Data Driven "Low-Fi" Covariances

Filling in missing and unlikely covariances in modern evaluations

Kyle Wendt
Lawrence Livermore National Laboratory

WPEC SG 50
September 14, 2020

LLNL-PRES-814499
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

What to do when *useful* covariances are missing?

- Find and yell at the evaluator:
  — "Why is this covariance such #%##$&%&^"

- Request/fund a re-evaluation with a full quantification of uncertainties and systematic errors
  — Expensive without investments in automating the evaluation process.

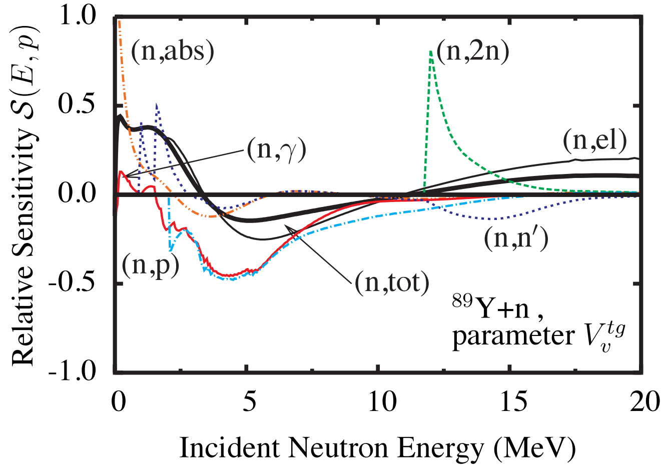
- Construct simple "ad hoc" covariances based on
  - Differences between existing evaluation libraries.
  - Comparison of mean values with spreads of experimental data
  - Model-dependence between channels
    - Clone covariance pattern in library for neighbors in this nuclear region.
    - For example, elastic and inelastic are commonly anti-correlated.

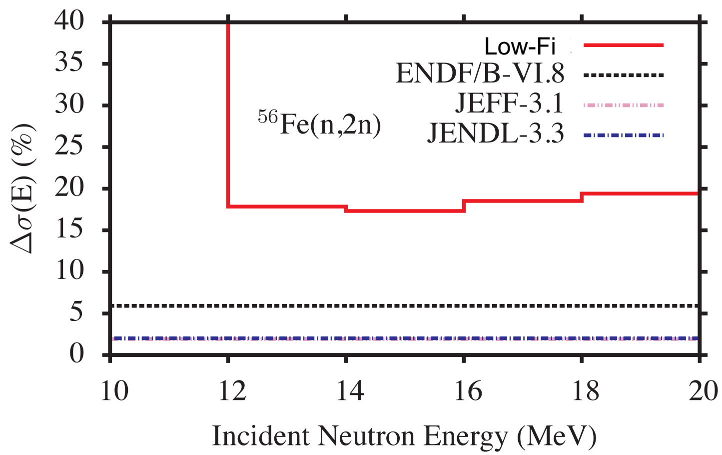
- Use low-fidelity ("Low-Fi") covariances described by Little *et al* (2008):
  - Fills in gaps in ENDF/B-VII.0

- Use "machine learning" like approaches to generate covariances

The Low-Fi Approach (For Fast Region)

Little *et al* (2008):
M. T. Pigni *et al* (2009)

