Nuclear data for high-fidelity, high performance reactor modelling and simulation

Benoit Forget

Massachusetts Institute of Technology
Department of Nuclear Science and Engineering
bforget@mit.edu

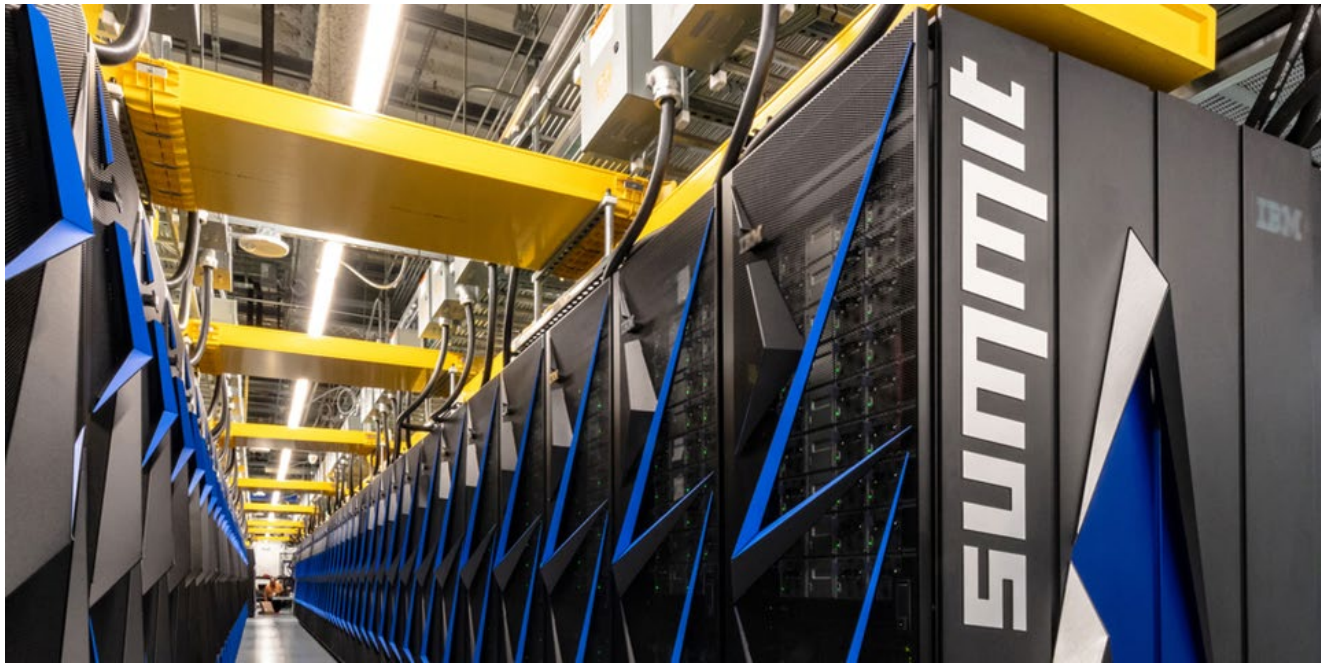
IAEA Workshop on
Computational Nuclear Science and Engineering
May 23, 2022

About Me

- Joined MIT in 2008
- Professor and Associate Department Head of Nuclear Science and Engineering
- Founded the Computational Reactor Physics Group
  - Major highlights:
    - OpenMC, an open source Monte Carlo code (lead developer Paul Romano)
    - OpenMOC, an open source Method of Characteristic transport code
    - BEAVRS benchmark, full core PWR with first 2 cycles of flux core mapping data

Our goal

- One of our primary goals has been the development of high-fidelity neutron transport methods for **full core nuclear reactor simulations**
  - Leverage high performance computing
  - Improve data representation
  - Reduce memory footprint
  - Develop novel algorithms for improved efficiency



Many of the roadblocks and bottlenecks identified for improving performance were tied to nuclear data

Outline

- **Part I: Nuclear data for high-fidelity Monte Carlo simulations**
  - **Nuclear Data Requirements**
  - **Nuclear Data Options**
  - **Limitations and Opportunities**

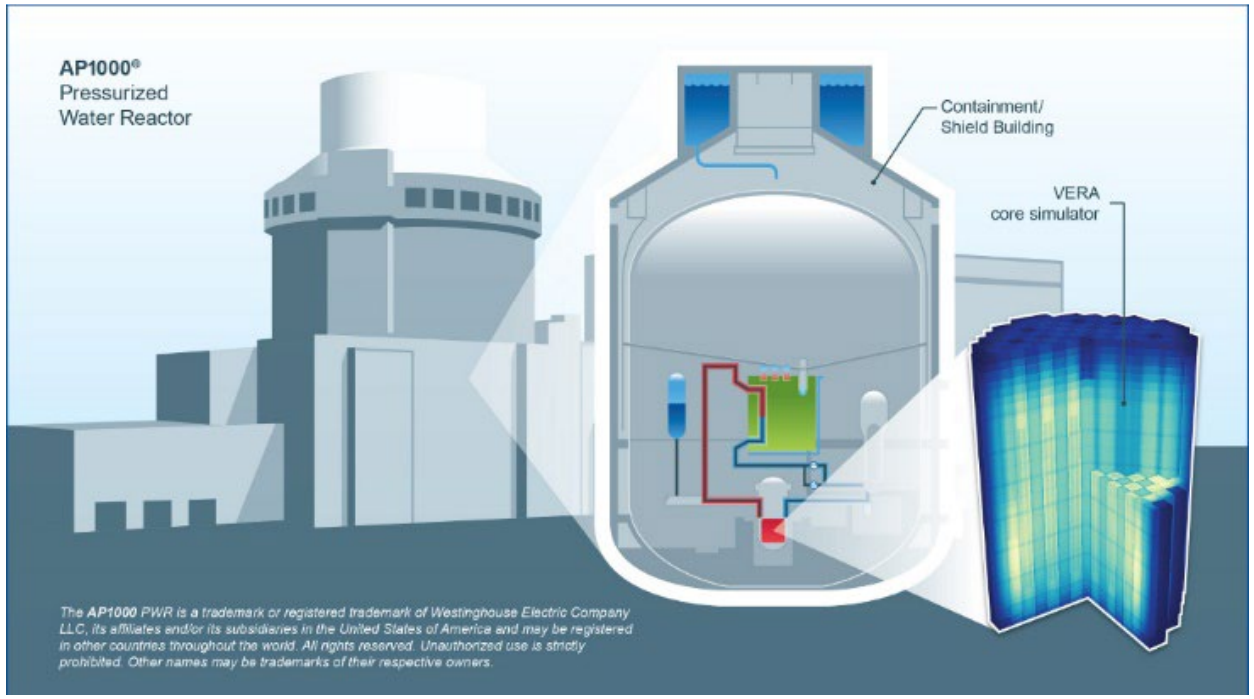
- Part II: Generating high-fidelity nuclear data for deterministic calculations
  - Transport cross-section
  - Equivalence Factors
  - Limitations and Opportunities

