Med-Fi Covariances: caveats, with Optical Potentials as an example

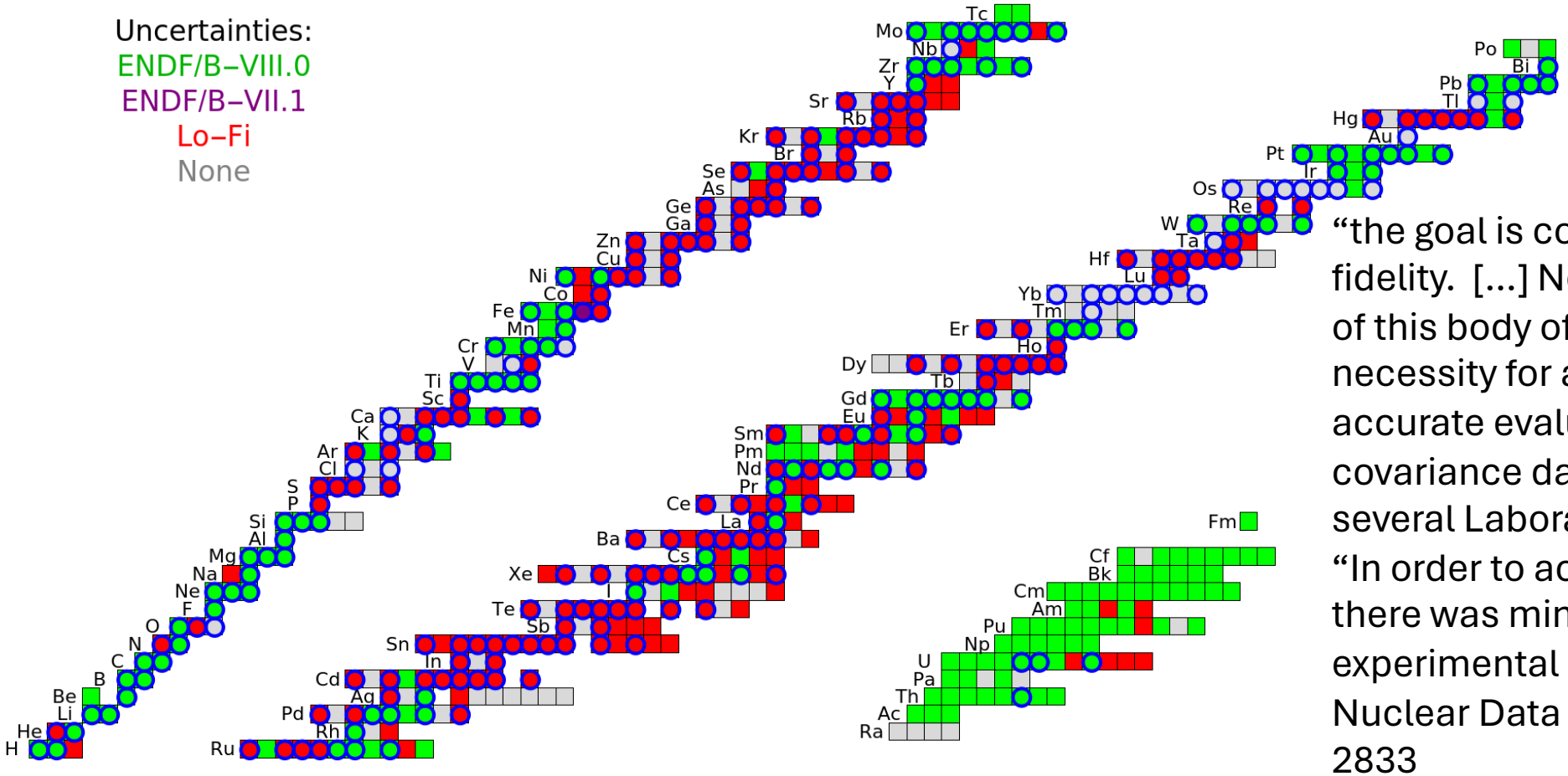
K. Kravvaris, with input from C. D. Pruitt and R. J. Casperson

LLNL-PRES-860934
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Distribution of data with select covariances: do not stray far from the line of stability

Evaluations that include (n,el), (n,n'), (n,2n), and (n,γ) covariances:



"the goal is completeness, not high fidelity. [...] Neither would the existence of this body of data remove at all the necessity for a more methodical and accurate evaluation of important covariance data, such as is underway at several Laboratories."

