

# A Search for the Neutrinoless Double Beta Decay of Xenon-136 with Improved Sensitivity from Denoising

Clayton G. Davis

April 3, 2014

# Outline

$\beta\beta 2\nu$  and  $\beta\beta 0\nu$  Decay

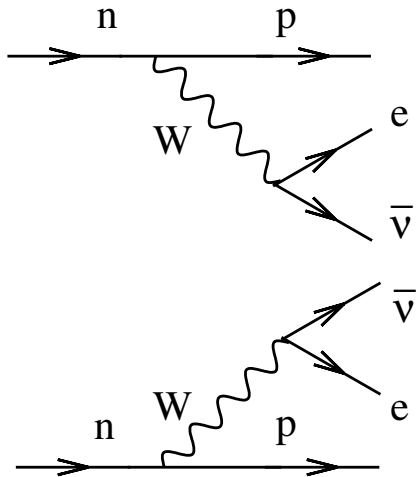
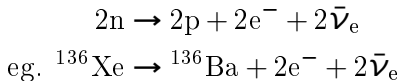
The EXO-200 Detector

Denoising

Results

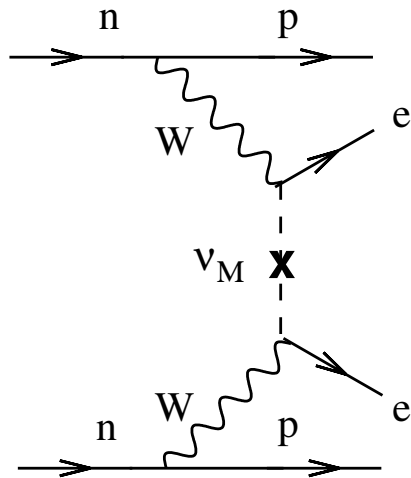
## What is Double-Beta Decay?

Feynman diagram for  $\beta\beta 2\nu$  decay. Equivalent to two single- $\beta$  decays:



Avignone et al., RMP 2008.

## What is Double-Beta Decay?



Avignone et al., RMP 2008.

Feynman diagram for  $\beta\beta 2\nu$  decay. Equivalent to two single- $\beta$  decays:

$$2n \rightarrow 2p + 2e^- + 2\bar{\nu}_e$$

eg.  $^{136}\text{Xe} \rightarrow ^{136}\text{Ba} + 2e^- + 2\bar{\nu}_e$

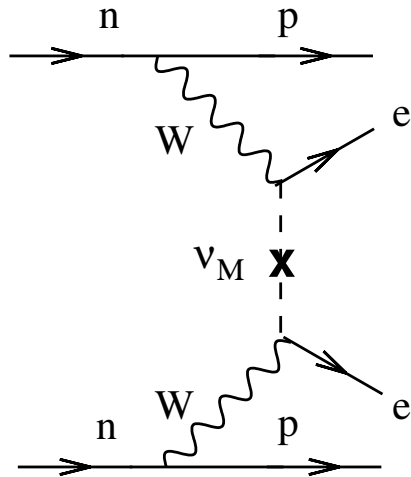
Feynman diagram for  $\beta\beta 0\nu$  decay. Neutrinos annihilate each other:

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

eg.  $^{136}\text{Xe} \rightarrow ^{136}\text{Ba} + 2e^-$

$\beta\beta 2\nu$  is allowed in the Standard Model;  $\beta\beta 0\nu$  is not.

## Implications of Double-Beta Decay



Avignone et al., RMP 2008.

- ▶ Lepton number changes:

$$\Delta L = +2$$

- ▶ Neutrinos can convert to their own antiparticle:

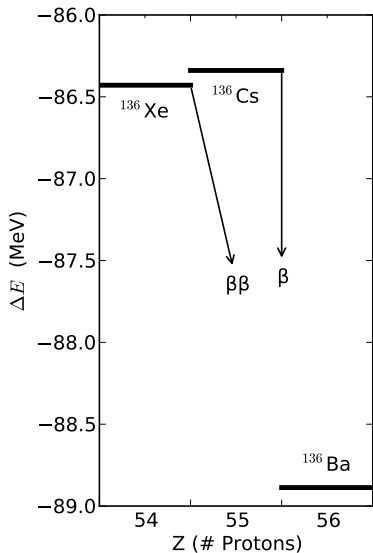
$$\bar{\nu}_R \rightarrow \nu_L$$

- ▶ Neutrinos have mass through a Majorana interaction:

$$-\frac{m_L}{2} (\bar{\Psi}_L^c \Psi_L + \bar{\Psi}_L \Psi_L^c)$$

$$-\frac{m_R}{2} (\bar{\Psi}_R^c \Psi_R + \bar{\Psi}_R \Psi_R^c)$$

## The $A = 136$ Isobar



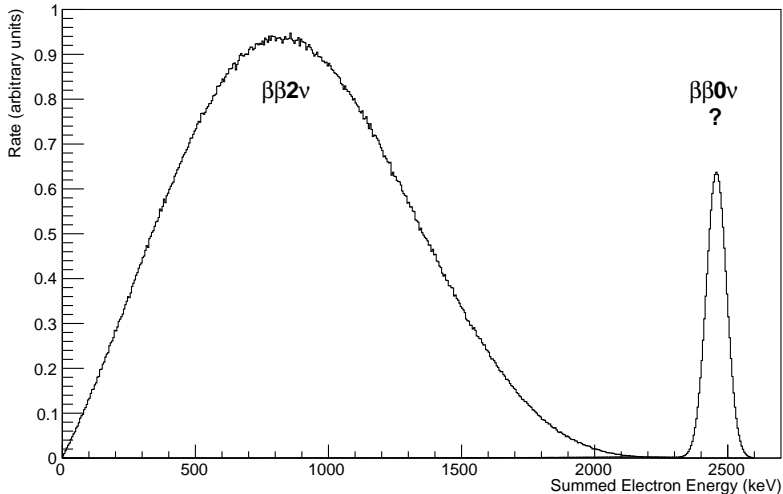
$^{136}\text{Cs}$  undergoes single- $\beta$  decay.

$^{136}\text{Xe}$  cannot, due to energy conservation – but it can  $\beta\beta$  decay through  $^{136}\text{Cs}$  to  $^{136}\text{Ba}$ .

The Q-value of  $^{136}\text{Xe} \rightarrow ^{136}\text{Ba}$  is  $2457.83 \pm 0.37$  keV, shared between all final products of the decay.

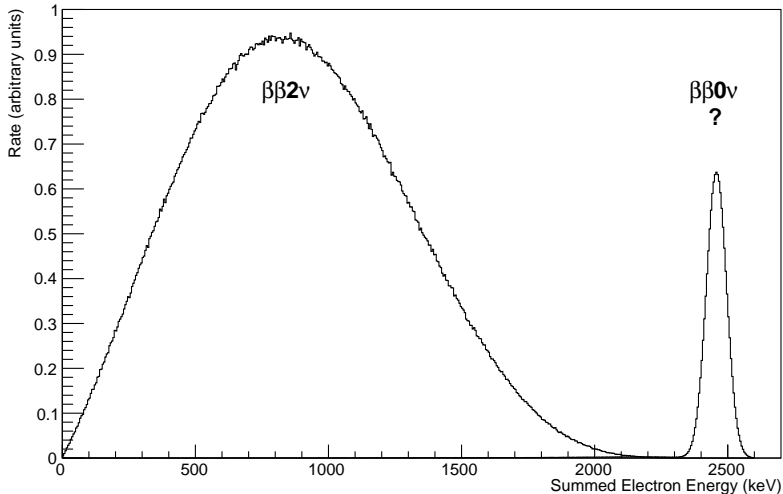
We observe energy in electrons; energy in neutrinos is lost.

## Ideal Double-Beta Energy Spectrum



$^{136}\text{Xe}$   $\beta\beta 2\nu$  produces a smooth energy spectrum; “missing” energy carried off by neutrinos.

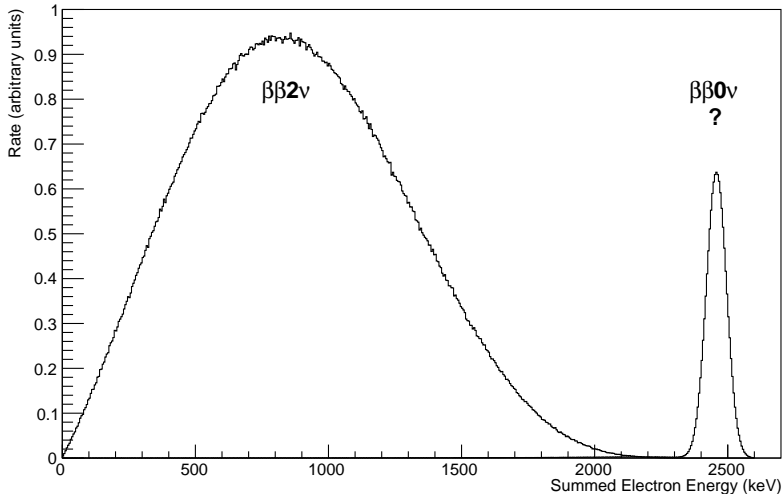
## Ideal Double-Beta Energy Spectrum



$^{136}\text{Xe}$   $\beta\beta 0\nu$  has no neutrinos, so no “missing” energy;  
mono-energetic peak at  $Q = 2458$  keV.

