

Augmenting the pipeline with sparse Bayesian modeling

Noah A.W. Walton, Denise Neudecker,
Mike J. Grosskopf, Scott A. Vander Wiel

02/11/2026

LA-UR-26-20777

Sparse Bayesian modeling isolates meaningful phenomena in complex data

Opportunity

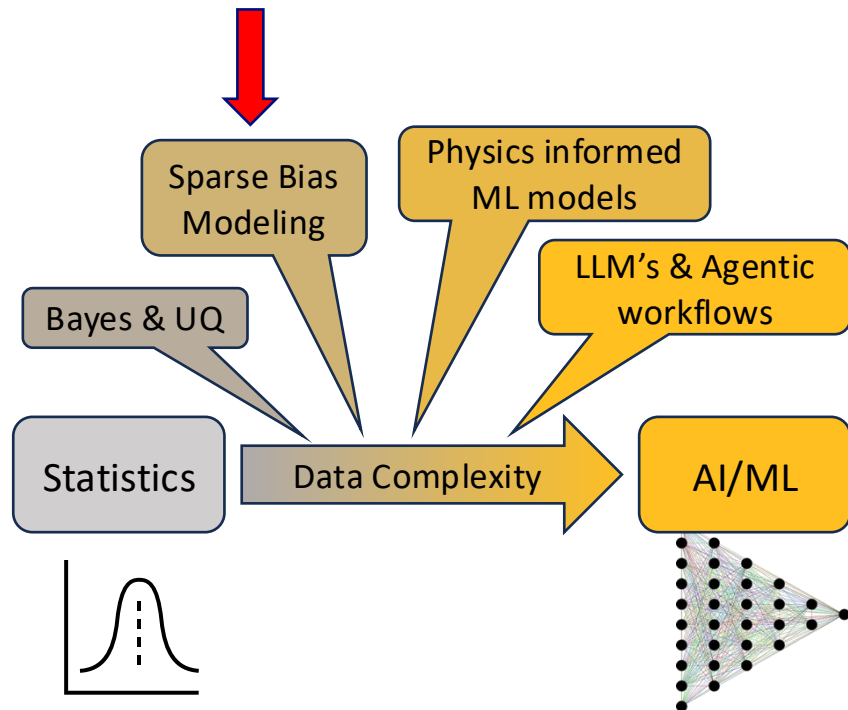
- Could be useful tool along the nuclear data pipeline
 - Success story: AIACHNE experimental bias identification
 - Concept is general, other opportunities?

Needs for further adoption/application

- AI/ML-ready experimental database (features)
- Largely enabled by advances in hierarchical modelling and PPLs
 - Solutions may require differentiable physics models
 - Ready made tools

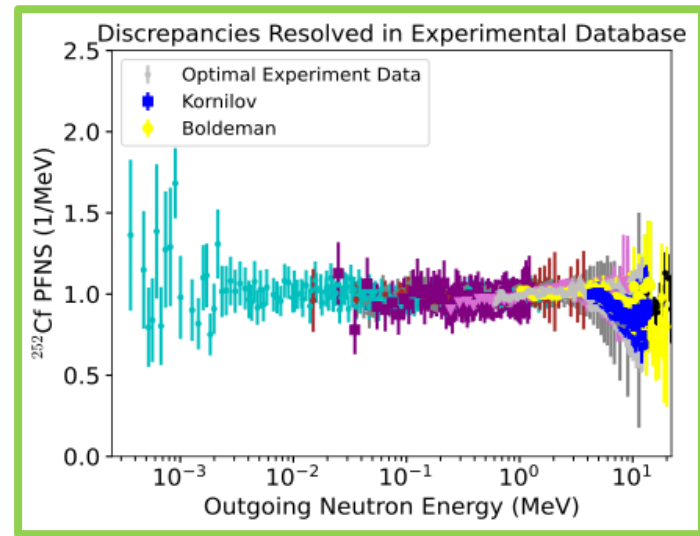
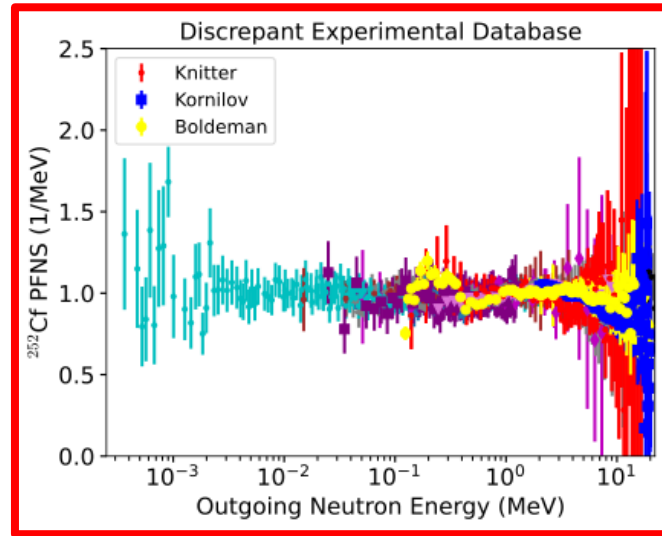
Sparse Bayesian modeling isolates meaningful phenomena in complex data

