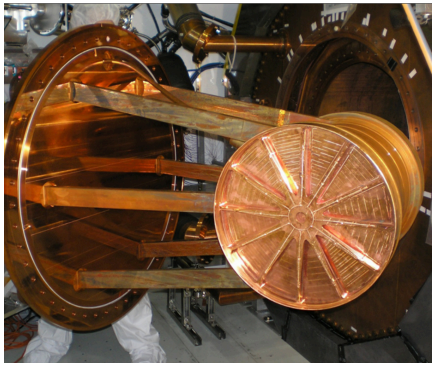
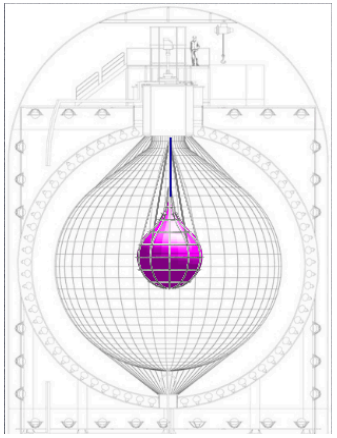
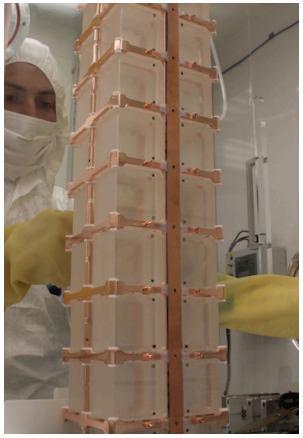
Statistical issues for neutrinoless double beta decay

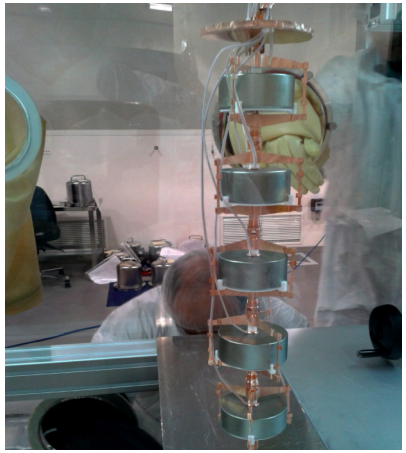
David Moore, Yale University

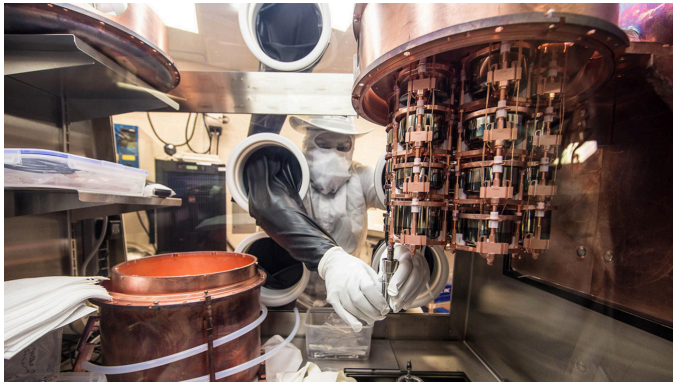
PhyStat-nu, Fermilab, September 20, 2016

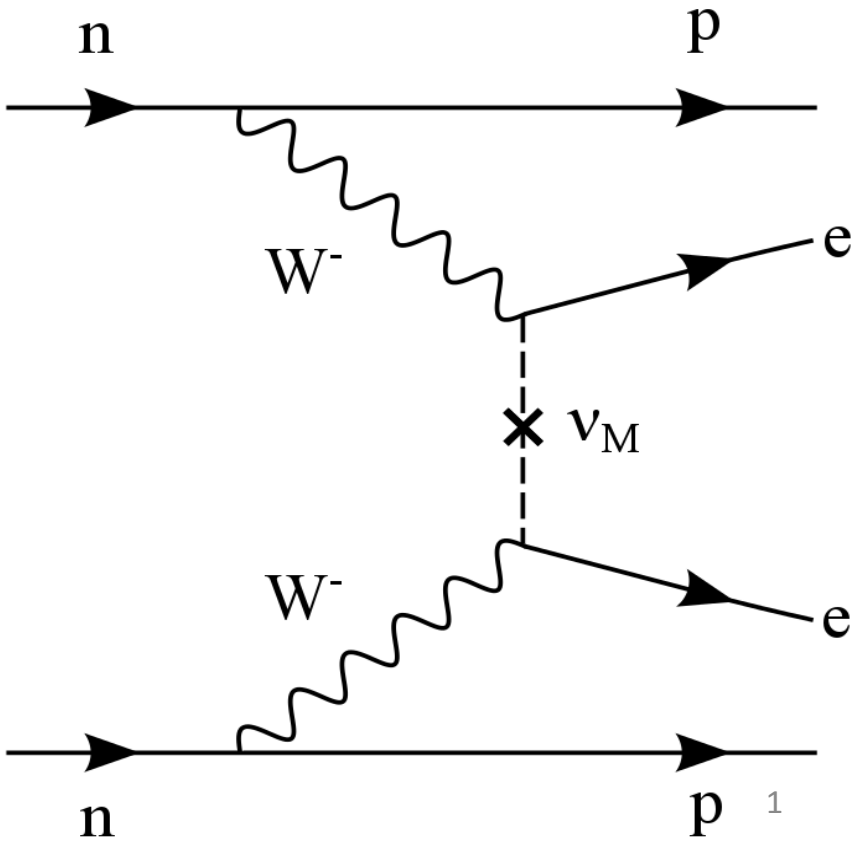








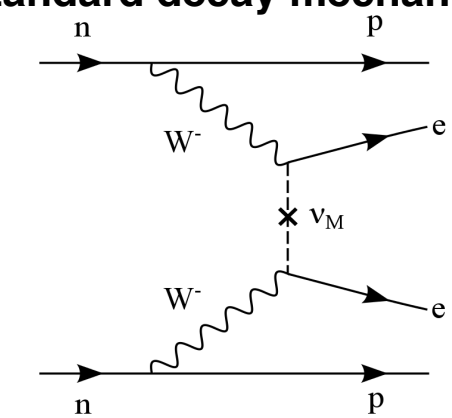




Introduction

- Neutrinoless double beta decay (0νββ) is an extremely rare decay ($T_{1/2} > 10^{26}$ yr)
- Can occur if neutrinos are Majorana particles ($\nu = \bar{\nu}$)
- Corresponding Standard Model process (2νββ) has been observed in several isotopes
- Statistical issues:
  - Low number of events
  - Background model systematics
  - Combination between isotopes

Standard decay mechanism:



Measured 2νββ half-lives:

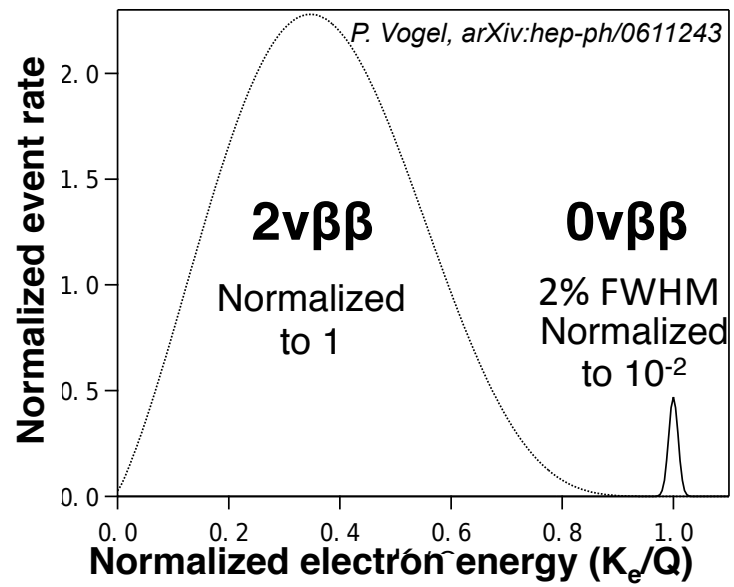
Isotope	Average/best half-life (years)
^{48}Ca	$4.4^{+0.6}_{-0.5} \times 10^{19}$ (79)
^{76}Ge	$1.84^{+0.14}_{-0.10} \times 10^{21}$ (28)
^{82}Se	$(0.92 \pm 0.07) \times 10^{20}$ (79)
^{96}Zr	$(2.35 \pm 0.21) \times 10^{19}$ (54)
^{100}Mo	$(7.1 \pm 0.4) \times 10^{18}$ (79)
^{100}Mo (0^+_1)	$5.9^{+0.8}_{-0.6} \times 10^{20}$ (79)
^{116}Cd	$(2.8 \pm 0.2) \times 10^{19}$ (79)
^{128}Te	$(1.9 \pm 0.4) \times 10^{24}$ (79)
^{130}Te	$(7.0 \pm 1.4) \times 10^{20}$ (61)
^{136}Xe	$(2.30 \pm 0.12) \times 10^{21}$[b]
^{150}Nd	$(9.11 \pm 0.68) \times 10^{18}$ (62)
^{150}Nd (0^+_1)	$1.33^{+0.45}_{-0.26} \times 10^{20}$ (79)
^{238}U	$(2.0 \pm 0.6) \times 10^{21}$ (79)
^{130}Ba, ECEC (2ν)	$(2.2 \pm 0.5) \times 10^{21}$ (79)

