IAEA 60 Years
Atoms for Peace and Development

Nuclear data evaluation with Bayesian networks

Georg Schnabel
g.schnabel [at] iaea.org

Nuclear Data Section
Division of Physical and Chemical Sciences NAPC
Department for Nuclear Sciences and Applications
IAEA, Vienna

New York Scientific Data Summit
29 October 2021

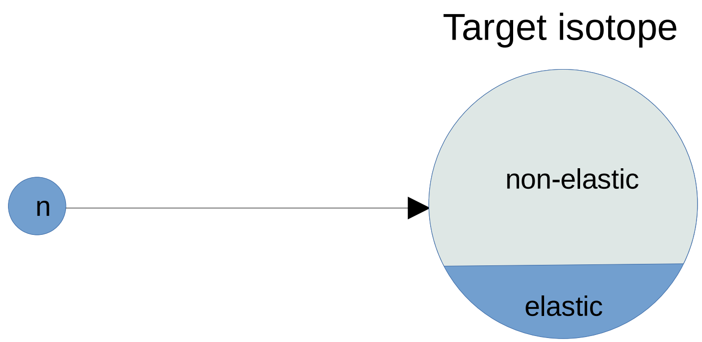
Short bio

- Studied physics at TU Vienna
- PhD in nuclear data evaluation 2015
- Postdoc at CEA Saclay (2015-2018) and Uppsala University (2018-2019)
- Since 2020 nuclear physicist in Nuclear Data Section at IAEA dealing with nuclear data library projects and code development

Nuclear data

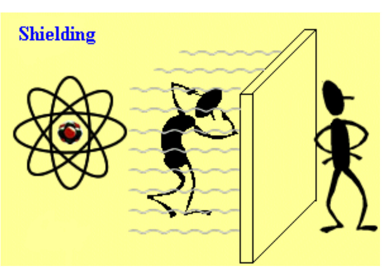
PSI Gantry 2 facility

Probabilities of various nuclear interactions involving the atomic nuclei, e.g., cross sections.

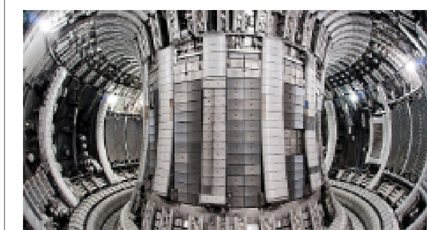


Relevant for:

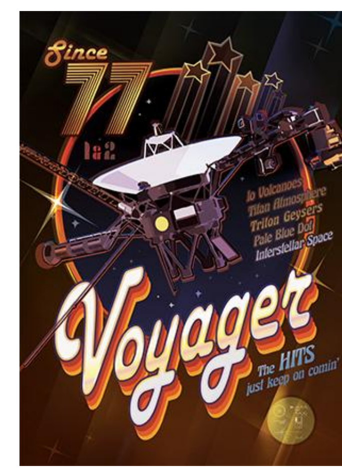
- Reactor physics
- Radiation dosimetry
- Radiation protection
- Radioactive waste management
- Astrophysics
- Nuclear medicine
- Fusion research
- ...



