Recent Work on Thermal Scattering Covariance Generation and Compression

[1]Aaron G. Tumulak, [2]Dorothea Wiarda, [2]Andrew M. Holcomb, [1]Brian C. Kiedrowski

[1]Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, MI

[2]Nuclear Energy and Fuel Cycle Division, Oak Ridge National Laboratory, Oak Ridge, TN

CSEWG Nuclear Data Week(s) – November 15, 2021

Table of Contents

Motivation

- How does is a quantity f affected by uncertainty in $x_1, \ldots, x_N$?
- Expand f about $\tilde{f} = f(\tilde{x}_1, \ldots \tilde{x}_N)$:

$$f(x_1, \ldots, x_N) = \tilde{f} + \frac{\partial f}{\partial x_1}(x_1 - \tilde{x}_1) + \ldots + \frac{\partial f}{\partial x_N}(x_2 - \tilde{x}_N)$$

- Treating $x_1, \ldots, x_N$ as random variables, we have f as a weighted sum of random variables with variance

$$\mathrm{Var}(f) = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \mathrm{Var}(x_i) + \sum_{i \neq j} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \mathrm{Cov}(x_i, x_j) \; .$$

- The partial derivatives $\partial f / \partial x_i$ are *sensitivities* and $\mathrm{Cov}(x_i, x_j)$ is the *covariance* between x_i and x_j.

Motivation

The Sandwich Formula

$$\mathrm{Var}\left(f\right) = \boldsymbol{s}^{\mathsf{T}} \boldsymbol{A} \boldsymbol{s}$$

where

$$\boldsymbol{s} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x} \end{pmatrix} \text{ and } \boldsymbol{A} = \begin{pmatrix} \mathrm{Cov}\left(x_1, x_1\right) & \cdots & \mathrm{Cov}\left(x_1, x_N\right) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}\left(x_N, x_1\right) & \cdots & \mathrm{Cov}\left(x_N, x_N\right) \end{pmatrix}$$

- End users compute $\boldsymbol{s}$.
- Evaluators provide covariances $\boldsymbol{A}$.
- The Evaluated Nuclear Data File format, ENDF-6 [1] provides file formats covariances for *some* types of nuclear data.

Thermal Scattering

Incoherent Inelastic Scattering

$$\sigma\left(E \to E', \mu\right) = \frac{\sigma_b}{2k_B T}\sqrt{\frac{E'}{E}}e^{-\frac{\beta}{2}}S\left(\alpha, \beta\right)$$

where

$$\alpha = \frac{E' + E - 2\mu\sqrt{EE'}}{Ak_B T} \text{ and } \beta = \frac{E' - E}{k_B T}.$$

- At energies (approximately less than $5\,\text{eV}$), thermal scattering takes place.
- The thermal scattering kernel $S\left(\alpha, \beta\right)$ is two dimensional data. The double differential scattering cross section $\sigma\left(E \to E', \mu\right)$ is actually three dimensional data.
- ENDF-6 does not currently specify a format for storing $S\left(\alpha, \beta\right)$ covariance.

Generalized Nuclear Data Format

- New data format [2] to eventually supersede ENDF-6
- Previous work developed a format in the Generalized Nuclear Data format for storing $S(\alpha, \beta)$ covariances.

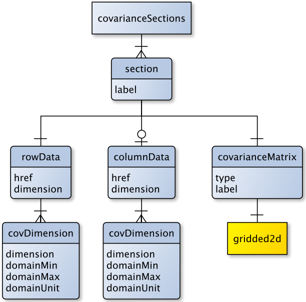


Figure: Structure of covariance data for thermal scattering data. From [3].

SCALE Implementation

- Implementation performed in AMPX [4] at GNDS Data Structures Level.
- Covariance for $S(\alpha, \beta)$ data is four dimensional.
- α and β indices are flattened to a linear index.
- Underlying data structure is a symmetric two dimensional matrix.

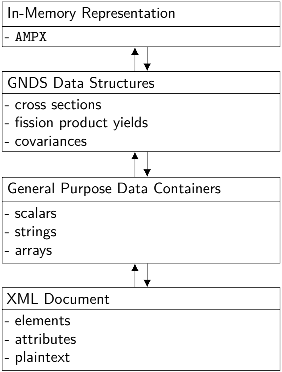


Figure: Hierarchy of abstraction layers for nuclear data in SCALE.

SCALE Implementation

- The base class `Covariance` is not necessarily symmetric to account for cross-covariance across different reactions.
- Existing classes updated to inherit from common base classes.

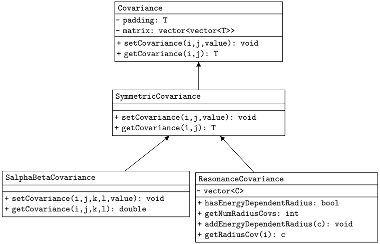


Figure: Covariance class inheritance

Generating $S(\alpha, \beta)$ Covariance Data

- Covariance data was generated from 1000 random realizations of H in H_2O.
- Full covariance matrix had $(182 \times 259)^2 \approx 2 \times 10^9$ entries.
- Full uncompressed covariance matrix occupied about 17.78 GB of storage as HDF5 file.
- GNDS format supports more than just α, β parameters. Adding one more dimension with just 10 points brings data storage into TB range.
- GNDS format supports basic compression methods. Sparsity and symmetric are immediately applicable.