Assumptions

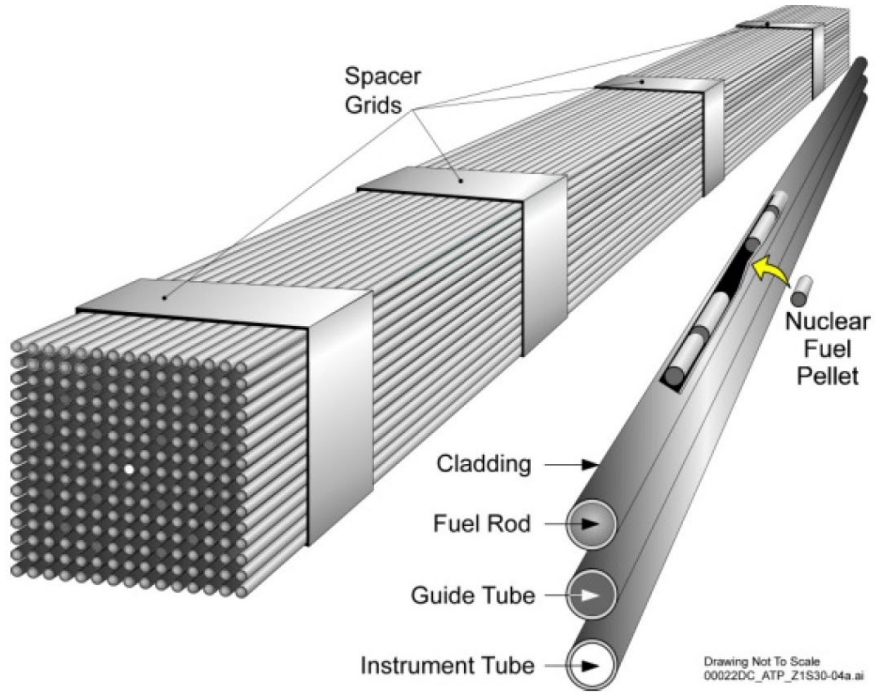
- When preparing this talk, I had to assume some level of knowledge, I thus assumed that most of you knew something about
  - Nuclear cross-sections
  - Neutron slowing down
  - Criticality
  - Multigroup cross-sections
  - Transport equation
  - Diffusion equation

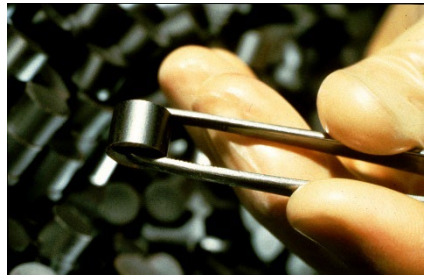
Part I: Nuclear data for high-fidelity Monte Carlo simulations

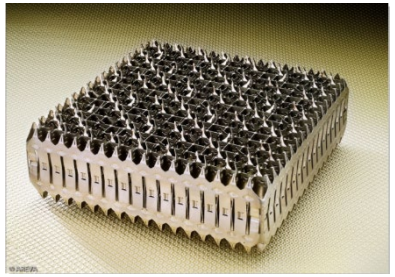
The Big Picture



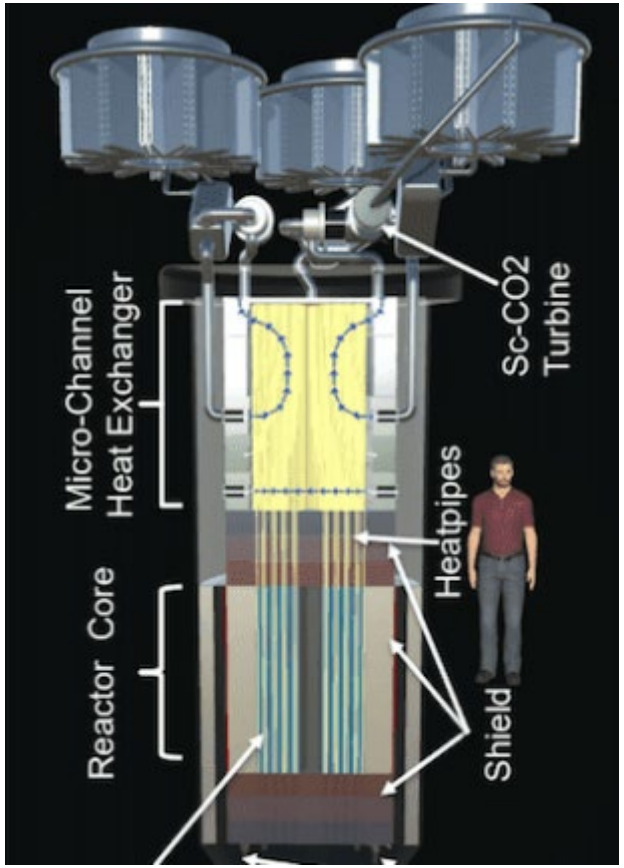
Oak Ridge National Laboratory, CASL News Release, 2014.



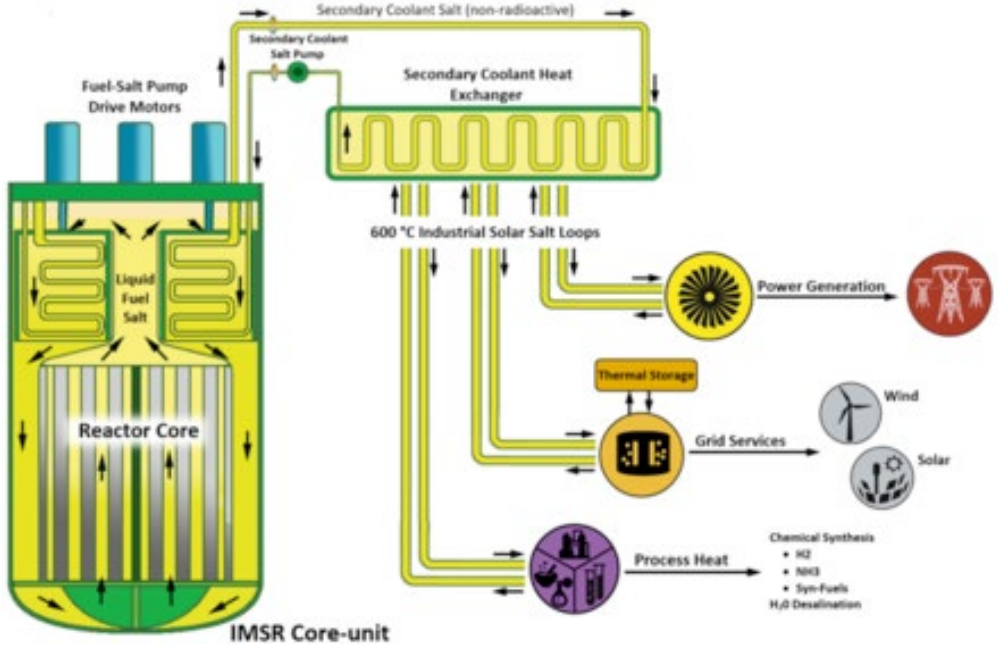




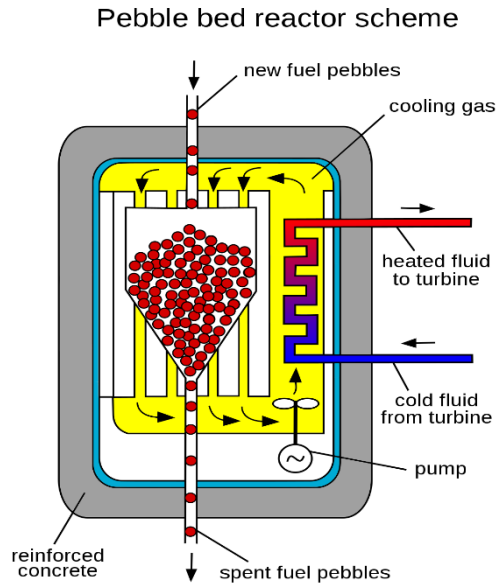
Beyond LWRs

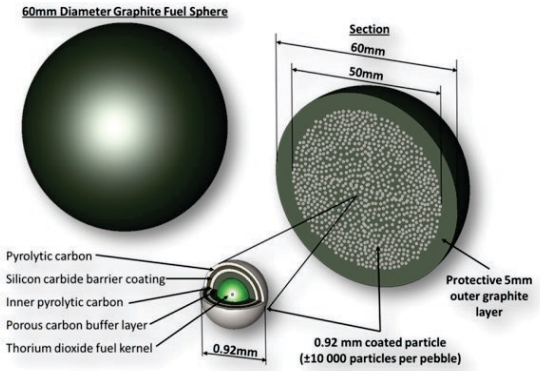


Microreactor



Molten Salt Reactor





Pebble Bed Reactor

Current state-of-the-art

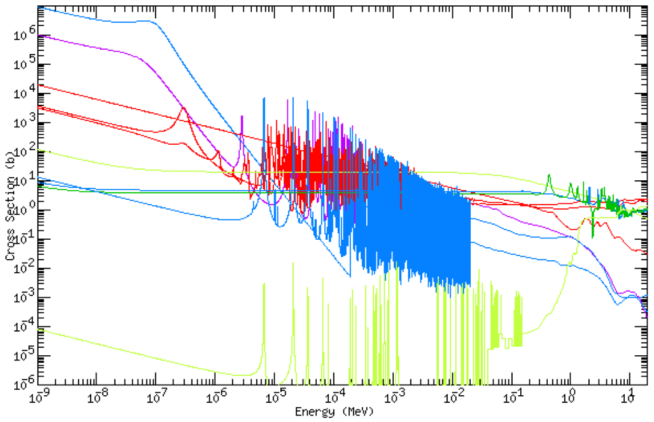
- Current methodologies rely on many-levels of approximation that have been extensively validated against experiments and operating nuclear fleet
  - Currently licensed methods are highly accurate for the current fleet of reactors
  - Most experiments were performed in the 60's and 70's

