Pacific Northwest
NATIONAL LABORATORY

SuperCDM

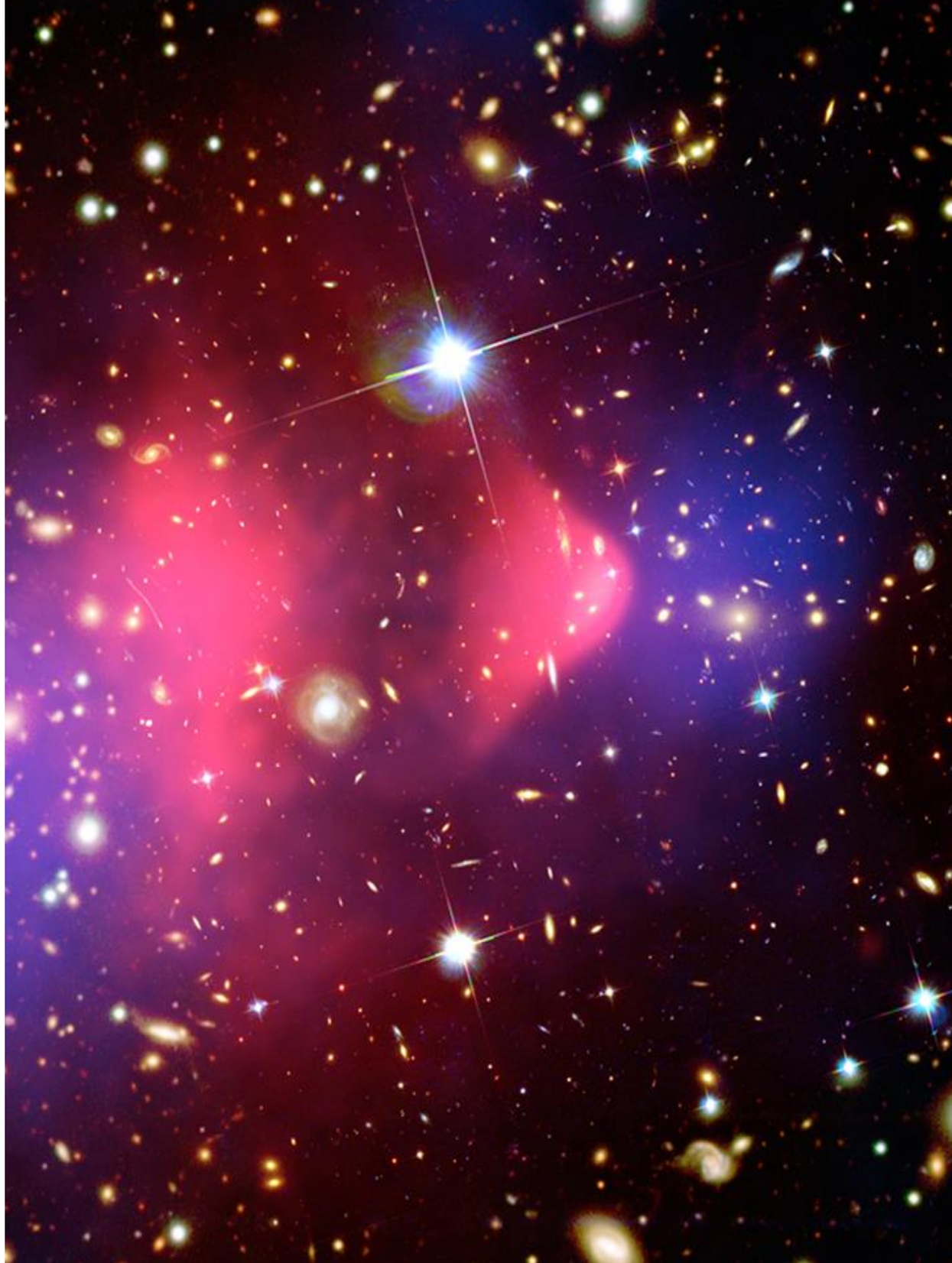
Blinding Strategies for Dark Matter Searches

Ben Loer

Pacific Northwest National Laboratory

U.S. DEPARTMENT OF ENERGY BATTELLE

PNNL is operated by Battelle for the U.S. Department of Energy

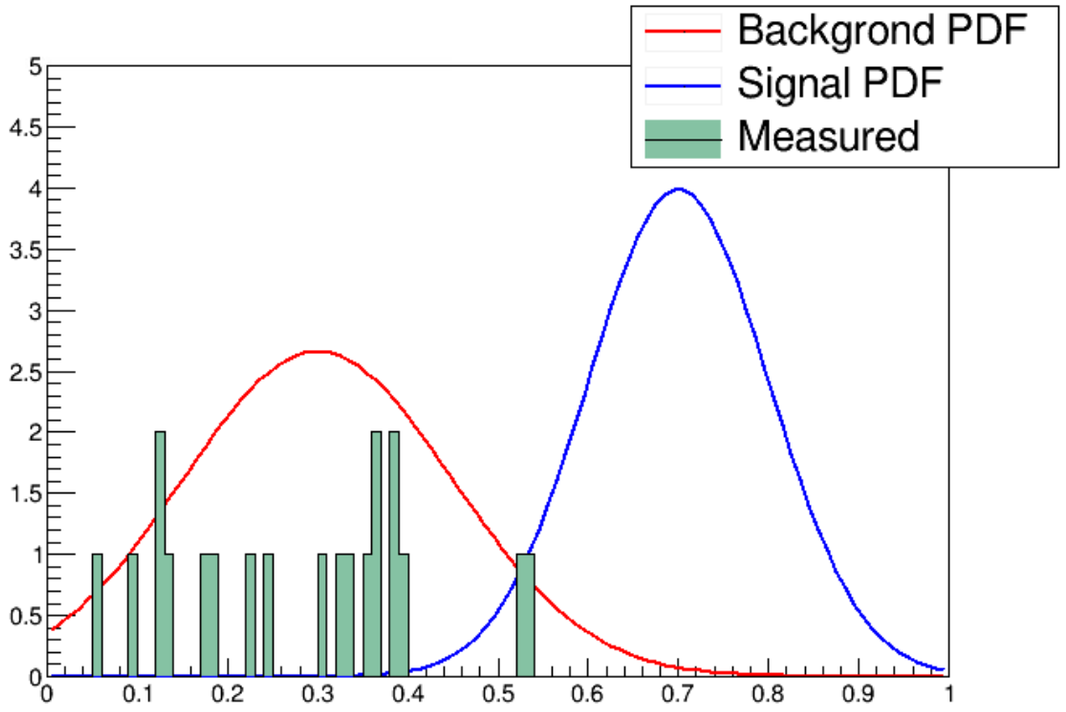


Pacific Northwest NATIONAL LABORATORY

Outline

- Why do a blind analysis?
- Blinding strategies
  - Strengths and weaknesses
  - Experimental anecdotes
- Technical implementations
  - CDMSlite salting
  - ADMX signal injection
- Conclusions

“If efficiency is calculated correctly, there is no bias”



- Cut and count analysis
  - Where to put the cut?
  - Signal efficiency only changes by 2% to include/exclude both events, based on PDF
- Likelihood analysis
  - Susceptible to tuning initial fit parameters, stopping when the answer “looks right”
- *May* be possible to correctly recalculate true efficiency accounting for “biased” selection using e.g. toy Monte Carlo

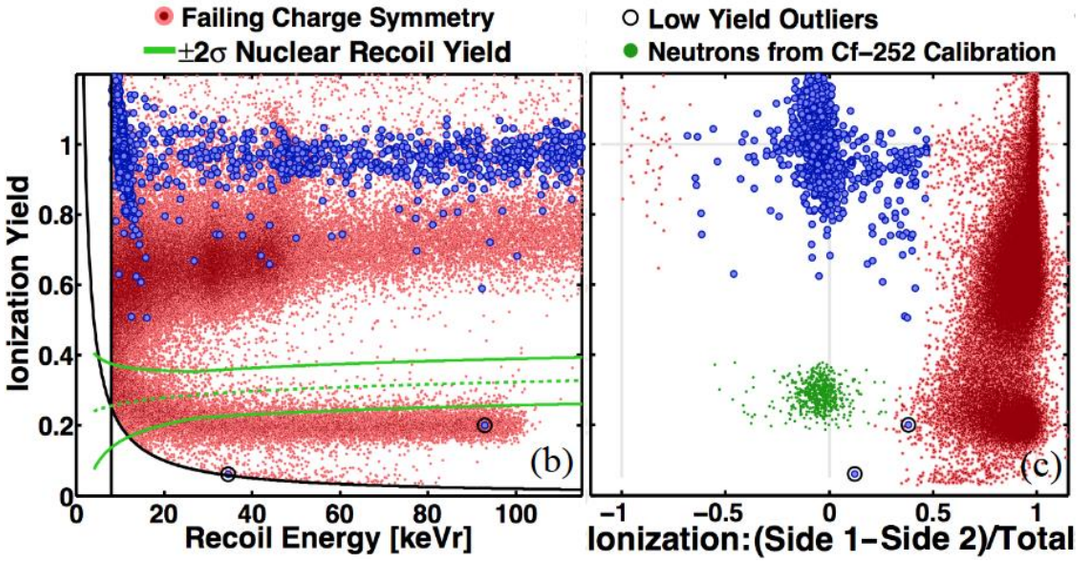
Blinding Strategies

- **Data division / prescaling**: look at only a fraction of the data
- **Box blinding**: avoid looking at signal region of interest until analysis finalized
- **Hidden parameter**: Perturb a key variable in an unknown fashion
- **Salting**: Add unknown number of fake signal-like events
- **Hiding**: Remove unknown number of events

