UQ Developments at the IAEA Nuclear Data Section and Perspectives

Georg Schnabel

IAEA Nuclear Data Section

Nuclear Data for the Next Decade (Paris, 10 March 2026)

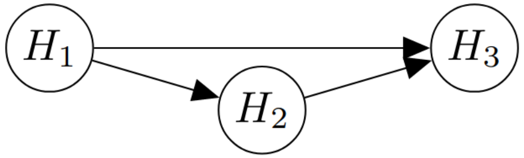
Outline

Developments towards making nuclear data evaluation easier, more flexible and more efficient

- Graphical models
- Differentiable programming
- Probabilistic programming

Bayesian networks

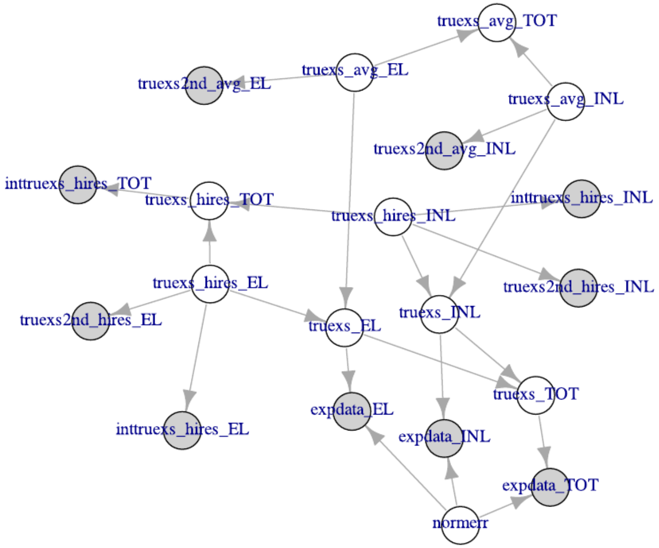
$$P(H_1, H_2, H_3) = P(H_1)P(H_2 \mid H_1)P(H_3 \mid H_1, H_2)$$



Thomas Bayes
(wrong picture)

Pierre-Simon Laplace

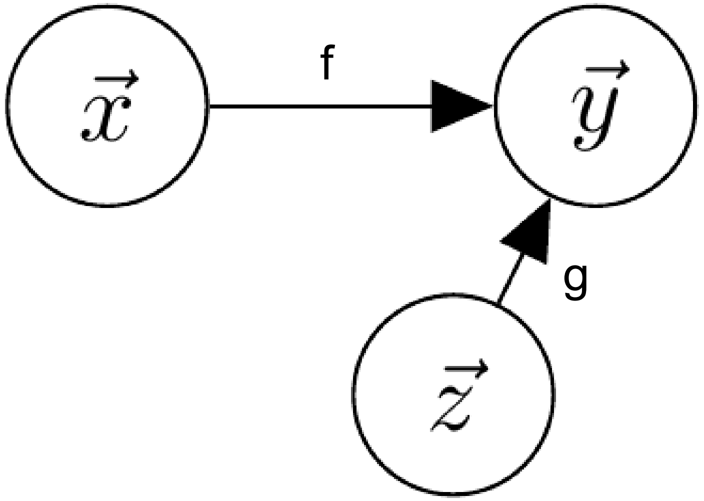
… build models by composing simple building blocks
… similar to how it is done for neural networks



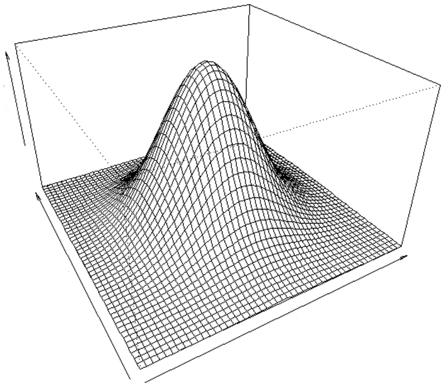
Judea Pearl*

* Better Than Bacon – Judea Pearl at NIPS 2013

Basic links



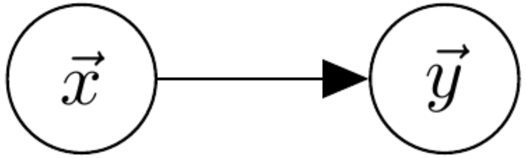
$$\vec{x} \sim \mathcal{N}(\vec{\mu}, \mathbf{\Sigma})$$



$$\vec{y} = f(\vec{x}) + g(\vec{z})$$

Because of $\boxed{\vec{y} = f(\vec{x})}$ the distribution of $\vec{y}$ is not necessarily multivariate normal