• New reactors promise much higher levels of heterogeneities.
• Experimental facilities in nuclear are increasingly costly and require very long lead times.

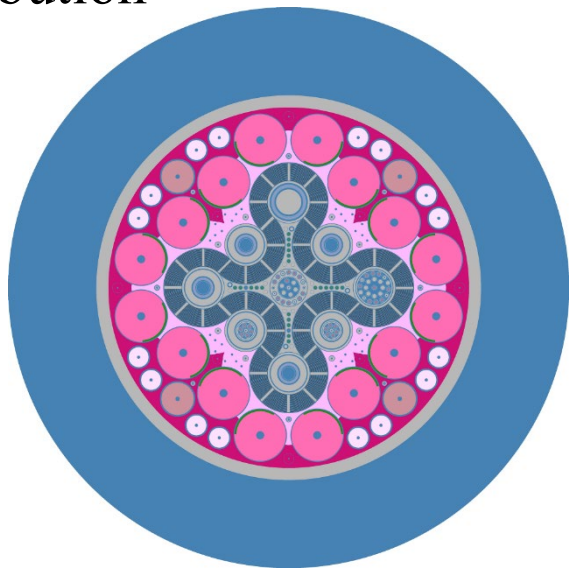
High-fidelity simulations are necessary to reduce the need for costly experiments for future nuclear reactor technologies

Why do we need high-fidelity Monte Carlo?

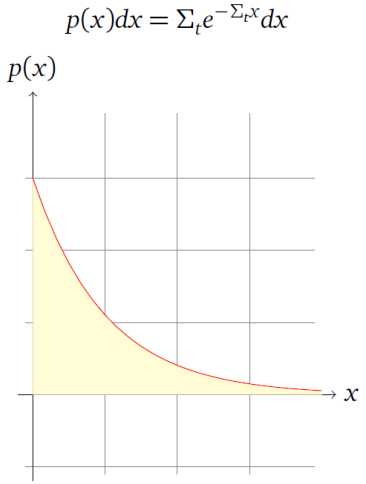
- Monte Carlo methods "faithfully" track neutrons through their lifetime
  - High-fidelity representation of the nuclear data
  - High-fidelity representation of the geometry
  - High-fidelity representation of the fission and scattering source distribution