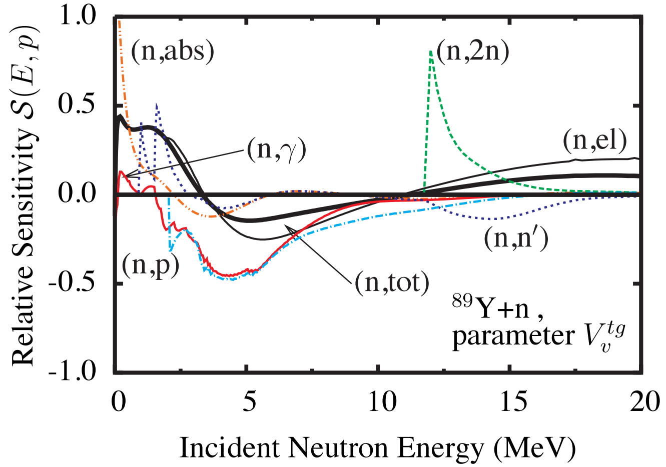


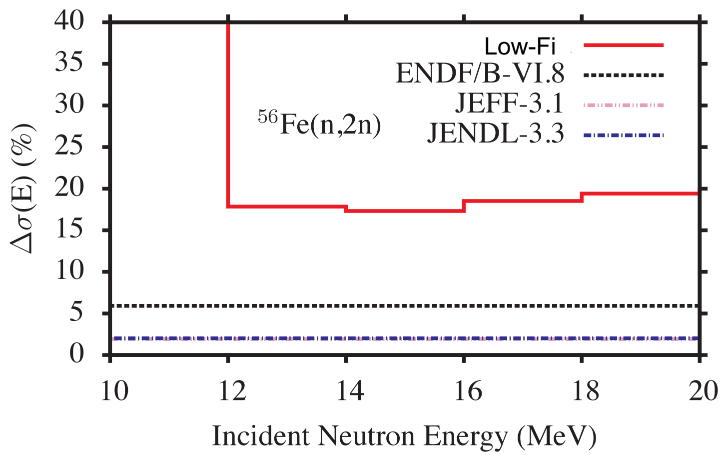


- Based on model variation.
  - Low-Fi's parameter variations built on **intuition** of Low-Fi collaboration
  - Extending to ENDF-VIII requires codification of that intuition.
- No cross correlations
- No direct connection back to measured data.
- Targeted quick approximate covariances to fill out library.

The Low-Fi Approach (For Fast Region)

Little *et al* (2008):
M. T. Pigni *et al* (2009)





- Based on model variation.
  - Low-Fi's parameter variations built on intuition of Low-Fi collaboration
  - Extending to ENDF-VIII requires codification of that intuition.

Model + **Intuition** Driven
Surrogate Covariances

Towards data driven "Low-Fi" covariances

Abstractly, an evaluation with a covariance matrix represents a way to sample a set of (nearly) continuous functions that are distributed pointwise as a multivariate Gaussian*

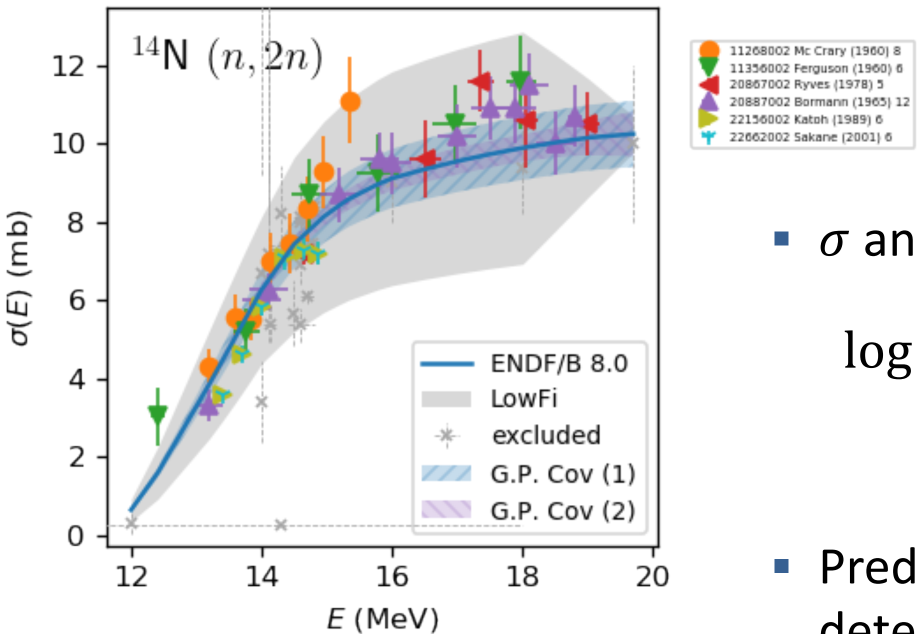
- **i.e. A Gaussian Process**

$$F(x) \sim GP\left(\mu(x), K(x', x)\right)$$

- $\mu(x)$ is the average or mean function
  - The evaluation + interpolation rules
- $K(x', x)$ is the covariance kernel of the functions
  - The covariance suite + interpolation rules
- $\mu(x), K(x', x)$ are often parameterized

*This is only approximately true, physical cross sections and distributions are positive definite, Gaussian stochastic variables are not.