Ann. Rev. Nucl. Part. Sci, 63, 503 (2013)

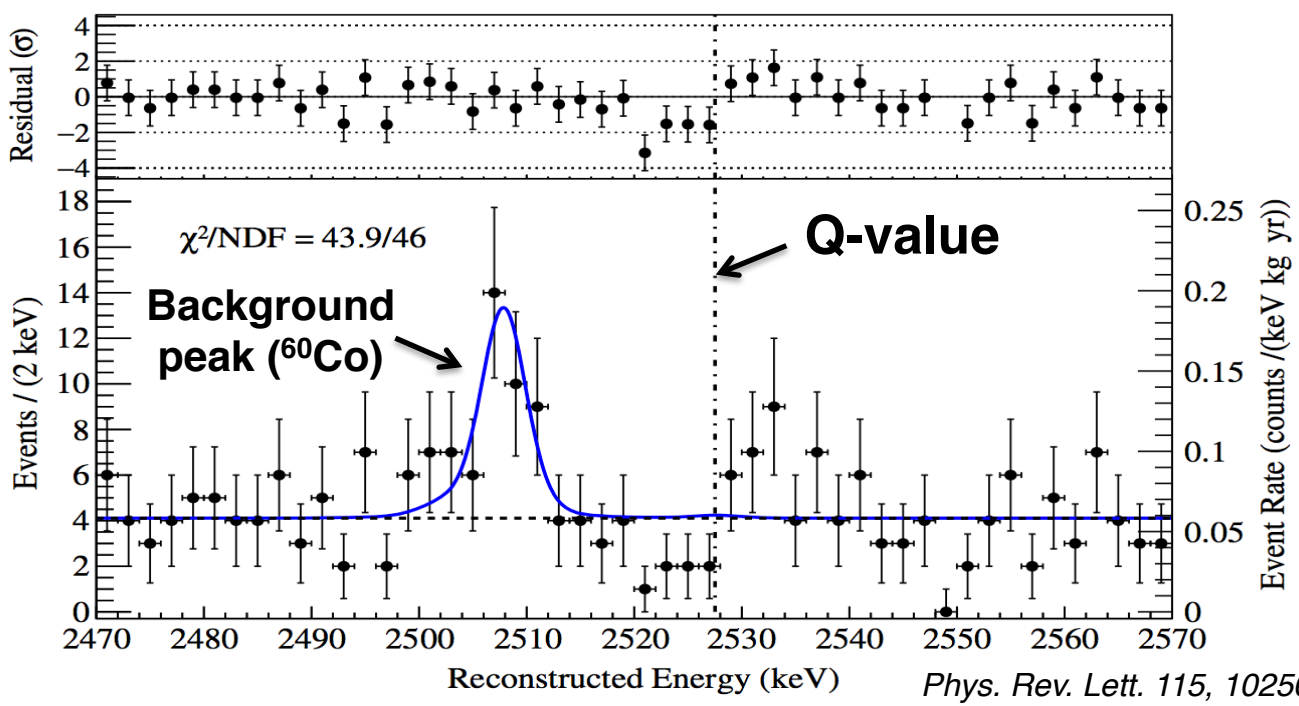
0νββ detection techniques

- Signature of 0νββ would be a peak at the full decay energy (Q-value)
- "Canonical" detection technology:
  - Use high resolution detector to search for peak at known energy
  - Ge ionization detectors (GERDA, Majorana), cryogenic bolometers (CUORE)

ββ energy spectrum:



Energy spectrum near Q-value, CUORE-0 (2015):

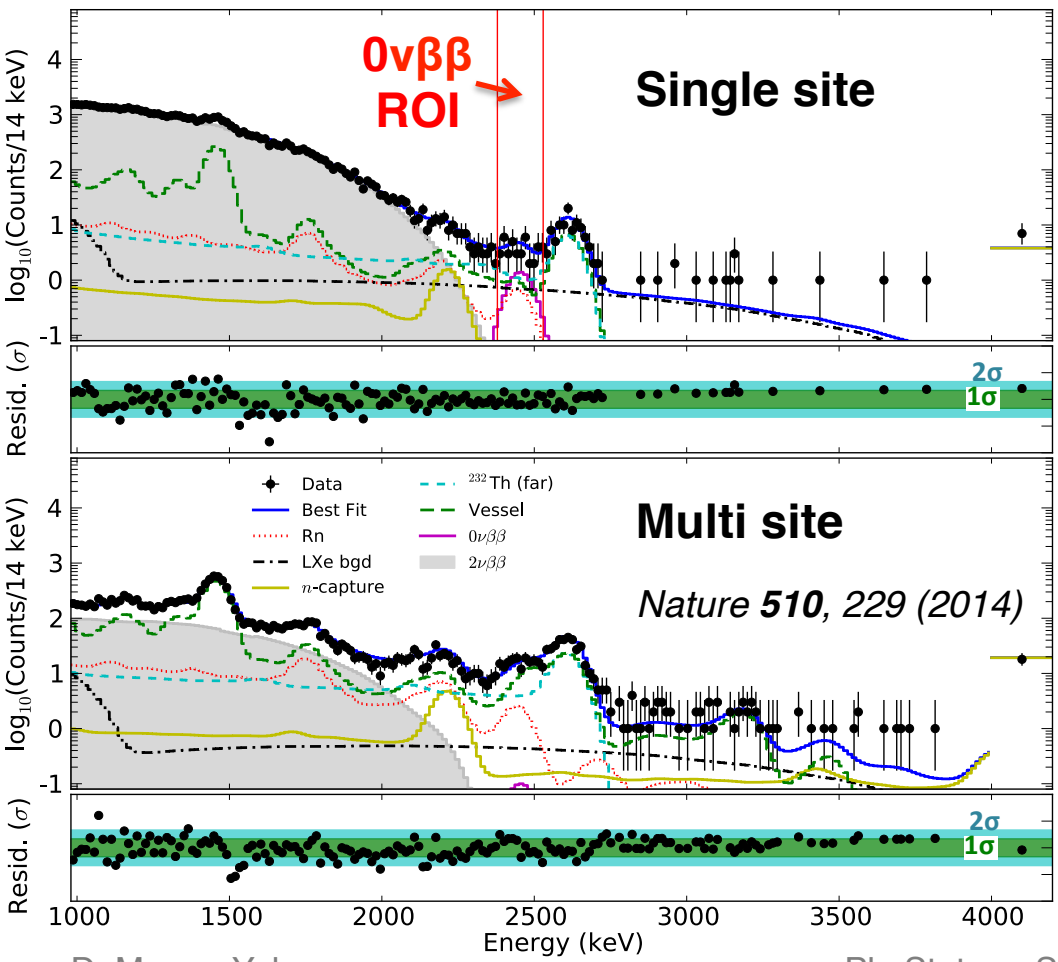


Phys. Rev. Lett. 115, 102502 (2015)

0νββ detection techniques (cont'd)