- Augments traditional evaluation framework
- Maintains interpretability
 - Closer to statistics than AI
- Sparsity implies most parameters have some default value while a few can extend to other values



AIACHNE used sparsity to model unknown experimental biases in ^{252}Cf PFNS data

Neutron Data
Standard



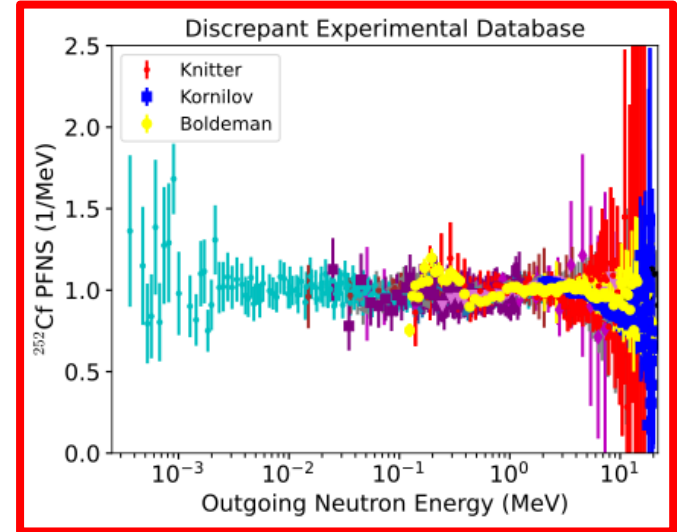
Resolving these discrepancies significantly improved the evaluation and reduced uncertainty by up to 25%

Sparsity encodes domain knowledge as Bayesian prior

Domain Knowledge

- Some data are discrepant, but we can't rigorously determine which ones
 - All have the possibility to be biased, but only a few are
- Bias is likely related to physics phenomenon unique to some data

- Priors are often hierarchical distributions
- PPLs allow us to effectively sample from it



Domain knowledge informs bias representation and level of sparsity

- Sparse bias model:
 - Gaussian basis functions
 - Logarithmic in energy
 - Multiplicative
 - Active groups are related to experimental metadata

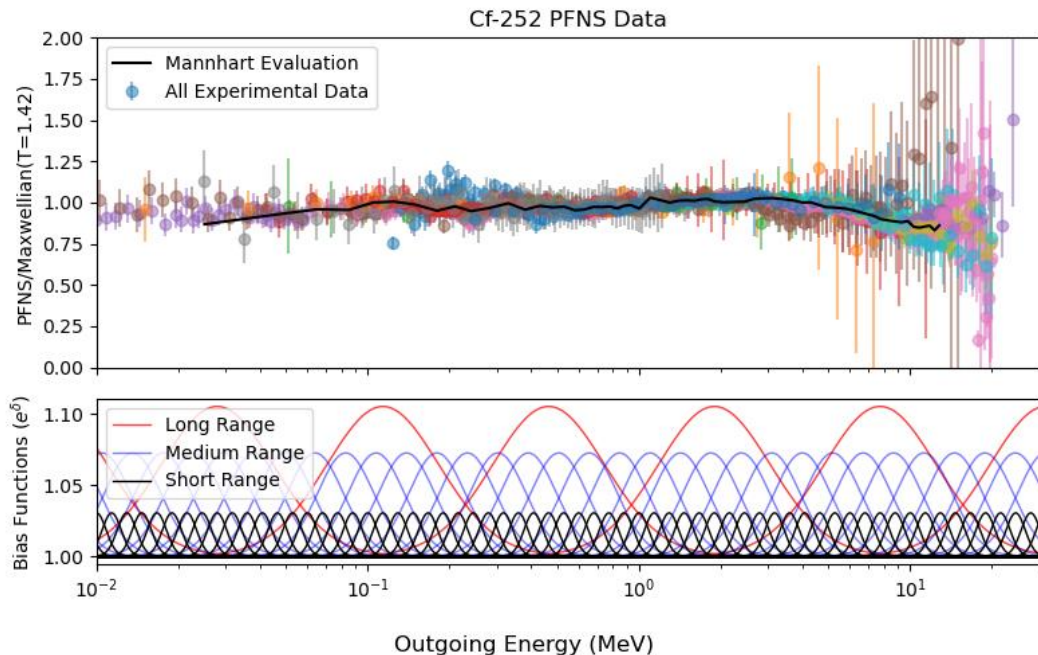
$$\mathbf{y} = \mathbf{D}\boldsymbol{\sigma} \cdot \mathbf{e}^{\boldsymbol{\delta}} + \boldsymbol{\varepsilon}$$

