Nuclear Instruments and Methods in Physics Research A 743 (2014) 39–43

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Challenges and solutions for random sampling of parameters with extremely large uncertainties and analysis of the ^{232}Th resonance covariances

Gašper Žerovnik [a,*], Andrej Trkov [a], Luiz C. Leal [b]

[a] Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[b] Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6354, USA

ARTICLE INFO

Article history:
Received 15 October 2013
Received in revised form
8 January 2014
Accepted 8 January 2014
Available online 18 January 2014

Keywords:
Resonance integral
Self-shielding
Resonance parameter
Covariance matrix
Random sampling

ABSTRACT

Covariance data in the existing evaluated nuclear data libraries often include large relative uncertainties and mathematical inconsistencies, which arise especially in combination with random sampling. The ^{232}Th evaluation from the ENDF/B-VII.1 library has been taken as an example. Possible solutions for mathematically impossible correlation matrices with negative eigenvalues and too low correlation coefficients between inherently positive parameters with large relative uncertainties are proposed. Convergence of the random sampling for lognormal distribution with extremely high relative standard deviations is slow by nature. Using weighted sampling, single parameters or a limited number of correlated parameters with large uncertainties can be sampled. Efficient sampling of a large number of correlated parameters with extremely large relative uncertainties remains unsolved.

1. Introduction

In contemporary evaluated nuclear data libraries (e.g. ENDF/B-VII.1 [1]) covariance data are included for the most materials, important for nuclear applications, and physical parameters. In principle, in cases when mean values and covariance matrices represent the only available information about a group of coupled parameters, multivariate normal or lognormal distributions should be assigned for location or scale parameters, respectively [2]. Due to its simplicity and other favourable properties, normal distribution is most commonly adopted for uncertainty propagation, either deterministic or probabilistic. However, in case that physical parameters such as cross-sections are inherently positive, and difficulties can be encountered when randomly sampling these parameters if normal distributions are assumed and the uncertainties are large [3,4]. Furthermore, it has been pointed out recently [5] that the correlation coefficients corresponding to multivariate normal and lognormal distributions are not equivalent. As opposed to the normal distribution, there are certain limitations for correlation coefficients for lognormal distribution due to lack of both shape invariance and symmetry. Methods for random sampling according to the multivariate lognormal distribution have been developed [5,6]. These methods are analytically exact, they circumvent the fundamental issues raised above and enable consistent random sampling of correlated inherently positive parameters. However, in very unfavourable conditions, like parameters with extremely large relative uncertainties (e.g. 1000%) or a badly conditioned covariance/correlation matrix, due to purely numerical reasons efficiency of these random sampling methods might be compromised, both in terms of accuracy and performance, i.e. the speed of convergence.

Such conditions are simulated by a realistic reference case. Resonance covariance data for ^{232}Th from the ENDF/B-VII.1 library [1] have been analyzed. Several points regarding the consistency of the covariance matrix and its usage for estimating the uncertainty of integral parameters are addressed in this paper.

2. Methodology

2.1. Resonance parameters

The incident neutron ^{232}Th evaluation [7] from the ENDF/B-VII.1 library [1] has been chosen as an example due to several reasons. The corresponding resolved resonance covariance file includes significant flaws, which might be circumvented (i.e. ignored) in deterministic sensitivity analysis but become apparent in combination with random sampling. Some widths of narrow *p*-wave resonances have extremely high relative uncertainties, up to 5000%! The use of lognormal distribution is highly recommended in order to avoid

* Corresponding author. Tel.: +38 615885326; fax: +38 615885454.
E-mail addresses: gasper.zerovnik@ijs.si (G. Žerovnik), andrej.trkov@ijs.si (A. Trkov), leallc@ornl.gov (L.C. Leal).

0168-9002/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nima.2014.01.012

sampling of negative values. Furthermore, some small resonances have been added for a better description of the interference effects. In fact, most of the resonances with extremely large uncertainties are placed statistically. It is known that these resonances exist, however they cannot be resolved from the experimental data. Since their exact position is not known, the statistical distribution of resonance parameters based on average values obtained from an energy region with good average resonance parameters values is used [8]. The consequent almost perfect linearity of the cumulative number of resonances as a function of incident neutron energy is shown in Fig. 1. In general, statistically placed resonances in the resolved resonance range can be a good method to correct for missing resonances, however in future evaluations such artificial resonances should strictly be "flagged" in order to avoid possible confusion.

