

Gaussian Mixture Models (GMM) for data evaluations: - a phonon density of states (PDOS) case study

G. Arbanas, C. Chapman, J. McDonnell, I. Al-Qasir, K. Ramic, J. Brown, D. Wiarda, L. Leal,
U. Mertzyurek, W. Wieselquist

Nuclear Energy and Fuel Cycle Division
Oak Ridge National Laboratory (ORNL)

Cross Section Evaluation Working Group (CSEWG) Meeting, Nuclear Data Week
Brookhaven National Laboratory, Brookhaven, New York; January 6-9, 2026

(Gaussian) Mixture Model (GMM) as a Framework for evaluation of non-normal PDFs:

- GMM PDF is a weighted sum of normal PDFs → non-normal PDF
 - GMM parameters: weight, mean, and (co)variance of each normal PDF
 - 1) randomly select a Gaussian according to the weights, 2) sample from that Gaussian
 - Alternative MM employ other kinds of basis PDF: step, triangle, etc.
- GMM parameters may be made to obey normal PDF
 - Nevertheless, the GMM is still a non-normal PDF
- Enables use of General Linear Least Squares (GLLS) for such GMM
 - Including evaluation of covariances of *all* GMM parameters
- Apply it to Phonon density of states (PDOS), a non-normal PDF
 - compact evaluated covariance matrix for thermal neutron scattering

Phonon Density of States as a *scaled symmetric GMM*

- A scaling function $(2\beta)\sinh(\beta/2)$ ensures PDOS $\sim \beta^2$ as $\beta \rightarrow 0$
 - Ensures convergence of Debye-Waller & multi-phonon convolution integrals
 - *cf.* a non-scaled GMM in [1] with an adjustment (next slide)
- PDOS symmetry $\rho(\beta)=\rho(-\beta)$ built-in by *symmetric pairs* of normal PDFs
 - Convolution integral $(-\infty, \infty)$ of 2 normal PDFs is a normal PDF: $\mu_1+\mu_2, \sigma_1^2+\sigma_2^2$
 - Multi-phonon convolution \rightarrow multinomial sum of normal PDFs
 - Analytical expressions could be used to test numerical integration routine
- Scaling and symmetry enables derivation of NJOY TSL analytically
 - Derived gradients of DDXS w.r.t. GMM & LEAPR parameters for the GLLS [3]
 - TSL covariance matrix derived from the GMM parameter covariance matrix
 - LEAPR [2] parameters assumed to be normally distributed

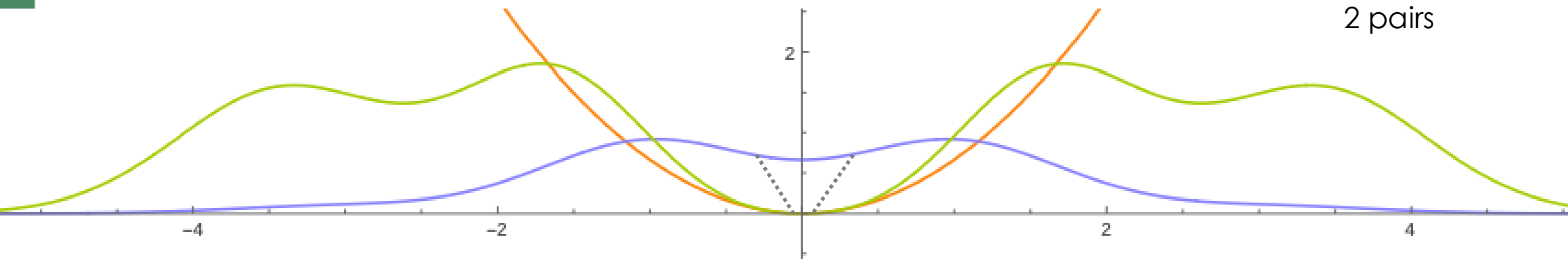
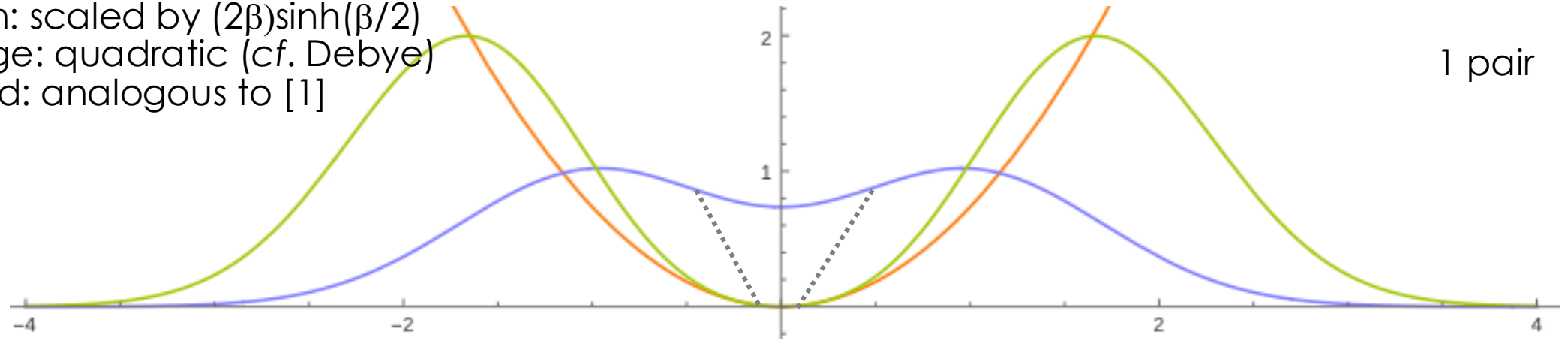
Illustration: scaled symmetric Gaussian pairs