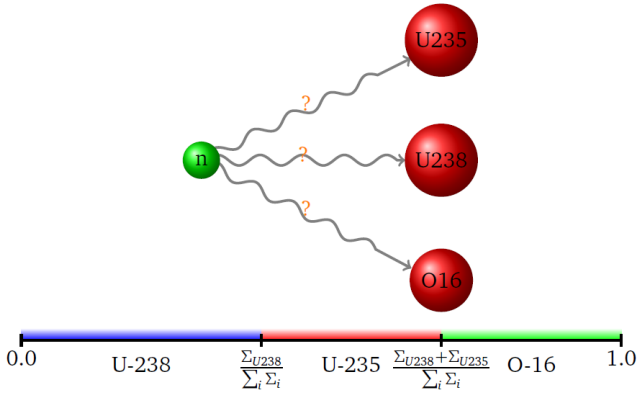
10 orders of magnitude in Energy



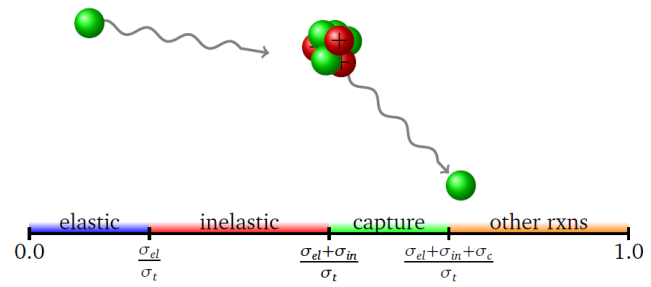
ATR Geometry in OpenMC



Path sampling



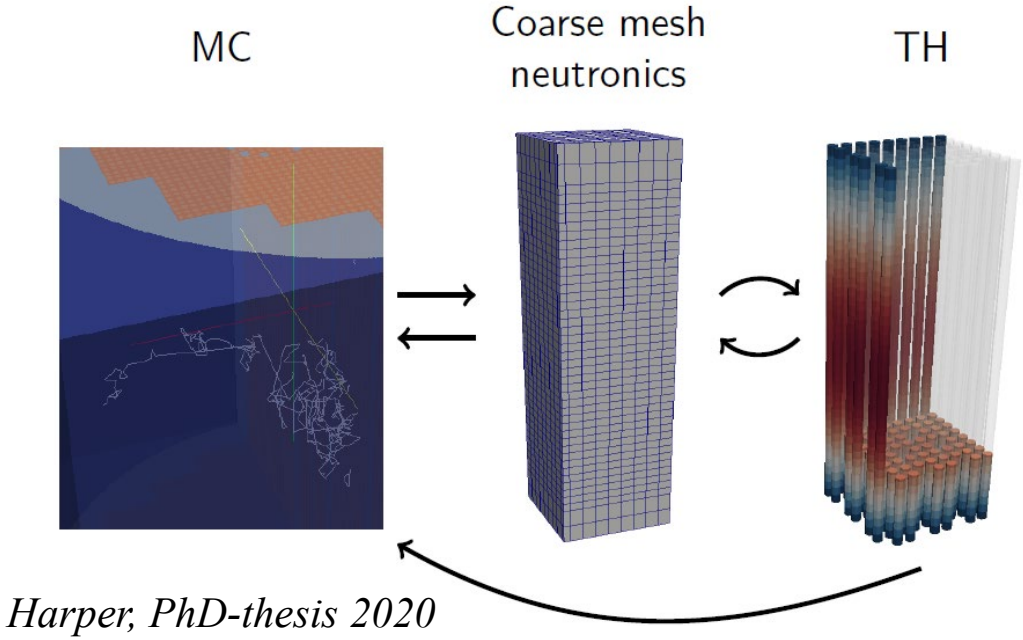
Nuclide sampling



Collision sampling

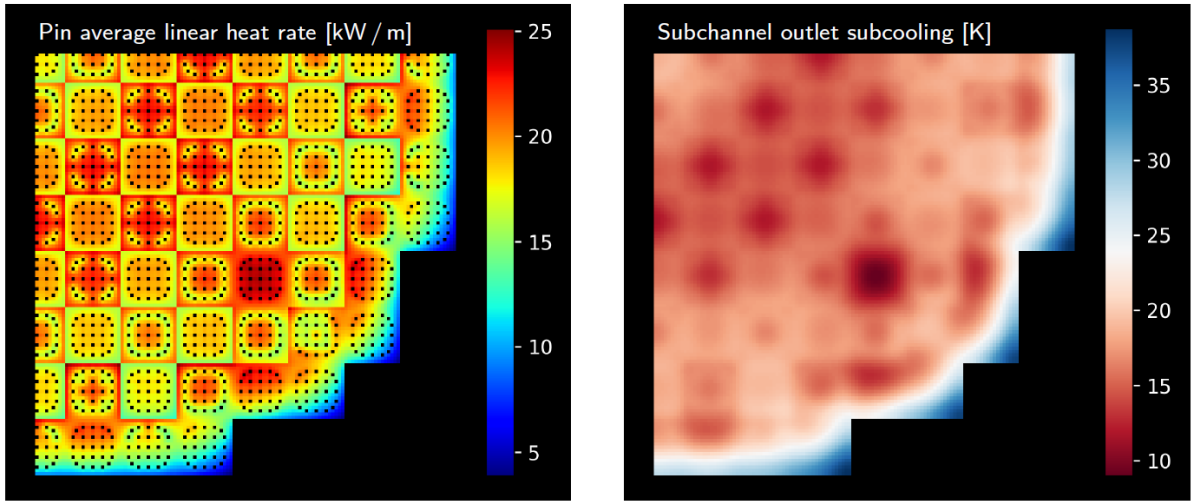
Multiphysics applications

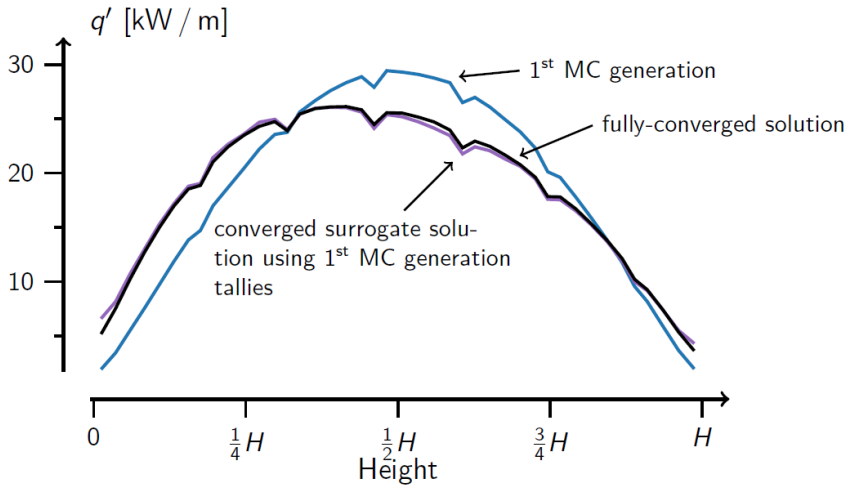
- High-fidelity simulations beyond benchmarking require at the very least thermohydraulic feedback
  - Example of 3D 1/4 PWR between OpenMC and subchannel
  - Power distribution shifts radially and axially based on temperature feedback



Harper, PhD-thesis 2020

$$\Sigma(T,\rho) \approx \Sigma(T_{\text{ref}},\rho_{\text{ref}}) + \frac{\partial\Sigma}{\partial T}(T - T_{\text{ref}}) + \frac{\partial\Sigma}{\partial\rho}(\rho - \rho_{\text{ref}})$$





Data Requirements (LWR Example)

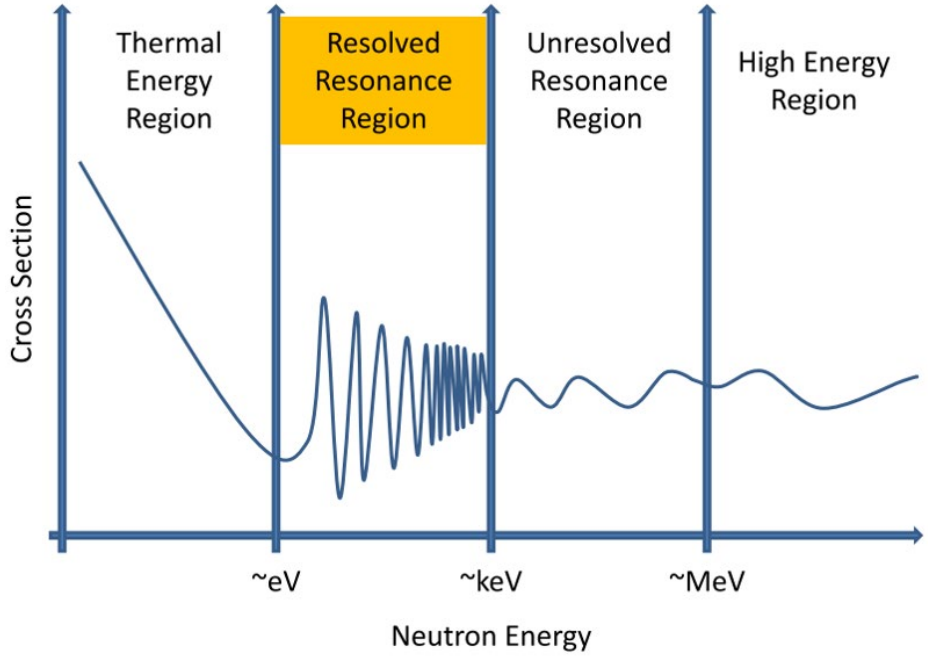
- Our goal is to predict the power in every single fuel pellet as a function of time
  - PWR has ~20,000,000 fuel pellets
  - Every pellet has a **different average temperature**
  - Every pellet has a **unique temperature profile**
  - Fuel transmutes over its lifetime (each pellet resides ~5 years in the core)
    - We must track ~6 different reactions that can occur in the fuel.
    - We must follow ~300 nuclides being consumed and produced in the fuel.

Each event requires accessing 100's of GBs of nuclear data!
Each time step requires ~3-5TB of data to be stored!

Detailed knowledge enables better fuel utilization, improves understanding of safety margins that can lead to a reduction of conservatism and improves predictability of the system.

Current status – Nuclear Data in MC

- Most Monte Carlo codes rely on pre-processed data stored in large tables to capture
  - Energy dependence
  - Angular dependence
  - Temperature
  - Probabilistic nature of the URR
  - Energy-angle correlation of the thermal range
- These data structures can become quite prohibitive for multi-physics calculations
  - On the order of 1GB per temperature point

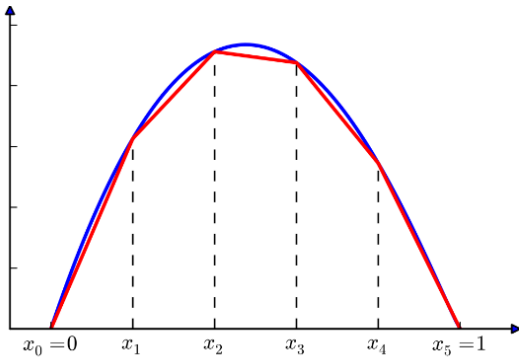


Our approach to bring the physics closer to the simulation and reduce the reliance on massive random access table lookups.