What does it mean to be data driven



$$K(x', x) = \sigma^2 f(x) f(x') e^{-\frac{(x-x')^2}{2l^2}}$$

- σ and l are hyperparameters trained on data.

$$\log p(y_d|M) = -\frac{1}{2} y_d \widetilde{K}^{-1} y_d - \frac{1}{2} \log|K| + C$$

$$\widetilde{K} = K + \Sigma_d$$

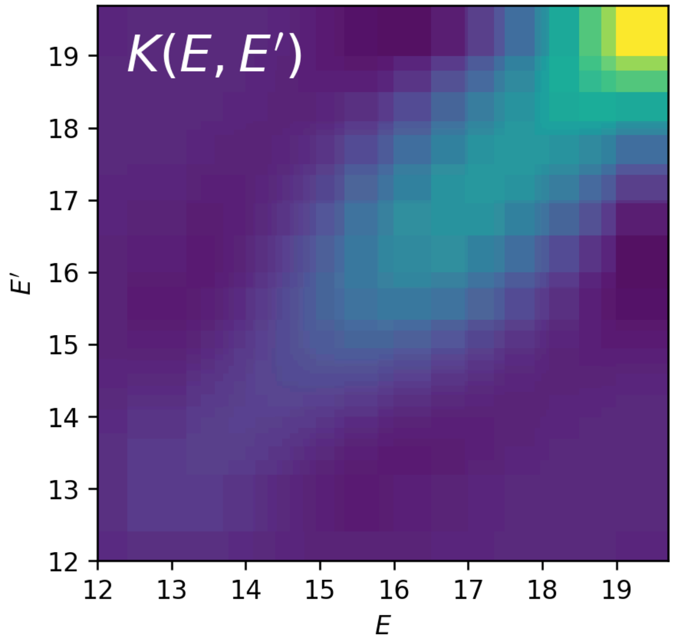
- Predicted effective mean and covariance are determined by MVN conditioned on training data

$$\bar{\mu}(x) = \mu(x) + K(x, \boldsymbol{x_d}) \widetilde{K}(\boldsymbol{x_d}, \boldsymbol{x'_d})^{-1} (\boldsymbol{y_d} - \mu(\boldsymbol{x_d}))$$

$$\bar{K}(x, x') = K(x, x') - K(x, \boldsymbol{x_d}) \widetilde{K}(\boldsymbol{x_d}, \boldsymbol{x'_d})^{-1} K(\boldsymbol{x_d}, x)$$

The surrogate covariance function is determined fully by data (x_d, y_d) and the predicted mean of the evaluation ($\mu(x)$), instead of parameter variations.

What does it mean to be data driven



$$K(x', x) = \sigma^2 f(x) f(x') e^{-\frac{(x-x')^2}{2l^2}}$$

- σ and l are hyperparameters trained on data.

$$\log p(y_d|M) = -\frac{1}{2} y_d \widetilde{K}^{-1} y_d - \frac{1}{2} \log|K| + C$$

$$\widetilde{K} = K + \Sigma_d$$

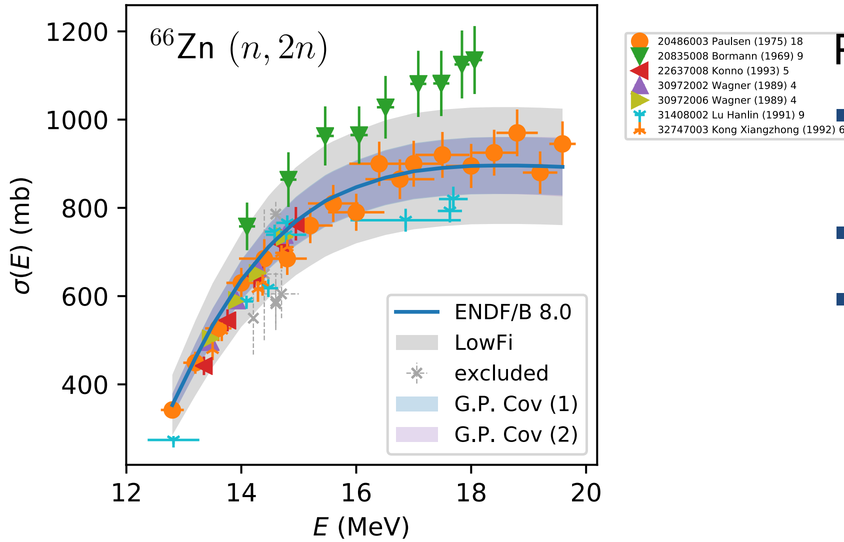
- Predicted effective mean and covariance are determined by MVN conditioned on training data

$$\bar{\mu}(x) = \mu(x) + K(x, \boldsymbol{x_d}) \widetilde{K}(\boldsymbol{x_d}, \boldsymbol{x'_d})^{-1} (\boldsymbol{y_d} - \mu(x_d))$$

$$\bar{K}(x, x') = K(x, x') - K(x, \boldsymbol{x_d}) \widetilde{K}(\boldsymbol{x_d}, \boldsymbol{x'_d})^{-1} K(\boldsymbol{x_d}, x)$$

The surrogate covariance function is determined fully by data (x_d, y_d) and the predicted mean of the evaluation ($\mu(x)$), instead of parameter variations.