Data division / prescaling

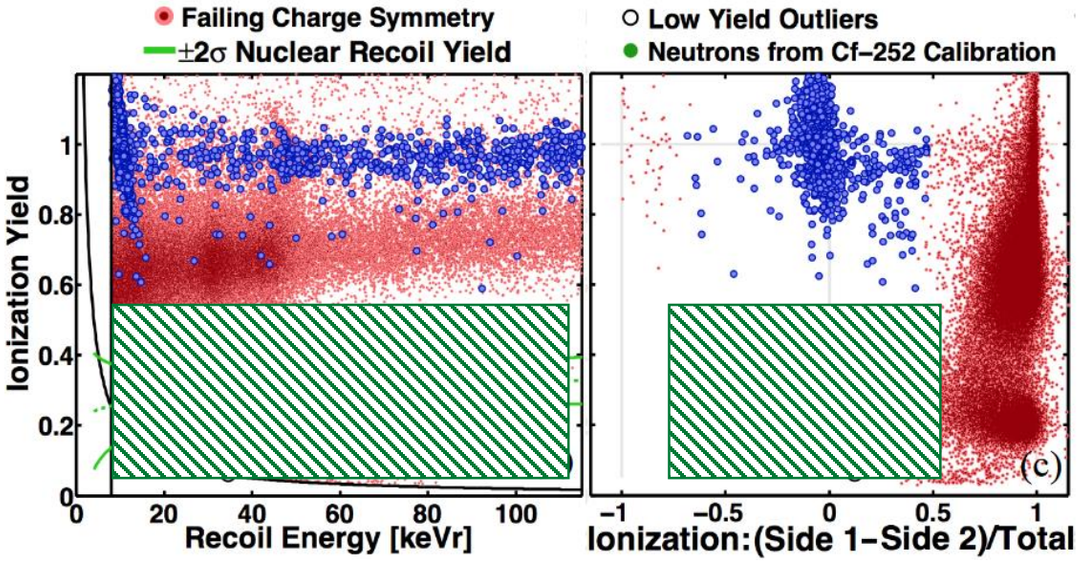
- Develop analysis on ~25% of data, only open remaining once finalized
- Easiest to implement
- No good rule to determine open fraction size
  - Large enough to have good chance to see rare backgrounds
  - Small enough to keep closed portion statistically independent, limit livetime loss
- Include or discard open portion? Both have been used
  - If include, *should* correct efficiencies based on expected level of statistical fluctuations. In practice, usually just choose “small enough” fraction
  - If discard, reduce sensitivity. Less important if background- or systematics-limited
- Does not protect against systematic bias, e.g.
  - background described equally well by two PDFs; choose one that gives expected signal
  - Tune initial fit parameters.

Box blinding



- Determine signal region in some plane via calibration, simulation, “pre-science” data
- Hide all data in a larger region around signal
- Gold standard for rare event searches
- Use sidebands to update models of behavior in blind region
  - Particularly critical to handle differences to calibration from e.g. high trigger rates
- Tension between
  - Want tight box to get most information from sidebands
  - Loose box prevents premature rejection of possibly valid signal region
- Developing entire analysis on calibration/simulation is an extreme limit

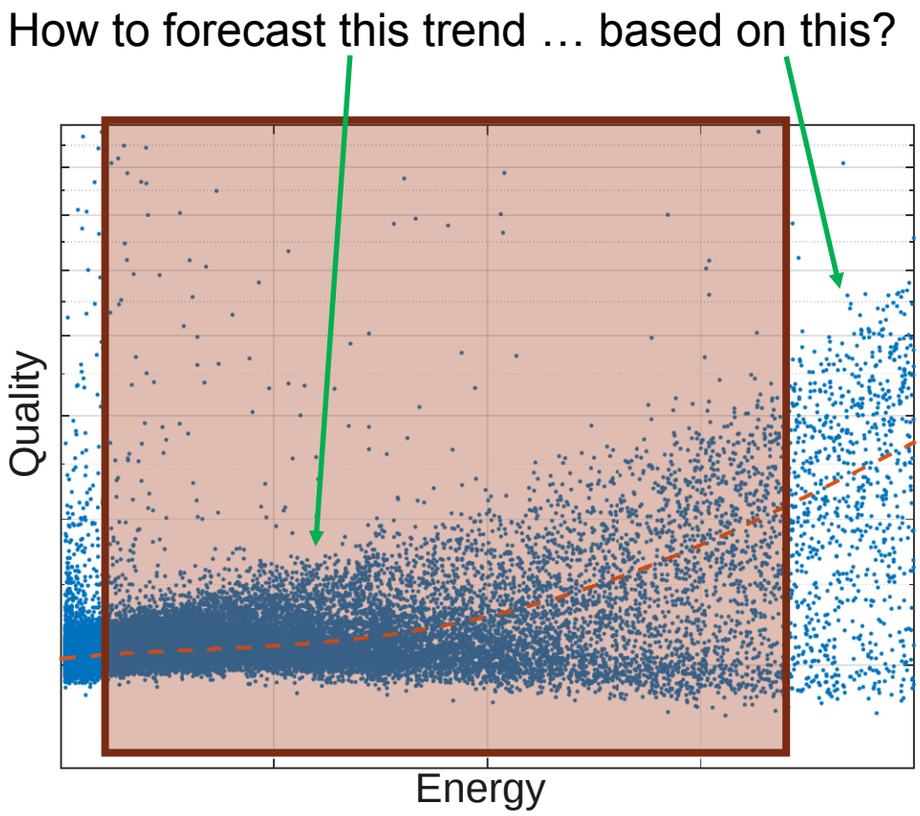
Box blinding



- Determine signal region in some plane via calibration, simulation, “pre-science” data
- Hide all data in a larger region around signal
- Gold standard for rare event searches
- Use sidebands to update models of behavior in blind region
  - Particularly critical to handle differences to calibration from e.g. high trigger rates
- Tension between
  - Want tight box to get most information from sidebands
  - Loose box prevents premature rejection of possibly valid signal region
- Developing entire analysis on calibration/simulation is an extreme limit

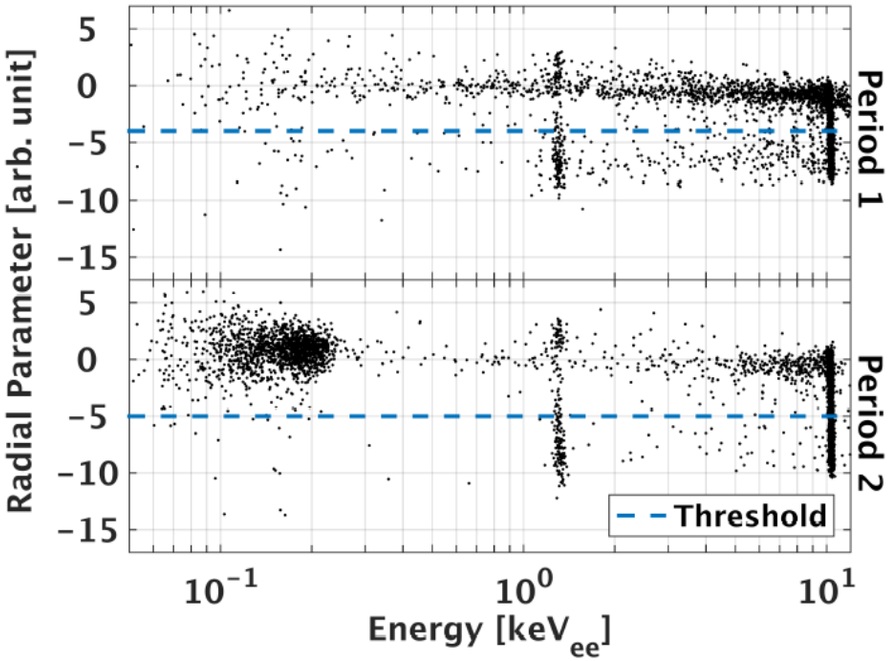
Concerns with box blinding

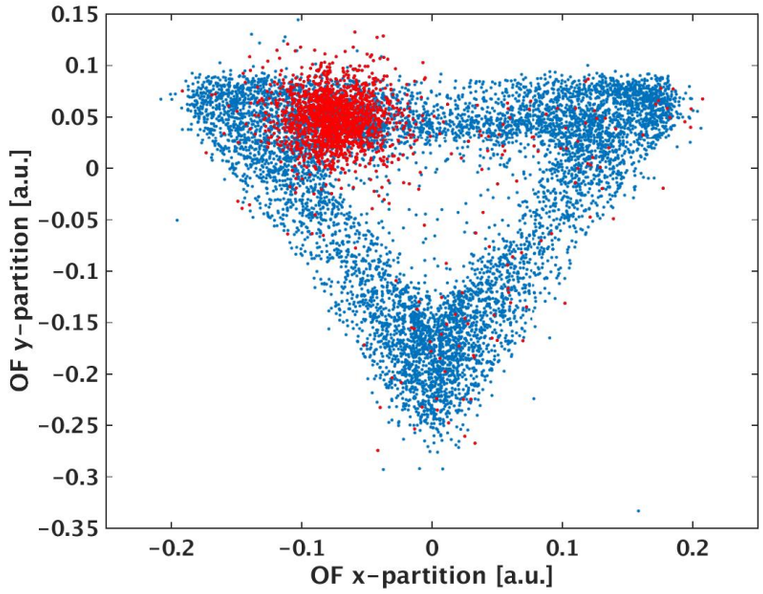
- Where to draw the box? Need a good data model before analysis starts
- Box too large if search is open-ended
  - E.g. relevant energy range for 1 vs 100 GeV WIMP
- If no good discrimination variable, sidebands very limited
  - Currently the case for most sub-GeV dark matter searches
- Unexpected backgrounds
- Managing data quality