$\boldsymbol{\delta} = \mathbf{B}\boldsymbol{\gamma} = \text{relative bias}$

$\mathbf{B} = \text{bias basis matrix}$

$\boldsymbol{\gamma} = \text{bias coefficients}$

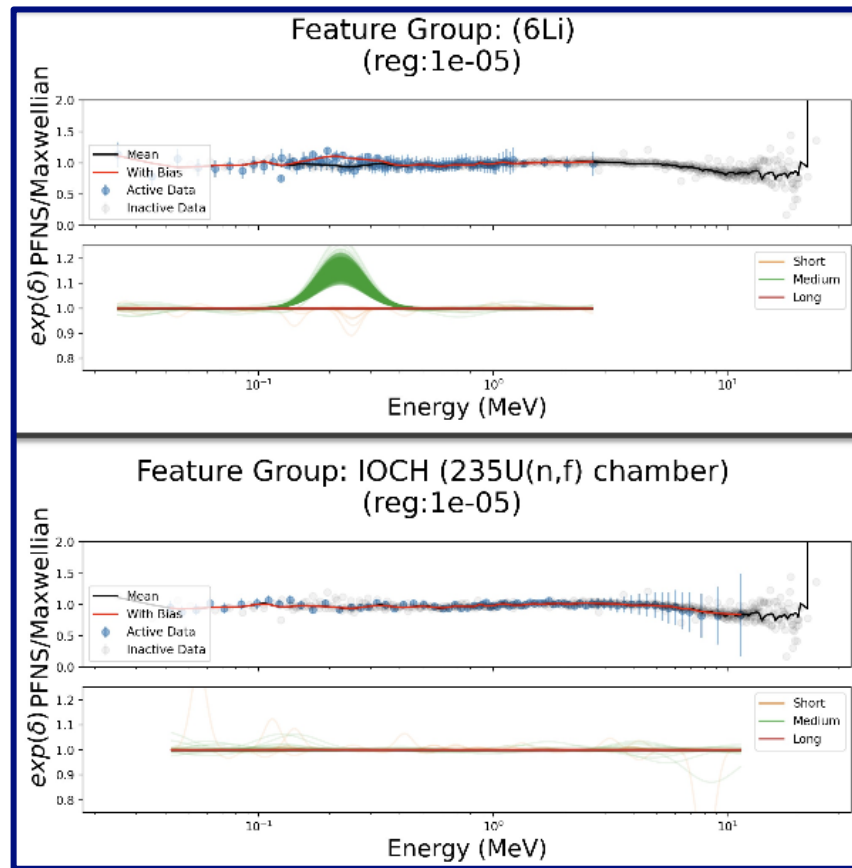
$\cdot = \text{element-wise product}$



Experimental metadata guide active/inactive groups

- Experimental features of active data give indicate possible physics root cause of the bias
- Sparsity shrinks away most bias terms unless justified by the data

Identified 4 biases in data base, corresponding physics root causes were investigated further by experts & simulations



Sparse Bayesian modeling can be a powerful tool to learn from complex data

Opportunity

- Specific experimental bias example worked well
 - Open-source code base w/STAN PPL
- Sparse inference is generally applicable along the pipeline
 - Other experimental data types
 - Phenomenological physics parameters?

Needs for further adoption/application

- AI/ML-ready experimental database (features for bias)
- Largely enabled by advances in hierarchical modelling and PPLs
 - Solutions may require differentiable physics models
 - Ready made tools

Acknowledgements

Research reported in this presentation was in part supported by the U.S. Department of Energy LDRD program at Los Alamos National Laboratory under project number(s) and 20240878PRD4.

Research reported in this publication was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under the Nuclear Data InterAgency Working Group Research Program.

Could do more on **WHAT** the model is

- Feature groups
- Horseshoe prior on bias terms
- Math behind it

- AIACHNE example
 - What is the root cause (physics) of discrepancies between trustworthy experimental data?