The correlation matrix of the resonance parameters is given in sparse (compact) format (LCOMP=2 in ENDF-6 [9]). The initial correlation matrix is inconsistent due to

- The initial "correlation matrix" has some negative eigenvalues, therefore strictly mathematically speaking, the matrix is not a correlation matrix since it does not have all the required properties (i.e. it is not positive-semidefinite). When propagating uncertainties to integral quantities using the conventional deterministic sensitivity analysis (e.g. the first order approximation, the "sandwich formula"), there is no need for correlation/covariance matrix inversion or diagonalization, therefore the issue of negative eigenvalues might be overlooked, either intentionally or by accident. On the other hand, when resonance parameters are randomly sampled using covariance matrix with negative eigenvalues, samples with non-zero imaginary components are produced with non-zero probability. In the case of the ^{232}Th evaluation from the ENDF/B-VII.1 library, the nearest correlation matrix (with non-negative eigenvalues) was generated using the method by Higham [10]. The new correlation matrix is nearest to the old matrix according to the Frobenius norm [11]. In the discussion below, this step has always been performed. Use of this method is highly recommended for any analogous case.
- Some of the uncertainties of strictly positive quantities (i.e. resonance widths) in the file are very large. Sampling according to a normal distribution would frequently yield negative values which are unphysical. Therefore, the use of lognormal distribution is recommended. However, we often observe correlation coefficients approaching the value −1, which is possible only for symmetric distributions [5]. Some of the anti-correlations are too large for lognormal distribution. In the case of ^{232}Th evaluation from ENDF/B-VII.1 library, the evaluators prepared the correlation matrix which is consistent with multivariate normal distribution. Therefore, we transformed the correlation coefficients C_{ij} as if they were generated for symmetric (normal) distribution to correlation coefficients for lognormal distribution [5]:

$$C_{ij}^{(ln)} = \frac{\langle x_i \rangle \langle x_j \rangle}{\sigma_{x,i}\sigma_{x,j}} \left(\exp\left[C_{ij}^{(n)} \sqrt{\ln\left(\frac{\sigma_{x,i}^2}{\langle x_i \rangle^2}+1\right) \ln\left(\frac{\sigma_{x,j}^2}{\langle x_j \rangle^2}+1\right)} \right] - 1 \right), \quad (1)$$

where $\langle x_i \rangle$ and $\sigma_{x,i}$ are the mean value and the standard deviation of the parameter x_i, respectively. For more general recommendations the reader is advised to consult Ref. [5].

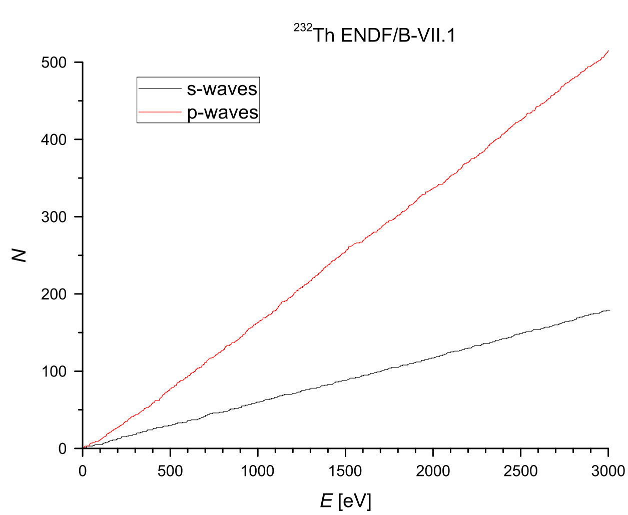

Fig. 1. Cumulative number N of ^{232}Th resonances as a function of incident neutron energy E.

2.2. Random sampling procedure

A Monte Carlo method has been used to calculate resonance cross-sections and their uncertainties considering the full covariance matrices. The mean resonance parameter values and their covariance matrix are read from the selected ENDF file. Details of the applied sampling procedure can be found in Refs. [5,6]. It was verified that the employed sampling method is consistent with information stored in the source evaluation. By random sampling around the mean values and considering correlations according to the information in the covariance matrix, a number of *perturbed* cross-section files in ENDF-6 format are obtained. Each of the perturbed ENDF files is then processed to investigate the influence of the uncertainties of resonance parameters on integral observables (in particular on the resonance integral *RI*). For calculations in this paper, 1000 perturbed ENDF files have been produced.

2.3. Convergence of the random sampling of lognormally distributed parameters

The lognormal distribution is defined by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2 x^2}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0 \quad (2)$$

with

$$\sigma = \sqrt{\ln\left(1 + \frac{\sigma_x^2}{\langle x \rangle^2}\right)}, \quad (3)$$

$$\mu = \ln\langle x \rangle - \frac{\sigma^2}{2}, \quad (4)$$

where $\langle x \rangle$ and σ_x represent the mean value and standard deviation, respectively.

The speed of convergence of the distribution moments as a function of the number of random samples n strongly depends on the shape of the distribution, i.e. the ratio $\sigma_x/\langle x \rangle$, which means the relative uncertainty or standard deviation. If the relative uncertainty is increased, the shape of the distribution approaches loguniform $(1/x)$ on a progressively broader interval around the mean value. Due to the high tail of the loguniform distribution the convergence of the mean value (Fig. 2) and especially the standard deviation (Fig. 3) of the sampled parameters becomes very slow. If the relative uncertainty is below around 100%, 1000 samples suffice to achieve an accuracy within a few percent. However, for larger relative uncertainties, the required number of samples soon increases by several orders of magnitude. For example, for relative uncertainty of 1000% the standard deviation is still far from the converged one even after 10^5 generated samples.