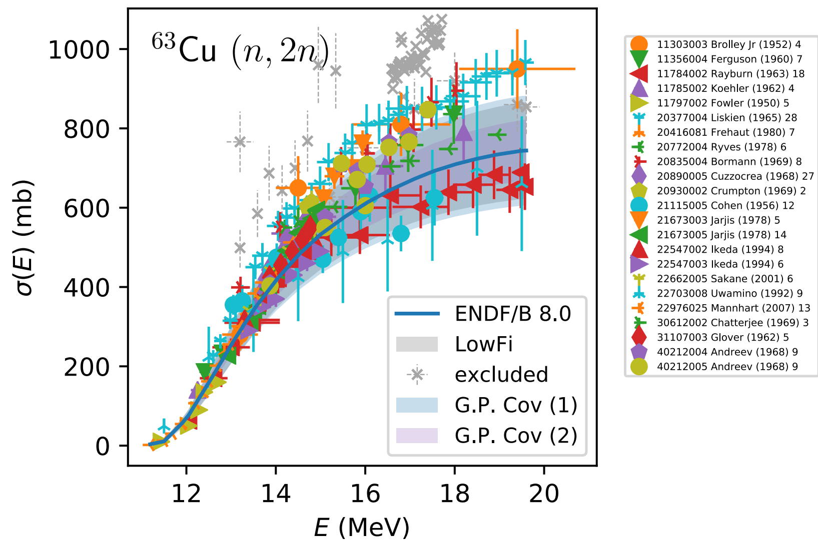
Gaussian Processes to Make Up Covariances



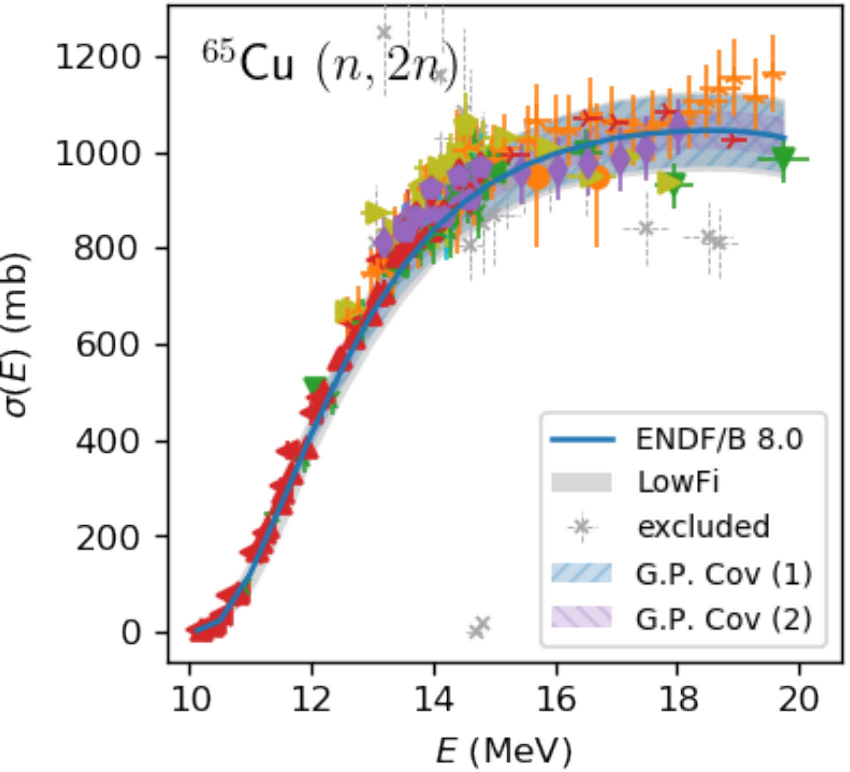
Purely data driven

- No evaluation code is used, extremely fast and simple to run.
- Can be applied directly by a user.
- Depends critically on quality of data used.

