ events due to the high-energy and modular structure, meaning that α particles can only deposit energy in one crystal.

Finally, for our pulse shape and light yield cuts the efficiency is evaluated using γ peaks in $\mathcal{M}_1$ and summed $\mathcal{M}_2$ energy which are a clean sample of signal-like events. For each prominent γ peak, we fit the energy distribution of events passing and failing each cut to a Gaussian plus linear

TABLE IV. Efficiency for analysis cuts, the energy independent efficiencies are computed assuming the cut efficiency does not depend on energy. For the extrapolated efficiency in $\mathcal{M}_1$ data at 3034 keV, we fit the efficiency as a function of energy to a first-order polynomial and extrapolate this to 3034 keV, this is only relevant for the $\mathcal{M}_1$ signatures with 3034 keV energy.

Cut	Efficiency energy independent [%]	Efficiency extrapolated [%]	Method of evaluation
Pileup	95.7 ± 1.0		Noise events
Multiplicity	99.55 ± 0.07		^{210}Po
LD anti-coincidence	99.98 ± 0.02		^{210}Po
Delayed coincidence	99.16 ± 0.01		^{210}Po
Muon veto	99.62 ± 0.07		^{210}Po
Light cut ($\mathcal{M}_1$)	99.4 ± 0.4	99.7 ± 0.8	Gamma peaks ($\mathcal{M}_1$)
Light cut $\mathcal{M} > 1$	97.7 ± 1.8		Gamma peaks ($\mathcal{M}_2$)
PCA <23 ($\mathcal{M}_1$)	99.2 ± 0.3		Gamma peaks ($\mathcal{M}_1$)
PCA <9 ($\mathcal{M}1$)	95.2 ± 0.5	94.3 ± 1.5	Gamma peaks ($\mathcal{M}_1$)

background. From the measured number of events N_{pass}, N_{fail} (events in the Gaussian peak only for the two fits) we compute the efficiency $\varepsilon = N_{\text{pass}}/(N_{\text{pass}} + N_{\text{fail}})$. We estimate numerically the uncertainty on ε by sampling from the measured uncertainties on N_{pass}, N_{fail}. We observe an energy independent efficiency for all of our cuts (shown in column 2 of Table IV). The energy range for the $\mathcal{M}_1$ 3034 keV category of $0\nu\beta\beta$ decay lies outside the range of our γ peaks. Similar to Ref. [31] we use a first-order polynomial to extrapolate the PCA and light yield cut efficiency to 3034 keV, accounting for the possibility of a slight energy dependence of the cut (shown in column 3 of Table IV). We summarize the cut efficiencies in Table IV and use them as nuisance parameters with Gaussian priors in our Bayesian analysis.

V. BAYESIAN ANALYSIS

We perform a Bayesian analysis to extract decay rates. In particular, we use an extended unbinned maximum likelihood fit implemented using the Markov chain Monte Carlo (MCMC) of the Bayesian analysis toolkit (BAT) [62] for all categories except for two $0\nu\beta\beta$ decay categories, for which we use a binned fit due to their exceptionally high statistics. We use a fit of the summed data of all 19 detectors and 9 datasets.

Two separate fits are created, one to extract the decay rates of $2\nu\beta\beta$ decay and another for the $0\nu\beta\beta$ decay. Both fits are implemented in the same framework. In each category (index r) we model our experimental data as

$$f_r(E) = \varepsilon_{a,r} n_{2^+} f_{2^+,r}(E) + \varepsilon_{a,r} n_{0^+} f_{0^+,r}(E)$$
$$+ n_{r,\text{bkg}} f_{r,\text{bkg}}(E) + \sum_{p=1}^{5} n_{r,p} f_p(E), \quad (7)$$

where

(i) n_{0^+}, n_{2^+} are the observed number of counts to 0_1^+, 2_1^+ excited states;

(ii) $f_{0^+,r}(E)$, $f_{2^+,r}(E)$ are the phenomenological functions describing the signal shape of category r, normalized to unity over the sum of all categories;

(iii) $\varepsilon_{a,r}$ is the analysis efficiency for category r (described in Sec. VI);

(iv) $f_{r,\text{bkg}}(E)$ is a function describing the background in category r, either exponential or flat for categories with low statistics;

(v) $n_{r,p}$ are the number of background counts from known γ lines (a subset of 511, 583, 609, 2448, and 2505 keV) with index p and $f_p(E)$ is the model of this spectral shape, a single Gaussian.

We show in Fig. 7 the fit of the $2\nu\beta\beta$ decay category three and the contribution from the exponential background, signal, and background peaks. Our likelihood function is

$$\ln(\mathcal{L}) = \sum_{n=1}^{N} \ln[f_{r(n)}(E_n)]$$
$$+ \sum_{r=1}^{C} -\lambda_r + n_r \ln(\lambda_r) - \ln(n_r!)$$
$$+ \sum_{b=1}^{B} \sum_{i=1}^{N_{\text{bin}}} -\lambda_i + n_{i,b} \ln(\lambda_{i,b}) - \ln(n_{i,b}!), \quad (8)$$
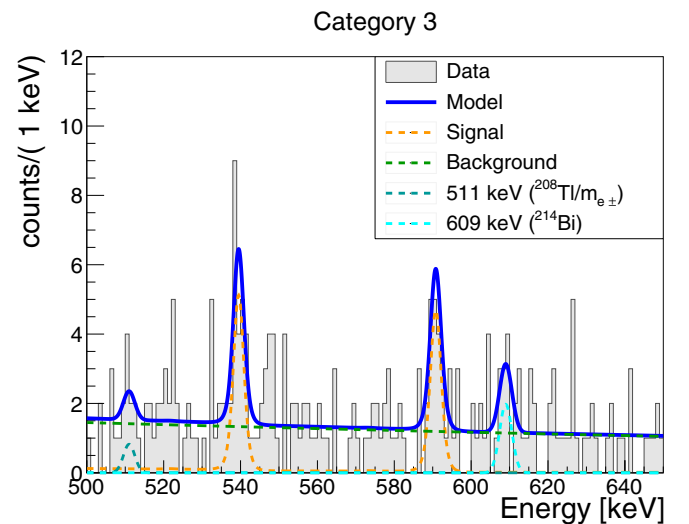

FIG. 7. Fit of the $2\nu\beta\beta$ decay category three. The experimental data are binned for visualization and we show the best-fit reproduction of our model and the contributions from the signal peaks, background peaks, and the exponential background.

where the first sum is over the events in the experimental data (unbinned categories), $r(n)$ is the category of event n, C is the number of categories, λ_r is the total predicted number of events in category r, and n_r is the observed number of events. The last sums are a binned likelihood where the sum b runs over the B categories which use a binned likelihood. These are the categories 0 and 1 for the $0\nu\beta\beta$ decay where a large number of events make an unbinned fit computationally unfeasible. λ_i is the expectation value for bin i and $n_{i,b}$ is the number of events in experimental data. We use 0.5 keV bins for these fits, much smaller than the energy resolution over the full spectral range.

We use BAT to sample the full posterior distribution of the parameters of interest $\vec{\theta} = (n_{0_1^+}, n_{2_1^+})$ and nuisance parameters $\vec{v}$, which include all of the background components (both the exponential or flat background and the γ peaks) and also parameters of systematic uncertainties (see Sec. VI):

$$p(\vec{\theta}, \vec{v}|\mathcal{D}) = \frac{P(\mathcal{D}|\vec{\theta}, \vec{v})P(\vec{\theta}, \vec{v})}{P(\mathcal{D})}, \quad (9)$$

where $\mathcal{D}$ represents the data. We use uninformative (flat) priors on all background model components and parameters of interest n_{0+}, n_{2+}. We define observables for the decay rates to e.s. s, (where s is 0^+ or 2^+):

$$\Gamma_s = \frac{n_s W}{\varepsilon_{\text{cont},s} Mt \times 1000 N_a}, \quad (10)$$

where W is the molecular mass for enriched $\text{Li}_2\,^{100}\text{Mo}\,\text{O}_4$, N_a is the Avogadro constant, Mt is the exposure (in kg yr), η is the isotopic abundance of ^{100}Mo and $\varepsilon_{\text{cont},s}$ is the containment efficiency for the decay to e.s. s, i.e., the total probability a simulated event is in one of the categories. The half-life is then given by $\ln(2)/\Gamma$. For each step of the Markov chains the values of Γ_{0+}, Γ_{2+} are stored and used to compute the marginalized posterior distributions. We include all of our systematic uncertainties as nuisance parameters in our analysis as described in Sec. VI.

VI. SYSTEMATIC UNCERTAINTIES

A. Energy resolution and bias

We account for the uncertainty in the energy resolution of our peaks. Both the terms r_1 and r_0 in the resolution scaling function [$R(E)$, see Sec. IV A] are given Gaussian priors based on the measured uncertainty from the fit to γ-peak resolution. A multivariate prior is not necessary since the correlation is very low (-0.2%) because the r_0 term is fixed very well by noise events (random triggers used to estimate energy resolution at 0 keV). At each stage in the Markov chain this resolution is applied to all the signal components and background peaks, by adjusting the resolution of the functions $f_{0+,r}(E)$, $f_{2+,r}(E)$, $f_{r,\text{bkg}}(E)$.