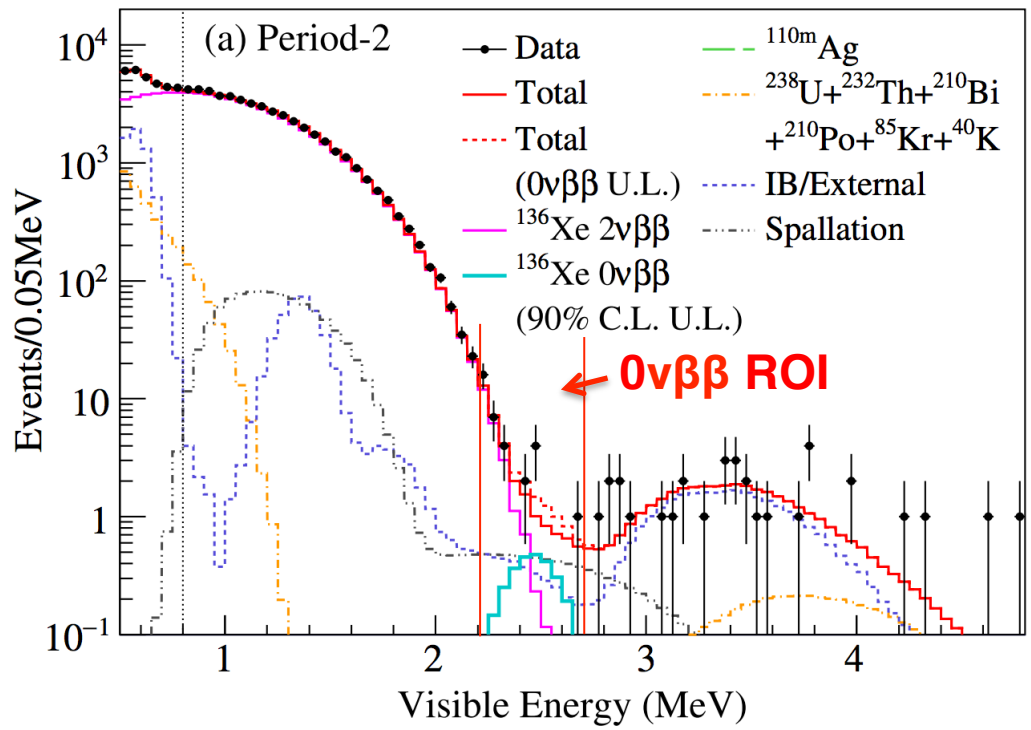
- Alternative technique uses large, self-shielded detectors (EXO, KamLand-Zen, SNO+)
- Poorer energy resolution, but multidimensional fit to event location, topology, and energy used to constrain 0νββ

EXO-200 spectrum (2014):



*Nature **510**, 229 (2014)*

KamLAND-Zen spectrum (2016):

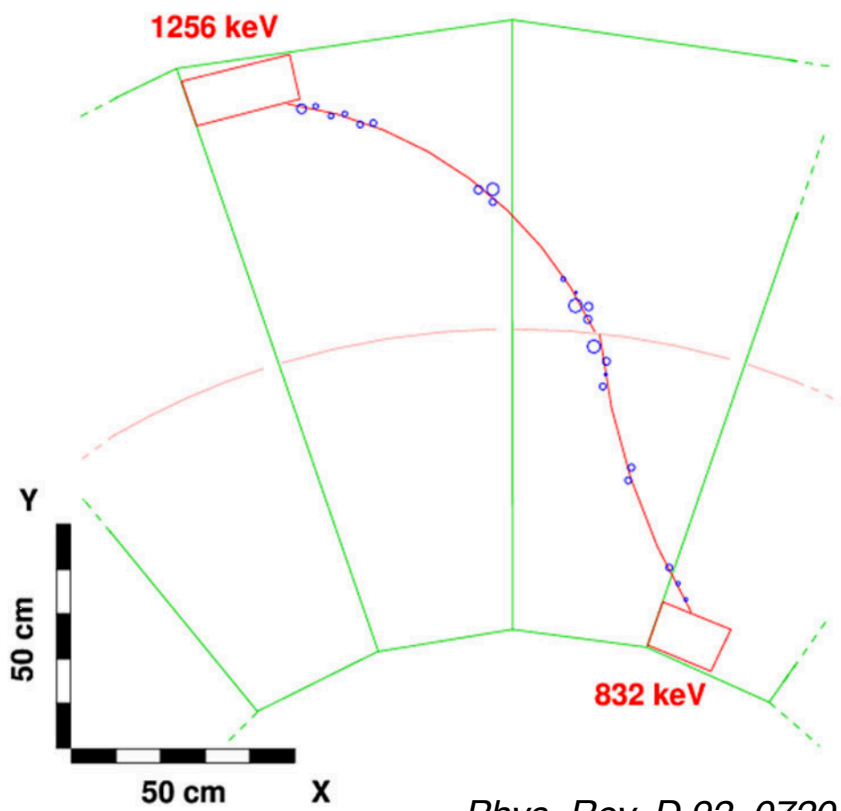


*PRL **117**, 082503 (2016)*

0νββ detection techniques (cont'd)

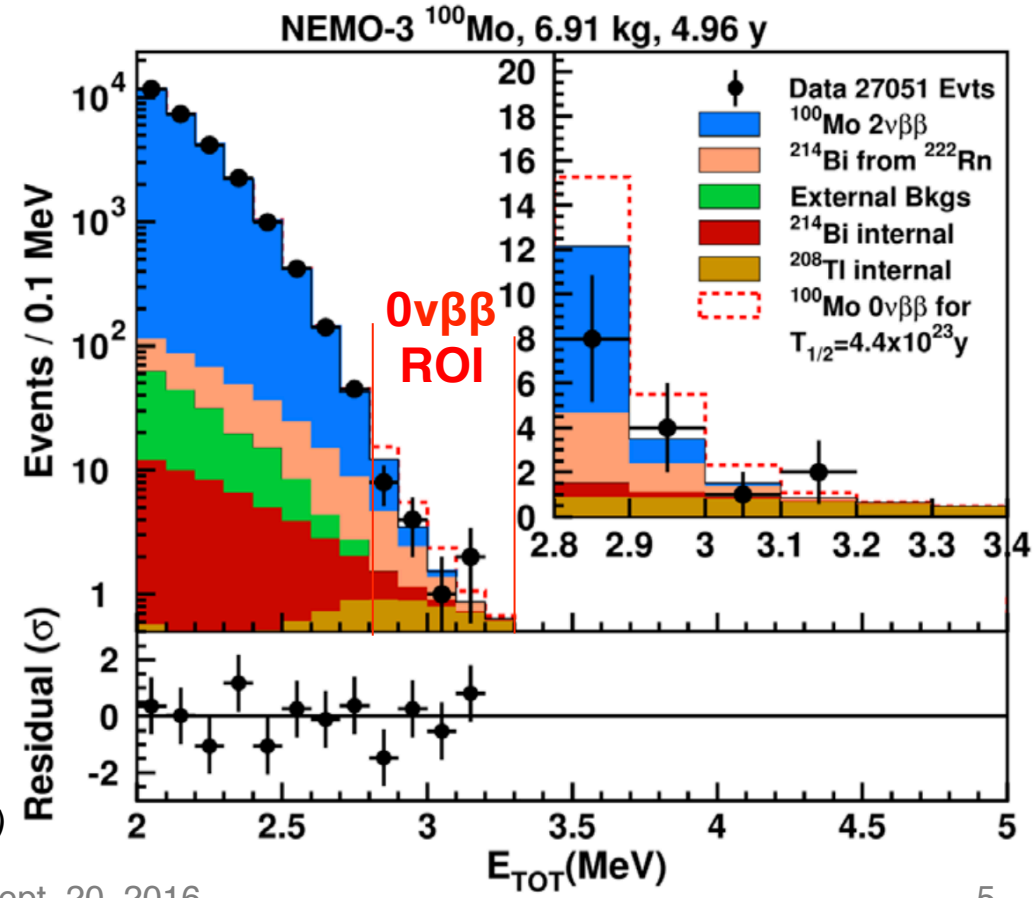
- Final technique uses detailed topological reconstruction to identify characteristic double track of ββ decay (e.g. SuperNEMO, NEXT)
- Background rejection not perfect, so typically still require spectral fits that are similar to large, non-tracking detectors

Example of topological reconstruction (NEMO-3):



Phys. Rev. D 92, 072011 (2015)

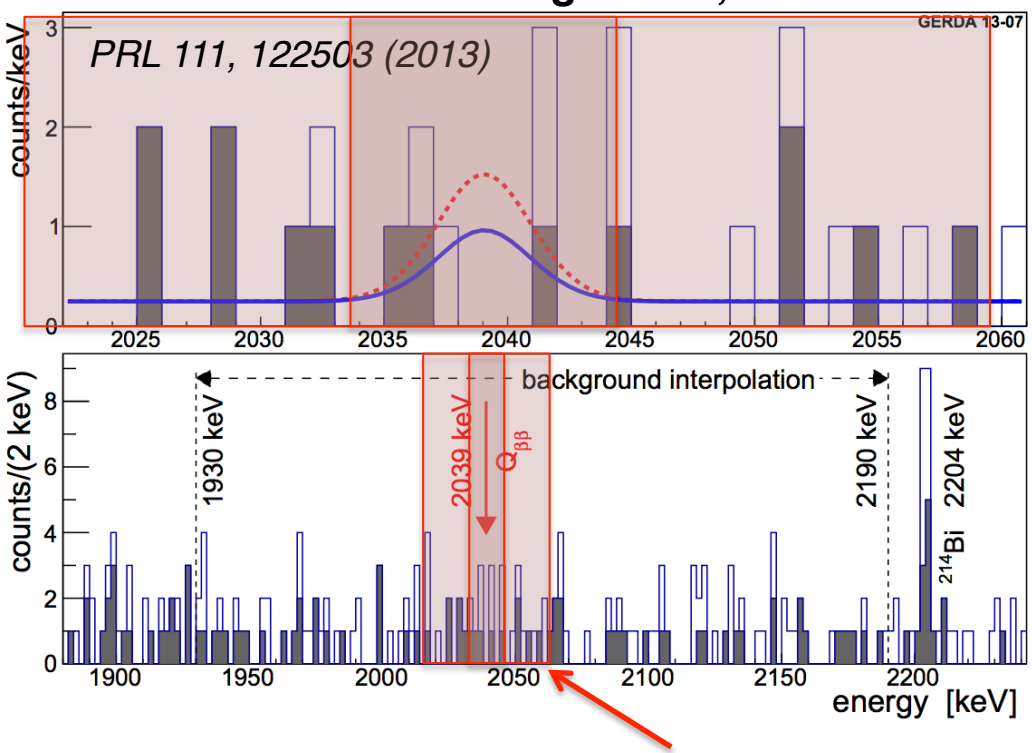
Energy spectrum after cuts NEMO-3 (2015):