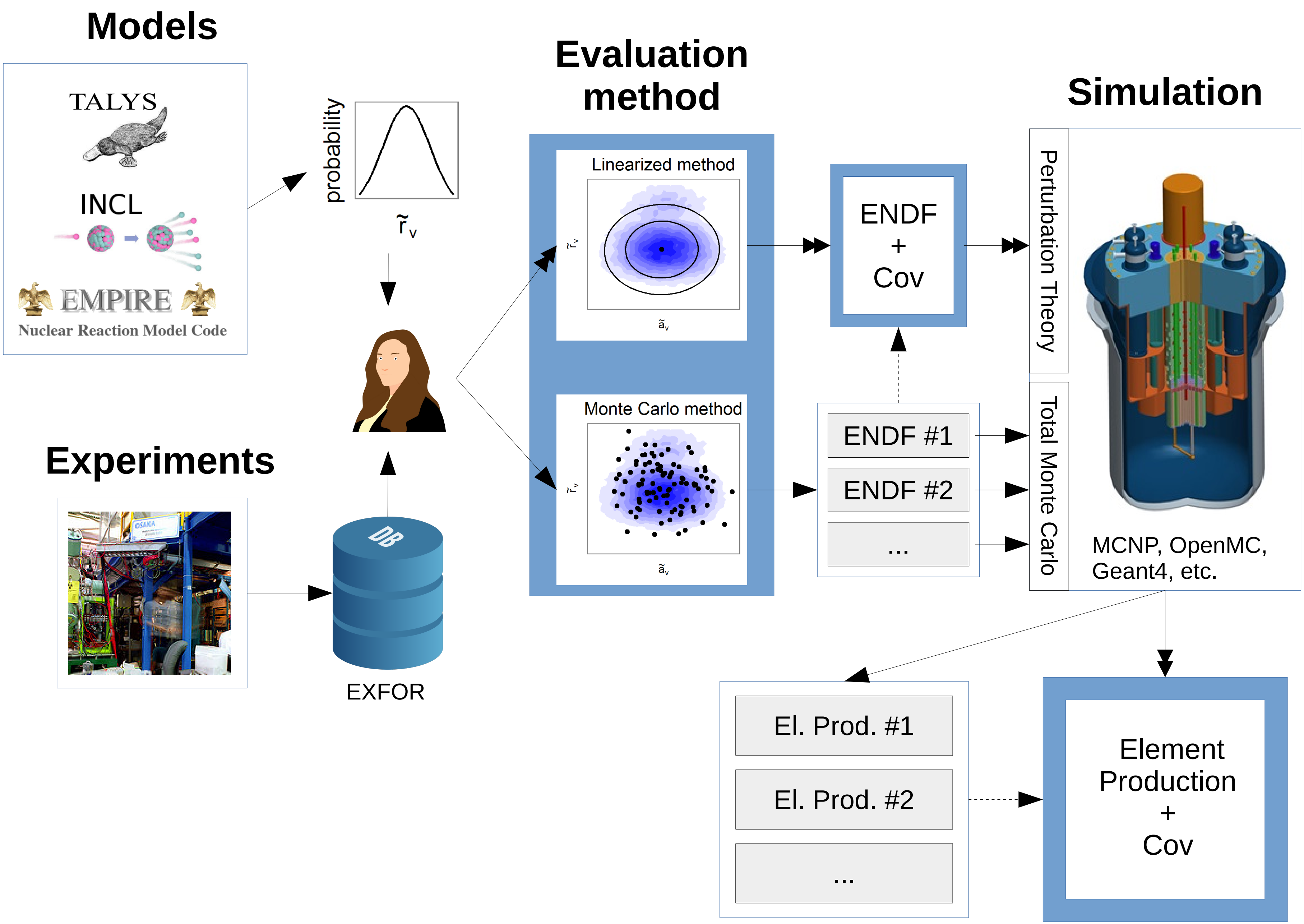
Palisades Nuclear Generating Stations



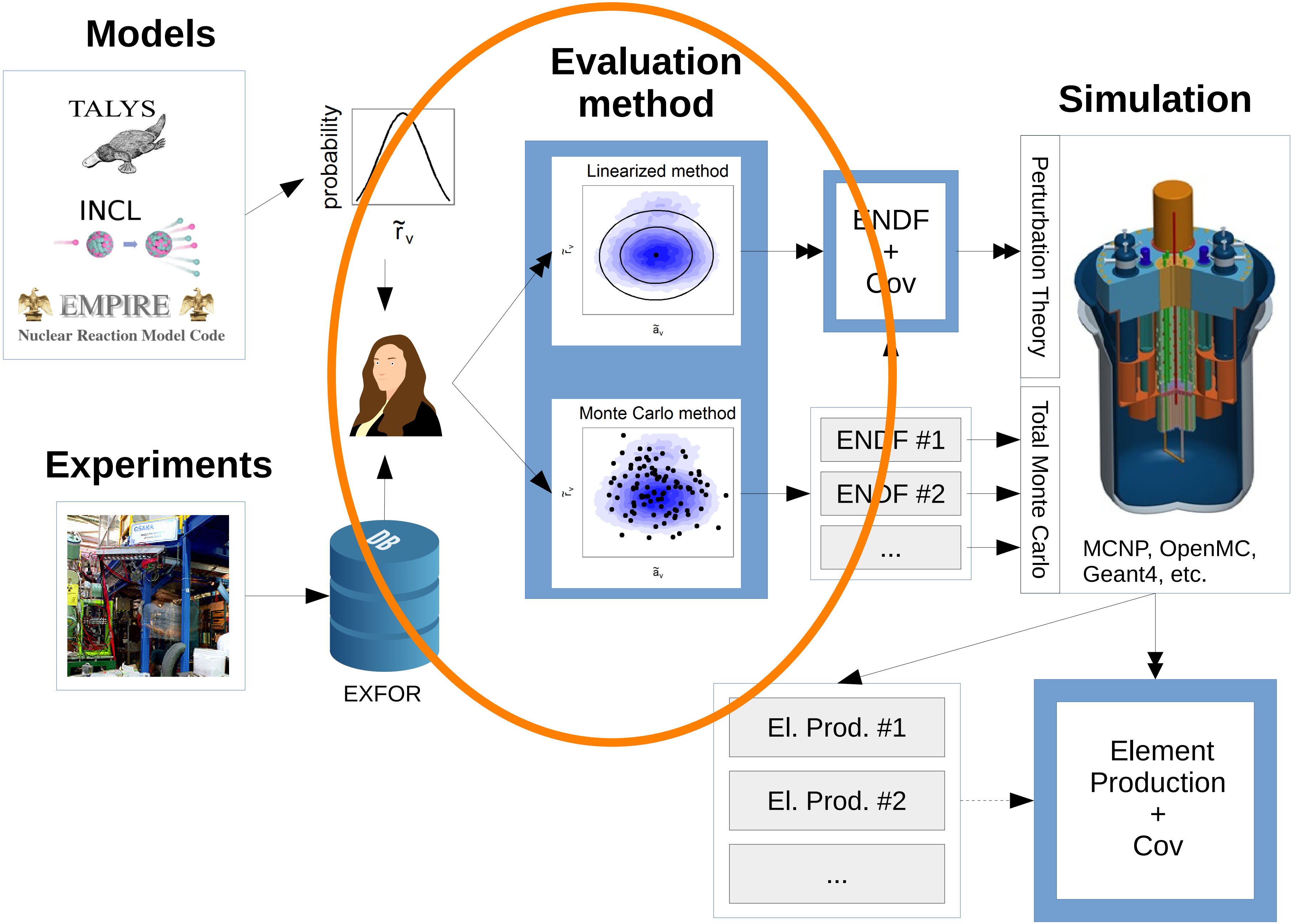
Joint European Torus



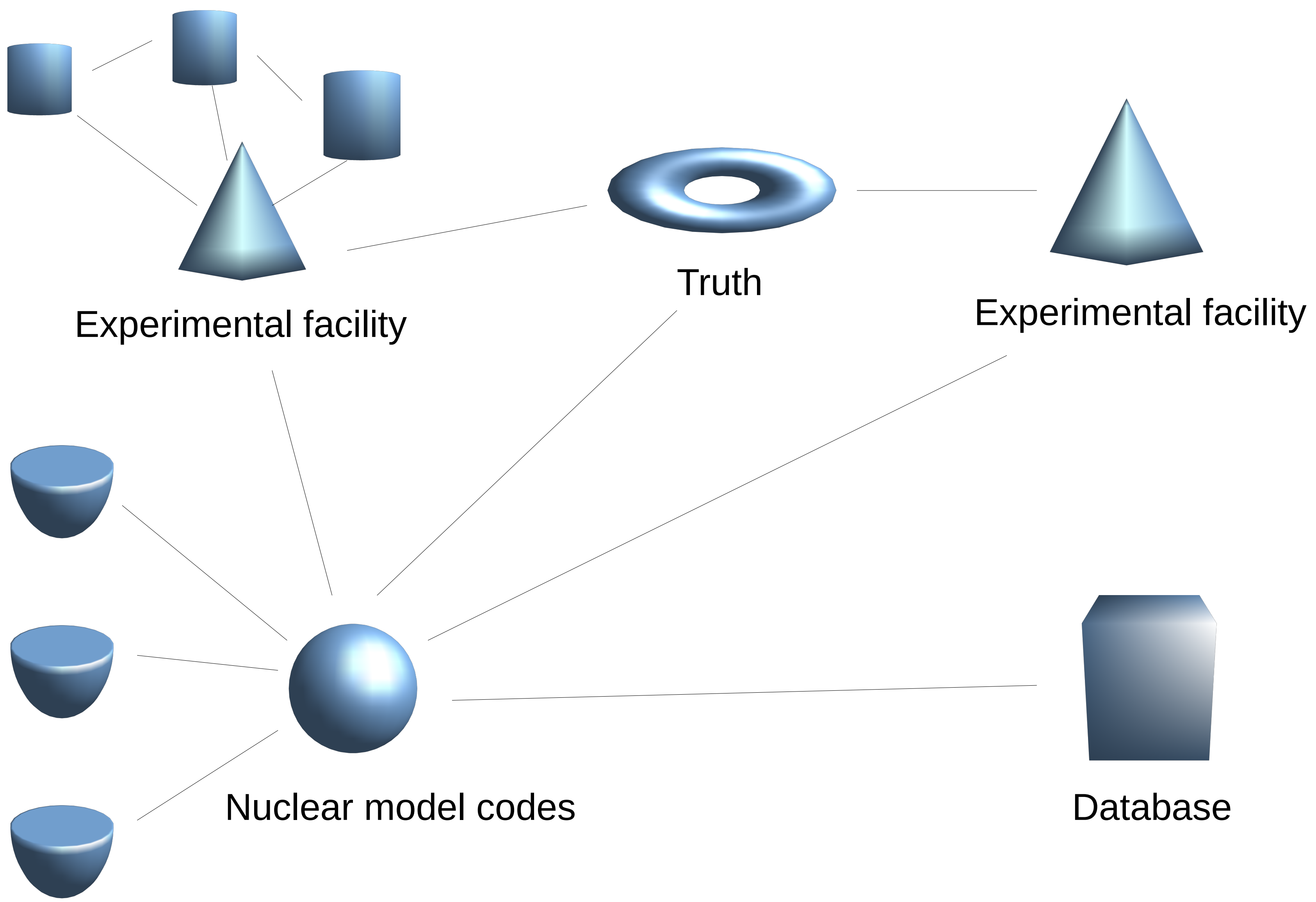
Nuclear data evaluation



Nuclear data evaluation