CDMSlite Run 2 unexpected background

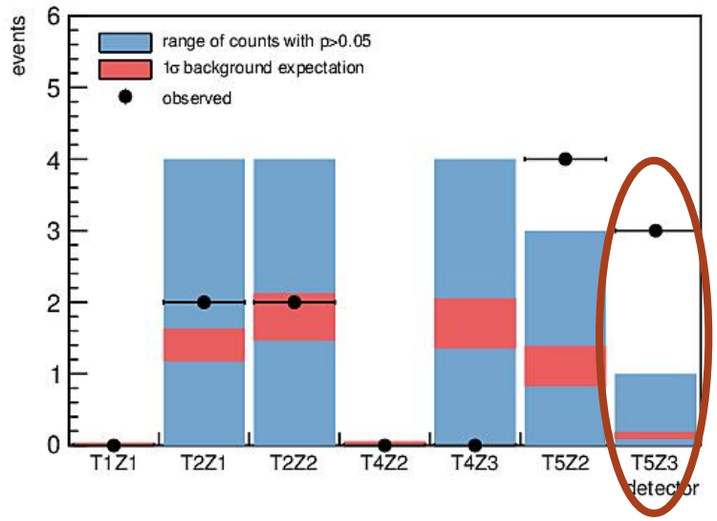
- Source of clustered events in signal region spontaneously appeared
- Would not have been visible in box blinding analysis

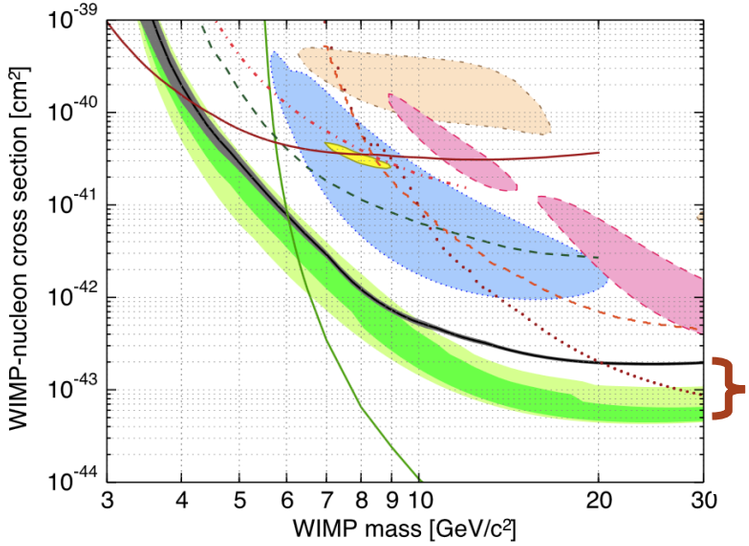




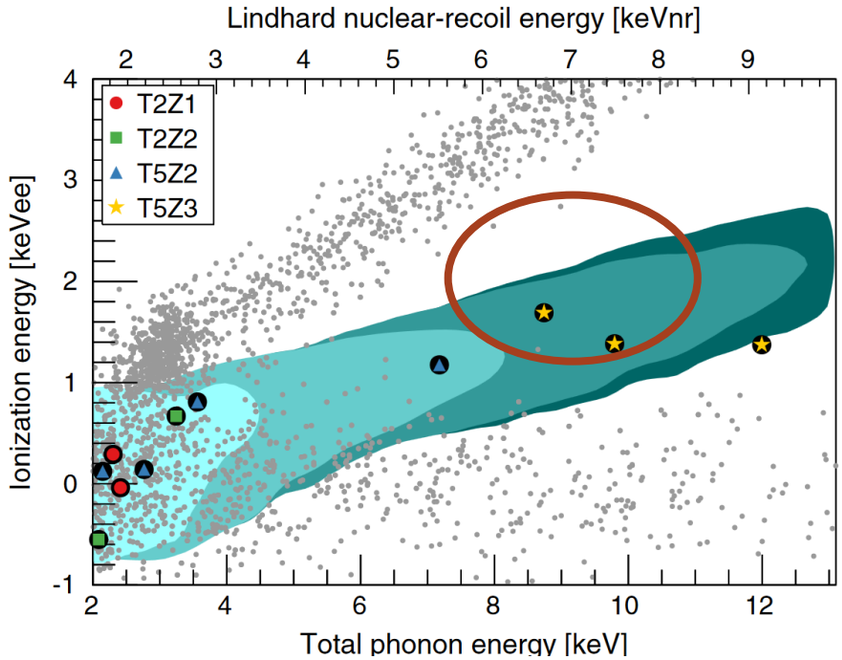
SuperCDMS Soudan unexpected background

- Significant excess of events seen in detector with shorted ionization guard electrode
- Events appear to be of good quality
- On post-inspection, likely background model incorrectly treats grounded electrode



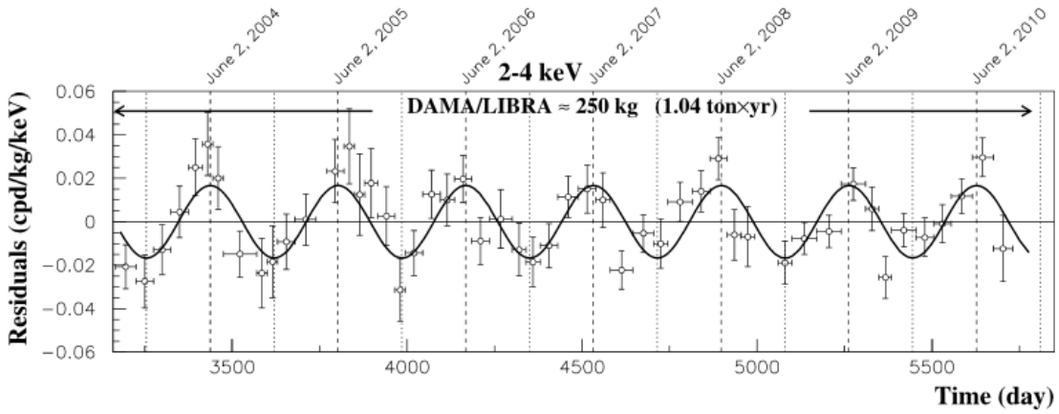


Sensitivity gap from T5Z3 events



Hidden parameter

- Modify a measured variable with a hidden but known scale or offset
  - E.g. scale recorded particle size (mass) in Millikan oil drop experiment
- OR hide a parameter needed for a final calculation step
  - E.g. If making precision measurement on “known” cross section, hide dataset livetime
- Difficult to apply to a rare event search
- Scale and offset timestamps could be an effective way to blind frequency and phase for annual modulation search



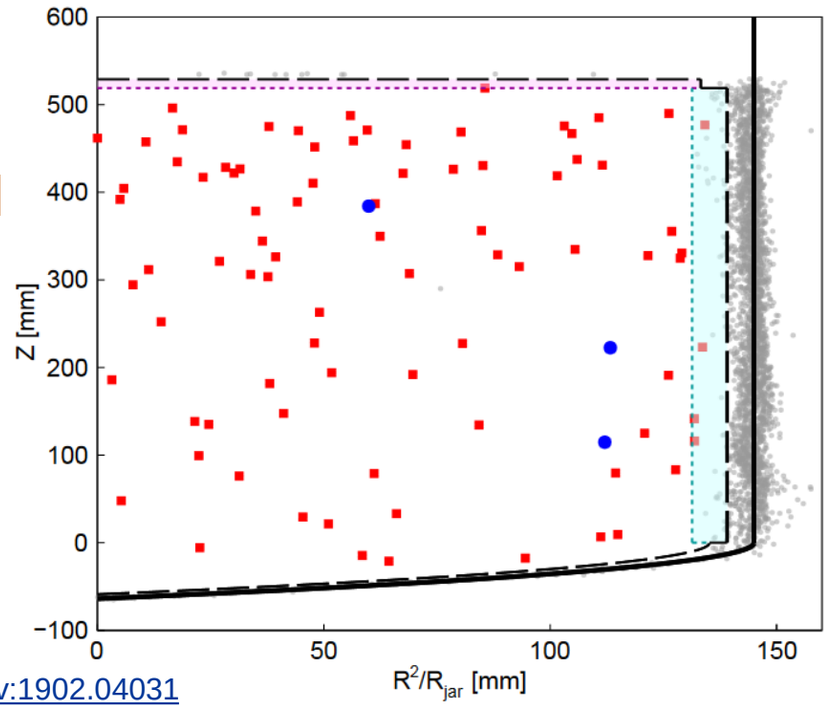
R. Bernabei et al. (DAMA/LIBRA), Eur. Phys. J.C73,2648 (2013).