Blinding

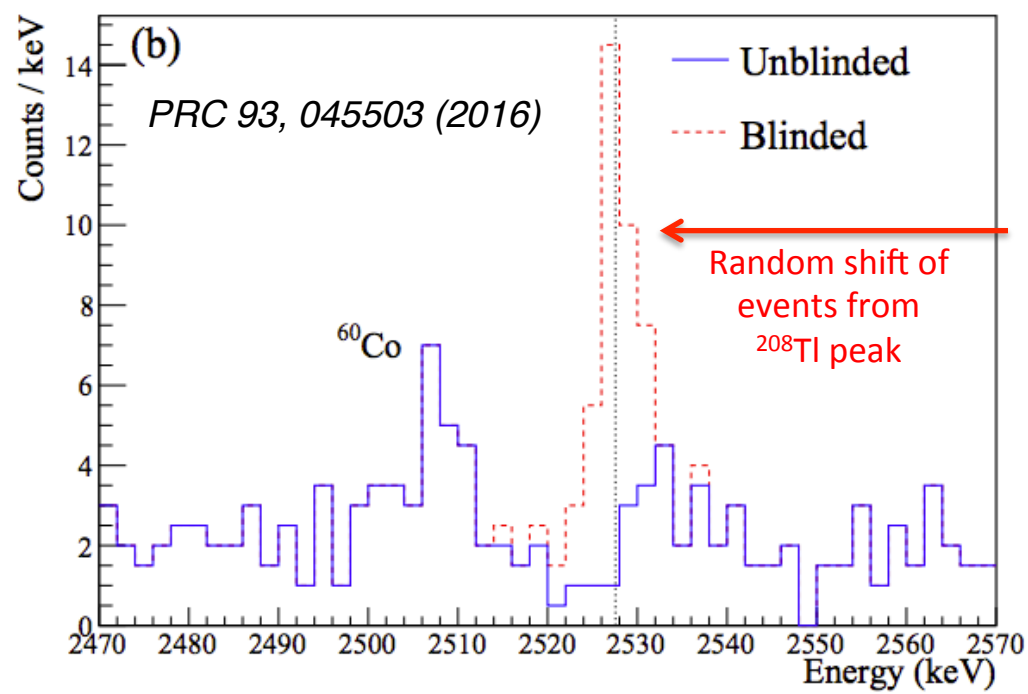
- All experiments are looking for a small number of events (typically <10 per dataset), so most implement blinding to avoid any statistical bias
- Standard method removes all events in the 0νββ ROI
- Alternative method "salts" an artificial peak at the Q-value, which allows data quality issues to be identified

"Standard" blinding of ROI, GERDA:



Events in ROI removed

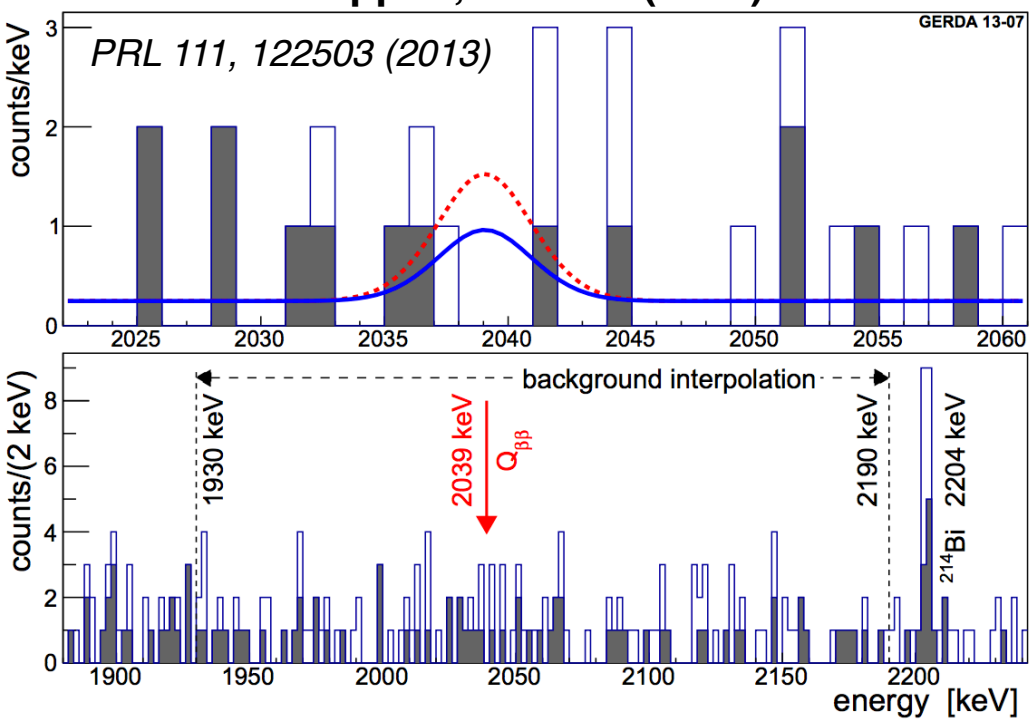
Example of "salting" technique, CUORE:



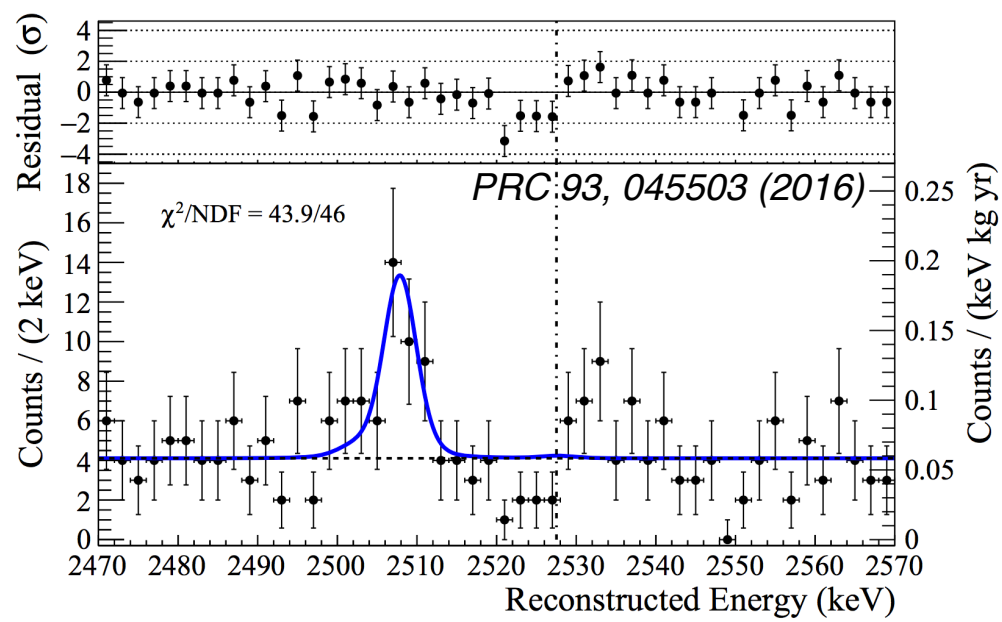
High resolution spectral fits

- Experiments typically perform a likelihood fit to a Gaussian peak (at known energy) on top of a flat background
- Perform likelihood fit, profiling over nuisance parameters (energy scale error, efficiency error, etc.)
- Often report frequentist limits as well as Bayesian limits with flat prior on counts

0νββ fit, GERDA (2013):



0νββ fit, CUORE-0 (2015):



Frequentist: $T_{1/2} > 2.1 \times 10^{25}$ yr (90% CL)
Median sensitivity = 2.4×10^{25} yr