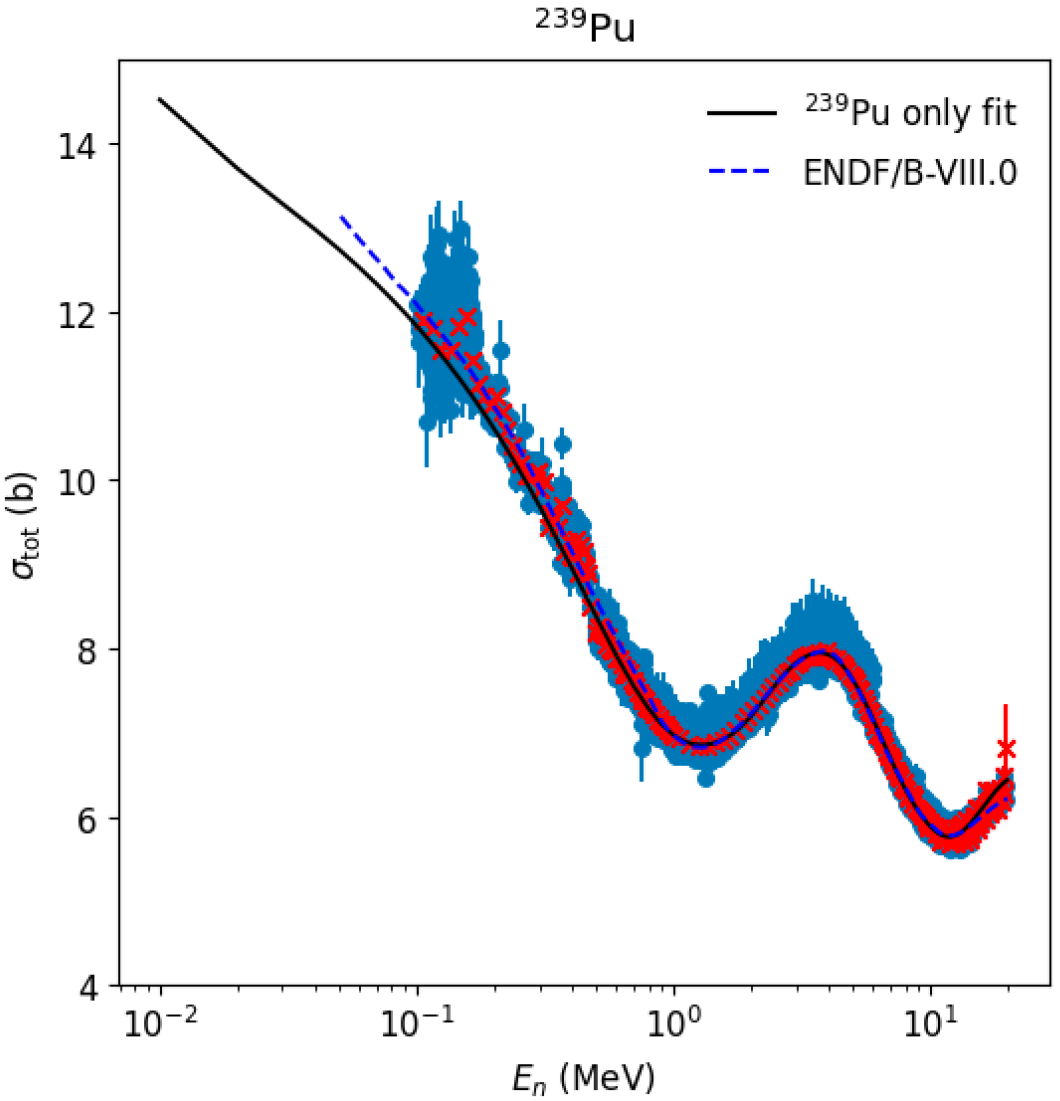
"In order to accomplish this large task, there was minimal utilization of experimental data."

Nuclear Data Sheets 109 (2008) 2828–2833

Image credit to Robert Casperson

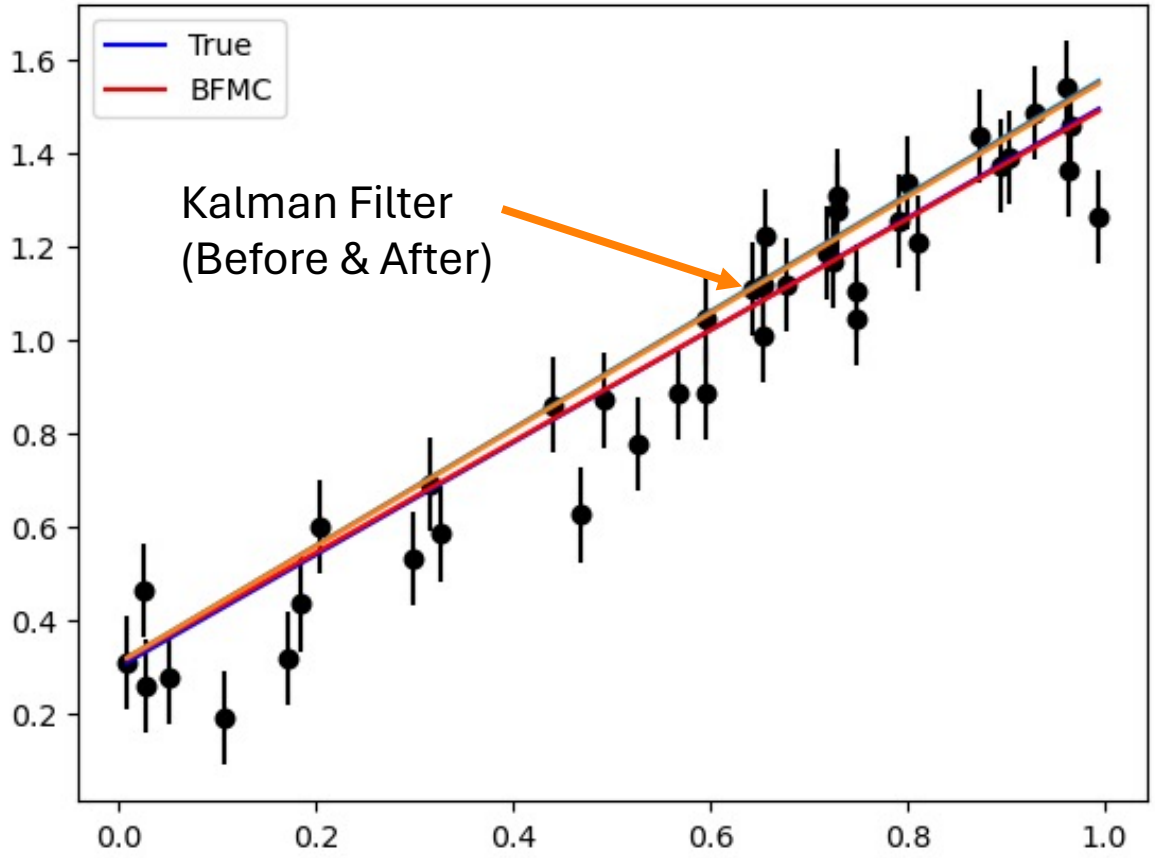
What does it mean in practice? In regions where theory works well data can be used to gauge parameters and uncertainties