Salting / data removal

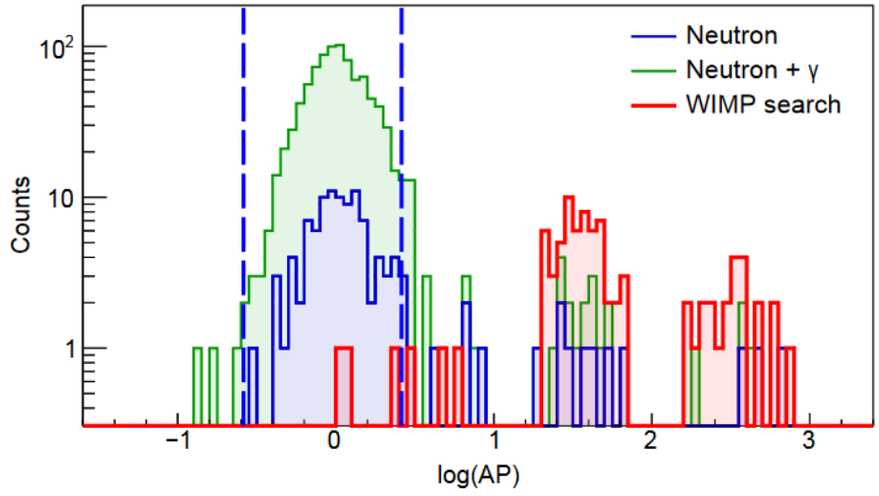
- Add fake signals to combat bias toward overly strong limits
  - E.g. keep adding cuts until signal ROI is empty
- Hide a fraction of events to combat bias toward overly strong signal claims
  - E.g. “Fix” cuts if they reject “too much” signal
- Best when used together
- Salting requires enough knowledge of signal characteristic to fool analyzers, before analysis starts
- Sources for fake events:
  - Unsubtracted background: e.g. SNO phase 2 “hole” in muon veto
  - Calibration events: e.g. LUX
  - Simulated events: e.g. CDMSlite (sort of)
  - Signal injection: e.g. ADMX

PICO-60
Hidden Discriminat

- Primary signal/background identifier for PICO is acoustic parameter (AP)
- For PICO-60 analyses, AP is hidden until final step
- Effectively “salted” analysis with background (alpha) events
- Required tuning AP cut on calibration, small subset of open WIMP search data
- Lose a potential handle to identify data quality problems early

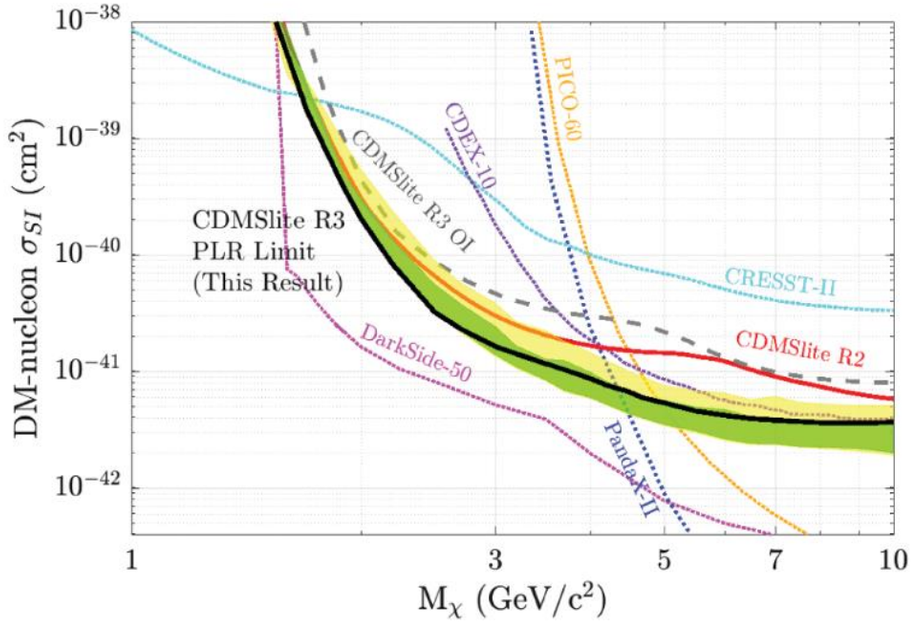


arXiv:1902.04031



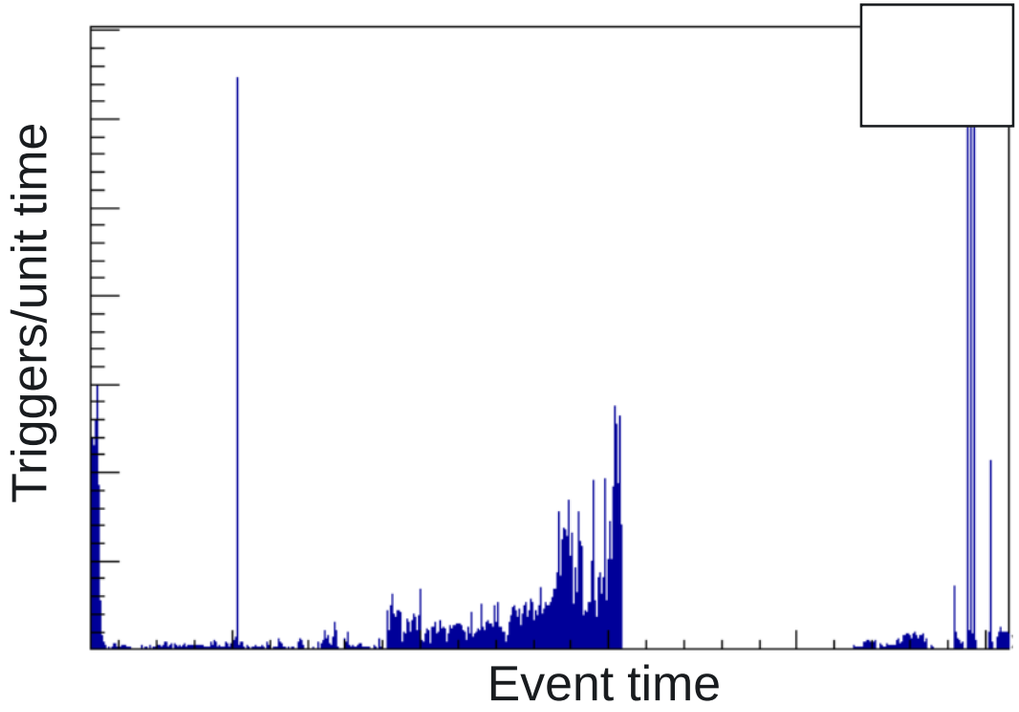
CDMSlite Run 3

- Try two new (to us) techniques:
  - Salting for blind analysis
  - Profile likelihood result
- Modeled heavily on LUX approach
- Couldn't use calibration data directly due to subtle differences with WIMP search data
- Instead construct fake pulses from templates and sampled baselines
- Complicated by sequential event IDs: have to replace events, not insert new ones



CDMSlite Salting method

- Pick a WIMP mass, estimate our 90% CL predicted sensitivity, calculate how many events would be in the final spectrum, call that N
- Randomly select between N and 3N events in raw data to replace with salt
- Weight selection to be uniform in time
  - NOT the same thing as uniform in event number!



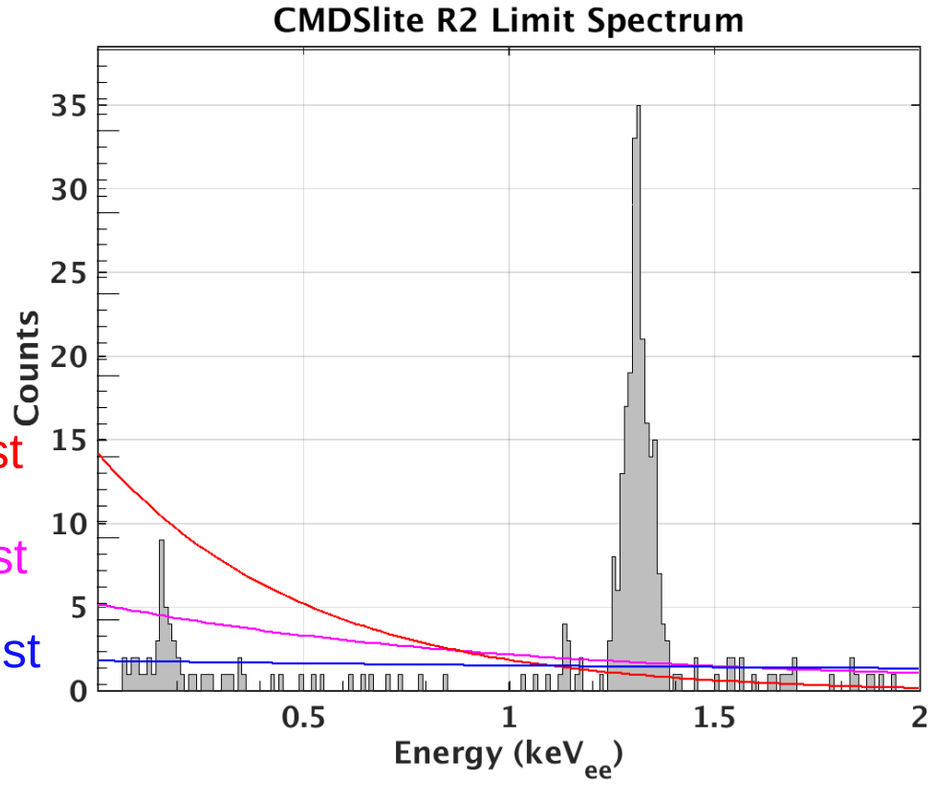
CDMSlite Salt Spectrum

- Assign energy from exponential + constant spectrum
- Slope varies from 0.5-2 keV, expo/constant varies from 1/3 to 3