We also account for the uncertainty in the energy scale. Our energy bias is parametrized as a second-order polynomial as described earlier in Sec. IV A. We adjust the mean position of all peaks in our signal and background template functions ($f_{0+,r}(E)$, $f_{2+,r}(E)$, f_{bkg}) by this bias. We add a nuisance parameter to the model accounting for the

uncertainty on the bias, which is given a Gaussian prior with 0.07 keV σ, and vary the position of each peak (in a correlated way) by this amount. For the $0\nu\beta\beta$ decay analysis the peaks cover a wide range in energy and we take a conservative approach varying the position of all the peaks by the largest uncertainty of 0.3 keV (from 3034 keV).

B. Analysis efficiency

We account for the uncertainty in the analysis efficiency, or the probability that a signal like event would pass all cuts. These cut efficiencies are included as nuisance parameters in our model which are given Gaussian priors based on the estimated uncertainty (see Sec. IV B). A nuisance parameter is included for each cut, including two separate parameters for the constant and extrapolated light cut, and the PCA < 9 cut. We then predict the efficiency in category r as

$$\varepsilon_r = \prod_{c=1}^{11} \varepsilon_c^{p(\mathcal{M}(r),c)}, \quad (11)$$

where the product c runs over all cuts (both energy independent and extrapolated, see Table IV) and $p(\mathcal{M}(r), c)$ is a power which represents how many times a cut was applied to an event. For example, for PSD the cut is applied for both pulses so the efficiency is raised to the power of the multiplicity.

C. Containment efficiency

The final systematic uncertainty we account for is the containment efficiency from MC simulations. In particular, this is related to the accuracy of our GEANT4 MC simulations which can broadly be divided into two parts: the accuracy of the simulated geometry and the implemented physics process. For the simulated geometry we vary the amounts of the three main materials in the experimental setup (see Fig. 2):

(i) the dimensions of the LMO crystals (and therefore density since the mass is very well known);
(ii) the thickness of the copper holders;
(iii) the thickness of the Ge LDs.

We run a set of simulations varying the LMO dimensions [by ± 100 (200) μm or around $\pm 0.2\%$ (0.4%)], the Cu holder thickness [by ± 100 (200) μm or $\pm 2.5\%$ (5%)] and the Ge LD thickness [by ± 10 (20) μm or around $\pm 6\%$ (12%)]. We also account for the possible inaccuracy of GEANT4 physics cross sections by running MC simulations varying both the Compton effect and photoeffect cross sections by $\pm 5\%$ which is a conservative estimate of GEANT4 cross-section accuracy [63,64]. For each systematic effect (Cu thickness, LMO dimension, Ge thickness, Compton or photoeffect cross-section scale) we introduce a parameter v_i. For every effect, decay, category (index r) and peak (index p) we parametrize the variation in peak containment efficiency. This is the product of the overall containment efficiency and the Gaussian peak probability (see Sec. III D). We use first-order polynomial fits to obtain $d\varepsilon_{\text{cont},r,p}/dv_i$. We show the results of all these tests in Appendix C. We observe that the most

significant effects are the photoeffect cross section (higher scale tends to increase containment) and the Compton cross section (higher tends to decrease containment). The geometrical uncertainties provide a much smaller effect. Effects where the variation in containment efficiency is consistent with zero (within 2σ) are not included. Based on the parameters $\vec{v}$ we predict the efficiency for category r and peak p as

$$\varepsilon_{\text{cont},r,p} = \varepsilon^0_{\text{cont},r,p} + \sum_{c=1}^{5} v_c d\varepsilon_{\text{cont},r,p}/d(v_c), \qquad (12)$$

where the sum c runs over systematic effects, and $\varepsilon^0_{\text{cont},r,p}$ is the default efficiency (for the MC simulation without varying any parameters). In our fit, the nuisance parameters accounting for uncertainties of copper thickness and LMO dimension are given Gaussian priors with a standard deviation of 100 μm; for the Ge thickness we use 20 μm, and we use 5% for the Compton and photoeffect physics cross sections. At every step of the Markov chain the amplitude of each peak is adjusted according to these parameters. The normalization of the fit functions are adjusted accordingly and we recompute the containment efficiency for both 0^+_1 and 2^+_1 signals.

D. Choice of $2\nu\beta\beta$ decay to 0^+_1 state model

The final systematic uncertainty we consider is the $2\nu\beta\beta$ to 0^+_1 state decay model. By default we use the SSD model; however, we also run MC simulations and a fit using the HSD model. We then consider the difference between these two results as a systematic on the $2\nu\beta\beta$ decay to 0^+_1 state decay rate.

VII. RESULTS

A. Toy Monte Carlo sensitivity and bias

To validate our fitting routine and predict the median exclusion sensitivity we use an ensemble of pseudo-experiments (toy MC). We generate 2000 toy datasets for both $0\nu\beta\beta$ and $2\nu\beta\beta$ decay analysis based on the background parameters from the fit to data (see Secs. VII B 1 and VII B 2). We assume the $2\nu\beta\beta \rightarrow 0^+_1$ signal from the best fit to data and zero signal of $2\nu\beta\beta \rightarrow 2^+_1$ or $0\nu\beta\beta$ decay modes. We sample from a Poisson distribution for the number of counts of each component in each toy experiment. We fit each dataset to determine the distribution of possible limits and therefore the median sensitivity for each decay. This is shown in Figs. 8 and 9. We extract median exclusion sensitivities $\hat{T}$ at 90% credibility interval (c.i.) for the three decays which have not yet been measured of

$$\hat{T}^{2\nu \rightarrow 2^+_1}_{1/2} = 3.8 \times 10^{21} \text{ yr}, \qquad (13)$$

$$\hat{T}^{0\nu \rightarrow 2^+_1}_{1/2} = 3.1 \times 10^{23} \text{ yr}, \qquad (14)$$

$$\hat{T}^{0\nu \rightarrow 0^+_1}_{1/2} = 9.8 \times 10^{22} \text{ yr}. \qquad (15)$$

We also use these toys to verify the fitting procedure is not biased. We fit our toy datasets for $2\nu\beta\beta$ decay analysis with both a model including the 2^+_1 signal called $H_0 = H(0^+_1 + 2^+_1 + B)$ and one including only the 0^+_1 signal, $H_1 = H(0^+_1 +$
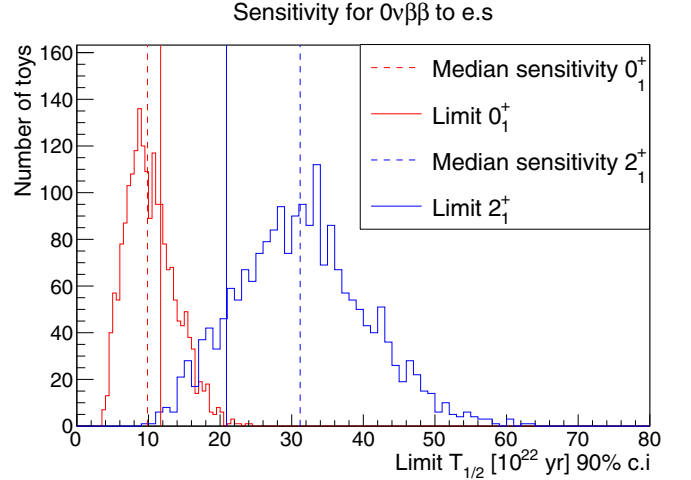

FIG. 8. Sensitivity for $0\nu\beta\beta$ decay to both 0^+_1 and 2^+_1 excited states obtained from an ensemble of 2000 toy experiments. We show the median sensitivity (dashed line) and the limit obtained on real data (solid lines). In both cases the limits on real data are consistent with those expected from toy experiments.

