Statistical Issues & Methods in Neutrinoless Double-β Decay Experiments

Matteo Agostini

Munich Technical University (TUM)

PhyStat-nu, CERN, Jan 22-25, 2019
Neutrinoless Double-$\beta$ Decay ($0\nu\beta\beta$)

From the point of view of Nuclear Physics:

\[(A, Z) \leftrightarrow (A, Z + 2) + 2e^-\]

- second order nuclear transition
- atomic number $Z$ increased by two units
Neutrinoless Double-$\beta$ Decay ($0\nu\beta\beta$)

From the point of view of Nuclear Physics:

$$(A, Z) \leftrightarrow (A, Z + 2) + 2e^-$$

- second order nuclear transition
- atomic number $Z$ increased by two units

From the point of view of Particle Physics:

$$2n \leftrightarrow 2p + 2e^-$$

- channel depends on new physics
- 2 leptons produced w/o balancing anti-leptons
Neutrinoless Double-$\beta$ Decay ($0\nu\beta\beta$)

If $0\nu\beta\beta$ decay is discovered:

- first observation of “matter creation”
- matter-antimatter asymmetry
- neutrinos are their own antiparticle
- origin neutrino masses
A portal to Physics beyond the Standard Model

Decay probability proportional to coherent sum of involved mechanisms:

\[ \Gamma = \frac{1}{T_{1/2}} = G(Q, Z) \times \sum_i M_i \times \eta_i \]

- **Phase Space Factor**
- **Nuclear Physics**
- **Propagator**

- **light/heavy neutrinos**
- **right-handed current**
- **gluino / R-parity**

A portal to Physics beyond the Standard Model

Decay probability proportional to coherent sum of involved mechanisms:

\[ \Gamma = \frac{1}{T_{1/2}} = G(Q, Z) \times \sum_i M_i \times \eta_i \]

Propagator factor:

- Connects decay probability to energy scale of new physics
- \( T_{1/2} \) is for \( 0\nu\beta\beta \) decay what the collision energy is for LHC
- Scale of new physics is unknown and a signal can be around the corner -> important to increase the sensitivity on \( T_{1/2} \)

Strongest constraints

\( T_{1/2} > 10^{26} \text{ yr} \)

i.e. more than a million trillion times the age of the Universe!

M. Agostini (TU Munich)
Experiment design

Detection approach:

- source in detector active material
- detector acts as a calorimeter
- additional observables used to constrain background and systematics

Signal:

- spread due to energy resolution
- peak at Q-value (often Gaussian)

Background:

- typically flat if resolution is <0.1%
- not connected to physics mechanism generating the signal -> hard to model!
Statistical analysis

Search for a peak at fixed position

- $\lambda_S$ signal expectation (typically very small)
- $\lambda_B$ background expectation
- no look elsewhere
- background control sample (on/off problem)

1. Counting:

$$\mathcal{L} = \text{Poisson}(N_{\text{tot}}^{\text{obs}} | \lambda_S + \lambda_B) \cdot \text{Poisson}(N_{B}^{\text{obs}} | \tau \lambda_B)$$

- $N_{\text{tot}}$ number of cts in signal window
- $N_B$ number of cts in control window
- $\tau$ ratio of windows size
- gives bulk of results
- window optimization is important

2. Binned ML spectral fit:

- full info on energy distribution
- goodness of background model

3. Extended Unbinned ML spectral fit:

- CPU convenient for sparse histogram
- not sensitive to background modeling
Statistical issues

Frequentist framework:
➢ physical border ($\lambda_s \geq 0$)
➢ best fit in unphysical space when background under-fluctuates
➢ empty intervals emerge to preserve coverage
➢ stronger limits for higher background

Bayesian framework:
➢ low statistic $\rightarrow$ strong impact of the prior

Large deviations between different techniques!
# Frequentist methods

<table>
<thead>
<tr>
<th>Two-sided test statistic:</th>
<th>Confidence interval construction:</th>
<th>Coverage issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>➢ $\chi^2$</td>
<td>➢ Wilks’ or Feldman Cousins (FC)</td>
<td>➢ intrinsic in discrete problem</td>
</tr>
<tr>
<td>➢ profile likelihood</td>
<td>➢ $\lambda_s$ bounded to positive values to avoid empty intervals</td>
<td>➢ Wilks’ overcovers for $\lambda_s &lt; \sqrt{\lambda_B}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>➢ FC overcovers for small $\lambda_s$</td>
</tr>
</tbody>
</table>

