Blinding and Salting in Dark Matter Searches

Maria Elena Monzani

SLAC
NATIONAL ACCELERATOR LABORATORY

Kavli Institute for Particle Astrophysics and Cosmology

PHYSTAT Dark Matter 2019
Stockholm, July 31 – August 2

Mitigating Experimenter's Bias

The first principle is that you must not fool yourself – and you are the easiest person to fool.

Mitigating Experimenter's Bias

The first principle is that you must not fool yourself – and you are the easiest person to fool.

Bias is an error in decision making:

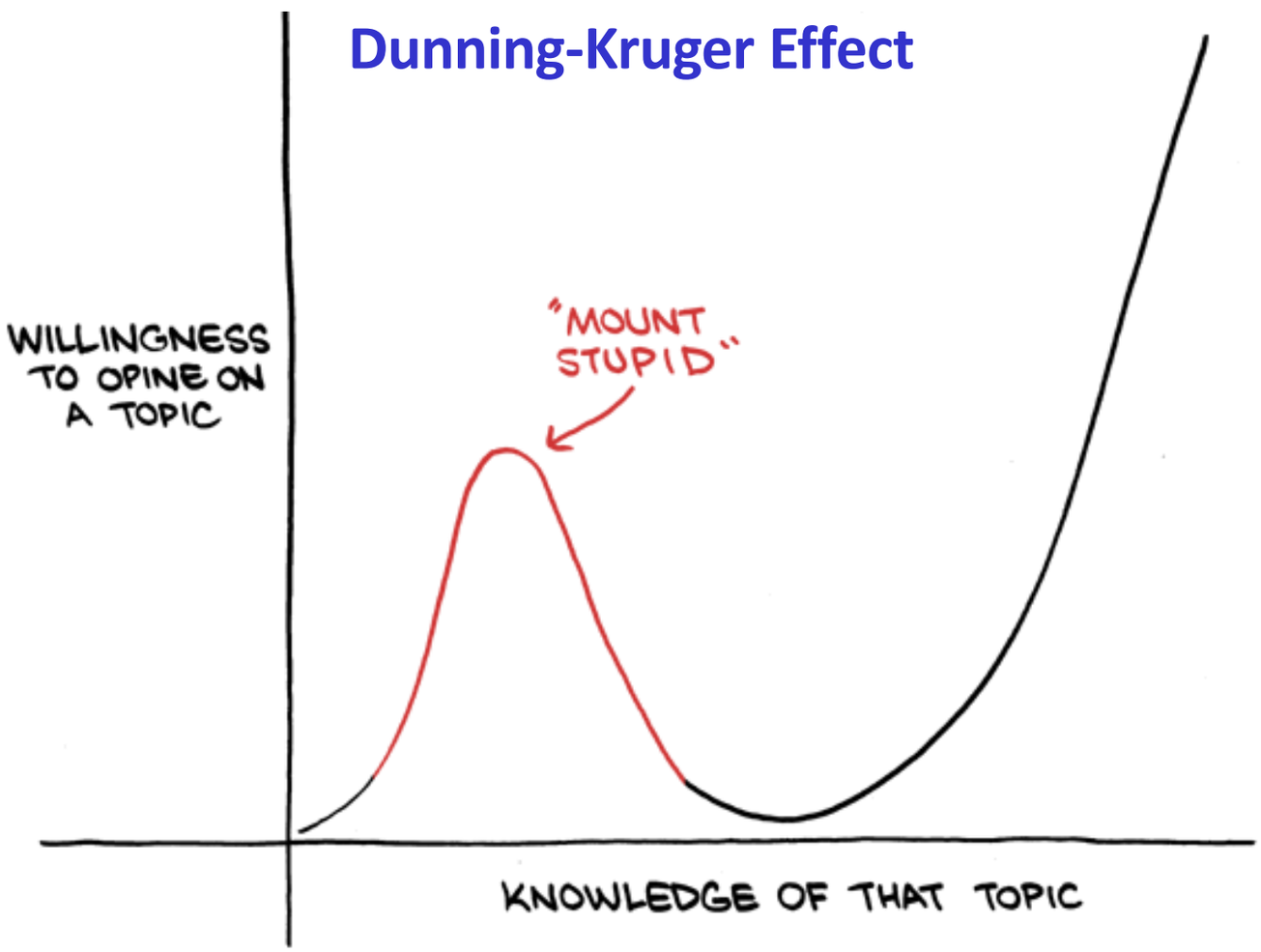
- **it usually functions as a cognitive shortcut ("I've seen this before")**
- **affects the criteria we use to evaluate experimental results**
- **is largest when stakes are higher**

Mitigating Experimenter's Bias

The first principle is that you must not fool yourself – and you are the easiest person to fool.

Bias is an error in decision making:

- **it usually functions as a cognitive shortcut ("I've seen this before")**
- **affects the criteria we use to evaluate experimental results**
- **is largest when stakes are higher**

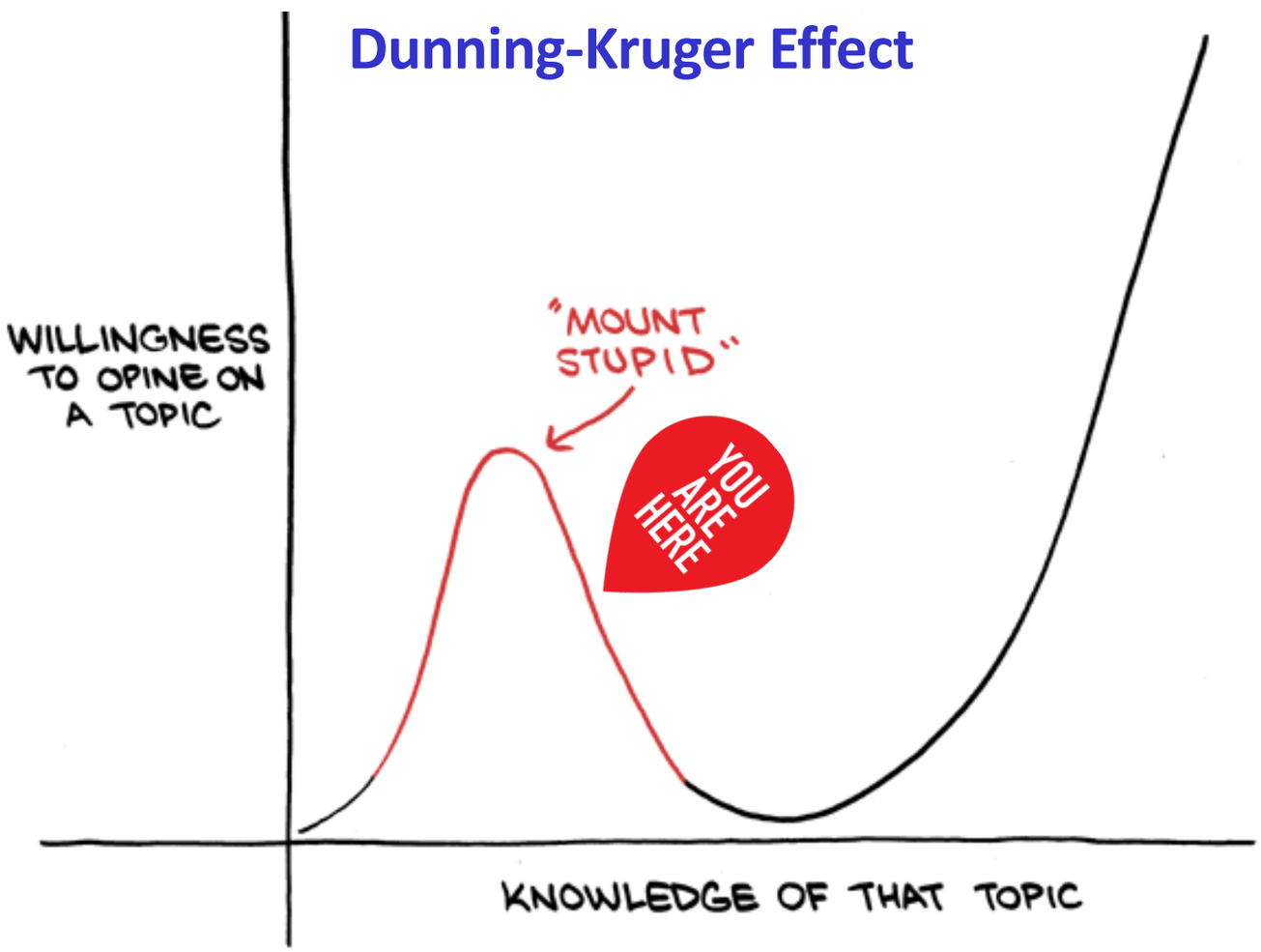


Mitigating Experimenter's Bias

The first principle is that you must not fool yourself – and you are the easiest person to fool.

Bias is an error in decision making:

- **it usually functions as a cognitive shortcut ("I've seen this before")**
- **affects the criteria we use to evaluate experimental results**
- **is largest when stakes are higher**



Millikan-Fletcher Experiment, 1909

Established the discrete nature of the electron's charge

Measured value 6% smaller than currently accepted (>5σ discrepancy!)

Root cause of the discrepancy: incorrect value for the viscosity of air

Early example of confirmation bias in subsequent experiments

Millikan-Fletcher Experiment, 1909

Established the discrete nature of the electron's charge

Measured value 6% smaller than currently accepted (>5σ discrepancy!)

Root cause of the discrepancy: incorrect value for the viscosity of air

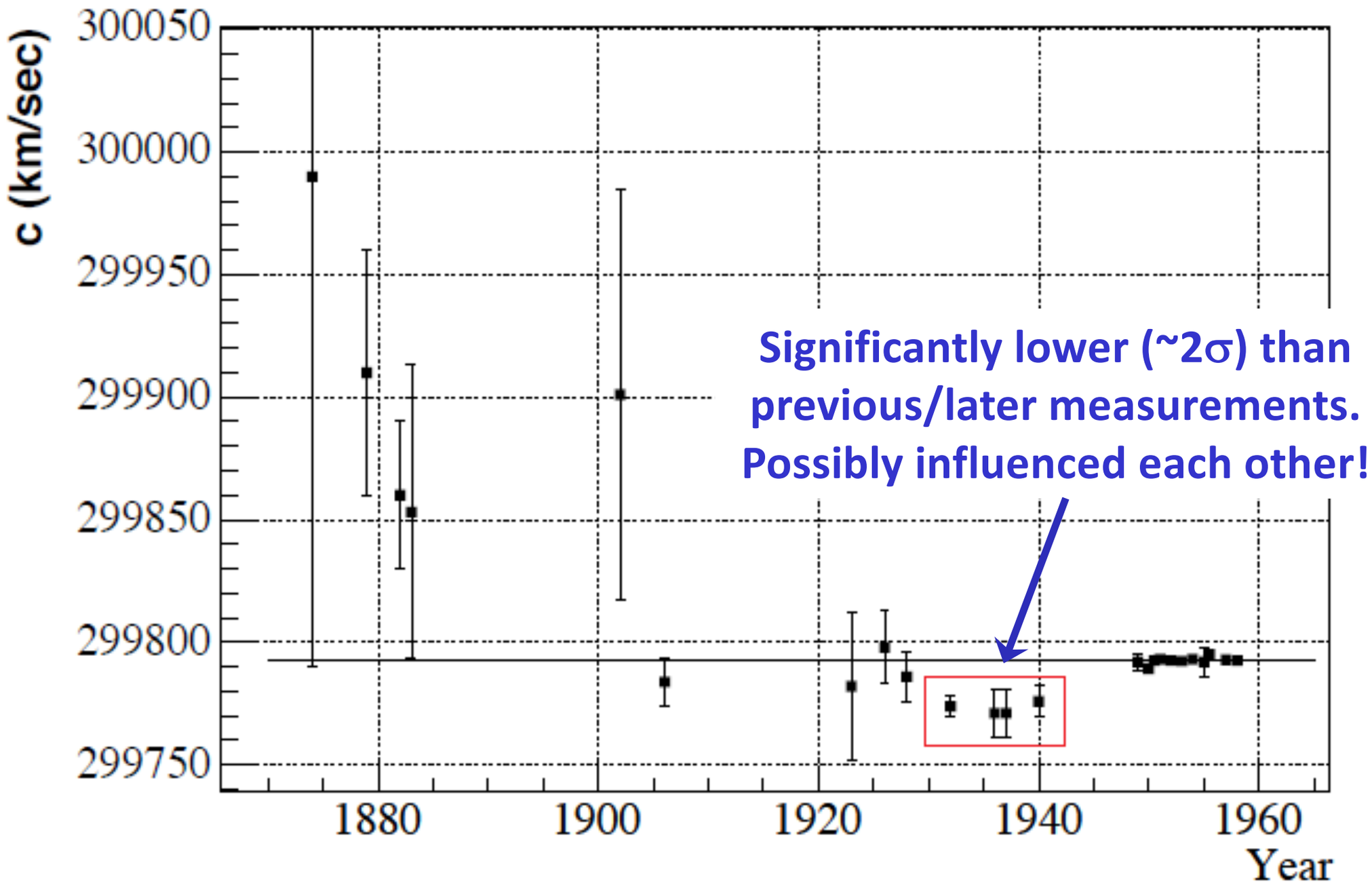
Early example of confirmation bias in subsequent experiments

In his speech *Cargo Cult Science* (58), Richard Feynman warns that

> It's a thing that scientists are ashamed of—this history—because it's apparent that people did things like this: When they got a number that was too high above Millikan's, they thought something must be wrong—and they would look for and find a reason why something might be wrong. When they got a number closer to Millikan's value they didn't look so hard. . .
>
> The first principle is that you must not fool yourself—and you are the easiest person to fool.

The Speed of Light Saga

Measured value of the Speed of Light, 1875 - 1960



Significantly lower (~2σ) than previous/later measurements. Possibly influenced each other!

