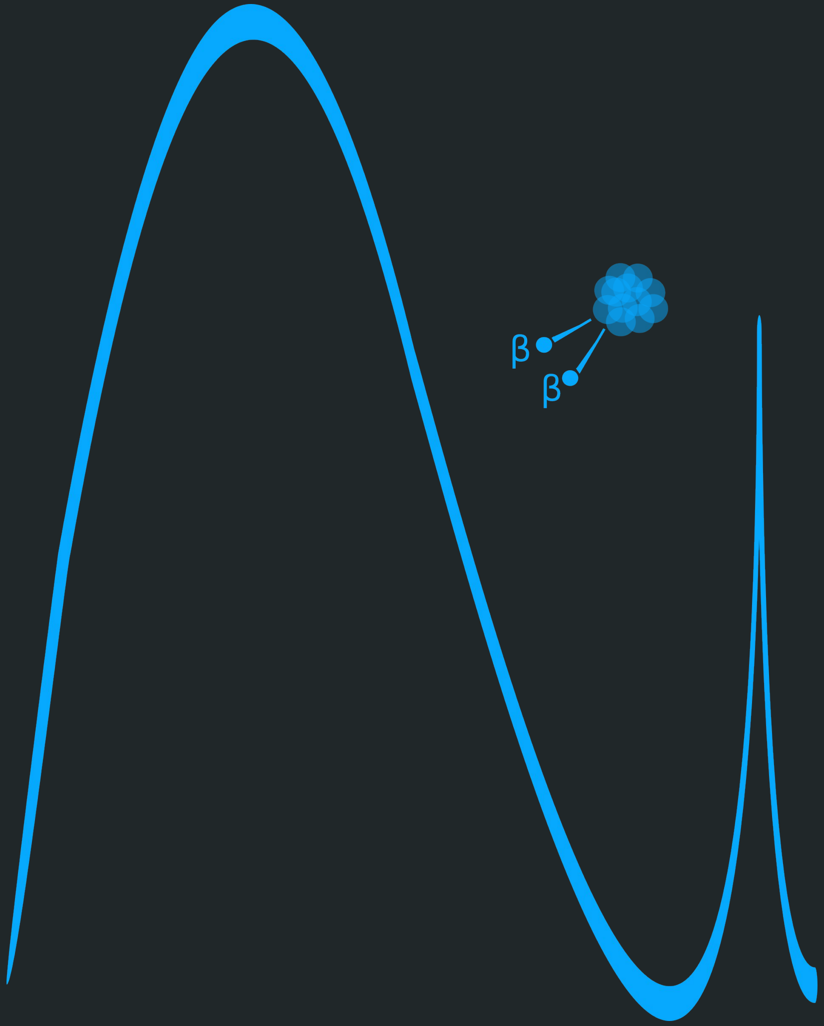


Statistical Inference in Double-β Decay Searches

Matteo Agostini

Technical University Munich

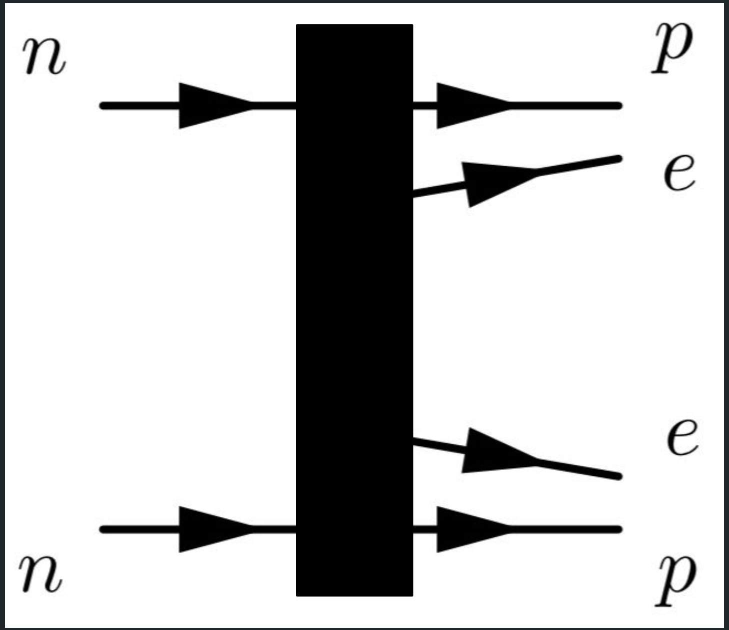
PHYSTAT Dark Matter
Stockholm, Jul 31 - Aug 2, 2019

Neutrinoless Double-β Decay ($0\nu\beta\beta$)

Nuclear transition in which:

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

- channel depends on new physics
- 2 leptons produced w/o balancing anti-leptons
  - **matter creating** process
  - complementary to proton decay
  - matter-antimatter asymmetry
- possible only if neutrinos are their own antiparticle
  - origin of neutrino masses

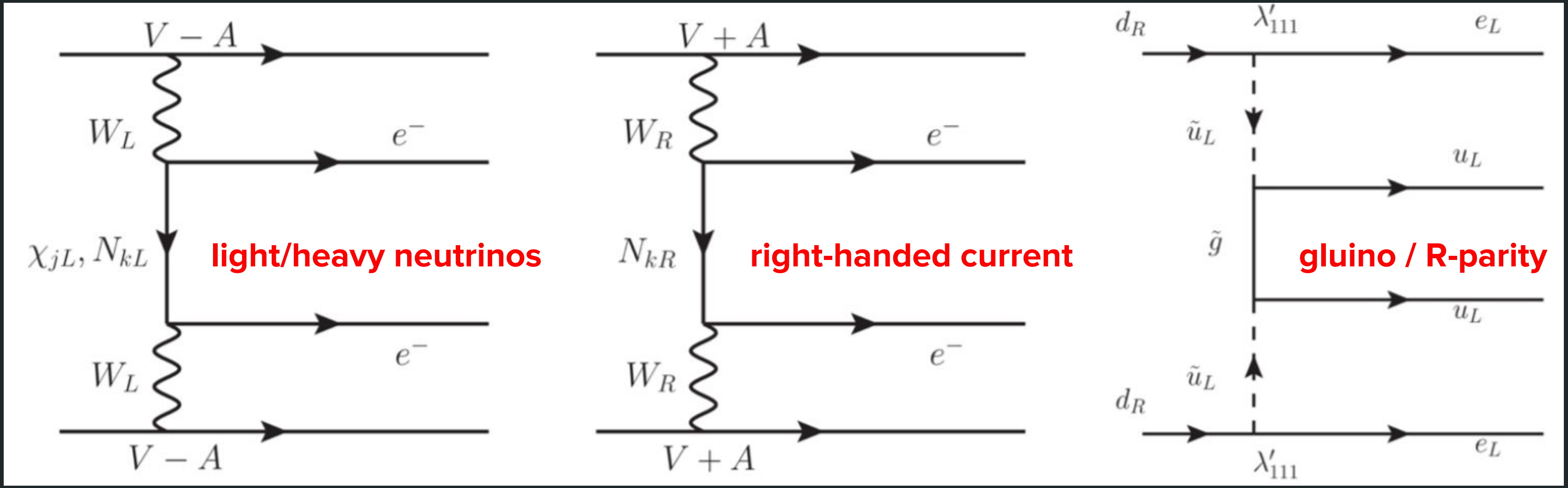


A portal to Physics beyond the Standard Model

Decay probability proportional to coherent sum of involved mechanisms:

$$\Gamma = \frac{1}{T_{1/2}} = \left|\sum_i G_i \times \mathcal{M}_i \times \eta_i\right|^2$$

Phase Space Factor — Nuclear Physics — Propagator



[Faessler et al, PRD, 83, 11 (2011), 113003]

A portal to Physics beyond the Standard Model

Decay probability proportional to coherent sum of involved mechanisms:

$$\Gamma = \frac{1}{T_{1/2}} = \left| \sum_i G_i \times \mathcal{M}_i \times \eta_i \right|^2$$

Phase Space Factor — Nuclear Physics — Propagator

- $T_{1/2}$ is connected to neutrino physics and other BSM processes (heavy sterile neutrinos, SUSY, ...)
- $T_{1/2}$ is for **$0\nu\beta\beta$** decay what the **collision energy** is for **LHC**
- Experiments: $T_{1/2} > 10^{26}$ **yr**, i.e. more than a **million trillion** times the age of the Universe!

Experimental Searches - The Signal

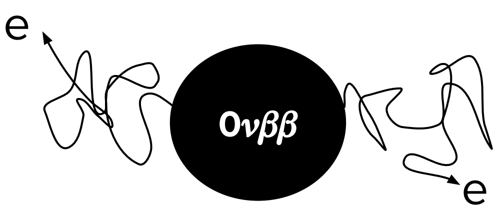
Observables for (single-isotope) experiments:

- $T_{1/2}$
- daughter isotope status
- electron energy and angular correlations

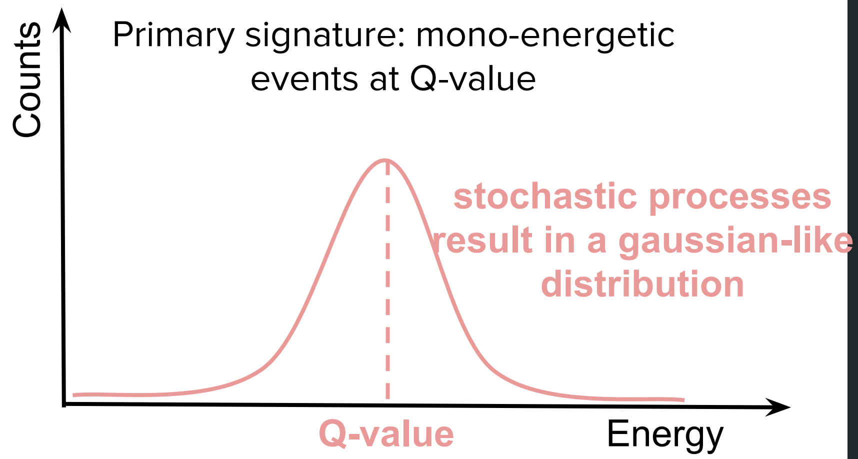
Currently most sensitive searches are based on colorimetric approach:

- source is the detector active material
- efficiency maximized
- full energy measured

$0\nu\beta\beta$: $(A, Z) \rightarrow (A, Z + 2) + 2e^-$



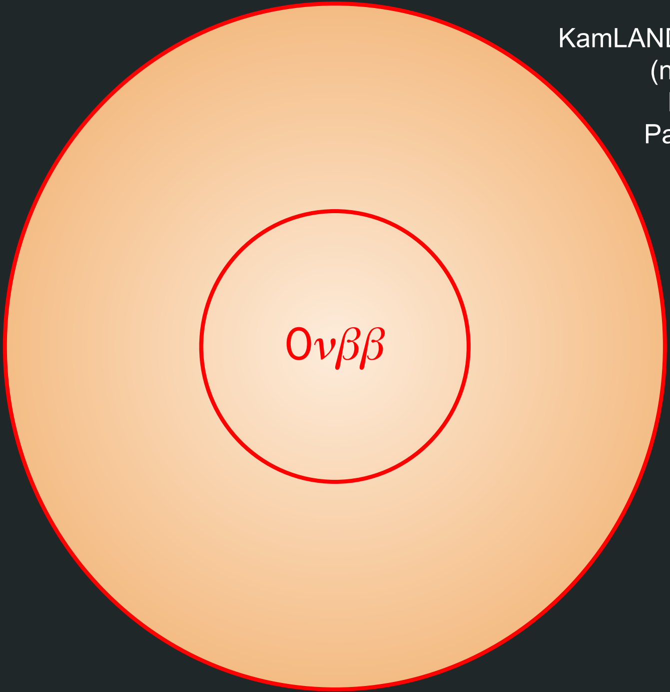
Sum of electron energy equal to Q-value



$T_{1/2} > 10^{26}$ yr $\rightarrow$ 1 event/yr in 10^3 moles of isotope

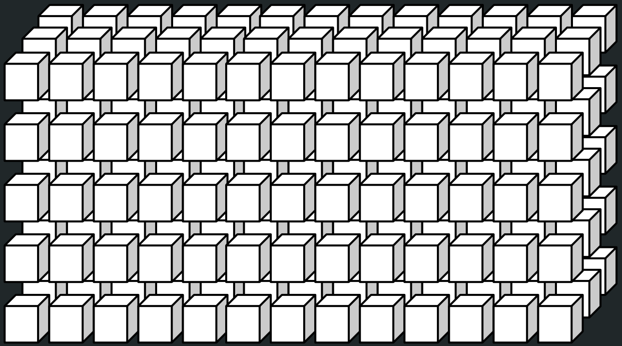
Detector Design

KamLAND-Zen
(n)EXO
NEXT
PandaX



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



CUORE
CUPID
AMORE
Majorana
GERDA

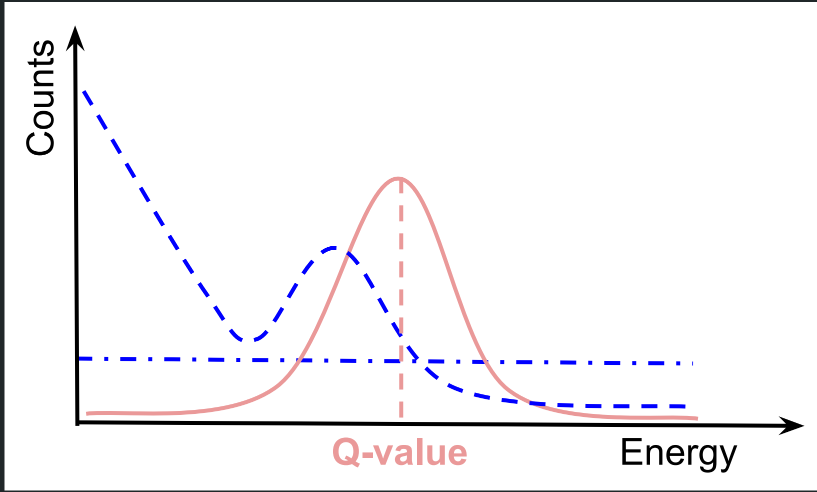
Cryogenic Bolometers or Semiconductor detectors:

- many crystals of isotopically enriched material
- detector granularity
- 0.1% energy resolution

Experimental Searches - The Background

Residual Backgrounds:

- ➢ gamma-rays due to radioactive isotopes in the material surrounding
- ➢ radioactive isotopes within the detector or on its surface
- ➢ cosmic rays
- ➢ …



Various observables to constrain the background:

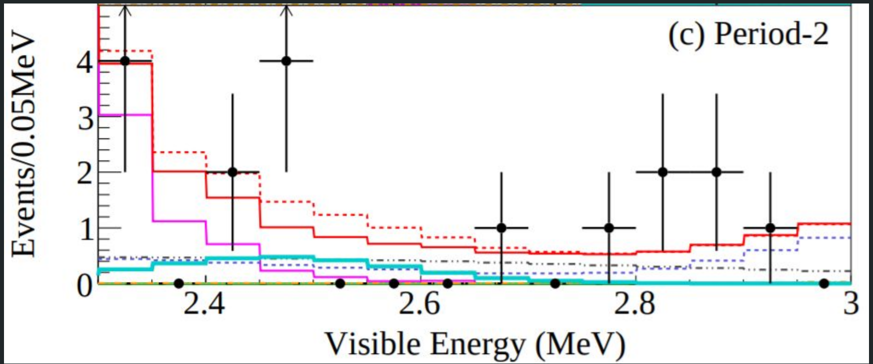
- ➢ event location and topology
- ➢ particle identification
- ➢ time correlations
- ➢ …

The background is however not connected to physics mechanism generating the signal

- ➢ hard to model -> systematic uncertainty
- ➢ if energy resolution is good enough, the background becomes approximately flat

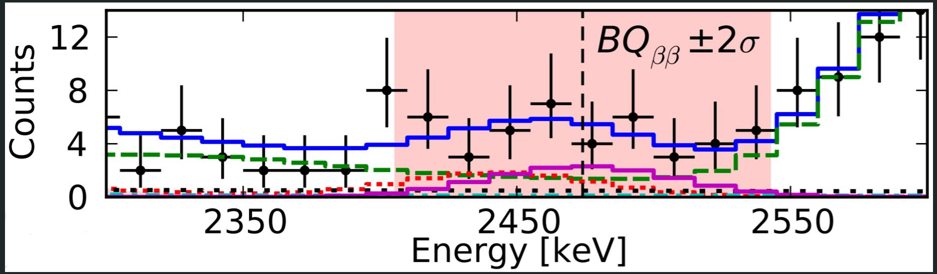
Liquid/Gas vs Solid Detectors

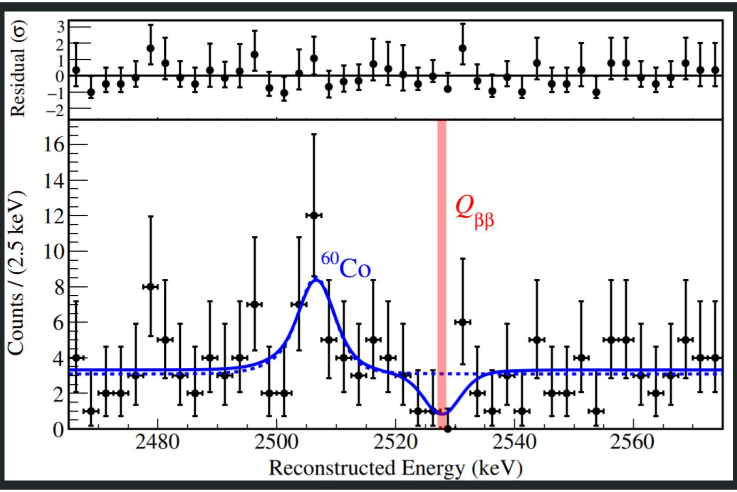
[Phys.Rev. Lett. 117 (2016) 109903]
[Phys.Rev. Lett. 120 (2018) 072701]
[Phys.Rev. Lett. 120 (2018) 132501]



KamLAND-Zen
B=O(10) cts

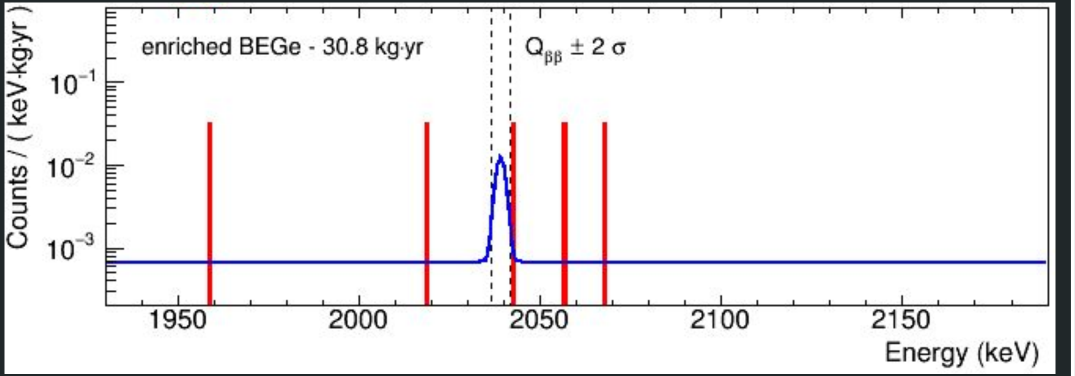
EXO $B=O(10)$ cts





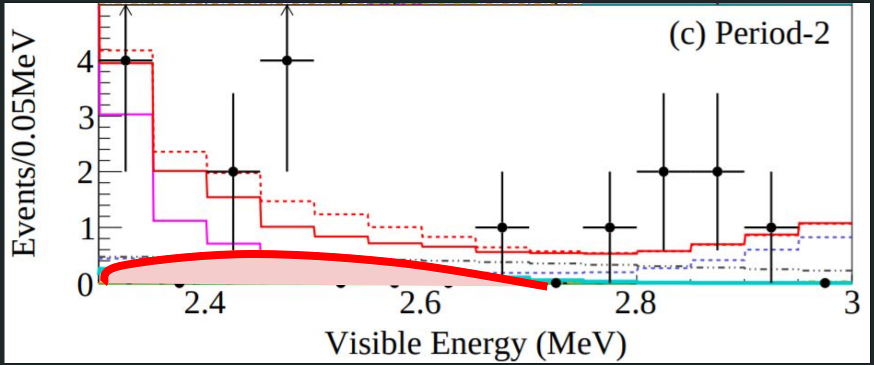
CUORE: $B=O(10)$ cts

GERDA: $B=O(0.1)$ cts



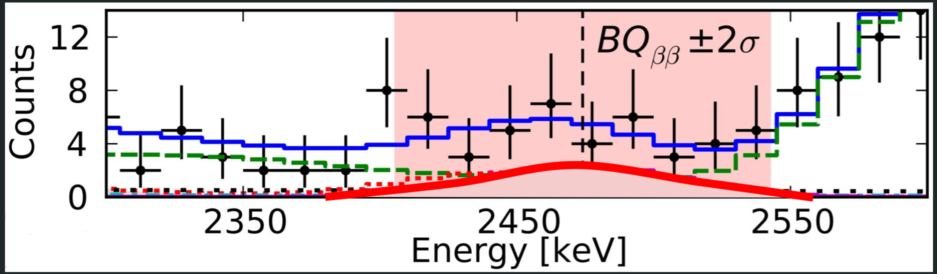
Liquid/Gas vs Solid Detectors

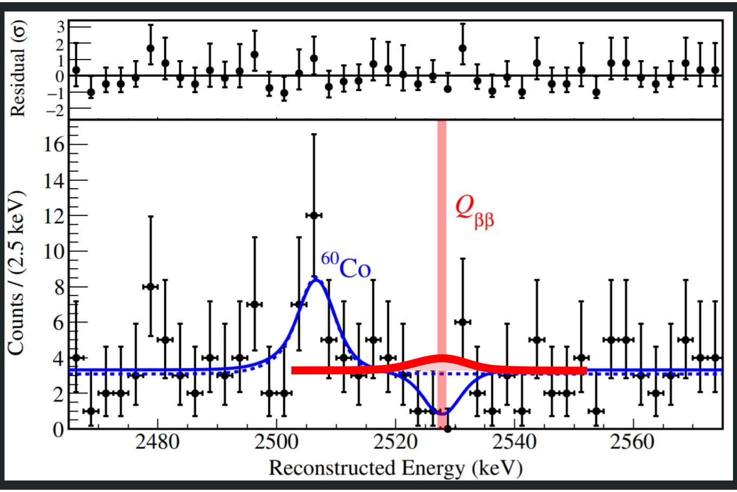
[Phys.Rev. Lett. 117 (2016) 109903]
[Phys.Rev. Lett. 120 (2018) 072701]
[Phys.Rev. Lett. 120 (2018) 132501]



KamLAND-Zen
B=O(10) cts

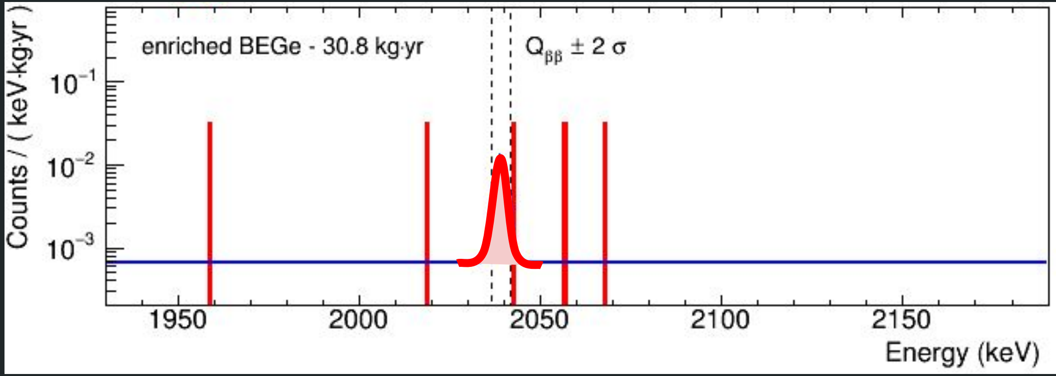
EXO $B=O(10)$ cts





CUORE: $B=O(10)$ cts

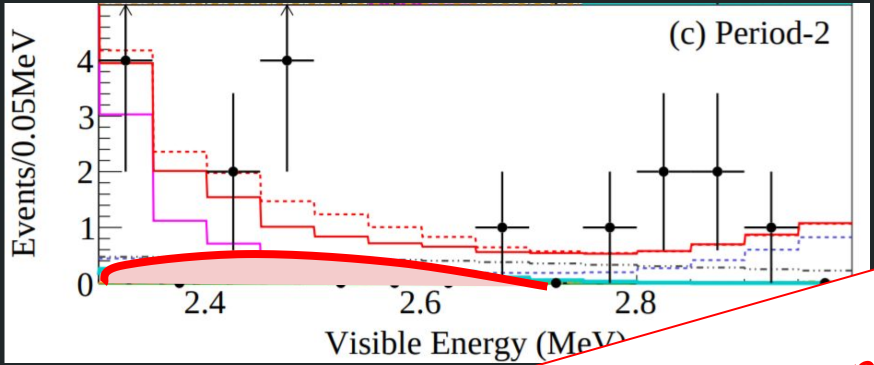
GERDA: $B=O(0.1)$ cts



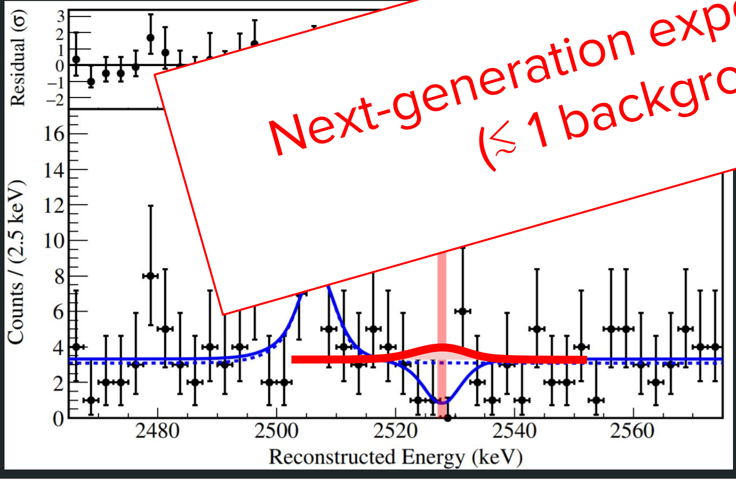
[Phys.Rev. Lett. 117 (2016) 109903]
[Phys.Rev. Lett. 120 (2018) 072701]
[Phys.Rev. Lett. 120 (2018) 132501]

