Bayesian Optimisation using microscopic and integral measurements to infer nuclear data parameters

Daan Houben, Mathieu Hursin

December 8, 2025

Abstract

test 123

1 Introduction

Nuclear data is considered as the major source of uncertainty in some reactor observables, most notably effective multiplication factor (k_{eff}). The nuclear data available in evaluated nuclear data libraries, such as cross sections, neutron multiplicities, angular distributions or fission neutron energy spectrum, are a result of a complex fitting procedure including theoretical models, microscopic experiments and expert judgement. Integral experiments are then used to assess the performance of the nuclear data. In this work, a Bayesian Optimization (BO) framework is proposed which enables consolidating microscopic energy dependent measurements with integral experiments to estimate nuclear data parameters.

The BO is performed using a Markov Chain Monte Carlo (MCMC) method in which surrogates are used to evaluate the likelihoods. For the microscopic energy dependent measurements, the SAMMY v8.1.0 resonance fitting tool [1] is employed. While SERPENT v2.2.2 [2], a Monte Carlo neutron transport code, is used to quantify the integral response. Surrogates are trained by evaluating random samples drawn in the input space in the high-fidelity models. The methodology is tested on two case studies, being ^{53}Cr and ^{238}U. Since microscopic experiments often provide many points and integral experiments only provide a single points, care is given to analyze how different assumptions regarding the likelihood evaluation affect the posterior.

2 Background and mathematical motivation

2.1 Bayesian Optimisation setup

The main objective of this paper is to infer nuclear data parameter(s) from a set of both microscopic energy dependent an integral experiments. Microscopic energy dependent measurements, further noted as microscopic measurements, are measurements of single energy neutron properties. Often such measurements result from neutron Time-Of-Flight (nTOF) facilities in which the energy of the neutron is derived from the time it takes for the neutron to reach a target. Typical of these measurements is the many measurement points that are obtained. In contrast, integral measurements such as criticality experiments, only provide one value which is representative of a group of nuclides, reactions and energies.

According to Bayes theorem, the posterior (updated) probability density finction (PDF), $P(\theta|\mathrm{data})$, is proportional to the likelihood of observing the parameter(s) θ given the data multiplied by the prior belief of the parameter(s):

$$P(\theta|\mathrm{data}) \propto P(\mathrm{data}|\theta) \cdot P(\theta) \tag{1}$$

For brevity, we shall refer to the likelihood as $\mathcal{L}(\theta) = P(\mathrm{data}|\theta)$.

2.2 Markov Chain Monte Carlo (MCMC)

To calculate the posterior distribution, various techniques can be derived from Bayes theorem. Popular techniques include Generalized Linear Least Squares (GLLS), Bayesian Monte Carlo (BMC), MOCABA, etc. In this paper, an algorithm belonging to the family of Markov Chain Monte Carlo (MCMC) techniques is selected. As MCMC is a name for a family of techniques, first the general principle is explained, after which the Emcee algorithm used in this paper is detailed.

In most MCMC algorithms, the unnormalized posterior is evaluated to generate samples:

$$P^*(\theta|\text{data}) = \mathcal{L}(\theta) \cdot P(\theta) \tag{2}$$

where $P^*(\theta|\text{data})$ represents the posterior up to a normalizing constant. The objective is to construct a Markov chain $\{\theta_0, \theta_1, \ldots, \theta_N\}$ such that the stationary distribution of the chain converges to the posterior distribution $P(\theta|\text{data})$. Under the assumption of Normally distributed prior ($\mathcal{N}(\theta_0, \Sigma)$) and likelihood ($\mathcal{N}(\mathbf{y}_{\text{true}}, \mathbf{A})$), the unnormalized posterior probability evaluated at θ can be rewritten as

$$P^*(\theta|\text{data}) = \frac{1}{\sqrt{(2\pi)^m \det \mathbf{A}}} \exp\left[-\frac{1}{2}(\mathbf{f}(\theta) - \mathbf{y}_{\text{exp}})^T \mathbf{A}^{-1}(\mathbf{f}(\theta) - \mathbf{y}_{\text{exp}})\right] \times$$
$$\frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left[-\frac{1}{2}(\theta_0 - \theta)^T \Sigma^{-1}(\theta_0 - \theta)\right] \tag{3}$$

in which the first term represents the probability of observing θ given the measurements, with $\mathbf{y}_{\text{exp}}$ is the vector describing the m measurement points, $\mathbf{f}(\theta)$ is the vector containing the model responses for the vector θ and $\mathbf{A}$ is the covariance matrix describing the measurement points of size $(m \times m)$. The prior probability evaluated at θ can then be calculated using the prior belief of the n parameters θ_0, with covariance matrix Σ of size $(n \times n)$.