Klein-Roodman, 2005

A Successful "Blind" Experiment

Frank Dunnigton's e/m measurement, 1932

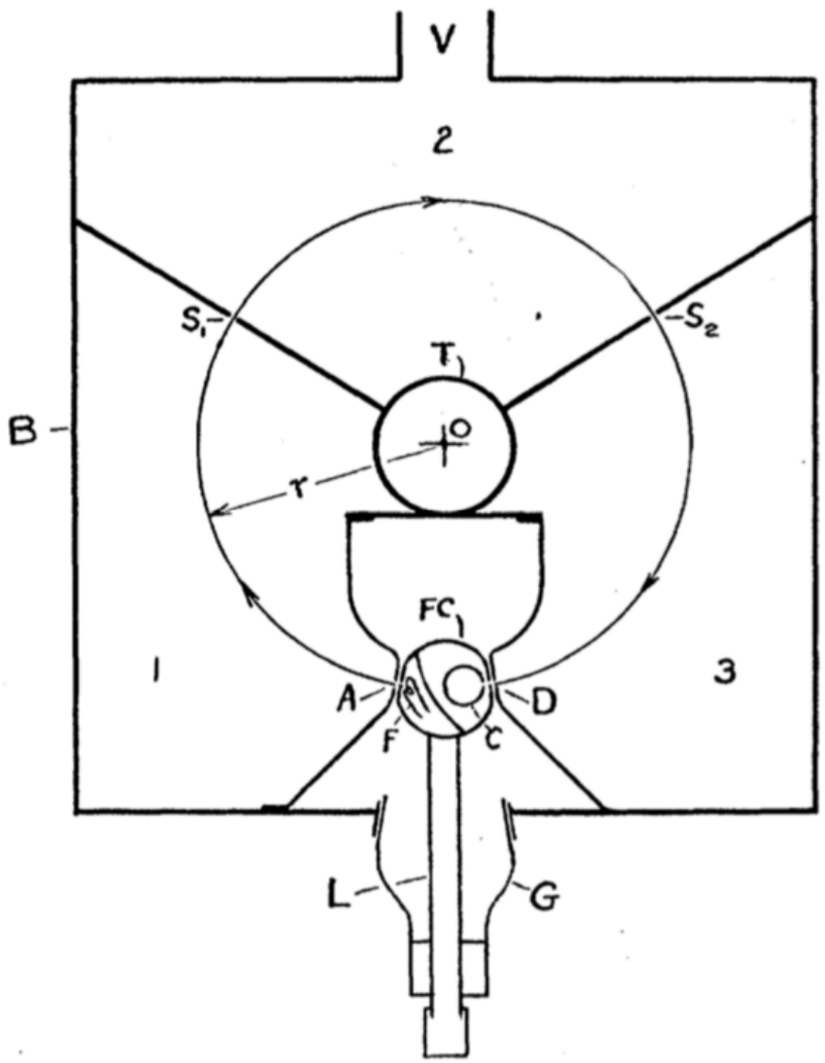


FIG. 1. Schematic diagram of measuring tube.

- **Two previous measurements disagreed**
- **e/m was proportional to the angle between electron source and detector**
- **Dunnington asked his machinist to arbitrarily choose an angle around 340°**
- **Only when the analysis was complete, did he accurately measure the hidden angle.**

First known blind measurement in particle physics!

A Successful "Blind" Experiment

Frank Dunnigton's e/m measurement, 1932

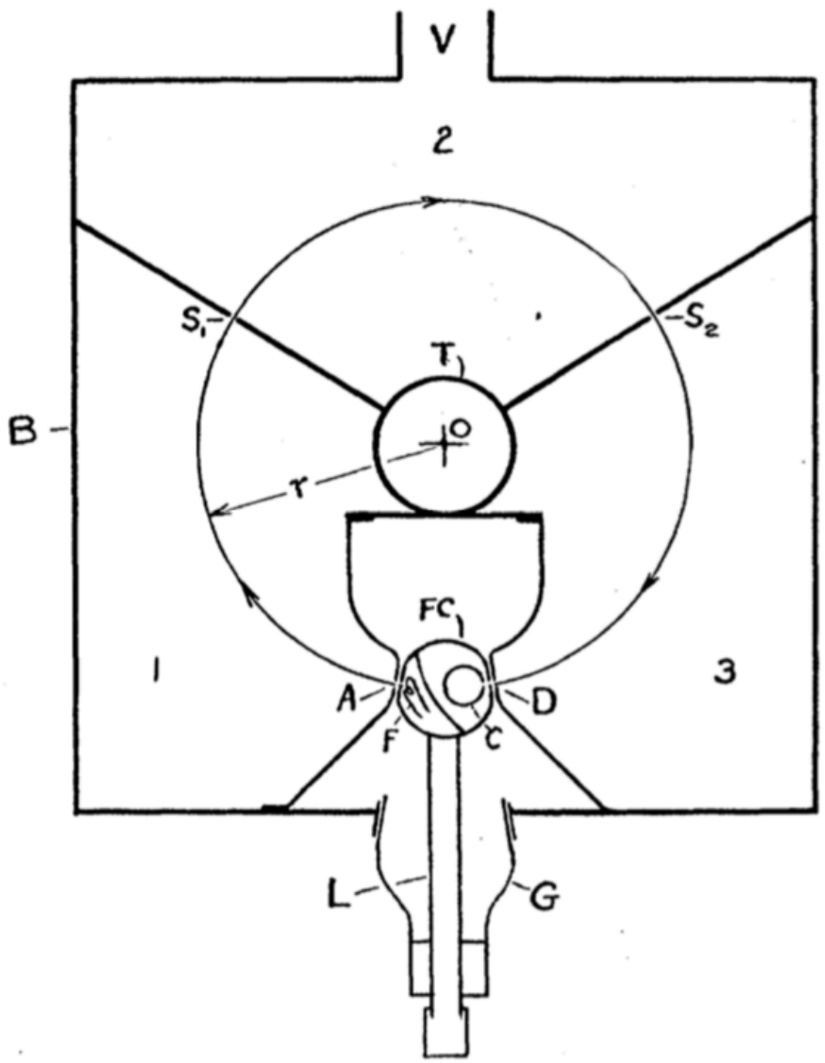


FIG. 1. Schematic diagram of measuring tube.

First known blind measurement in particle physics!

- **Two previous measurements disagreed**
- **e/m was proportional to the angle between electron source and detector**
- **Dunnington asked his machinist to arbitrarily choose an angle around 340°**
- **Only when the analysis was complete, did he accurately measure the hidden angle.**

"it is easier than is generally realized to unconsciously work toward a certain value. One cannot, of course, alter or change natural phenomena... but one can, for instance, seek for those corrections and refinements which shift the results in the desired direction. Every effort has been made to avoid such tendencies in the present work".

Dunnington, Phys.Rev. 43, pp. 404-416, (1932)

A Recent Success Story: LIGO

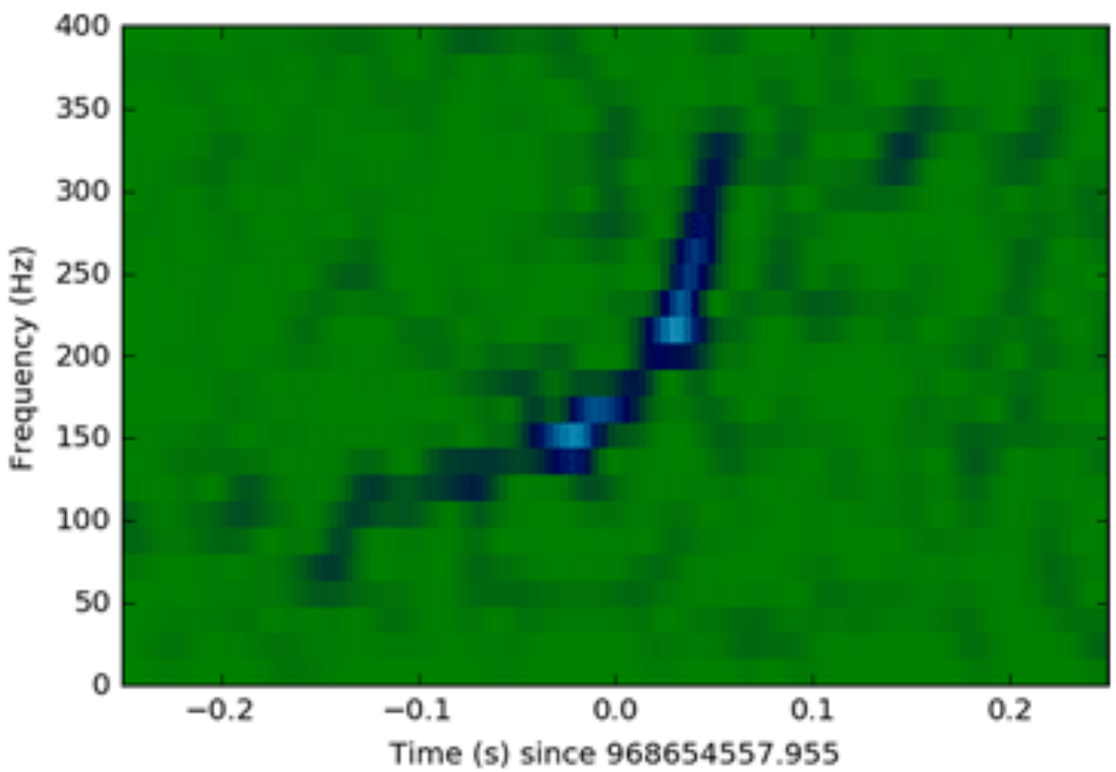
Famous "blind injection" exercise from September 16, 2010:

- **Treated like a real event, including alerting telescopes for follow-up**
- **Team performed a complete analysis, all the way to discovery paper**
- **"Envelope" opened on March 14, 2011 revealing the blind injection**

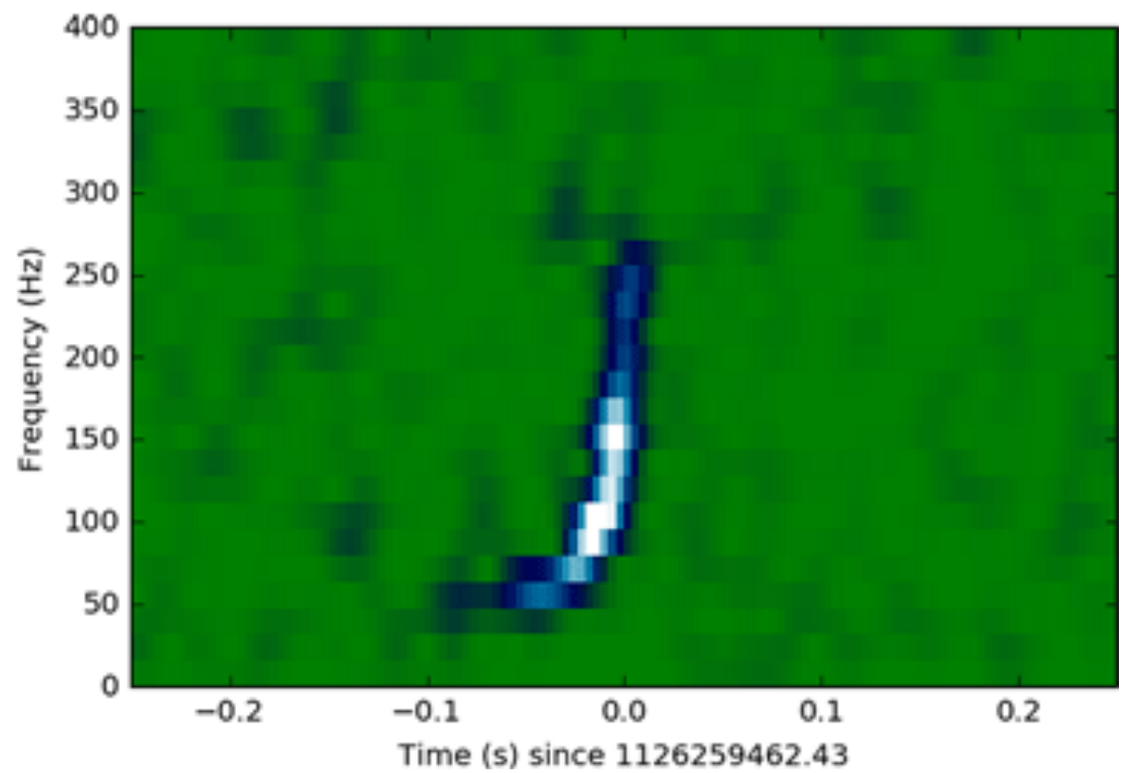
A Recent Success Story: LIGO

Famous "blind injection" exercise from September 16, 2010:

- **Treated like a real event, including alerting telescopes for follow-up**
- **Team performed a complete analysis, all the way to discovery paper**
- **"Envelope" opened on March 14, 2011 revealing the blind injection**



GW100916, blind injection



GW150914, discovery of gravitational waves!

Discovery based on one event!!! Minimize confirmation bias at all cost!!!