Liquid/Gas vs Solid Detectors

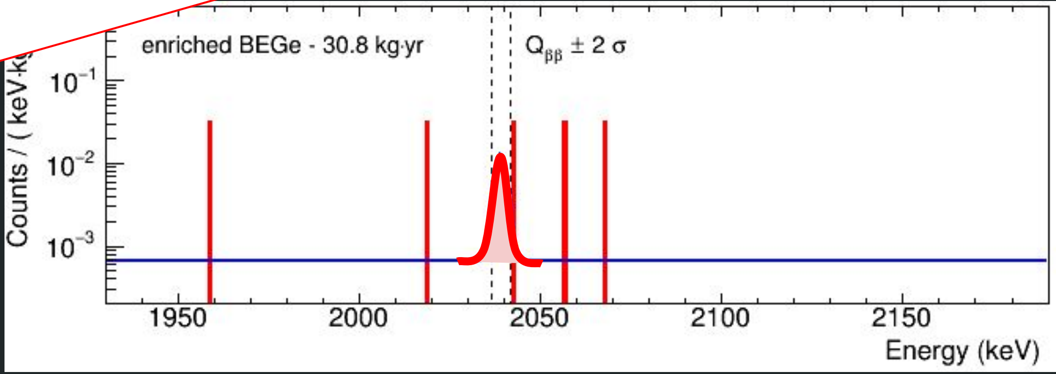
KamLAND-Zen
B=O(10) cts



EXO $B=O(10)$ cts



GERDA: $B=O(0.1)$ cts



Next-generation experiments aim to be "background-free"
($\lesssim$ 1 background event in the peak region)

Statistical Inference used in 0νββ Experiments

Question:	Statistical Task	Frequentist Techniques	Bayesian Techniques
What is the most plausible $T_{1/2}$ value (i.e. peak amplitude)?	Point Estimation	Maximum likelihood estimators	Mode of Posterior
Is there a signal? With which significance the no-signal hypothesis can be rejected?	Hypothesis test: $H_0 : 1 / T_{1/2} = 0$ $H_1 : 1 / T_{1/2} \neq 0$	Generalized likelihood ratio tests (two-sided profiled likelihood)	Bayes factors or posterior distributions
Which set of $T_{1/2}$ values is compatible with the data?	Interval Estimation	Inversion of likelihood-ratio tests	Smallest/Central interval of posterior

- ➢ Other Frequentist techniques have been used in the past but we are now converging to a standard
- ➢ Chi-square test statistic and distribution still used in combination with coverage checks
- ➢ Bayesian techniques are becoming increasingly popular

Statistical Problem

Search for a peak over some background:

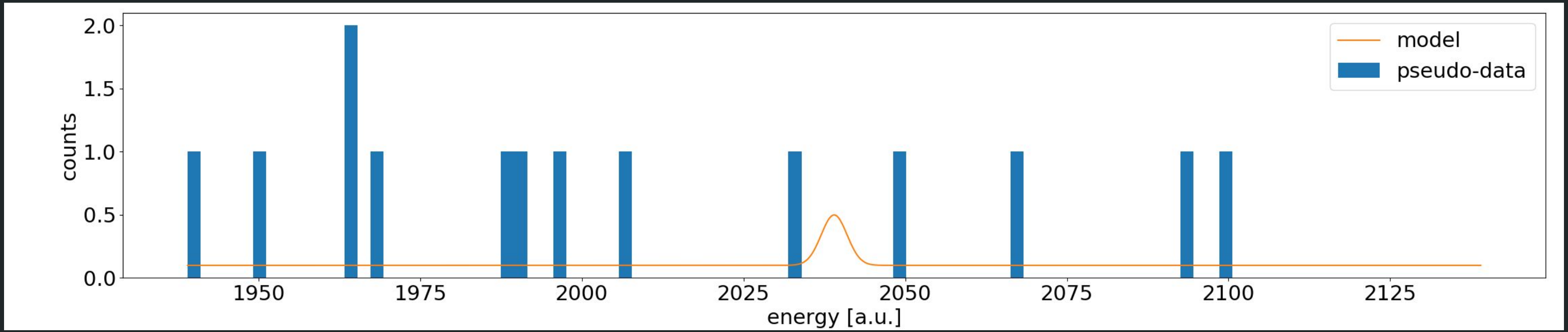
- ➢ basically a **counting experiment**
- ➢ side-bands for background (**on/off** problem)
- ➢ signal at known position (**no look-elsewhere**)

Primary random variables:

- ➢ Energy of the events
- ➢ Total number of events

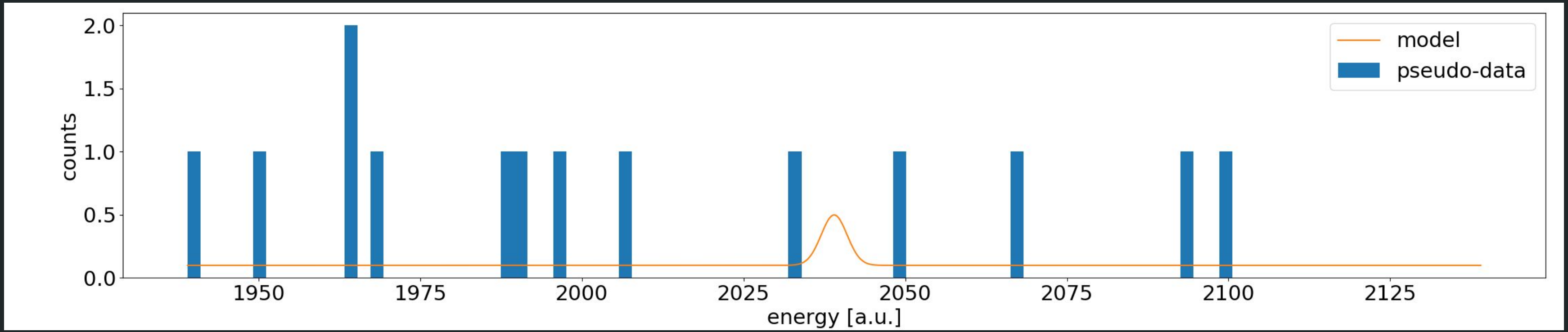
Main parameters and allowed range:

- ➢ expected number of signal events: $\lambda_s \geq 0$
- ➢ expected number of bkg events: $\lambda_b \geq 0$



Random Variables and Likelihoods

Type of Analysis	Random Variables	Likelihood function
Counting	bin content: $\{N_{s+b}, N_b\}$	$L = \text{Poiss}(N_{s+b} \mid \lambda_s+\lambda_b) \cdot \text{Poiss}(N_b \mid \tau\, \lambda_b)$
Extended Unbinned	event energies $\{E_1, E_2, \ldots , E_M\}$ and total number of events M	$L = \text{Poiss}(M \mid \lambda_s+\lambda_b) \cdot \prod_i \text{PDF}(E_i \mid \lambda_s, \lambda_b)$
Binned	bin content: $\{N_1, N_2, \ldots , N_k\}$	$L = \prod_i \text{Poiss}(N_i \mid (\lambda_s+\lambda_b)_i)$



Random Variables and Likelihoods

Type of Analysis	Random Variables	Likelihood function
Counting	bin content: $\{N_{s+b}, N_b\}$	$L = \text{Poiss}(N_{s+b} \mid \lambda_s+\lambda_b) \cdot \text{Poiss}(N_b \mid \tau\, \lambda_b)$
Extended Unbinned	event energies $\{E_1, E_2, \dots, E_M\}$ and total number of events M	$L = \text{Poiss}(M \mid \lambda_s+\lambda_b) \cdot \prod_i \text{PDF}(E_i \mid \lambda_s, \lambda_b)$
Binned	bin content: $\{N_1, N_2, \dots, N_k\}$	$L = \prod_i \text{Poiss}(N_i \mid (\lambda_s+\lambda_b)_i)$

Profile likelihood ratio test statistic

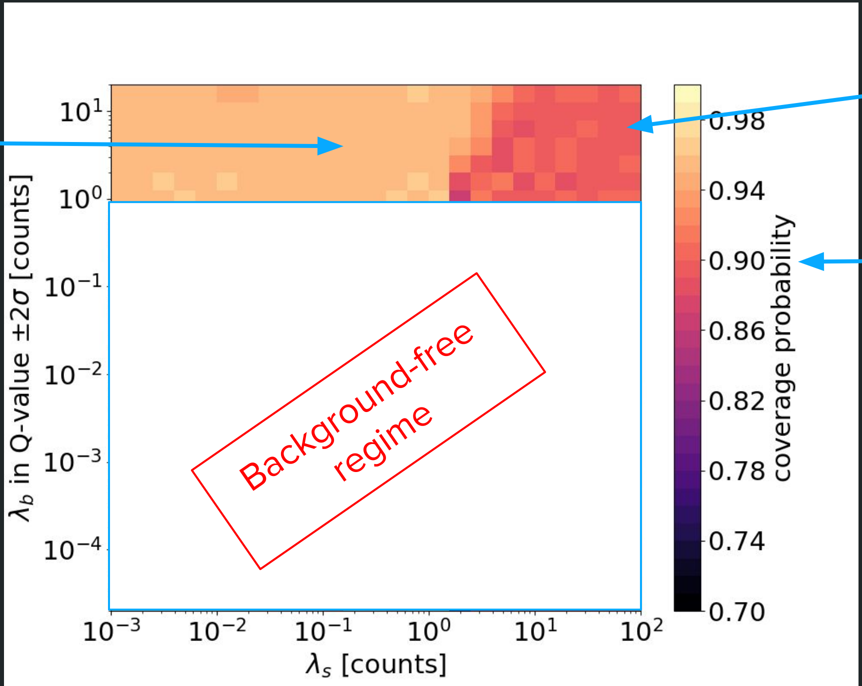
$$T(\lambda_s = X) = -2 \ln \frac{\sup_{\lambda_b} \mathcal{L}(\lambda_s = X, \lambda_b)}{\sup_{\lambda_s, \lambda_b} \mathcal{L}(\lambda_s, \lambda_b)}$$

Challenges of a Frequentist Construction

Coverage map for a 90% C.L. test assuming Wilks' asymptotic distributions (chi-square with 1 dof)

$\lambda_b > 1$
$\lambda_s < \text{sqrt}(\lambda_b)$

Physical border ($\lambda_s \geq 0$) creates overcoverage



$\lambda_s > 1$ and
$\lambda_s > \text{sqrt}(\lambda_b)$

coverage okay

exact coverage

[MA and Hans Niederhausen, in preparation]

(chi-square with 1 dof)