Examples of useful (linear) mappings

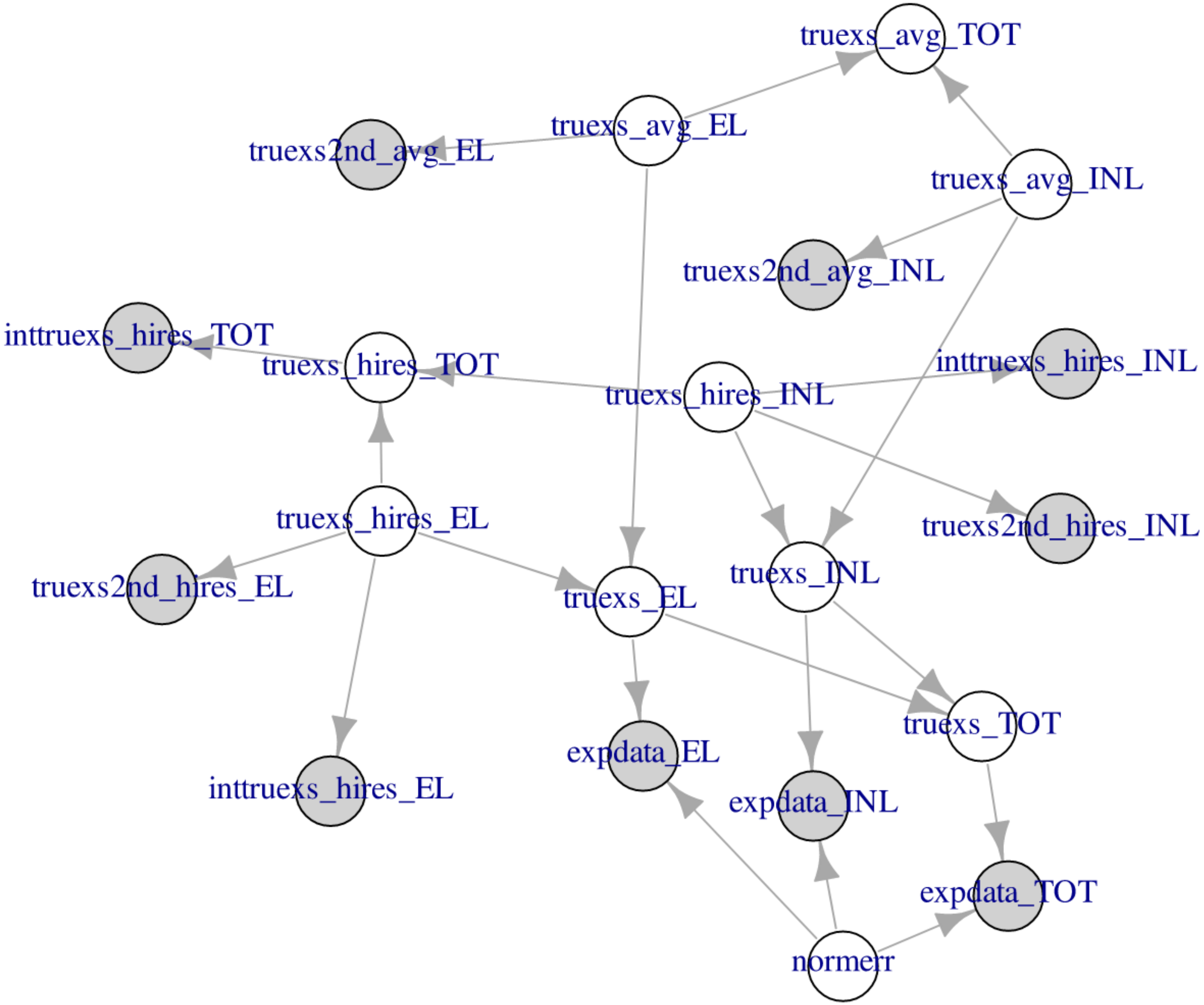


Linear interpolation
(e.g., model mesh to experimental mesh)

$$y_j = \left(\frac{E_{i+1} - E'_j}{E_{i+1} - E_i}\right) x_i + \left(\frac{E'_j - E_i}{E_{i+1} - E_i}\right) x_{i+1}$$
$$\text{if } \; E_i \leq E < E_{i+1}$$

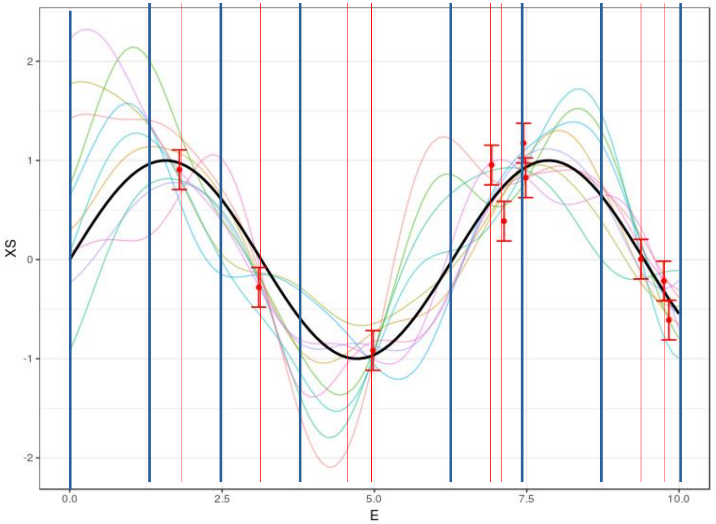
Convolution
(e.g., model mesh to experimental mesh with finite energy resolution)

No-model evaluation Fe-56 between 1 and 2 MeV (PoC)



G. Schnabel, R. Capote, A. Koning, D.A. Brown, “Nuclear Data Evaluation with Bayesian Networks,” arXiv:2110.10322 (2021)

Sparse GP construction



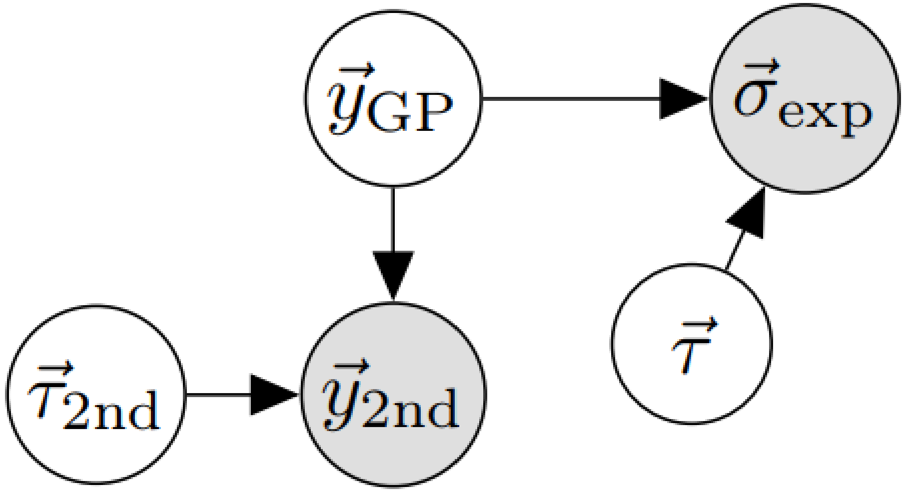
$$\Delta_i = \frac{y_{i+1} - y_i}{E_{i+1} - E_i}$$

$$\Delta_i^2 = \frac{\Delta_{i+1} - \Delta_i}{E_{i+1} - E_i}$$

Pseudo-observations (y_{2nd})