3*exp(-x/0.5)+const

exp(-x)+const

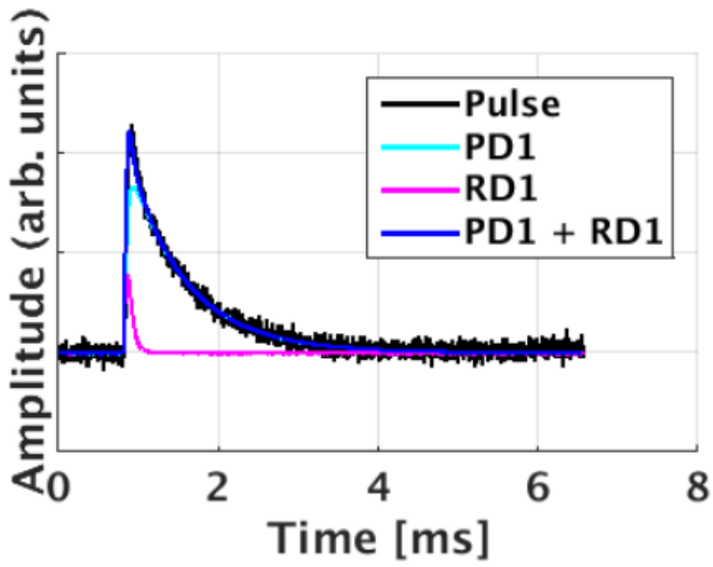
1/3*exp(-x/2)+const

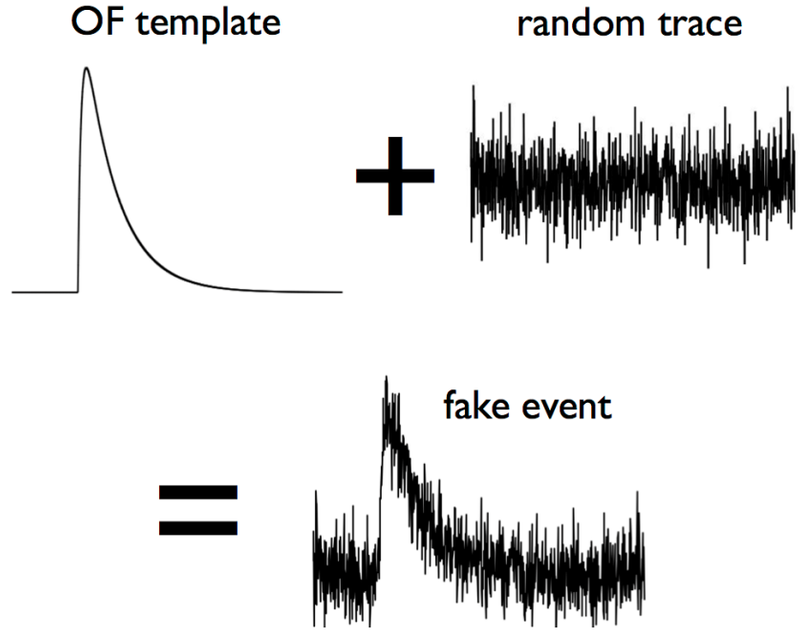


Salt amplitude exaggerated

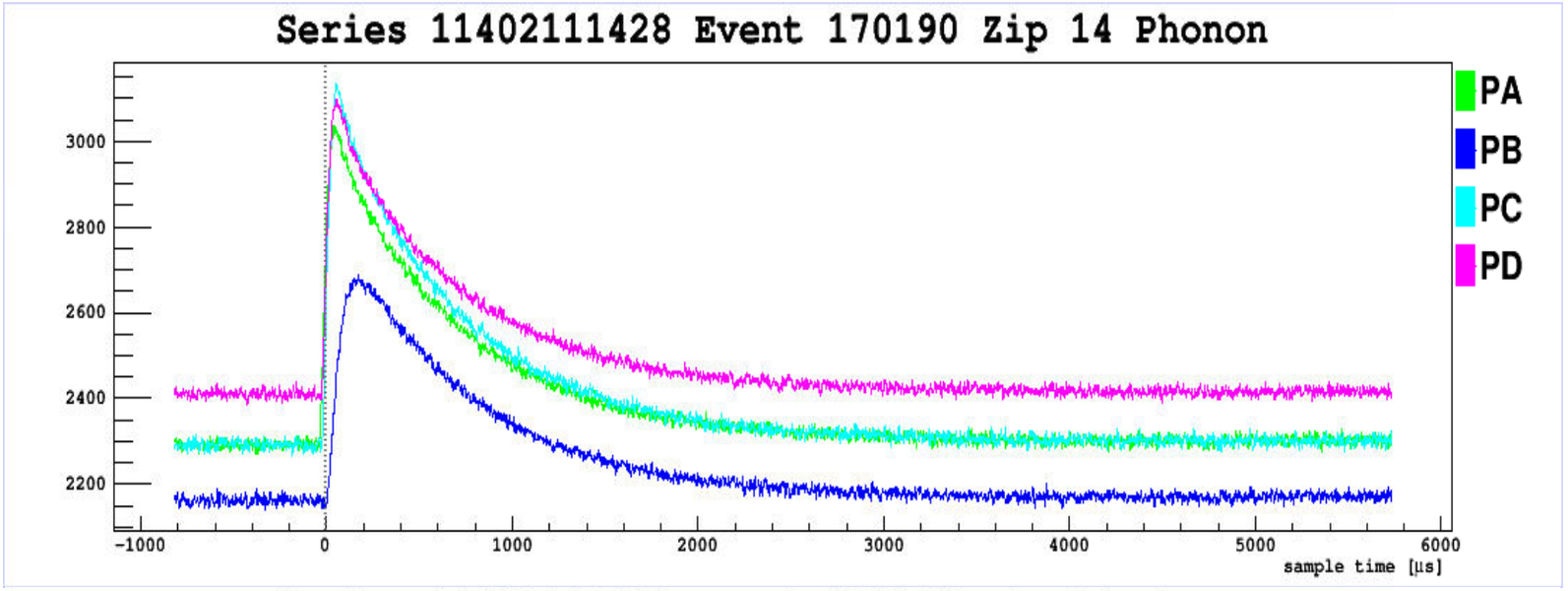
Creating fake pulses

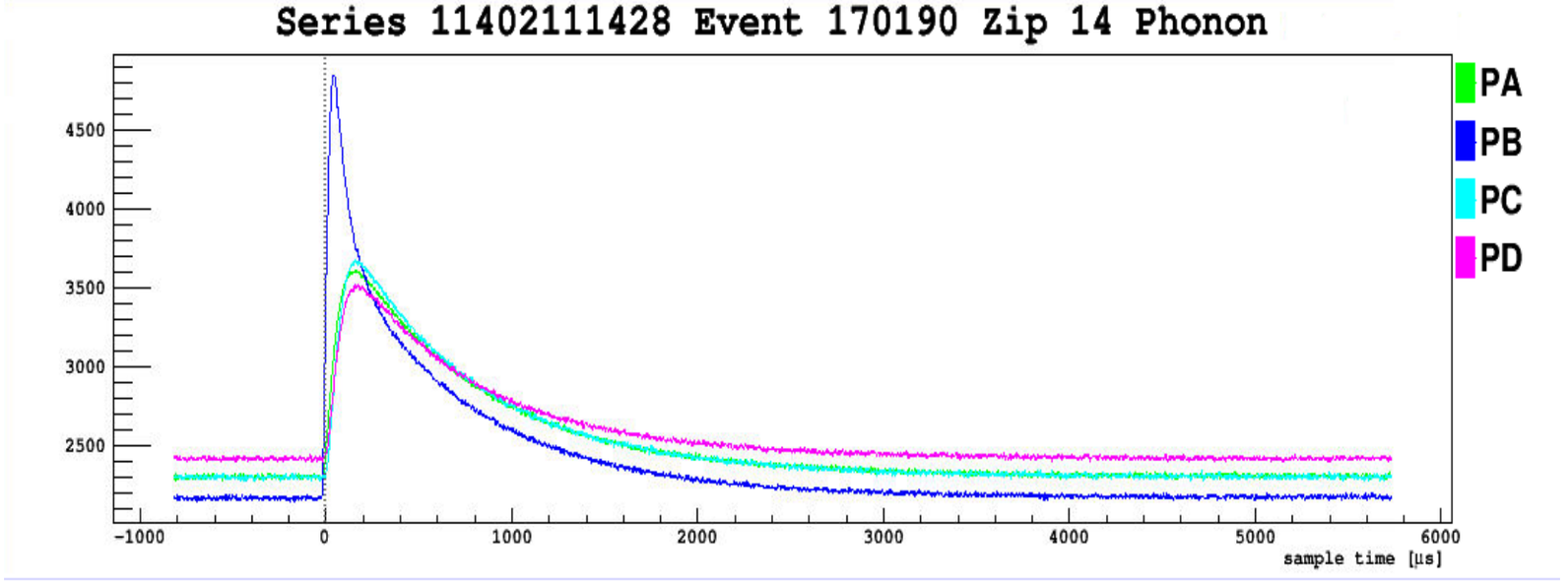
- Select calibration events with similar energy
  - Want to avoid modeling distributions as much as possible
  - Lots of energy dependence, so tight energy selection
- Fit each channel of calibration events to 2 templates (average + residual)
- Scale template amplitudes to new target energy (keeping same trigger delay values)
- Add to empty trigger to sample noise