Temperature dependence of Resolved Resonance Range

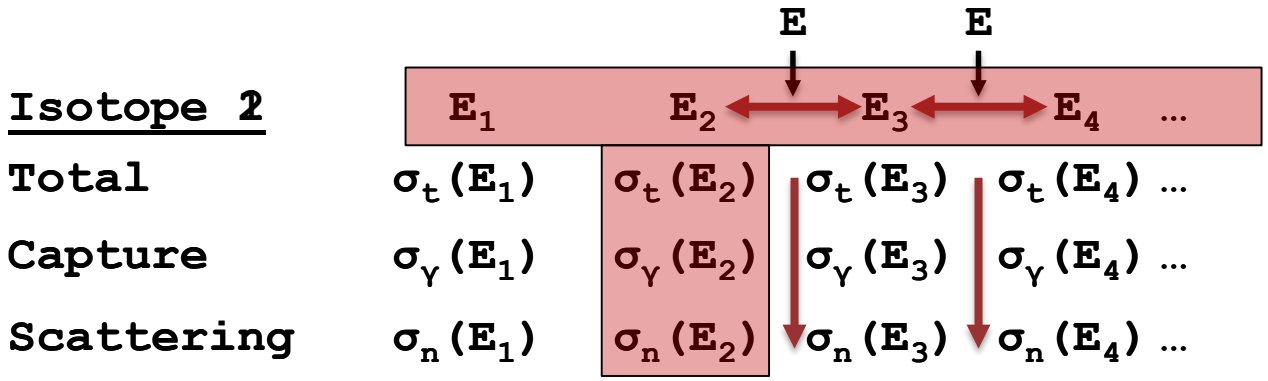
- Monte Carlo simulations are commonly used as reference calculations at fixed temperatures
- Cross sections are pre-generated at fixed temperatures using the BROADR (SIGMA1 algorithm) module of NJOY
- Nuclear data is commonly represented (in the resonance range) using models representative of the R-matrix theory (SLBW, MLBW, RM, …)
  - Requires only a few parameters per resonance (E_0,Γ, Γ_n, Γ_f …)
  - However, SIGMA1 requires a linearization of the data to perform the convolution integral

$$v\sigma(v,T) = \int d\mathbf{V}|\mathbf{v}-\mathbf{V}|\sigma_0(|\mathbf{v}-\mathbf{V}|)P(\mathbf{V},T)$$

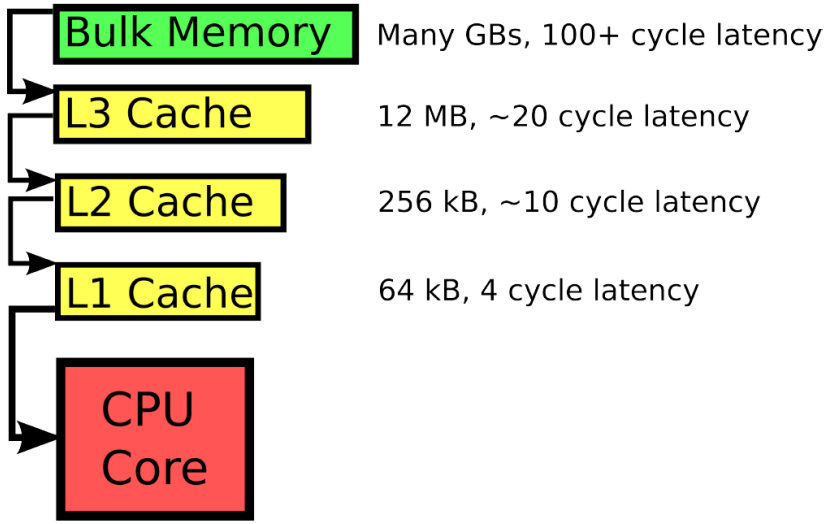


Nuclear Data Reconstruction

- Every isotope has its own energy grid for each temperature
  - Requires a binary search over 1000's – 100,000's of points



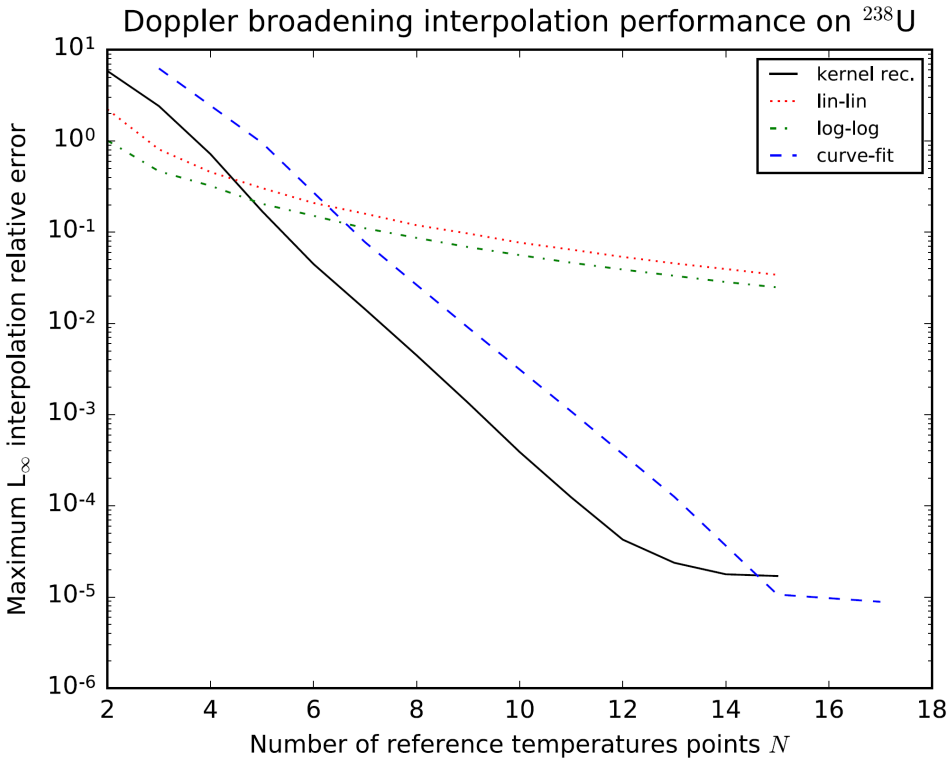
- Simple algorithmic fixes
  - Unionized grid over all temperatures and isotopes
    - Global or local (cell-based)
    - Serpent has an option for a global unionized grid
  - Hash table to accelerate search
    - MCNP6 and OpenMC use a hash table



The further you reach for data, the slower your code becomes!

Options in the Resolved Resonance Range

- SIGMA1
  - Linearize data and solve Solbrig kernel analytically over an energy band
  - Some are exploring off-loading this operation to GPU
- Stochastic Mixing
  - Randomly sample between bounding temperatures to mimic interpolation
  - Requires temperature spacings on the order of 10-50K for good accuracy
- Kernel Reconstruction
  - Reconstruct the Solbrig kernel effect using ~10 temperatures from which to randomly sample. Weights are determined analytically through an error minimization process.



Ducru et al, JCP, 2017

Options in the Resolved Resonance Range

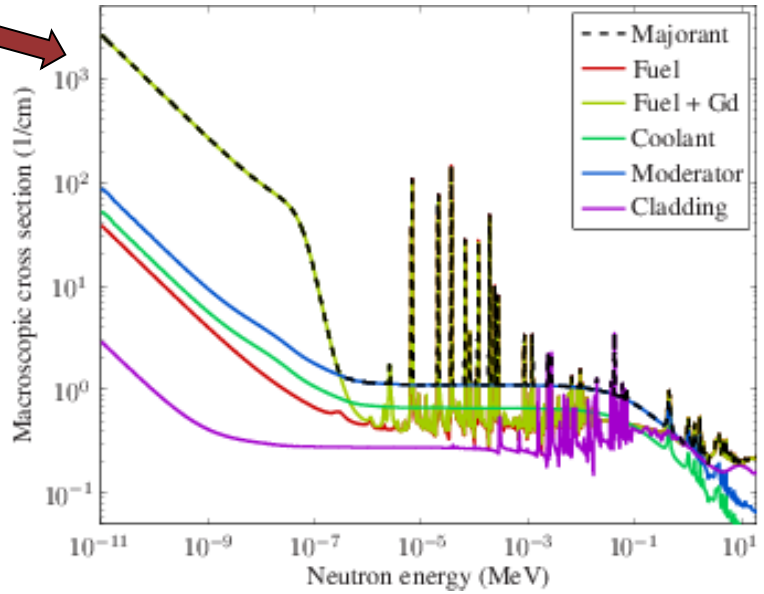
- Gauss-Hermite Quadrature
  - Replace convolution integral by a Gauss-Hermite quadrature

$$\int_{-\infty}^{\infty} dz e^{-z^2} f(z) \approx \sum_{k=1}^{N} w_k f(z_k)$$

- Polynomial fitting (e.g. MCNP)
  - High order fit across many temperatures

$$\sigma_\gamma(T, E_g) \approx \sum_{i=1}^{N} \frac{a_{g,i}}{T^{i/2}} + \sum_{i=1}^{N} b_{g,i} T^{i/2} + c_g$$