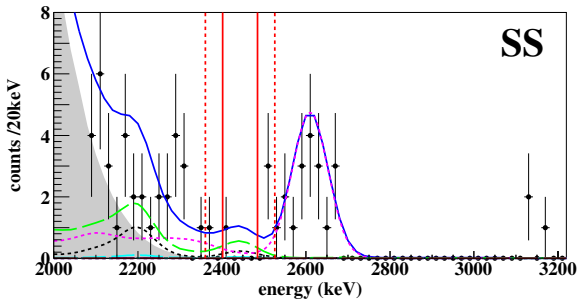
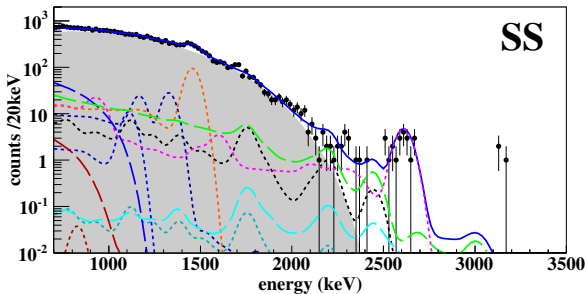


## Ideal Double-Beta Energy Spectrum



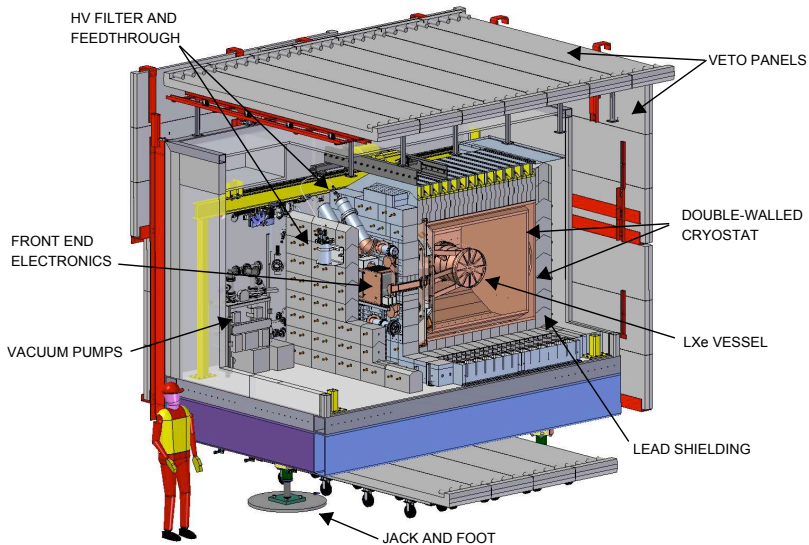
If the  $\beta\beta 0\nu$  peak exists, neutrinos have Majorana mass; peak height gives a measurement of that mass.

## Observed Energy Spectrum (May 2012)

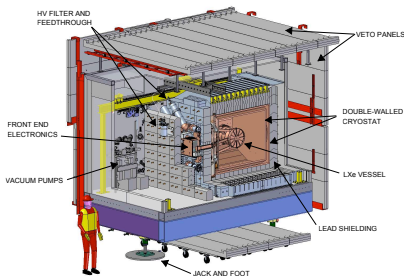


32.5 kg-yrs  
(vs. 99.8  
kg-yrs for  
the present  
analysis)

# The EXO-200 Detector

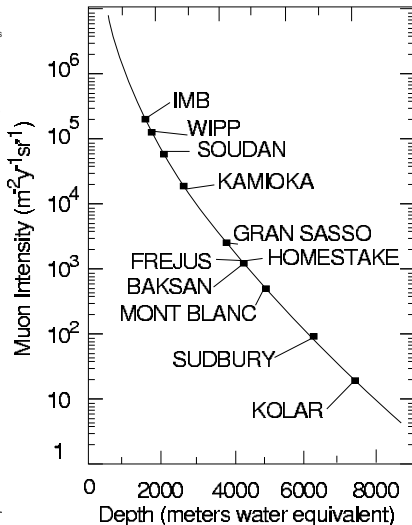


# The EXO-200 Detector



To search for rare decays, low background is key:

- ▶ Clean (low-radioactivity) materials surrounding TPC.
- ▶ Deep underground to avoid cosmogenics.

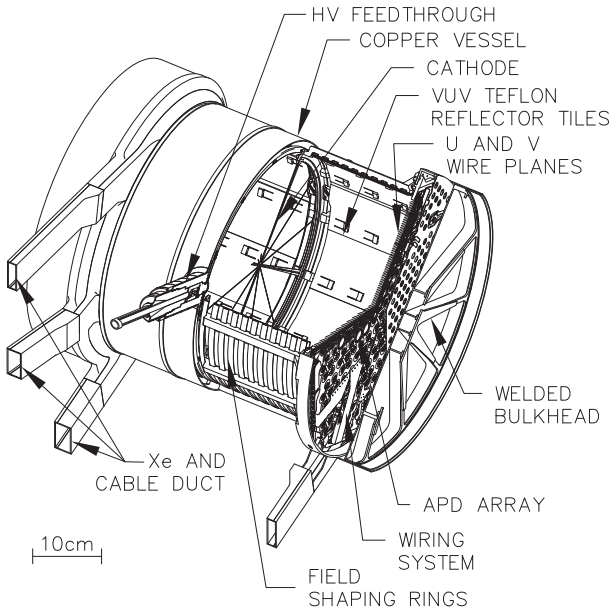


Esch et al., NIM A 2005.

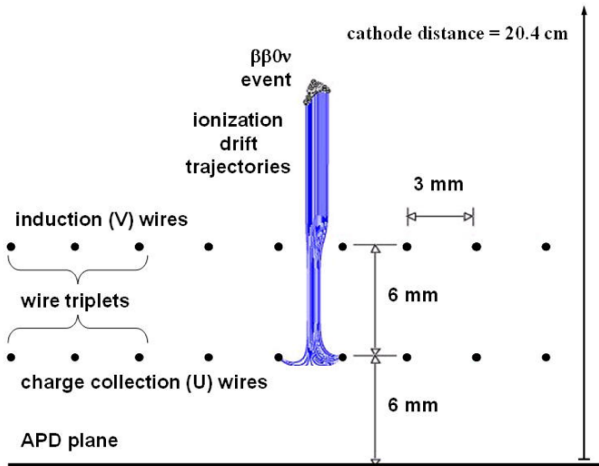
# EXO-200 TPC

110 kg of liquid xenon in active volume, enriched to 80.6% in  $^{136}\text{Xe}$ , contained in a time projection chamber (TPC).

Xenon continuously circulates through purifiers outside of the cryostat.

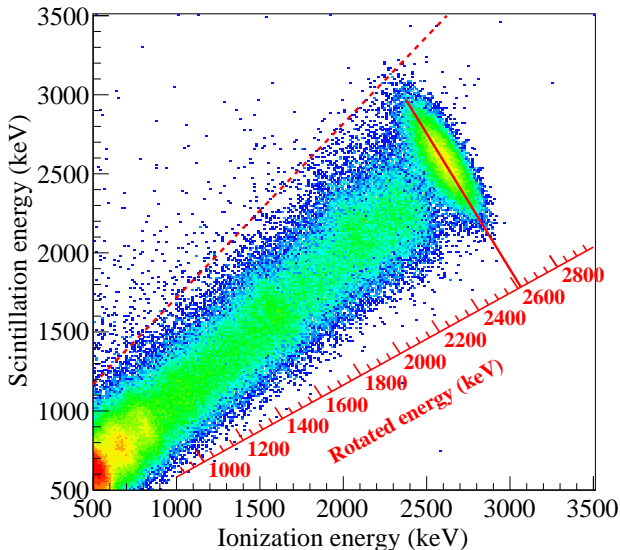


# EXO-200 TPC



Charge drifts under an electric field and is collected by wires on the anodes. Light is observed by APDs behind the wires.