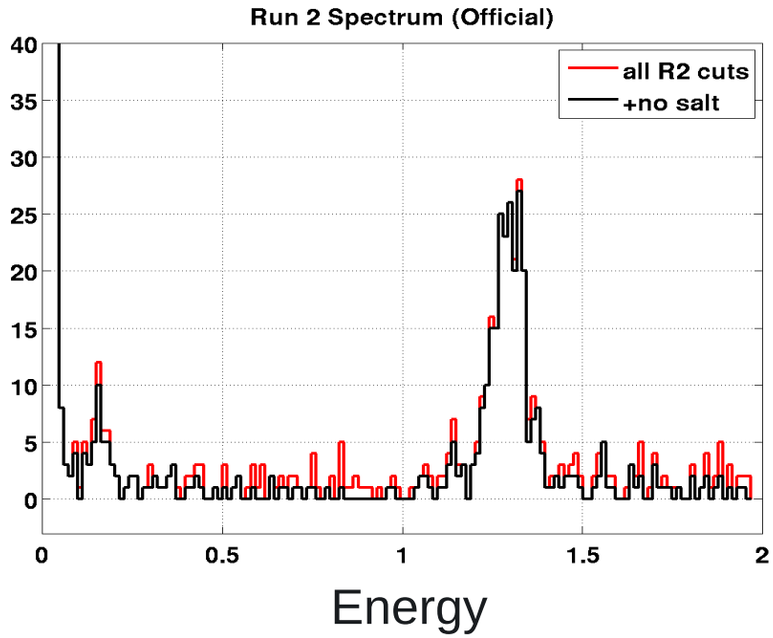
Can you spot the difference?

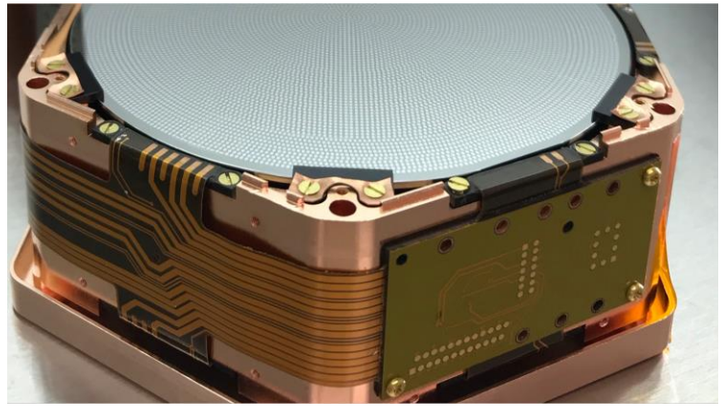




CDMS Salting and the future

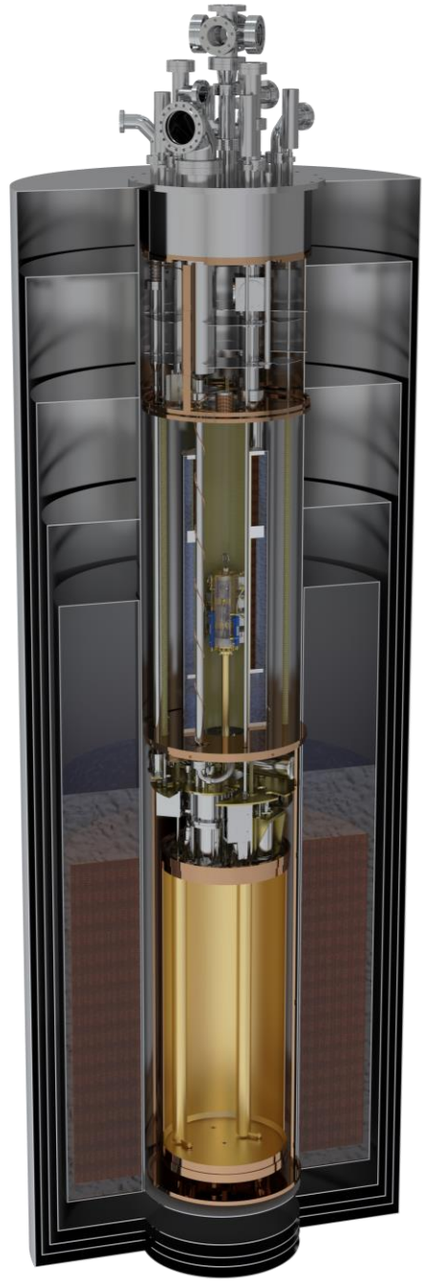
- Run 3 salted data was successfully generated, analysis published this year
- SuperCDMS SNOLAB detectors are brand new technology, not yet well enough understood to attempt salting
- Plan to use data division: open and reject 25% of science dataset for earliest analysis
- New DAQ will insert empty headers into raw data to make inserting salt events easier





ADMX

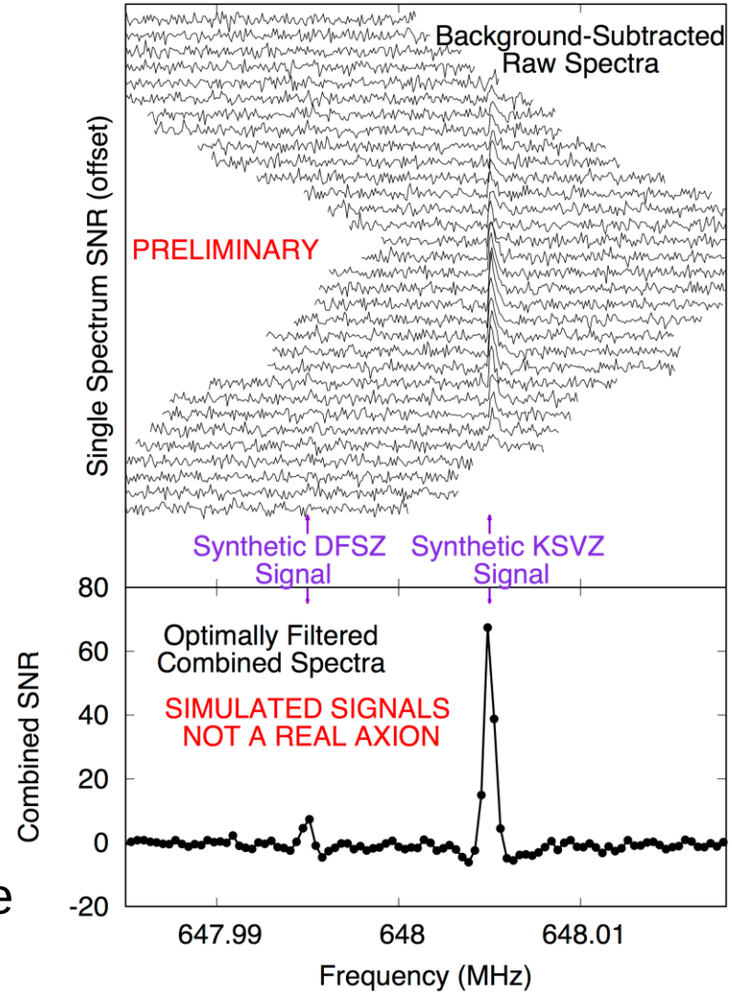
- **A**xion **D**ark **M**atter E**x**periment
- Axion haloscope
- Search for axion coupling to B-field photons
- High Q resonator in strong B field
- Signal is a spike in power at frequency of axion mass



Many thanks to Noah Oblath for the following slides

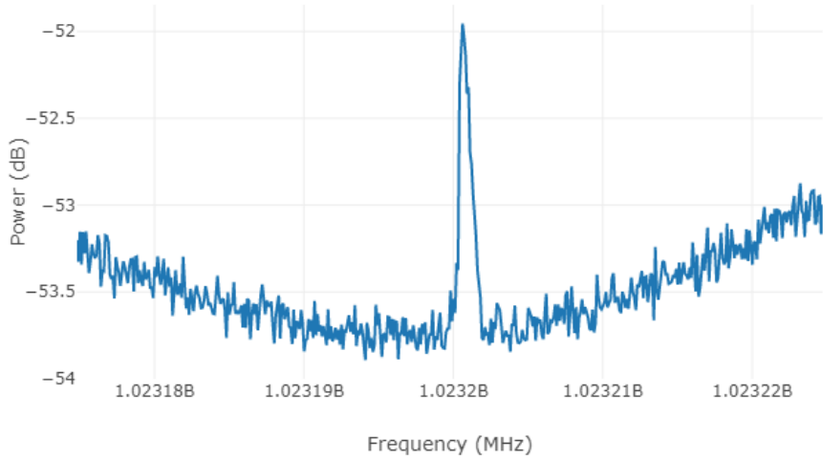
Context: ADMX Operations

- Live analysis
  - Cavity frequency scanned, pausing for 100 seconds at each frequency
  - Regions with power above trigger threshold are flagged as potential candidates
    - Could be statistical anomalies, external RF leakage, synthetic injected axions, or real detected axions
  - Candidates are rescanned to see if they persist
  - For persistent candidates, perform confirmation tests:
    - Switch to resonant mode that doesn't couple to axions
    - Turn B field down (axion power scales as B^2)
- Offline analysis
  - Vary bin size to look for higher-frequency structure
  - High resolution analysis looks for ultra-sharp lines