![Two-sided test statistic](image1.png) ![Confidence interval construction](image2.png)
“Deep Poisson” regime

Spectral fit under Wilks’ approximation

Threshold of test statistic for a 90% CL derived through a FC construction (GERDA Phase II)

ChisquareQuantile(90%,1)=2.7

Background-free condition
Target of many experiments
### Sensitivity Definitions

<table>
<thead>
<tr>
<th>CL</th>
<th>Concept</th>
<th>How to compute</th>
</tr>
</thead>
<tbody>
<tr>
<td>limit setting</td>
<td>90%</td>
<td>Assuming there is no signal, what is the expected upper limit on the signal expectation?</td>
</tr>
<tr>
<td>signal discovery</td>
<td>99.7% (3σ)</td>
<td>Assuming there is a signal, how strong does it have to be to make a discovery?</td>
</tr>
</tbody>
</table>

[Consistent with prescription from Cowan, Cranmer et al]
Counting Experiment Sensitivity

Simple calculation for signal + known background:

- based on distributions of expected frequency of observations
- Poisson CDF approximated with gamma functions

90% CL limit setting

3σ signal discovery

M. Agostini (TU Munich)
Counting Experiment Sensitivity

Simple calculation for signal + known background:

➢ based on distributions of expected frequency of observations
➢ Poisson CDF approximated with gamma functions

Background free regime:

➢ limit setting sensitivity: flattens for a background expectation of .7 cts (median of $H_0$ is at 0)
➢ signal discovery sensitivity: flattens for a background expectation of 0.003 (99.7% quantile of $H_0$ is at 0)
➢ 1 count is a discovery for low enough background

[M.A., G Benato and J A Detwiler, PRD 96, 053001 (2017)]
Counting Experiment Sensitivity

Simple calculation for signal + known background:

- based on distributions of expected frequency of observations
- Poisson CDF approximated with gamma functions

Background free regime:

- **limit setting sensitivity**: flattens for a background expectation of .7 cts (median of $H_0$ is at 0)
- **signal discovery sensitivity**: flattens for a background expectation of 0.003 (99.7% quantile of $H_0$ is at 0)
- 1 count is a discovery for low enough background

M. Agostini (TU Munich)
Counting Experiment Sensitivity

Comparison with a FC construction:

- likelihood ratio
- test statistic distributions from toy MC
- median significance (not mean!)

Features:

- jumps in coverage due to discrete nature of problem
- not monotonic functions -> apparent sensitivity improvement when increasing background
- “better than background free” regime
Target materials

35 isotopes, 9 actually used for $0\nu\beta\beta$ searches:

Different detection techniques available for different isotopes

Variety of the field is important for convincing discovery
Liquid/Gas vs Solid Detectors

Loaded scintillator detectors or Xe Time Projection Chambers

- $0\nu\beta\beta$ isotope mixed in the liquid/gas material
- self-shielding from external background
- volume fiducialization

Cryogenic Bolometers or Semiconductor detectors:

- many crystals of isotopically enriched material
- detector granularity
- per mill energy resolution

CUORE
CUPID
AMORE
Majorana
GERDA
Liquid/Gas vs Solid Detectors

KamLAND-Zen
B=O(10) cts

EXO
B=O(10) cts

BQ_{\beta\beta} \pm 2\sigma

CUORE: B=O(10) cts

GERDA: B=O(0.1) cts

enriched BEGe - 30.8 kg-\text{yr}

Q_{\beta\beta} \pm 2\sigma

Counts / (keV-kg-yr)
Liquid/Gas vs Solid Detectors

(c) KamLAND-Zen
B=O(10) cts

EXO
B=O(10) cts

BQ_{ββ} ± 2σ

(c) CUORE:
B=O(10) cts

Q_{ββ}

(c) GERDA:
B=O(0.1) cts

enriched BEGe - 30.8 kg·yr
Q_{ββ} ± 2σ
### KamLAND-Zen

<table>
<thead>
<tr>
<th>Location</th>
<th>Kamioka, Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isotope</td>
<td>$^{136}$Xe [$Q_{\beta\beta} = 2458$ keV]</td>
</tr>
<tr>
<td>Technology</td>
<td>Xe-loaded liquid scintillator</td>
</tr>
<tr>
<td>Isotope Mass</td>
<td>350 kg</td>
</tr>
<tr>
<td>$0\nu\beta\beta$ efficiency</td>
<td>16%</td>
</tr>
<tr>
<td>Resolution [$\sigma$]</td>
<td>100-120 keV</td>
</tr>
<tr>
<td>Latest results</td>
<td>$T_{1/2} &gt; 1.1 \cdot 10^{26}$ yr (90% CL)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>$T_{1/2} &gt; 5.6 \cdot 10^{25}$ yr (90% CL)</td>
</tr>
</tbody>
</table>