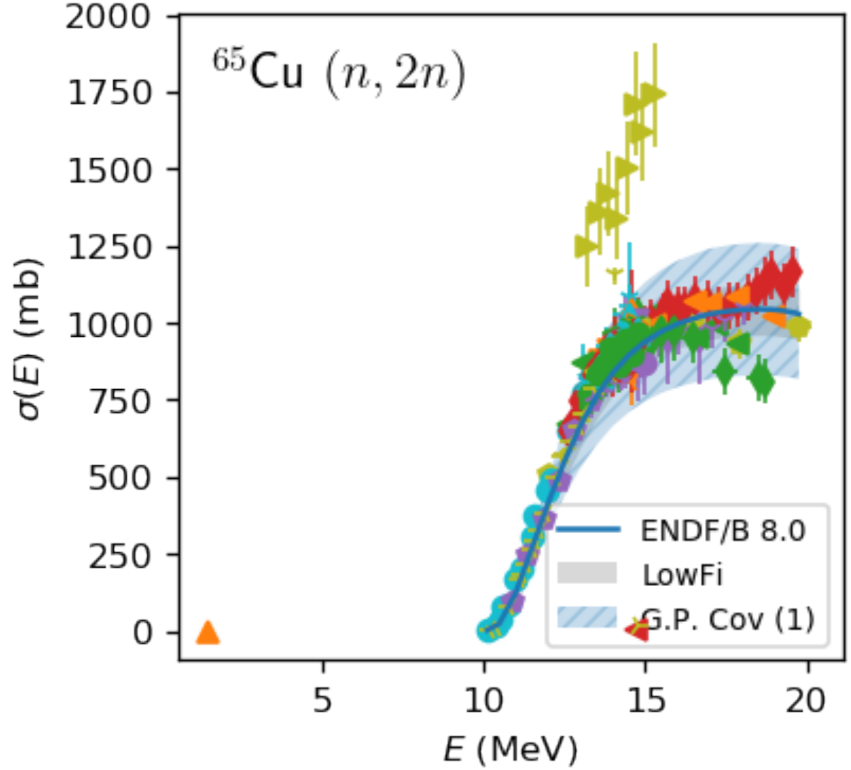
- Presently, EXFOR is hand processed, *correctness* validated, and extremal values are potentially excluded.
  - Is this a task a user should be trusted with? **Likely not.**
- Pre-curation of EXFOR would allow for such a tool for be full automated.



Data Curation is critical to reasonable error bars



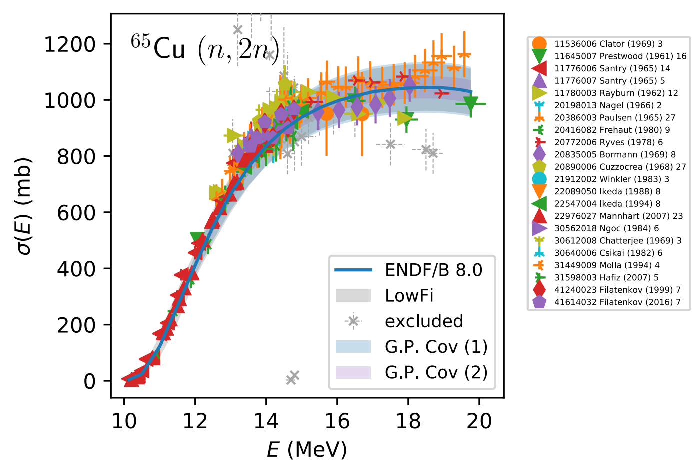
With light curation, hyperparameters tend to adopt reasonable values, $l = \sim 3$ MeV



Without any curation, hyperparameters explore or collapse, $l = \sim 15$ MeV

Conclusions

- Modern nuclear data libraries have many inadequate or incorrect covariances; Limits uncertainty analysis of applications that consume nuclear data.
- Present solutions to supplement covariances
  - Ad hoc mix and match from nearby evaluations.
  - Low-fidelity "fill-in" covariances capture model variations
- Potential future solutions
  - Gaussian processes provide a formalism to extract "data driven" covariances for fast region of cross sections.
  - More general machine learning can replace concept of covariances completely.
    - But this would require "new" evaluations.
- We need reliable (correctly encoded) lightly curated inputs to enable such solutions.



Lawrence Livermore
National Laboratory