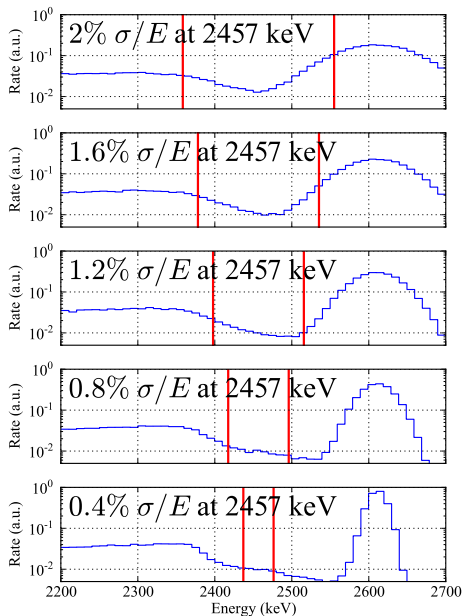
## Energy from Ionization and Scintillation



Energy is independently measured from scintillation and ionization.

They are anticorrelated – better energy resolution from both together than either independently.

## Simulation



## Primary Backgrounds: $^{232}\text{Th}$ , $^{238}\text{U}$ , and $^{137}\text{Xe}$

Energy resolution is measured as  $\sigma/\text{mean}$  of a peak at the Q-value. Typically 1.5-2% for us.

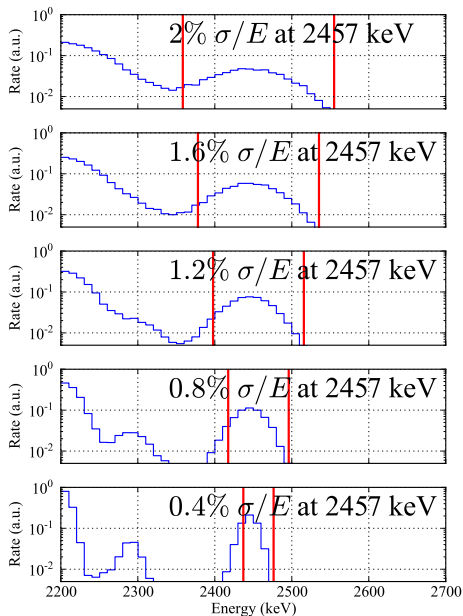
Impact of background goes like integral of pdf between red  $2\sigma$  lines.

$^{232}\text{Th}$  gamma line at 2615 keV, so energy resolution has strong impact down to about 1.2%.

Beyond that, impact on  $^{232}\text{Th}$  is much less.



## Simulation

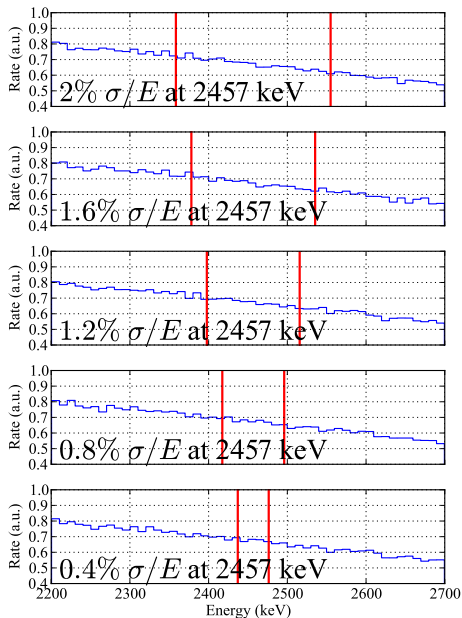


## Primary Backgrounds: $^{232}\text{Th}$ , $^{238}\text{U}$ , and $^{137}\text{Xe}$

$^{238}\text{U}$  has a 2448-keV gamma line, indistinguishable from 2457-keV Q-value except with extremely good resolution.

So, even down to 0.4% energy resolution, most of the  $^{238}\text{U}$  peak at 2448 keV is still within our energy window. Impact of resolution on  $^{238}\text{U}$  backgrounds is small.

## Simulation

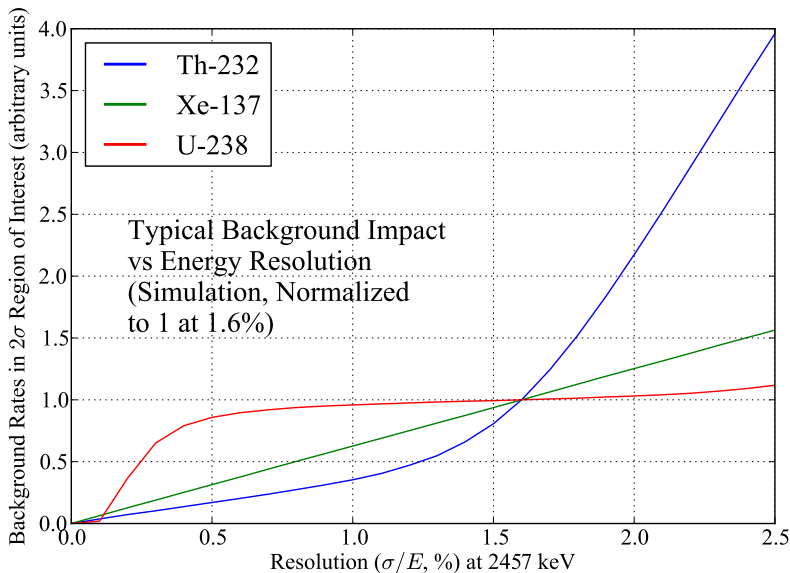


## Primary Backgrounds: $^{232}\text{Th}$ , $^{238}\text{U}$ , and $^{137}\text{Xe}$

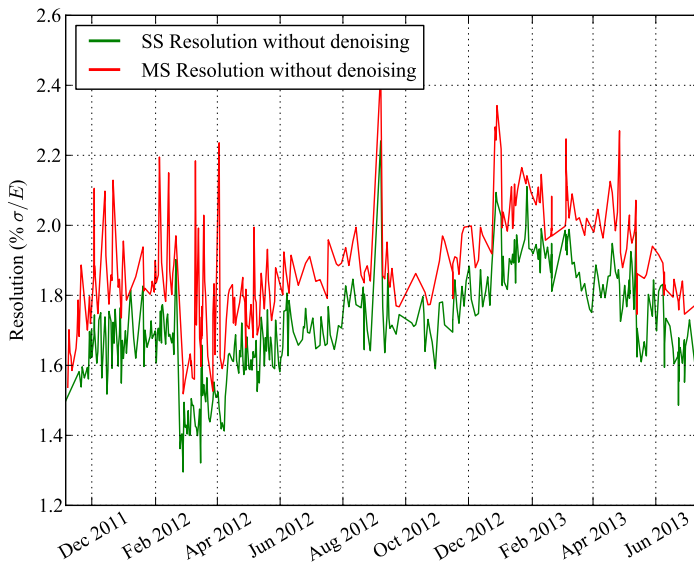
$^{137}\text{Xe}$  is a cosmogenic background from  $^{136}\text{Xe}$  which  $\beta$ -decays; endpoint of 4173 keV.

$^{137}\text{Xe}$  spectrum is smooth around Q-value, so background impact goes linearly with resolution.