	Min	Max	Points	Interpolation
α	0.0005	632.9	182	Lin-Log
β	0.0	158.1	259	Lin-Log

Table: Example TSL Data Dimensions

Principal Component Analysis

- Decompose the covariance matrix $\boldsymbol{A}$ into a unitary matrix of eigenvectors $\boldsymbol{U}$ and diagonal matrix of eigenvalues $\boldsymbol{\Lambda}$ by

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\intercal \,.$$

- If $\boldsymbol{A}$ is of rank r, then there are r eigenvectors making up the columns of $\boldsymbol{U}$ and r eigenvalues making up the diagonal entries $\boldsymbol{\Lambda}$:

$$\boldsymbol{U} = \begin{pmatrix} | & & | & | & & | \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r & \mathbf{0} & \cdots & \mathbf{0} \\ | & & | & | & & | \end{pmatrix} \quad \boldsymbol{\Lambda} = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0\right)$$

- A low rank approximation of $\boldsymbol{A}$ can be created by keeping only $\tilde{r} < r$ eigenvalues in $\boldsymbol{\Lambda}$ and discarding the $r - \tilde{r}$ smallest eigenvalues.

Principal Component Analysis

- The amount of information lost is

$$\frac{\left\|\boldsymbol{A}-\tilde{\boldsymbol{A}}\right\|_2}{\|\boldsymbol{A}\|_2}=\frac{\sqrt{\sum_{i=\tilde{r}+1}^{r}\lambda_i}}{\sqrt{\sum_{i=1}^{r}\lambda_i}}=\sqrt{\frac{\sum_{i=\tilde{r}+1}^{r}\lambda_i}{\sum_{i=1}^{r}\lambda_i}}.$$

- Significant amount of information retained by keeping only a few eigenvalues

$\tilde{r}$	Information Lost	Memory Usage (MB)
0	100.00%	0.0
1	17.34%	0.3771
6	0.94%	2.263
121	0.000997%	45.63
1000	0.00%	377.1

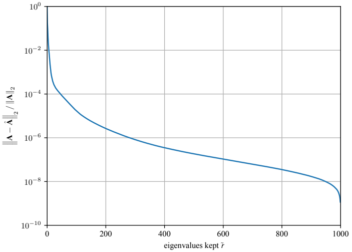
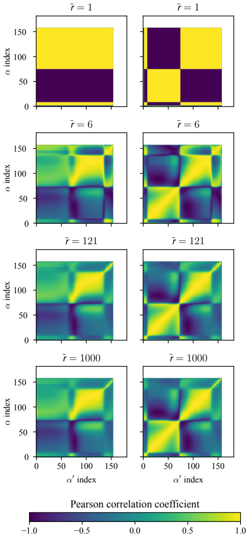


Figure: Frobenius norm of residual matrix $\boldsymbol{A}-\tilde{\boldsymbol{A}}$.

Principal Component Analysis

- Each sub-block has fixed β values and changing α values.
- The left column shows off-diagonal submatrices (corresponding to β indices $[116, 96]$).
- The right column shows on-diagonal submatrices (corresponding to β indices $[116, 116]$).
- Each row shows the approximated submatrix obtained by keeping the first $\tilde{r}$ largest eigenvalues.
- Values for which $S(\alpha, \beta)$ is undefined have no covariance



Conclusions

- Data structures for representing thermal scattering kernel covariances were implemented in AMPX.
- Covariance matrices will be quite large for practical problems.
- GNDS format already supports some methods of compression.
- Principal Component Analysis can take advantage of large singular values to further compress matrices.

Future Work

- Generate temperature dependent $S(\alpha, \beta, T)$ from perturbed probability density of states.
- Compute sensitivities for a benchmark problem.
- Propagate uncertainties in $S(\alpha, \beta, T)$ through benchmark problem.
- Quantify uncertainty due to uncertainty in $S(\alpha, \beta, T)$.

Acknowledgements

- This work is performed under NEUP Project 18-150008: Development of Thermal Inelastic Scattering Covariance Data Capabilities with Demonstration of Light Water Evaluation.

References

M. HERMAN,
"ENDF-6 Formats Manual Data Formats and Procedures for the Evaluated Nuclear Data File ENDF/B-VI and ENDF/B-VII,"
BNL-90365-2009, Brookhaven National Laboratory (BNL) National Nuclear Data Center (2009).

WPEC Subgroup 38,
"Specifications for the next generation nuclear data hierarchy," 2016.

A. G. TUMULAK, H. PARK, V. SOBES, W. S. YANG, and B. C. KIEDROWSKI,
"Progress on the Development of Thermal Scattering Covariance Formats and Processing Tools,"
Trans. Am. Nucl. Soc., **121**, 1354 (2019).

D. WIARDA, M. DUNN, N. GREENE, C. CELIK, and L. PETRIE,
"AMPX-6: A Modular Code System for Processing ENDF/B,"
ORNL/TM-2016/43, Oak Ridge National Laboratory (2016).