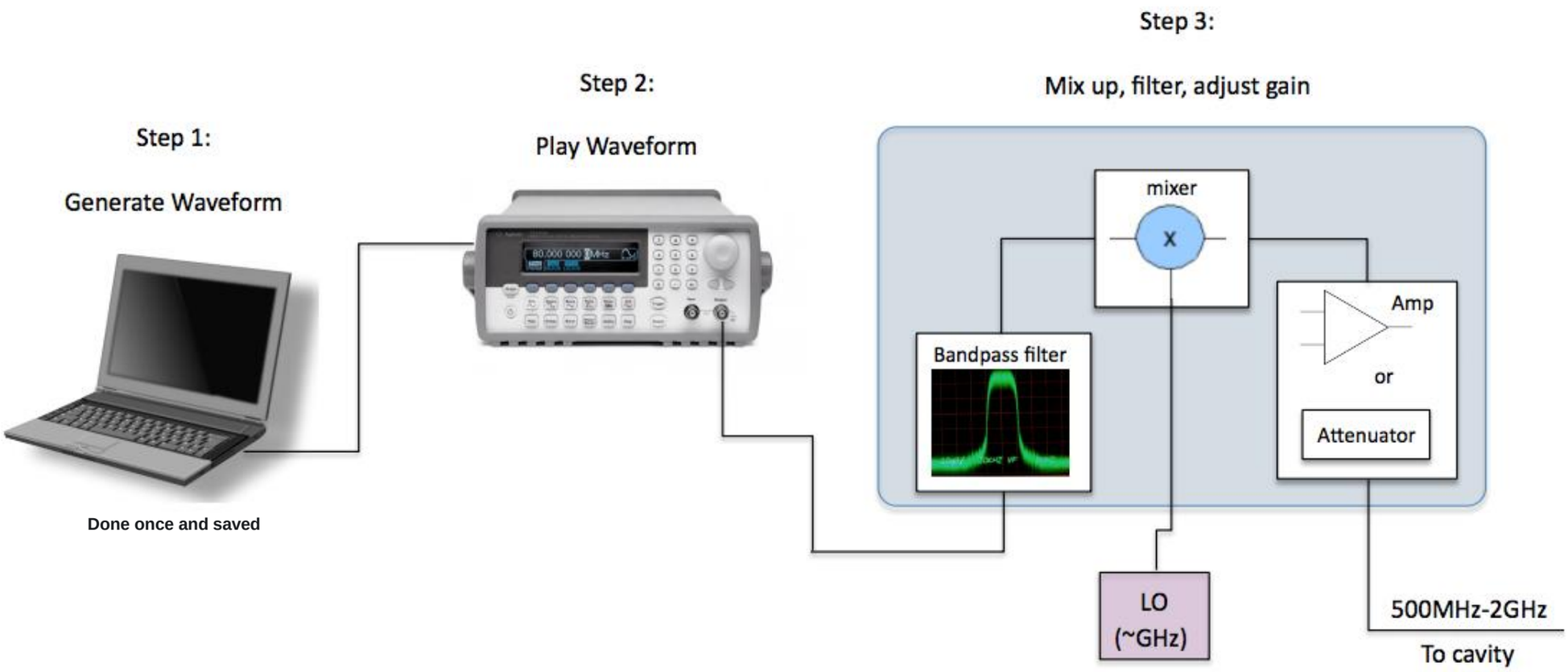


Blind Signal Injections at ADMX

- Purpose: to blind operations of the experiment & validate the candidate identification procedure
- Problem: we need to minimize the possibility of biases influencing our data-taking
  - Our data can be seen live
  - The data-taking procedure involves manual interactions
- Two injection categories are used
  - Primary blinding
    - Several (~4) per 10 MHz, on average
    - Unblinded after candidate search
  - Secondary blinding
    - Once per 1-2 months, on average
    - Unblinded before the magnet-ramp procedure