B). We compute

$$r_i = \frac{n^i_{0^+} - \bar{n}_{0^+}}{\sigma(n_{0^+})}, \qquad (16)$$

where $\bar{n}_{0^+}$ is the marginalized mode from the fit to real data used to generate the toys, $n^i_{0^+}$ is the marginalized mode in toy i, and $\sigma(n_{0^+})$ is the estimated error (central 68% c.i.). The distribution of r_i for both models is shown in Fig. 10. We observe a clear bias in r_i for the H_0 model with a mean of -0.58σ, while r_i is unbiased for the H_1 model (mean of -0.005σ). Hence, we decide to use the H_1 model for the case that no evidence of a decay to 2^+_1 state is found. This bias is due to non-negative Bayesian priors on the rate of 2^+_1 decay
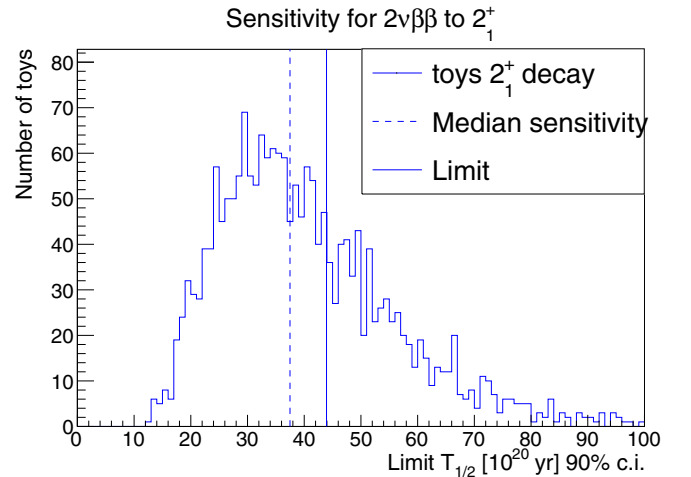

FIG. 9. Sensitivity for $2\nu\beta\beta$ decay to 2^+_1 excited state obtained from an ensemble of 2000 toy experiments. As with $0\nu\beta\beta$ decay we show the median sensitivity (dashed line) and the limit obtained on real data (solid line). Again the limit on real data is consistent with the toy experiments.
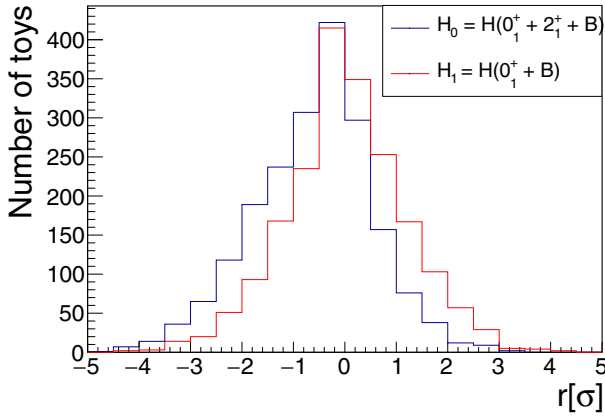
FIG. 10. The distribution of r_i [see Eq. (16)] for both H_0 and H_1 models for our fits to toy data. We observe a clear bias for the H_0 model while the H_1 model is unbiased.

which allow the rate of 2_1^+ to fluctuate only up. Since the 0_1^+ and 2_1^+ rates are anticorrelated this causes the 0_1^+ rate to be biased.

B. Fit of $2\nu\beta\beta$ with $H_0 = H(0_1^+ 2_1^+ + B)$ model

We run the $2\nu\beta\beta$ decay fit with both $2\nu\beta\beta$ decay to 2_1^+ and 0_1^+ e.s. signals and background components, i.e., the H_0 model. We can use this fit to determine if we observe evidence for the decay to 2_1^+ state and to set a limit if no evidence is found. We consider first a fit including all possible background peaks, the *maximal model*. We then repeat the fit removing any background peak where the central 68% confidence interval contains zero counts and replacing the exponential background with a constant if the slope is compatible with zero (within 1σ). We call this *the minimal model* and we use this model for statistical inference. We show the marginalized posterior distribution of the decay rate Γ for $2\nu\beta\beta$ decay to 2_1^+ state (including all systematics) in Fig. 11
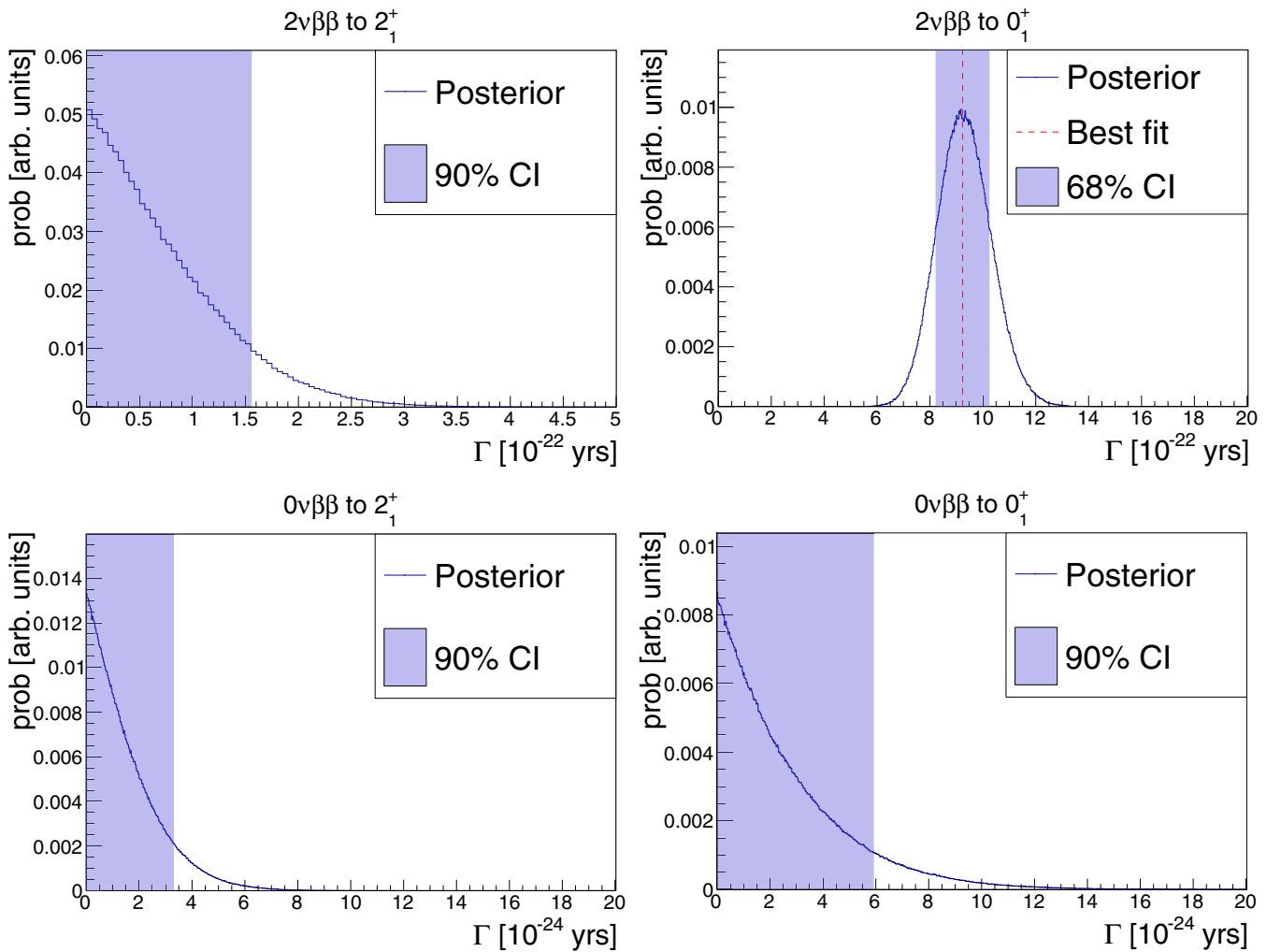
FIG. 11. Posterior probability distributions for the decay rates, $\Gamma_{2\nu}$ to 2_1^+ state (top left) and 0_1^+ state (top right). The mode for 2_1^+ decay is at zero rate and so we find no evidence for this decay and we extract the measurement of 0_1^+ mode using the $H(0_1^+ + B)$ model. Bottom left and right are the posteriors of $\Gamma_{0\nu}$ to the 2_1^+ and 0_1^+ states, respectively, again showing no evidence of the decay.

(left). We see that the mode is at zero rate and therefore find no evidence for $2\nu\beta\beta$ to 2_1^+ state. We correspondingly set a limit (including all systematics) of

$$T_{1/2}^{2\nu\to2_1^+} > 4.4 \times 10^{21} \text{ yr (90\% c.i.).} \quad (17)$$

This is the most stringent constraint on this process in ^{100}Mo, an improvement of $\approx 80\%$ over the previous constraint, 2.5×10^{21} yr [42].

1. Fit of $2\nu\beta\beta$ with $H_1 = H(0_1^+ + B)$ model

Since we find no evidence of the $2\nu\beta\beta$ decay to 2_1^+ state and toy experiments show that including this parameter in our fit would bias our measurement, we run the $2\nu\beta\beta$ decay $H_1 = H(0_1^+ + B)$ fit without the 2_1^+ decay contribution. We find that this model is able to describe our data in all seven categories very well (see Fig. 7 for an example and all fits in Appendix D). We verify the goodness of fit using our ensemble of toy experiments. For each fit we extract the *global mode* or the set of parameters $\vec{\theta}$ with the maximum probability. We use the posterior probability of these parameters:

$$k = p(\vec{\theta} \,|\mathcal{D}), \quad (18)$$

as a test statistic to quantify the quality of the fit by comparing the probability distribution obtained in toy experiments $p(k')$ to the fit on real data. We extract the p value:

$$p = \int_0^k p(k')dk' = 0.49, \quad (19)$$

indicating that the model describes the data well.