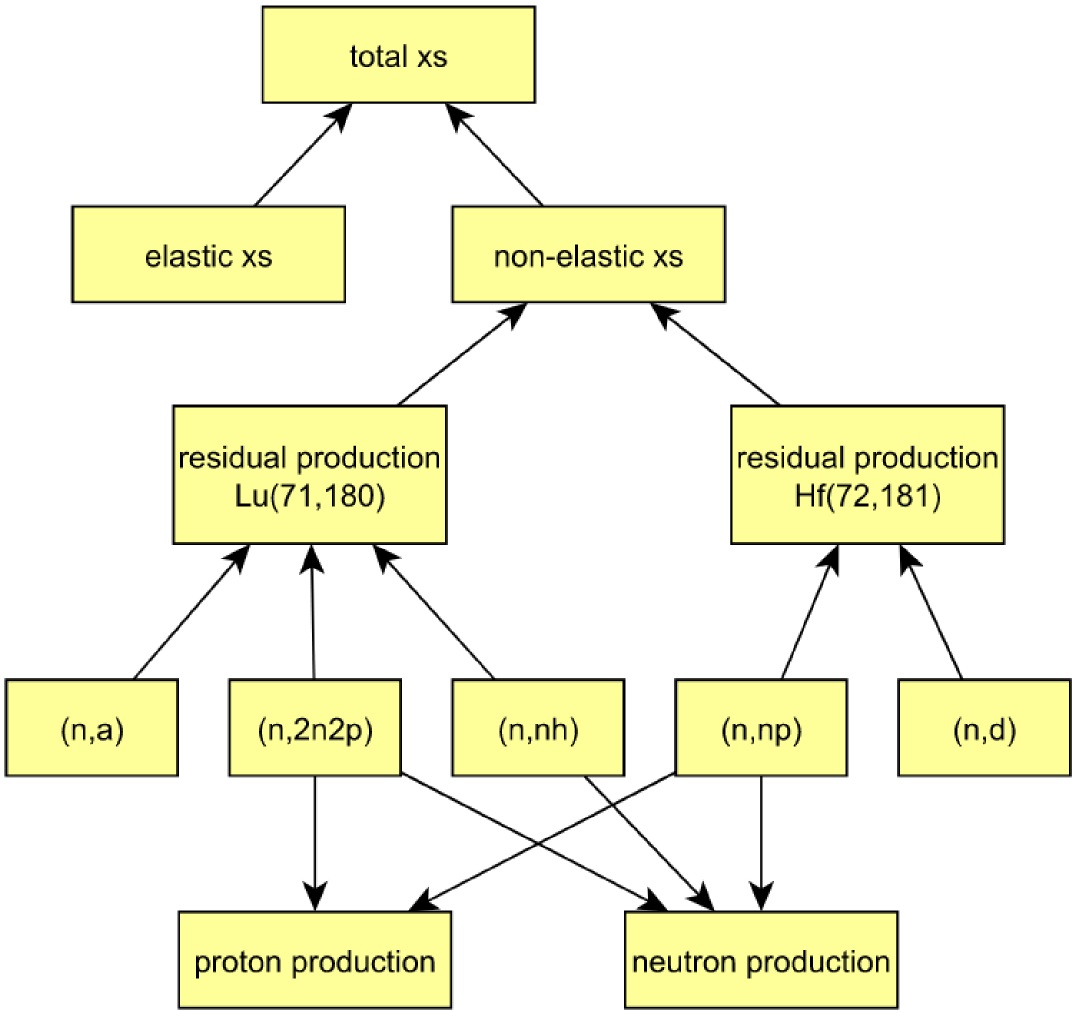


Another perspective on the nuclear data evaluation process

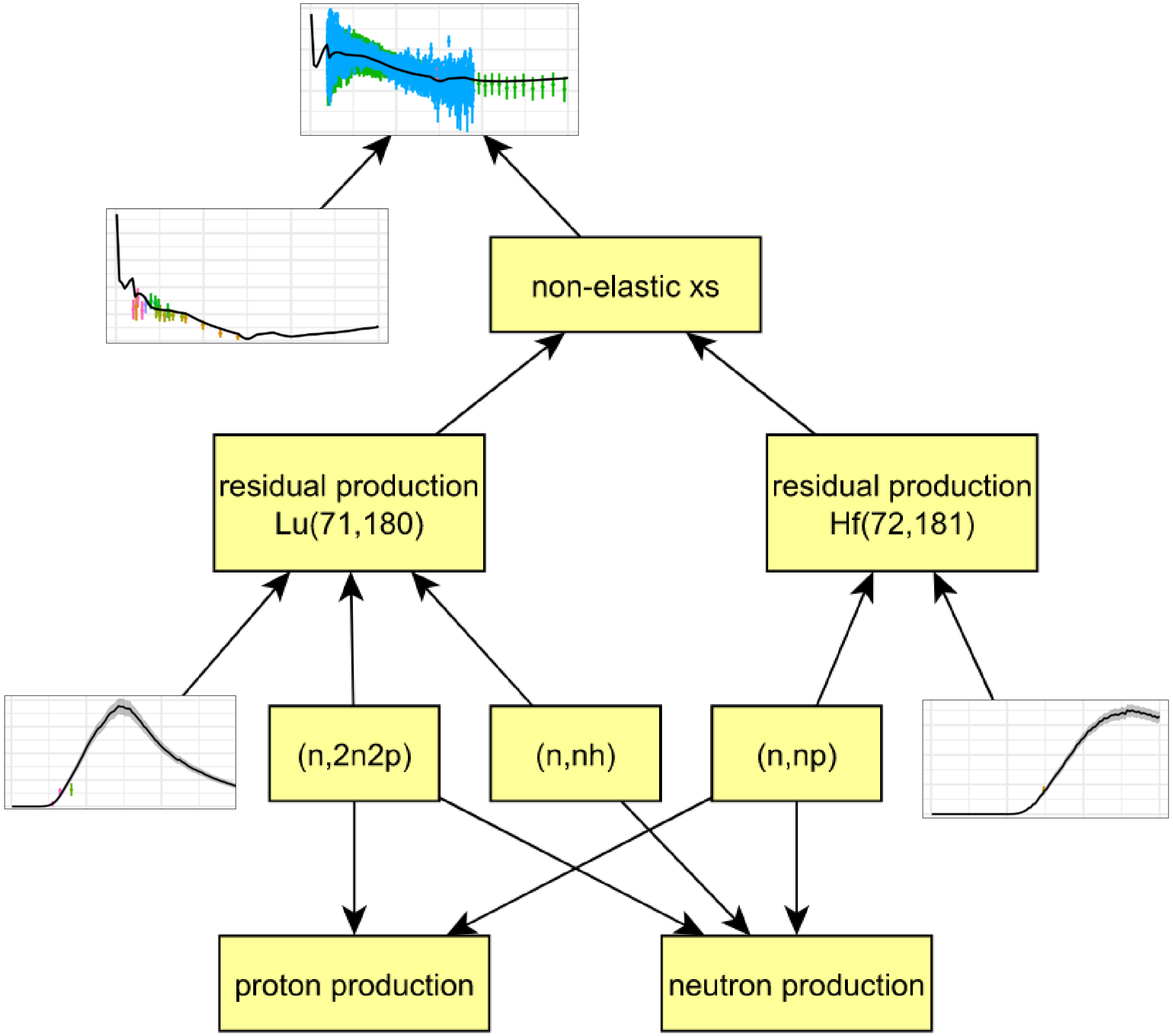


Keywords: meta-analysis, sensor fusion, digital twins

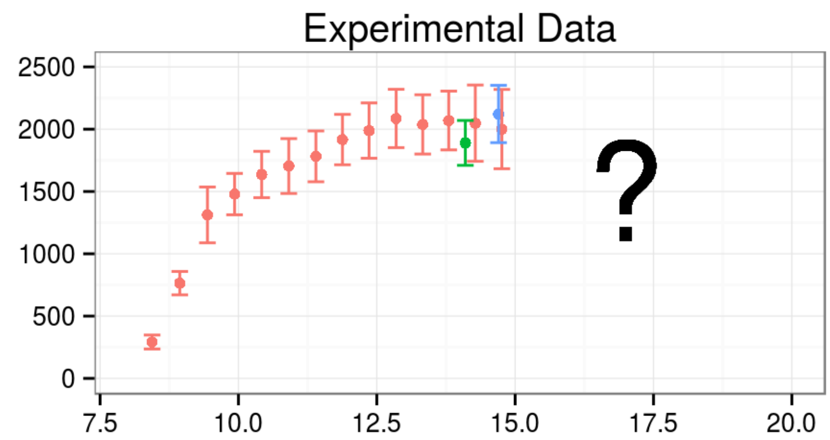