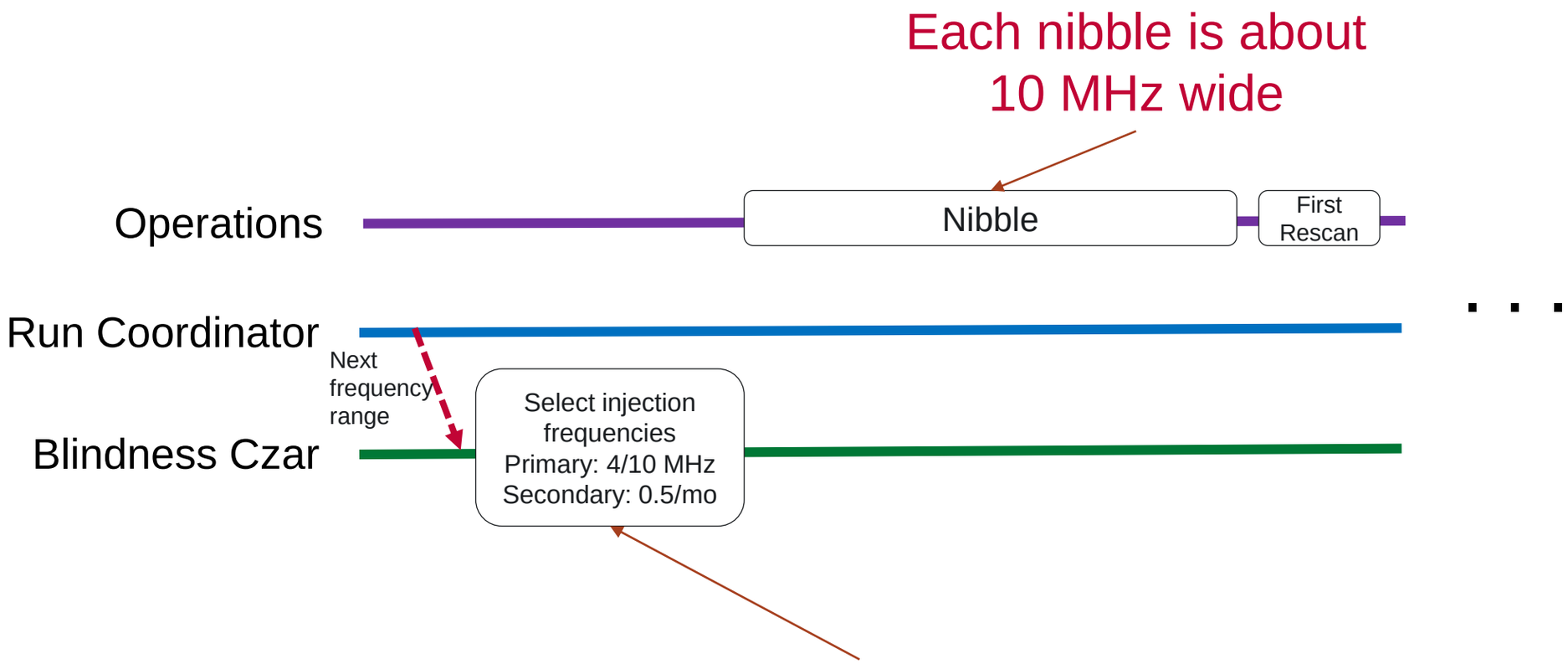


ADMX Blind Injection System

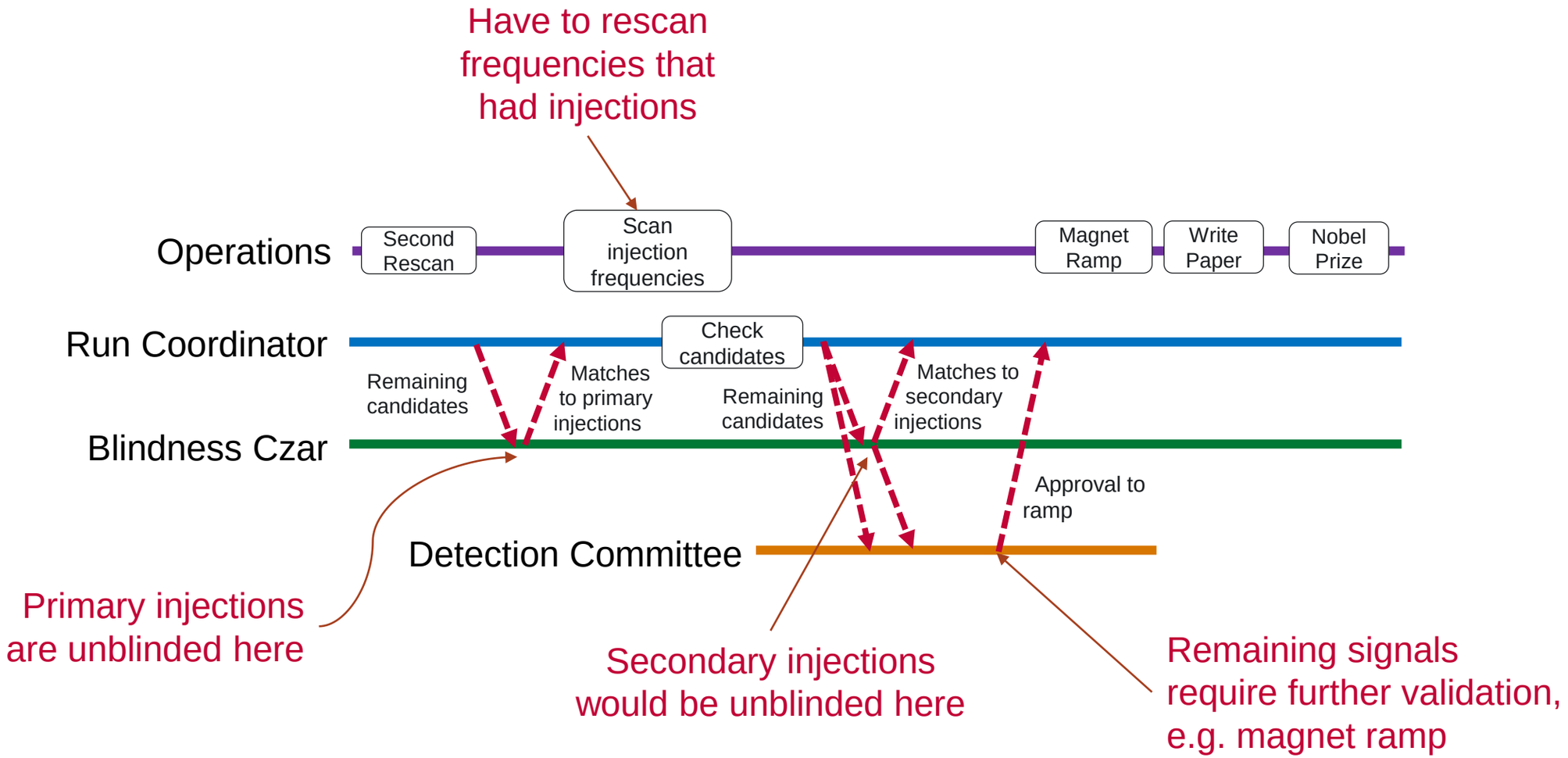


- Signals are played by an arbitrary waveform generator, filtered, and mixed up to the appropriate frequency
- Signals are then injected directly into the cavity via a second, weakly-coupled, antenna

ADMX Blind Injection Procedure I



ADMX Blind Injection Procedure II



ADMX Other Details

- This is the only blinding procedure in use.
- Each injection impacts runtime, so they're used judiciously.
- Secondary injections have a large runtime impact, so they're pretty rare.
- Signal frequencies and strengths are random.
- Signal strengths are around that of a DFSZ axion, up to about the strength of a KSVZ axion
- Experiment controls website has a signal injection page visible only to the Blindness Czar
- Truth information is recorded automatically in a database that other collaborators do not look at
- It's also recorded in a spreadsheet that includes extra information such as when unblinding was performed and the detected-candidate information

Pacific Northwest NATIONAL LABORATORY

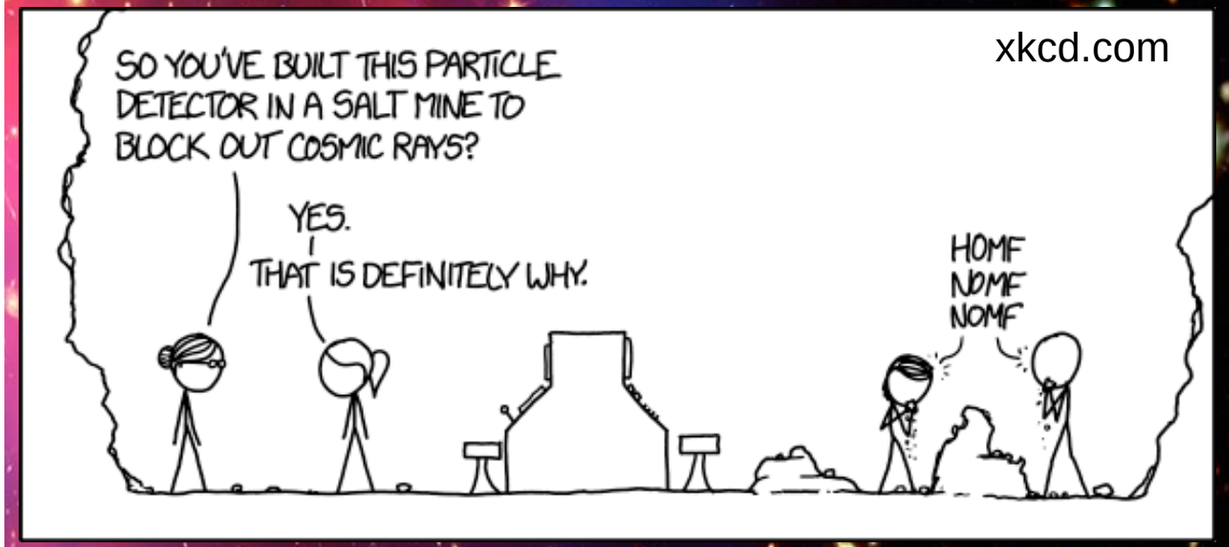
Conclusions

- Variety of possible blinding techniques
- May be limited by technology, analysis methodology, ease of implementation
- Most still admit some possibility for biased analyses
- Surprisingly hard to find “recipes”, e.g.
  - How much salt to add?
  - What fraction of dataset to open in prescaled analysis?

Pacific Northwest
NATIONAL LABORATORY

SuperCDM

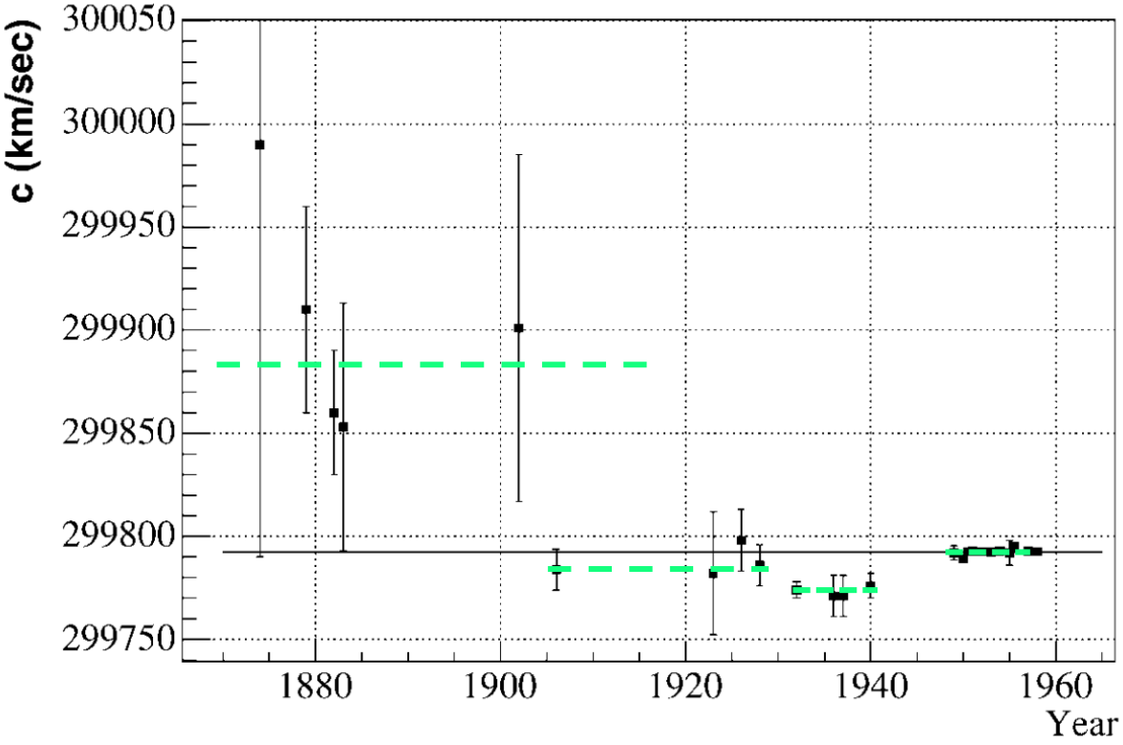
Thank you



Further Reading

- Klein and Roodman, “Blind Analysis In Particle Physics.” Annu. Rev. Nucl. Part. Sci. 2005. 55:141–63. doi: 10.1146/annurev.nucl.55.090704.151521
- Oser, “Blind Analysis – or – The Answer’s Not in the Back of the Book.” UBC Colloquium, 2010. https://www.phas.ubc.ca/~oser/blind_colloq.pdf

Why blind analysis?



Klein and Roodman, *Ann. Rev. Nucl. and Part. Systems*, **55**, 141-163 (2006)

- Likely confirmation bias in the history of speed of light measurements
- Results tend to agree much better with previous measurement than accepted value
- But does it apply to discovery searches with no prior value?

Fallacy: Multiple independent analyses

- Can be great for many reasons!
- Competitive drive, encourage original thinking
- Reduces “group think”
- But still easy to independently manipulate analyses toward expected result
  - E.g. if expecting to set limit, keep inventing cuts to remove “background” in signal ROI

Example: Blinding calibration edge with hidden scale

- Calibration with mono-energetic neutrons produces spectrum with soft shoulder
- Could apply hidden scaling to energy variable to prevent tuning fit to expected value