- Fit optical model potential to reproduce data.
- Reproduce ENDF result (obtained using generalized least squares)
- Post-fit, use various methods (Kalman Filter/Backward-Forward Monte Carlo) to obtain covariances by mixing theory with data.



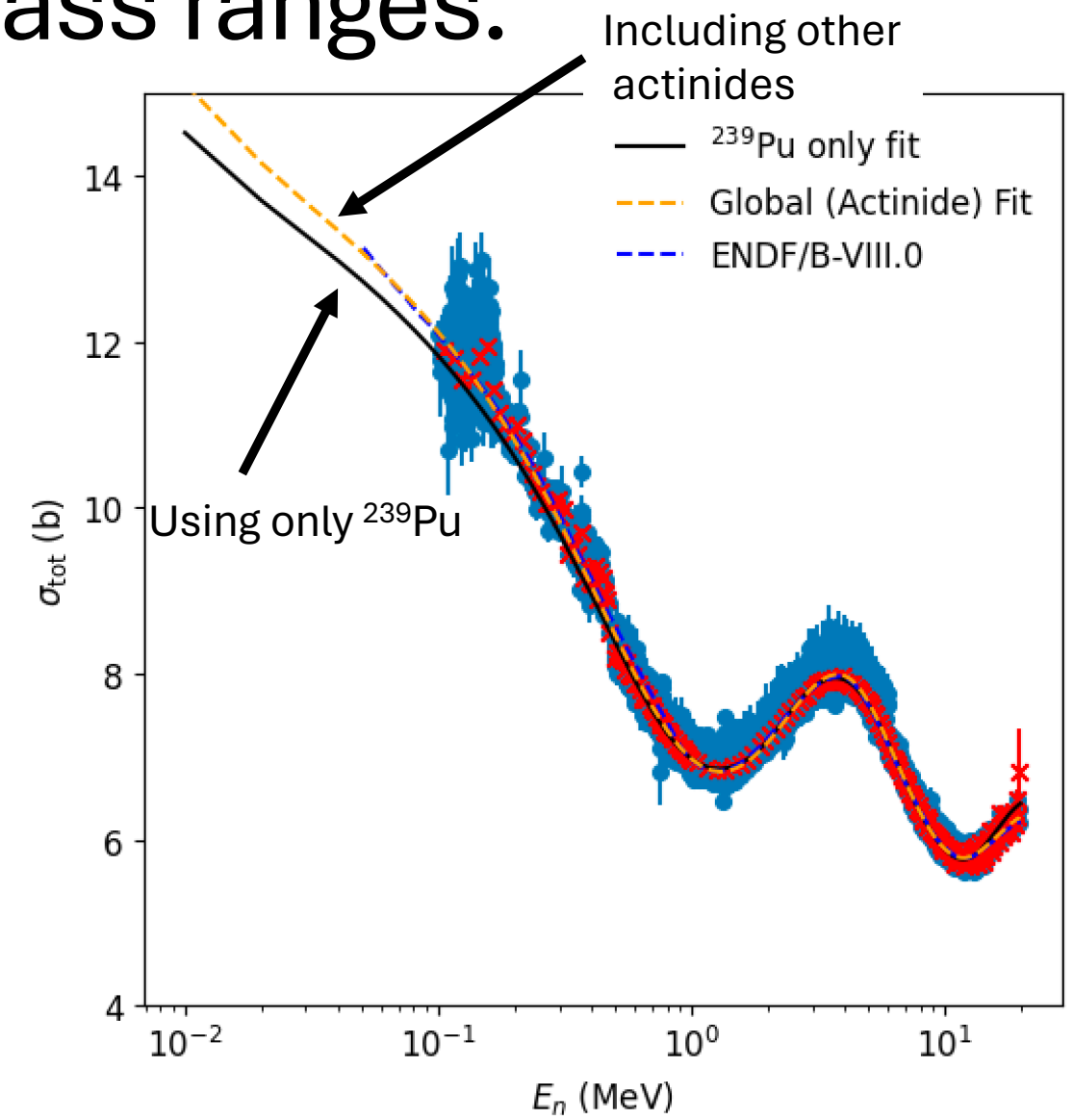
In a perfect world, with data statistically distributed, both methods seem to work

- BFMC mean is closer to the true value post-data introduction.
- Multiple realizations, each get assigned a different weight according to agreement to data.
- But how do pathogenies in the data manifest/propagate to the evaluation?