Challenges of a Frequentist Construction

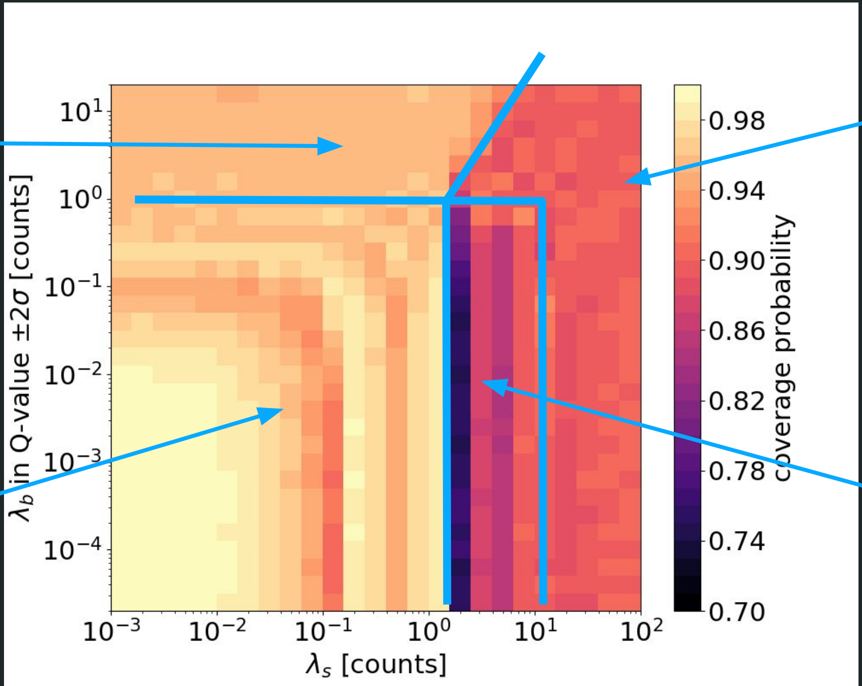
Coverage map for a size-10% test assuming Wilks' asymptotic distributions (chi-square with 1 dof)

$\lambda_b > 1$
$\lambda_s < \text{sqrt}(\lambda_b)$

Physical border ($\lambda_s \geq 0$) creates overcoverage

$\lambda_b < 1$
$\lambda_s < 1$

coverage jumps due to the discrete number of cts



[MA and Hans Niederhausen, in preparation]

$\lambda_s > 1$ and
$\lambda_s > \text{sqrt}(\lambda_b)$

coverage okay

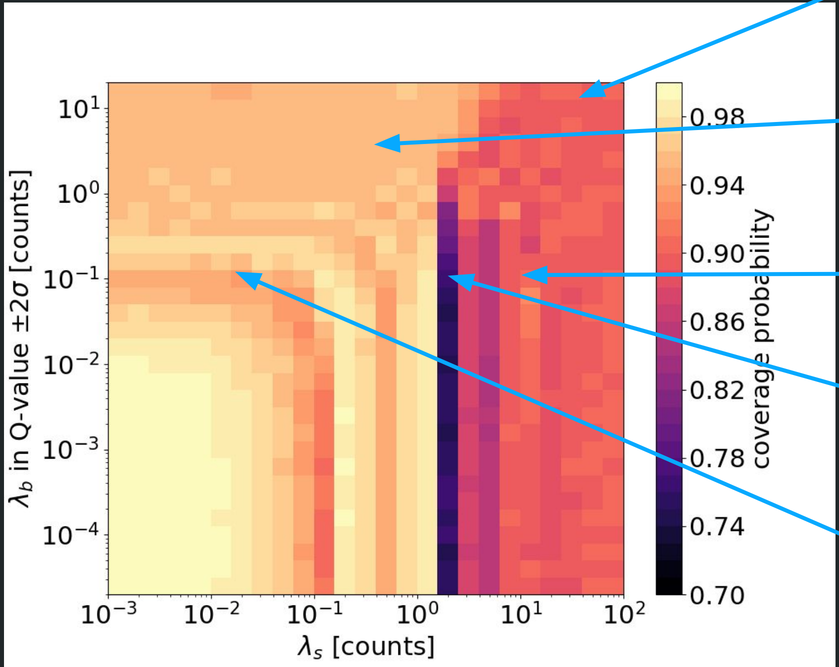
$\lambda_s \sim 1$ and $\lambda_b < 1$

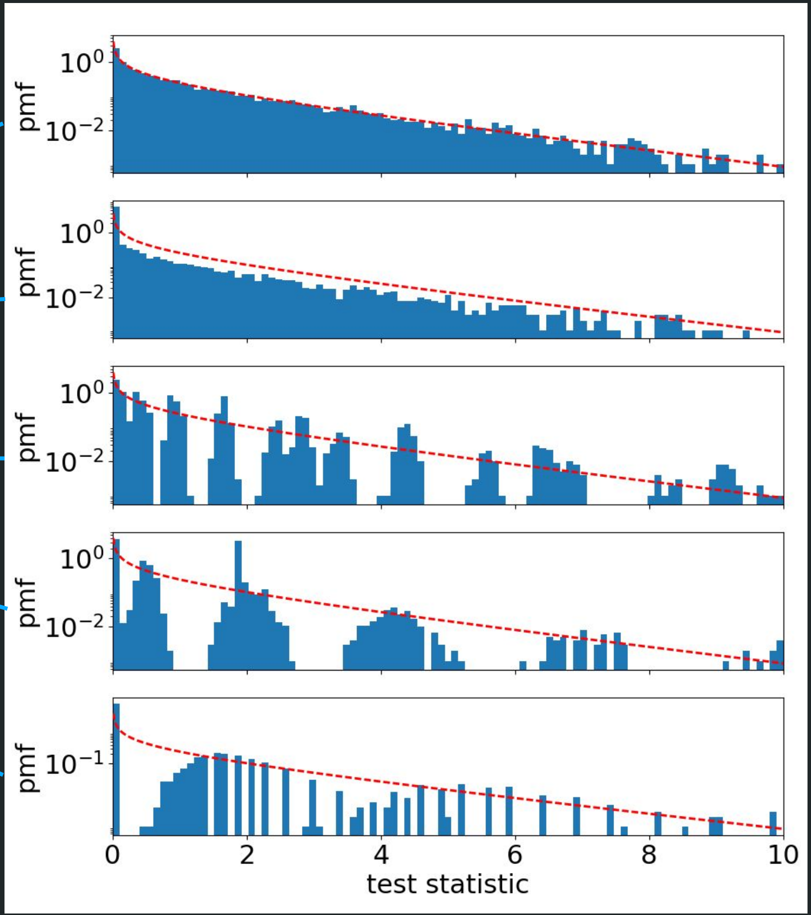
strong undercoverage due to integer number of cts

Parameter space of interest of next-generation experiments!

Challenges for a Frequentist construction

Coverage map for a size-10% test assuming Wilks' asymptotic distributions (chi-square with 1 dof)





[MA and Hans Niederhausen, in preparation]

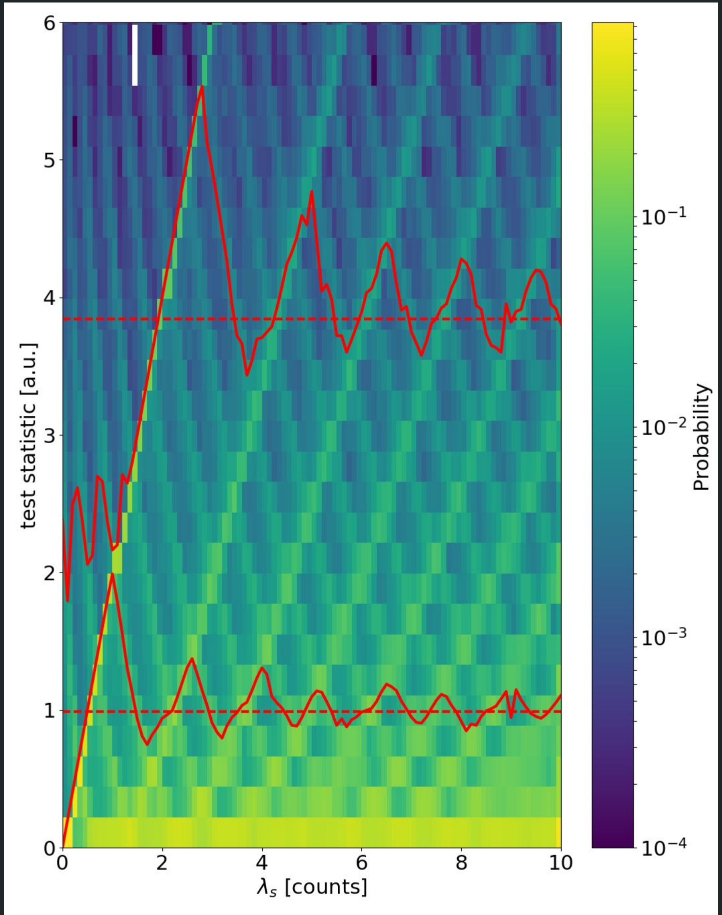
Issues related to Test Statistic Distribution

- ➢ Monte Carlo construction becomes mandatory
- ➢ Test statistic distribution can depend also on nuisance parameters
  - ➢ possibly needed to construct threshold as a function of parameters of interest and nuisance parameters
  - ➢ how to handle p-values?

[FC, Phys.Rev. D57 (1998) 3873-3889]
[Bodhisattva, Walker, Woodroofe, Statist.Sinica 19 (2009) 301-314]

threshold for 95% CL

threshold for 68% CL



[MA and Hans Niederhausen, in preparation]

Sensitivity and Discovery Power

	CL	Concept	How to compute
limit setting	90%	Assuming there is no signal, what is the expected upper limit on the signal expectation?	find signal expectations that would be rejected with a **median** significance of 90% CL assuming the no-signal hypothesis
signal discovery	99.7% (3σ)	Assuming there is a signal, how strong does it has to be to make a discovery?	find signal expectations for which the no-signal hypothesis would be rejected with a **median** significance $\geq 3\sigma$

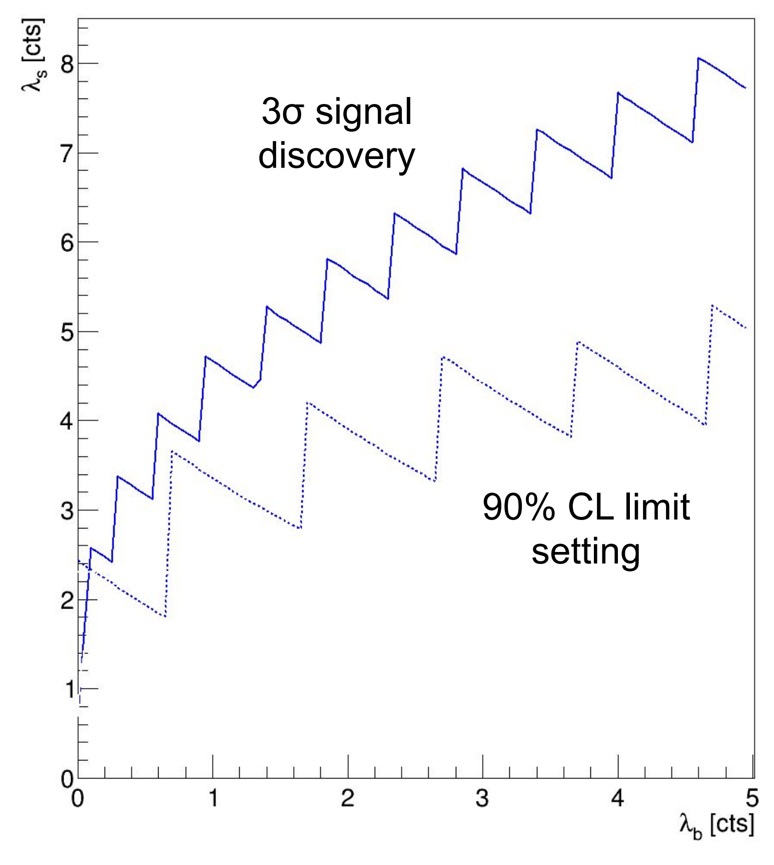
Counting Experiment Sensitivity

Construction à la FC:

- ➢ likelihood ratio with fixed bkg expectation
- ➢ test statistic distributions from pseudo-data
- ➢ median significance (not mean!)