Blue: unscaled

Green: scaled by $(2\beta)\sinh(\beta/2)$

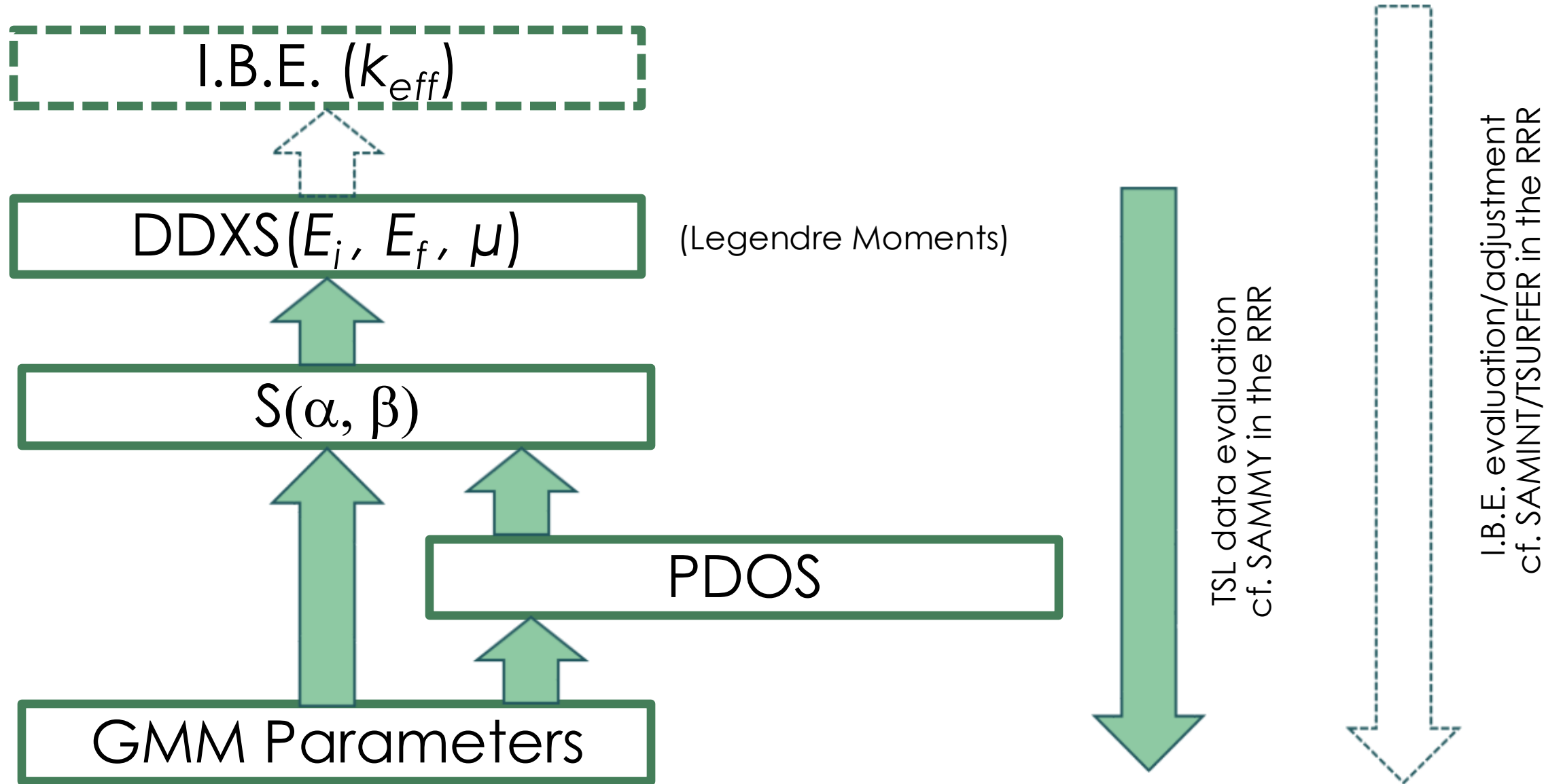
Orange: quadratic (cf. Debye)

Dotted: analogous to [1]



Energy transfer
(Arb. Units)