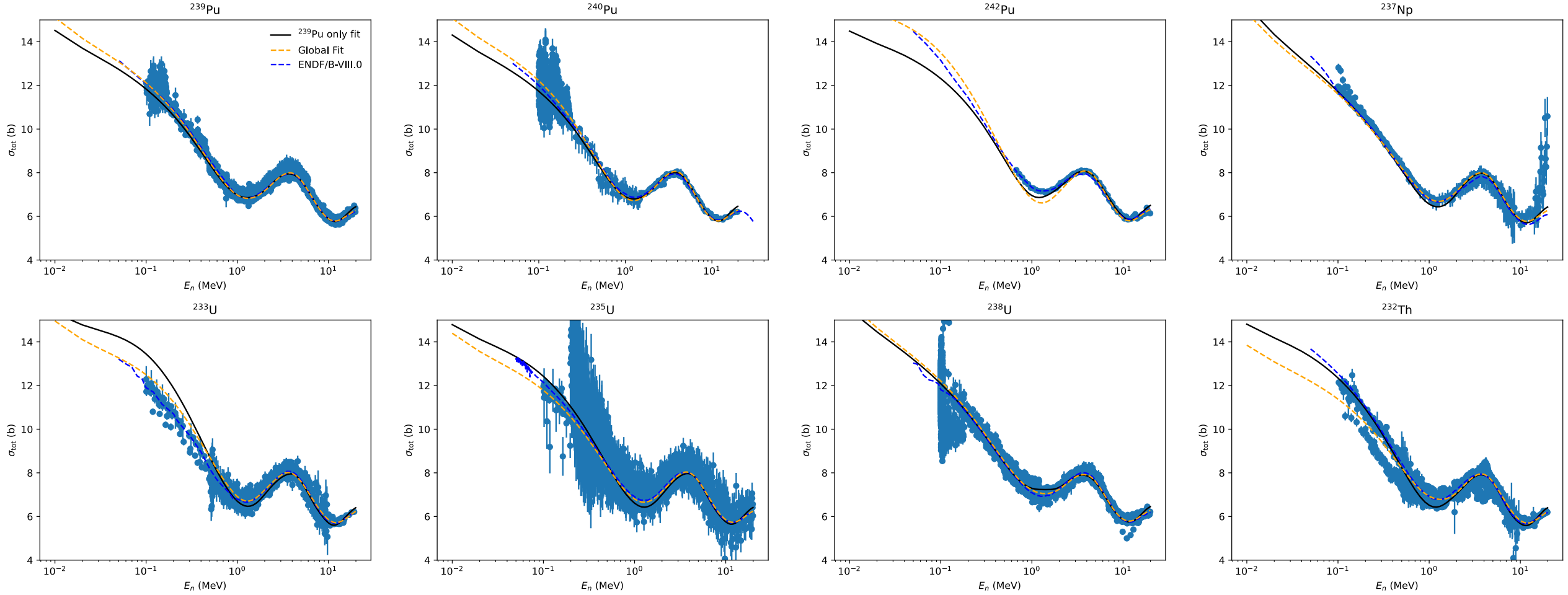


Most codes can handle inputs that are correlated across small mass ranges.

- Example: optical potential parameters:
  Linear with mass/isospin asymmetry
- Hauser-Feshbach calculations:
  Same level density/strength function/(fission) parameters for a given nucleus regardless if it is reached by neutron absorption, (γ,n), (n,2n) etc.