“Truth” - System of reactions

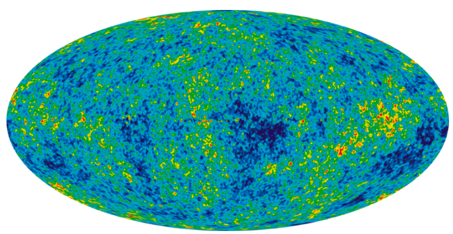


"Truth" - System of reactions



Experimental data



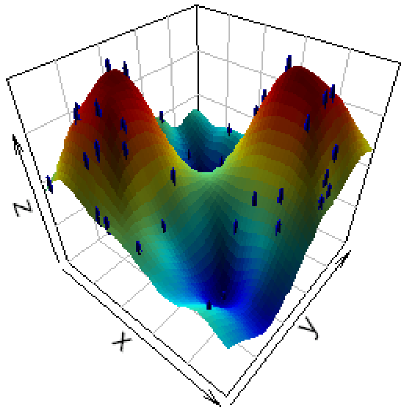


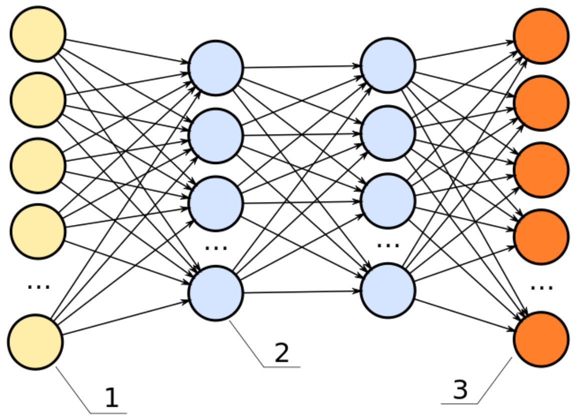
background noise

Sample*
(thickness, density, impurities)

detector

Bayesian statistics vs neural networks





Bayesian statistics …

… allows inference in sophisticated probabilistic models

… inference is a computational challenge (e.g., MCMC)