Hierarchy of TSL covariance matrices

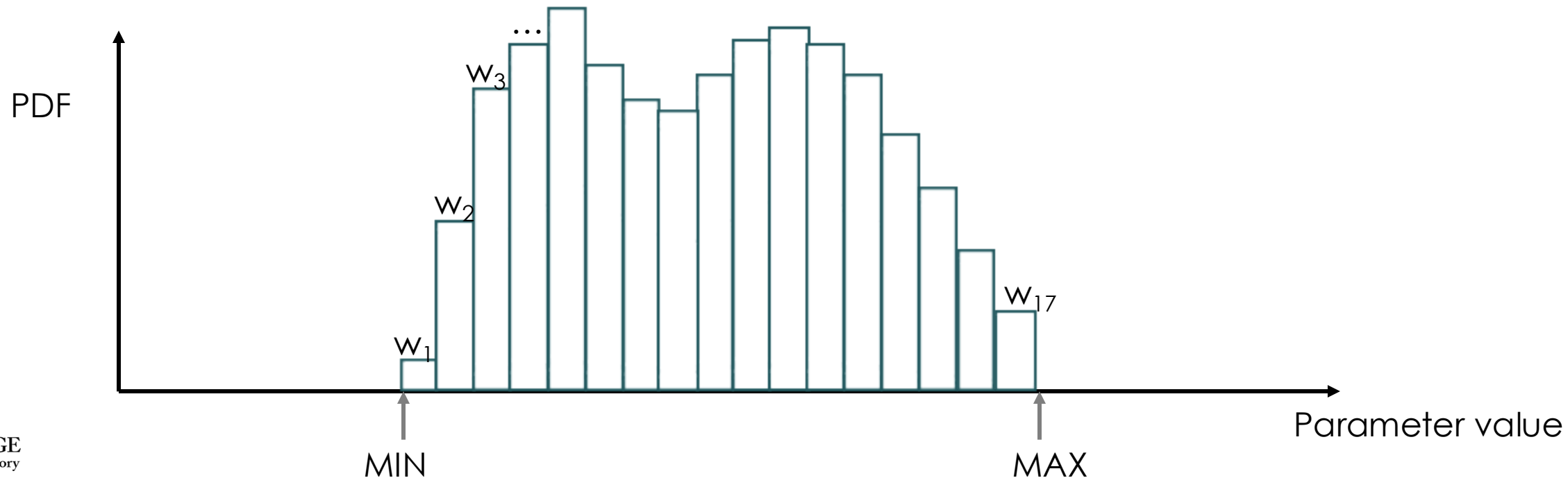


Gaussian Mixture Models (GMM) for Nuclear Data (ND)

- ND models beyond thermal are functions of many parameters
 - In contrast to the $S(\alpha, \beta) =$ a functional of one PDOS
- GMM of multivariate model is a sum of multivariate Gaussian PDF
 - cf. RRR evaluations represented by a single multivariate Gaussian
- Complicated because each multivariate GMM parameterized by:
 - weight, mean values (a vector!), covariance matrix(!)
- GMM parameters obey a Gaussian PDF to be evaluated via GLLS
 - It entails evaluation of covariance of GMM covariance matrix elements
 - A simpler alternative may be to reduce GMM to a histogram (weights only)

"Histogram" MM: parameter PDF specified by weights alone

- Weights obey a Gaussian PDF: their mean and covariance
 - Weights may be evaluated by the GLLS
- Bin structure chosen to accommodate the shape of the PDF
 - Parameter minimum and maximum values constrained may be enforced
- Sensitivity and Uncertainty Quantification still feasible



Summary and Outlook

- Scaled symmetric GMM could be used for evaluation of Phonon Density of States and thermal neutron scattering data by a Generalized Linear Least Squares (GLLS)
 - covariance matrix of thermal neutron double-differential scattering cross sections may be computed from the evaluated covariance matrix of GMM parameters by a linear approximation.
 - Enables simultaneous evaluations of PDOS+RRR data via GLLS
 - Enables simultaneous adjustment of PDOS+RRR by IBE data via GLLS
 - Generalizes the SAMMY/TSURFER/SAMINT concept in the RRR toward PDOS+RRR
 - Bayesian (Metropolis-Hastings) Monte Carlo to quantify non-linearities
- GMMs may, in principle, be useful for nuclear data evaluations, but the multivariate aspect of models makes it complicated
 - Alternative (e.g. histogram, ...) mixture models may be comparable but simpler for evaluation of non-normally distributed parameters

Acknowledgements

- This work is supported under the framework of DOE / US Nuclear Regulatory Commission (NRC) collaboration for Criticality Safety support for commercial-scale HALEU fuel cycles project (DNCSH)
- This work was supported by the Nuclear Criticality Safety Program, funded and managed by the National Nuclear Security Administration for the Department of Energy.

References:

- [1] Lance Maul, José Ignacio Márquez Damián, George Braoudakis, Mark Ho, Guan Heng Yeoh, "Perturbation scheme for estimating uncertainties in thermal scattering cross sections of water", *Annals of Nuclear Energy*, **121**, (232-249) 2018
<https://doi.org/10.1016/j.anucene.2018.07.020>
- [2] <https://docs.njoy21.io/projects/thermalScattering/en/latest/LEAPR/index.html>
- [3] Brown, Jesse M., et al. "Bayesian Optimization Framework for *Imperfect* Data or Models." , ORNL/TM-2022/2448, Apr. 2022.
<https://doi.org/10.2172/1874643>