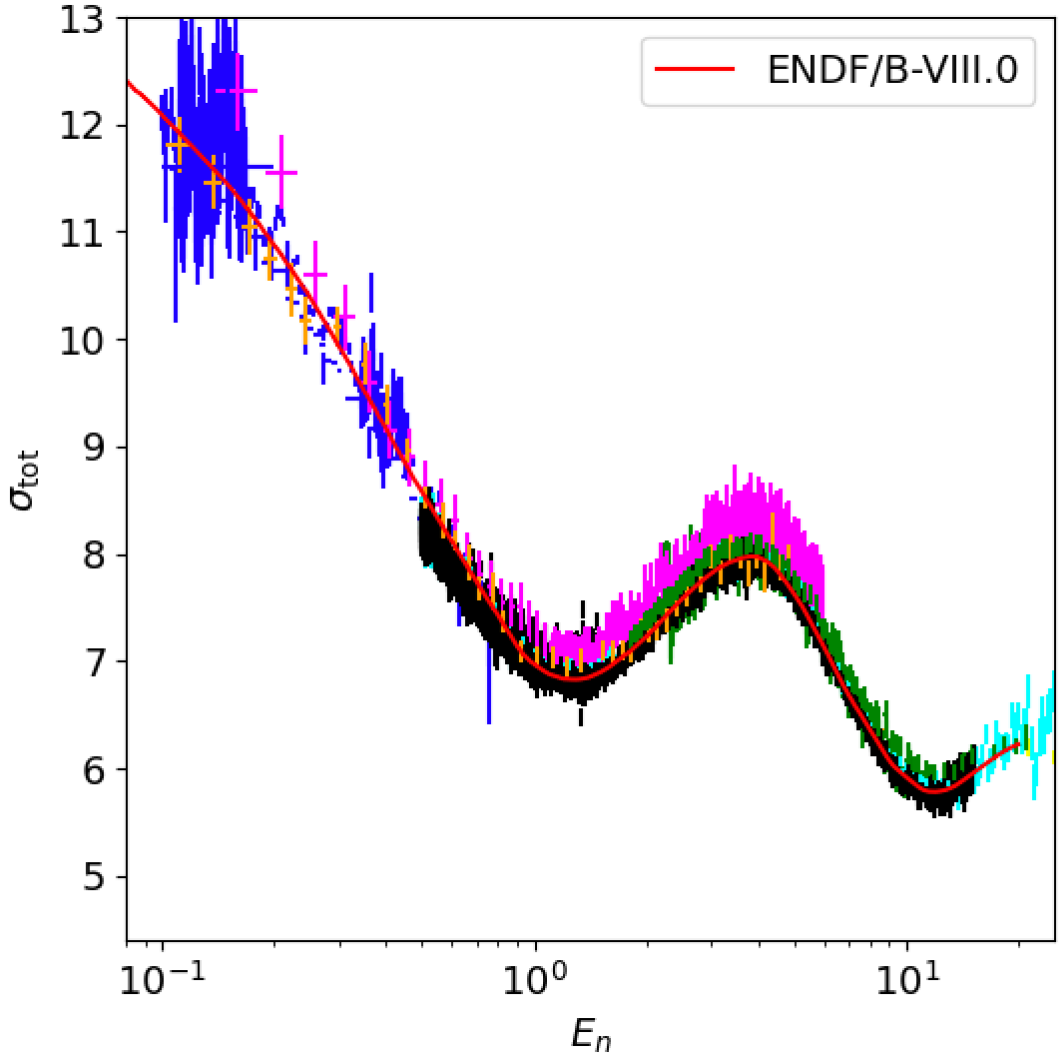
Result is a more robust fit, that also extends across multiple nuclei. Can we extend data to discern mass dependence ?



Need: correlated data sets are highly valuable.

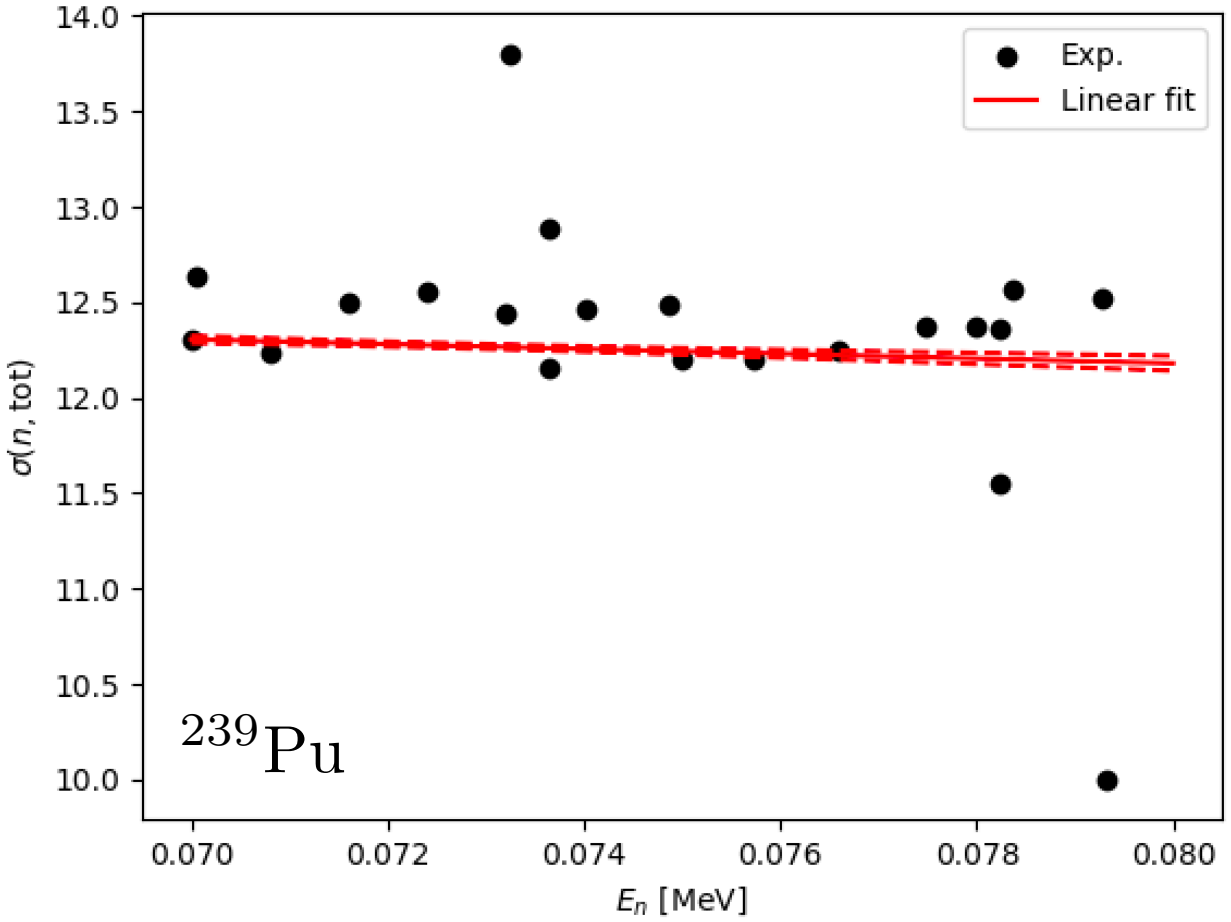
If data is abundant, we can try to ignore theory

- ^{239}Pu(n,tot) cross section has multiple measurements across various energy ranges.
- Clear normalization disagreements between various experiments.
- ENDF evaluation is smooth in this case, not necessarily the case for other actinides.