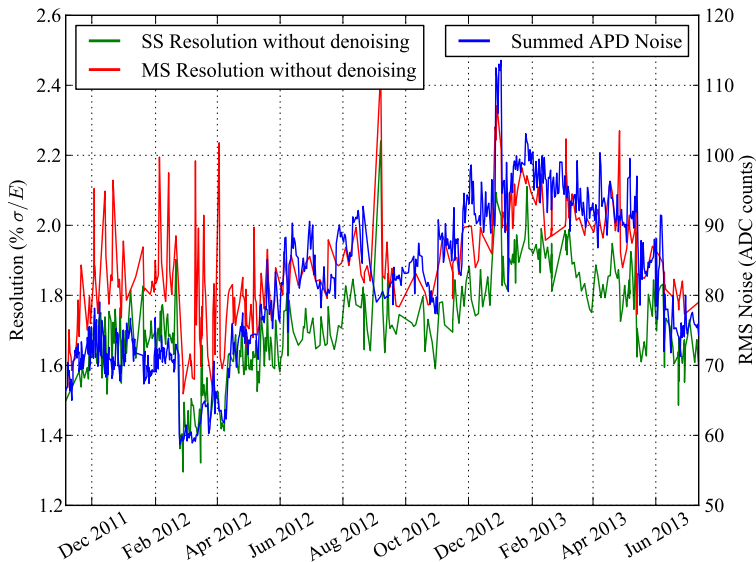
## Backgrounds vs. Resolution



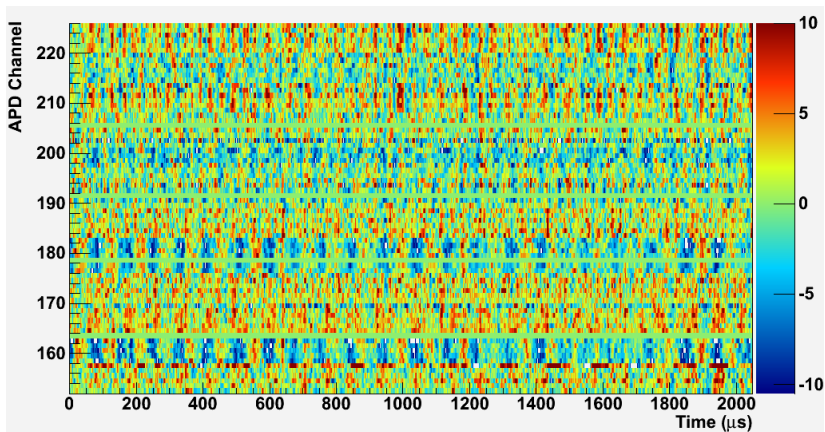
# Time Variation of Resolution



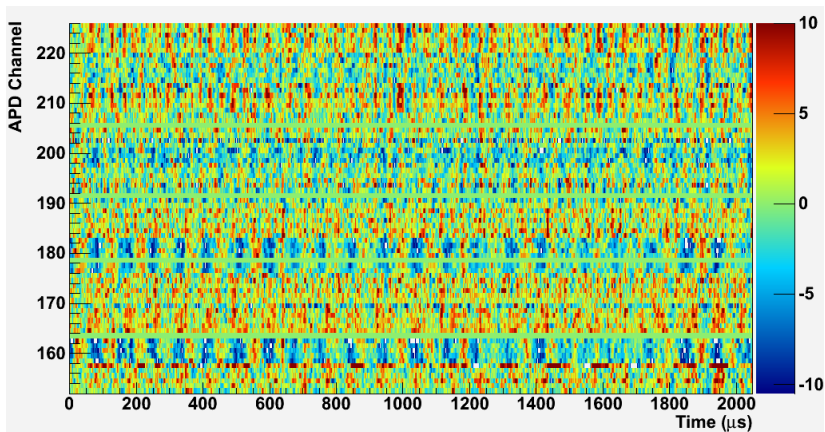
# Time Variation of Resolution



## APD Noise is Correlated across Channels



## APD Noise is Correlated across Channels



Correlated noise  $\Rightarrow$  offline denoising of some sort should work!

## Types of Noise

Three types of noise in the scintillation measurements:

- ▶ Electronic noise.
- ▶ Photon fluctuations.
- ▶ Gain fluctuations.

A “denoising” algorithm should minimize all three.



## Types of Noise

Three types of noise in the scintillation measurements:

- ▶ Electronic noise.                      ← Additive; use fewer channels.
- ▶ Photon fluctuations.
- ▶ Gain fluctuations.

A “denoising” algorithm should minimize all three.

## Types of Noise

Three types of noise in the scintillation measurements:

- ▶ Electronic noise. ← Additive; use fewer channels.
- ▶ Photon fluctuations. ← Poissonian; use more channels.
- ▶ Gain fluctuations. ← Poissonian; use more channels.

A “denoising” algorithm should minimize all three.

## Types of Noise

Three types of noise in the scintillation measurements:

- ▶ Electronic noise. ← Additive; use fewer channels.
- ▶ Photon fluctuations. ← Poissonian; use more channels.
- ▶ Gain fluctuations. ← Poissonian; use more channels.

A “denoising” algorithm should minimize all three.

Most denoising algorithms transform an input waveform to an output waveform, with pulses amplified and additive noise reduced.

Here we’re trying to reduce the impact of noise which is correlated with pulses. That means traditional denoising won’t accomplish what we need.

So, “denoising” for us means producing a noise-tolerant estimate of the scintillation energy.

## Types of Denoising

There are also three approaches to reducing the impact of noise:

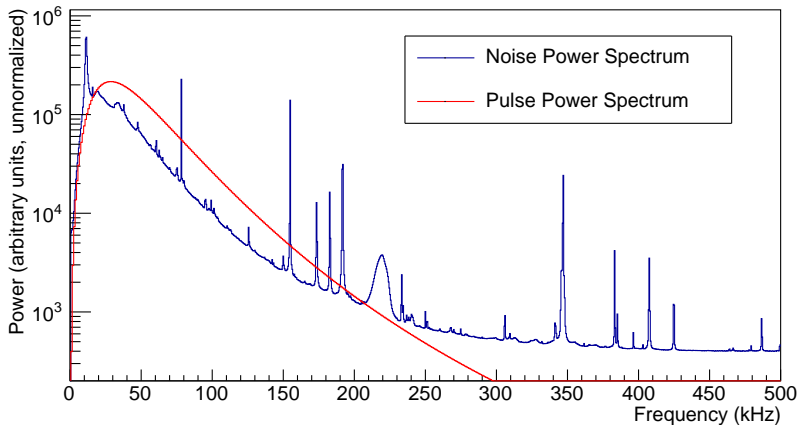
- ▶ Frequency weighting.
- ▶ Channel weighting.
- ▶ Exploit noise correlations.

# Types of Denoising

There are also three approaches to reducing the impact of noise:

- ▶ Frequency weighting.
- ▶ Channel weighting.
- ▶ Exploit noise correlations.

Noise and pulse power spectra.

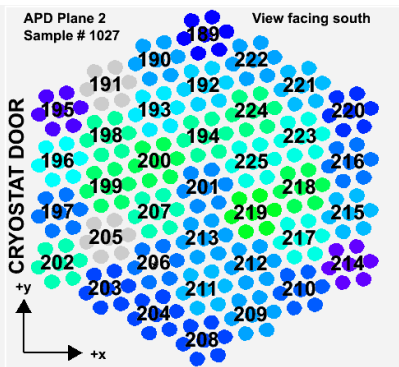
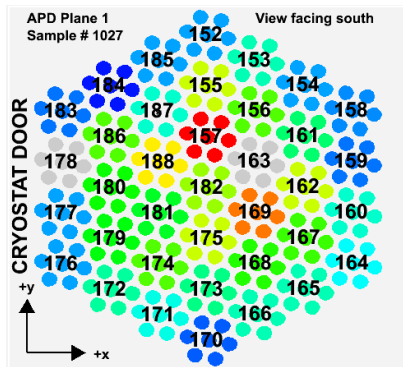


## Types of Denoising

There are also three approaches to reducing the impact of noise:

- ▶ Frequency weighting.
- ▶ Channel weighting.
- ▶ Exploit noise correlations.

APD pulse magnitudes.

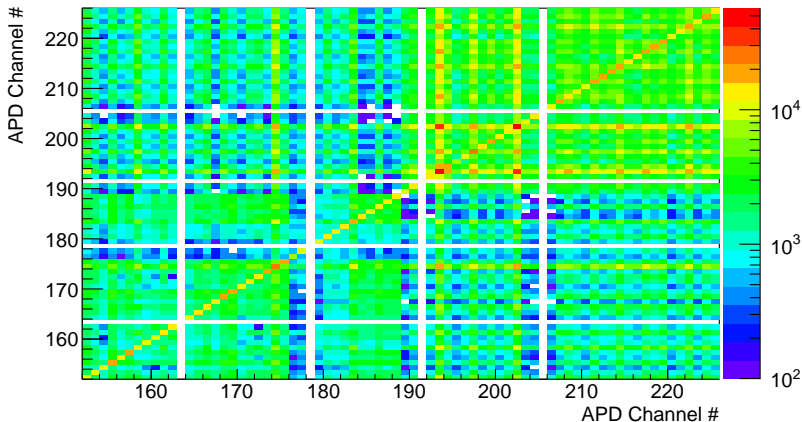


## Types of Denoising

There are also three approaches to reducing the impact of noise:

- ▶ Frequency weighting.
- ▶ Channel weighting.
- ▶ Exploit noise correlations.

Noise correlations at 78.6 kHz.



## Waveform Model

When there is one energy deposit in the xenon, the APD waveform  $X_i[f]$  (in frequency space) on channel  $i$  can be modeled:

$$X_i[f] = M_i Y_i[f] + N_i[f], \text{ where:}$$

- ▶  $Y_i[f]$  is the unit-magnitude template pulse for channel  $i$ .
- ▶  $M_i$  is the magnitude of the pulse observed on channel  $i$ , including the scintillation energy and Poisson fluctuations in photon statistics and gain.
- ▶  $N_i[f]$  is the additive (electronic) noise on channel  $i$ .