Neural networks …

… scale to huge datasets

… are not that easily amenable to UQ

… are composed of simple building blocks

Best of both worlds*

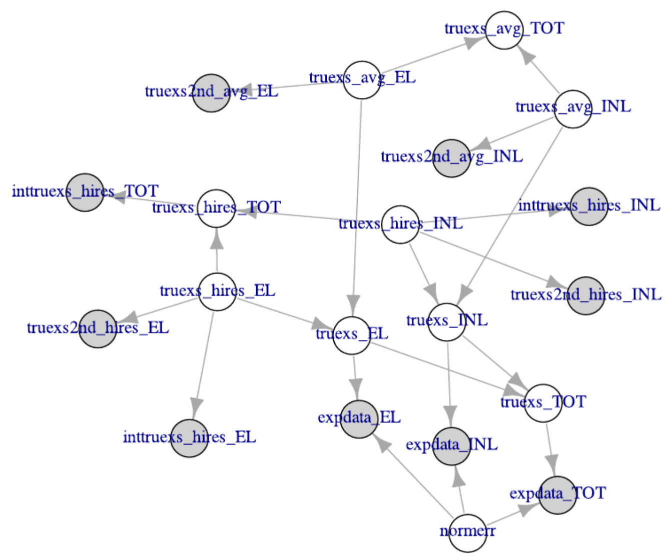
Bayesian networks …

… use Bayesian inference

… build models by composing simple building blocks

… similar to how it is done for neural networks

Thomas Bayes



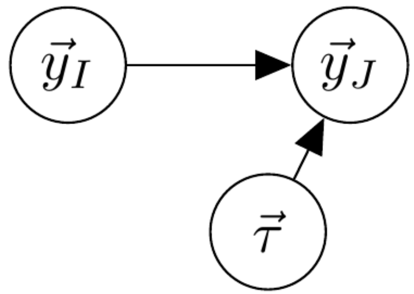
Judea Pearl**

Pierre-Simon Laplace

* at least for nuclear data evaluation

** Better Than Bacon – Judea Pearl at NIPS 2013

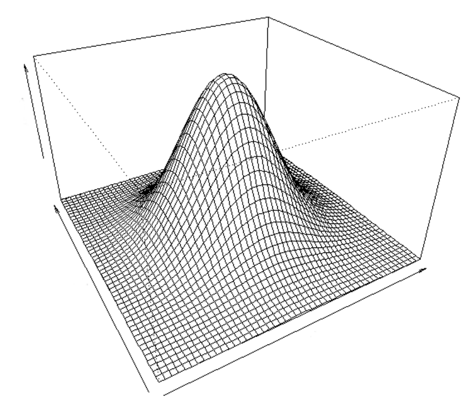
Basic building block



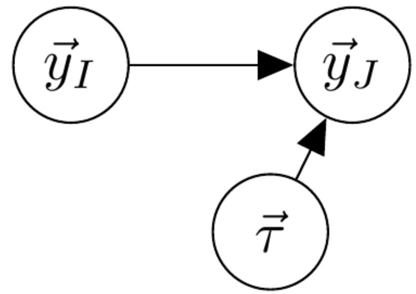
$$\vec{y}_J = \vec{y}_{\mathrm{ref},J} + \mathbf{T}\,(\vec{y}_I - \vec{y}_{\mathrm{ref},I}) + (\vec{\tau} - \vec{\tau}_{\mathrm{ref}})$$

$$\vec{y}_I \sim \mathcal{N}(\vec{u}_I, \mathbf{U}_{I,I})$$

$$\vec{\tau} \sim \mathcal{N}(\vec{u}_J, \mathbf{U}_{J,J})$$



Versatile building block



Gaussian process

Spline

Linear interpolation

$$\vec{y}_J = \vec{y}_{\mathrm{ref},J} + \mathbf{T}(\vec{y}_I - \vec{y}_{\mathrm{ref},I}) + (\vec{\tau} - \vec{\tau}_{\mathrm{ref}})$$

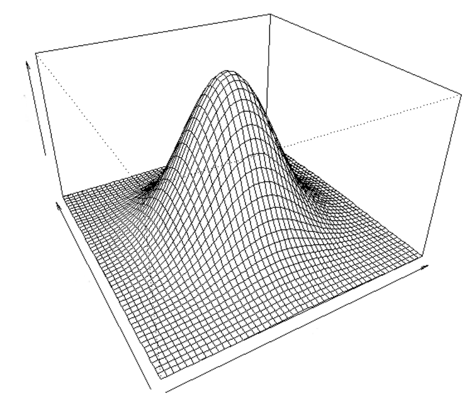
Fourier

Convolution

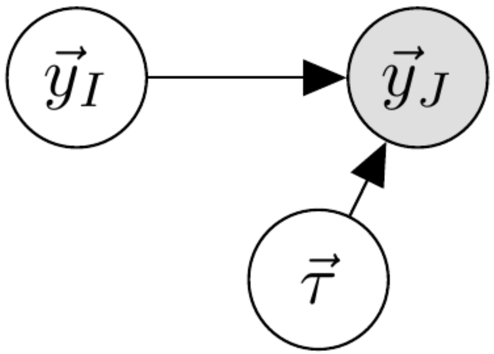
Linearized nuclear model

$$\vec{y}_I \sim \mathcal{N}(\vec{u}_I, \mathbf{U}_{I,I})$$

$$\vec{\tau} \sim \mathcal{N}(\vec{u}_J, \mathbf{U}_{J,J})$$



Bayesian inference



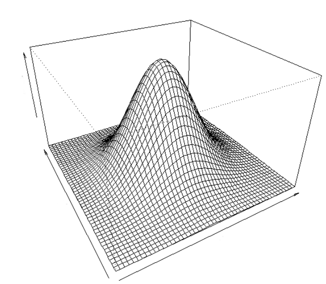
Posterior

$$\vec{y}_I \sim \mathcal{N}(\vec{u}'_I, \mathbf{U}'_{I,I})$$

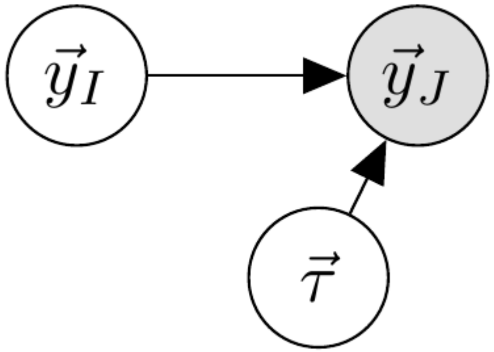
Analytic update equations

$$\vec{u}'_I = \vec{y}_{\mathrm{ref},I} + \mathbf{U}'_{I,I}\Big(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}(\vec{r} - \vec{y}_{\mathrm{ref},J}) + \mathbf{U}^{-1}_{I,I}(\vec{u}_I - \vec{y}_{\mathrm{ref},I})\Big)$$

$$\mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}\mathbf{S}_{J,I} + \mathbf{U}^{-1}_{I,I}\right)^{-1}$$



