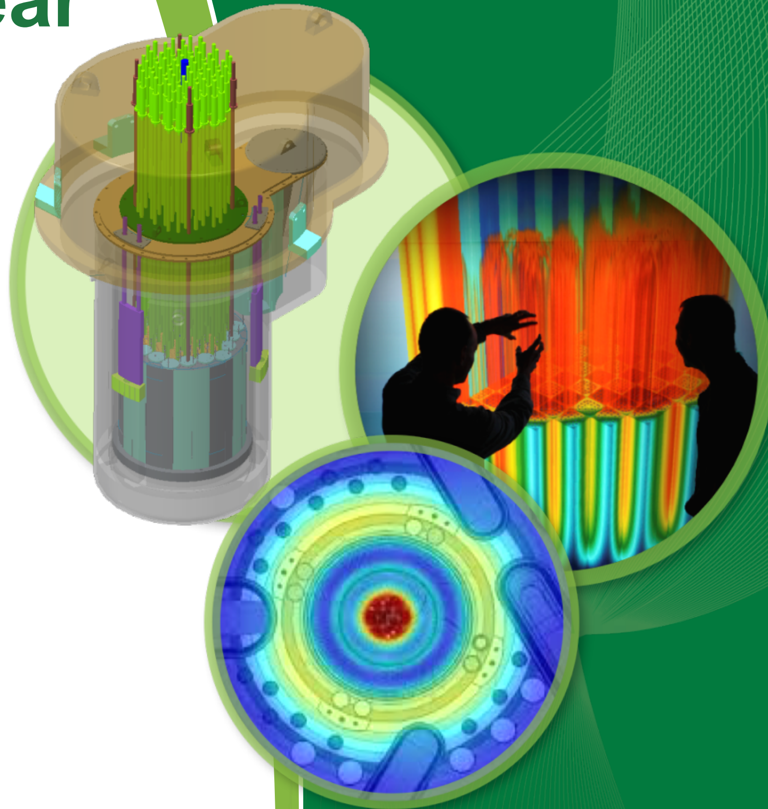
Advances in ORNL Nuclear Data Evaluation Methods



Goran Arbanas (ORNL)
Andrew M. Holcomb (ORNL)
Marco T. Pigni (ORNL)
Dorothea Wiarda (ORNL)
Vladimir Sobes (ORNL)
Christopher W. Chapman (ORNL)

Nuclear Data and Criticality Safety Group
Reactor and Nuclear Systems Division
Nuclear Science and Engineering Directorate

ORNL is managed by UT-Battelle
for the US Department of Energy

Cross Section Evaluation Working Group Meeting
Brookhaven National Laboratory, November 6-8, 2017

Outline

- Motivation and background
- SAMMY modernization update
- Bayesian generalized data optimization method
- Phenomenological Dirac relativistic R-matrix formalism
- Thermal neutron scattering, $S(\alpha,\beta)$, evaluation framework
- Conclusions and outlook

Motivation and background

- DOE's programmatic pursuit of improved nuclear data (ND) and uncertainties (variances and covariances)
  - Nuclear Criticality Safety, Nuclear Engineering & Applications, Science
- SAMMY modernization approved by the US Nuclear Criticality Safety Program (NCSP)
  - Modernize the software quality assurance (SQA) programming framework (C++ APIs)
  - Modernize the evaluation methods
- Need for simultaneous, consistent evaluation of differential and integral benchmark experimental data
  - Revisit application of Bayes' Theorem to ND
  - Improve consistency between differential and integral uncertainties
  - Evaluated cross sections are now post-adjusted to fit integral data

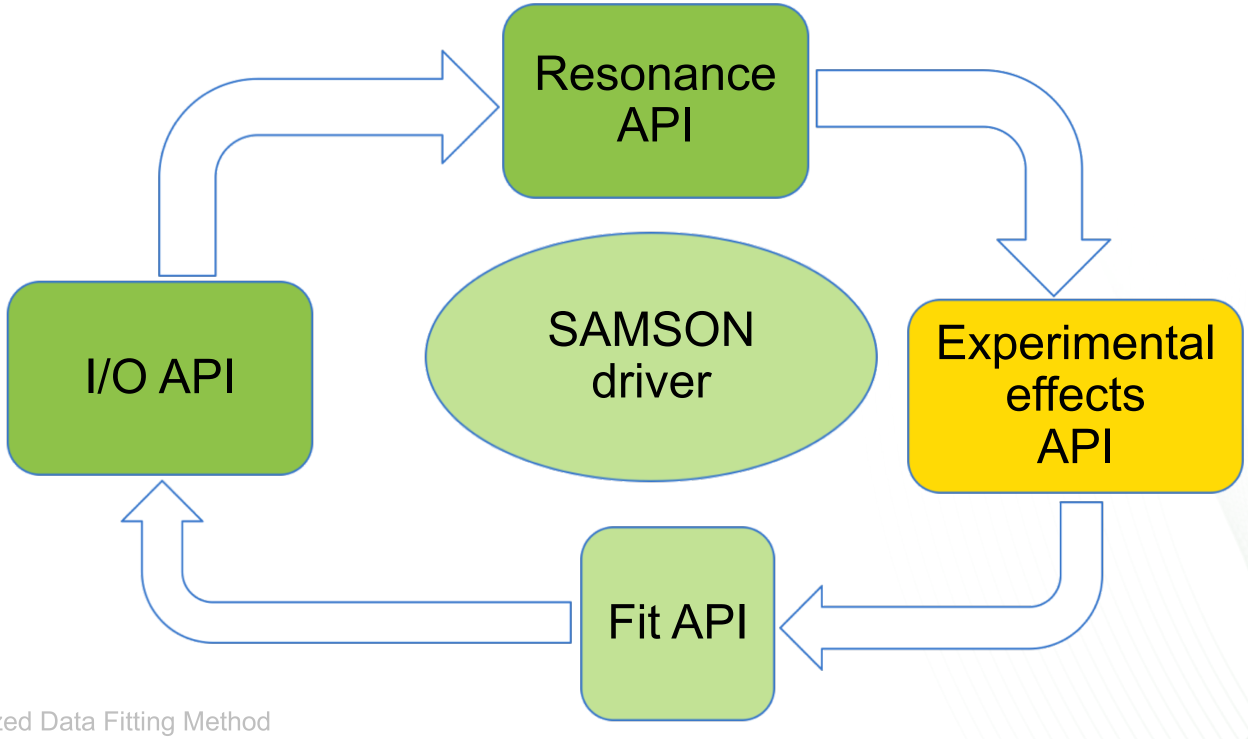
SAMMY 8.1 released by RSICC April 2017:

- SAMINT: integral benchmark experiments inform research parameter evaluations (Implemented by Vlad Sobes)
- SAMMY integrated into SCALE SQA in AMPX footsteps
  - Automated cmake/ctest suite, revision control repository, FogBugz
  - Platforms supported Linux/gfortran, Mac/gfortran, Windows/ifort
- New detector resolution functions in collaboration with Rensselaer Polytechnic Institute (RPI)
- Updated physical constants; SAMMY and SAMRML compute consistently now
- Corrected a misplaced index that was causing incorrect matrix multiplication for non-diagonal data covariance matrix
- Implemented several other bug fixes
- Obtained license for distributing COULCWF with SAMMY/AMPX/SCALE

OAK RIDGE National Laboratory

SAMSON high-level application programming interface (API) diagram

- Defines APIs before implementation
  - Enables variety of methods for each API
  - Leverages input/output (I/O) and Resonance API from C++ SAMRML
    - SAMMY PARameter and GND file reader/writer under development
    - To replace SAMMY I/O routines



Fit API: GLS, Bayesian Monte Carlo

GLS

Bayesian Monte Carlo

- Generalized least squares (GLS)
  - Nuclear Engineering Science Laboratory Synthesis (NESLS) summer intern Jinghua Feng implemented a prototype
  - Andrew Holcomb ported the prototype into the FitAPI
  - Compact expressions by Froehner were used (Sect. 2.2 of JEFF Report 18, 2000)
- Bayesian generalized data optimization method (CW2017)
  - NESLS summer intern Jinghua Feng implemented a prototype
  - Prototype avoids Peelle's Pertinent Puzzle (PPP) issue plaguing GLS of reduced data
- Fit differential and/or integral benchmark experiment data performed simultaneously instead of adjusting differential data retroactively