To complete our model, we need to understand the distributions of the random variables:

- ▶ How  $M_i$  depends on the unknown energy  $E$ .
- ▶ How magnitudes  $M_i$  and  $M_j$  are correlated.
- ▶ How electronic noise  $N_i[f]$  and  $N_j[f]$  are correlated.



## Electronic Noise

To measure the electronic (additive) noise, we need waveforms with no pulse.

Fortunately, EXO-200 is a low-background detector: most of the time all it measures is noise.

So, this is fairly easy: use our noise data to measure all of the pairwise noise correlations.

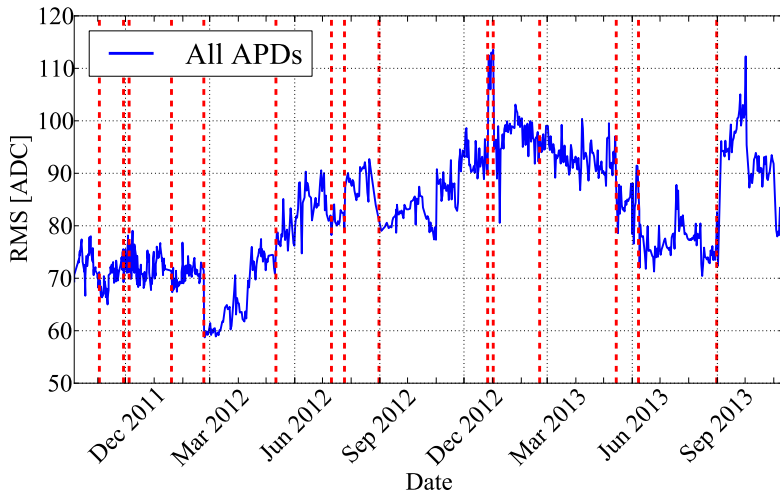
Explicitly, we measure:

$$\left\langle N_i^R[f] N_j^R[f] \right\rangle, \left\langle N_i^R[f] N_j^I[f] \right\rangle, \text{ and } \left\langle N_i^I[f] N_j^I[f] \right\rangle,$$

where  $N_i^R[f]$  and  $N_i^I[f]$  are the real and imaginary parts of  $N_i[f]$ .

## Electronic Noise

Main detail: the electronic noise changes over time (mostly in discontinuous steps).

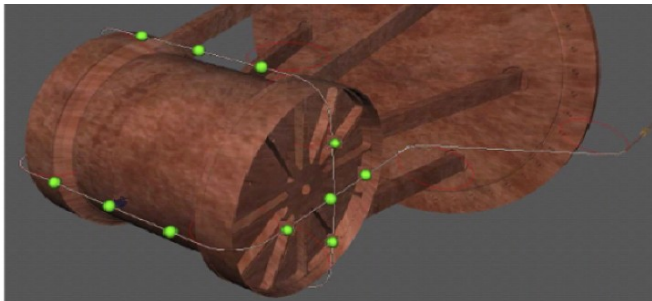


## Lightmaps

The expected pulse height  $M_i$  on a channel  $i$  from a deposit at position  $\vec{x}$  and time  $t$  is described by a lightmap  $L_i(\vec{x}, t)$ :

$$\langle M_i \rangle = L_i(\vec{x}, t)E.$$

To measure  $L_i(\vec{x}, t)$ , we need a known-energy deposit for all position and time bins. But we don't have that much data: calibration sources give us good statistics, but aren't near every position.



## Lightmaps

The expected pulse height  $M_i$  on a channel  $i$  from a deposit at position  $\vec{x}$  and time  $t$  is described by a lightmap  $L_i(\vec{x}, t)$ :

$$\langle M_i \rangle = L_i(\vec{x}, t)E.$$

To reduce the amount of statistics needed, we make an approximation that  $L_i(\vec{x}, t)$  is separable into spatial and temporal components:

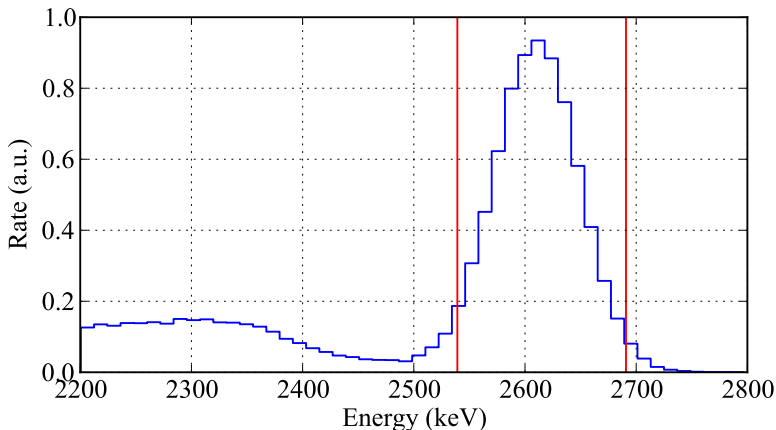
$$L_i(\vec{x}, t) = R_i(\vec{x})S_i(t).$$

So, rather than needing one known-energy deposit per position per time, we need one per position during the whole history of EXO-200.

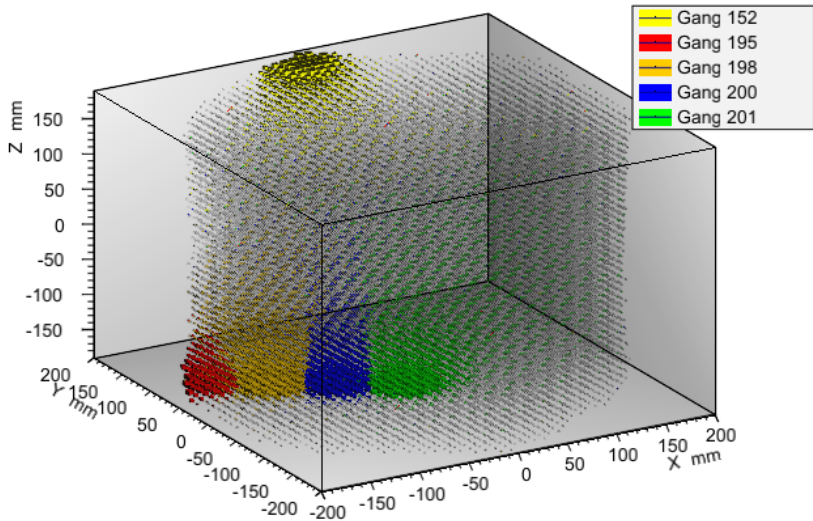
(We do this assuming that APD gain varies over time, but light collection does not, and gain does not depend on the position of the deposit.)

## Lightmaps

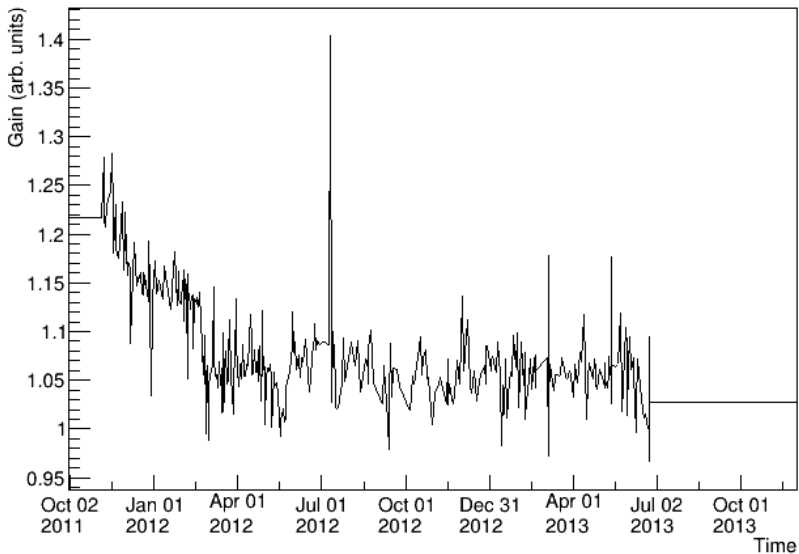
With this simpler form for the lightmap, we can measure  $R_i(\vec{x})$  and  $S_i(t)$  by combining all thorium source calibration events from the  $^{208}\text{Tl}$  2615-keV gamma line. It is a clean, well-isolated peak:



# Lightmaps: $R_i(\vec{x})$



## Lightmaps: $S_i(t)$ for $i = 152$



## Light Collection Noise

The expected pulse height  $M_i$  on a channel  $i$  from a deposit at position  $\vec{x}$  and time  $t$  is described by a lightmap  $L_i(\vec{x}, t)$ :

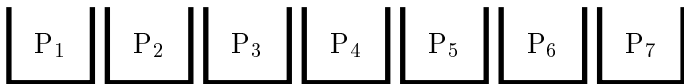
$$\langle M_i \rangle = L_i(\vec{x}, t)E.$$

What about covariances of  $M_i$ ?