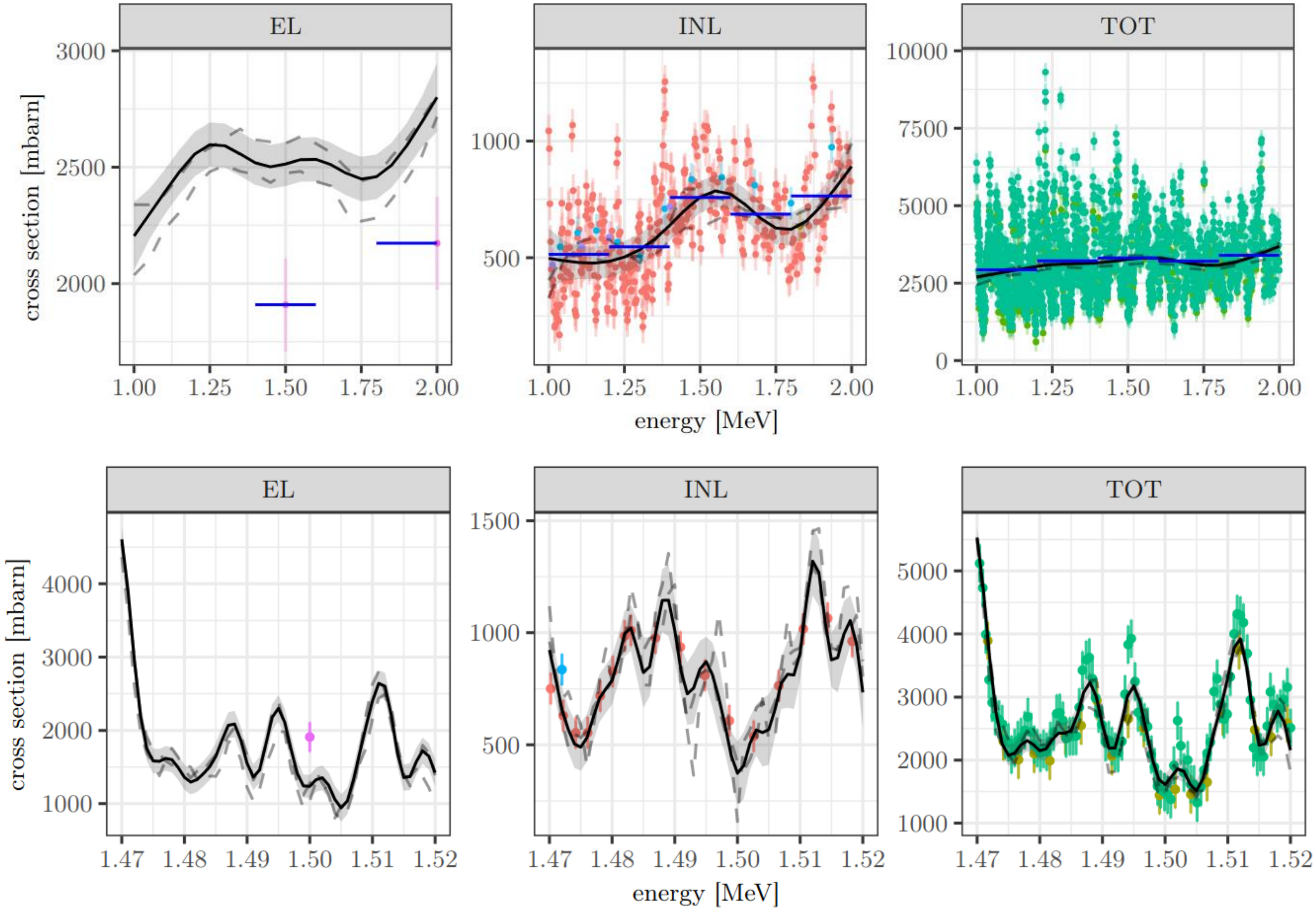
$$\Delta_i^2 \sim \mathcal{N}(0, \delta_i^2)$$