Blinding Techniques in Particle Physics

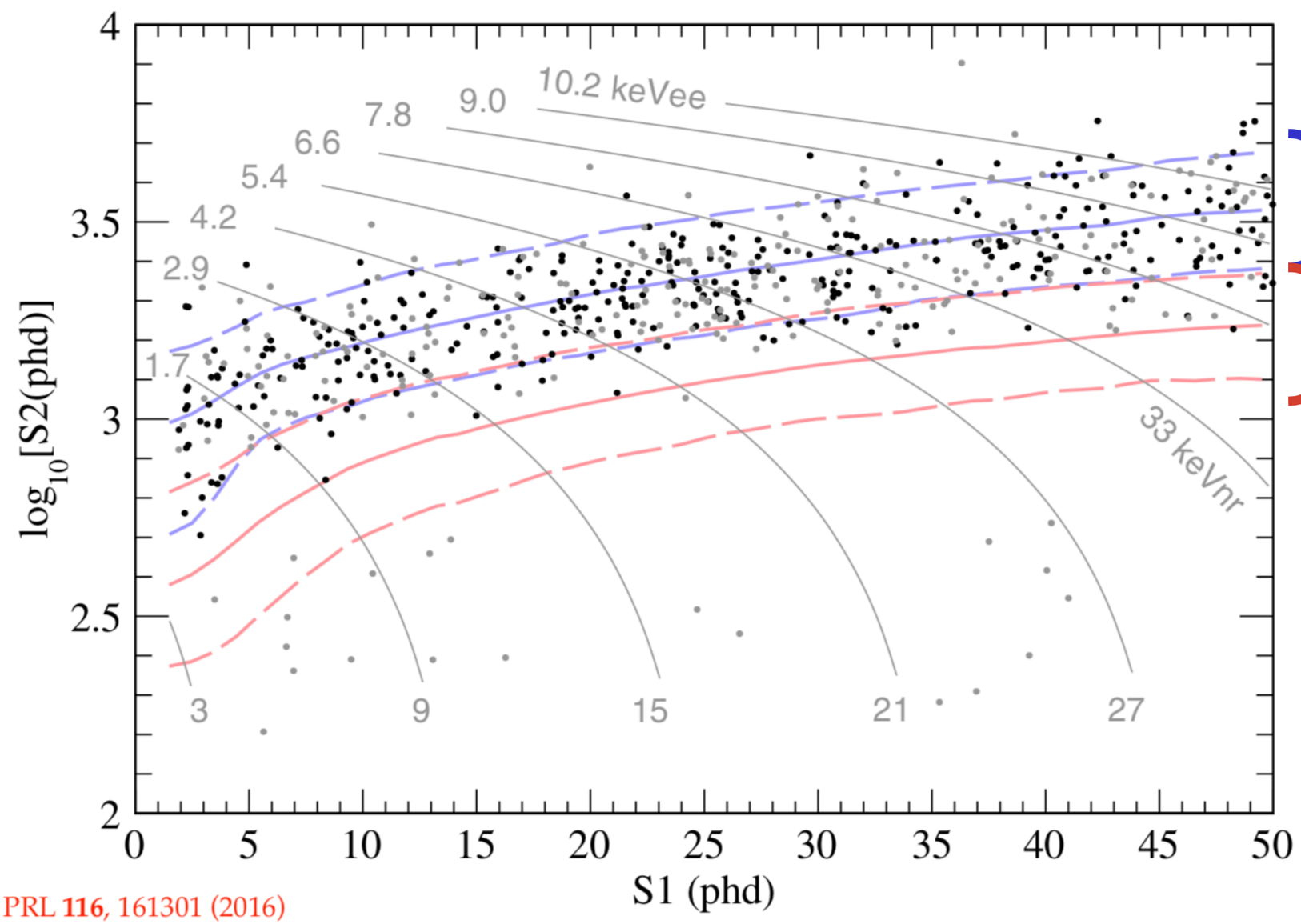
Hidden Signal Box: well suited to rare event searches

-> a subset of the data, containing the potential signal, is kept hidden until all aspects of the analysis are complete

Blinding Techniques in Particle Physics

Hidden Signal Box: well suited to rare event searches

-> a subset of the data, containing the potential signal, is kept hidden until all aspects of the analysis are complete



Electronic recoils (Bkg)

Nuclear Recoils (Signal)

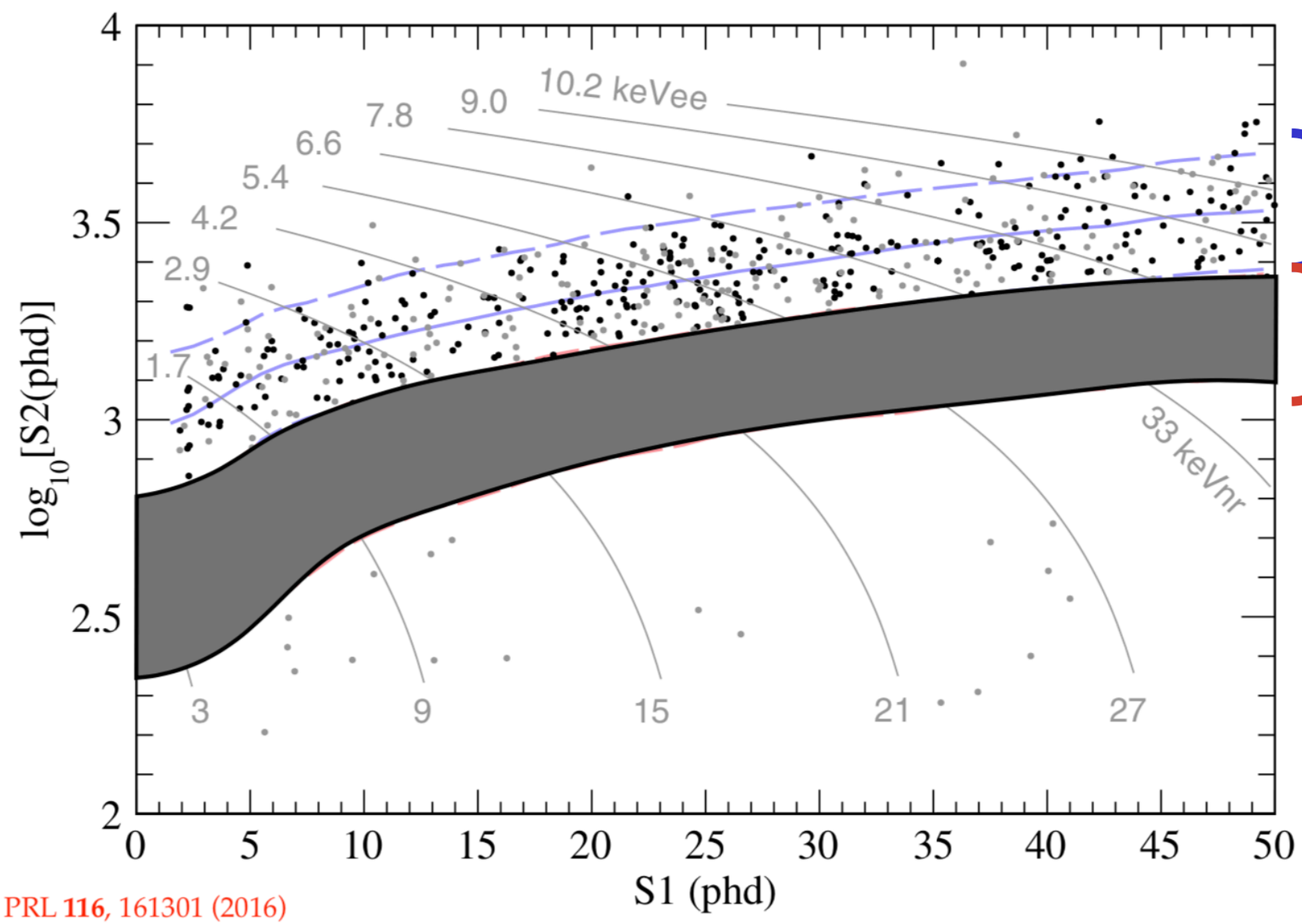
LUX's first science data (95 live days, non-blind)

PRL **116**, 161301 (2016)

Blinding Techniques in Particle Physics

Hidden Signal Box: well suited to rare event searches

-> a subset of the data, containing the potential signal, is kept hidden until all aspects of the analysis are complete



PRL **116**, 161301 (2016)

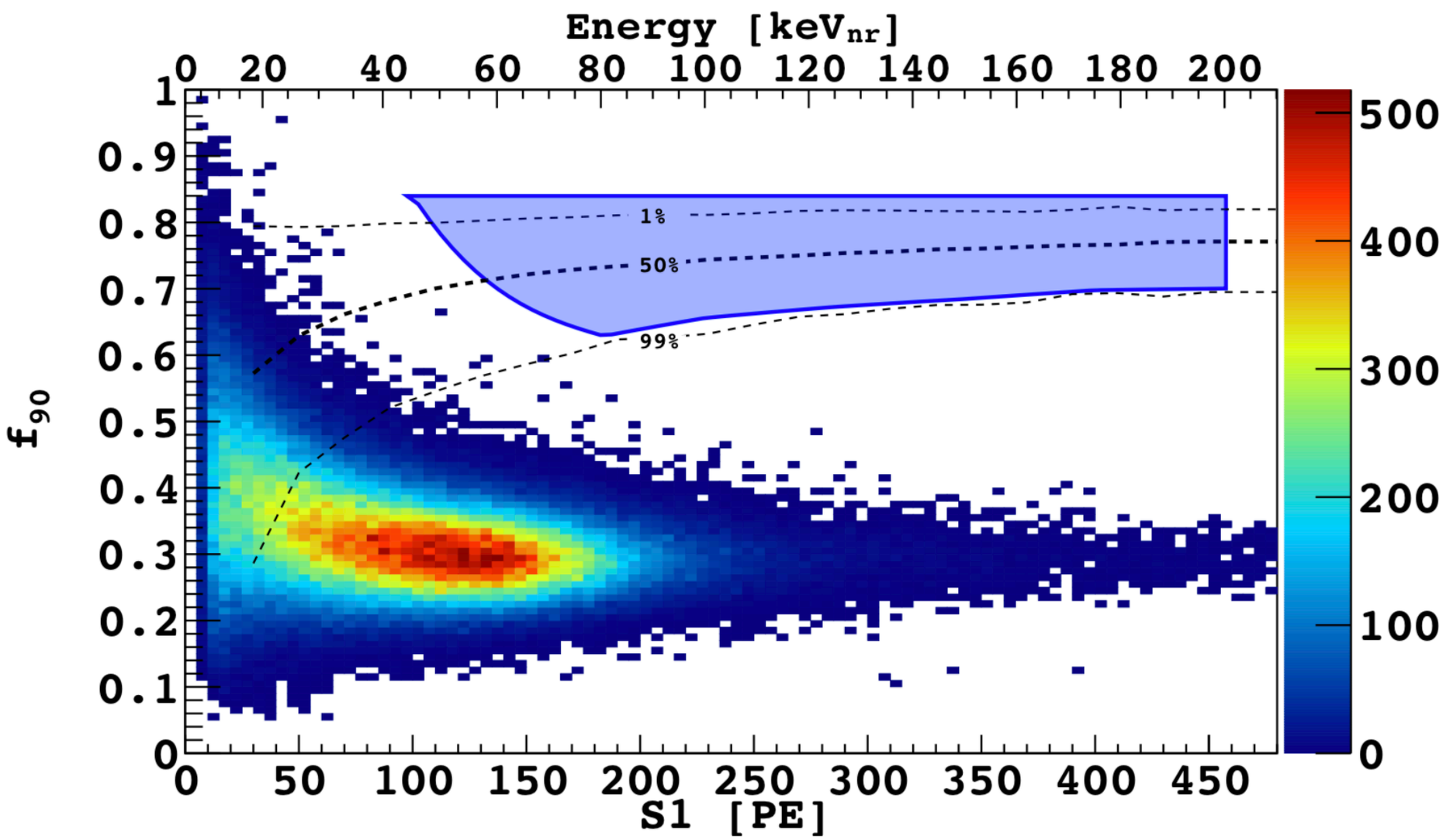
Electronic recoils (Bkg)

Nuclear Recoils (Signal)

Classic Blinding box: hide from analyzers all/most events in the signal region

A Successful "Hidden Signal Box" DM Search

DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon, Phys.Rev. D 98, 102006 (2018)



"With the analysis cuts applied and the data fully unblinded, no events are observed in the pre-defined DM search region"

Most/other DM experiments had to rely on post-unblinding cuts!

Caveats on "Hidden Signal Box" Searches

A "hidden box" means one is also blind to rare backgrounds.

From Klein-Roodman, 2005: ***"a hidden signal box approach assumes that the characteristics of the backgrounds are known well enough that nothing unexpected will be discovered when the signal box is opened"*.**

Caveats on "Hidden Signal Box" Searches

A "hidden box" means one is also blind to rare backgrounds.

From Klein-Roodman, 2005: *"a hidden signal box approach assumes that the characteristics of the backgrounds are known well enough that nothing unexpected will be discovered when the signal box is opened"*.

Many examples of blind searches in WIMP experiments, which have "discovered" un-modeled, pathological backgrounds after unblinding:

- **XENON10: PRL 100 (2008) 021303 (arXiv:0706.0039)**
- **XENON100: PRL 109 (2012) 181301 (arXiv:1207.5988)**
- **ZEPLINIII: PRD 80 (2009) 052010 (arXiv:0812.1150)**
- **SuperCDMS: PRL 112 (2014) 241302 (arXiv:1402.7137)**

Can we think of other techniques of bias mitigation?

Other Blinding Techniques

a. **Hidden Answer Methods:** best for experiments measuring a single precise parameter (Dunnigton's e/m measurement)
   - Specific techniques: hidden detector parameters; hidden offset; hidden offset/asymmetry; divided analyses

Klein-Roodman, 2005

Other Blinding Techniques

a. **Hidden Answer Methods:** best for experiments measuring a single precise parameter (Dunnigton's e/m measurement)
   - Specific techniques: hidden detector parameters; hidden offset; hidden offset/asymmetry; divided analyses

b. **Adding or Removing Events:** appropriate for "counting" experiments -> spoil the event count in an unknown way
   - This includes "salting", "spiking" and "blind injections" LIGO-style. SNO used a combination of add/remove

Klein-Roodman, 2005

Other Blinding Techniques

a. **Hidden Answer Methods:** best for experiments measuring a single precise parameter (Dunnigton's e/m measurement)
   - Specific techniques: hidden detector parameters; hidden offset; hidden offset/asymmetry; divided analyses

b. **Adding or Removing Events:** appropriate for "counting" experiments -> spoil the event count in an unknown way
   - This includes "salting", "spiking" and "blind injections" LIGO-style. SNO used a combination of add/remove

c. **Data Prescaling:** the analysis is developed on a prescaled fraction of the data set and then applied to the remainder.
   - Most blind searches in DM experiments are actually *some* version of Data Prescaling (arXiv:1907.11485)