OAK RIDGE National Laboratory

Fit API: Preliminary interface

Interface
Fit
setData Set an instance of Data interface **initialize** After setting data object initialize internal data structures **execute** Do the actual fitting **finalize** Clean up any internal resources

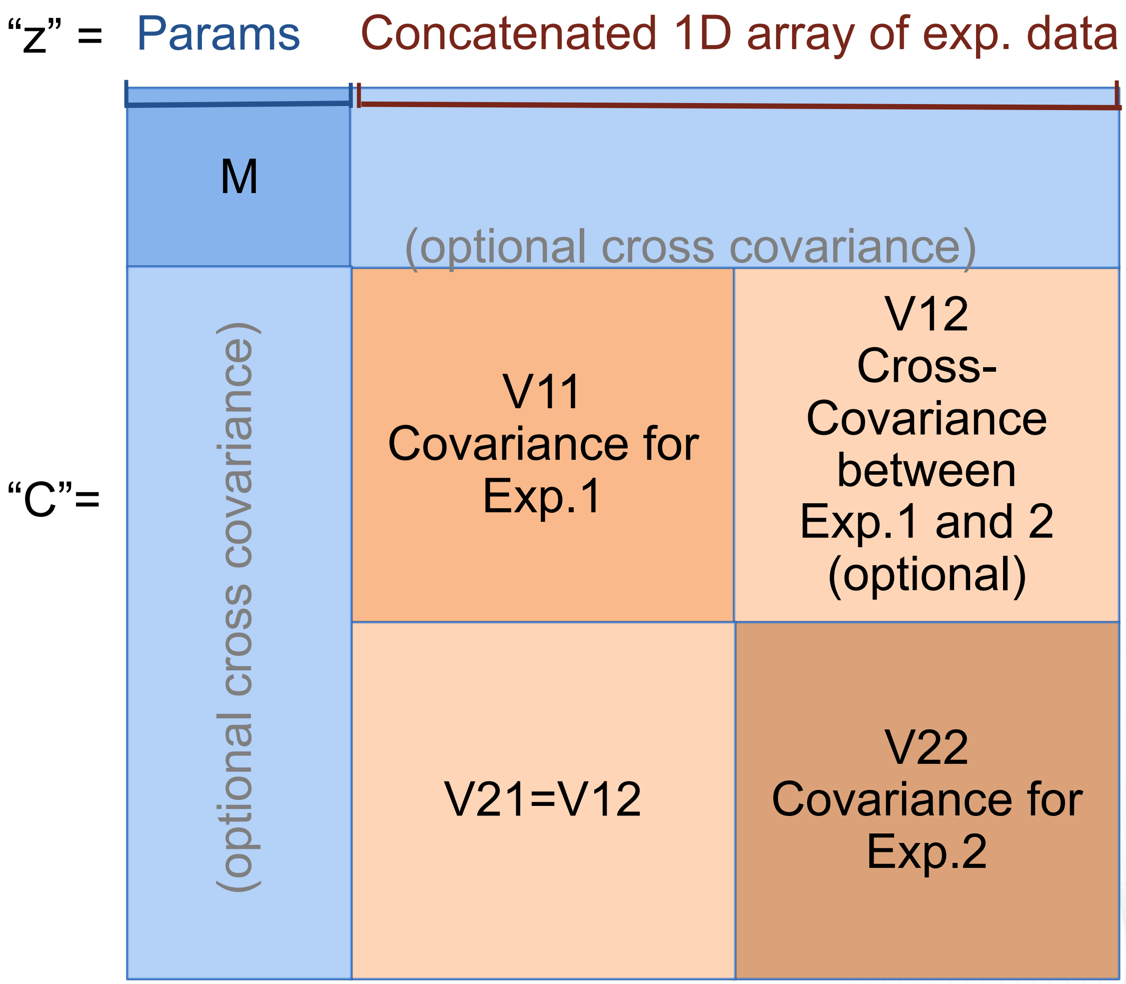
Interface
Array
getNumDim Get the number of dimensions **getSize(int dim)** Get the array size for dimension m **getValue(int i1, int i2, …)** Get the value for the indicated indices. In C++ we would pass in a vector of length getNumDim **setValue(int i1,int i2, …_** Set value

Interface
Data
getNumberParams Get the number of parameters **getNumData** Get the number of experimental data **getData** Get the list of experimental data (1-dim Array) **getParam** Get the list of initial params (1-dim Array) **getCovMatrix** Get the full covariance matrix (2-dim Array) **getTheory** Get theoretical values based on current parameters (1-dim Array) **setParam** Set the current parameters (1-dim Array) **setCovMatrix** Set the full covariance matrix (2-dim Array)

- Actual instances are instantiated by a factory class
- Data will have a method to obtain the derivatives (2-dim Array: getNumberParams x getNumData); there will be a function that computes derivatives numerically
- Fit calls setParams, getTheory, setCovMatrix repeatedly in the course of fitting the data

Fit API: GLS implementation

- Parameters and experimental data cast into 1D array by implementation of data
  - for generic use inside SCALE framework
  - Froehner's formulation and notation:



Fit API: GLS implementation

GLS

- Generalized least squares (GLS)
  - Compact expressions by Froehner (Sect. 2.2 of JEFF Report 18, 2000)

    (Warning: Nondiagonal data covariance matrix [DCM] may lead to PPP syndrome)

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \left[\mathbf{S}(\mathbf{x}_n)^\dagger \mathbf{C}^{-1} \mathbf{S}(\mathbf{x}_n)\right]^{-1} \mathbf{S}(\mathbf{x}_n)^\dagger \mathbf{C}^{-1} \left[\zeta - \mathbf{z}(\mathbf{x}_n)\right]$$

- Implementation uses *cpp-array* library (CPC **185**,1681, 2014)
  - Transparently parallelized via BLAS library (Intel MKL, cuBLAS, MAGMA)
  - Compact notation for matrix operations, e.g. parameter set update
  - P = P + inv(transpose(S) * inv(C) * S) * transpose(S) * inv(C) * (Π – T)
- BLAS advantages: drastically speeds up large matrix operations in SAMMY and shortens code (Arbanas, Dunn, Wiarda, M&C2011)

Experimental effects (EE) API

- Convolution of Doppler broadening, target, and detector effects, each one implenting the EE API:

Doppler broadening: FGM, DDXS, S(a,b) BROADEN/AMPX	Neutron transport: SHIFT API (DBRC, LH, multiple scatter.)	Resolution function via Monte Carlo N-Particle (MCNP) for example

- SHIFT API for on-the-fly neutron transport aspects
  - To enable fitting integral benchmark experiments (IBEs)
  - Developed for SCALE by Cihangir Celik (ORNL Nuclear Data and Criticality Safety) during FY2017
  - Message passaging interface (MPI) enabled
  - Could use MCNP input
- In principle, the entire experimental setup could be simulated
  - Fitting to raw data may be possible; varying opinions
  - Raw data may become publicly available (needed for Bayesian MC)

Relevant quotes from Edwin T. Jaynes (1922-1998)

- "… every Bayesian problem is open ended; no matter how much analysis you have completed, this only suggests still *other kinds of prior information* that you might have had, and therefore still more interesting calculations that need to be done to get still deeper insight into the problem."

