Statistical methods for the new data release of GERDA

M. Agostini [2], A. Caldwell [1], O. Schulz [1], L. Shtembari [1] on behalf of the GERDA collaboration

[1] Max-Planck-Institut für Physik, München, Germany
[2] Technische Universität München, Physik-Department, München, Germany

INTRODUCTION

The GERDA experiment [1] is searching for neutrinoless double-β decay ($0\nu\beta\beta$) of ^{76}Ge operating bare detectors, made of germanium with an enriched ^{76}Ge fraction, in liquid argon [2]. The signal would be a peak at $Q_{\beta\beta} = 2039.061 \pm 0.0007\ keV$. Under the assumption that no signal could be claimed, we will present the models and the procedures used to derive a preliminary limit on $T_{1/2}^{0\nu}$ (the $0\nu\beta\beta$ decay half-life) based on the Phase-II final data.

PRELIMINARIES

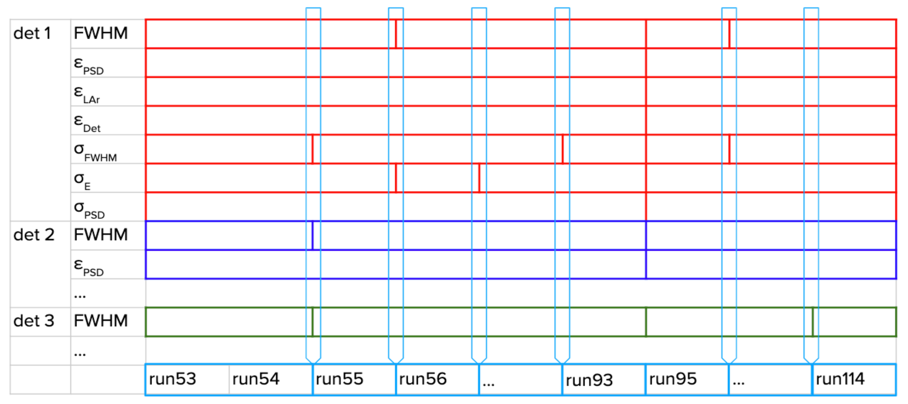
The analysis proceeds as follows:

- events with a reconstructed energy in the interval $Q_{\beta\beta} \pm 25\ keV$ are not analysed but only stored on disk
- continuous monitoring of detectors
- freezing of analysis procedure and parameters
- blinded events are processed.
- data analysis of events detected in the analysis window $1930 - 2190\ keV$ excluding the 2 gamma line regions [1]

Close monitoring of each detector allows us to:

- obtain values of parameters that affect the data analysis and model selection later described:
  - efficiencies
  - energy resolution (reported as Full Width Half Maximum, FWHM)
- create data partitions:
  - cut data with respect to different detectors
  - cut data with respect to time windows that share the same constant parameters

With these cuts we identify 383 different partitions.



Events detected in the analysis window (excluding those in the blinded window) are used to investigate the shape of the background distribution.

We test the hypothesis of a flat background by means of a test-statistic derived from Order-Statistic, which models the distribution of ranked statistical samples.

The goodness-of-fit of order k is estimated with the p-value derived from the distribution of the sum of the k smallest spacings in the energy spectrum.

The flat background hypothesis is not rejected by this test.

LIKELIHOOD

The parameter of interest for our analysis is the strength of a possible $0\nu\beta\beta$ decay signal-rate: $S = 1/T_{1/2}^{0\nu}$.

The number of expected $0\nu\beta\beta$ events in the ith partition P_i as a function of S is given by:

$$\mu_i^S = \ln 2\ (N_A/m_a)\ \epsilon_i \mathcal{E}_i S$$

where:

- ϵ_i: signal efficiency of the ith partition. Accounts for:
  - fraction of ^{76}Ge in the detector material
  - fraction of the detector active volume
  - efficiency of the analysis cuts (PSD and LAr)
  - probability that all the energy from the $0\nu\beta\beta$ decay is deposited in the active volume
- $\mathcal{E}_i$: exposure
  - total exposure of Phase-II is $103.7\ kg \cdot yr$
- m_a: ^{76}Ge molar mass

The total number of expected background events as a function of the background index $\mathrm{BI_i}$ is:

$$\mu_i^B = \mathcal{E}_i\ \mathrm{BI_i}\ \Delta E$$

where $\Delta E = 240\ keV$ is the width of the energy region around $Q_{\beta\beta}$ used for the fit.