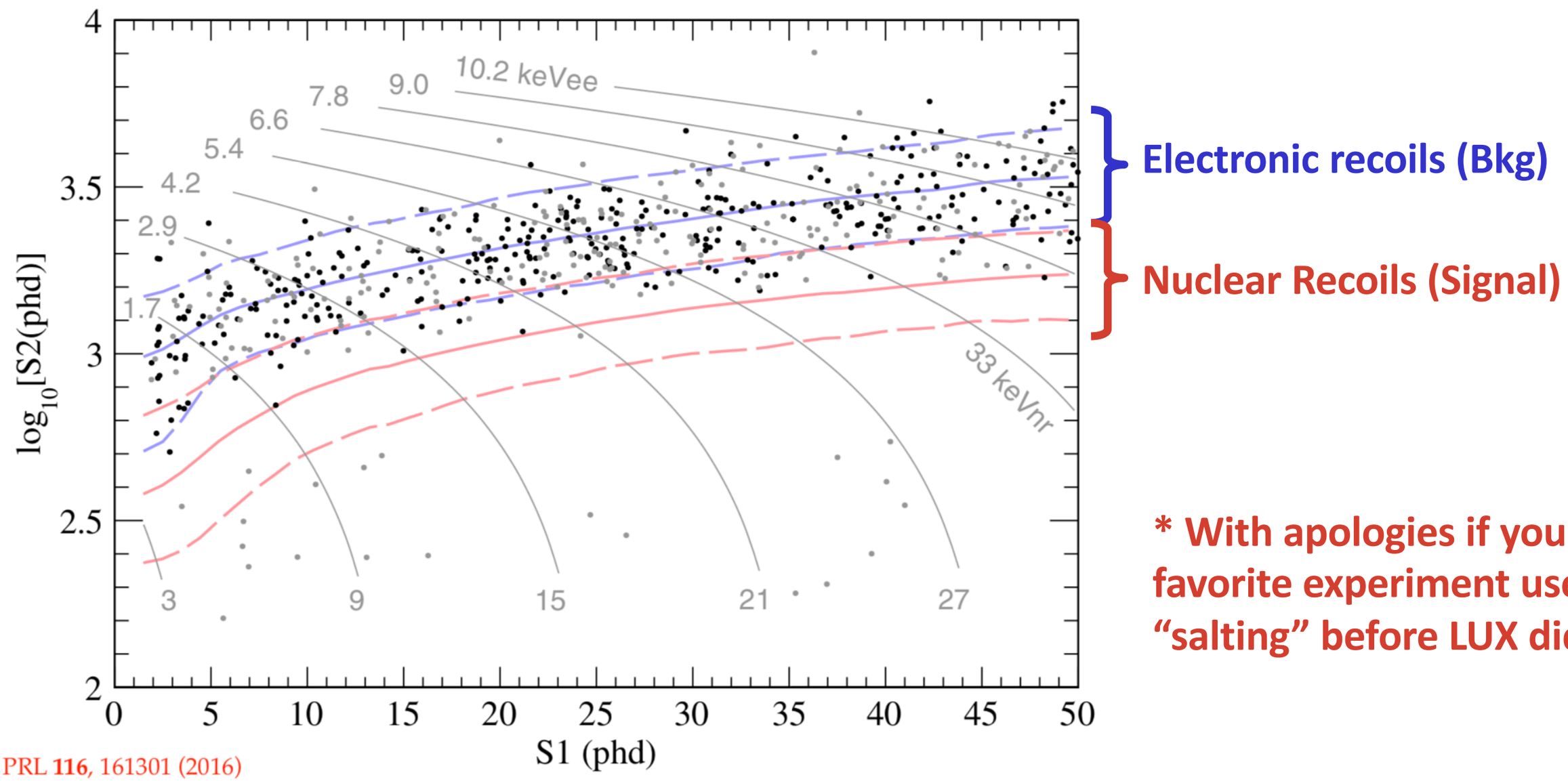
Klein-Roodman, 2005

Salting in LUX*

Added a number of signal-like events to the search box

Number of "salt" events unknow to analyzers, event list kept in a DB maintained by non-physicists only

Allow access to background events in WIMP search box



PRL **116**, 161301 (2016)

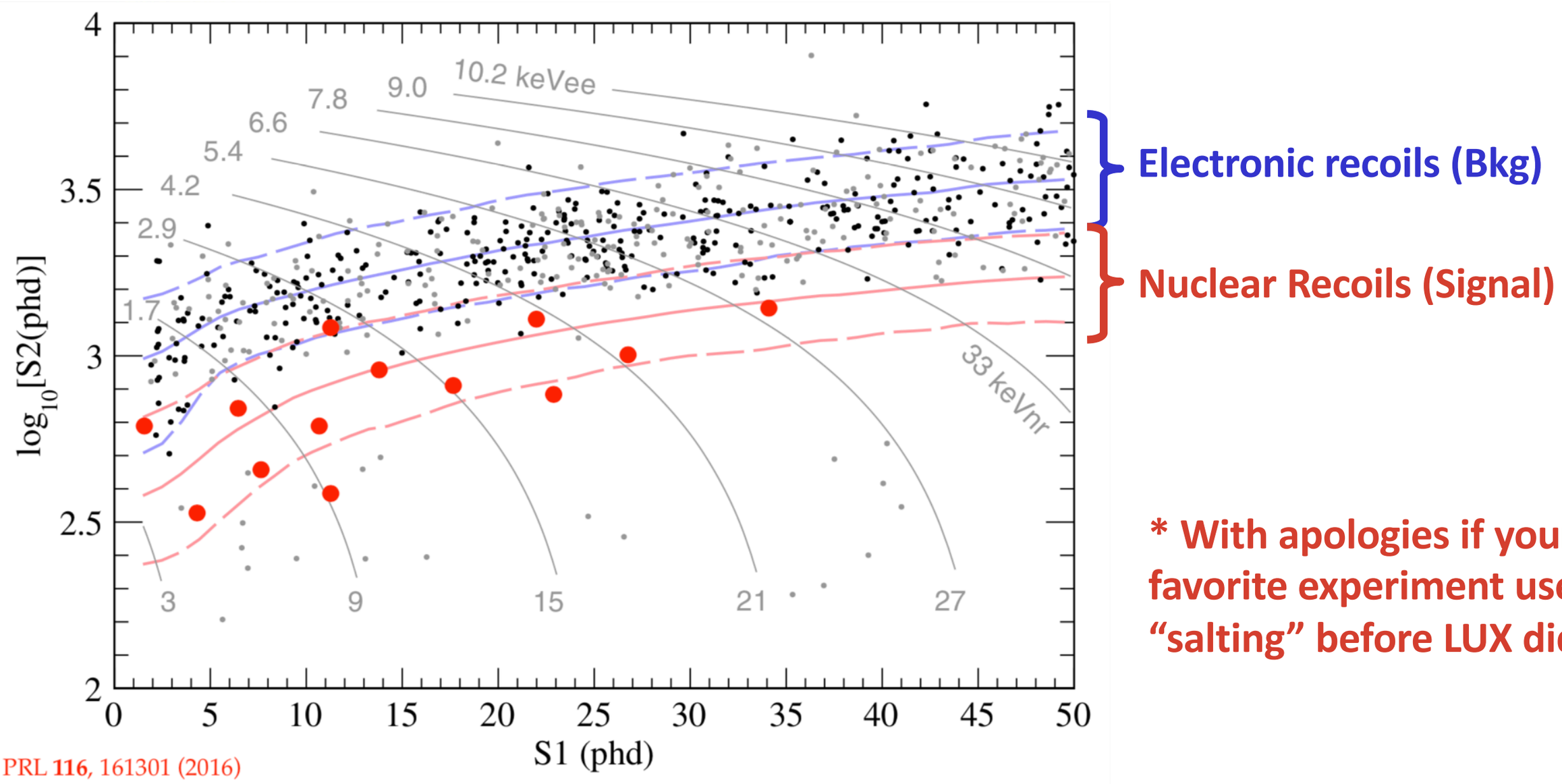
*** With apologies if your favorite experiment used "salting" before LUX did.**

Salting in LUX*

Added a number of signal-like events to the search box

Number of "salt" events unknow to analyzers, event list kept in a DB maintained by non-physicists only

Allow access to background events in WIMP search box

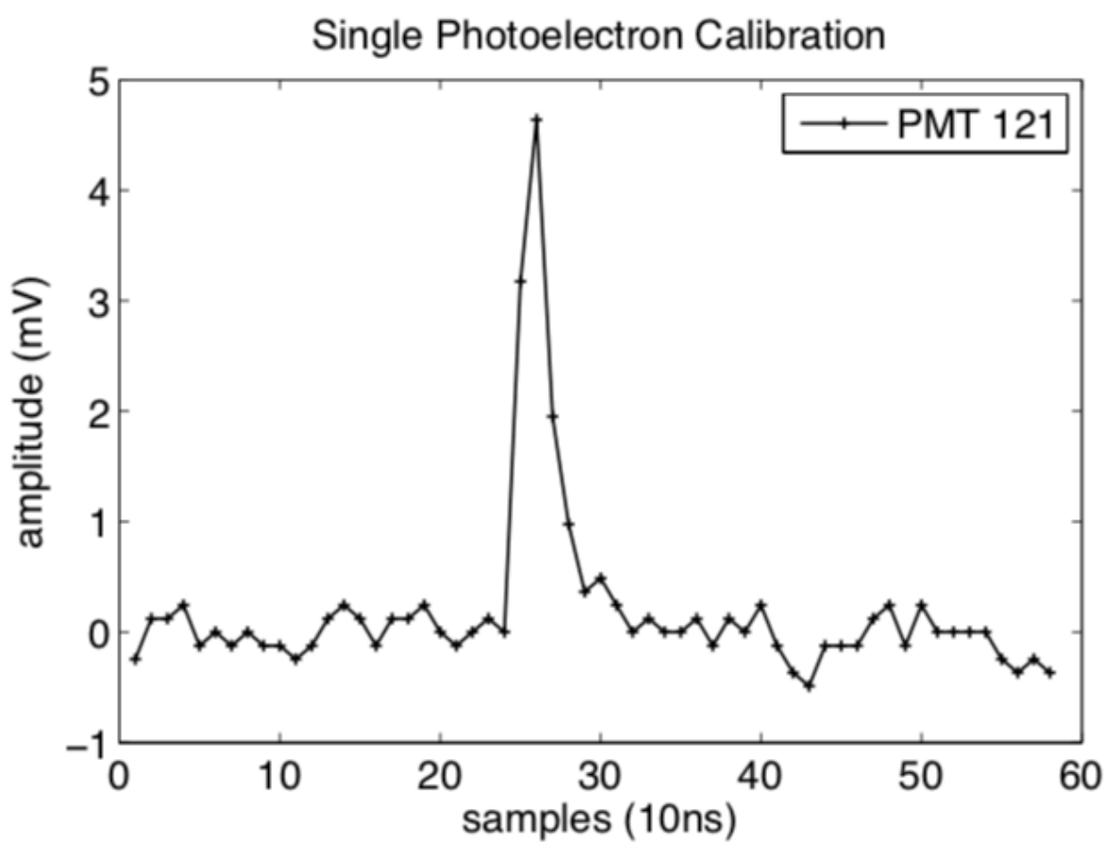


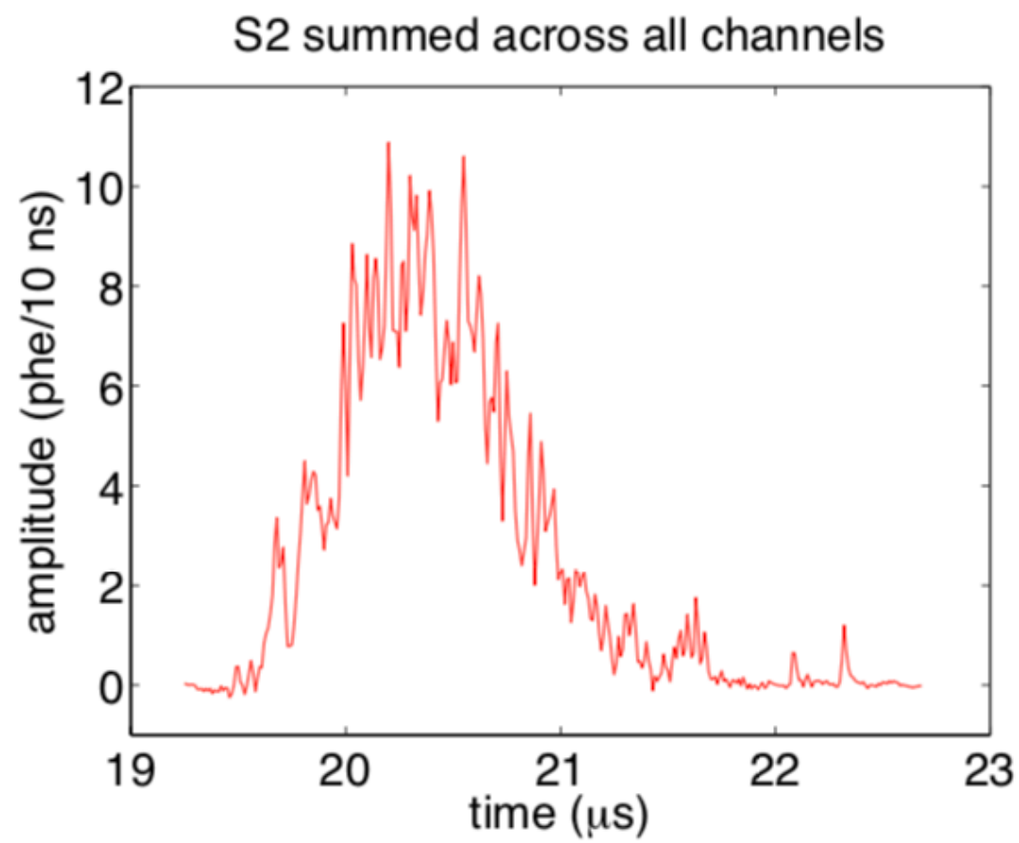
PRL **116**, 161301 (2016)

*** With apologies if your favorite experiment used "salting" before LUX did.**

LUX Implementation of Salting

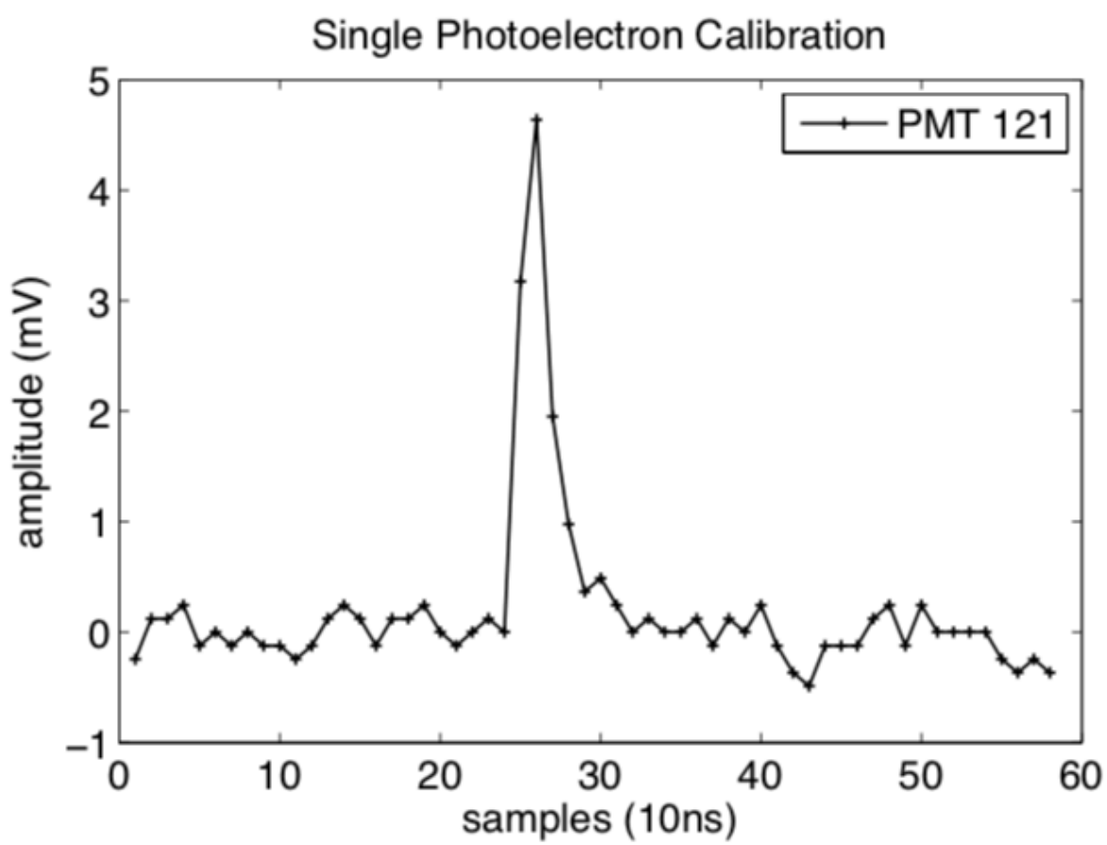
Salt injected in the LUX datastream at the waveform level