We extract the posterior distribution on $\Gamma_{2\nu\to0_1^+}$ from the fit on the data and we extract the central 68.3% c.i. on the decay rate as

$$\Gamma_{2\nu\to0_1^+} = (9.2 \pm 1.0 \text{ (stat. + syst.)}) \times 10^{-22} \text{ yr}^{-1}. \quad (20)$$

As the systematics are allowed to float freely in our fit this error is the sum of the statistical and systematic components. To evaluate the statistical error, we repeat the fit fixing all nuisance parameters connected to systematics, obtaining

$$\Gamma_{2\nu\to0_1^+} = (9.1 \pm 1.0 \text{ (stat.)}) \times 10^{-22} \text{ yr}^{-1}. \quad (21)$$

A difference in the error compared with the statistical plus systematic fit is only observed after the first digit. We also run a fit using the HSD model, which leads to

$$\Gamma_{2\nu\to0_1^+}^{\text{HSD}} = (8.9 \pm 1.0 \text{ (stat. + syst.)}) \times 10^{-22} \text{ yr}^{-1}, \quad (22)$$

a 4% decrease in the decay rate. Under the assumption that statistical and systematic errors add in quadrature we obtain

$$\Gamma_{2\nu\to0_1^+} = \left(9.2 \pm 1.0 \text{ (stat.)} \, {}^{+0.3}_{-0.5} \text{ (syst.)}\right) \times 10^{-22} \text{ yr}^{-1}, \quad (23)$$

or converting to half-life:

$$T_{1/2}^{2\nu\to0_1^+} = \left(7.5 \pm 0.8 \text{ (stat.)} \, {}^{+0.4}_{-0.3} \text{ (syst.)}\right) \times 10^{20} \text{ yr.} \quad (24)$$

This is a new independent measurement of this decay, with total uncertainty consistent with the previous leading measurement [42].

2. $0\nu\beta\beta$ decay fit

From the fit for $0\nu\beta\beta$ decay we find no evidence of decay to either the 0_1^+ or 2_1^+ excited states. The best fit reproductions of the experimental data are shown in Appendix D. The posterior distributions of $\Gamma_{0\nu\to0+}, \Gamma_{0\nu\to2+}$ are shown in Fig. 11 (bottom left and right), this leads to limits (both at 90% c.i. including all systematic uncertainties) of

$$T_{1/2}^{0\nu\to2_1^+} > 2.1 \times 10^{23} \text{ yr (90\% c.i.),} \quad (25)$$

$$T_{1/2}^{0\nu\to0_1^+} > 1.2 \times 10^{23} \text{ yr (90\% c.i.).} \quad (26)$$

These are new leading limits on these processes, a factor of 1.3 stronger than previous limits from [39] in both cases, despite a factor of ≈ 6 lower exposure of ^{100}Mo.

3. $2\nu\beta\beta$ decay spectral shape

The CUPID-Mo source equals detector geometry also allows us to investigate the $2\nu\beta\beta \to 0_1^+$ spectral shape. This provides a concrete demonstration of a method for this analysis which can be applied to future experiments with large statistics. Our analysis described in Sec. III does not reconstruct the $\beta\beta$ spectral shape directly, so we develop a procedure to extract this shape. For simplicity this analysis is only applied to $\mathcal{M}_2$ data which is the main contribution to the sensitivity.

We select all events that have either E_1 or E_2 in the range $[539.5 - 1.75, 539.5 + 1.75]$ keV and $[590.8 - 1.75, 590.8 + 1.75]$ keV. This choice of the interval is the FWHM energy resolution, roughly optimal for maximizing the sensitivity $S/\sqrt{B}$. We note there is some ambiguity in this selection when $E_1 \approx E_2$ (when the $\beta\beta$ has the same energy as the γ) however the excellent CUPID-Mo energy resolution makes this negligible. We then define $E_{\beta\beta,540}, E_{\beta\beta,591}$ as the energy which does not contain the peak for both 539.5 and 590.8 keV peaks respectively and we construct histograms of these energies. This procedure is applied to both the data, signal MC simulations, and to the preliminary background model reconstruction of the data.

To compare the SSD and HSD models we use a simultaneous binned Bayesian likelihood fit of the two $\beta\beta$ spectra. We use three components; background, 0_1^+ signal and 2_1^+ signal and 50 keV bins. Using the SSD model, this fit reconstructs the half-life as

$$T_{1/2}^{2\nu\to0_1^+} = (7.3 \pm 1.0 \text{ (stat.)}) \times 10^{20} \text{ yr,} \quad (27)$$

$$T_{1/2}^{2\nu\to2_1^+} > 3.6 \times 10^{21} \text{ yr (90\% c.i.).} \quad (28)$$

These are consistent with the values from Secs. VII B and VII B 1 and should be considered a cross-check of this more robust analysis which does not depend on the quality of the background model. We show this fit reproduction in Fig. 12. Repeating the fit using MC simulations obtained with the HSD model, we extract the evidence for both models $p(\mathcal{D}|\text{HSD})$ and $p(\mathcal{D}|\text{SSD})$ and therefore probabilities for the HSD or SSD models of

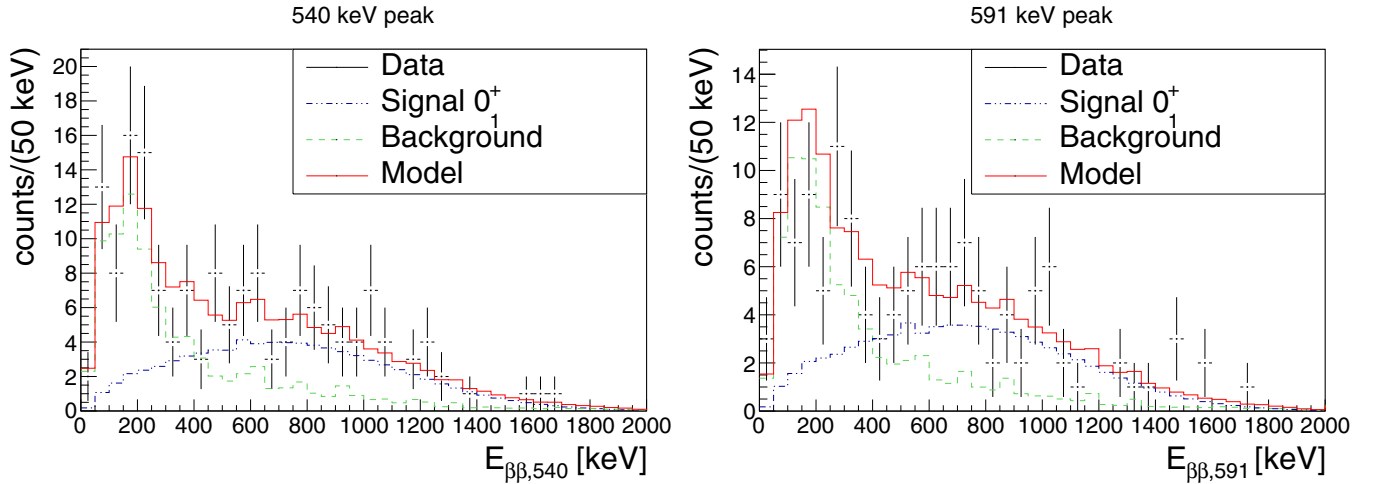$$p(\text{SSD}|\mathcal{D}) = 0.51, \quad p(\text{HSD}|\mathcal{D}) = 0.49. \quad (29)$$

FIG. 12. Plots showing the $2\nu\beta\beta$ decay to e.s. $\beta\beta$ spectral shape for 539.5 keV peak (left) and 590.8 keV peak (right). We overlay the best-fit prediction. We find that in the region >600 keV, the spectrum is dominated by $2\nu\beta\beta \rightarrow 0_1^+$.

This indicates that the CUPID-Mo data is not able to differentiate between the two models. However, this analysis provides a method which could be used to differentiate these two models in a future experiment with larger statistics.

VIII. DISCUSSION

A. Matrix element for $2\nu\beta\beta$ decay to 0_1^+

From our measurement of $2\nu\beta\beta \rightarrow 0_1^+$ excited state we can measure experimentally the nuclear matrix element for this process based on

$$\Gamma_{2\nu} / \ln(2) = G_{2\nu} \left| M_{2\nu}^{\text{eff}} \right|^2. \tag{30}$$

The phase-space factor is given by $(60.55\text{–}65.18) \times 10^{-21}$ 1/yr depending on whether the HSD or SSD model is used [65]. Therefore the matrix element (assuming the SSD model) is given by

$$M_{2\nu \rightarrow 0_1^+}^{\text{eff}} = \sqrt{\frac{\Gamma}{[\ln(2)G]}} \pm \frac{1}{2\sqrt{\ln(2)G}} \frac{\Delta\Gamma}{\sqrt{\Gamma}}$$

$$= 0.143 \pm 0.008 \text{ (stat. + syst.).} \tag{31}$$

We compare the theoretical predictions of dimensionless $M_{2\nu}^{\text{eff}} = g_A^2 M_{2\nu}$ calculated assuming an unquenched value of $g_A = 1.27$. These are 0.395, 0.595, 0.185 for the shell model [66], microscopic interacting boson model [27], and quasiparticle random-phase approximation [67], respectively. This shows the decay rate is reduced relative to theoretical predictions.