- Target Motion Sampling (e.g. SERPENT)
  - Sample target velocity at collision site and apply rejection sampling
- Windowed Multipole (e.g. OpenMC)
  - Transform the resonance parameters to an equivalent representation in complex space and perform convolution integral analytically
  - For performance benefits, broadening only performed over a surrounding energy window



Viitanen et al, NSE, 2012

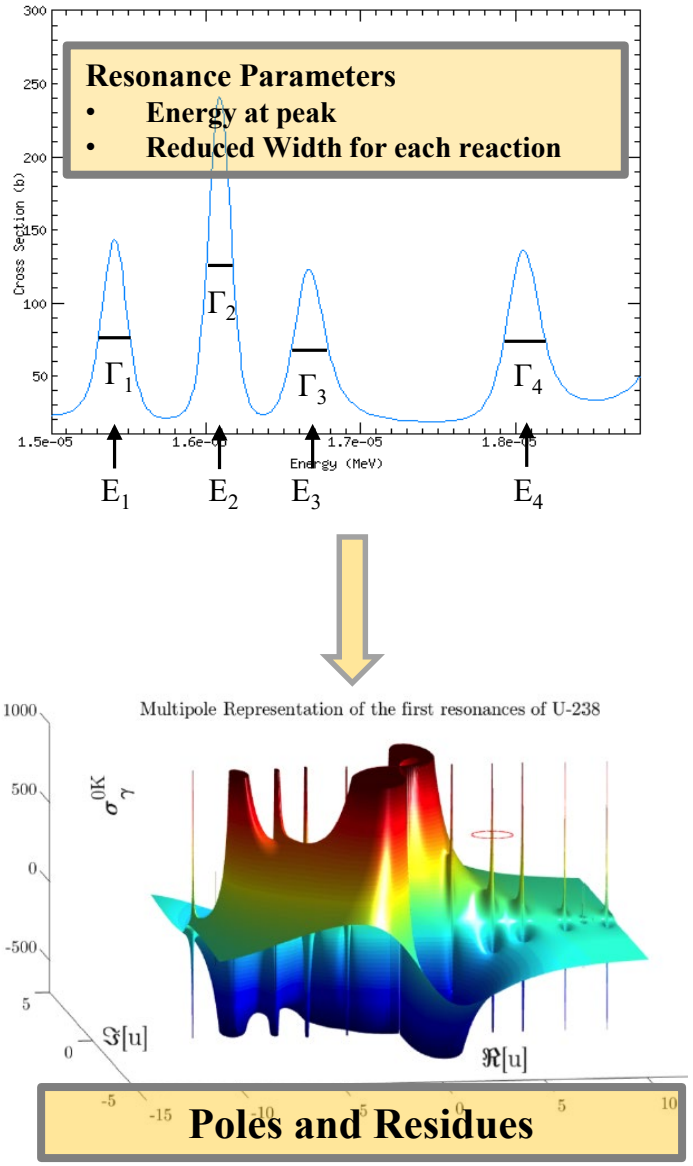
Multipole Formalism

- Developed by R.Hwang in 1987
  - Recognized that R-matrix formulation yielded a meromorphic function on which a partial fraction decomposition could be performed
  - *(E,Γ)* real parameters are transformed into *(p,r)* complex parameters

$$\sigma = \frac{1}{u^2}\sum_j \Re\left[\frac{r_j}{p_j - u}\right]$$

  - And the convolution integral yields

$$\sigma(u,T) = \mathfrak{D}\mathfrak{B}_{T_0}^{T}\left[\sigma(u,T_0)\right] = \frac{1}{2u^2\sqrt{\xi}}\sum_j \Re\left[ir_j\sqrt{\pi}W(z_j^0) - \frac{r_j}{\sqrt{\pi}}C\left(\frac{p_j}{\sqrt{\xi}}, \frac{u}{2\sqrt{\xi}}\right)\right]$$

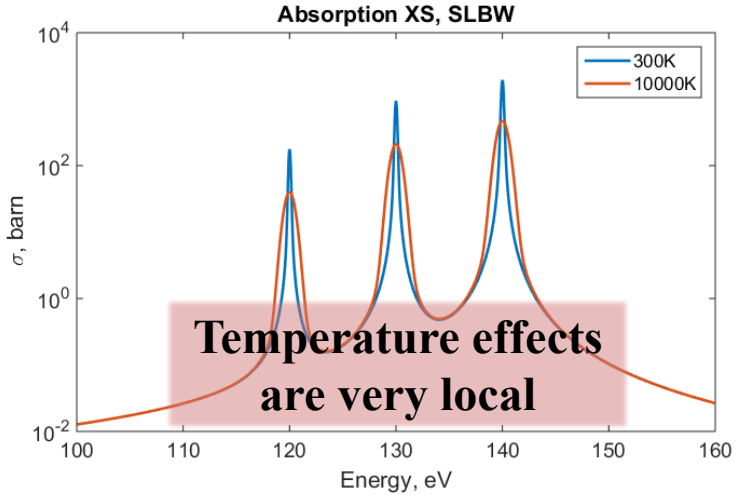


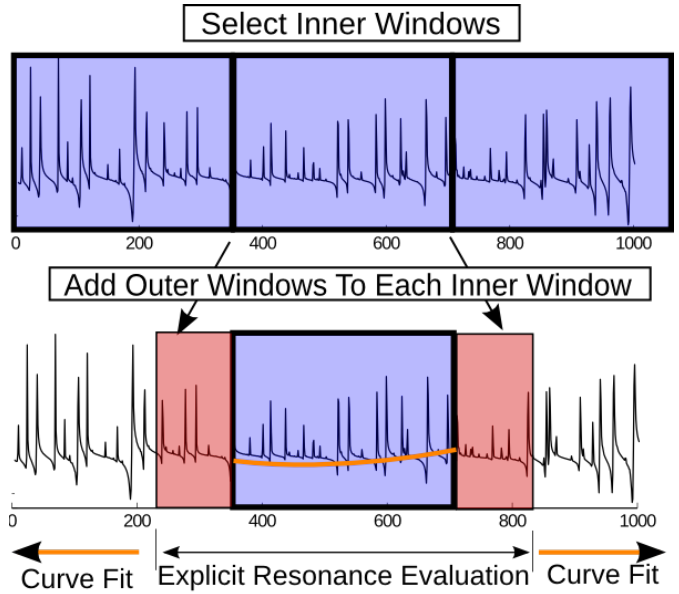
Josey and Ducru, JCP, 2016

Windowed multipole method

- Key observation was made that Doppler broadening effects are local
  - Far away resonances contribute to the local cross section but they exhibit little to no temperature dependence
- Windowed multipole method creates a system of inner windows and pointers to minimize the number of Faddeeva function evaluations
  - Far away resonances are fitted to a low order polynomial

$$\sigma(E) = \frac{1}{E} \sum_{j \in \mathcal{W}(E)} \mathrm{Re}\left[\frac{-ir_j}{p_j - \sqrt{E}}\right] + \sum_{n \geq -2}^{N} a_n (\sqrt{E})^n$$