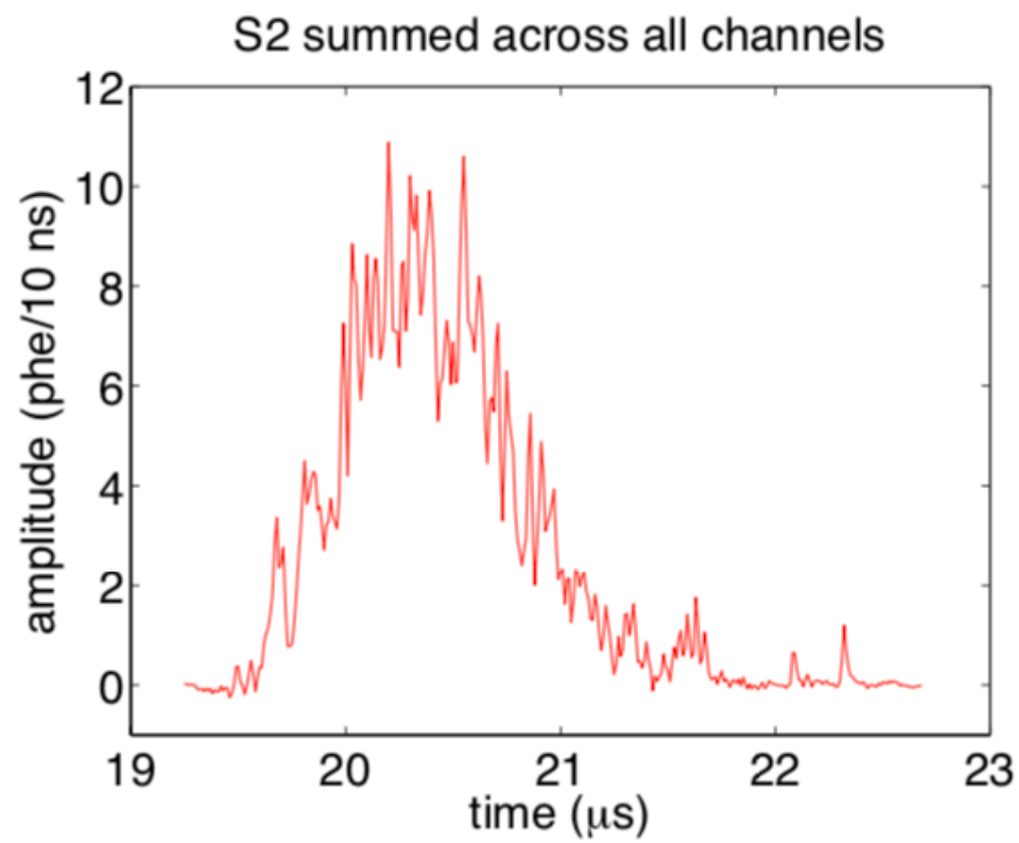




LUX Implementation of Salting

Salt injected in the LUX datastream at the waveform level

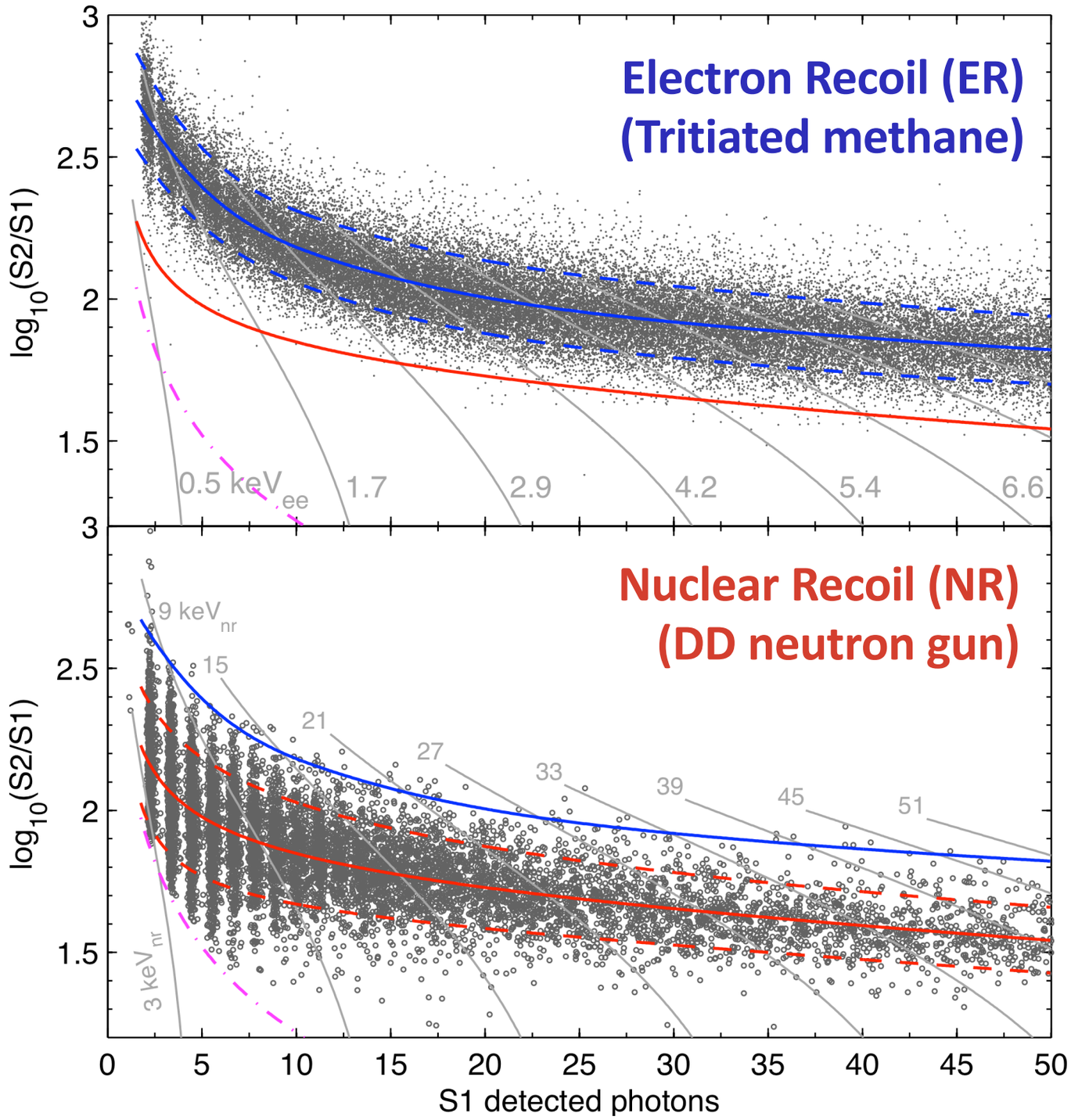




- **"Chimera" events created from uncorrelated S1 and S2 in calibration data**
- **Collaboration-wide Turing test to ensure nobody could identify the salt**
- **Key to unsalting only known to two people, external to the collaboration**

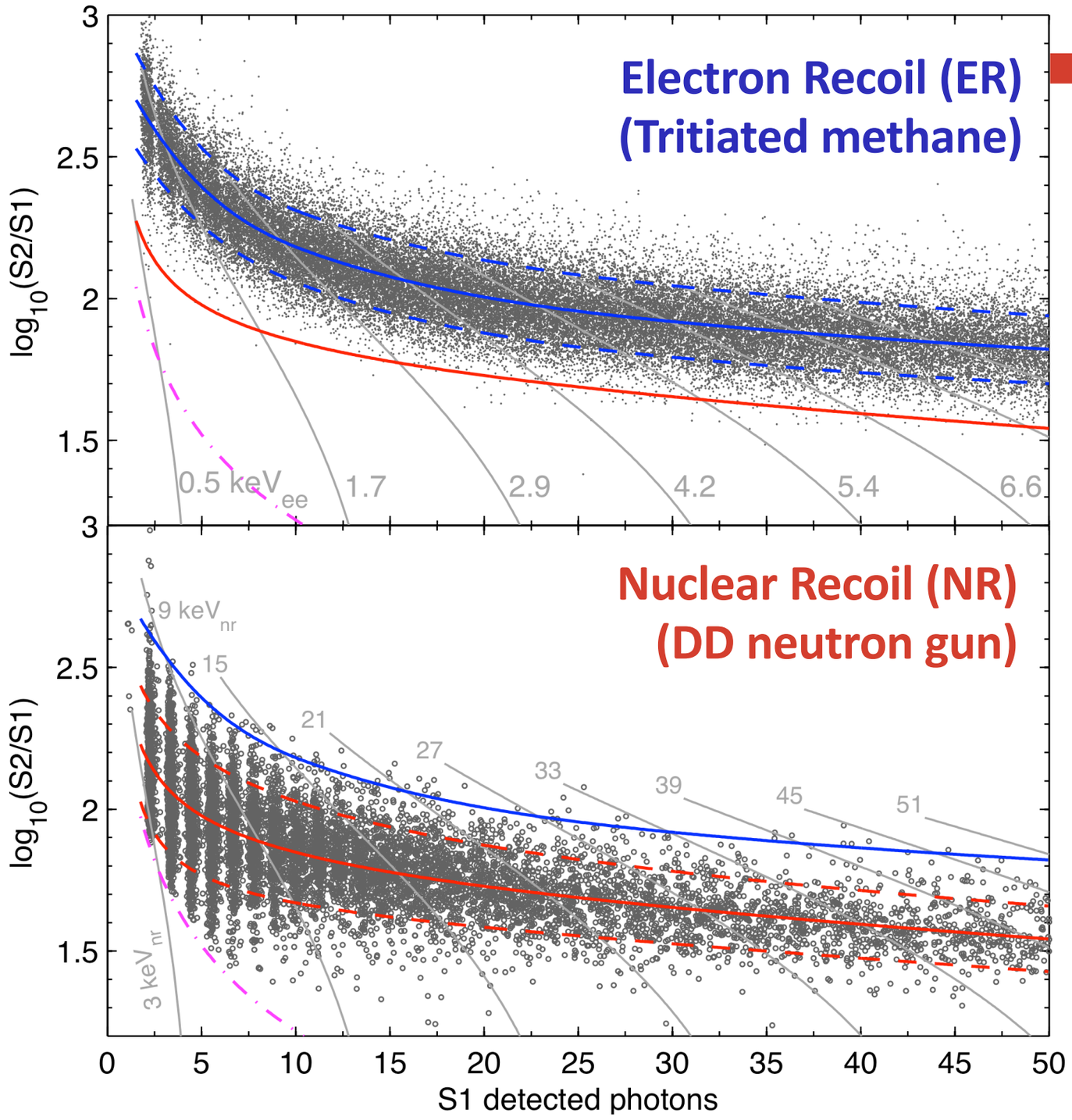
More details in: PRL 118, 021303 (2017)

High Statistics Calibrations in LUX



Phys. Rev. Lett. 116, 161301 (2016)

High Statistics Calibrations in LUX

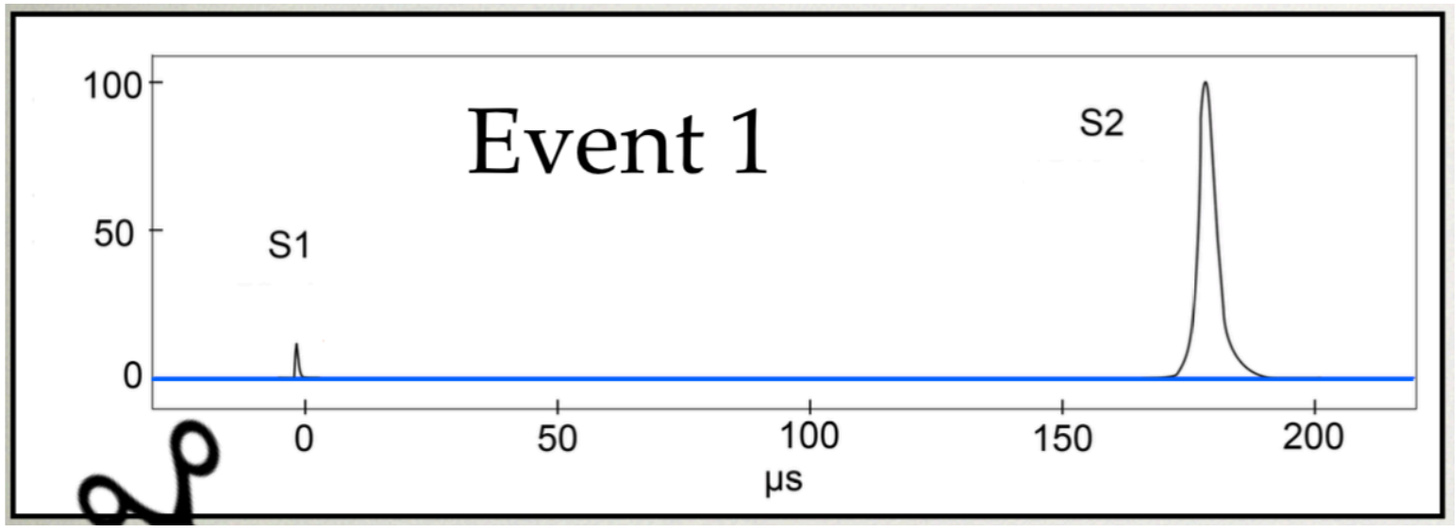


Phys. Rev. Lett. 116, 161301 (2016)

Excellent option for generating salt. Much more realistic than MC data.