B. Effective Majorana neutrino mass $\langle m_{\beta\beta} \rangle$

The limits on $0\nu\beta\beta \rightarrow 0_1^+$, despite the lower phase space, can be used to set a limit on $m_{\beta\beta}$. We use the phase-space factor $(3.162 \times 10^{-15}$ yr^{-1} from Ref. [65]) and the NMEs from Refs. [27,66] to obtain

$$m_{\beta\beta} < (11\text{–}15) \text{ eV,} \tag{32}$$

depending on the NMEs used. While this is still several orders of magnitude above the constrains from the g.s. decay this can be improved significantly in future for a large experiment with minimized dead material.

C. Bosonic neutrinos

Double beta decays to $J \neq 0$ excited states can be sensitive to a bosonic contribution to the neutrino wave function [33]. In particular, under the SSD hypothesis, the predicted half-lives for $2\nu\beta\beta$ to the 2_1^+ e.s. are 2.4×10^{22}, 1.7×10^{23} yr for bosonic and fermionic neutrinos, respectively. The current limit of 4.4×10^{21} yr is still below these predictions.

D. SSD vs HSD models

Both NEMO-3 and CUPID-Mo have demonstrated that SSD hypothesis describes the experimental data very well for the $2\nu\beta\beta$ to $0_{\text{g.s.}}^+$ decay in ^{100}Mo [56,57]. However, for the decay to 0_1^+ state the mechanism (SSD or HSD approximation) is not known. CUPID-Mo is only the second experiment to reconstruct the $2\nu\beta\beta$ to 0_1^+ $\beta\beta$ spectral shape. However, the present level of statistics is insufficient to distinguish the two modes. The excellent energy resolution and low background rates mean this will be possible in a future experiment. In this paper, we demonstrate a method to quantify numerically this and obtained the first result on the compatibility of these two models with data.

IX. CONCLUSION

In this paper we have presented a new analysis of $0\nu\beta\beta$ and $2\nu\beta\beta$ transitions of ^{100}Mo to the first two (2_1^+ and 0_1^+) excited states of ^{100}Ru using the full exposure of CUPID-Mo. This analysis exploits the information available for a source equals detector experiment, where both β and γ's can be measured. A measurement of ^{100}Mo $T_{1/2}^{2\nu \rightarrow 0_1^+}$ was obtained:

$$T_{1/2}^{2\nu \rightarrow 0_1^+} = \left(7.5 \pm 0.8 \text{ (stat.)} {}_{-0.3}^{+0.4} \text{ (syst.)}\right) \times 10^{20} \text{ yr.} \tag{33}$$
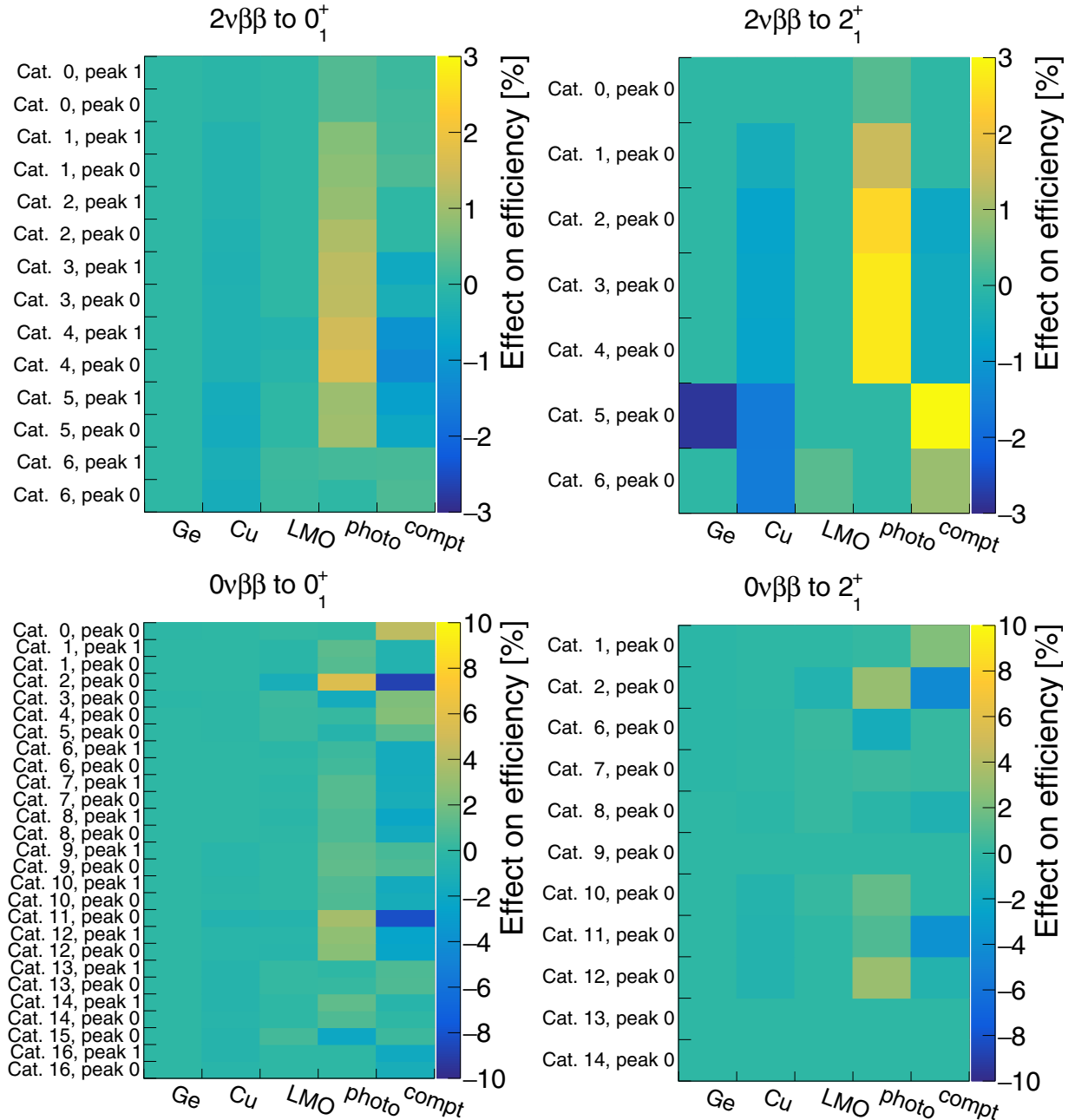
FIG. 13. Plots showing the systematic uncertainty on containment efficiency for each peak, category, decay mode, and systematic effect. The z-axis scale is the percentage effect. Top left shows the $2\nu\beta\beta$ decay to the 0_1^+ state and the right shows decay to the 2_1^+ state. Bottom left show the $0\nu\beta\beta$ decay to the 0_1^+ state, while the bottom right shows the $0\nu\beta\beta$ decay to the 2_1^+ state.

For the other three decay modes, no evidence was found and we extract the limits:

$$T_{1/2}^{2\nu \to 2_1^+} > 4.4 \times 10^{21} \text{ yr (90\% c.i.)}, \tag{34}$$

$$T_{1/2}^{0\nu \to 0_1^+} > 1.2 \times 10^{23} \text{ yr (90\% c.i.)}, \tag{35}$$

$$T_{1/2}^{0\nu \to 2_1^+} > 2.1 \times 10^{23} \text{ yr (90\% c.i.)}. \tag{36}$$

These are the leading limits on these processes. The sensitivity was limited by the small size of the array and large amount of dead material which results in a low containment efficiency.

Future experiments such as CUPID [13] or CROSS [68] will feature a more tightly packed array, much lower amounts of dead material, and a much larger exposure. This will lead to a significantly improved sensitivity with a precision measurement of $2\nu\beta\beta$ decay to 0_1^+ state, distinction between the SSD and HSD mechanisms of the decay and the possibility to measure the $2\nu\beta\beta$ decay to 2_1^+ state.

ACKNOWLEDGMENTS

APPENDIX A: OPTIMIZATION OF CATEGORIES

As explained in Sec. III, some categories are divided up by their projected out energies. In particular, for the $2\nu\beta\beta$ decay analysis the difference in change of spin and Q values, $Q_{2\beta,0_1^+} = 1904.1$ keV and $Q_{2\beta,2_1^+} = 2494.9$ keV, leads to an expected difference in the $E_{\beta\beta}$ spectral shape. This can be used to help reduce correlation between the decays to these two states.