Features:

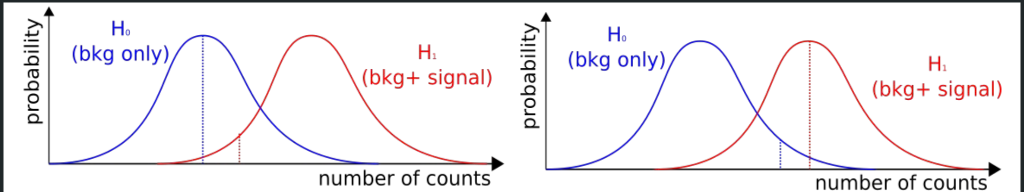
- ➢ jumps in coverage due to integer counts
- ➢ not monotonic functions -> apparent sensitivity improvement when increasing background
- ➢ “better than background free” regime



Counting Experiment Sensitivity

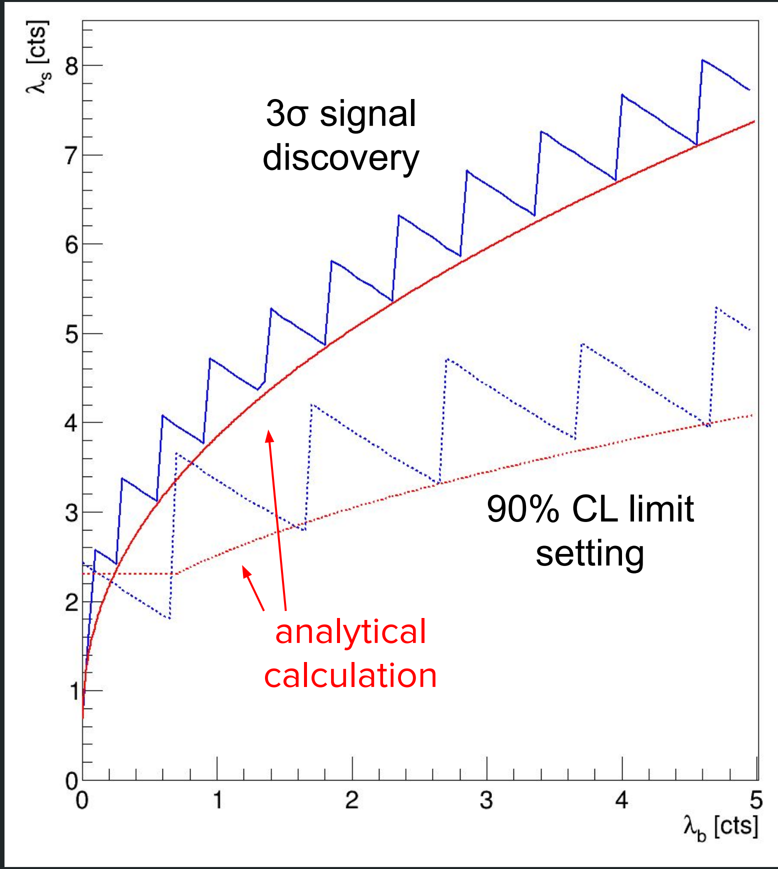
Simple analytical method:

- ➢ based on distributions of expected frequency of observations
- ➢ computed directly using poisson CDF described through gamma functions



Features:

- ➢ touches toy-MC line in the point with exact coverage
- ➢ becomes constant for small signal expectations (median or 3σ quantile of the distribution for H_0 are at zero)

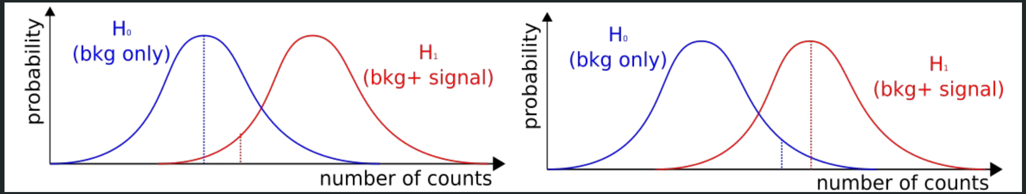


[**M.A.**, G Benato and J A Detwiler, PRD 96, 053001 (2017)]

Counting Experiment Sensitivity

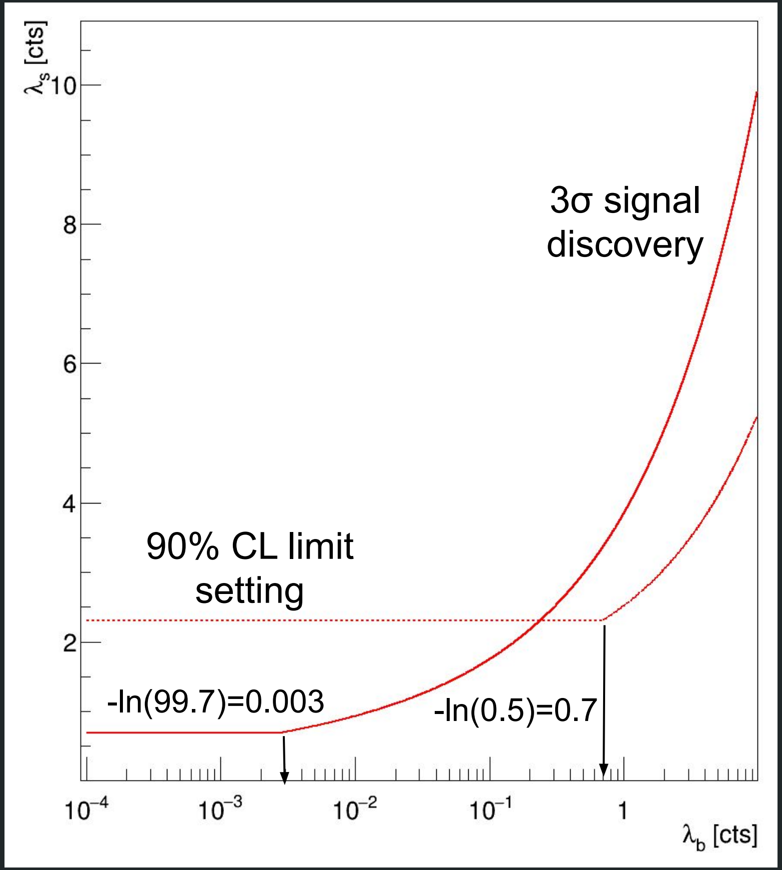
Simple analytical method:

- ➢ based on distributions of expected frequency of observations
- ➢ computed directly using poisson CDF described through gamma functions



Features:

- ➢ touches toy-MC line in the point with exact coverage
- ➢ becomes constant for small signal expectations (median or 3σ quantile of the distribution for H_0 are at zero)

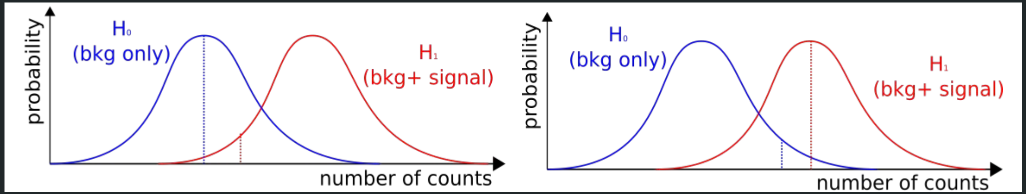


[**M.A.**, G Benato and J A Detwiler, PRD 96, 053001 (2017)]

Counting Experiment Sensitivity

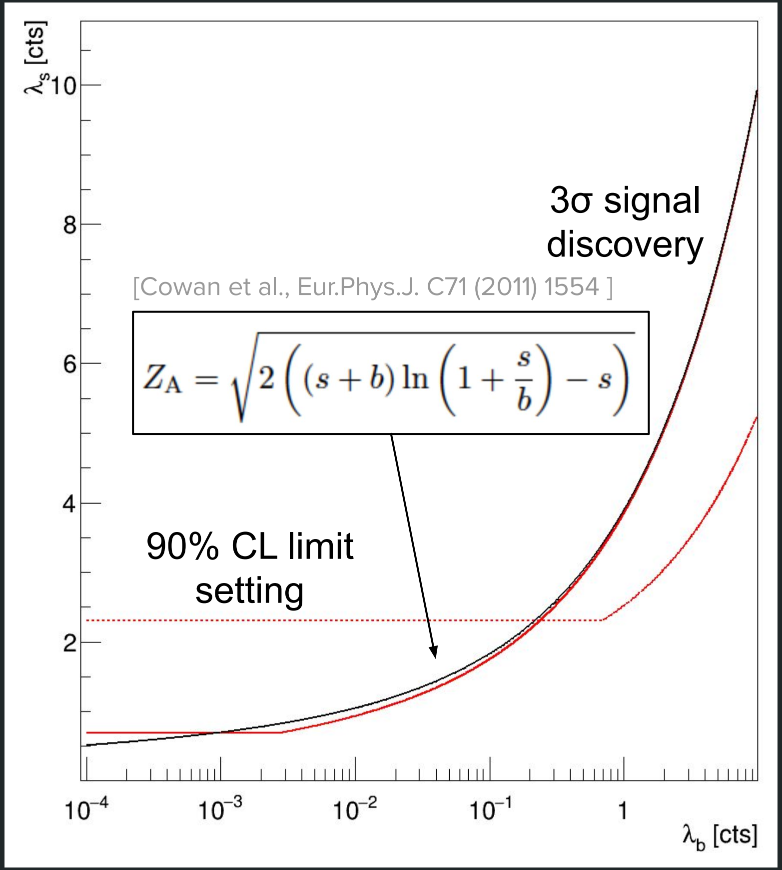
Simple analytical method:

- ➢ based on distributions of expected frequency of observations
- ➢ computed directly using poisson CDF described through gamma functions



Features:

- ➢ touches toy-MC line in the point with exact coverage
- ➢ becomes constant for small signal expectations (median or 3σ quantile of the distribution for H_0 are at zero)



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

Bayesian vs Frequentist

Issues with Bayesian construction:

- ➢ data are weakly informative and results strongly depend on prior
- ➢ Flat and Jeffreys priors can give a hint of the spread of the results
- ➢ Scale-invariant log-prior and other typical choices might lead to not-normalizable posteriors

Large spread of the results from different methods:

- ➢ results quoted for multiple
- ➢ blind analysis is almost the standard

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

Statistical Method in the last PRL of the MAJORANA DEMONSTRATOR	$T_{1/2}$ lower limit 90% prob [10^{25} yr]	$T_{1/2}$ lower limit sensitivity [10^{25} yr]
Counting	1.6	
Unbinned likelihood fit	1.9	2.1
Unbinned likelihood fit & CLs	1.5	1.4
Bayesian flat prior	1.6	
Bayesian Jeffreys prior	2.6	

Systematic Uncertainties

- Statistical uncertainty can affect the result by a factor 2 or 3
- Systematic uncertainties typically affect the result by $\lesssim$10%
- Accounted by nuisance parameters and pull terms (auxiliary data) or priors
- Sources:
  - background modeling
  - energy scale and resolution
  - signal detection efficiency (active volume & analysis cuts)
-

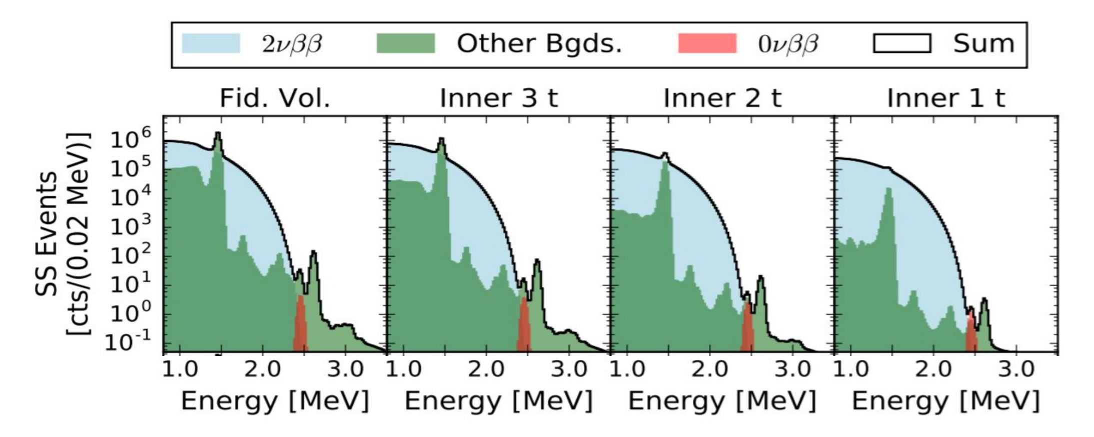
Background modeling is troublesome in case of a **discovery based on 1 single event**:

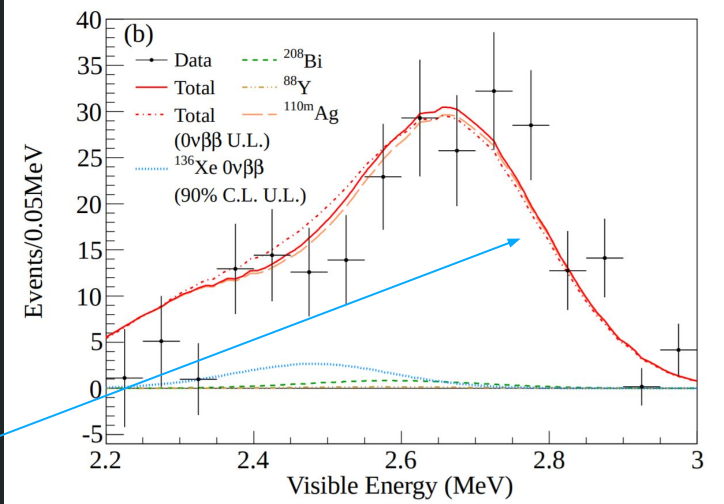
- Gas/Liquid detectors
  - complicated background modeling
  - all components considered?
  - shapes correct within uncertainties?
- Solid state detectors
  - granular design -> many pixels
  - is background homogenous?
  - how to create data sets?