Bayesian inference



Posterior

$$\vec{y}_I \sim \mathcal{N}(\vec{u}'_I, \mathbf{U}'_{I,I})$$

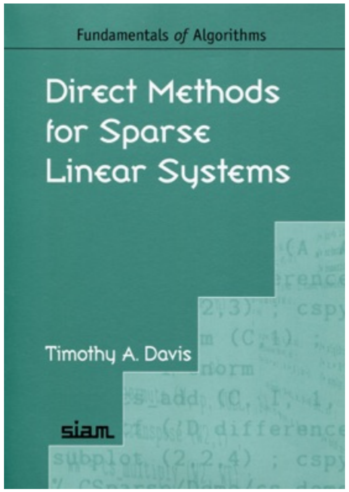
Analytic update equations

(aka Generalized Least Squares (GLS))

$$\vec{u}'_I = \vec{y}_{\mathrm{ref},I} + \mathbf{U}'_{I,I}\Big(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}(\vec{r} - \vec{y}_{\mathrm{ref},J}) + \mathbf{U}^{-1}_{I,I}(\vec{u}_I - \vec{y}_{\mathrm{ref},I})\Big)$$

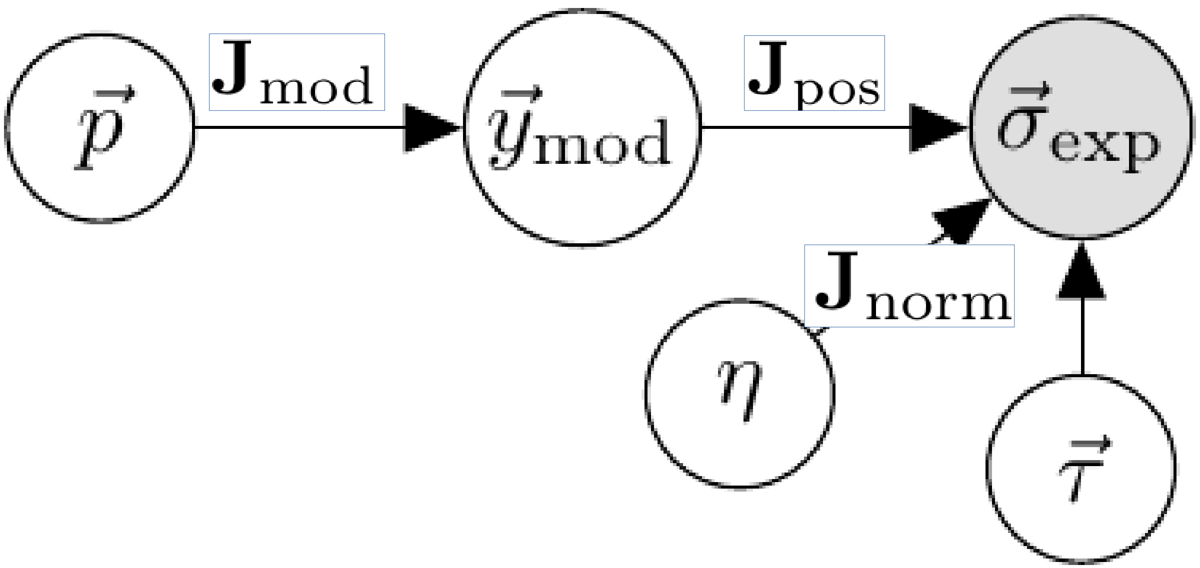
sparse

$$\mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}\mathbf{S}_{J,I} + \mathbf{U}^{-1}_{I,I}\right)^{-1}$$



SuiteSparse / CHOLMOD

Composability – nested relationships



apply chain rule to get a compound Jacobian matrix

$$\mathbf{S} = \begin{pmatrix} \mathbb{1} & \mathbf{0} & \mathbf{J}_{\mathrm{pos}}\mathbf{J}_{\mathrm{mod}} & \mathbf{J}_{\mathrm{norm}} \\ \mathbf{0} & \mathbb{1} & \mathbf{J}_{\mathrm{mod}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbb{1} \end{pmatrix}$$

can be done automatically: “automatic differentiation”

Framework flexible enough?