**Frequentist likelihood fit**
- Multivariate: E vs R
- Wilks’ approximation tested with toy MC
<table>
<thead>
<tr>
<th>Location</th>
<th>Isotope</th>
<th>Technology</th>
<th>Isotope Mass</th>
<th>0νββ efficiency</th>
<th>Resolution $[\sigma]$</th>
<th>Latest results</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIPP, New Mexico, USA</td>
<td>$^{136}$Xe $[Q_{\beta\beta} = 2458 \text{ keV}]$</td>
<td>TPC with liquid Xe</td>
<td>76 kg</td>
<td>80%</td>
<td>34 keV</td>
<td>$T_{1/2} &gt; 1.8 \times 10^{25} \text{ yr (90% CL)}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$T_{1/2} &gt; 3.7 \times 10^{25} \text{ yr (90% CL)}$</td>
</tr>
</tbody>
</table>

Frequentist binned likelihood fit:
- multivariate (energy, position, TMVA observables)
- Wilks’ approximation valid (coverage tested)
CUORE

<table>
<thead>
<tr>
<th>Location</th>
<th>LNGS, Italy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isotope</td>
<td>$^{130}$Te [$Q_{\beta\beta} = 2527 \text{ keV}$]</td>
</tr>
<tr>
<td>Technology</td>
<td>Cryogenic calorimeters</td>
</tr>
<tr>
<td>Isotope Mass</td>
<td>206 kg</td>
</tr>
<tr>
<td>$0\nu\beta\beta$ efficiency</td>
<td>68%</td>
</tr>
<tr>
<td>Resolution [$\sigma$]</td>
<td>3.3 keV</td>
</tr>
<tr>
<td>Latest results</td>
<td>$T_{1/2} &gt; 1.5 \cdot 10^{25}$ yr (90% CL)</td>
</tr>
<tr>
<td></td>
<td>$T_{1/2} &gt; 0.7 \cdot 10^{25}$ yr (90% CL)</td>
</tr>
</tbody>
</table>

Bayesian:
- flat prior
- profiling instead of marginalization

Frequentist:
- bounded profile likelihood
- Wilks approximation

[Phys. Rev. Lett. 120, 132501 (2018)]

1000 detectors
GERDA

<table>
<thead>
<tr>
<th>Location</th>
<th>LNGS, Italy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isotope</td>
<td>$^{76}$Ge [$Q_{\beta\beta} = 2039$ keV]</td>
</tr>
<tr>
<td>Technology</td>
<td>Semiconductor Ge detectors</td>
</tr>
<tr>
<td>Isotope Mass</td>
<td>35 kg</td>
</tr>
<tr>
<td>$0\nu\beta\beta$ efficiency</td>
<td>65%</td>
</tr>
<tr>
<td>Resolution [$\sigma$]</td>
<td>1.3 keV</td>
</tr>
</tbody>
</table>

**Latest results**

- $T^{1/2}_{1/2} > 0.9 \times 10^{26}$ yr (90% CL)
- $T^{1/2}_{1/2} > 1.1 \times 10^{26}$ yr (90% CL)

**Sensitivity**

- $T^{1/2}_{1/2} > 0.9 \times 10^{26}$ yr (90% CL)
- $T^{1/2}_{1/2} > 1.1 \times 10^{26}$ yr (90% CL)

**Frequentist:**

- extended unbinned likelihood
- profile likelihood
- FC construction (only for best fit value of nuisance parameters)

**Bayesian:**

- flat prior

[Nature 544 (2017) 47]
How to split data in subsets? One data set for each detector?