[nEXO pre-CDR, arXiv:1805.11142]

Background Modeling

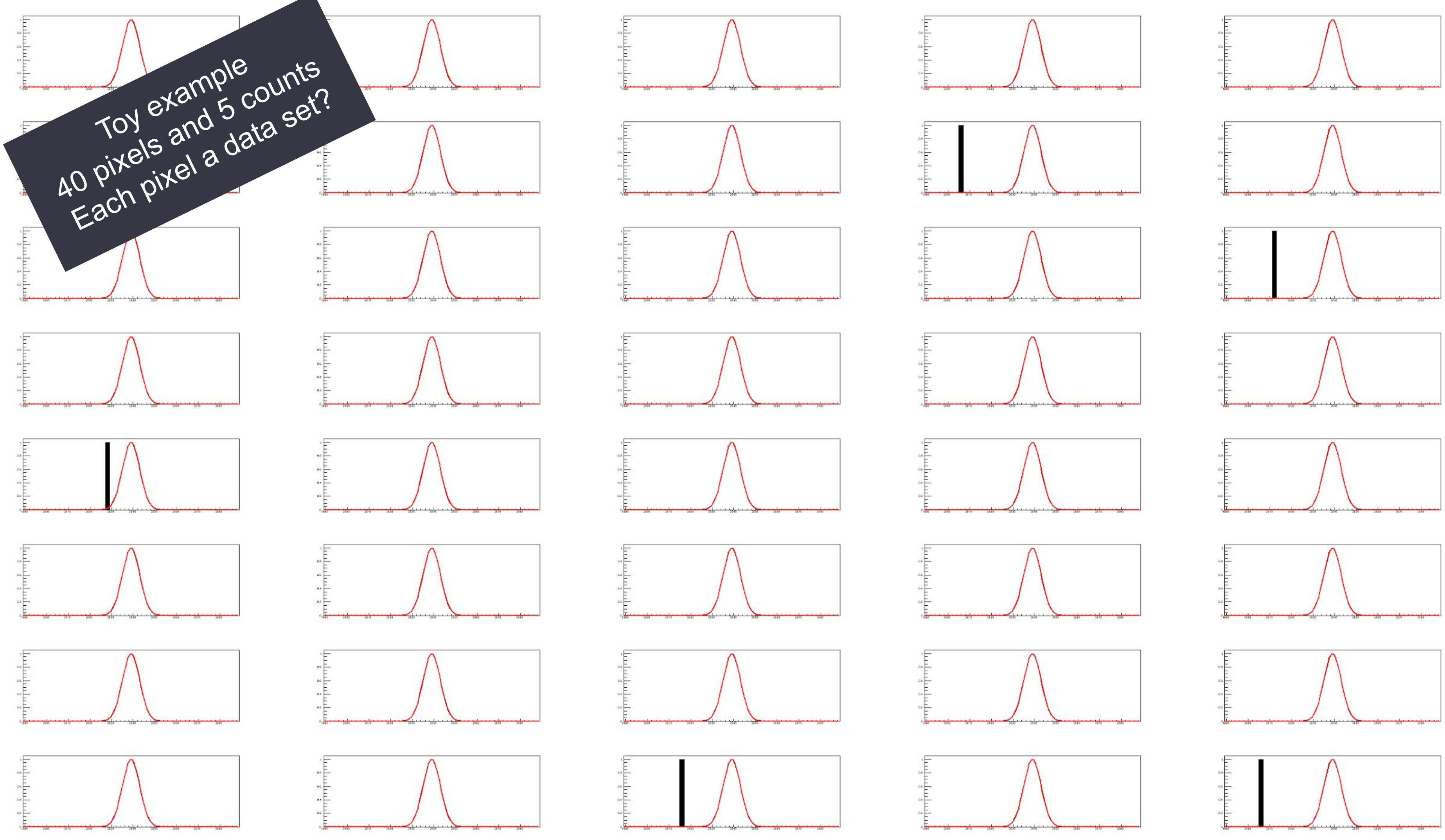
- Gas/Liquid detectors constrain the background using multivariate analyses (event topology, position, pulse shape)
- Contribution due to gamma-rays from radioactivity in the material around the target isotope is under control
- how to exclude other backgrounds due to radioactive isotopes moving within the detector, e.g. Rn-222?





unexpected Ag-110m

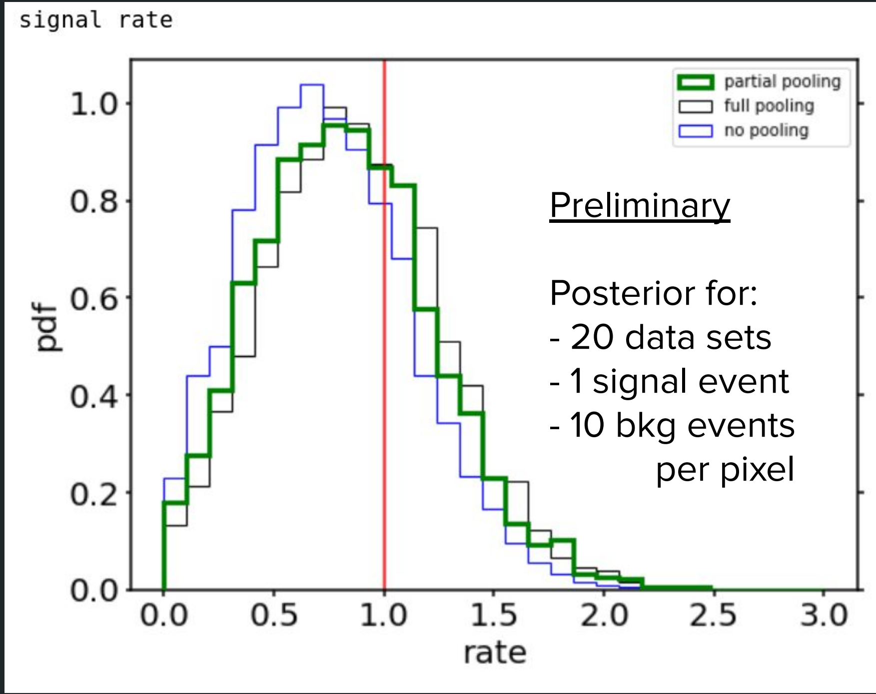
[KamLAND-Zen, Phys.Rev.Lett.110]


Toy example
40 pixels and 5 counts
Each pixel a data set?

Inhomogeneous Background Levels

The problem can be addressed using hierarchical models:

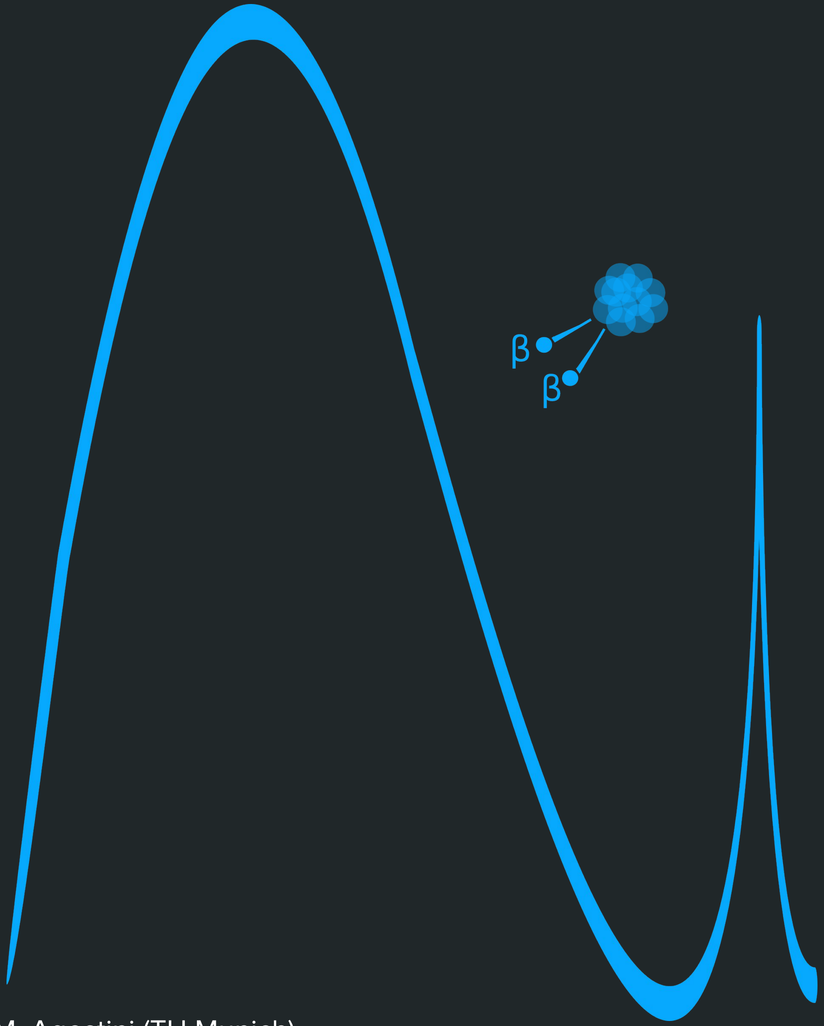
- ➢ each data set has a $\lambda_s^i \geq 0$ and a $\lambda_b^i \geq 0$
- ➢ λ_s^i are fully correlated (common signal)
- ➢ λ_b^i can be
  - ➢ non correlated
  - ➢ partially correlated
  - ➢ fully correlated
- ➢ partial correlations can have the form e.g. of a Gaussian with centroid and variance defined by the data set itself + pull terms



[MA and Hans Niederhausen, in preparation]

Under study:

- ➢ impact of partial vs full pooling
- ➢ dependence of results from form of correlations and pull terms
- ➢ probability distribution of test statistic and dependence on nuisance parameters



Outlook

- $0\nu\beta\beta$ decay is a portal to new physics and experiments aim to be **background-free**
  - claim a **discovery** based on a **single event**

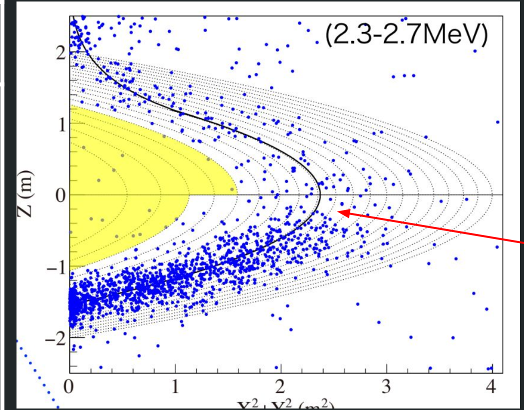
- Search for a peak with background still poses challenges in the “Deep Poisson” regime:
  - popular asymptotic methods are not valid
  - test statistic distributions might depend on nuisance parameters
  - data set definition not trivial

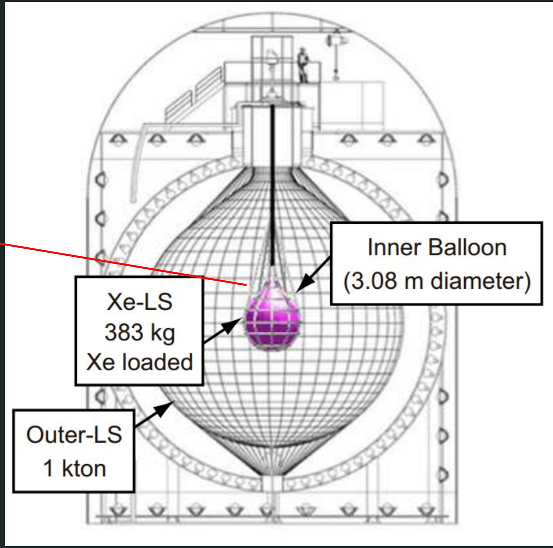
- Important to shift focus towards a discovery analysis and define in advance how to deal with the background modeling systematics