⇨ *model defect* and its covariance introduced as priors

- "The *difficulties are never mathematical*; at no point do we encounter any mathematical problem that could not be dealt with by an undergraduate…"

⇨ Bayes' theorem is straightforward for given priors

from "Straight Line Fitting – a Bayesian Solution" (1999), which chronicles Bayesian straight line fitting during his lifetime (http://bayes.wustl.edu/etj/articles/leapz.pdf)

Principles and notation used in derivation

- Distinction is maintained between the variables' values and their *expectation* values for both *prior* **and** *posterior*
  - Prior generalized data expectation values $<z>=(<P>, <D>)$
  - PDF of a particular value $z=(P,D)$ is $p(z|<z>,C,T(.))$
  - Posterior expectation values denoted by a prime: $<z>'=(<P>', <D>')$
- *Symmetry* of any normal PDFs was enforced:

  $(z-<z>)^{+}C^{-1}(z-<z>)$, where $C=<(z-<z>)\,(z-<z>)^{+}>$
- The model, its defect, and defect covariance introduced as priors, which has important implications in the application of Bayes' theorem

$$T(\cdot), \langle\delta\rangle, \quad \mathbf{\Delta} \equiv \langle(\delta - \langle\delta\rangle)(\delta - \langle\delta\rangle)^{\mathsf{T}}\rangle$$

OAK RIDGE National Laboratory

Definitions used in derivation

- Generalized data, $z \equiv (P, D)$, and its covariance matrix $\mathbf{C}$

$$
\begin{aligned}
\mathbf{C} &\equiv \langle (z - \langle z \rangle)(z - \langle z \rangle)^{+} \rangle \\
&= \begin{pmatrix} \langle (P - \langle P \rangle)(P - \langle P \rangle)^{+} \rangle & \langle (P - \langle P \rangle)(D - \langle D \rangle)^{+} \rangle \\ \langle (D - \langle D \rangle)(P - \langle P \rangle)^{+} \rangle & \langle (D - \langle D \rangle)(D - \langle D \rangle)^{+} \rangle \end{pmatrix} \\
&\equiv \begin{pmatrix} \mathbf{M} & \mathbf{W} \\ \mathbf{W}^{+} & \mathbf{V} \end{pmatrix}
\end{aligned}
$$

- Generic Bayes' theorem:

$$p(\alpha|\beta\gamma) = p(\alpha|\gamma\beta) \propto p(\alpha|\gamma)p(\beta|\alpha\gamma)$$

- Generic Product Rule of Probability Theory:

$$p(\alpha\beta|\gamma) = p(\alpha|\beta\gamma)p(\beta|\gamma)$$

where α ,β and γ will be specified for each particular application

Direct derivation in generalized data notation

- Application of Bayes' theorem

$$p(\alpha|\beta\gamma) = p(\alpha|\gamma\beta) \propto p(\alpha|\gamma)p(\beta|\alpha\gamma)$$

with the following substitutions:

$$\alpha \to z$$

New prior: ⇨ $\beta \to T(\cdot), \langle\delta\rangle, \mathbf{\Delta}$ $\qquad \mathbf{\Delta} \equiv \langle(\delta - \langle\delta\rangle)(\delta - \langle\delta\rangle)^{\mathsf{T}}\rangle$

$$\gamma \to \langle z\rangle, \mathbf{C},$$

yields a posterior PDF for ***generalized*** data:

Posterior $\propto$ Prior $\times$ Likelihood

$$p(z|\langle z\rangle, \mathbf{C}, T(\cdot), \langle\delta\rangle, \mathbf{\Delta}) \propto p(z|\langle z\rangle, \mathbf{C}) \times p(T(\cdot), \langle\delta\rangle, \mathbf{\Delta}|z, \langle z\rangle, \mathbf{C})$$

(Normalized) generalized data posterior PDF

- Enables computation of arbitrary posterior expectation values:

$$\langle f(z)\rangle' = \int dz f(z) p(z|\langle z\rangle, \mathbf{C}, T(\cdot), \langle\delta\rangle, \mathbf{\Delta})$$

$$\langle z'\rangle \longrightarrow f(z') = z' = (P', D')$$

$$\mathbf{C}' \longrightarrow f(z') = (z' - \langle z'\rangle)(z' - \langle z'\rangle)^{+}$$

- Integration is over all posterior values of z'=*(P',D')*
  - Analogous to integrating out "nuisance parameters" (Jaynes 1999)
  - Posterior expectation values <z'>=(<P'>,<D'>) ≠ <z>=(<P>,<D>)
    - <D'>≠<D>; posterior experimental data are informed by the model
    - This appears to be an organic aspect of Bayes' theorem, not an adjustment
  - Evaluated model covariance matrix can be computed to ***all*** orders
  - Use *f(z')* below; *cf.* 1st order approximation (dT/dP')M'(dT/dP'), and V'

$$f(z') = (T(P') - \langle T(P')\rangle)(T(P') - \langle T(P')\rangle)^{+}$$

New Bayesian derivation outline

Bayes' Theorem

↓

Bayesian Generalized Data Fitting Method

↓

A Form Specialized for Normal PDFs

↓

Constrained Generalized Least Squares (CGLS)

↓

Generalized Least Squares (GLS)

The GLS in ND is at the bottom of this approximation hierarchy

Hierarchy of approximations I: posterior PDFs

- The most general posterior generalized data PDF is

$$p(z|\langle z\rangle, \mathbf{C}) \; \times \; p(T(\cdot), \langle\delta\rangle, \mathbf{\Delta}|z, \langle z\rangle, \mathbf{C})$$

- Assuming normal forms of PDFs, where $\tilde{\delta} \equiv T(P) - D$,

$$\mathcal{N}(z|\langle z\rangle, \mathbf{C}) \; \times \; \mathcal{N}(\tilde{\delta}|\langle\delta\rangle, \mathbf{\Delta})$$

- Taking the limit $\langle\delta\rangle \to 0$ and $\mathbf{\Delta} \to \mathbf{0}$

$$\mathcal{N}(z|\langle z\rangle, \mathbf{C}), \quad \text{constraint} \;\; T(P) = D$$

- Application of constraint into prior yields UMC-B w/o constraint

$$\mathcal{N}(z|\langle z\rangle, \mathbf{C}) \qquad z \equiv (P, D) \to (P, T(P))$$

Hierarchy of approximations II: cost functions

- The most general PDF in normal form, where $\tilde{\delta} \equiv T(P) - D$,

$$Q(z) \equiv (z - \langle z \rangle)^{\intercal} \mathbf{C}^{-1} (z - \langle z \rangle) + (\tilde{\delta} - \langle \delta \rangle)^{\intercal} \mathbf{\Delta}^{-1} (\tilde{\delta} - \langle \delta \rangle)$$

becomes CGLS in the limit $\langle \delta \rangle \to 0 \text{ and } \mathbf{\Delta} \to \mathbf{0}$

$$Q(z) \approx (z - \langle z \rangle)^{\intercal} \mathbf{C}^{-1} (z - \langle z \rangle), \quad \text{constraint } T(P) = D$$

- Application of constraint *before* cost minimization yields GLS

$$\begin{aligned} Q(z) \approx \quad & (P - \langle P \rangle)^{\intercal} \mathbf{M}^{-1} (P - \langle P \rangle) \\ & + (T(P) - \langle D \rangle)^{\intercal} \mathbf{V}^{-1} (T(P) - \langle D \rangle) \end{aligned}$$

Testing favors CGLS over GLS: χ^2 minimization

- CGLS has been used in high energy physics for 50+ years (APLCON; no PPP when used for systematic or data reduction parameters
- V. Blobel (U. Hamburg, DESY) compiled cases for which CGLS finds more reasonable fits than GLS or χ^2 minimization
  - http://www.desy.de/~blobel/apltalk.pdf
  - 2-D fitting with line, parabola, uncertainties in x- and y-dimensions
  - PPP appears in χ^2-min. but ***not*** in CGLS (p.17, 18)
  - Related to non-diagonal reduced DCM computed at experimental vs. theoretical values (N. Larson, SAMMY)

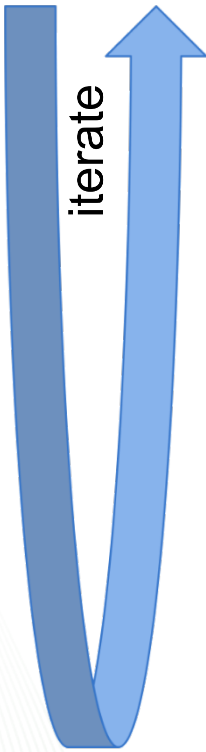
Constrained iterative Monte Carlo algorithm

- Iterate on posterior distribution until convergence is attained
  - "*k*," random sample label from posterior normal distribution $\mathcal{N}(\langle z'\rangle, \mathbf{C}')$
  - ω_k is the Monte Carlo weight for computing expectation values

$$\omega_k = \frac{p(z'_k|\langle z\rangle, \mathbf{C})}{p(z'_k|\langle z'\rangle, \mathbf{C}')} e^{-\frac{1}{2}(D'_k - T(P'_k))^+ \mathbf{X}'^{-1}(D'_k - T(P'_k))}$$