➢ 40 detectors, 5 counts in analysis window
➢ 1 parameter of interest ($\lambda_s$ common for all sets)
➢ 40 nuisance parameters (one $\lambda_B$ for each subset)
➢ $\lambda_B$ is either completely or mostly unconstrained

Statistical / Physical difficulties:
➢ profiling / FC construction / marginalization quickly not feasible
➢ how to claim for a signal when background is unconstrained
Discovery analysis with 1 candidate event

Conceptual problem:

- focus on a single signal-candidate event
- this event comes with a set of experimental parameters
- we feel like the analysis should use those parameters and not average values over a data set
- splitting data in more sets would reduce the difference between average parameters and those of a single event but leads to proliferation of nuisance parameters associated to empty data sets and larger uncertainties on their central values

Current approach:

- find a trade-off between number of data sets and variance of the average parameters
- blinded analysis (including blinded data set definition) avoid biases in the process
Systematics

Main systematics:

➢ background shape (liquid/gas experiments)
➢ energy scale and resolution
➢ signal detection efficiency
   ➢ active volume
   ➢ analysis cuts
➢ ...

Approaches:

➢ nuisance parameters / pull terms
➢ fit with different models
➢ toy-MC generated under different models

Impact of systematic uncertainties

➢ not informative data -> impact determined by shape of priors or pull term
➢ 1% level for solid state detectors (no counts to pull nuisance parameters)
➢ 5-10% level for gas/liquid experiments
➢ different in case of signal

EXO-200 systematic uncertainties:

<table>
<thead>
<tr>
<th>Source</th>
<th>Phase I</th>
<th>Phase II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal detection efficiency</td>
<td>3.0%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Background errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spectral shape agreement</td>
<td>2.1%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Background model</td>
<td>5.6%</td>
<td>5.9%</td>
</tr>
<tr>
<td>Energy scale and resolution</td>
<td>1.5%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Total</td>
<td>6.2%</td>
<td>6.2%</td>
</tr>
</tbody>
</table>
most experiments quote results from multiple methods and give enough info to reproduce the analysis

- sensitivity always reported (sometimes also for Bayesian methods)

- blind analysis is almost the standard

- frequentist intervals still used as Bayesian intervals (even when Bayesian interval is available)

- sensitivity computed for the no signal hypothesis, more interesting to quote discovery power

<table>
<thead>
<tr>
<th>Statistical Method in the last PRL of the MAJORANA DEMONSTRATOR</th>
<th>$T_{1/2}$ lower limit [10^{25} yr]</th>
<th>$T_{1/2}$ lower limit sensitivity [10^{25} yr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counting (FC)</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Unbinned likelihood fit (FC)</td>
<td>1.9</td>
<td>2.1</td>
</tr>
<tr>
<td>Unbinned likelihood fit &amp; CLs</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Bayesian flat prior</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Bayesian Jeffreys prior</td>
<td>2.6</td>
<td></td>
</tr>
</tbody>
</table>

[Phys. Rev. Lett. 120, 132502 (2018)]

M. Agostini (TU Munich)
Outlook

➢ 0νββ decay is a portal to new physics and experiments are interested in signals at the edge of their sensitivity

➢ The search for a peak with background keeps on posing new challenges: (too) many methods available, very different results

➢ Many experiments operate in the “Deep Poisson” regime where even a simple concept as sensitivity requires attention

➢ Important to shift focus towards a discovery analysis and define in advance how to dial with it a discovery based on a single count
Backup slides
$0\nu\beta\beta$ and $\nu$ Mass Origin

Black Box theorem:
$0\nu\beta\beta$ operator can be rearranged into a $\nu-\bar{\nu}$ oscillation (i.e. a Majorana mass term)

If $0\nu\beta\beta$ decay is discovered:
- neutrinos are their own antiparticle
- neutrinos can have a Majorana mass
- neutrino small masses can be explained through see-saw models

M. Agostini (TU Munich) [Schechter, Valle, PRD 25 (1982) 2951]
Sensitivity of an experiment

➢ 3σ signal discovery sensitivity
assuming there is a signal, find $H_1$ such that $H_0$ is rejected at 3σ in 50% of the cases

➢ 90% CL limit setting sensitivity
assuming there is no signal, find $H_1$ that is rejected in at 90% CL in 50% of the cases
Analytical computation of sensitivity

Signal discovery

- Find the number of counts $C_{3\sigma}$ such that: $\text{CDF}(C_{3\sigma} \mid B) = \text{erf}(3/\sqrt{2})$
- Solve: $\text{CDF}(C_{3\sigma} \mid S_{3\sigma} + B) = 50\%$
- $C_{3\sigma}$ is an integer: $S_{3\sigma}$ has discrete jumps ➔ Approximate the Poisson CDF with the upper incomplete gamma function so that the above equations can be inverted with standard numerical methods

Limit Setting:

- Find the median number of cts expected from bkg only $C_{\text{med}}$: $\text{CDF}(C_{\text{med}} \mid B) = 50\%$
- Solve: $\text{CDF}(C_{\text{med}} \mid S_{90\%CL} + C_{\text{med}}) = 10\%$