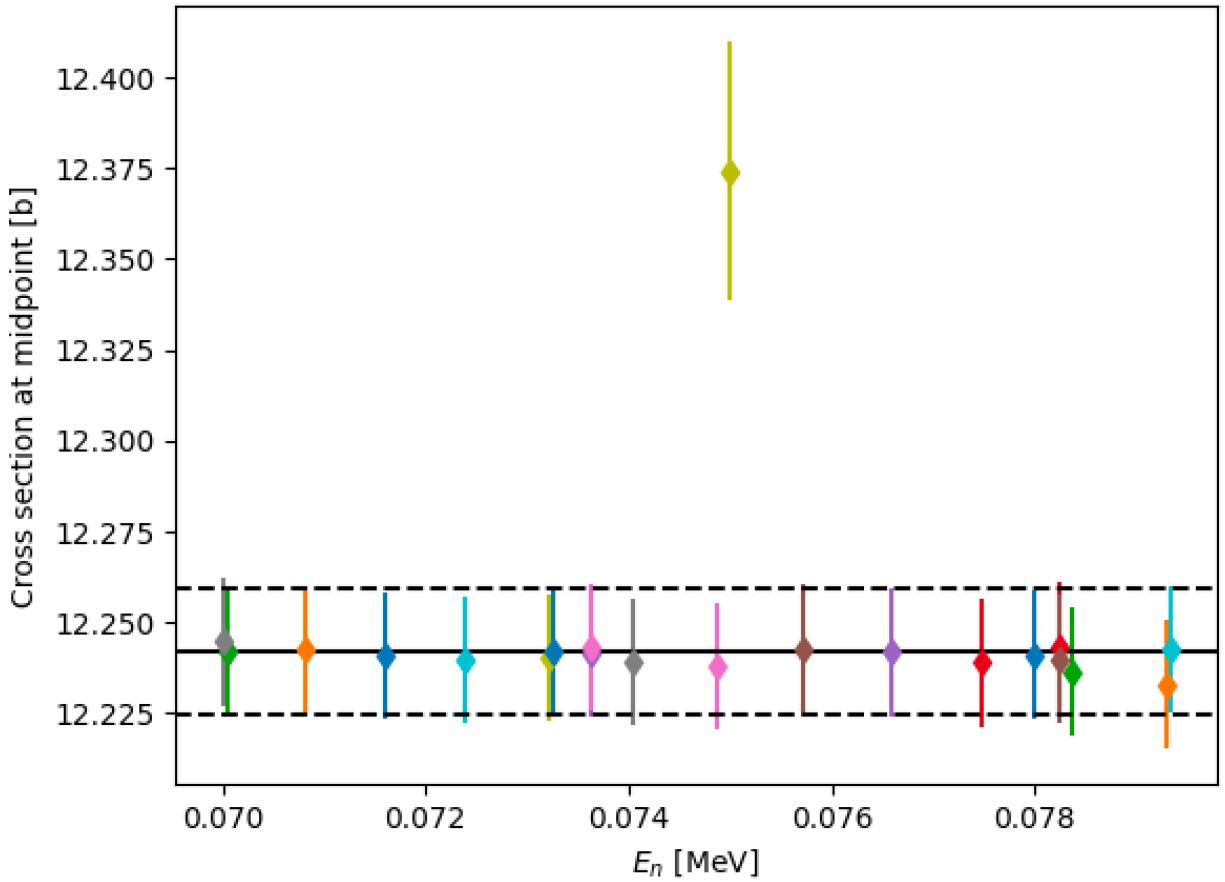
By binning the data in small energy increments we can discern driving parts

- Make small (10 keV) bins in energy and fit a straight line.
- The uncertainties **are** plotted here.
- Line can be evaluated at mid-bin as a reference.
- Clearly, while the data could follow the linear model, statistically speaking, they don't agree with it (>3σ discrepancies).