$$\mathrm{cov}(\Delta_i^2, \Delta_j^2) = 0 \text{ for } i \neq j$$

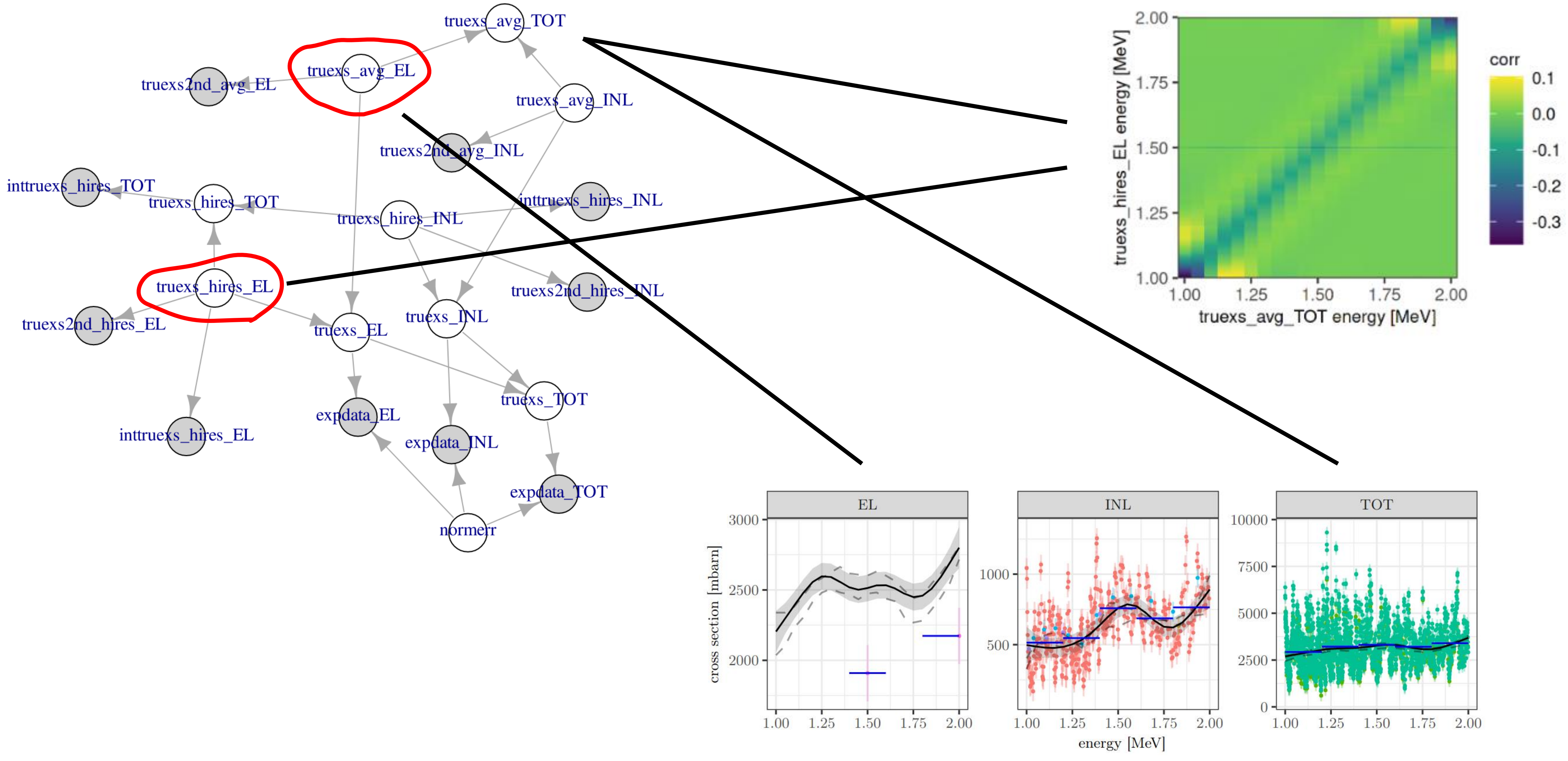


Related and more sophisticated (thanks to Jeremias Knoblauch):
F. Lindgren et al, "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach," Journal of the Royal Statistical Society (2011)

No-model evaluation Fe-56 between 1 and 2 MeV (PoC)

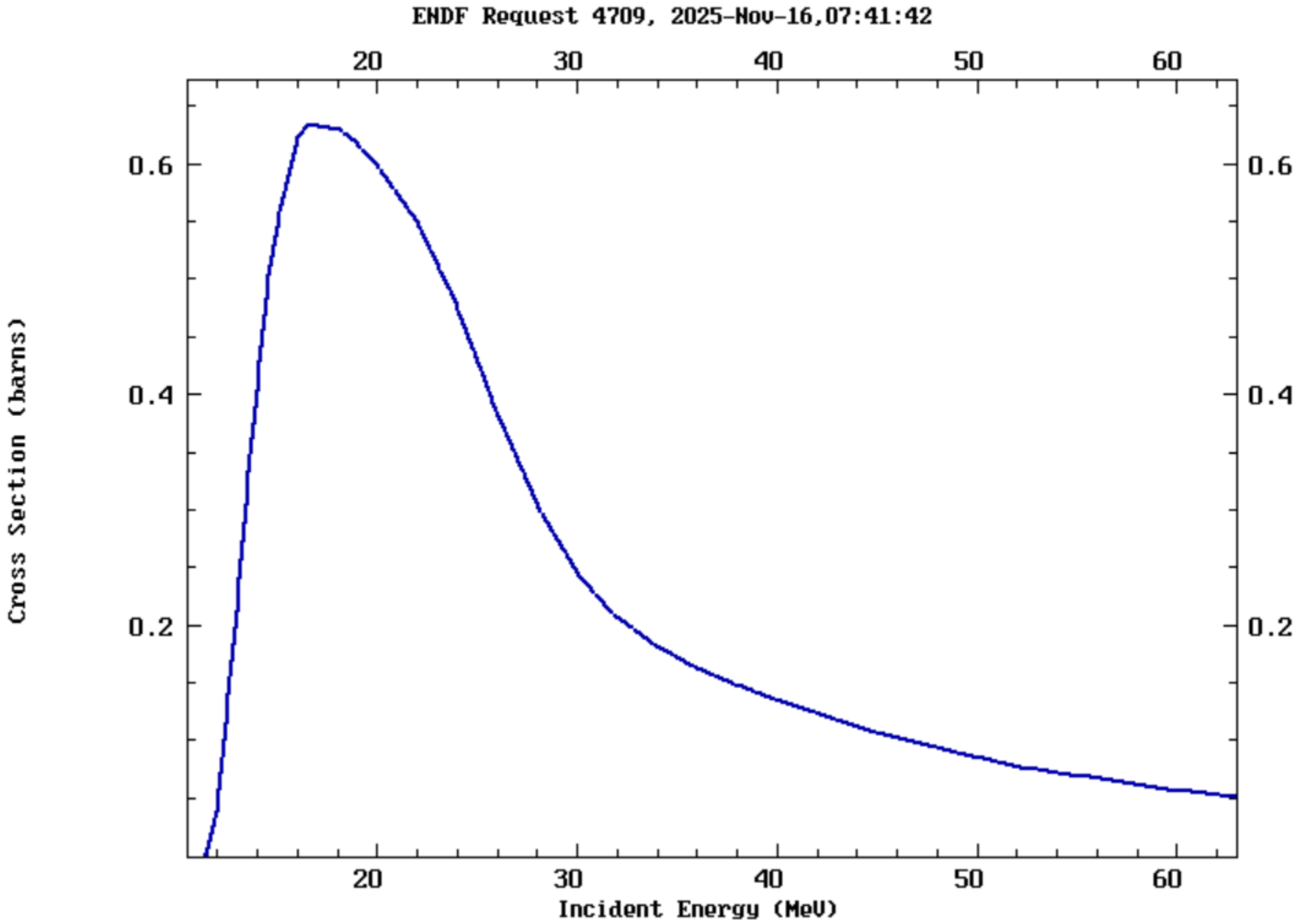


No-model evaluation Fe-56 between 1 and 2 MeV (PoC)