For constraint $\langle c(\cdot)\rangle = 0$

$$\langle z'\rangle = \sum_k \hat{\omega}_k z'_k \qquad \mathbf{C}' = \langle (z' - \langle z'\rangle)(z' - \langle z'\rangle)^+\rangle$$



- Resample using SVD of C':

$$z' = \langle P', T(P')\rangle + \mathcal{V}\mathcal{E}^{1/2}\mathcal{R} \qquad \text{where} \qquad \mathbf{C}' = \mathcal{V}\mathcal{E}\mathcal{V}^+$$

  - $\mathcal{R}$ is a *(#P + #D)* × *(#k)* matrix of $\mathcal{N}$(0,1) normal random numbers

Oak Ridge National Laboratory

Conclusions and future applications

- Generalized data fitting method based on the Bayes' theorem
  - Accommodates model defect via constraint on expectation values
  - Implemented via iterative Monte Carlo method (J. Feng for SAMMY)
  - Its special cases:
    - Constrained [generalized] least squares (CLS) of V. Blobel in APLCON
    - Generalized linear least squares (GLLS) [constrained] TSURFER WHISPER
    - But NOT the $\chi^2 = (D-T(P'))V^{-1}(D-T(P'))$, which may yield different results (Blobel)
- Comparison of CLS/GLS to quantify accuracy of approximations
- Consideration of log normal random distributions (I. Kodeli)
- Enables multidimensional sampling such as cross sections *and* energies
- The new method could be used for consistent simultaneous optimization of differential and integral benchmark experiments
  - To simultaneously fit differential and integral parameters
  - To yield more complete and consistent evaluations

Phenomenological Dirac R-matrix formalism

- Originally derived for calculable R-matrix, but expressed in a form that could be used for phenomenological fitting
  - Ph.D. Thesis (2011): http://scholarworks.wmich.edu/dissertations/411
- Boundary condition is determined by the channel radius
- Compare to approximations and the non-relativistic R-matrix

J. Grineviciute and Dean Halderson PHYSICAL REVIEW C **85**, 054617 (2012)

$$R_{cc'} = \sum_{\mu} \gamma_{\mu c}\gamma_{\mu c'}/(E_\mu - E)$$

$$\mathbf{S} = \mathbf{v}^{1/2}(\mathbf{Fo} - \mathbf{RGo})^{-1}(\mathbf{F_I} - \mathbf{RG_I})\mathbf{v}^{-1/2} \qquad T_{cc'} = i(\delta_{cc'} - S_{cc'})/2$$

$$\langle f\rangle_{\alpha\sigma M_A,\alpha'\sigma' M'_A} = \frac{1}{k}\sum \sqrt{4\pi(2\ell+1)}C^{\ell 1/2 j}_{0\sigma m} C^{J_A j J_B}_{M_A m M_B} C^{J'_A j' J_B}_{M'_A m' M_B} C^{\ell' 1/2 j'}_{m'_\ell \sigma m'} \times i^{(\ell-\ell')} e^{i(\delta'_\kappa + \delta'_{\kappa'})} T_{\alpha J_A \ell j J_B, \alpha' J'_A \ell' j' J'_B} Y_{\ell' m'_\ell}(\mathbf{k}').$$

$$\frac{d\sigma}{d\Omega}(\theta) = \frac{1}{2(2J_A+1)} \sum_{\sigma\sigma' M'_A M'_A} \left| \langle f_c\rangle_{\sigma\sigma'}\delta_{J_A\alpha M_A, J'_A\alpha' M'_A} + \langle f\rangle_{\alpha\sigma M_A,\alpha'\sigma' M'_A}\right|^2$$

OAK RIDGE National Laboratory

ORNL S(a,b) evaluation framework overview

- The objective is to combine experimental double differential scattering data and model parameters to yield the best estimate of double differential cross section (DDCS) and *uncertainties*
- Data and simulation fit is achieved using the Unified Monte Carlo (UMC) [1] method
- Simulations are constrained by physical properties of material
- Framework tested on light water
  - Data collected from ORNL Spallation Neutron Source (SNS)
  - RPI collaboration
- Validated using benchmarks from the International Criticality Safety Benchmark Evaluation Project (ICSBEP) handbook
- C. Chapman's Ph.D. https://smartech.gatech.edu/handle/1853/58693

Framework specifics: simulation

- Ran simulations of TIP4P/2005f [2] water in a box in the classical molecular dynamics (MD) code GROMACS [3]
- Computed density of states using trajectories from GROMACS
- Intermediate structure factor calculated using Gaussian approximation found in Abe and Tasaki [4]

$$F(q,t) = \exp\left[-\frac{q^2}{2}\frac{\hbar}{M}\int_0^\infty d\omega\,\frac{g(\omega)}{\omega}\left\{\coth\left(\frac{\hbar\omega}{2k_BT}\right)(1-\cos(\omega t)) - i\sin(\omega t)\right\}\right]$$

$$S(\alpha,\beta) = k_BTe^{\frac{-E}{2k_BT}}\big(S(q,E)\big) = k_BTe^{\frac{-E}{2k_BT}}\left(\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}F(q,t)e^{\frac{-iEt}{\hbar}}dt\right)$$

OAK RIDGE National Laboratory

Framework specifics: simulation

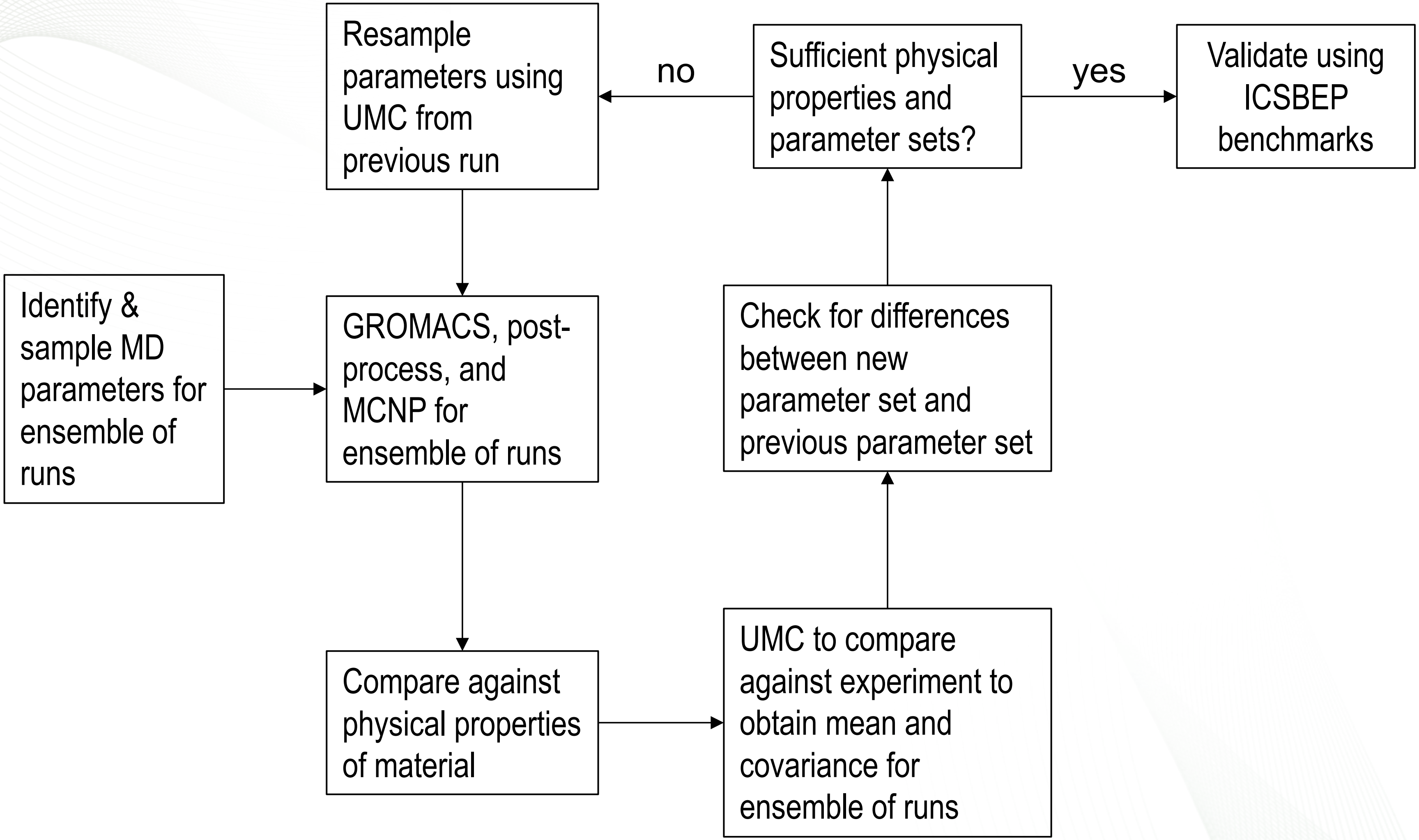
- Calculated scattering law using code developed for this project:

  Intermediate structure factor → dynamic structure factor → scattering law

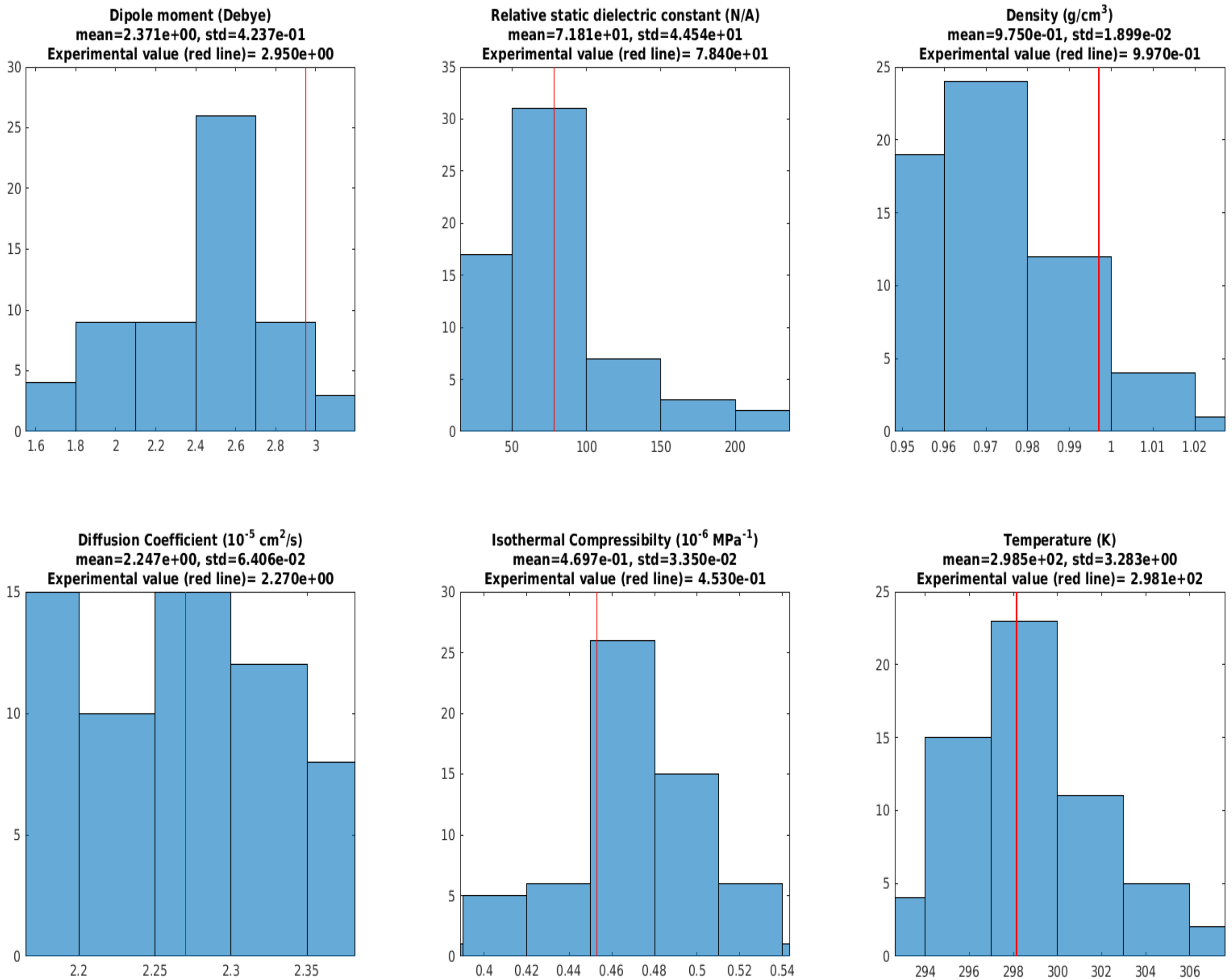
- Ran simplified MCNP code
- DDCS convoluted with SNS detector resolution function

$$\frac{d^2\sigma}{dE_f d\Omega} = \frac{\sigma_b}{4\pi k_B T}\sqrt{\frac{E_f}{E_i}}\, e^{\frac{-\beta}{2}} S(\alpha, \beta)$$