Each data partition P_i is fitted with an unbinned likelihood function where we assume a flat distribution for the background and a Gaussian distribution for the signal:

$$\mathcal{L}_i(P_i|S,\mathrm{BI_i}) = \prod_{j=1}^{n_i} \frac{1}{(\mu_i^B + \mu_i^S)} \left[\mu_i^B \frac{1}{\Delta E} + \mu_i^S \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(\frac{-(E_j - Q_{\beta\beta})^2}{2\sigma_i^2} \right) \right]$$

where:

- n_i: total number of events observed in the ith partition
- E_j: individual event energies in the ith partition
- $\sigma_i = FWHM_i/(2\sqrt{2\ln 2})$: energy resolution in ROI
  - the average FWHM across partitions is $3.29\ keV$

The total likelihood is constructed as the product of all $\mathcal{L}_i$ weighted with the Poisson term:

$$\mathcal{L}(\boldsymbol{P}|S,\mathbf{BI}) = \prod_{i=1} \left[\frac{\exp\left(-(\mu_i^B + \mu_i^S)\right)(\mu_i^B + \mu_i^S)^{n_i}}{n_i!} \cdot \mathcal{L}_i(P_i|S,\mathrm{BI_i}) \right]$$

where $\boldsymbol{P} = \{P_1 \dots P_i \dots\}$ and $\mathbf{BI} = \{\mathrm{BI_1} \dots \mathrm{BI_i} \dots\}$.

The statistical analysis is performed in a Bayesian framework where the combined posterior probability density function is calculated according to Bayes' theorem:

$$\mathcal{P}(S,\mathbf{BI}|\boldsymbol{P}) \propto \mathcal{L}(\boldsymbol{P}|S,\mathbf{BI})\mathcal{P}(S) \prod_i \mathcal{P}(\mathrm{BI_i})$$

MODELS

The detectors in the experiment can be subdivided in three different types:

- *BEGe*: Broad Energy Geranium detectors
- *Coax*: Coaxial detectors
- *Inv-Coax*: Inverted Coaxial detectors

The background index (BI) can be treated in 3 different ways, which gives rise to 3 different models:

- ***Single background index:*** there is only one background index for all detector types

$$BI \sim Uniform$$

- ***Uncorrelated background indices:*** each detector type has its own independent BI in the data analysis

$$BI_i \sim Uniform$$

- ***Correlated Background indices:*** each detector type has a different BI but they are all correlated

The inclusion of correlations amongst detector types requires the use of a hierarchical model:

$$\sigma_{BI} \sim Uniform$$
$$m_{BI} \sim Uniform$$
$$BI_i \sim LogNormal\left(\ln(m_{BI}) - \frac{\sigma_{BI}^2}{2}, \sigma_{BI} \right)$$

where σ_{BI} and m_{BI} are the hyperparameters that control the distribution of the background indices for each detector type and represent the standard deviation and mean respectively.

Comparing the performance of each model with simulated experiments using synthetic data we find that:

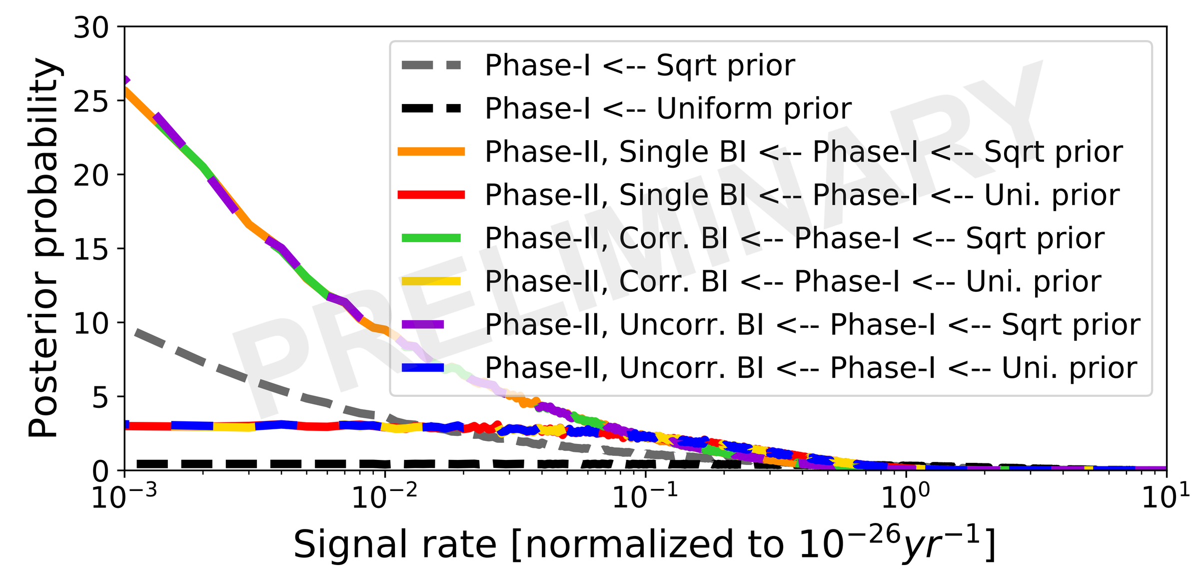
- the single BI model, on average, has a stronger discovery power, since it can reconstruct a larger portion of an injected signal event
- the uncorrelated BI model, on average, sets a stronger $T_{1/2}^{0\nu}$ limit with no injected signal events
- the correlated BI model can reproduce the behaviour of the two extreme cases and offers intermediate discovery power and limit setting capabilities
- The (median) sensitivity of all models assuming no signal and using a uniform prior for S is $T_{1/2}^{0\nu} > 1.4 \cdot 10^{26}$ yr (90% CI)

The statistical analysis was performed using the Julia package BAT.jl [5].