Photon collection by the APDs has a multinomial distribution:

**\*** N photons

With photon collection  $\langle P_i \rangle = N \cdot f_i$ ,  
covariances are  $\text{cov}(P_i, P_j) = N \cdot (f_i \delta_{ij} - f_i f_j)$ .



Photons randomly deposit on APDs (or are not observed).



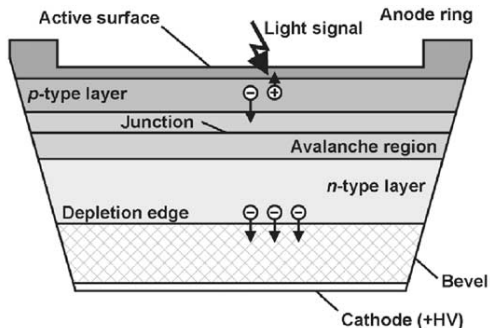
## Light Collection Noise

The expected pulse height  $M_i$  on a channel  $i$  from a deposit at position  $\vec{x}$  and time  $t$  is described by a lightmap  $L_i(\vec{x}, t)$ :

$$\langle M_i \rangle = L_i(\vec{x}, t)E.$$

What about covariances of  $M_i$ ?

We treat an APD as having two operational phases:



- ▶ Active surface:  
photons are converted to photoelectrons (with a quantum efficiency).
- ▶ Avalanche region:  
photoelectrons are amplified by avalanche (APD gain).

Both processes produce Poissonian fluctuations.

## Denoising: An Optimization Problem

So, the waveforms  $X_i[f]$  are modeled by:

$$X_i[f] = M_i Y_i[f] + N_i[f],$$

and we have full descriptions of the random variables  $M_i$  and  $N_i[f]$ .

Linear real-valued energy estimators  $\hat{E}$  take the form:

$$\hat{E} = \sum_{if} A_i[f] X_i^R[f] + B_i[f] X_i^I[f],$$

where  $A_i[f]$  and  $B_i[f]$  are parameters we can adjust.

Problem: minimize the mean square error,  $\langle (E - \hat{E})^2 \rangle$ .

Constraint: estimator is unbiased,  $\langle E - \hat{E} \rangle = 0$ .

# Denoising: An Optimization Problem

The solution is a big matrix equation

## Denoising: An Optimization Problem

The solution is a big matrix equation which we write in a compact way:

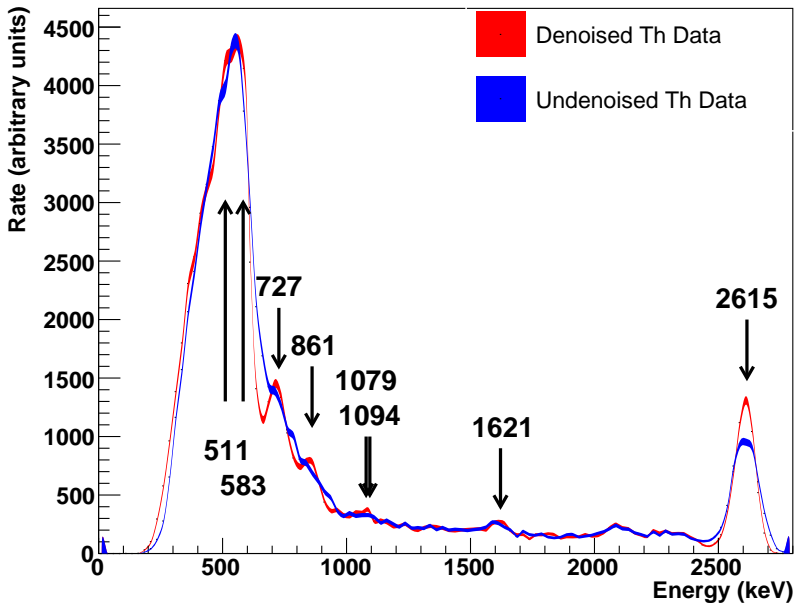
$$\begin{pmatrix} \mathbf{N} + \mathbf{P} & \mathbf{C}^\top \\ \mathbf{C} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The LHS matrix is a  $143,291 \times 143,291$  sparse matrix with roughly twenty million non-zero entries.

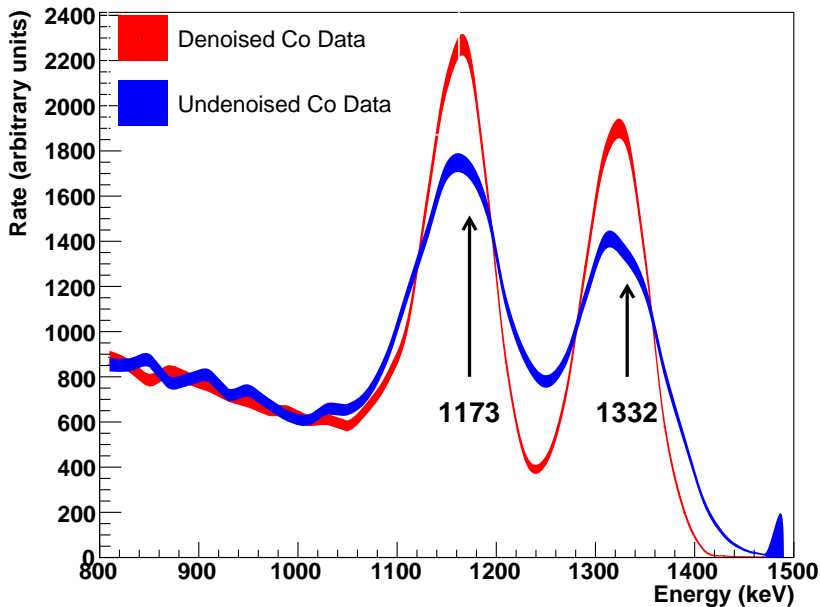
We need to solve this system once for each event, and there are hundreds of millions of events. Fortunately this is an embarrassingly parallel problem, so we can just find a big computing system and do it.

NERSC, at LBNL, was used for this analysis; we required roughly 50,000 core-hours to do this processing.

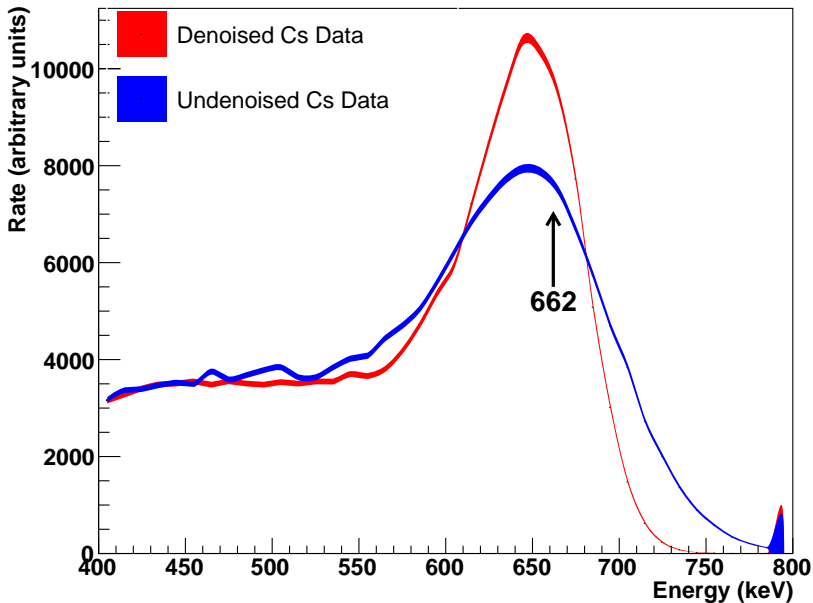
## $^{228}\text{Th}$ Source Spectrum (Before and After)