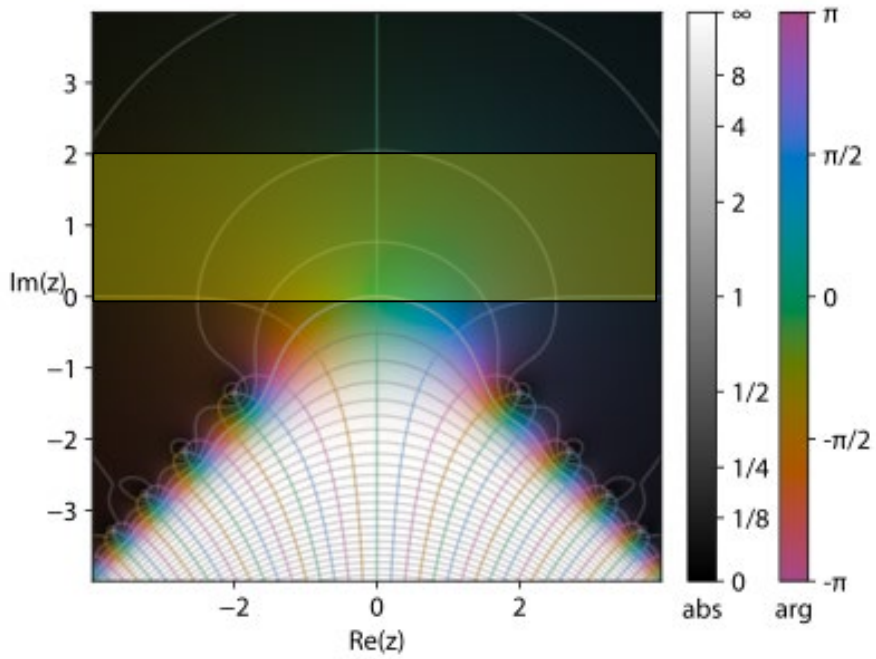
For any point in the blue inner window

Josey and Ducru, JCP, 2016

Faddeeva function

$$\sigma(E) = \frac{1}{E}\sum_{j\in\mathcal{W}(E)} \Re\left[ir_j\sqrt{\pi}W(z)\right] + \sum_{n\geq -2}^{N} a_n(\sqrt{E})^n$$

$$z = \frac{\sqrt{E}-p_j}{2\sqrt{\xi}} \qquad \xi = \frac{k_b T}{4A} \qquad W(z) = e^{-z^2}\mathrm{erfc}(-iz)$$



- Recently developed a low order rational approximation highly accurate in range of interest
  - 8th order approximation
  - No branching (i.e. no if-else statements)

$$w(z) = \frac{\sum_{k=0}^{n-1} a_k \tilde{z}^k}{\sum_{l=0}^{n} b_l \tilde{z}^l}$$

Forget et al, PHYSOR 2022

	Single Temperature Particles/s	Overhead	Multi Temperature Particles/s	Overhead
ACE	176329	-	163968[1]	-
Original	158939	9.8%	152927	6.8%
8th order	172303	2.2%	163476	0.3%

[1] Stochastic interpolation

Options in the Resolved Resonance Range

- SIGMA1
- Stochastic Mixing
- Kernel Reconstruction
- Gauss-Hermite Quadrature
- Polynomial fitting
- Target Motion Sampling
- Windowed Multipole
  - Additionally, WMP allows for arbitrary order analytical derivatives in T (*Ducru et al, 2022*)

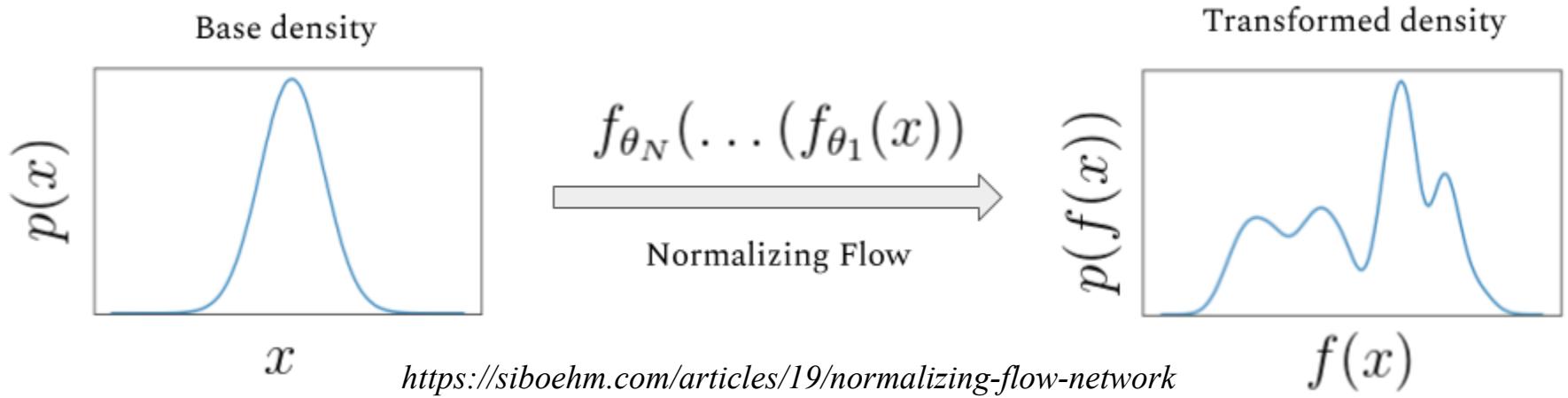
Method	Memory*	Efficiency*
Single Temp. ACE	1	1
Stochastic Mixing / Interpolation	~100	~1.5-2
Kernel Reconstruction	**~10**	**~2**
Gauss-Hermite Quadrature	~1	~10-15
Polynomial Fitting	~20	~1.1-1.3
Target Motion Sampling	~2	~4-10
Windowed Multipole	**~0.5-0.7**	**~1.0-1.1**

* Estimated by the lecturer (lower is better)

Opportunities

- Exposing the codes to more physics facilitates the integration of UQ methodologies with direct feedback on evaluations
  - Can we embed nuclear data uncertainty in a Monte Carlo simulations?
  - Can we provide valuable feedback to evaluators on where larger source of uncertainties are coming from?
- Neural network representations of complex data structures
  - Can advancements in data sciences provide a new path to data representation that can be both accurate and efficient?
- Modern computing architecture
  - Can we leverage power of GPU architectures to enable large steady-state and transient simulations?

Normalizing Flows for Thermal Scattering



https://siboehm.com/articles/19/normalizing-flow-network

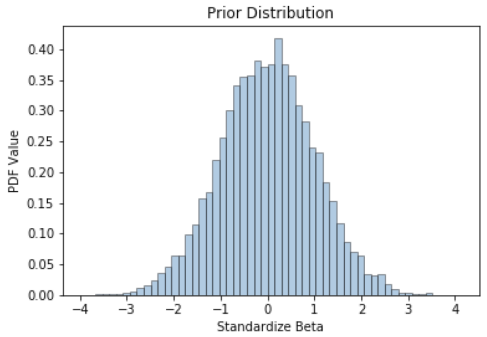
- Normalizing flows map a simple distribution function through a series of transformations to a complex distributions function
  - Complex distributions is represented as a spline where the coefficients and transformation are learned.
  - The process can be inverted such that one can randomly sample a point from the simple distribution and map this sample to the complex distribution.

This was applied to thermal neutron scattering!

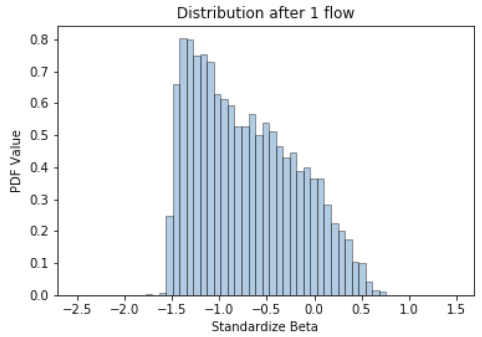
Results looked promising!