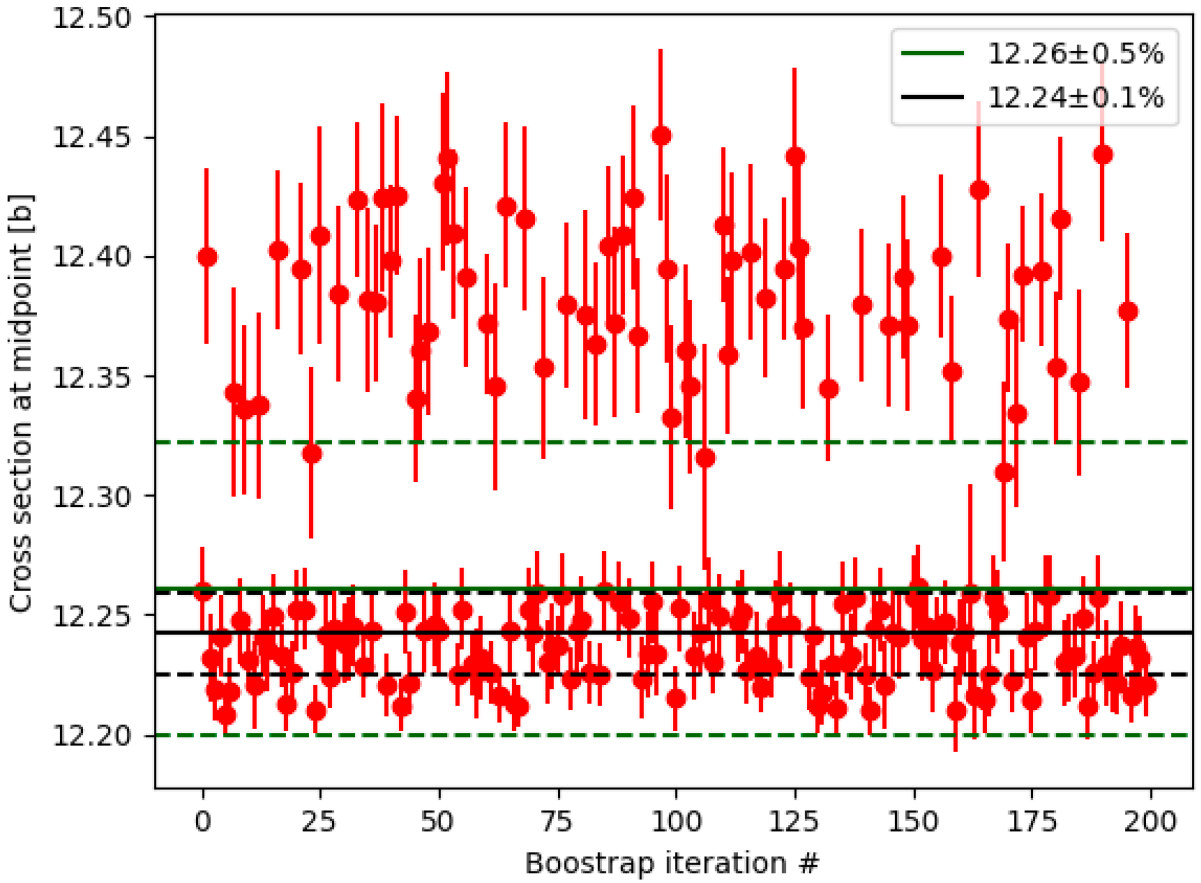
We can re-do the fit excluding each data point and see how they compare

- Single data point dominates the fit (difference of ~1%).
- Uncertainty of prediction almost doubles.
- Looking back, it corresponds to a single point that has ~1 order of magnitude less uncertainty than the rest (from Harvey, 1988, σ_{tot}=12.2 b at 0.075 MeV).
- The rest of the measurements are a perturbation of 0.4% of the mean.