Bayesian (flat prior): $T_{1/2} > 1.9 \times 10^{25}$ yr (90% credible int.)
Median sensitivity = 2.0×10^{25} yr

Bayesian (flat prior): $T_{1/2} > 2.7 \times 10^{24}$ yr (90% credible int.)
Median sensitivity = 2.9×10^{24} yr

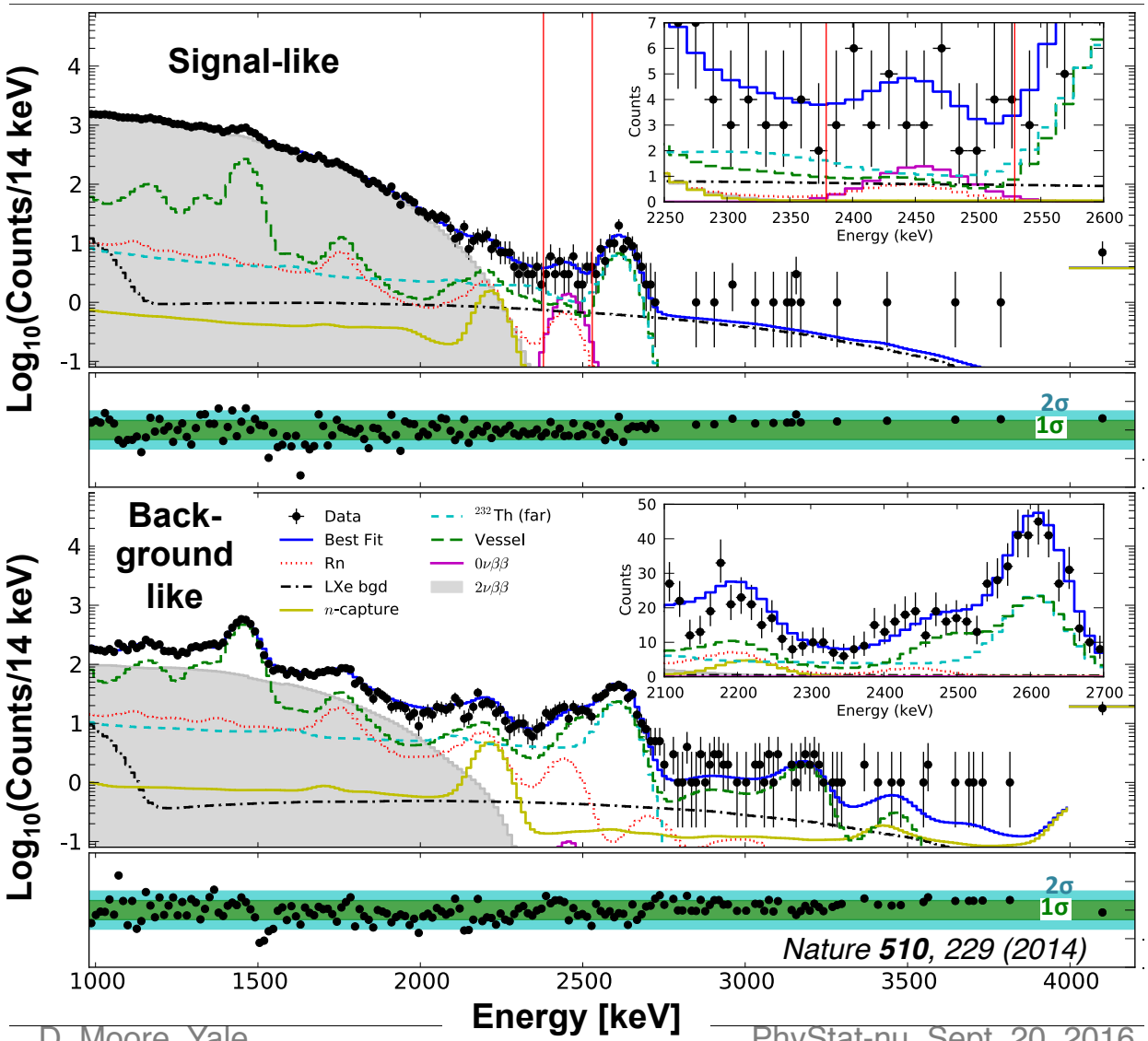
High resolution spectral fits (cont'd)

- Statistical issues for these sorts of "bump hunts" are well known in HEP
- For 0νββ, several simplifying factors:
  - Energy of the peak is known exactly (no "look elsewhere effect")
  - Cuts have very high signal efficiency (typically > 90%), however blinding still used by most experiments
  - Backgrounds are flat in this region (no strong background shape systematics)
  - Calibration lines are available at nearby energies to verify detector response, efficiencies

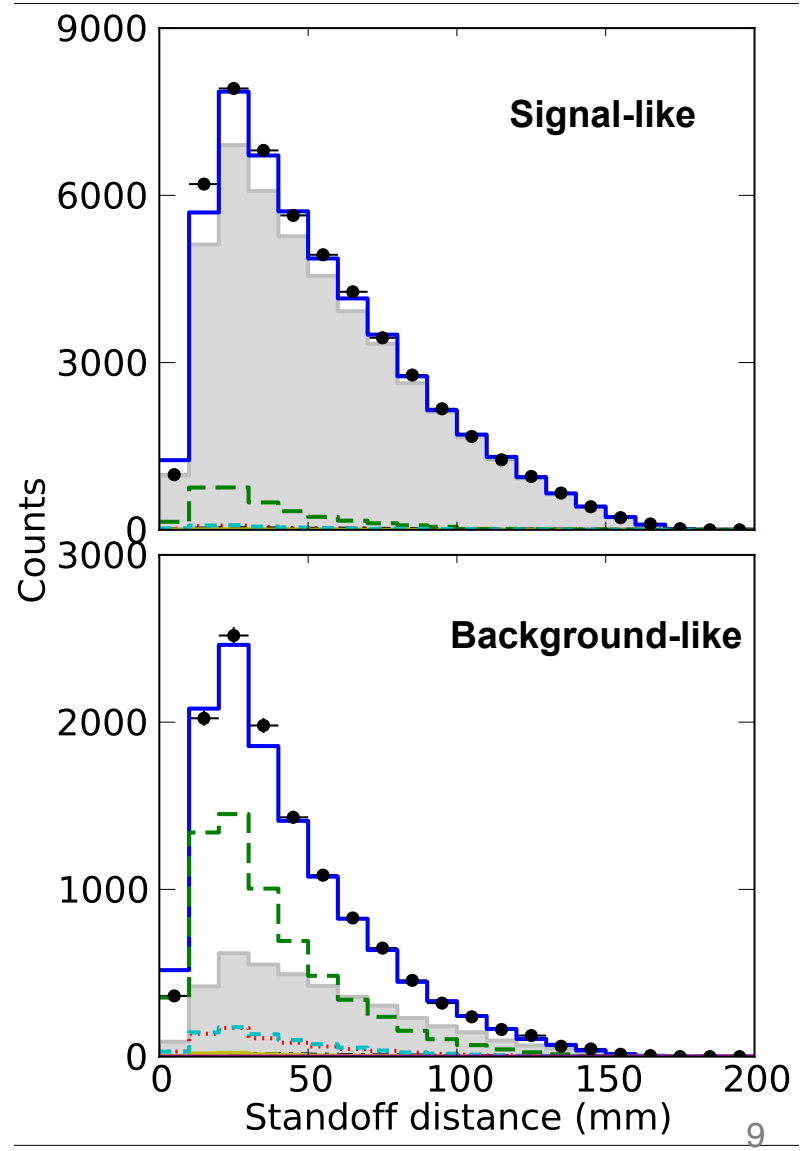
Large detectors

- For large, homogeneous detectors multiparameter likelihood fit over energy, position (and topology, if available) is performed

Best fit, energy projection, EXO-200:



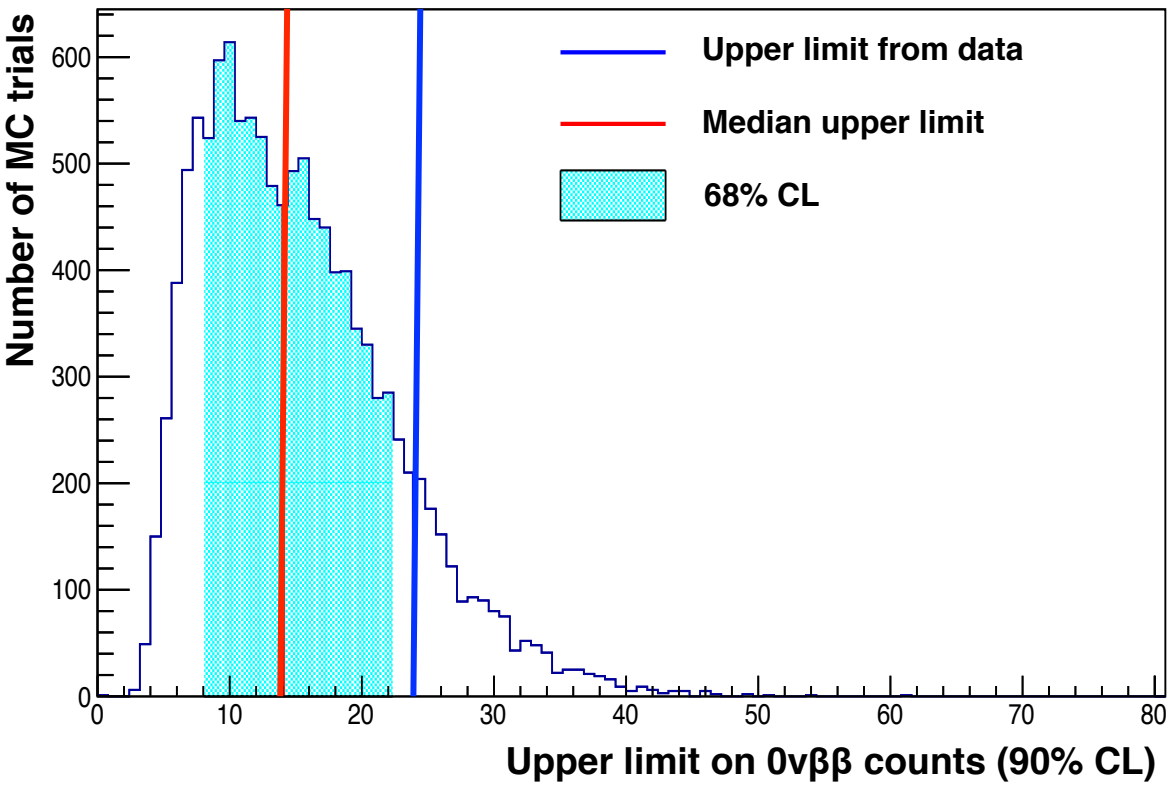
Best fit, position projection, EXO-200:



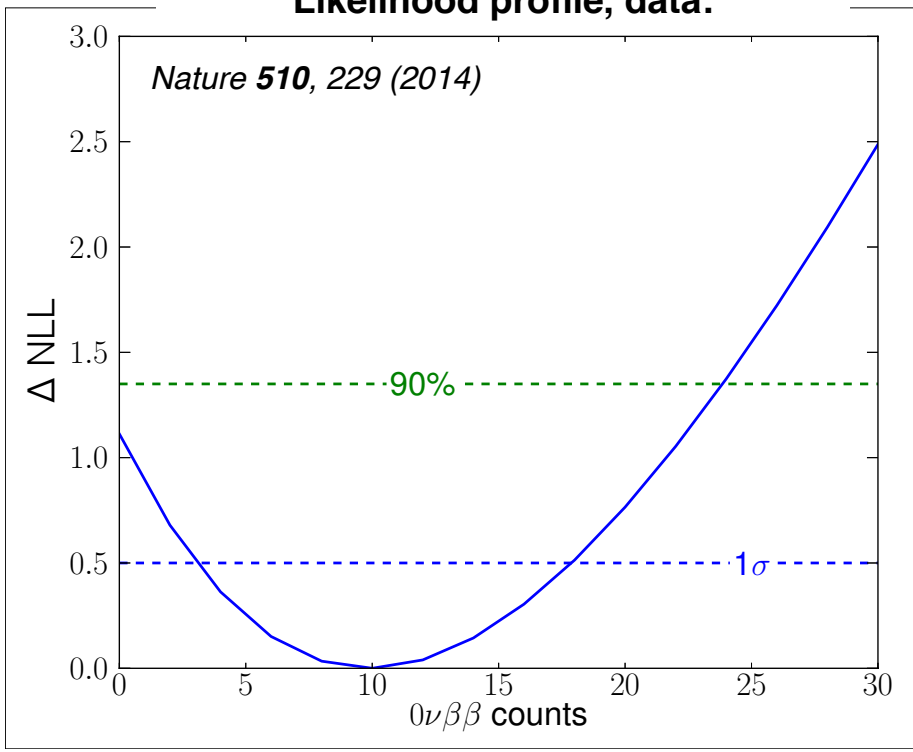
Large detectors (cont'd)

- Unlike high-resolution detectors, background models are less trivial
- Systematics are typically estimated through comparison of MC simulations to data from γ calibration sources and nuisance parameters profiled over
- Coverage verified from MC studies
- Expected sensitivity also determined from MC, low stats -> wide distribution

EXO-200 expected sensitivity, toy MC studies:



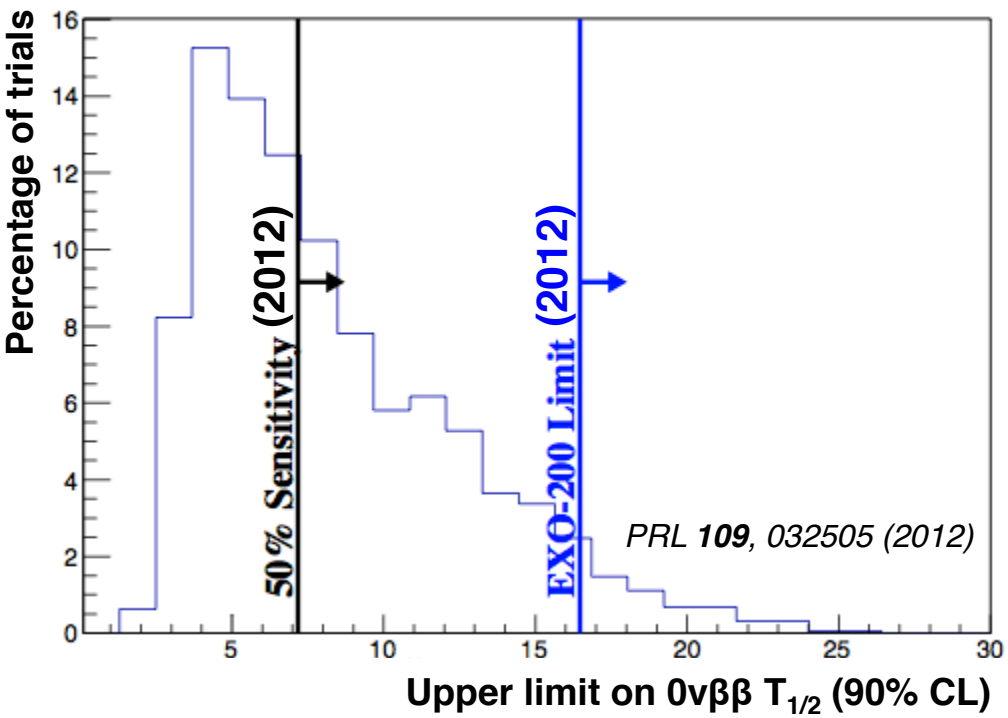
Likelihood profile, data:



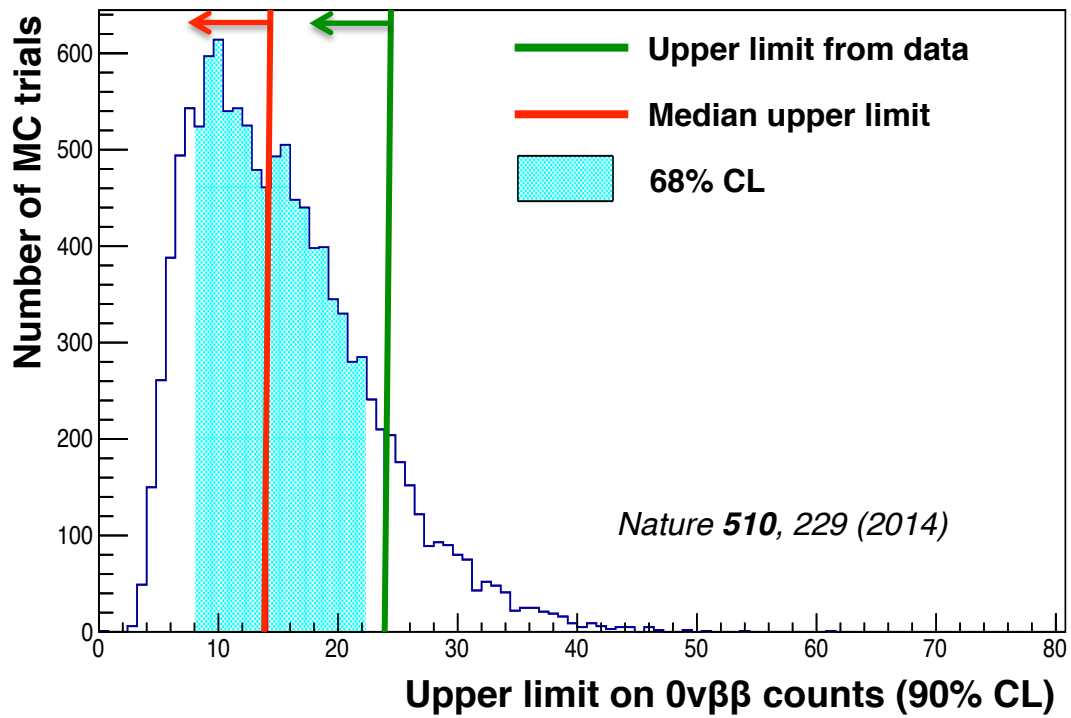
Statistical Issues

- Many statistical issues for 0νββ (typically <10 events/dataset) common to other rare event searches (see e.g. NIM A **774**, 103 [2015])
- Frequentist confidence intervals are most commonly presented, although some experiments report Bayesian credible intervals with flat prior
- Large statistical fluctuations common, e.g. EXO-200:

EXO-200 expected sensitivity, toy MC studies (2012):



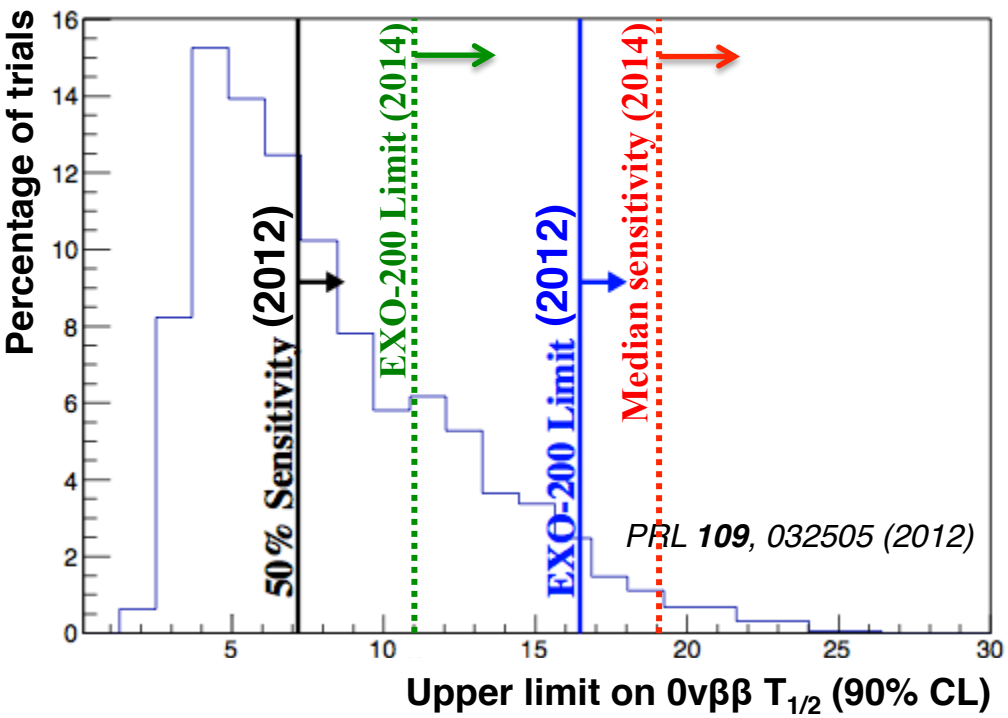
EXO-200 expected sensitivity, toy MC studies (2014):



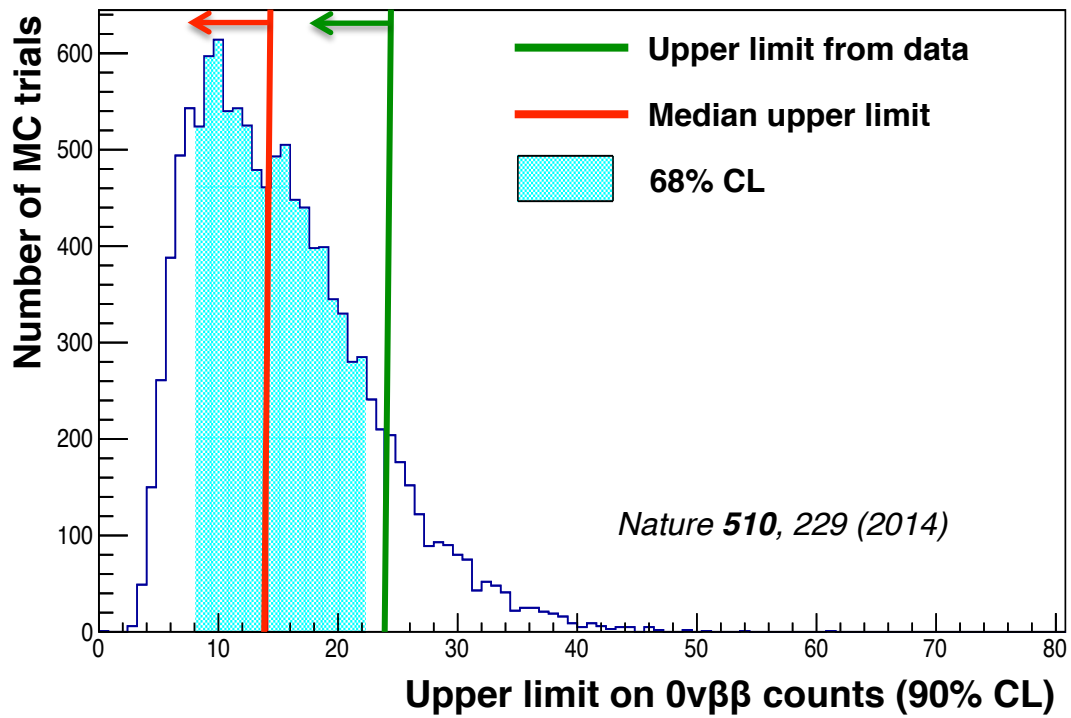
Statistical Issues

- Many statistical issues for 0νββ (typically <10 events/dataset) common to other rare event searches (see e.g. NIM A **774**, 103 [2015])
- Frequentist confidence intervals are most commonly presented, although some experiments report Bayesian credible intervals with flat prior
- Large statistical fluctuations common, e.g. EXO-200:

EXO-200 expected sensitivity, toy MC studies (2012):



EXO-200 expected sensitivity, toy MC studies (2014):



Isotope combinations

- Experiments measure $T_{1/2}$ but we are interested in the corresponding effective Majorana mass:

Majorana mass

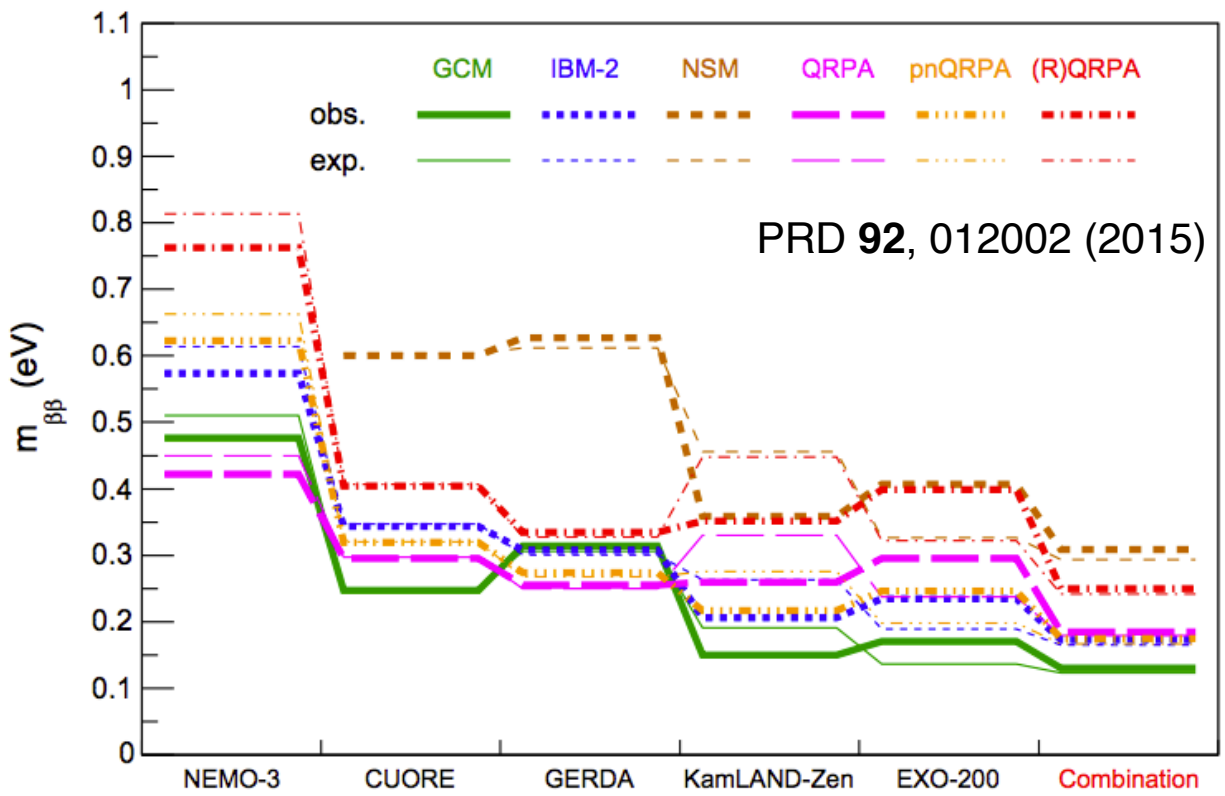
$$(T^{0\nu}_{1/2})^{-1} = G^{0\nu}|M^{0\nu}|^2(m_{\beta\beta}/m_e)^2$$