Standard algorithms, such as Metropolis-Hastings, propose a new state θ' based on a proposal distribution $q(\theta'|\theta_t)$ and accept it with probability $\alpha = \min(1, \frac{P^*(\theta')q(\theta_t|\theta')}{P^*(\theta_t)q(\theta'|\theta_t)})$. However, in high-dimensional nuclear data spaces where parameters may be highly correlated, standard proposal distributions often result in poor mixing.

For ease of tuning and implementation as well as future proofing, we employ the Affine Invariant Ensemble Sampler (AIES), as implemented in the *emcee* code [3]. In this algorithm, an ensemble of K "walkers" are propagated in parallel. The proposal step for a walker θ_k is based on the current position of a complementary walker θ_j from the ensemble:

$$\theta'_k = \theta_j + Z(\theta_k - \theta_j) \tag{4}$$

where Z is a scaling variable drawn from a distribution $g(z) \propto 1/\sqrt{z}$ on the interval $[1/a, a]$. This "stretch move" allows the algorithm to efficiently sample distributions with strong correlations without requiring manual tuning of the covariance matrix of the proposal distribution, making it straightfoward to implement.

2.3 Likelihood Formulation

The formulation of the likelihood function $\mathcal{L}(\theta)$ is one of the main challenges when combining integral and microscopic data, as the magnitude of data points differs by orders of magnitude, hereby possibly diluting the effect of the integral measurement. To try and understand better to which extend the diluting effect of the microscopic measurements affects the useability of the integral measurements, different possibilities to approach the covariance matrix of the measurement points are analyzed. First and foremost, the correct and most rigour method is to use the real covariance matrix describing the measurement points. However, this is often seen as too complex. Especially, calculating experimental correlations between integral experiements

and inbetween microscopic and integral experiments. In this regard, the first assumoption comes into play, i.e., there are no correlations between microscopic and integral experiments and no correlations between integral experiments. This assumption may not be valid when integral experiments from the same facility are used, therefore, we limit our analysis by only regarding integral experiments from different facilities. This assumption should be sufficiently valid regarding the correlations due to nuclear data will dominate. Now, it is possible to calculate the total likelihood $\mathcal{L}(\theta)$ for θ for the set of microscopic experiments J and integral experiments I using

$$\mathcal{L}(\theta) = \prod_{j \in J} \frac{1}{\sqrt{(2\pi)^{m_j} \det \mathbf{A_j}}} \exp\left[-\frac{1}{2}(\mathbf{f_j}(\theta) - \mathbf{y_j})^T \mathbf{A_j}^{-1}(\mathbf{f_j}(\theta) - \mathbf{y_j})\right] \times \\ \prod_{i \in I} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(f_i(\theta) - y_i)^2}{2\sigma_i^2}\right] \tag{5}$$

As probabilities are often very small numbers, the natural logarithm is taken to preserve numerical stability. The log-likelihood $\ln \mathcal{L}(\theta)$ then becomes

$$\ln \mathcal{L}(\theta) = -\frac{1}{2}\sum_{j \in J} \ln\left[(2\pi)^{m_j} \det \mathbf{A_j}\right] + -\frac{1}{2}\sum_{j \in J}\left[(\mathbf{f_j}(\theta) - \mathbf{y_j})^T \mathbf{A_j}^{-1}(\mathbf{f_j}(\theta) - \mathbf{y_j})\right] + \\ -\frac{1}{2}\sum_{i \in I} \ln\left[2\pi\sigma_i^2\right] + -\frac{1}{2}\sum_{i \in I} \frac{(f_i(\theta) - y_i)^2}{\sigma_i^2} \tag{6}$$

Now two approximations are analyzed, i.e. (1) all microscopic measurement points behave independently and (2) the correlation between all microscopic measurements points of a single experiment is 1 and there is no experimental correlation between the experiments.