Towards enriched modeling possibilities

Fe-56(n,2n) cross section



Monotonicity constraint

Mapping to first derivative

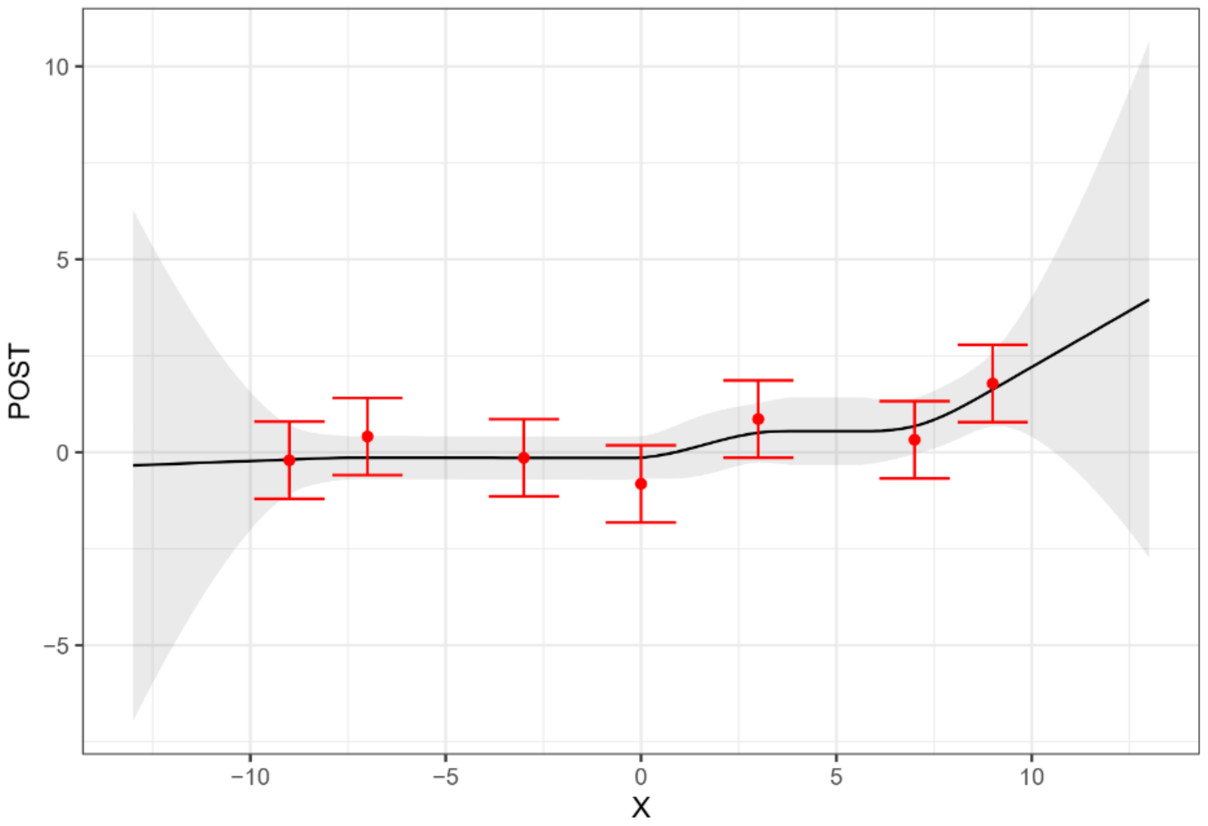
$$\Delta_i = \frac{y_{i+1} - y_i}{E_{i+1} - E_i}$$

Truncation

$$z_i = \min(0, \Delta_i)$$

Pseudo-observation

$$z_i \sim \mathcal{N}(0, \delta_i) \text{ with } \delta_i \ll 1$$



Quadratic penalty method

Decrease variance of pseudo-observation(s) δ_i

Solve with Levenberg-Marquardt algorithm

$$\mathbf{\Sigma}_{\text{mesh}} = \left(\mathbf{S}^T \mathbf{\Sigma}_{\text{exp}}^{-1} \mathbf{S} + \lambda \mathbf{D}\right)^{-1}$$

$$\vec{\sigma}_{\text{mesh}} = \vec{\sigma}_{\text{ref}} + \mathbf{\Sigma}_{\text{mesh}} \mathbf{S}^T \mathbf{\Sigma}_{\text{exp}}^{-1} \left(\vec{\sigma}_{\text{exp}} - \vec{\sigma}_{\text{pred}}\right)$$

Adjust λ according to improvement

LM in Nuclear Data Context:
P. Helgesson, H. Sjöstrand, "Fitting a defect non-linear model with or without prior, distinguishing nuclear reaction products as an example," Rev. Sci. Inst. (2017)

Quadratic programming

Find $\mathbf{x}$ to

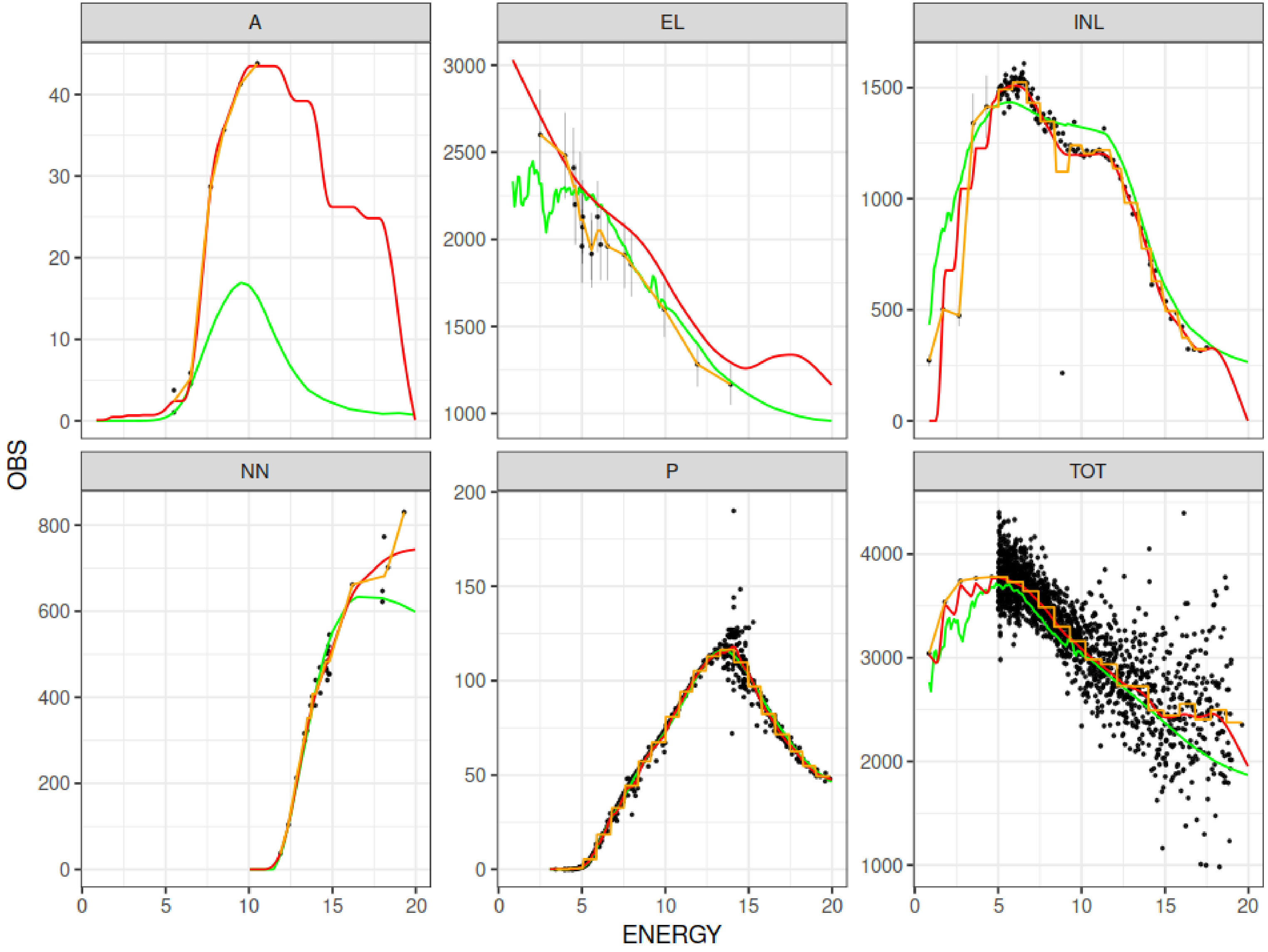
minimize $\frac{1}{2}\mathbf{x}^{\mathrm{T}}Q\mathbf{x} + \mathbf{c}^{\mathrm{T}}\mathbf{x}$