We use a preliminary background model fit to optimize these choices. We emphasize that this fit is only used for the optimization of the choice of categories and not in the Bayesian analysis. We first optimize a choice of $\beta\beta$ energies to minimize the expected measurement error ($\sqrt{S+B}/S$) for $2\nu\beta\beta$ to 0_1^+ e.s. decay. This optimization is performed on the 539.5 keV peak, however a consistent result is also obtained with 591 keV. A separate optimization is performed for the vertical line where the peak is in E_2 and a horizontal where the peak is in E_1 (see Fig. 4). This leads to accepting only events with $E_2 > 220$ keV in the first case, and $E_1 < 1900$ keV in the second.

Next we optimize the choice of energies, E_a, such that the beta energy with a γ in E_1 (horizontal line) is divided into two slices separated by E_a:

$$E_2 \in [220\,\text{keV} - E_a],\ [E_a - 500\,\text{keV}]. \quad (A1)$$

Similarly for the events with the γ in E_2 (vertical line) we divide into slices separated by the energies E_b, E_c:

$$E_1 \in [500\,\text{keV} - E_b],\ [E_b - E_c],\ [E_c - 1900\,\text{keV}]. \quad (A2)$$

This optimization maximizes the limit setting sensitivity for $2\nu\beta\beta$ to 2_1^+ e.s. We estimate this as $S/\sqrt{B}$ using the 539.5 keV peak, this is a proxy for the correlation between the two decays. These optimizations lead to $E_a, E_b, E_c = 410, 890, 1190$ keV, respectively.

We use a similar optimization scheme for the $0\nu\beta\beta$ decay analysis. In this case the relevant categories are horizontal lines with peaks in E_1 (see Fig. 4). We therefore make cuts only on the E_2 variable. This is performed separately for the peaks at ≈ 1900 keV and ≈ 2400 keV and results in the categories shown in Table II. Here we optimize by maximizing the approximate sensitivity to 0_1^+ decay for the 1904 keV peak, and the 2_1^+ decay for the ≈ 2400 keV peak.

APPENDIX B: SIGNAL MODEL FUNCTIONS

To model the signal shape as described in Sec. III we use phenomenological functions which are often used for modeling the shape of signals in cryogenic calorimeters (for example, see Refs. [57,58]). In particular we use a linear combination of

$$f_{\text{peak}}(E; \vec{p}, \mu, \vec{R}, \sigma) = \sum_{i=1}^{3} p_i \mathcal{N}(E, \mu, R_i \cdot \sigma(E)), \quad (B1)$$

$$f_{\text{xray}}(E; \mu, E_X, \sigma) = \mathcal{N}(E, \mu - E_X, \sigma(E)), \quad (B2)$$

$$f_{\text{step}}(E; E_{\min}, \mu, s, \sigma) = N\text{Erfc}\left(\frac{\mu - E}{\sqrt{2}\sigma}\right), \quad (B3)$$

$$f_{\text{step-neg}} = N\text{Erfc}\left(\frac{E - \mu}{\sqrt{2}\sigma}\right), \quad (B4)$$

$$f_{\text{bkg}}(E; s; E_{\min}, E_{\max}) = N(1 + sE), \quad (B5)$$

$$f_{\text{bump}} = \mathcal{N}(E, \mu, \sigma). \quad (B6)$$

Here $\mathcal{N}(E, \mu, \sigma)$ is a Gaussian with mean μ and standard deviation σ, N is a normalization coefficient and Erfc is the complementary error function. The first function models a peak, either γ or $\beta\beta$, the second models events where a Mo x-ray escapes the crystal. The third and fourth functions (step and step-neg.) are step functions to account for Compton-scattering events. The step function accounts for the Compton scatter of a single γ leading to a partial energy deposition and $E < E_{\text{photopeak}}$ and the step negative accounts for a combination of photoabsorption of one γ and Compton scatter of the other and $E > E_{\text{photopeak}}$. We include a linear background and a single Gaussian to model some features in the data where
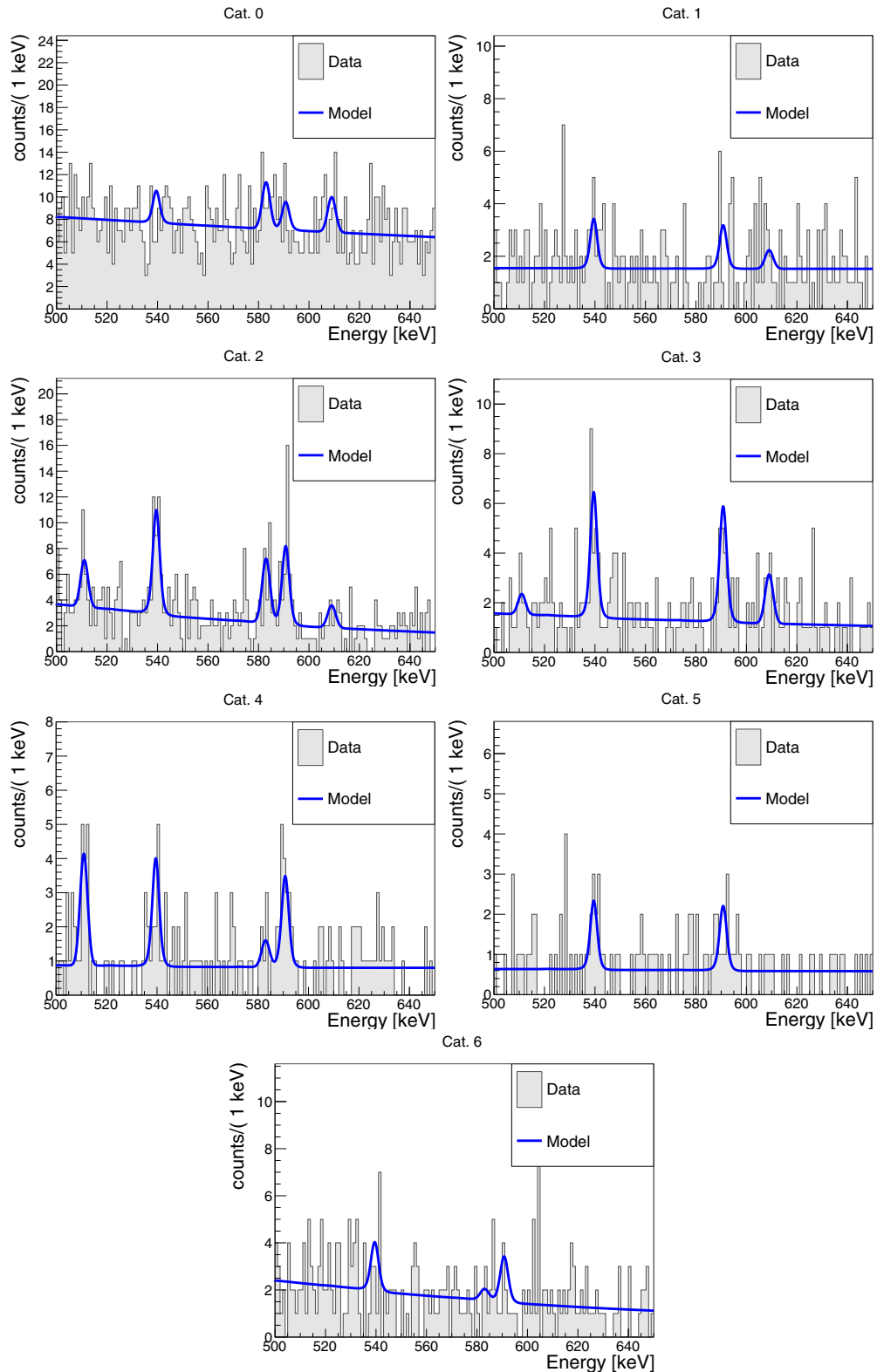
FIG. 14. Fits of the seven $2\nu\beta\beta$ decay regions of interest, the experimental data is shown in gray while the best fit (global mode) is shown in blue. We see clearly a signal in each of the regions of interest.
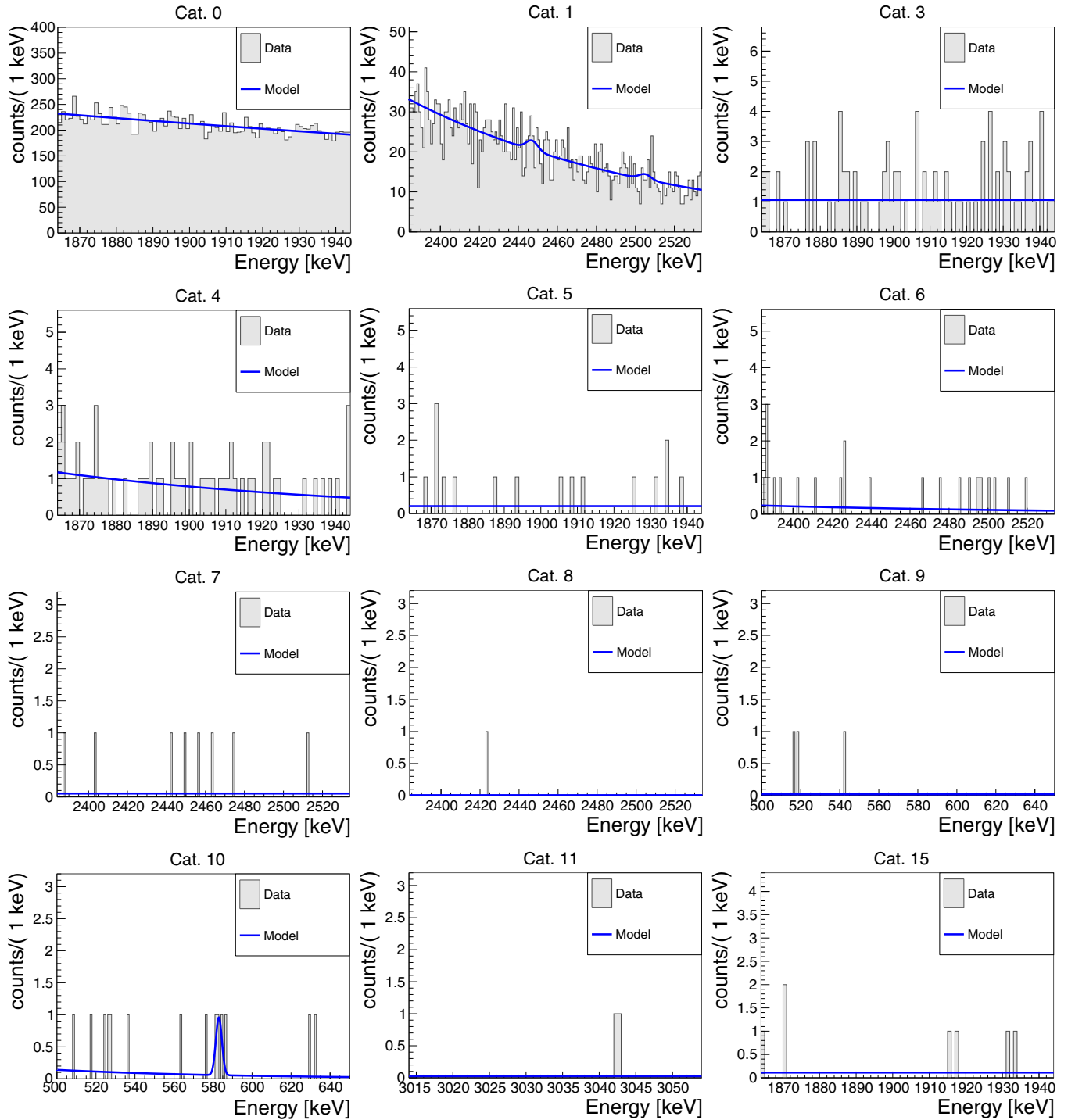
FIG. 15. Fits of the $0\nu\beta\beta$ decay regions of interest, the experimental data are shown in gray while the best fit (global mode) is shown in blue. We only show categories with nonzero counts in the fit region.