2.3.1 Independent microscopic measurement points

When all measurement points behave independently, each datapoint has the same weight as an integral experiment. In this sense, integral experiments will be diluted and their influence on the posterior is expected to be negligible. Nevertheless, the loglikelihood for independ microscopic experiments therefore becomes

$$\ln \mathcal{L}_J(\theta) = -\frac{1}{2}\sum_{j \in J}\sum_{e \in E_j} \ln\left[2\pi\sigma_{j,e}^2\right] + -\frac{1}{2}\sum_{j \in J}\sum_{e \in E_j}\left[\frac{(f_{j,e}(\theta) - y_{j,e})^2}{\sigma_{j,e}^2}\right] \\ = C - \frac{1}{2}\sum_{j \in J} \chi_j^2 \tag{7}$$

in which $\chi^2 = \sum \frac{(y_c - y_e)^2}{\sigma_e^2}$ is the chi-squared goodness-of-fit for the datapoints resulting from an experiment. Since the term $-\frac{1}{2}\sum_{j \in J}\sum_{e \in E_j} \ln\left[2\pi\sigma_{j,e}^2\right]$ does not depend on θ, it can be calculate beforehand and replaced by a constant C.

2.3.2 Fully correlated microscopic measurement points

Now the opposite is assumed and all microscopic measurement points are expected to be highly correlated. Then, we take the chi-squared per degree of freedom to replace

$$
\begin{aligned}
\ln \mathcal{L}_J(\theta) &= -\frac{1}{2} \sum_{j \in J} \frac{1}{N} \sum_{e \in E_j} \ln \left[2\pi\sigma_{j,e}^2\right] + -\frac{1}{2} \sum_{j \in J} \frac{1}{N} \sum_{e \in E_j} \left[\frac{(f_{j,e}(\theta) - y_{j,e})^2}{\sigma_{j,e}^2}\right] \\
&= -\frac{1}{2} \sum_{j \in J} 2\pi\bar{\sigma}_j{}^2 - \frac{1}{2} \sum_{j \in J} \chi_j^2/N \\
&= C - \frac{1}{2} \sum_{j \in J} \chi_{N,j}^2
\end{aligned}
\tag{8}
$$

here, $\bar{\sigma}_j$ is the average experimental uncertainty and $\chi_{N,j}^2$ is the chi-squared per degree of freedom for microscopic experiment j.

2.4 Surrogate modelling

Direct coupling of SAMMY and SERPENT within the MCMC loop is computationally prohibitive, as MCMC chains often require 10^4 to 10^6 evaluations to converge. Therefore, we construct Stochastic Spectral Surrogates (or Polynomial Chaos Expansions) to map the input nuclear data parameters $\theta \in \mathbb{R}^d$ to the observables of interest.

Let $f(\theta)$ be the high-fidelity code response (either SAMMY or SERPENT). We approximate $f(\theta)$ by a truncated expansion:

$$
\hat{f}(\theta) = \sum_{\alpha \in \mathcal{A}} c_\alpha \Psi_\alpha(\theta)
\tag{9}
$$

where Ψ_α are multivariate orthogonal polynomials (e.g., Legendre polynomials for uniform priors) corresponding to the multi-index α, and c_α are the coefficients determined via regression on a training set.

The training set $\{\theta^{(i)}, f(\theta^{(i)})\}_{i=1}^{N_{train}}$ is generated using Latin Hypercube Sampling (LHS) to ensure optimal space-filling properties. The total uncertainty matrix used in the likelihood evaluation, $\mathbf{C}_{tot}$, is then updated to include the interpolation error of the surrogate:

$$
\mathbf{C}_{tot} = \mathbf{C}_{exp} + \mathbf{C}_{code} + \mathbf{C}_{surrogate}
\tag{10}
$$

This ensures that the loss of precision due to the surrogate model is propagated into the final posterior width of the nuclear data parameters.

3 Results

3.1 Chromium-53

3.2 Uranium-238

4 Discussion

5 Conclusions and future work

References

[1] N. M. Larson, "Updated User's Guide for Sammy: Multilevel R-Matrix Fits to Neutron Data Using Bayes' Equations."

[2] J. Leppänen, M. Pusa, T. Viitanen, V. Valtavirta, and T. Kaltiaisenaho, "The Serpent Monte Carlo code: Status, development and applications in 2013," *Annals of Nuclear Energy*, vol. 82, pp. 142–150, Aug. 2015.

[3] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "¡tt¿emcee¡/tt¿: The mcmc hammer," *Publications of the Astronomical Society of the Pacific*, vol. 125, p. 306–312, Mar. 2013.