subject to $A\mathbf{x} \preceq \mathbf{b},$

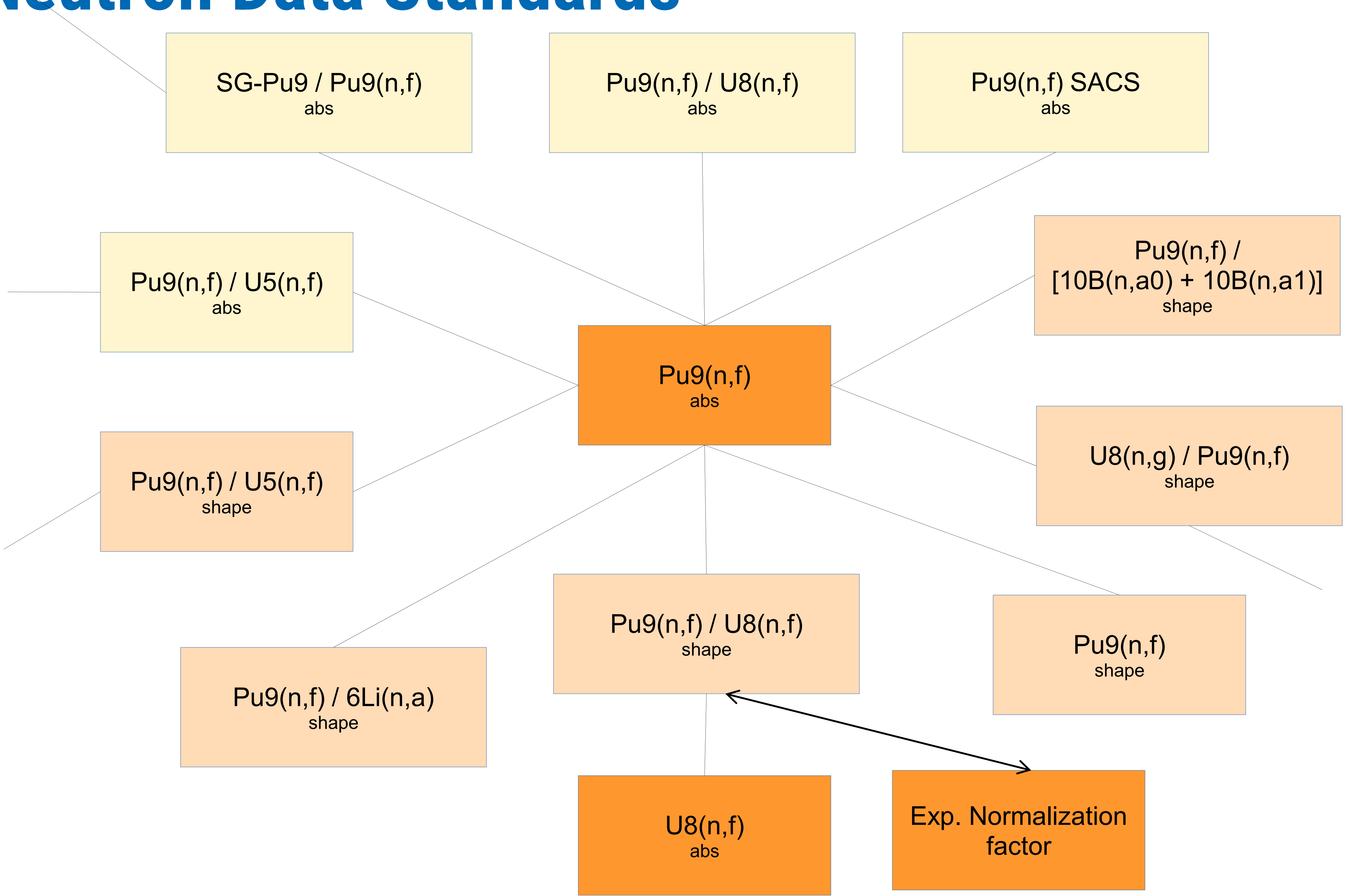
Goldfarb and Idnani, “A numerically stable dual method for solving strictly convex quadratic programs,” Mathematical Programming 27 (1983)

B. Stellato, G. Banjac et al, “OSQP: an operator splitting solver for quadratic programs,” Mathematical Programming Computation 12 (2020)