Forget, Alhajri, ANE, 2022

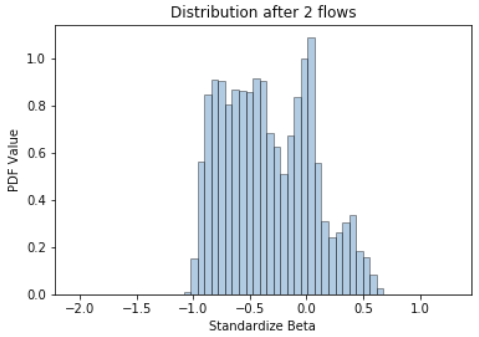
- A normal distribution in beta and alpha was mapped to the true thermal scattering kernel using 3 flows (i.e. 3 transformations)



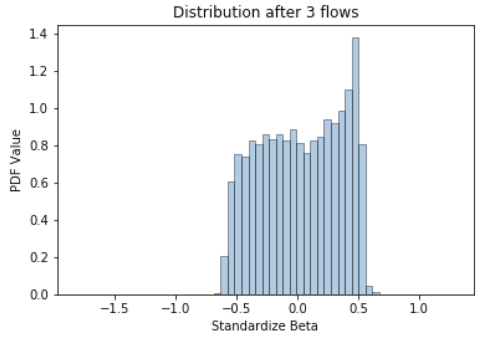
(a) Prior Distribution



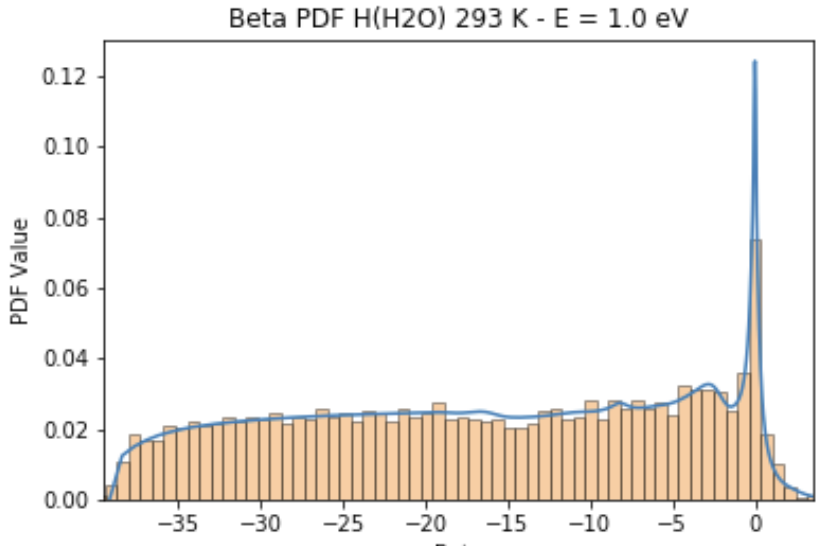
(b) After 1 flow

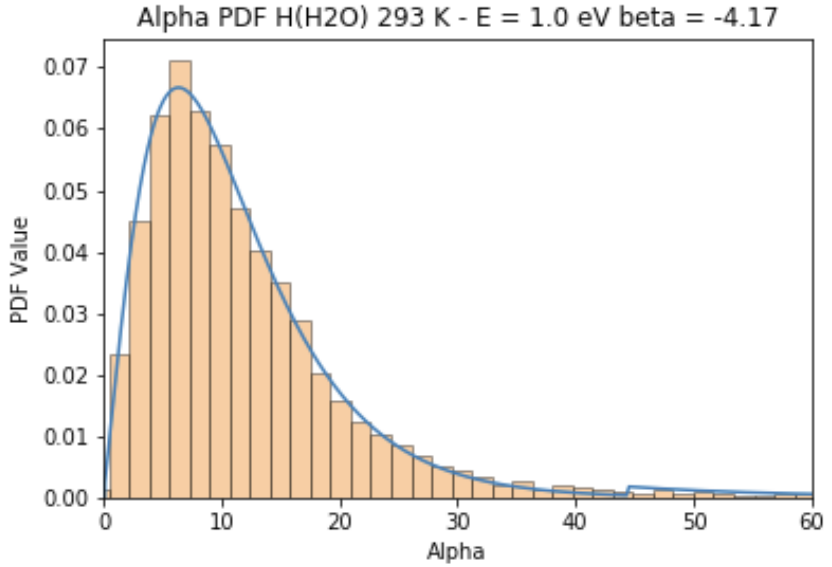


(c) After 2 flows



(d) After 3 flows





But performance was another story!

- The neural net to store the spline coefficients and transformations was quite small.
  - ~400 kB
- However, every query to sample an outgoing energy and angle requires accessing a large part of the memory.
  - Neural nets perform well when multiple samples are needed
  - Traditional MC algorithms follow neutrons from birth to death in a sequential manner, thus only 1 sample is needed at each collision

Table 1: Average Computational Cost per Sample

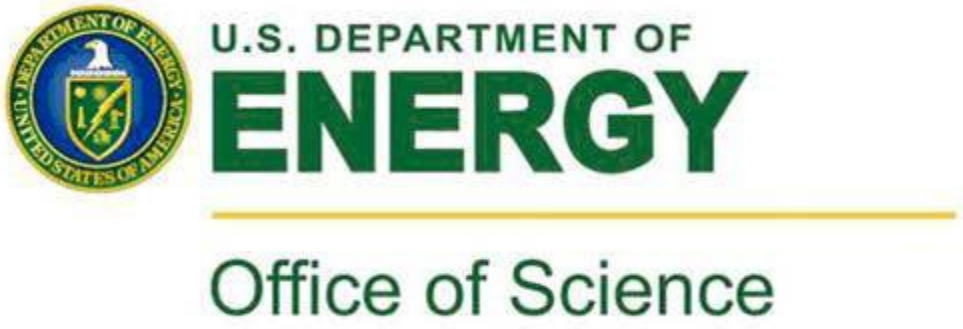
Samples	Avg. Time per sample (ms)
1	15
10	2.8
50	0.68
100	0.38
1000	0.12
10000	0.09

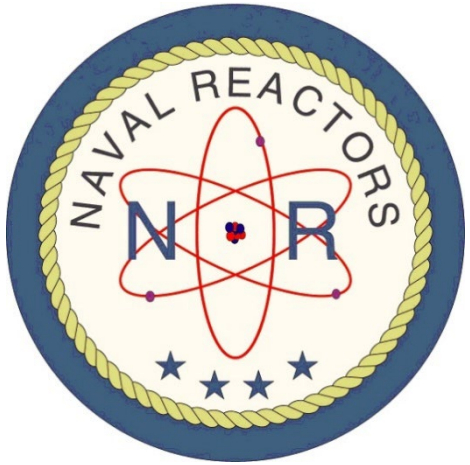
Still potential for GPU-based algorithms that prefer Event based MC!

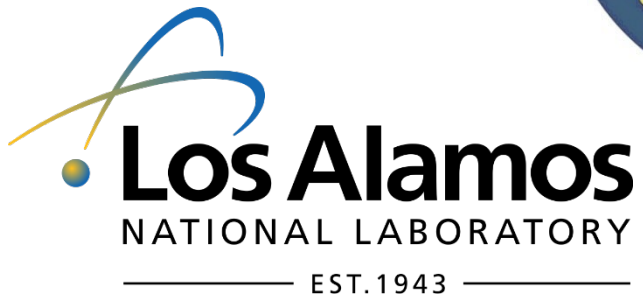
Summary - Stochastic

- High fidelity Monte Carlo simulations require large amounts of nuclear data, especially in coupled simulations where temperature must be accounted for
  - Many techniques exist that can accurately capture the temperature effects in the resolved resonance range
- Random access of nuclear data can hinder performance on modern computing architectures
- By default, most general Monte Carlo simulations tools still neglect some important temperature phenomena
  - Resonance upscattering
  - Thermal scattering

Acknowledgements







References

- S.Harper,Tally derivative based surrogate models for faster Monte Carlo Multiphysics, MIT, 2020.
- C.Josey, P.Ducru, B.Forget, K.Smith, Windowed Multipole for Cross-Section Doppler Broadening, JCP, 307, 2016.
- B.Forget, J.Yu, G.Ridley, Performance Improvements of the Windowed Multipole Formalism Using a Rational Fraction Approximation of the Faddeeva Function, International Conference on Physics of Reactors – PHYSOR, 2022.
- A.Trainer, B.Forget, Development of a Phonon-Based Sampling Method for Thermal Neutron Scattering Data, International Conference on Physics of Reactors – PHYSOR, 2022.
- B. Forget, A.Alhajri, Normalizing flows for thermal scattering sampling, Annals of Nuclear Energy, 170, 2022.

Questions?