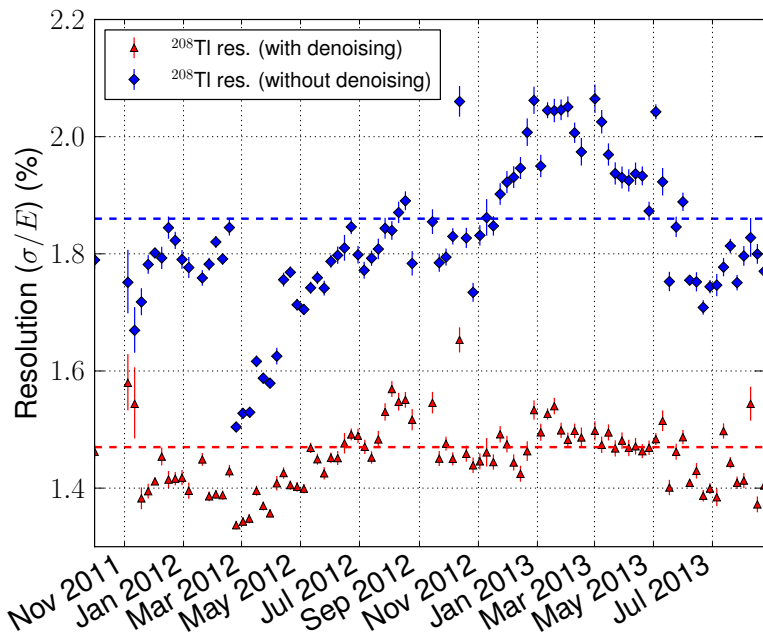
# $^{60}\text{Co}$ Source Spectrum (Before and After)



# $^{137}\text{Cs}$ Source Spectrum (Before and After)

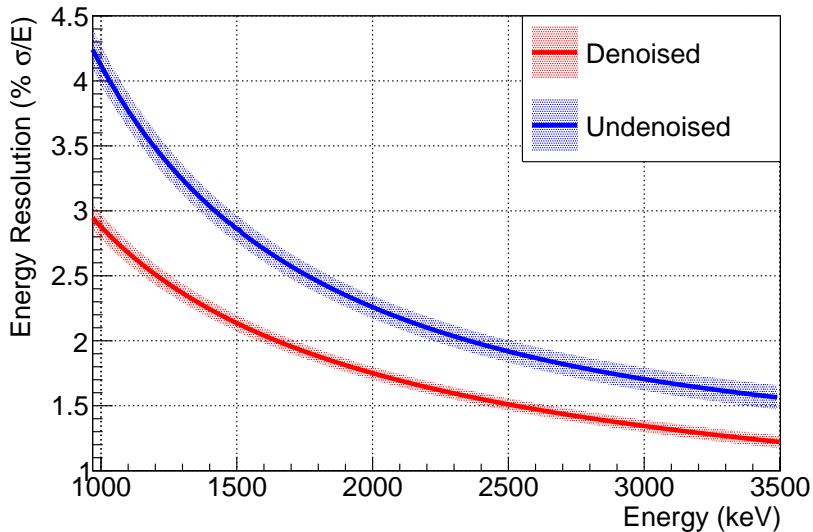


## Energy Resolution (Before and After)

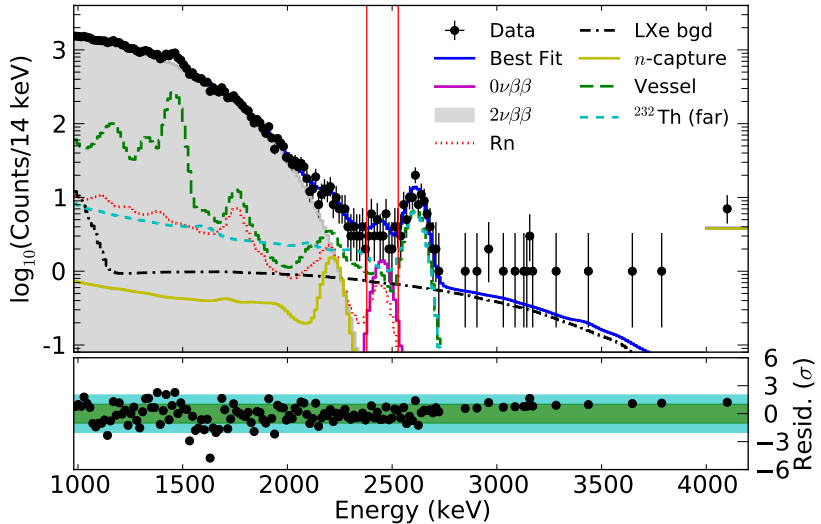




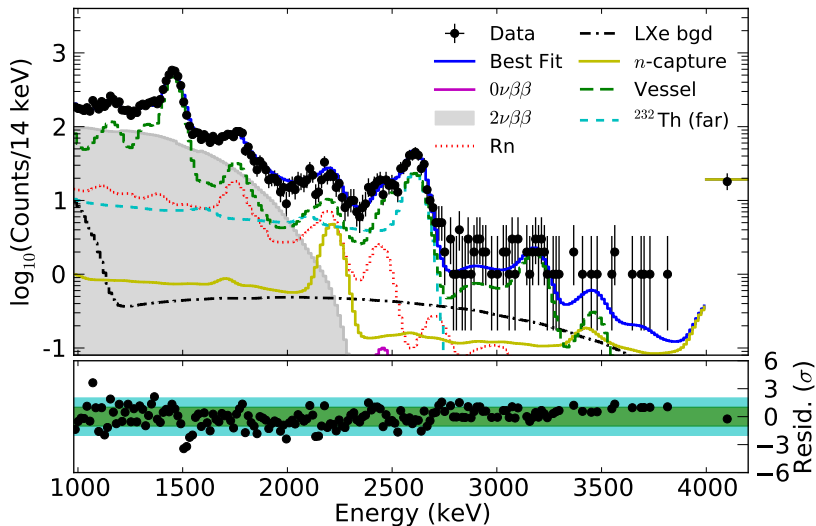
## Energy Resolution (Before and After)



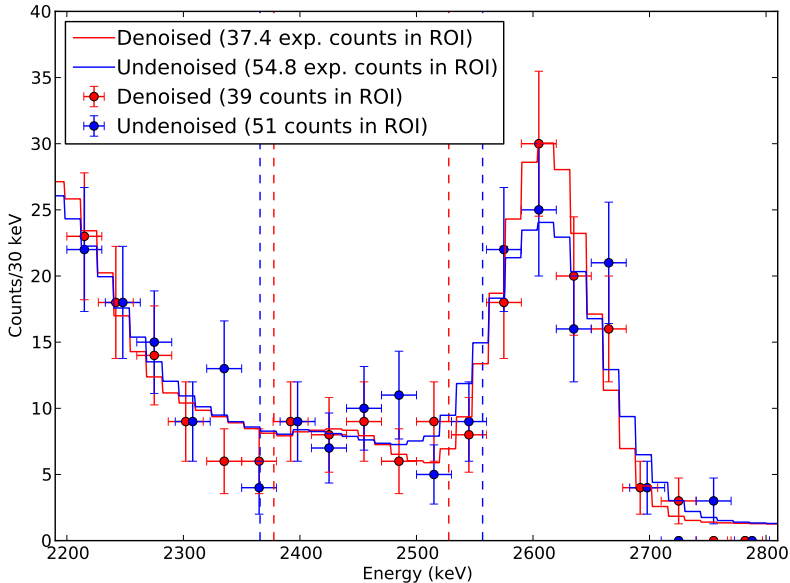
## Best Fit (Denoised)



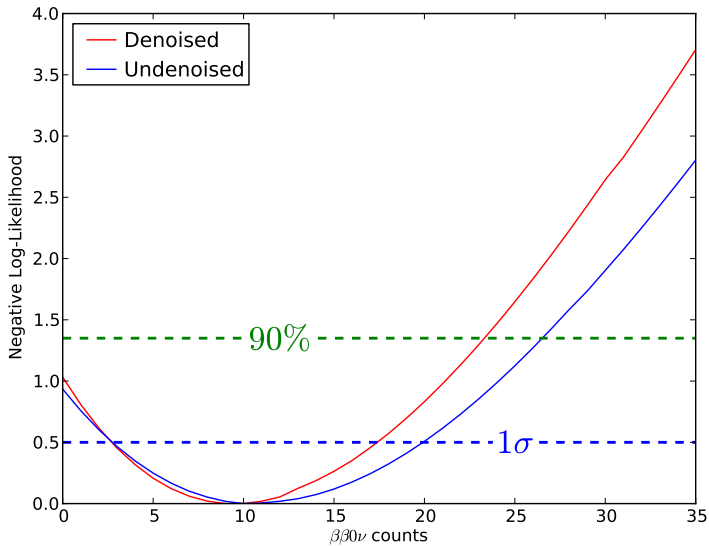
# Best Fit (Denoised), multi-site backgrounds