- **DD: correct S2/S1 ratio to mimic WIMP signal, but only produced at location of DD generator.**
- **Tritium injection: events are single scatters, and uniformly distributed in detector volume**
- **Downside of Tritium events: S2/S1 ratio inconsistent with WIMP signal -> need to create synthetic ("chimera") events with the appropriate S2/S1**

Turning Tritium into Salt

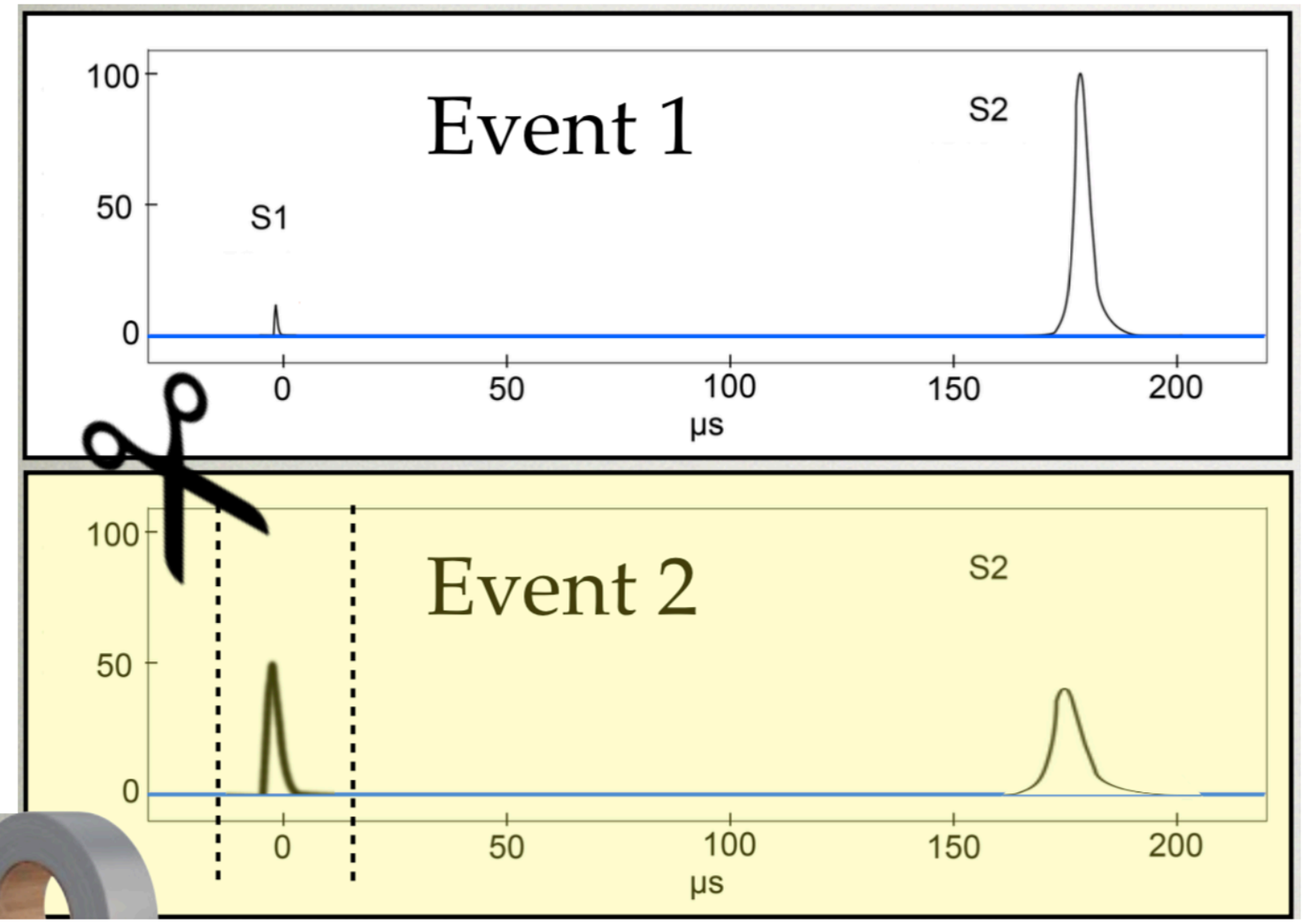


Recipe for making chimera events:

- **Find the expected S2/S1 ratio for WIMPs via simulation**
- **Take Tritium dataset, sequester from collaboration/analyzers**
- **Find a Tritium event with the targeted S2 size (Event 1)**

Cartoon courtesy of A. Manalaysay

Turning Tritium into Salt

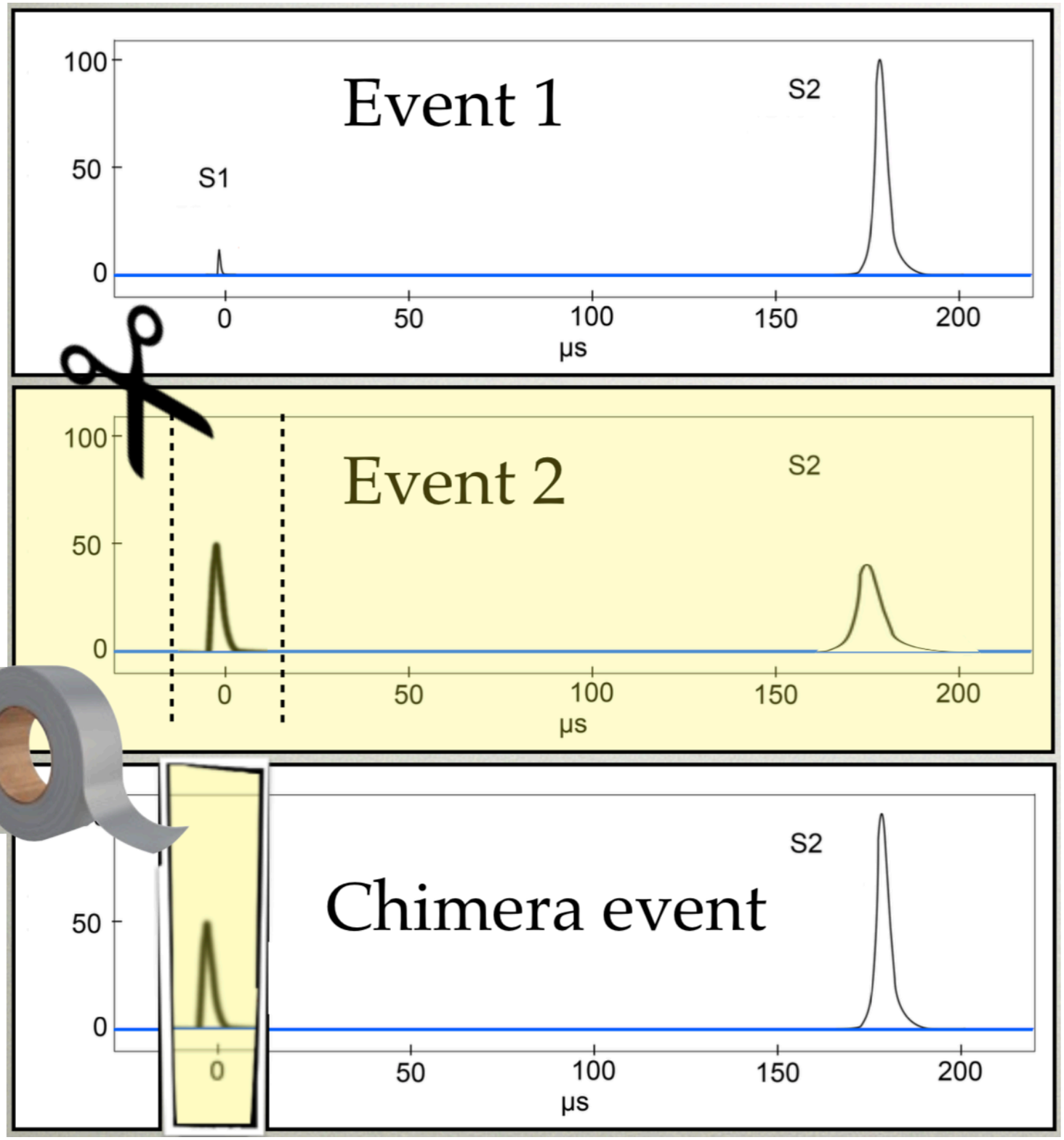


Recipe for making chimera events:

- **Find the expected S2/S1 ratio for WIMPs via simulation**
- **Take Tritium dataset, sequester from collaboration/analyzers**
- **Find a Tritium event with the targeted S2 size (Event 1)**
- **Find a Tritium event with the targeted S1 size (Event 2), from the same position as Event 1**

Cartoon courtesy of A. Manalaysay

Turning Tritium into Salt



Cartoon courtesy of A. Manalaysay

Recipe for making chimera events:

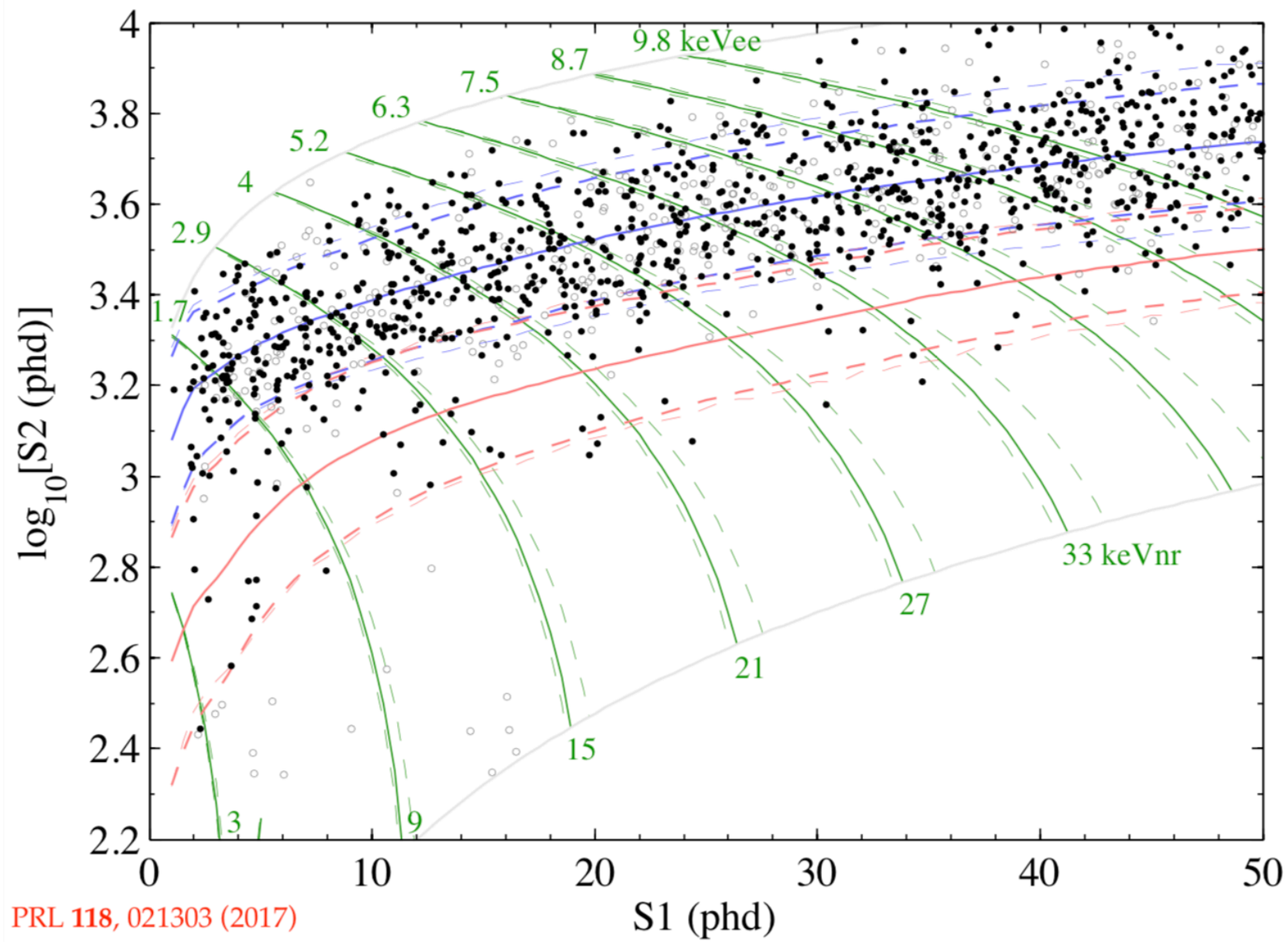
- **Find the expected S2/S1 ratio for WIMPs via simulation**
- **Take Tritium dataset, sequester from collaboration/analyzers**
- **Find a Tritium event with the targeted S2 size (Event 1)**
- **Find a Tritium event with the targeted S1 size (Event 2), from the same position as Event 1**
- **Rip out the S1 from Event 1, implant the S1 from Event 2**

The new event will contain all the detector effects/quirks as Event 1

LUX 332 live days Dark Matter Search

1. **Start with salted data, define analysis cuts, remove pathological events**

LUX's second science data (332 live days, with salt)