Backup slides

KamLAND-Zen

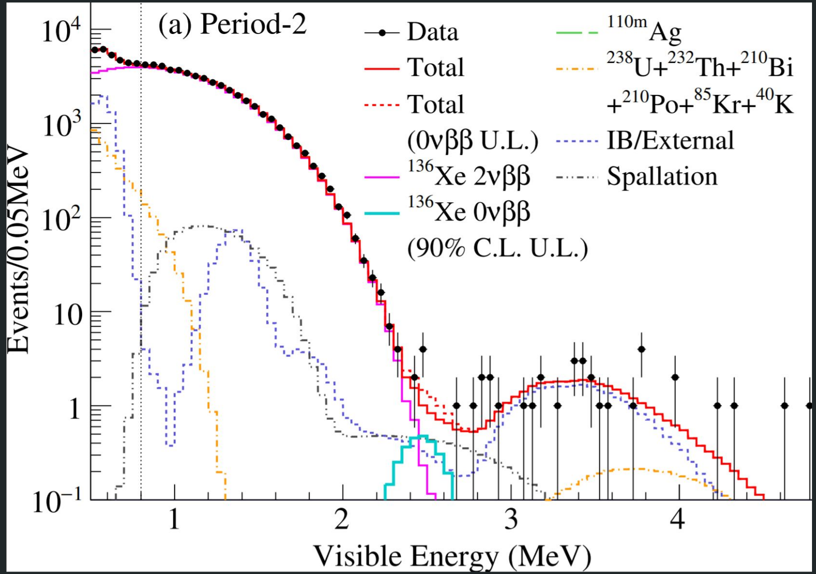
Location	Kamioka, Japan
Isotope	^{136}Xe [$Q_{\beta\beta}$=2458 keV]
Technology	Xe-loaded liquid scintillator
Isotope Mass	350 kg
$0\nu\beta\beta$ efficiency	16%
Resolution [σ]	100-120 keV
Latest results	$T_{1/2} > 1.1\cdot10^{26}$ yr (90% CL)
Sensitivity	$T_{1/2} > 5.6\cdot10^{25}$ yr (90% CL)

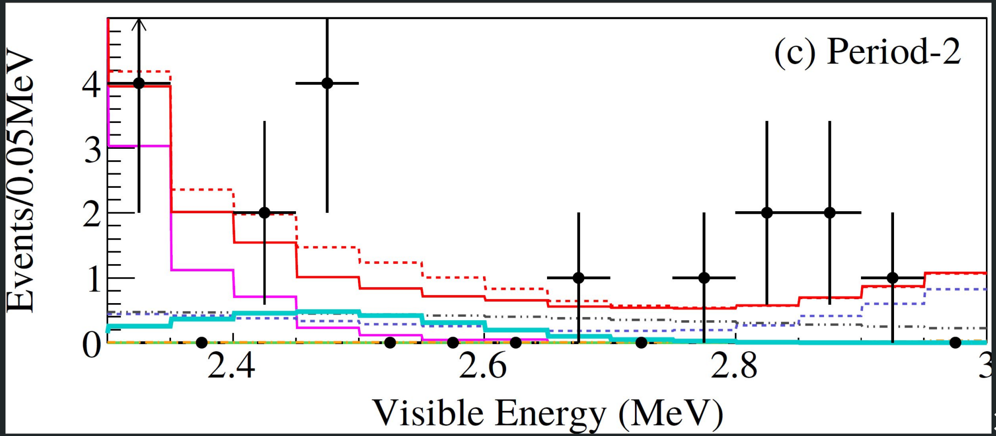




Frequentist likelihood fit

- Multivariate: E vs R
- Wilks' approximation tested with toy MC

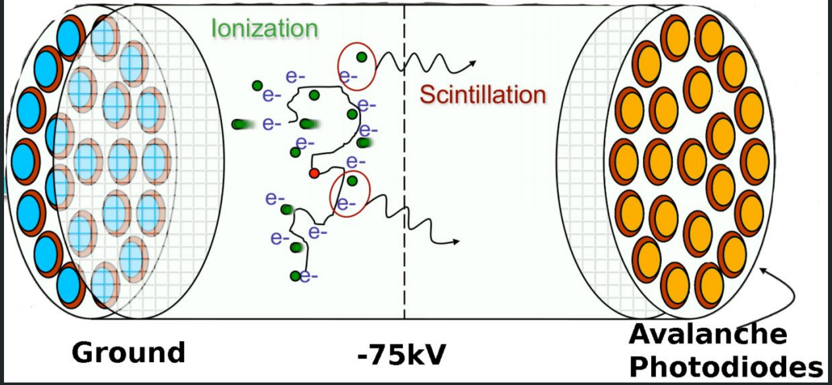




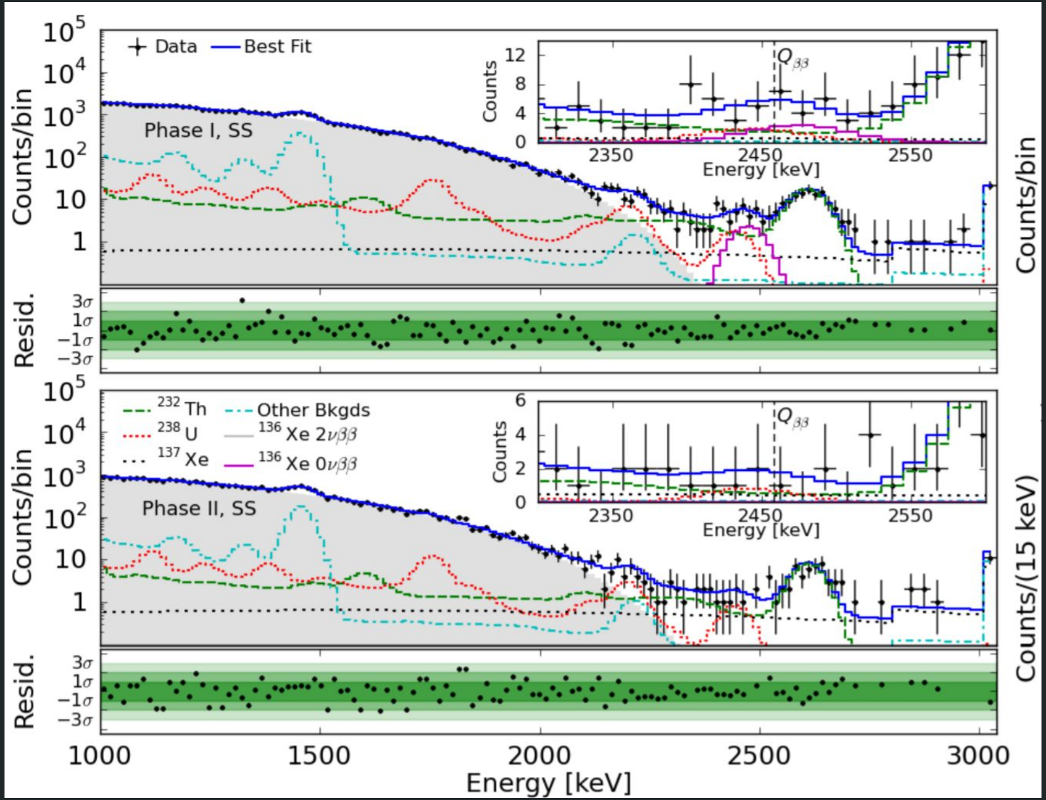
[Phys.Rev.Lett. 117 (2016) no.10, 109903]

EXO-200

Location	WIPP, New Mexico, USA
Isotope	^{136}Xe [$Q_{\beta\beta}$=2458 keV]
Technology	TPC with liquid Xe
Isotope Mass	76 kg
$0\nu\beta\beta$ efficiency	80%
Resolution [σ]	34 keV
Latest results	$T_{1/2} > 1.8 \cdot 10^{25}$ yr (90% CL)
Sensitivity	$T_{1/2} > 3.7 \cdot 10^{25}$ yr (90% CL)



[Phys.Rev.Lett. 120 (2018) no.7, 072701]

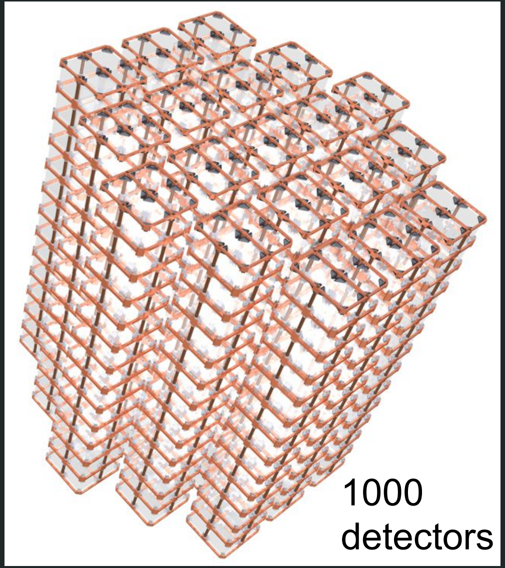


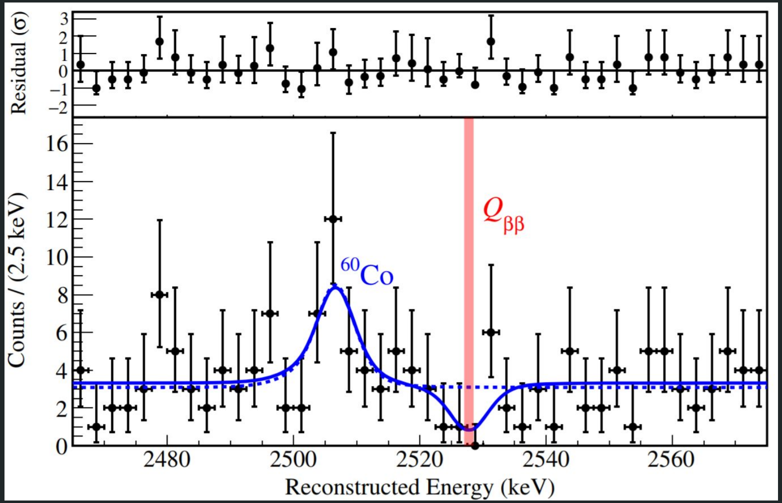
Frequentist binned likelihood fit:

- multivariate (energy, position, TMVA observables)
- Wilks' approximation valid (coverage tested)

CUORE

Location	LNGS, Italy
Isotope	^{130}Te [$Q_{\beta\beta}$=2527 keV]
Technology	Cryogenic calorimeters
Isotope Mass	206 kg
$0\nu\beta\beta$ efficiency	68%
Resolution [σ]	3.3 keV
Latest results	$T_{1/2} > 1.5\cdot10^{25}$ yr (90% CL)
Sensitivity	$T_{1/2} > 0.7\cdot10^{25}$ yr (90% CL)





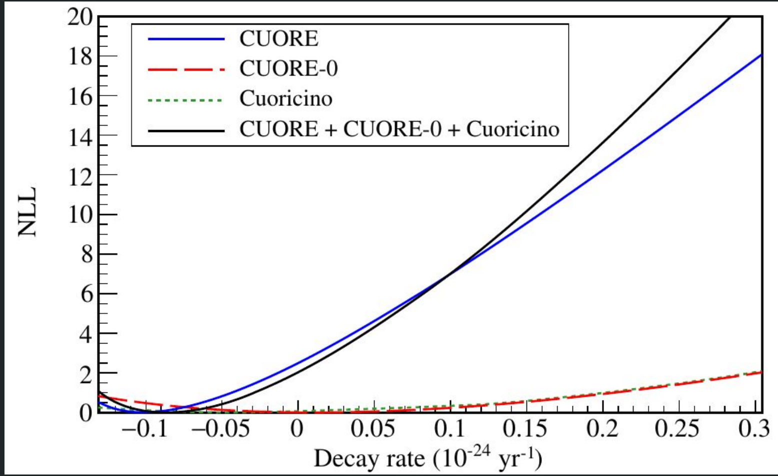
[Phys. Rev. Lett. 120, 132501 (2018)]

Bayesian:

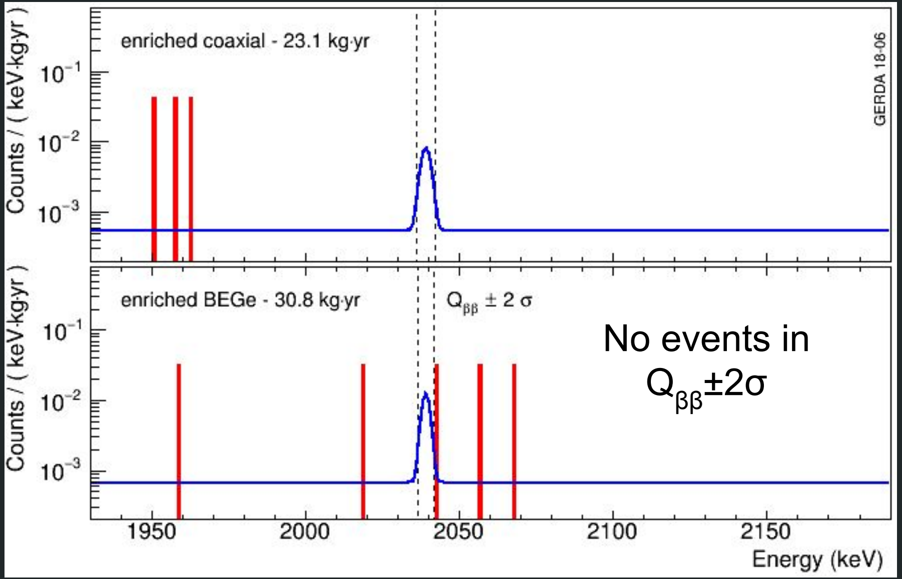
- ➢ flat prior
- ➢ profiling instead of marginalization