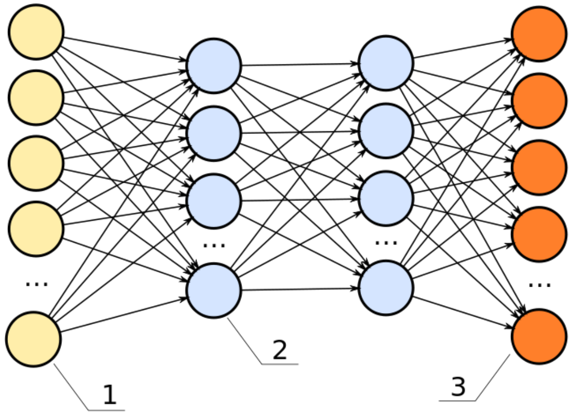
Testing Goldfarb & Idnani on Fe-56



Neutron Data Standards



Differentiable programming with Autodiff frameworks



TensorFlow

TORCH

theano

JAX

R. Kunitomo, G. Schnabel, “Exploring TensorFlow for Nuclear Data Evaluation,” IAEA-NDS-0245 (2024)

Different estimation approaches

TensorFlow + **TensorFlow Probability**

$$\log p(\vec{x} \,|\, \vec{d}) = -\frac{1}{2}\left(n \log(2\pi) + \log\det\Sigma + \chi^2(\vec{x})\right)$$

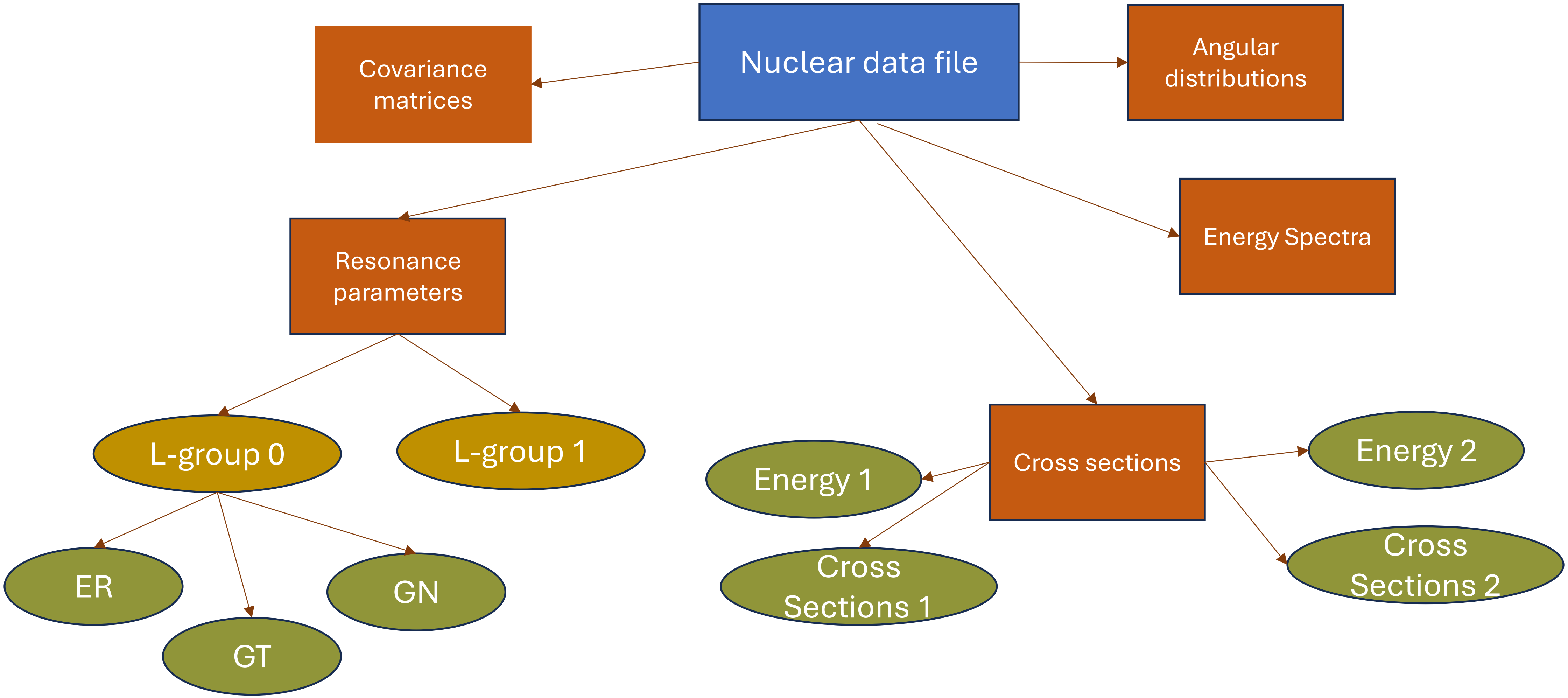
$$\chi^2(\vec{x}) = \left(\vec{d} - f(\vec{x})\right)^T \Sigma^{-1} \left(\vec{d} - f(\vec{x})\right)$$

$$\Sigma = \left[(f(\vec{x})^T f(x)\right] \odot \Sigma_{\mathrm{rel}}$$

Maximum Likelihood / Maximum A-Posteriori: Find x to maximize log p(x|d))
ChiSquare minimization: Find x to minimize χ^2
Generalized Least Squares: Iteratively apply GLS equation using Taylor expansion
Bayesian inference: Get sample from posterior distribution by MCMC (e.g. Hamiltonian Monte Carlo)