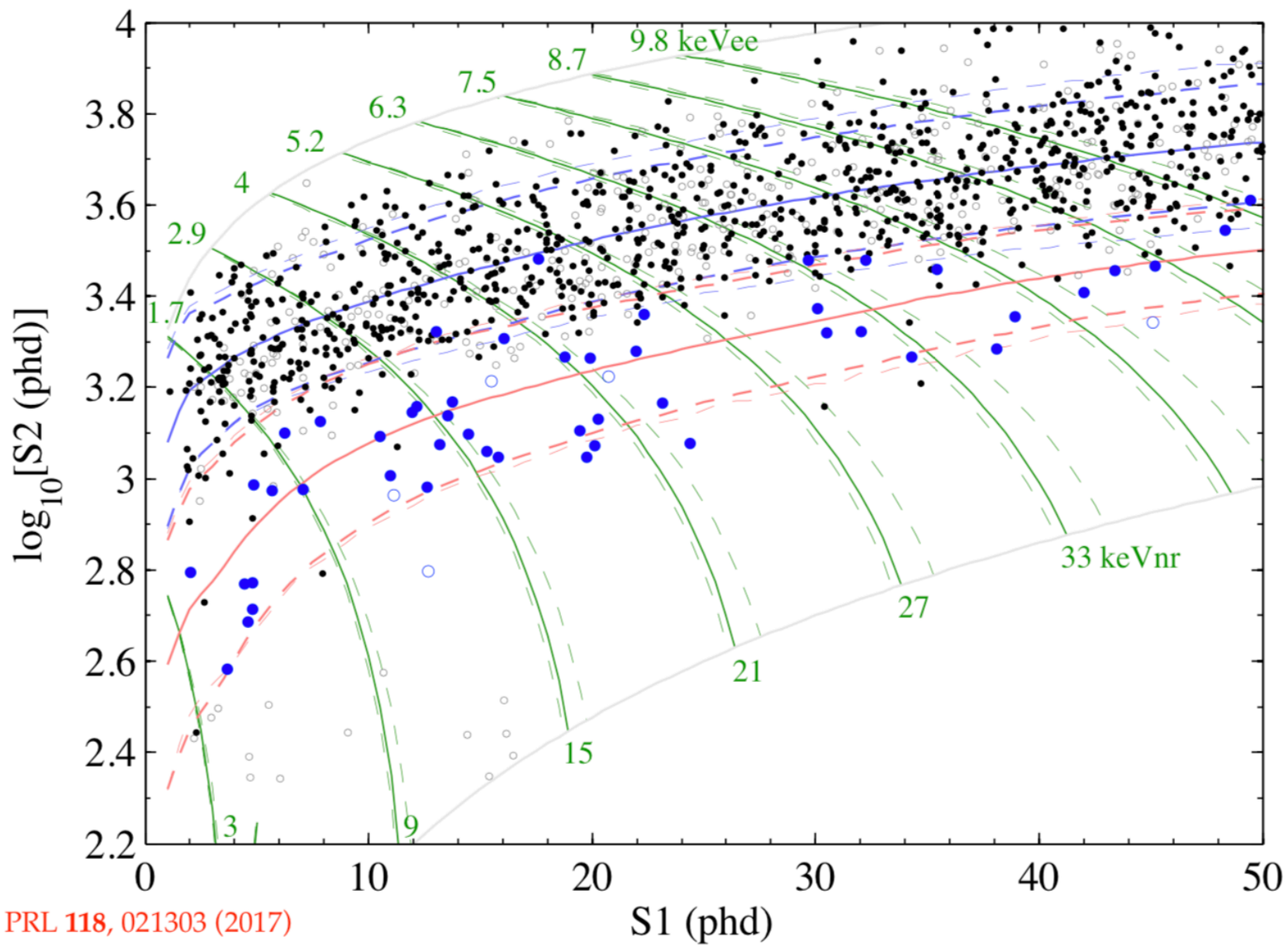


Phys. Rev. Lett. 118, 021303 (2017)

LUX 332 live days Dark Matter Search

1. Start with salted data, define analysis cuts, remove pathological events
2. Identify salt events (blue dots): cool, they lie in signal band, as expected

LUX's second science data (332 live days, with salt)

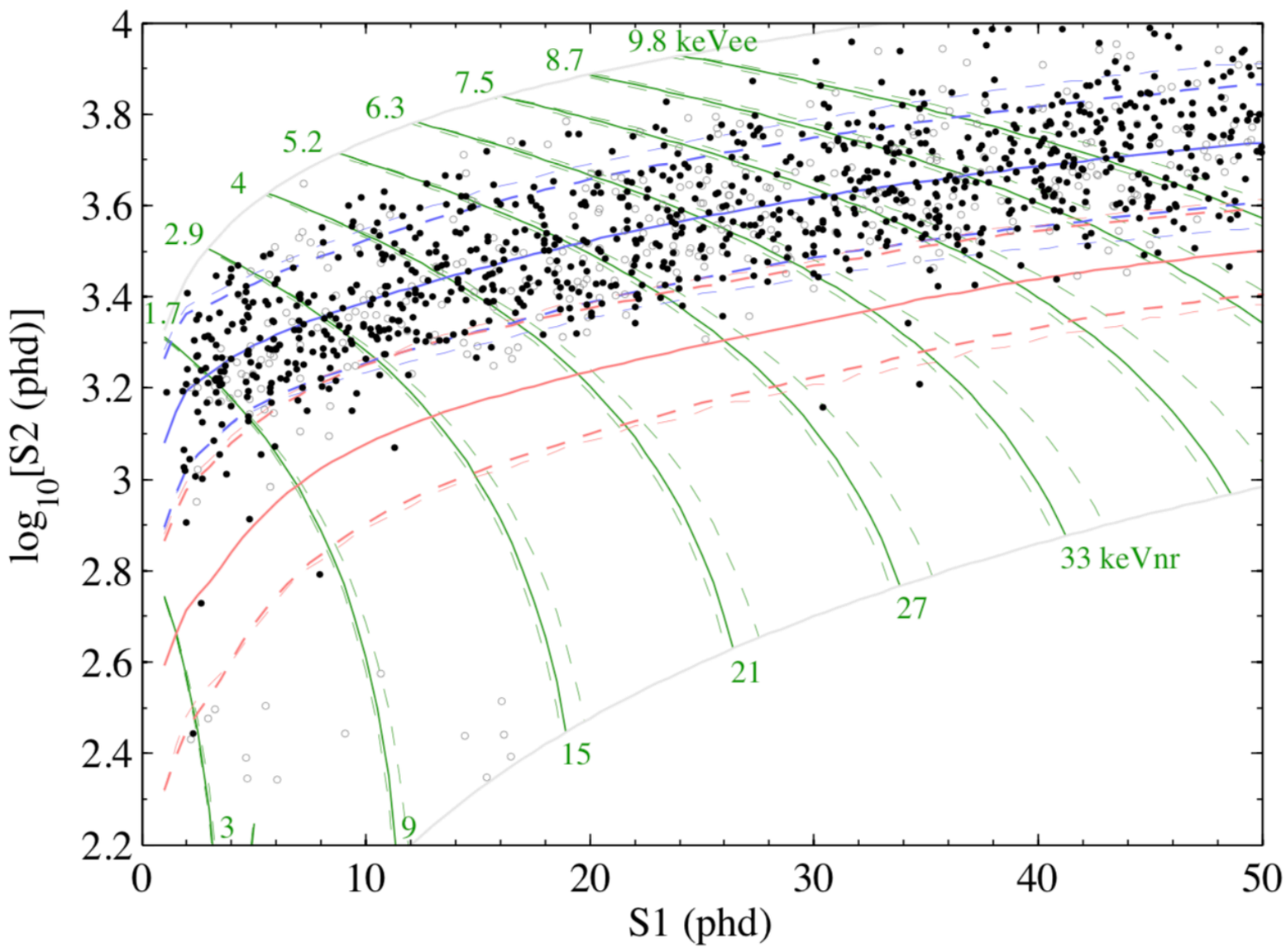


Phys. Rev. Lett. 118, 021303 (2017)

LUX 332 live days Dark Matter Search

1. Start with salted data, define analysis cuts, remove pathological events
2. Identify salt events (blue dots): cool, they lie in signal band, as expected

LUX's second science data (332 live days, unsalted)



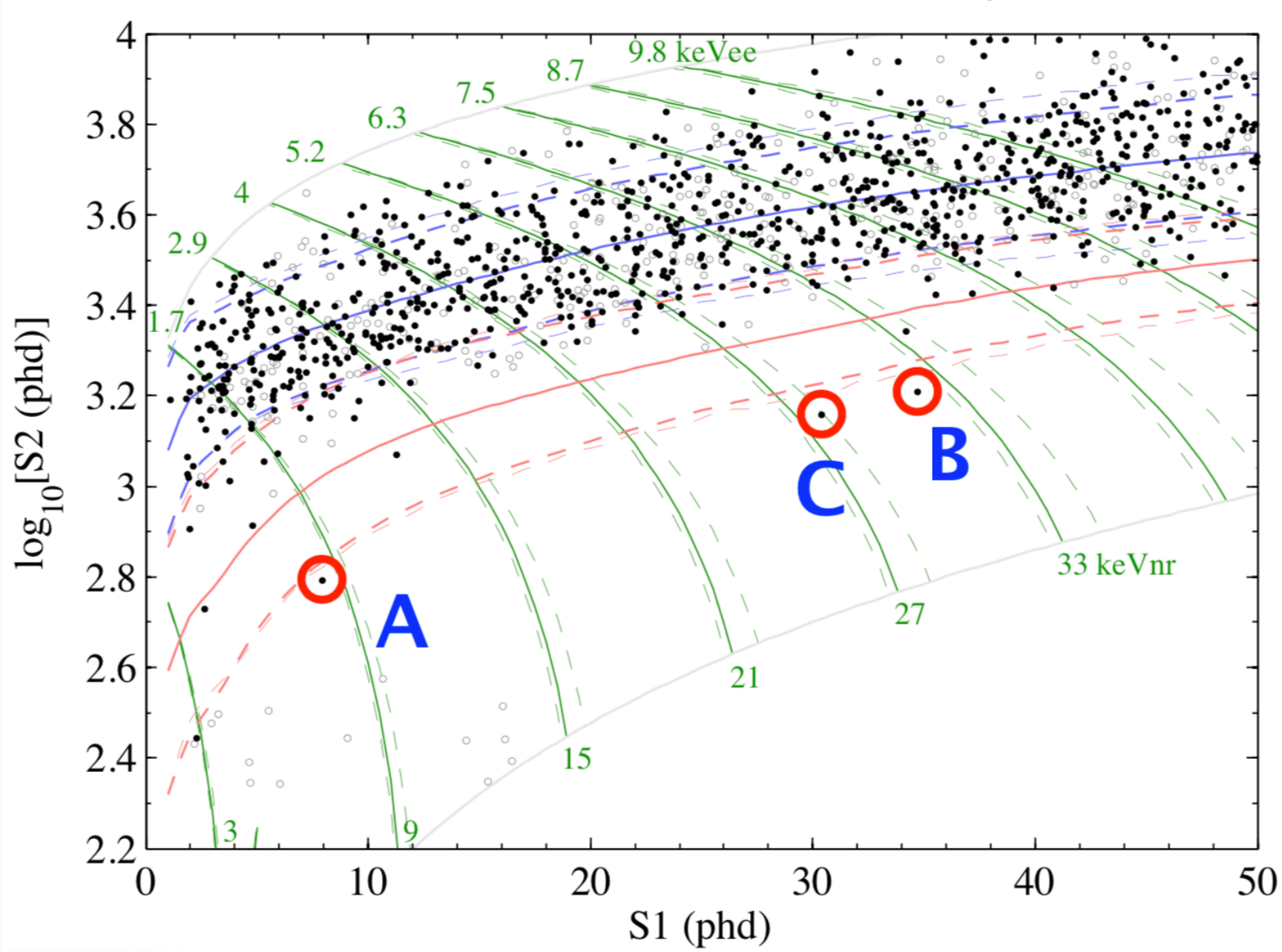
3. Remove salt!
4. Make exclusion plot

Phys. Rev. Lett. 118, 021303 (2017)

LUX 332 live days Dark Matter Search

1. Start with salted data, define analysis cuts, remove pathological events
2. Identify salt events (blue dots): cool, they lie in signal band, as expected

LUX's second science data (332 live days, unsalted)



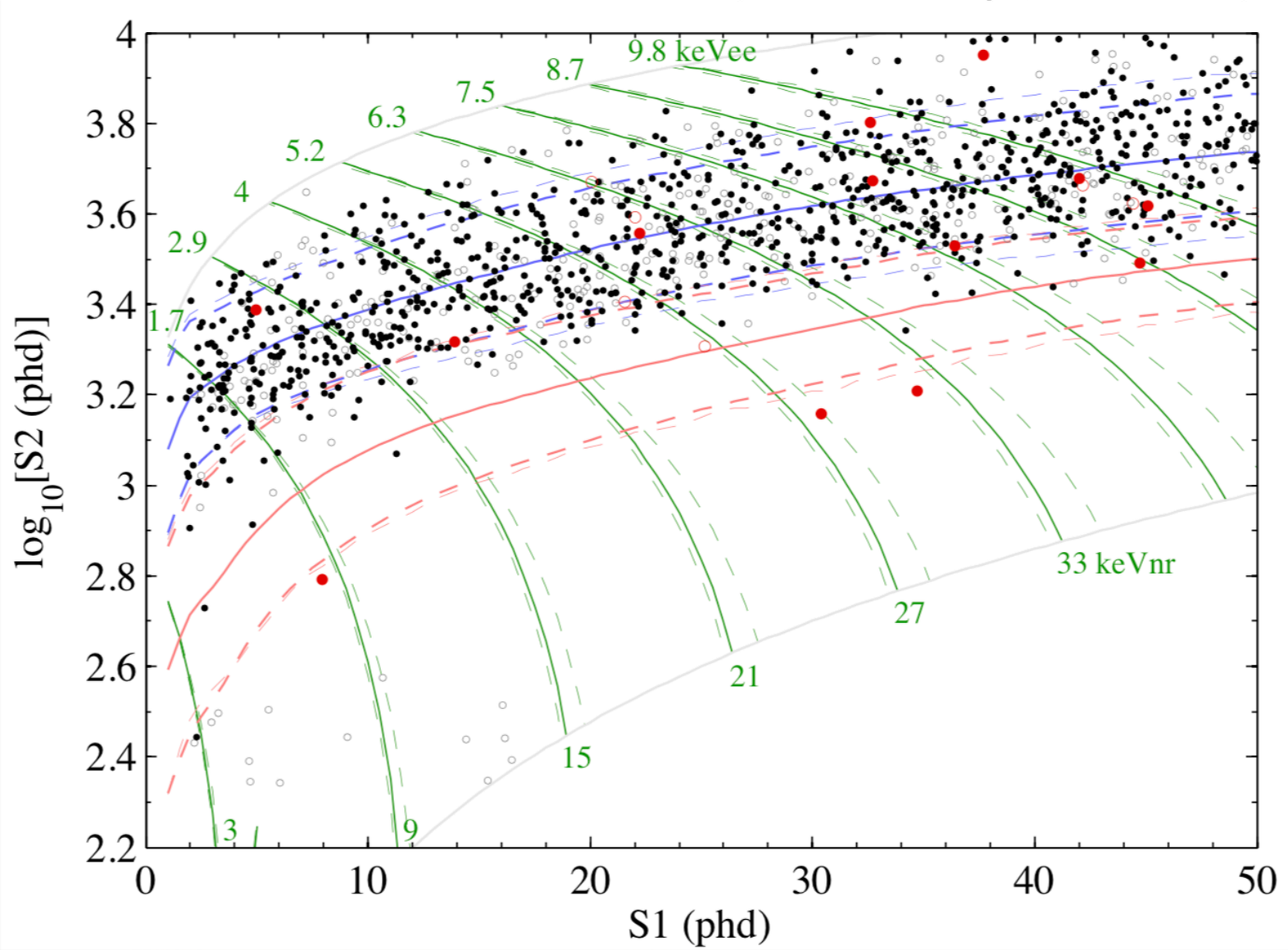
Phys. Rev. Lett. 118, 021303 (2017)

3. Remove salt!
4. ~~Make exclusion plot~~
4. Find anomalous backgrounds 😱 !!!
5. Add post-unsalting cuts (remove light leaks + gas events)

LUX 332 live days Dark Matter Search

1. Start with salted data, define analysis cuts, remove pathological events
2. Identify salt events (blue dots): cool, they lie in signal band, as expected

LUX's second science data (332 live days, unsalted)



Phys. Rev. Lett. 118, 021303 (2017)

3. Remove salt!
4. ~~Make exclusion plot~~
4. Find anomalous backgrounds 😱 !!!
5. Add post-unsalting cuts (remove light leaks + gas events)
6. Notice that the new cuts also remove 10 events in background band (red dots)
7. Could have been noticed sooner!!!