PRELIMINARY RESULTS

In the analysis window we detect 13 events. After a preliminary analysis of these data we cannot claim a signal. The following are preliminary results, obtained without the inclusion of systematic uncertainties in the analysis.

The posterior of the signal-rate S obtained from the Phase-I data analysis [3] is used as a prior for the Phase-II data analysis. The Phase-I analysis was performed using both a uniform and a $1/\sqrt{S}$ prior for S up to $10^{-25}\ yr^{-1}$.



The preliminary half-life limit extracted is the same for all models and it is a significant improvement with respect to the latest Phase-II data release [4]:

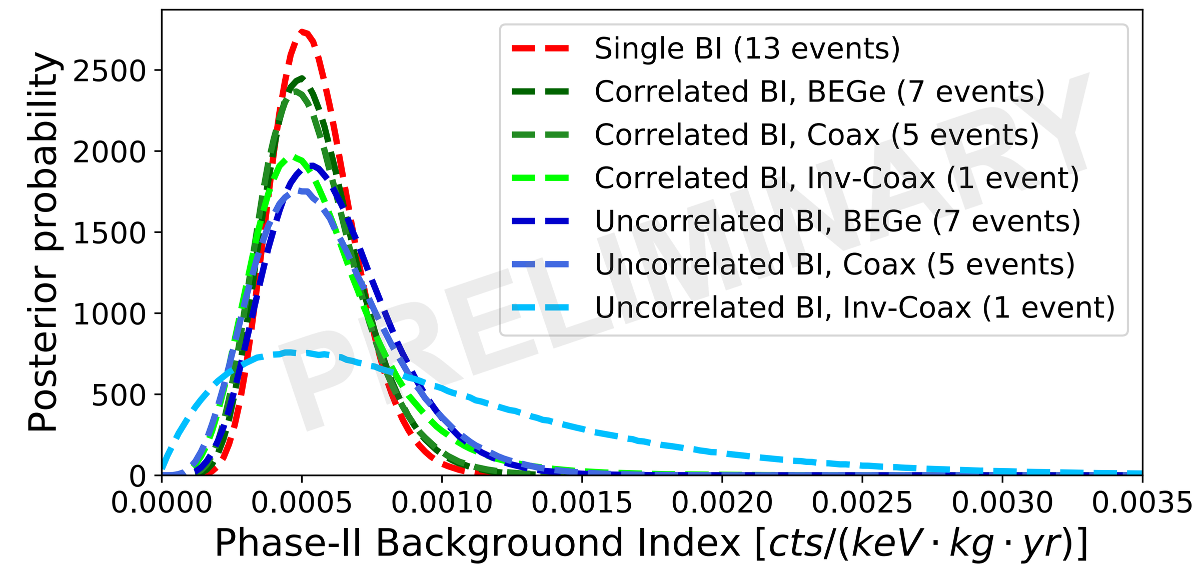
(Uniform prior) $\boldsymbol{T_{1/2}^{0\nu} > 1.4 \cdot 10^{26}\ \mathrm{yr}\ (90\%\ \mathrm{CI})}$

($1/\sqrt{S}$ prior) $\boldsymbol{T_{1/2}^{0\nu} > 2.3 \cdot 10^{26}\ \mathrm{yr}\ (90\%\ \mathrm{CI})}$

The most likely background index for the Phase-II analysis for the single BI model is:

(Phase-II) $\mathbf{BI} = 5.2^{+1.6}_{-1.4} \cdot 10^{-4} \left[\frac{\mathrm{cts}}{\mathrm{keV \cdot kg \cdot yr}}\right]$ **(68% SI)**

The other models yield similar results for the BIs.



REFERENCES

[1] Ackermann, K.-H. et al., The GERDA experiment for the search of $0\nu\beta\beta$ decay in ^{76}Ge, Eur. Phys. J. C 73, 2330 (2013)

[2] Agostini, M. et al., Background-free search for neutrinoless double-β decay of ^{76}Ge with GERDA. Nature 544, 47–52 (2017)

[3] Agostini, M. et al., Results on Neutrinoless Double-β Decay of ^{76}Ge from Phase I of the GERDA Experiment, Phys. Rev. Lett. 111, 12 (2013)

[4] Agostini, M. et al., Probing Majorana neutrinos with double-β decay, Science 365, 1445 (2019)

[5] Caldwell, A. et al., BAT.jl - A Bayesian Analysis Toolkit in Julia, doi: 10.5281/zenodo.2605312