- Multivariate normal distribution
  - Negative values are regarded possible
  - Tails not heavy enough? Too symmetric?
- Linearity assumption
  - Nuclear physics models are non-linear
  - Many non-linear interactions between variables

→ Not flexible enough (yet)

Non-linear relationships

- Permit non-linear relationships between nodes
- Embed GLS method in an iterative scheme* to obtain Maximum A Posteriori (MAP) estimate:

$$\mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}\mathbf{S}_{J,I} + \mathbf{U}^{-1}_{I,I}\right)^{-1} \quad \Longrightarrow \quad \mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{A,I}\mathbf{U}^{-1}\mathbf{S}_{A,I} + \lambda\mathbf{D}\right)^{-1}$$

* aka (modified) Levenberg-Marquardt algorithm

Non-linear relationships

- Permit non-linear relationships between nodes
- Embed GLS method in an iterative scheme* to obtain Maximum A Posteriori (MAP) estimate:

Adaptive control parameter

$$\mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{J,I}\mathbf{U}^{-1}_{J,J}\mathbf{S}_{J,I} + \mathbf{U}^{-1}_{I,I}\right)^{-1} \quad \Rightarrow \quad \mathbf{U}'_{I,I} = \left(\mathbf{S}^T_{A,I}\mathbf{U}^{-1}\mathbf{S}_{A,I} + \lambda\mathbf{D}\right)^{-1}$$

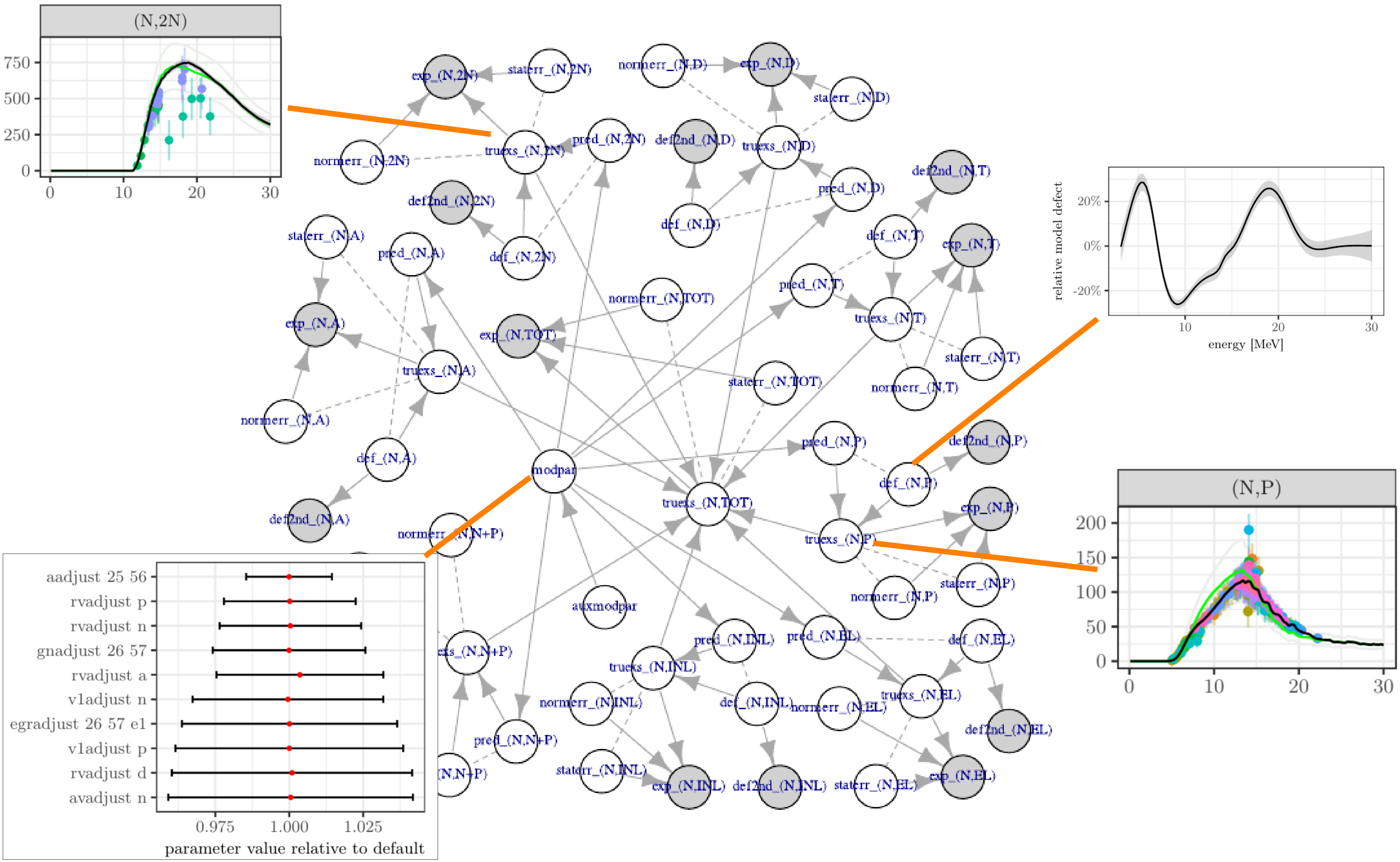
Enhanced modeling possibilities:

- Other distribution functions, e.g., log-normal distribution, via non-linear transformation

- Integration of more realistic relationships, e.g., non-linear physics model

* aka (modified) Levenberg-Marquardt algorithm

Nuclear data evaluation example



Summary

- Bayesian inference + network = Bayesian network
- Composability can be a great accelerator in the design of probabilistic models
- Simple distribution assumption (MVN) in combination with non-linear relationships yields a flexible yet tractable inference framework
- In the nuclear data evaluation context, we mostly deal with a system of functions linked by linear and non-linear relationships
- The future: link functions may be given by neural networks trained on lots of data if available
- Mathematical details and description of Bayesian network examples here:

G. Schnabel, R. Capote, A.J. Koning, D.A. Brown, “Nuclear data evaluation with Bayesian networks”, arXiv:2110.10322