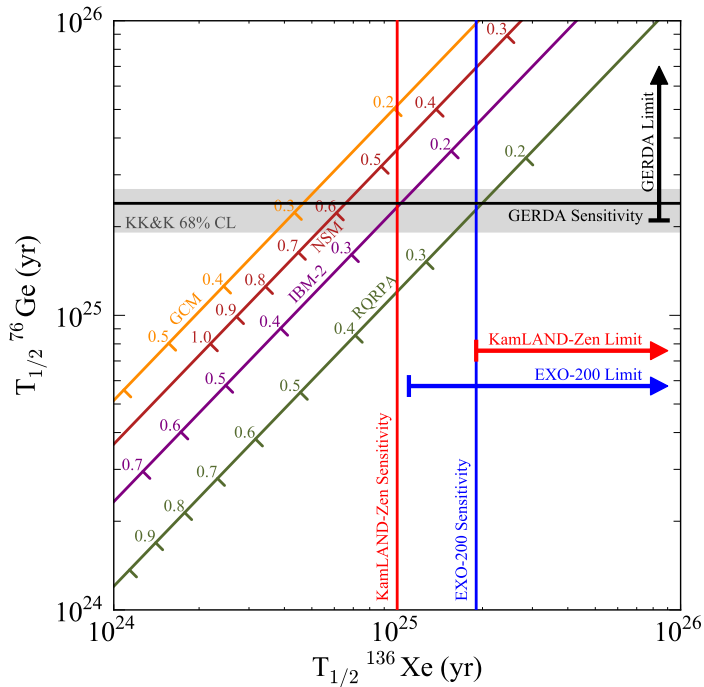


## Fits around $\beta\beta 0\nu$ (Before and After)



## Profile Likelihood (Before and After)





# The EXO-200 Collaboration

University of Alabama, Tuscaloosa AL, USA - D. Auty, T. Didberidze, M. Hughes, A. Piepke, R. Tsang  
University of Bern, Switzerland - S. Delaquis, G. Giroux, R. Gornea, T. Tolba, J-L. Vuilleumier  
California Institute of Technology, Pasadena CA, USA - P. Vogel  
Carleton University, Ottawa ON, Canada - V. Basque, M. Dunford, K. Graham, C. Hargrove, R. Killick, T. Koffas, F. Leonard, C. Licciardi, M.P. Roza, D. Sinclair  
Colorado State University, Fort Collins CO, USA - C. Benitez-Medina, C. Chambers, A. Craycraft, W. Fairbank, Jr., T. Walton  
Drexel University, Philadelphia PA, USA - M.J. Dolinski, M.J. Jewell, Y.H. Lin, E. Smith, Y.-R. Yen  
Duke University, Durham NC, USA - P.S. Barbeau  
IHEP Beijing, People's Republic of China - G. Cao, X. Jiang, L. Wen, Y. Zhao  
University of Illinois, Urbana-Champaign IL, USA - D. Beck, M. Coon, J. Ling, M. Tarka, J. Walton, L. Yang  
Indiana University, Bloomington IN, USA - J. Albert, S. Daugherty, T. Johnson, L.J. Kaufman  
University of California, Irvine, Irvine CA, USA - M. Moe  
ITEP Moscow, Russia - D. Akimov, I. Alexandrov, V. Belov, A. Burenkov, M. Danilov, A. Dolgolenko, A. Karelin, A. Kovalenko, A. Kuchnikov, V. Stekhanov, O. Zeldovich  
Laurentian University, Sudbury ON, Canada - B. Cleveland, A. Der Mesrobian-Kabakian, J. Farine, B. Mong, U. Wichoski  
University of Maryland, College Park MD, USA - C. Davis, A. Dobi, C. Hall  
University of Massachusetts, Amherst MA, USA - J. Abdollahi, T. Daniels, S. Johnston, K. Kumar, A. Pocar, D. Shy  
University of Seoul, South Korea - D.S. Leonard  
SLAC National Accelerator Laboratory, Menlo Park CA, USA - M. Breidenbach, R. Conley, A. Dragone, K. Fouts, R. Herbst, S. Herrin, A. Johnson, R. MacLellan, K. Nishimura, A. Odian, C.Y. Prescott, P.C. Rowson, J.J. Russell, K. Skarpaas, M. Swift, A. Waite, M. Wittgen  
Stanford University, Stanford CA, USA - J. Bonatt, T. Brunner, J. Chaves, J. Davis, R. DeVoe, D. Fudenberg, G. Gratta, S. Kravitz, D. Moore, I. Ostrovskiy, A. Rivas, A. Schubert, D. Tosi, K. Twelker, M. Weber  
Technical University of Munich, Garching, Germany - W. Feldmeier, P. Fierlinger, M. Marino  
TRIUMF, Vancouver BC, Canada - J. Dilling, R. Krucken, F. Retière, V. Strickland

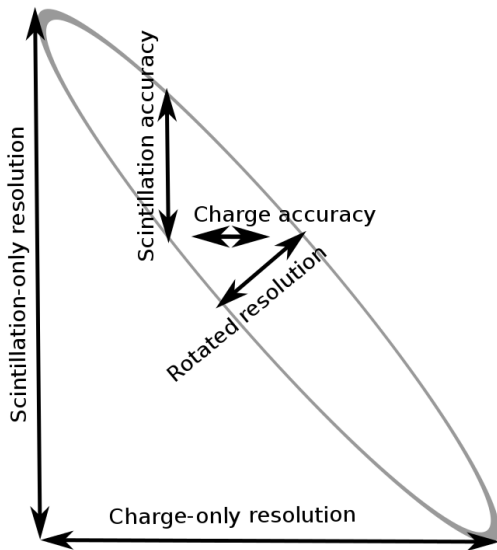
Thank You!

Questions?



# Backup Slides

## Anticorrelated Scintillation/Charge

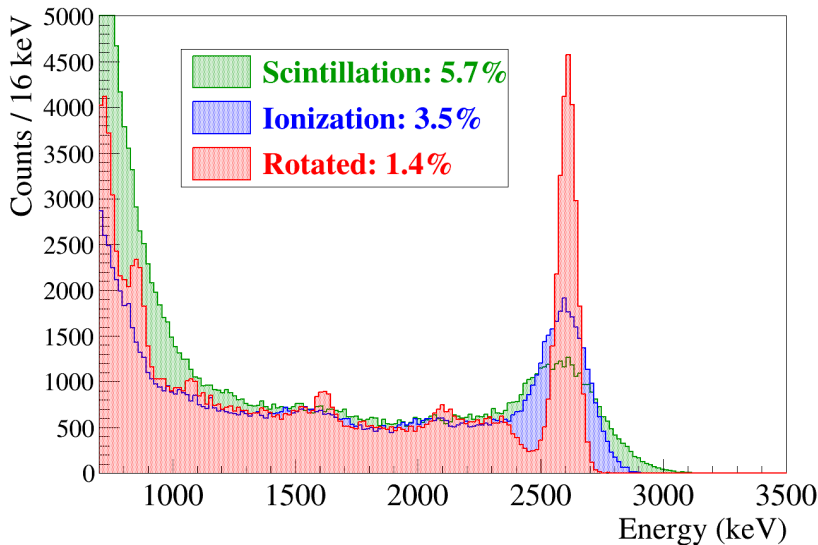


## Why not the Anscombe transformation?

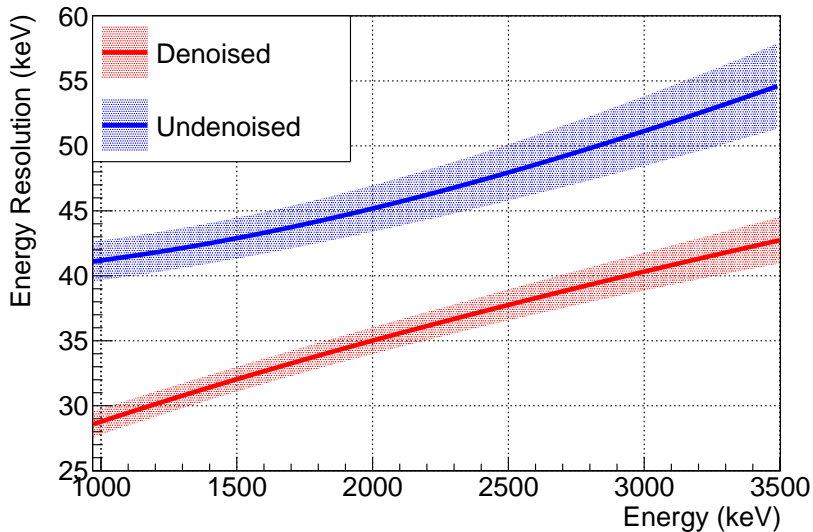
The Anscombe transformation (and generalizations) are designed to transform Poisson noise into Gaussian noise. However, for this purpose “noise” is defined to mean noise uncorrelated with pulse. We, on the other hand, are looking to understand noise which is correlated with the pulse, and then minimize its effect.

So, the Anscombe transformation doesn't really help us here.

## Impact of Rotated Energy



## Energy Resolution (Before and After)



# Mean Sensitivity