a diagonal line crosses the projected out energy leading to a very small but broader peak (bump). For example, see the category of $0\nu\beta\beta$ decay to 2_1^+ e.s. with $E_1 = 1904$ keV and $E_2 \in [500, 650]$ keV, here a diagonal line crosses the box at around 591 keV. This line is caused by a 2494 keV $\beta\beta$ in one crystal and a 540 keV γ shared between this crystal and another.

APPENDIX C: CONTAINMENT EFFICIENCY UNCERTAINTY

As mentioned in Sec. VI C, several sets of MC simulations were used to estimate the uncertainty in the containment efficiency. For each peak, category and systematic test the percentage change in the containment efficiency was extracted. These values are shown in Fig. 13.

APPENDIX D: FITS TO EACH CATEGORY

We show the best-fit reproduction of each category for both $2\nu\beta\beta$ and $0\nu\beta\beta$ decay analysis in Figs. 14 and 15. Note

that the spectra are binned for visualization purposes and we exclude plots with zero counts in the experimental data. For the $2\nu\beta\beta$ decay fits we use the $H_1 = (0_1^+ + B)$ model.

[1] R. Saakyan, Two-neutrino double-beta decay, Annu. Rev. Nucl. Part. Sci. **63**, 503 (2013).

[2] A. S. Barabash, Precise half-life values for two-neutrino double-β decay: 2020 review, Universe **6**, 159 (2020).

[3] A. S. Barabash, Double beta decay to the excited states: Review, AIP Conf. Proc. **1894**, 020002 (2017).

[4] S. Dell'Oro, S. Marcocci, M. Viel, and F. Vissani, Neutrinoless double beta decay: 2015 review, Adv. High Energy Phys. **2016**, 2162659 (2016).

[5] J. Schechter and J. W. F. Valle, Neutrinoless double-β decay in SU(2) × U(1) theories, Phys. Rev. D: Part. Fields **25**, 2951 (1982).

[6] W. H. Furry, On transition probabilities in double beta-disintegration, Phys. Rev. **56**, 1184 (1939).

[7] S. M. Bilenky and C. Giunti, Neutrinoless double-beta decay: A probe of physics beyond the standard model, Int. J. Mod. Phys. A **30**, 1530001 (2015).

[8] M. J. Dolinski, A. W. Poon, and W. Rodejohann, Neutrinoless double-beta decay: Status and prospects, Annu. Rev. Nucl. Part. Sci. **69**, 219 (2019).

[9] M. Fukugita and T. Yanagida, Barygenesis without grand unification, Phys. Lett. B **174**, 45 (1986).

[10] S. Davidson, E. Nardi, and Y. Nir, Leptogenesis, Phys. Rep. **466**, 105 (2008).

[11] F. F. Deppisch, L. Graf, J. Harz, and W. C. Huang, Neutrinoless double beta decay and the baryon asymmetry of the universe, Phys. Rev. D **98**, 055029 (2018).

[12] D. Q. Adams *et al.* (CUORE Collaboration), Search for Majorana neutrinos exploiting millikelvin cryogenics with CUORE, Nature (London) **604**, 53 (2022).

[13] W. R. Armstrong *et al.* (CUPID Interest Group), CUPID pre-CDR, arXiv:1907.09376.

[14] E. Armengaud *et al.* (CUPID-Mo Collaboration), The CUPID-Mo experiment for neutrinoless double-beta decay: performance and prospects, Eur. Phys. J. C **80**, 44 (2020).

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

[16] W. Rodejohann, Neutrinoless double-beta decay and neutrino physics, J. Phys. G **39**, 124008 (2012).

[17] G. Prézeau, M. Ramsey-Musolf, and P. Vogel, Neutrinoless double β decay and effective field theory, Phys. Rev. D: Part. Fields **68**, 034016 (2003).

[18] A. Atre, T. Han, S. Pascoli, and B. Zhang, The search for heavy Majorana neutrinos, J. High Energy Phys. 05 (2009) 030.

[19] M. Blennow, E. Fernandez-Martinez, J. Lopez-Pavon, and J. Menéndez, Neutrinoless double beta decay in seesaw models, J. High Energy Phys. 07 (2010) 096.

[20] M. Mitra, G. Senjanović, and F. Vissani, Neutrinoless double beta decay and heavy sterile neutrinos, Nucl. Phys. B **856**, 26 (2012).

[21] V. Cirigliano, W. Dekens, J. de Vries, M. L. Graesser, and E. Mereghetti, A neutrinoless double beta decay master formula from effective field theory, J. High Energy Phys. 12 (2018) 097.

[22] G. Benato, Effective Majorana mass and neutrinoless double beta decay, Eur. Phys. J. C **75**, 563 (2015).

[23] J. Engel and J. Menéndez, Status and future of nuclear matrix elements for neutrinoless double-beta decay: A review, Rep. Prog. Phys. **80**, 046301 (2017).

[24] P. K. Rath, R. Chandra, K. Chaturvedi, P. Lohani, P. K. Raina, and J. G. Hirsch, Neutrinoless $\beta\beta$ decay transition matrix elements within mechanisms involving light Majorana neutrinos, classical Majorons, and sterile neutrinos, Phys. Rev. C **88**, 064322 (2013).

[25] F. Šimkovic, V. Rodin, A. Faessler, and P. Vogel, $0\nu\beta\beta$ and $2\nu\beta\beta$ nuclear matrix elements, quasiparticle random-phase approximation, and isospin symmetry restoration, Phys. Rev. C **87**, 045501 (2013).