Blinding and Salting in Xenon-1T

Phys. Rev. Lett. 121, 111302 (2018): **"Event reconstruction and selection criteria were fixed prior to unblinding. However, four aspects of the models and statistical inference were modified after unblinding".**

Blinding and Salting in Xenon-1T

Phys. Rev. Lett. 121, 111302 (2018): "Event reconstruction and selection criteria were fixed prior to unblinding. However, four aspects of the models and statistical inference were modified after unblinding".

1: ER safeguard-related

the ER recombination parameterization, previously described, contains improvements implemented to solve a mis-modeling of the ER background in the NR ROI. The pre-unblinding parameterization included a sharp drop at ~1.5 keV_{ee}, which was sufficient for modeling the SR0 ^{220}Rn calibration data in [5] but caused an enhancement to the safeguard term in a post-unblinding fit of the larger statistics SR1 ^{220}Rn and DM search data.

2,3: neutron-releated

model to correctly describe events with enlarged S1s due to additional scatters in the charge-insensitive region below the cathode. These events comprise 13% of the total neutron rate in Table I. Third, we implemented the core mass segmentation to better reflect our knowledge of the neutron background's Z distribution, motivated again by the neutron-like event. This shifts the prob-

(neutron-X events; Z distribution of neutrons)

4: Rn calibrations and misc.

by 13% (4%). Fourth, the estimated signal efficiency decreased relative to the pre-unblinding model due to further matching of the simulated S1 waveform shape to ^{220}Rn data, smaller uncertainties from improved understanding and treatment of detector systematics, and correction of an error in the S1 detection efficiency nuisance parameter. This latter set of improvements was not influenced by unblinded DM search data.

Blinding and Salting in Xenon-1T

Phys. Rev. Lett. 121, 111302 (2018): "Event reconstruction and selection criteria were fixed prior to unblinding. However, four aspects of the models and statistical inference were modified after unblinding".

1: ER safeguard-related

the ER recombination parameterization, previously described, contains improvements implemented to solve a mis-modeling of the ER background in the NR ROI. The pre-unblinding parameterization included a sharp drop at ~1.5 keV_{ee}, which was sufficient for modeling the SR0 ^{220}Rn calibration data in [5] but caused an enhancement to the safeguard term in a post-unblinding fit of the larger statistics SR1 ^{220}Rn and DM search data.

2,3: neutron-releated

model to correctly describe events with enlarged S1s due to additional scatters in the charge-insensitive region below the cathode. These events comprise 13% of the total neutron rate in Table I. Third, we implemented the core mass segmentation to better reflect our knowledge of the neutron background's Z distribution, motivated again by the neutron-like event. This shifts the prob-

(neutron-X events; Z distribution of neutrons)

4: Rn calibrations and misc.

by 13% (4%). Fourth, the estimated signal efficiency decreased relative to the pre-unblinding model due to further matching of the simulated S1 waveform shape to ^{220}Rn data, smaller uncertainties from improved understanding and treatment of detector systematics, and correction of an error in the S1 detection efficiency nuisance parameter. This latter set of improvements was not influenced by unblinded DM search data.

Well... Could salt have helped with the post-unblinding changes?

Blinding and Salting in Xenon-1T

Phys. Rev. Lett. 121, 111302 (2018): **"Event reconstruction and selection criteria were fixed prior to unblinding. However, four aspects of the models and statistical inference were modified after unblinding".**

Well... Could salt have helped with the post-unblinding changes?

Wait for it...

Blinding and Salting in Xenon-1T

Phys. Rev. Lett. 121, 111302 (2018): **"Event reconstruction and selection criteria were fixed prior to unblinding. However, four aspects of the models and statistical inference were modified after unblinding".**

Well... Could salt have helped with the post-unblinding changes?

Wait for it...

"In addition to blinding, the data were also "salted" by injecting an undisclosed number and class of events in order to protect against fine-tuning of models or selection conditions in the post-unblinding phase. After the post-unblinding modifications described above, the number of injected salt and their properties were revealed to be two randomly selected AmBe events, which had not motivated any post-unblinding scrutiny".

No, not really 😱

Salting: the Panacea to all Background Problems?

1. Classic blinding assumes that the backgrounds are known well enough
   - Which is very rarely the case (see DarkSide)

2. In principle, salting gives access to ALL anomalous backgrounds, but...
   - Both LUX and Xenon-1T had to resort to post-unblinding/salting cuts
   - The reality is that people* are going to run out of time, patience, etc.
   - Especially in a highly competitive environment, such as right now

Salting: the Panacea to all Background Problems?

1. Classic blinding assumes that the backgrounds are known well enough
   - Which is very rarely the case (see DarkSide)

2. In principle, salting gives access to ALL anomalous backgrounds, but...
   - Both LUX and Xenon-1T had to resort to post-unblinding/salting cuts
   - The reality is that people* are going to run out of time, patience, etc.
   - Especially in a highly competitive environment, such as right now

OR IN OTHER WORDS:

What is "sufficient background modeling" or "sufficient event inspection"?

Is there a "blinding" technique that will "save" us from unknown unknowns?

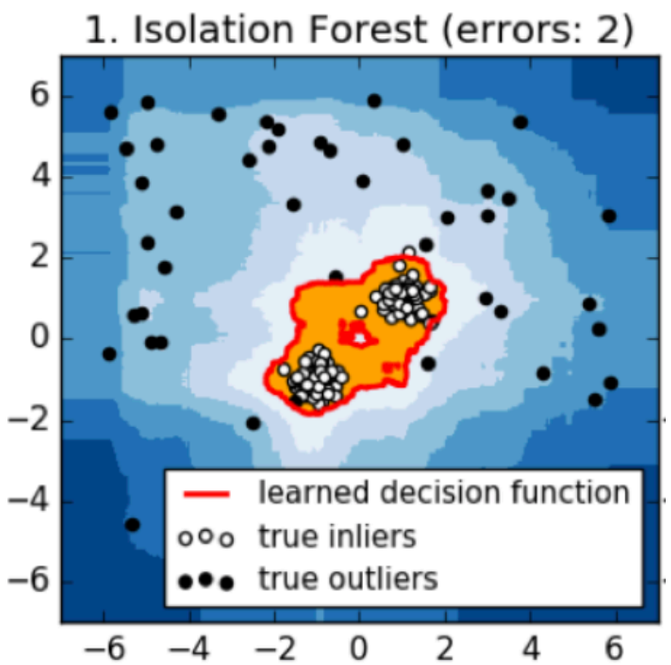
NB: our stated goal is "bias mitigation", not "bias elimination" (unrealistic)

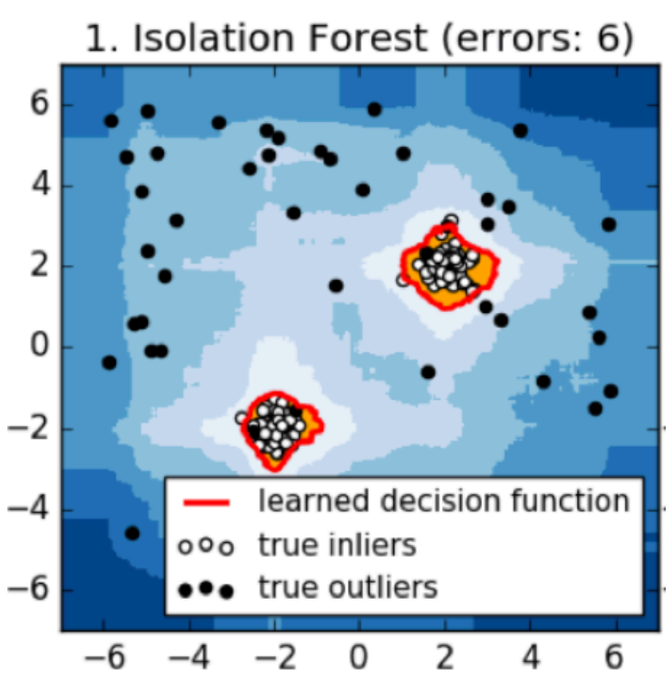
How to tackle Anomalous Backgrounds?

* What if we get rid of people? Is there something with infinite patience?

How to tackle Anomalous Backgrounds?

*** What if we get rid of people? Is there something with infinite patience?**

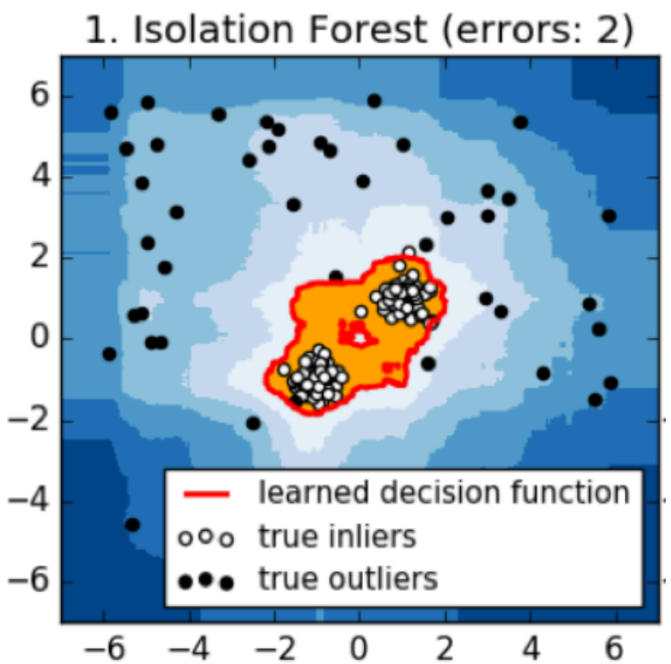


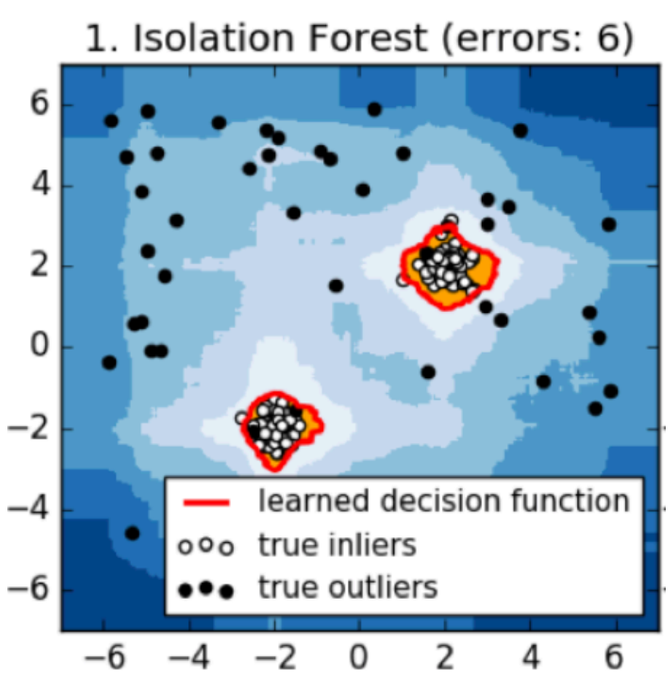


1. **Anomaly/Outlier detection is deeply/widely studied problem in Machine Learning**
   - **Widely used for security/intrusion detection (keystroke pattern of intruders is sufficiently different from regular users) – BRO at NERSC**
   - **Under study for medical applications 😱**

How to tackle Anomalous Backgrounds?

*** What if we get rid of people? Is there something with infinite patience?**





1. **Anomaly/Outlier detection is deeply/widely studied problem in Machine Learning**
   - **Widely used for security/intrusion detection (keystroke pattern of intruders is sufficiently different from regular users) – BRO at NERSC**
   - **Under study for medical applications 😱**

2. **Could we "borrow" one of these techniques for tagging anomalous backgrounds in DM searches?**
   - **I really don't know, but it'd be fun to try**
   - **And this is beyond the scope of my talk...**
   - **Let me know if you'd like to play with this?**

Summary and Conclusions

- **Bias mitigation crucial in high-stakes science**
  - **The DM community is mostly on board**

Summary and Conclusions

- **Bias mitigation crucial in high-stakes science**
  - **The DM community is mostly on board**
- **Multiple classes of blinding strategies for DM**
  - **Blind box (+ pre-scaling)**
  - **Salting/spiking/blind injections**
- **Blinding more vulnerable to rare backgrounds**
  - **But salting is not 100% successful so far**
  - **More accuracy in bkg/detector modeling**
  - **But how to deal with unknown unknowns?**
  - **"Bias mitigation" vs "bias elimination"**

Summary and Conclusions

- **Bias mitigation crucial in high-stakes science**
  - **The DM community is mostly on board**
- **Multiple classes of blinding strategies for DM**
  - **Blind box (+ pre-scaling)**
  - **Salting/spiking/blind injections**
- **Blinding more vulnerable to rare backgrounds**
  - **But salting is not 100% successful so far**
  - **More accuracy in bkg/detector modeling**
  - **But how to deal with unknown unknowns?**
  - **"Bias mitigation" vs "bias elimination"**
- **A discovery may be right around the corner**
  - **Everyone should be thinking about this**
  - **Machine Learning to the rescue?**

Backups

Pathological Events in LUX Run 2

Events A & B:

event outside the TPC, and light leaking through a gap close to the PMT array.

Event C:

timing/pattern consistent with gas scintillation emission (also outside the TPC volume).