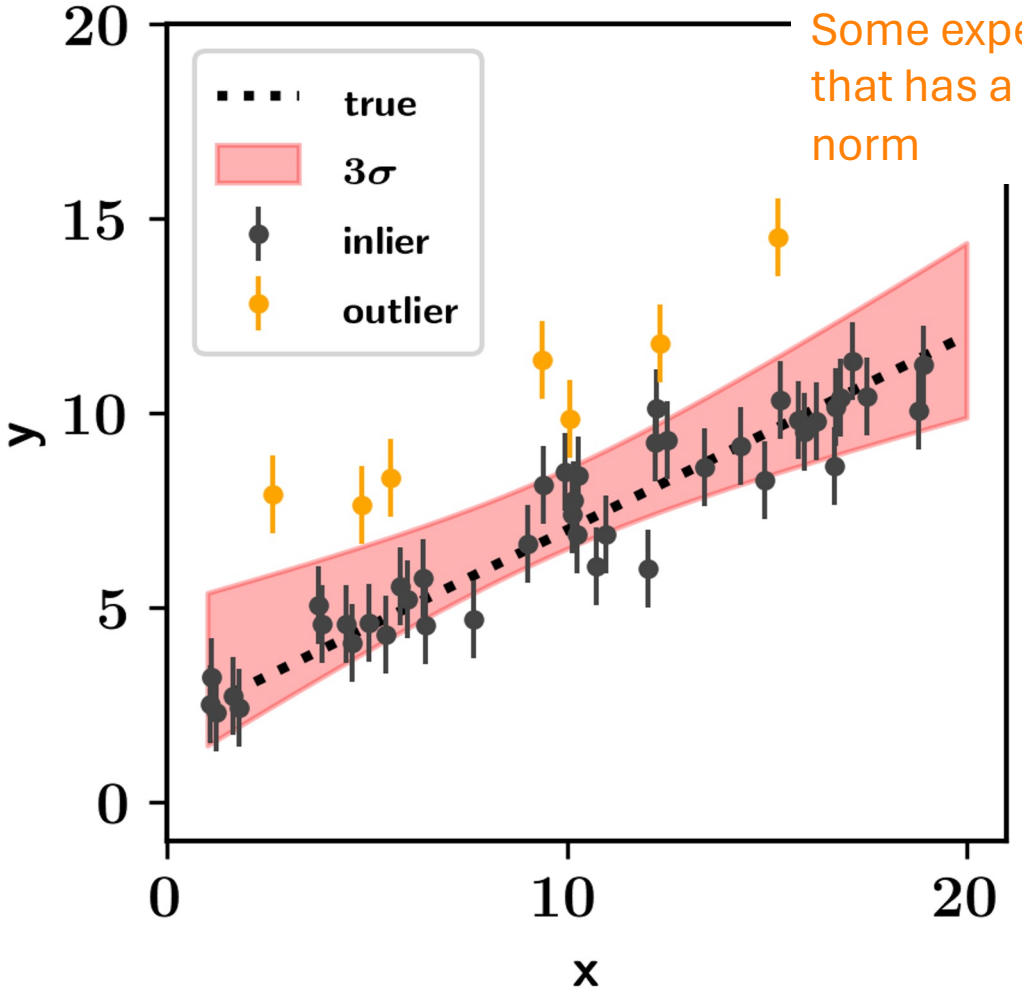


Bootstrapping also points to the same conclusions.

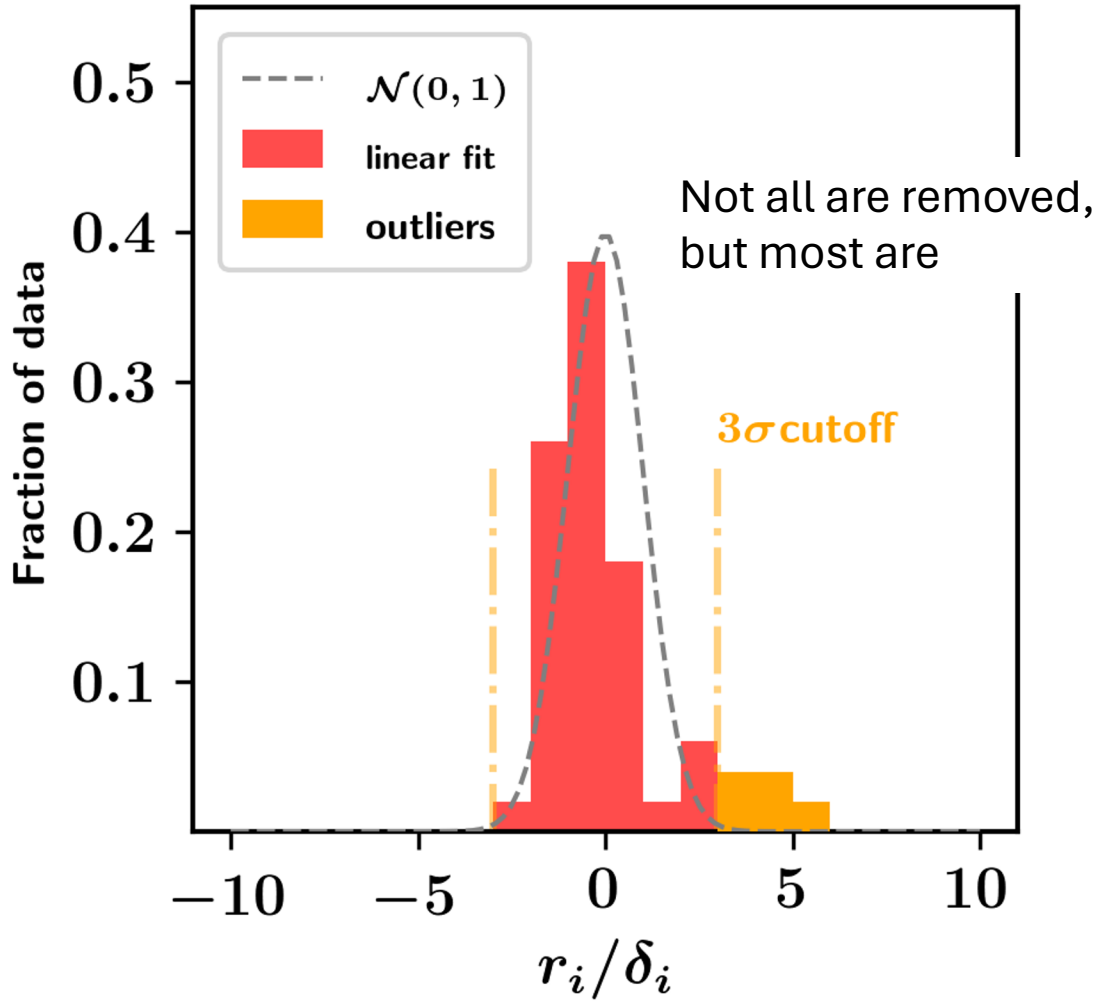
- Take a subset of data points, replicate them to have the same total number of data, repeat.
- Hover between the two values, but now the rest of the experiments have a larger weight.



Iterative outlier rejection can be automated



Some experiment that has a different norm



Not all are removed, but most are

Images credit to Cole Pruitt

Conclusions

- We use what data exists to constrain the model(s).
- Moving to regions of the chart vs specific nuclei represents new challenges.
- Should test data sets for consistency with one another: could be automated?
- Inflating error bars could also work (used by PDG).
- Is automatic exclusion a weapon of math destruction?
- Cases where data is sparse require further attention (for example, ^{239}Pu(n,2n))
- As is, only applies to fast region (>100 keV). Connect regions?