Frequentist:

- ➢ bounded profile likelihood
- ➢ Wilks approximation



GERDA	
Location	LNGS, Italy
Isotope	^{76}Ge [$Q_{\beta\beta}$=2039 keV]
Technology	Semiconductor Ge detectors
Isotope Mass	35 kg
$0\nu\beta\beta$ efficiency	65%
Resolution [σ]	1.3 keV
Latest results	$T_{1/2} > 0.9\cdot10^{26}$ yr (90% CL)
Sensitivity	$T_{1/2} > 1.1\cdot10^{26}$ yr (90% CL)



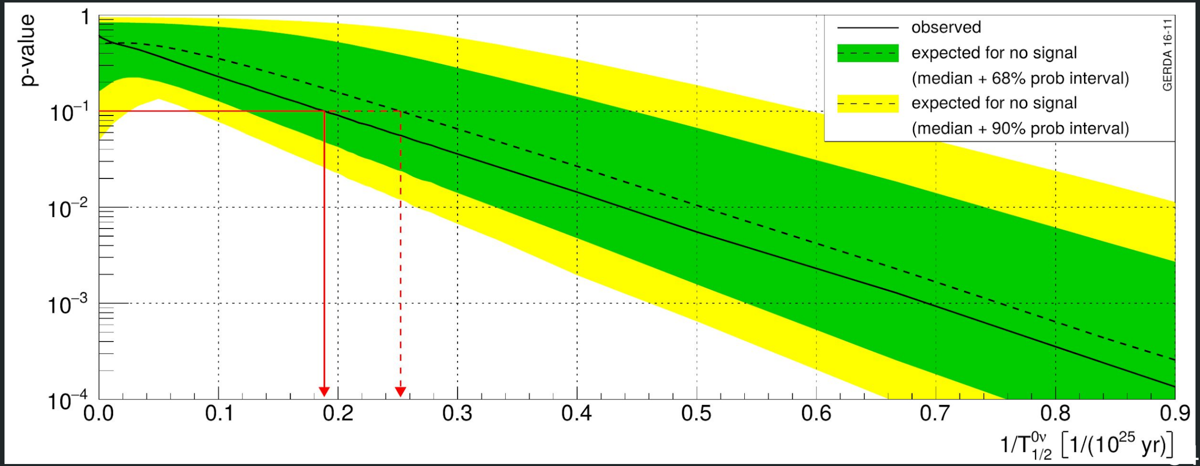
[Nature 544 (2017) 47]

Frequentist:

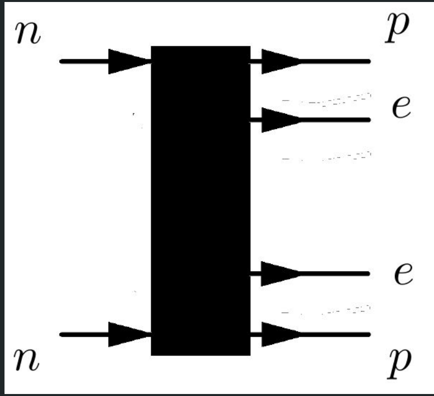
- ➢ extended unbinned likelihood
- ➢ profile likelihood
- ➢ FC construction (only for best fit value of nuisance parameters)

Bayesian:

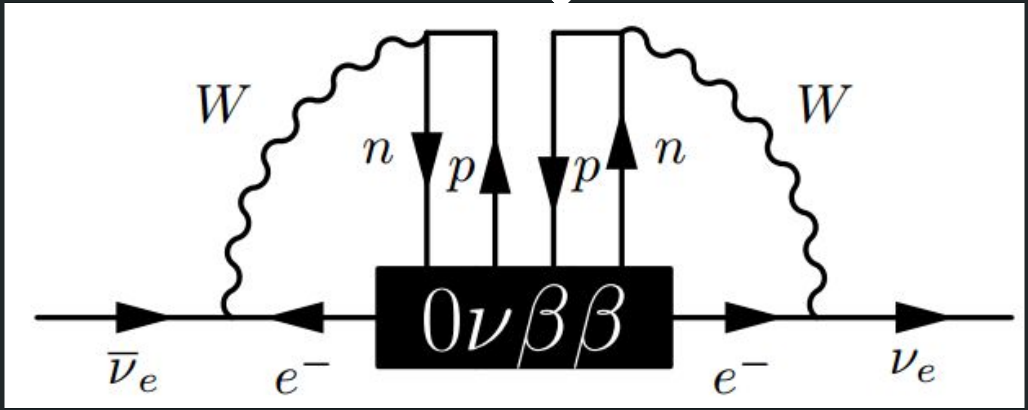
- ➢ flat prior



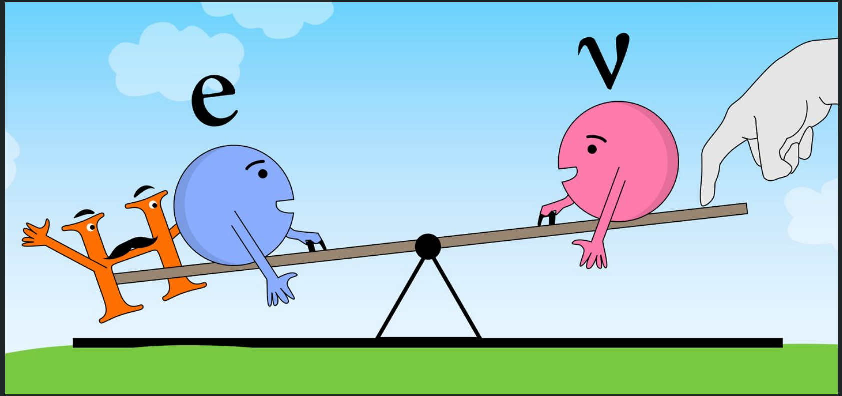
$0\nu\beta\beta$ and ν Mass Origin



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



[Schechter, Valle, PRD 25 (1982) 2951]



[www.symmetrymagazine.org]

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

Analytical computation of sensitivity

Signal discovery

- ➢ Find the number of counts $C_{3\sigma}$ such that: $CDF(C_{3\sigma}|B) = erf(3/\sqrt{2})$
- ➢ Solve: $CDF(C_{3\sigma} \mid S_{3\sigma} + B) = 50\%$
- ➢ $C_{3\sigma}$ 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_{med} : $CDF(C_{med}|B) = 50\%$
- ➢ Solve: $CDF(C_{med}| S_{90\%CL} + C_{med}) = 10\%$

[more in M.A. et al., Phys.Rev. D96 (2017) no.5)]

The counting experiment with a profile likelihood

if B is perfectly known:

B := background expectation

S := signal expectation

N := number of cts in ROI

L(s) = Pois(N|S+B)

t(s) = -2 [Pois (N|S+B) - Pois(N|S_{best}+B)]

with S_{best}=max(0,N-B)

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

τ := side band width / ROI width

(for GERDA: **τ** = 220/6 ~40)

M := number of cts in side band

L(s) = Pois(N|S+B) * Pois(M|tB)

t(s) =-2[LogPois(N|S+B_{cond})+LogPois(M|τB_{cond})

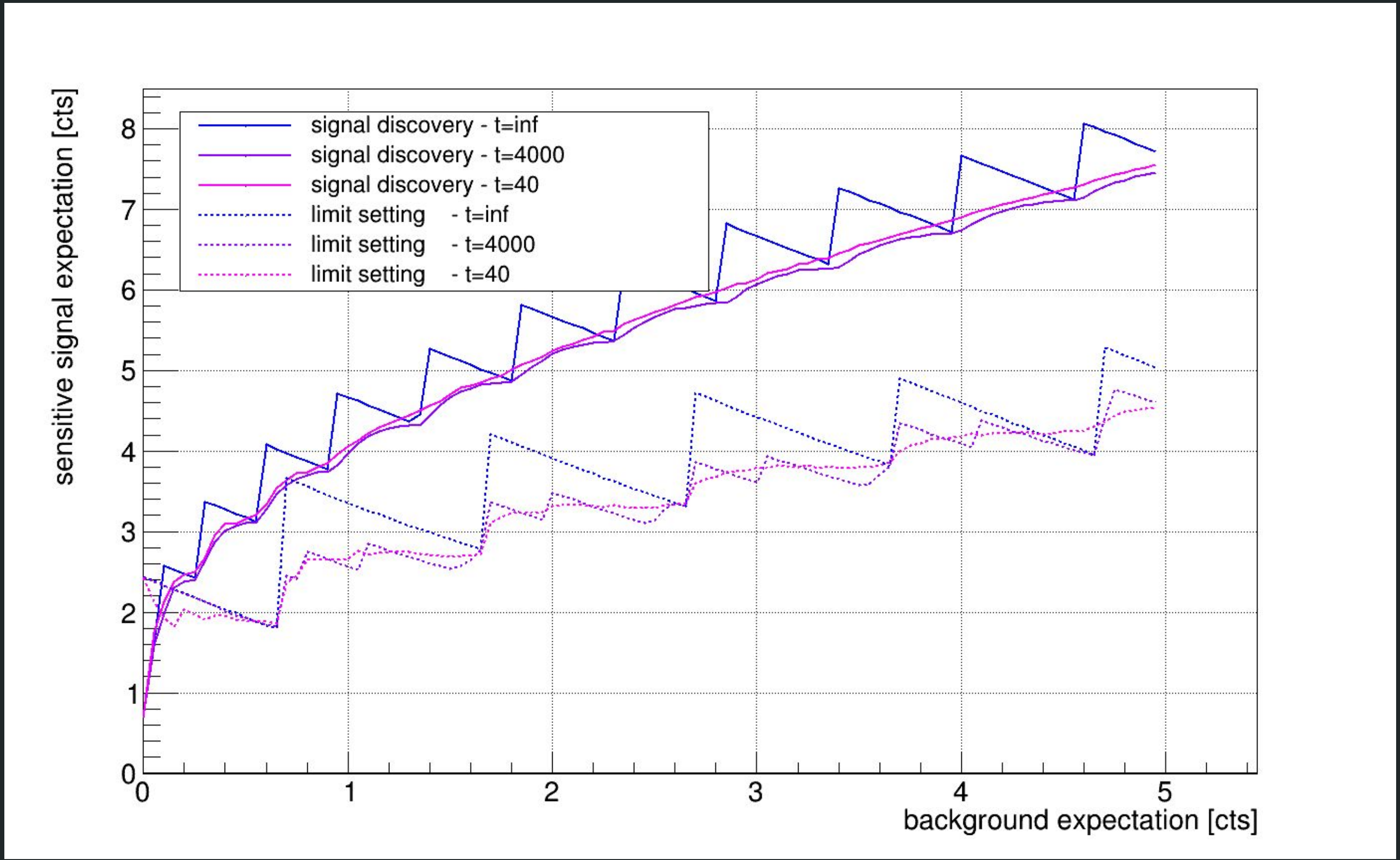
-LogPois(N|S_{best}+B_{best}) - LogPois(M|τB_{best})]

with: S_{best}=N-M/τ,

B_{best}= M/τ

B_{cond} = N+M-(1+t)S+sqrt((N+M-(1+t)S)^2 + 4(1+t)SM)]/ [2(1+t)]

Sensitivity with background uncertainty



A portal to Physics beyond the Standard Model

Effective field theory - General Expression for dim 5/7/9 operators:

$$\Gamma = \frac{1}{T_{1/2}} = K_1 \frac{m_{\beta\beta}^2}{m_e^2} + K_2 \left(\frac{v}{\Lambda}\right)^6 + K_3 \left(\frac{v}{\Lambda'}\right)^{10} + \cdots$$

Effective Majorana Mass
Function of neutrino masses, angles and phases

New Physics Scale
($\lambda \gtrsim 100$ TeV, $\lambda' \gtrsim 10$ TeV)

[Cirigliano et al. JHEP 12 082 (2017)]

- **$T_{1/2}$** is connected to neutrino physics and other BSM processes (heavy sterile neutrinos, SUSY, ...)
- **$T_{1/2}$** is for **$0\nu\beta\beta$** decay what the **collision energy** is for **LHC**
- Experimental constraints: **$T_{1/2} > 10^{26}$ yr**, i.e. more than a **million trillion** times the age of the Universe!