The counting experiment with a profile likelihood

if B is perfectly known:

\[ B := \text{background expectation} \]
\[ S := \text{signal expectation} \]
\[ N := \text{number of cts in ROI} \]

\[ L(s) = \text{Pois}(N|S+B) \]
\[ t(s) = -2 \left[ \text{Pois}(N|S+B) - \text{Pois}(N|S_{\text{best}}+B) \right] \]
with \( S_{\text{best}} = \max(0,N-B) \)

if B is derived from a side band or control region:

\[ \tau := \text{side band width / ROI width} \]
\[ \text{(for GERDA: } \tau = 220/6 \sim 40) \]
\[ M := \text{number of cts in side band} \]

\[ L(s) = \text{Pois}(N|S+B) \times \text{Pois}(M|\tau B) \]
\[ t(s) = -2 \left[ \log\text{Pois}(N|S+B_{\text{cond}}) + \log\text{Pois}(M|\tau B_{\text{cond}}) \\ - \log\text{Pois}(N|S_{\text{best}}+B_{\text{best}}) - \log\text{Pois}(M|\tau B_{\text{best}}) \right] \]
with: \( S_{\text{best}} = N-M/\tau \),
\[ B_{\text{best}} = M/\tau \]
\[ B_{\text{cond}} = N+M-(1+t)S+\sqrt{(N+M-(1+t)S)^2 + 4(1+t)SM)} / [2(1+t)] \]
Sensitivity with background uncertainty
<table>
<thead>
<tr>
<th>Major $0\nu\beta\beta$ Projects</th>
<th>current gen</th>
<th>mid-term</th>
<th>long-term</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gas/Liquid detector</strong></td>
<td>Liquid scintillator</td>
<td>KZ</td>
<td>KZ-800 SNO$^+$ phase I</td>
</tr>
<tr>
<td></td>
<td>Time Projection chambers</td>
<td>EXO NEXT-10</td>
<td>NEXT-100 PANDA-X-III</td>
</tr>
<tr>
<td><strong>Solid detectors</strong></td>
<td>Cryogenic Calorimeters</td>
<td>CUORE CUPID-0 AMORE</td>
<td>AMORE II</td>
</tr>
<tr>
<td></td>
<td>Ge semiconductor</td>
<td>GERDA MJD</td>
<td>LEGEND-200</td>
</tr>
<tr>
<td><strong>External detectors</strong></td>
<td>Magnetized tracking</td>
<td>NEMO</td>
<td></td>
</tr>
</tbody>
</table>

[Adapted from A Giuliani, Neutrino2018]
Probability Density from Global Fits

In absence of neutrino mass mechanisms or flavour symmetries that fix the value of the Majorana phases or drive \( m_{\text{light}} \) to zero, the probability distribution for \( m_{\beta \beta} \) is pushed to large values:

Flat prior for the Majorana phases \( \Rightarrow \) small \( m_{\beta \beta} \) values require a fine tuning of the parameters

[M.A., G Benato and J A Detwiler, PRD 96, 053001 (2017)]
Probability density from global fits

- data in the analysis: oscillations + $0\nu\beta\beta$ + (cosmology)
- bands shows deformation due to NME uncertainty
- $0\nu\beta\beta$ constraints on $m_{\text{lightest}}$ competitive with cosmology

Bulk of probability at reach with next generation experiments

see also [A Caldwell et al, Phys.Rev. D96 (2017) no.7, 073001]
Bayesian global fits

Flavour Models

Sterile Neutrinos

[King et al., JHEP 1312, 005 (2013)]
[M.A. et al., EPJ C76 (2016) no.4, 176]
[Mass et al., PRD 96, 053001 (2017)]
Normal vs Inverted ordering

T2K, PRD 96, 092006 (2017)

NOvA, Nu2018

- $\sin^2 \theta_{13} = 0.082$
- $\sin^2 \theta_{23} = 0.55$
- $\sin^2 \theta_{23} = 0.45$
- $\Delta m^2_{32} = -2.55 \times 10^{-3} \text{eV}^2$
- $\Delta m^2_{32} = +2.50 \times 10^{-3} \text{eV}^2$

1 $\sigma$
2 $\sigma$
Normal vs inverted ordering

- NOVA (no sensitive)
- T2K (no sensitive)

\[ \delta_{CP} = -\pi/2, \text{ NH} \]