Phase space factor

Nuclear matrix element (NME)

- Substantial theoretical uncertainty on NMEs makes comparison difficult
- Correlations and model errors largely unquantified
- Typical procedure is to use most and least extreme set for each isotope

Example of model-dependent NME uncertainty:



Summary

- Statistical challenges in 0νββ are common to many rare event searches
- Due to low statistics, care must be taken in presenting and interpreting statistical intervals reported by each experiment
- Both Frequentist and Bayesian techniques are used in recent literature, no community-wide consensus on precise statistical technique
- Substantial theoretical uncertainty in comparing different isotopes, which prevents a straightforward combination of all experimental results

Statistical Issues (cont'd)

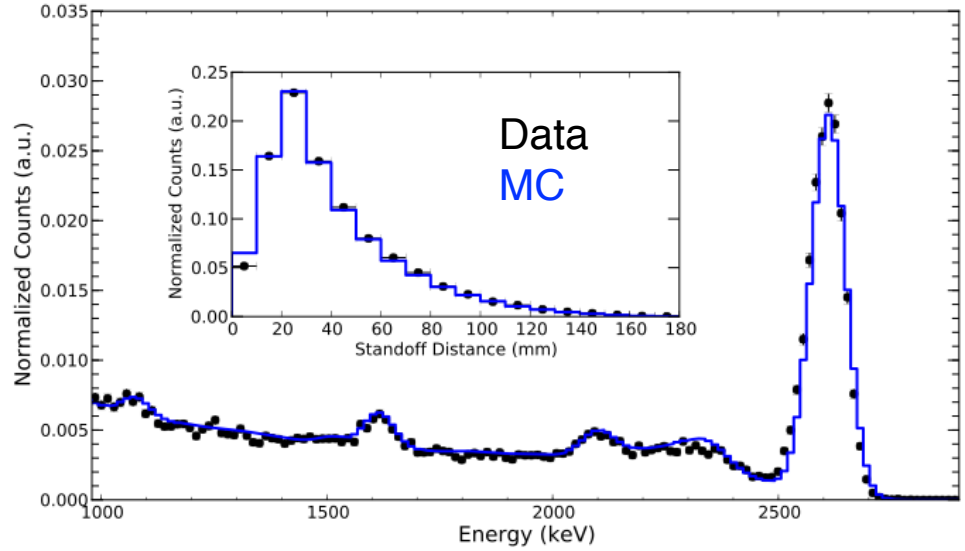
Selection of recent 0νββ half-life sensitivities and limits:

Experiment:	90% UL on $T_{1/2}$ [10^{25} yr]	Median sensitivity [10^{25} yr]	P-value from toy MCs	Reference
Large detectors (complex background models):				
EXO-200 (2012)	1.6	0.7	0.065	PRL **109**, 032505 (2012)
EXO-200 (2014)	1.1	1.9	0.85	Nature **510**, 229 (2014)
KamLAND-Zen (2013)	1.9	1.0	0.12	PRL **110**, 062502 (2013)
KamLAND-Zen (2016)	9.2	5.6	0.12	PRL **117**, 082503 (2016)
NEMO-3 (2015)	0.11	0.10		PRD **92**, 072011 (2015)
High resolution detectors (simpler background models):				
GERDA (2013)	2.1	2.4		PRL **111**, 122503 (2013)
GERDA (2016)	5.2	4.0		Presentation at Nu16
CUORE (2015)	0.27	0.29	0.547	PRL **115**, 102502 (2015)

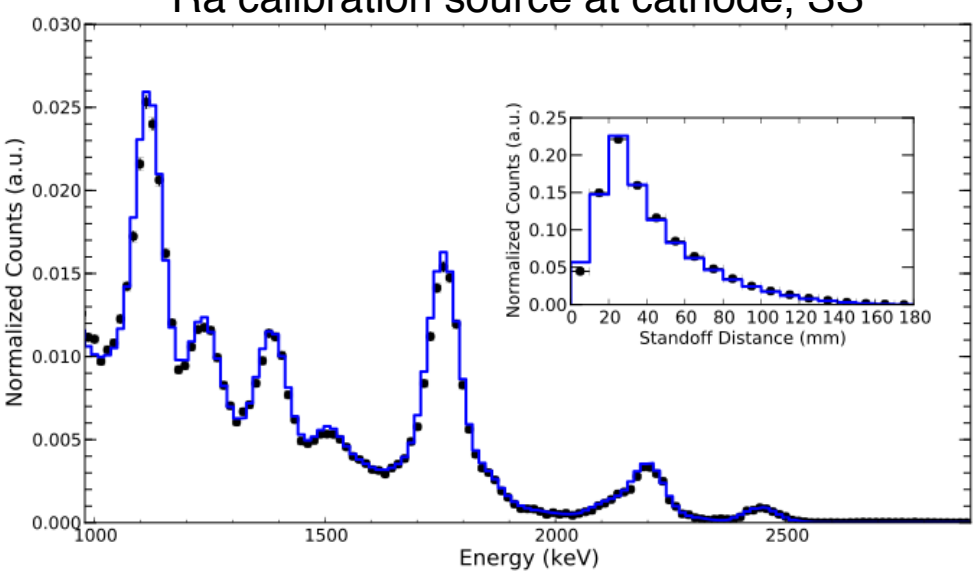
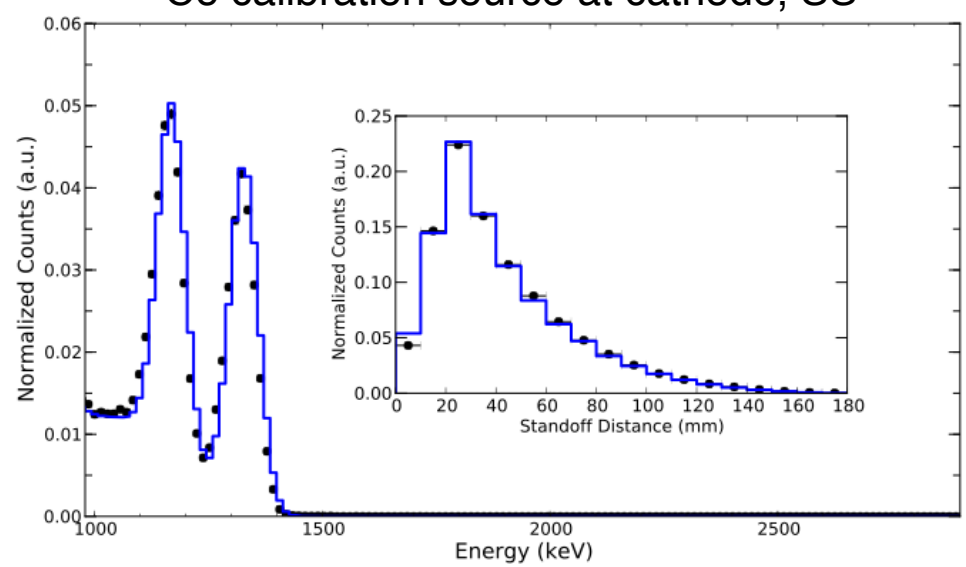
Shape systematics

- Systematic errors estimated from calibrations sources
- Background PDFs determined from MC
- Residual data/MC discrepancy propagated to systematic using toy MC studies

^{228}Th calibration source at cathode, SS



^{226}Ra calibration source at cathode, SS



^{60}Co calibration source at cathode, SS