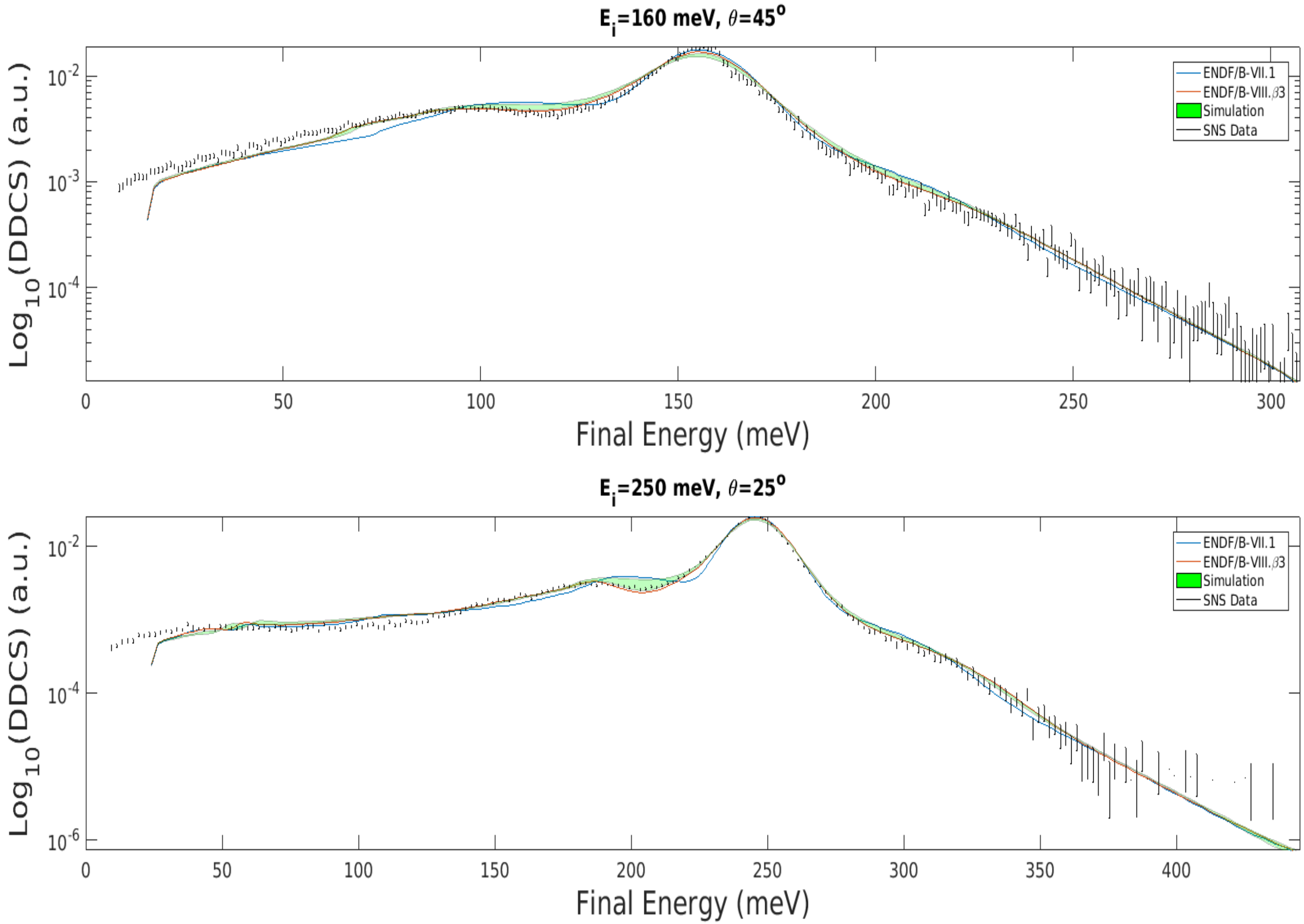
Framework outline



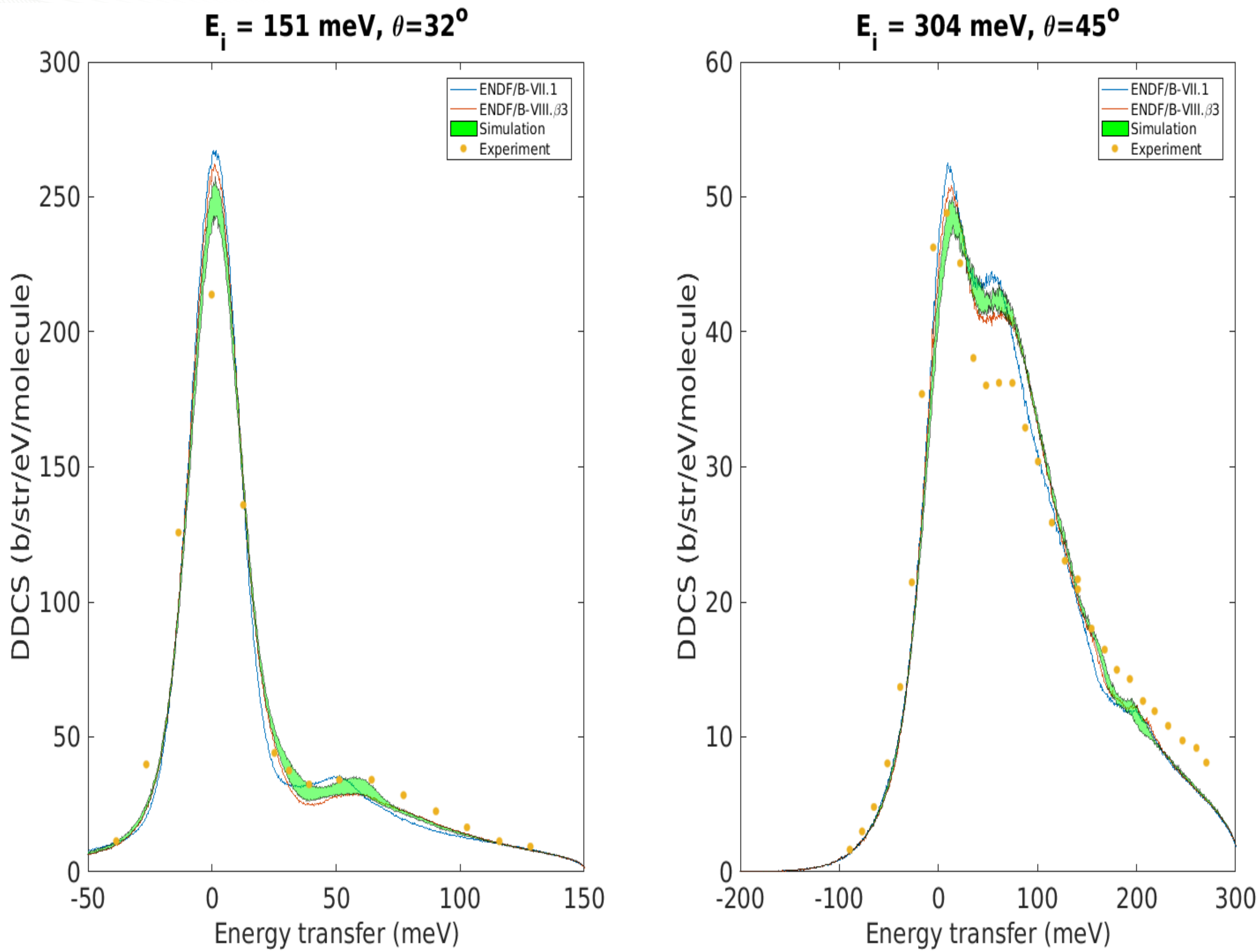
Results: properties of water



Results: DDCS – SNS

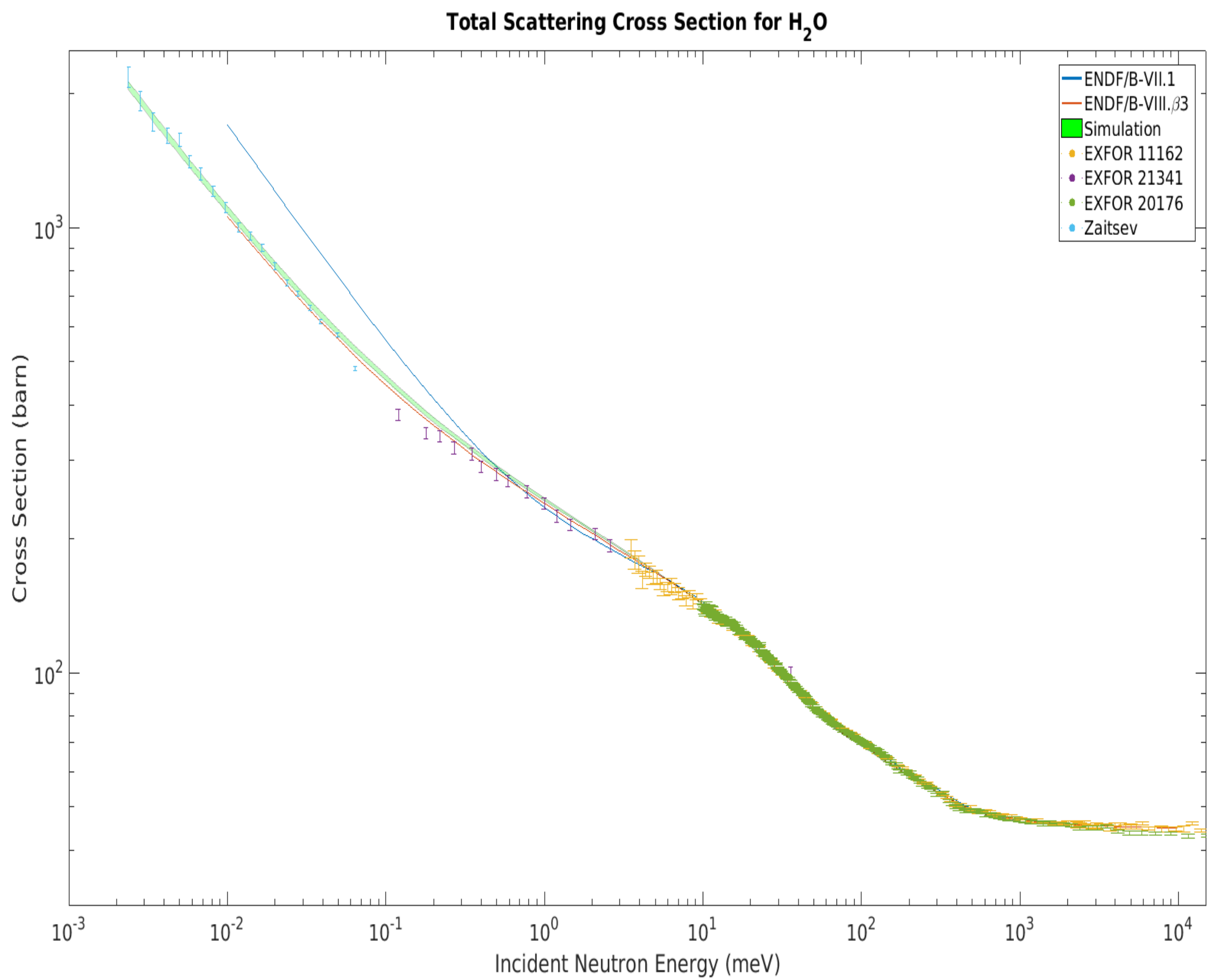


Results: DDCS – independent



OAK RIDGE
National Laboratory

Results: total cross section

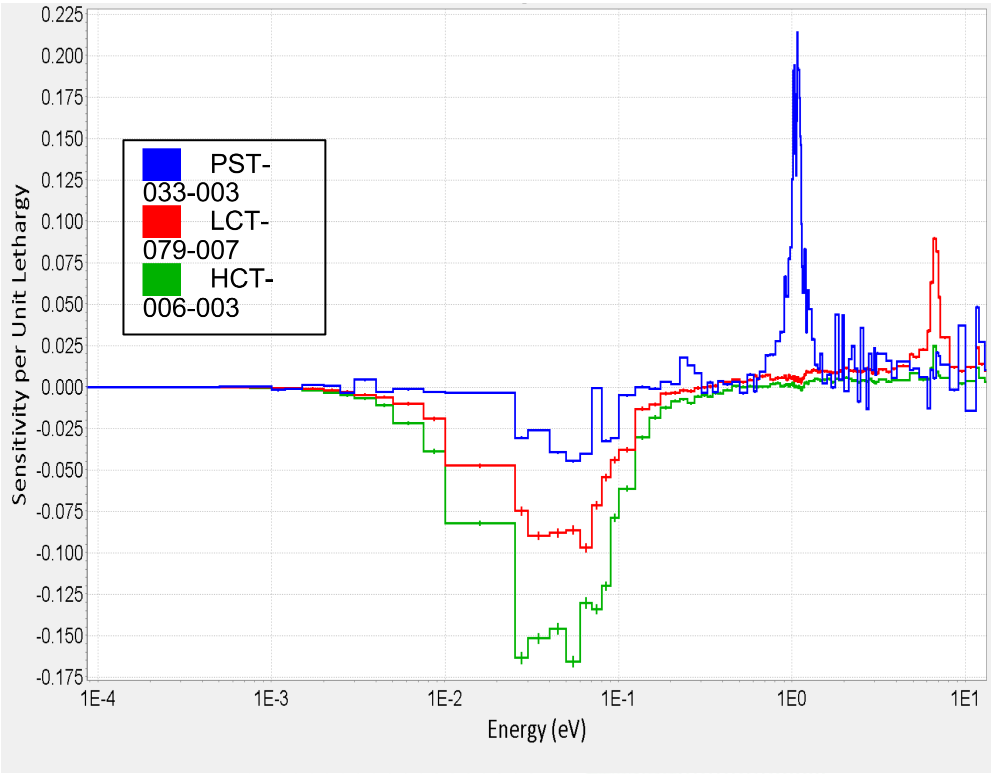


Results: ICSBEP benchmarks

Benchmarks

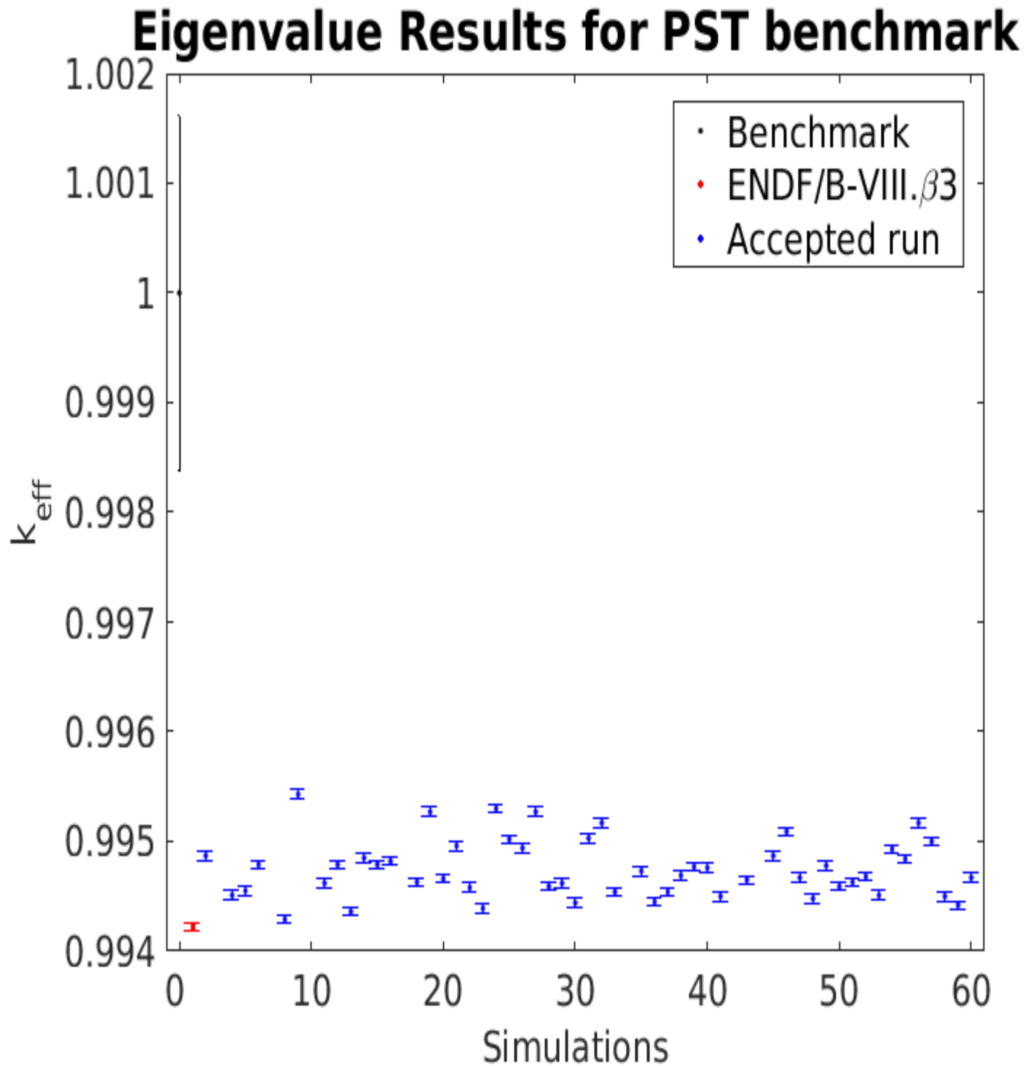
- PST-033-003: plutonium nitrate solution surrounded by water
- LCT-079-007: water moderated and reflected triangular pitched UO_2 fuel elements
- HCT-006-003: water moderated hexagonally pitched high enriched (80% ^{235}U) fuel rods

Sensitivity plot



Results: PST-033-003

ENDF Library	K_{eff}	St.Dev (pcm)	ΔK_{eff} (pcm)
Benchmark	1.00000	162	N/A
ENDF/B-VII.1	0.99349	4	651
ENDF/B-VIII.β3	0.99422	4	578
New XS	0.99483	4	517