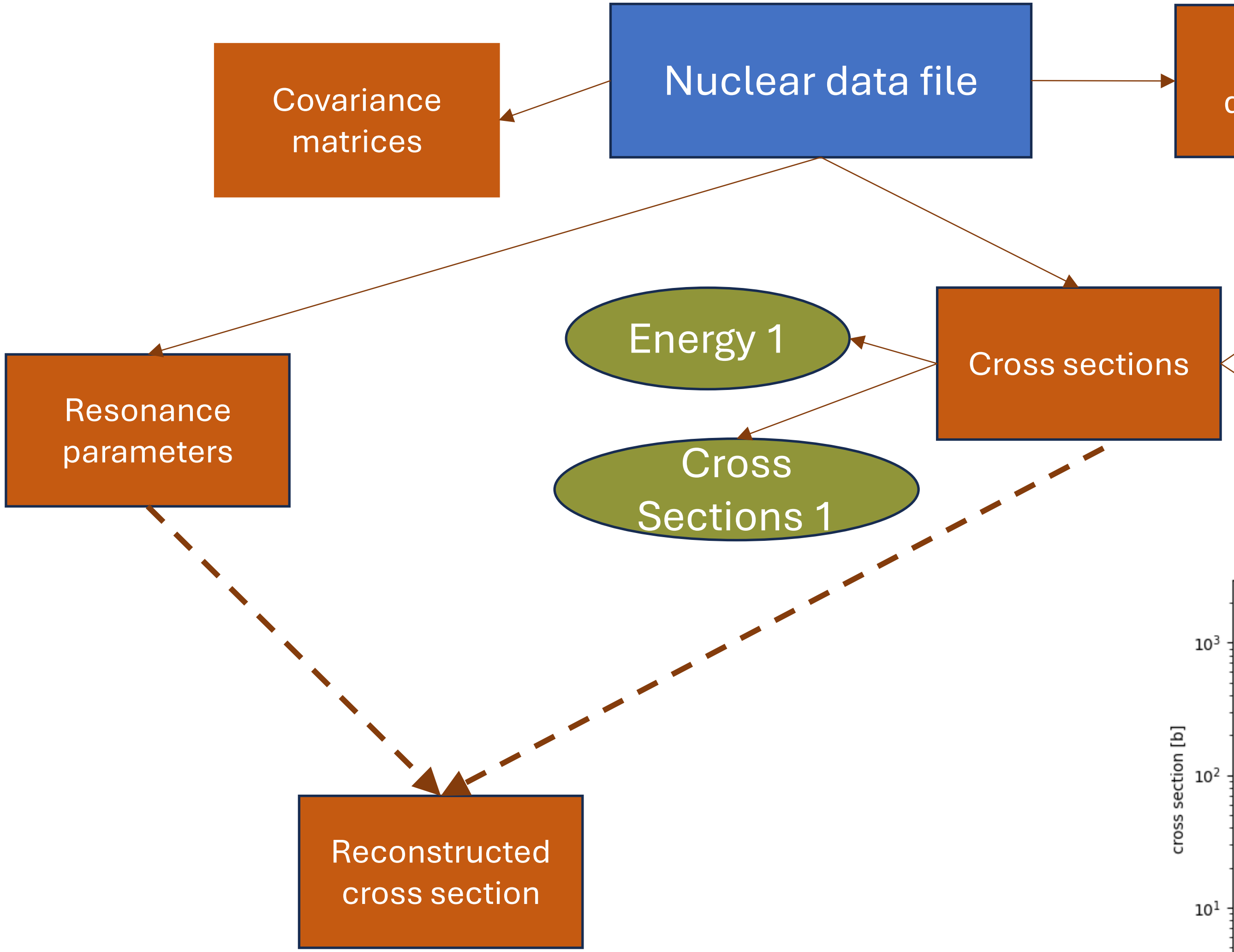
github.com/iaea-nds/gmapy (TF stuff on dev branch)

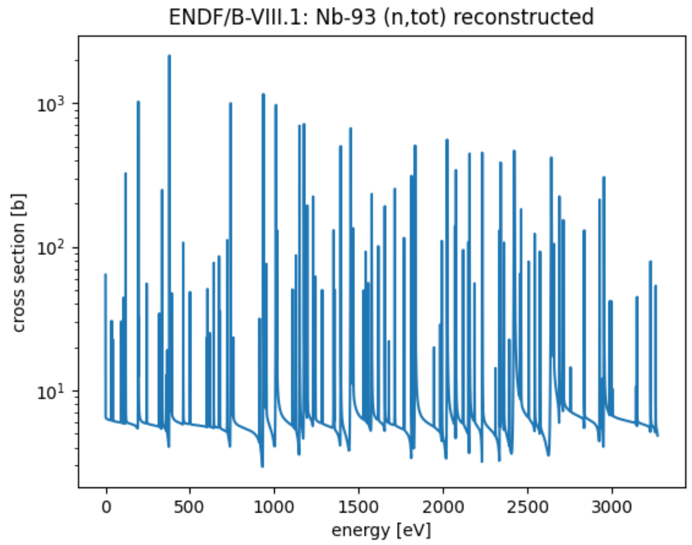
From files to knowledge graphs



G. Schnabel, D. L. Aldama, R. Capote, "How to explain ENDF-6 to computers: A formal ENDF format description language," 2312.08249v1 (2023)

From files to knowledge graphs





>60k E_{incs} with 140 resonances in about 10 s

All dxs/dE_{res} in about 15 s

Work by Daniel Lopez Aldama (IAEA Consultant)

Summary

- Graphical models are versatile (bottom-up modelling)
- Scalable Sparse GP constructions (GMRF <-> GP)
- Positivity, monotonicity constraints -> Quadratic programming
- Differentiable and Probabilistic programming (TensorFlow, TF Probability, JAX)
- From files to graphs (and differentiable transformations on graph nodes)

Some ideas for the next decade (IT infrastructure)

- Global nuclear data knowledge graph
- A reusable global library of differentiable transformations acting on the knowledge graph
- Global sensitivity studies and Bayesian inference (or any other kind of inference) connecting institutions and nuclear subdomains enabled by differentiable and probabilistic programming

Some ideas for the next decade (Statistics)

- More elaborate statistical models (Mixture models, hierarchical models in general)
- Sparsity-inducing priors (Laplace, spike-and-slab, horseshoe, etc.)
- Robust statistics (e.g. based on multivariate t-Student distribution, quantile regression)
- Non-parametric approaches beyond GP (e.g. tree-based models, BART)
- Further development of model compression techniques
- More systematic application of available inference algorithms (Approximate Bayesian computation (ABC), Pseudo-Marginal Likelihood, Variational Inference, MCMC sampling strategies (Hamiltonian Monte Carlo, NUTS, etc.))
- Prior-Data Fitted Networks (PFNs)

(Examples for the pertinent application of many of these approaches in nuclear data work exist, e.g. from LANL, LLNL, BAND, etc. and it would be wonderful if they become more mainstream)