[26] N. L. Vaquero, T. R. Rodríguez, and J. L. Egido, Shape and Pairing Fluctuation Effects on Neutrinoless Double Beta Decay Nuclear Matrix Elements, Phys. Rev. Lett. **111**, 142501 (2013).

[27] J. Barea, J. Kotila, and F. Iachello, $0\nu\beta\beta$ and $2\nu\beta\beta$ nuclear matrix elements in the interacting boson model with isospin restoration, Phys. Rev. C **91**, 034304 (2015).

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

[29] A. Gando *et al.* (KamLAND-Zen Collaboration), Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen, Phys. Rev. Lett. **117**, 082503 (2016).

[30] O. Azzolini *et al.* (CUPID-0 Collaboration), First Result on the Neutrinoless Double-β Decay of ^{82}Se with CUPID-0, Phys. Rev. Lett. **120**, 232502 (2018).

[31] C. Augier *et al.*, Final results on the $0\nu\beta\beta$ decay half-life limit of ^{100}Mo from the CUPID-Mo experiment, Eur. Phys. J. C **82**, 1033 (2022).

[32] A. Dolgov and A. Smirnov, Possible violation of the spin-statistics relation for neutrinos: Cosmological and astrophysical consequences, Phys. Lett. B **621**, 1 (2005).

[33] A. S. Barabash, A. D. Dolgov, R. Dvornicky, F. Simkovic, and A. Y. Smirnov, Statistics of neutrinos and the double beta decay, Nucl. Phys. B **783**, 90 (2007).

[34] J. H. Hubbell and S. M. Seltzer, Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients (version 1.4). [Online] Available: http://physics.nist.gov/xaamdi (National Institute of Standards and Technology, Gaithersburg, MD, 2004).

[35] M. J. Berger, J. S. Coursey, M. A. Zucker, and J. Chang, ESTAR, PSTAR, and ASTAR: Computer programs for calculating stopping-power and range tables for electrons, protons, and helium ions (version 1.2.3). [Online] Available: http://physics.nist.gov/Star (National Institute of Standards and Technology, Gaithersburg, MD, 2005).

[36] A. S. Barabash, A possibility for experimentally observing two-neutrino double beta decay, Zh. Eksp. Teor. Fiz. **51**, 181 (1990) [JETP Lett. **51**, 207 (1990)].

[37] A. S. Barabash *et al.*, Two neutrino double beta decay of ^{100}Mo to the first excited 0^+ state in ^{100}Ru, Phys. Lett. B **345**, 408 (1995).

[38] A. S. Barabash, V. I. Umatov, R. Gurriaran, F. Hubert, and P. Hubert, $2\nu\beta\beta$ decay of ^{100}Mo to the first 0^+ excited state in ^{100}Ru, Phys. Atom. Nucl. **62**, 2039 (1999).

[39] R. Arnold *et al.* (NEMO Collaboration), Measurement of double beta decay of ^{100}Mo to excited states in the NEMO 3 experiment, Nucl. Phys. A **781**, 209 (2007).

[40] M. F. Kidd, J. H. Esterline, W. Tornow, A. S. Barabash, and V. I. Umatov, New results for double-beta decay of ^{100}Mo to excited final states of ^{100}Ru using the TUNL-ITEP apparatus, Nucl. Phys. A **821**, 251 (2009).

[41] P. Belli *et al.*, New observation of $2\beta2\nu$ decay of ^{100}Mo to the 0_1^+ level of ^{100}Ru in the ARMONIA experiment, Nucl. Phys. A **846**, 143 (2010).

[42] R. Arnold *et al.*, Investigation of double beta decay of ^{100}Mo to excited states of ^{100}Ru, Nucl. Phys. A **925**, 25 (2014).

[43] P. Belli *et al.*, Double beta decay to excited states of daughter nuclei, Universe **6**, 239 (2020).

[44] E. Armengaud *et al.* (EDELWEISS Collaboration), Performance of the EDELWEISS-III experiment for direct dark matter searches, JINST **12**, P08010 (2017).

[45] E. Armengaud *et al.* (LUMINEU Collaboration), Development of ^{100}Mo-containing scintillating bolometers for a high-sensitivity neutrinoless double-beta decay search, Eur. Phys. J. C **77**, 785 (2017).

[46] E. E. Haller, N. P. Palaio, M. Rodder, W. L. Hansen, and E. Kreysa, NTD Germanium: A novel material for low temperature bolometers, in *Neutron Transmutation Doping of Semiconductor Materials*, edited by R. D. Larrabee (Springer US, Boston, 1984), pp. 21–36.

[47] E. Armengaud *et al.* (CUPID Collaboration), New Limit for Neutrinoless Double-Beta Decay of ^{100}Mo from the CUPID-Mo Experiment, Phys. Rev. Lett. **126**, 181802 (2021).

[48] D. Q. Adams *et al.* (CUORE Collaboration), Search for double-beta decay of ^{130}Te to the 0^+ states of ^{130}Xe with CUORE, Eur. Phys. J. C **81**, 567 (2021).

[49] C. Alduino *et al.* (CUORE Collaboration), Double-beta decay of ^{130}Te to the first 0^+ excited state of ^{130}Xe with CUORE-0, Eur. Phys. J. C **79**, 795 (2019).

[50] O. Azzolini *et al.* (CUPID-0 Collaboration), Search of the neutrino-less double beta decay of ^{82}Se into the excited states of ^{82}Kr with CUPID-0, Eur. Phys. J. C **78**, 888 (2018).

[51] J. Allison *et al.*, Recent developments in Geant4, Nucl. Instrum. Methods Phys. Res., Sect. A **835**, 186 (2016).

[52] E. Armengaud *et al.* (EDELWEISS Collaboration), Background studies for the EDELWEISS dark matter experiment, Astropart. Phys. **47**, 1 (2013).

[53] D. R. Hamilton, On directional correlation of successive quanta, Phys. Rev. **58**, 122 (1940).

[54] S. Chauvie *et al.*, Geant4 low energy electromagnetic physics, in *IEEE Symposium Conference Record Nuclear Science 2004* (IEEE, Rome, Italy, 2004), Vol. 3, pp. 1881–1885.

[55] F. Simkovic, P. Domin, and S. V. Semenov, The single state dominance hypothesis and the two neutrino double beta decay of ^{100}Mo, J. Phys. G **27**, 2233 (2001).

[56] R. Arnold *et al.* (NEMO-3 Collaboration), Detailed studies of ^{100}Mo two-neutrino double beta decay in NEMO-3, Eur. Phys. J. C **79**, 440 (2019).

[57] E. Armengaud *et al.* (CUPID-Mo Collaboration), Precise measurement of $2\nu\beta\beta$ decay of ^{100}Mo with the CUPID-Mo detection technology, Eur. Phys. J. C **80**, 674 (2020).

[58] C. Alduino *et al.* (CUORE Collaboration), Analysis techniques for the evaluation of the neutrinoless double-beta decay lifetime in ^{130}Te with CUORE-0, Phys. Rev. C **93**, 045503 (2016).

[59] O. Azzolini *et al.* (CUPID-0 Collaboration), Analysis of cryogenic calorimeters with light and heat read-out for double beta decay searches, Eur. Phys. J. C **78**, 734 (2018).

[60] E. Gatti and P. Manfredi, Processing the signals from solid-state detectors in elementary-particle physics, Riv. Nuovo Cimento **9**, 1 (1986).

[61] R. Huang *et al.* (CUPID Collaboration), Pulse shape discrimination in CUPID-Mo using principal component analysis, JINST **16**, P03032 (2021).

[62] A. Caldwell, D. Kollár, and K. Kröninger, BAT—The Bayesian analysis toolkit, Comput. Phys. Commun. **180**, 2197 (2009).

[63] G. Weidenspointner *et al.*, Validation of Compton scattering Monte Carlo simulation models, in *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference and Workshop on Room-Temperature Semiconductor Detectors* (IEEE, Seoul, South Korea, 2013).

[64] G. A. P. Cirrone, G. Cuttone, F. Di Rosa, L. Pandola, F. Romano, and Q. Zhang, Validation of the geant4 electromagnetic photon cross-sections for elements and compounds, Nucl. Instrum. Methods Phys. Res., Sect. A **618**, 315 (2010).

[65] J. Kotila and F. Iachello, Phase-space factors for double-β decay, Phys. Rev. C **85**, 034316 (2012).

[66] L. Coraggio, N. Itaco, G. DeGregorio, A. Gargano, R. Mancino, and F. Nowacki, Shell-model calculation of ^{100}Mo double-β decay, Phys. Rev. C **105**, 034312 (2022).

[67] P. Pirinen and J. Suhonen, Systematic approach to β and $2\nu\beta\beta$ decays of mass $A = 100$–136 nuclei, Phys. Rev. C **91**, 054309 (2015).

[68] I. C. Bandac *et al.* (CROSS Collaboration), The $0\nu2\beta$-decay CROSS experiment: Preliminary results and prospects, J. High Energy Phys. 01 (2020) 018.