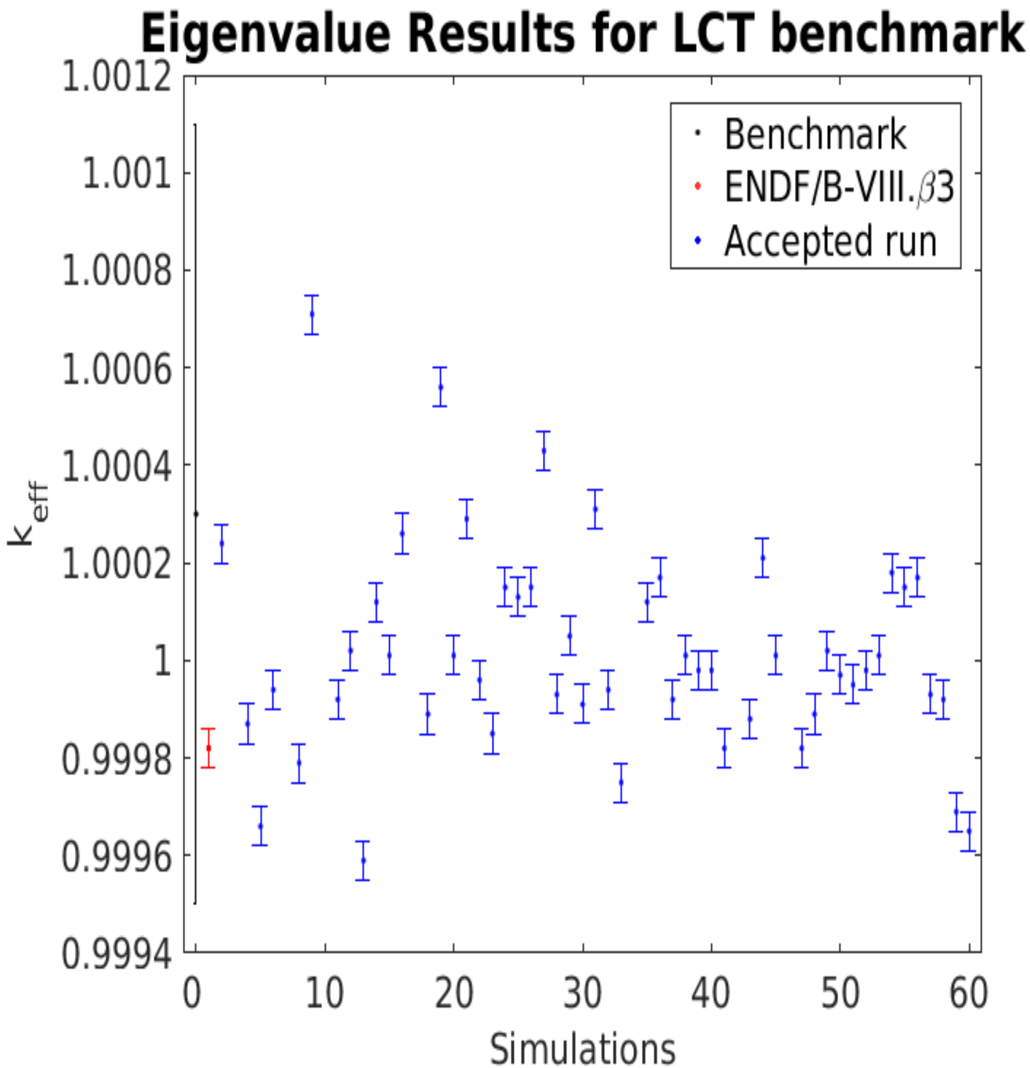


- Noticeable improvement
- Issues may be caused by materials other than light water

OAK RIDGE National Laboratory

Results: LCT-079-007

ENDF Library	K_{eff}	St.Dev (pcm)	ΔK_{eff} (pcm)
Benchmark	1.00030	80	N/A
ENDF/B-VII.1	0.99933	4	97
ENDF/B-VIII.β3	0.99982	4	48
New XS	1.00006	4	24

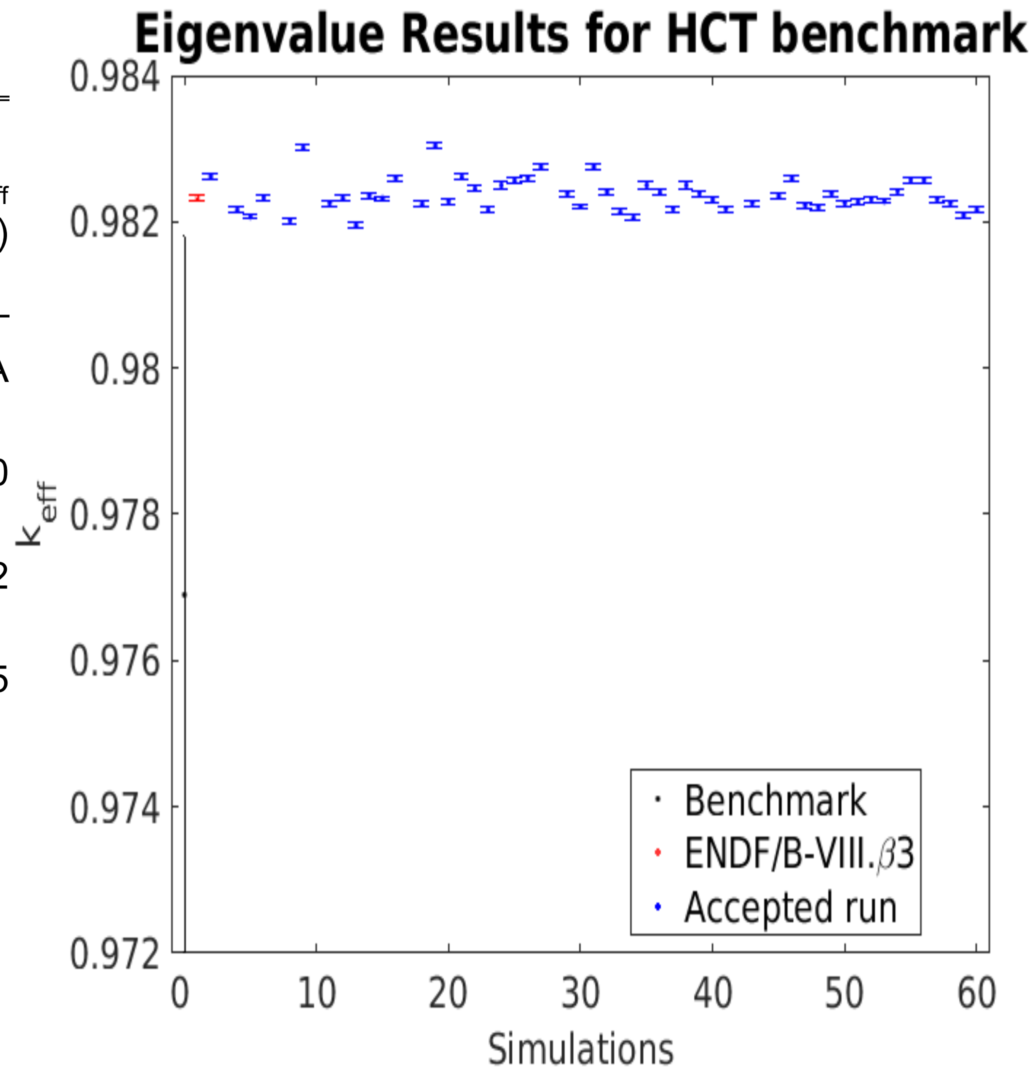


- Inconclusive improvement
- Results within error bounds of experimental uncertainty

OAK RIDGE National Laboratory

Results: HCT-006-003

ENDF Library	K_{eff}	St.Dev (pcm)	ΔK_{eff} (pcm)
Benchmark	0.97690	490	N/A
ENDF/B-VII.1	0.98190	4	-500
ENDF/B-VIII.β3	0.98232	4	-542
New XS	0.98245	4	-555



- Slightly worse results
- Still within 2 standard deviations of experiment

Conclusions

- Evaluation Framework was presented to generate thermal scattering law and uncertainties
- It showed sufficient agreement with the properties of light water
- It also showed good agreement with ENDF/B-VII.1 and ENDF/B-VIII.β3
- It has been validated against independent DDCS and ICSBEP benchmarks

Future work

- Other temperatures for light water: recent discussion about deficiencies at high temperatures
- Other models of light water
- Other materials
- Covariance generation and propagation
- ENDF format for $S(\alpha, \beta)$ covariance and uncertainties

S(a,b) Acknowledgments

- RPI: Yaron Danon, Kemal Ramic, and Carl Wendorff
- SNS: Garrett Granroth and Alexander Kolesnikov
- IRSN: Luiz Leal and Vaibhev Jaiswal
- CAB: Jose Ignacio Marquez Damian
- MD simulations performed at the National Energy Research Scientific Computing Center (NERSC) courtesy of Nuclear Data and Nuclear Theory Computing in the Nuclear Physics of the DOE Office of Science
- The work was funded by the DOE National Nuclear Security Administration's Nuclear Criticality Safety Program

OAK RIDGE National Laboratory