Evaluating Embedded Monte Carlo vs. Total Monte Carlo for Nuclear Data Uncertainty Quantification

Grégoire Biot[1,*],, *Dimitri* Rochman[2], *Pablo* Ducru[3], and *Benoit* Forget[1]

[1]Massachusetts Institute of Technology, 77 Mass. Avenue, Cambridge, MA 02139, United States
[2]Paul Scherrer Institute, Forschungsstrasse 111, 5232 Villigen PSI, Switzerland
[3]Raive, Inc., 6 Bellis Cir, Cambridge, MA 02140, United States

Abstract. The purpose of this paper is to compare a new method called Embedded Monte Carlo (EMC) to the well-known Total Monte Carlo (TMC) method for nuclear data uncertainty propagation. Indeed, the TMC methodology is based on the use of a large number of random samples of nuclear data libraries and performing separate Monte Carlo calculations for each random sample. Then, the computation of nuclear data uncertainty is based on the difference between the total uncertainty and the statistical uncertainty of each Monte Carlo simulation. This method can either be applied to MC and deterministic codes where there are no statistical uncertainties. The goal of EMC is to compute statistical uncertainty for each random sample by utilizing historical statistics instead of the batch statistics employed in TMC. Consequently, one large Monte Carlo simulation can be conducted where each batch represents a new random sample, thereby embedding the propagation of uncertainties within a single calculation and reducing computational expenses. This approach allows for the calculation of nuclear data uncertainty using history statistics in fixed source and eigenvalue calculations. This paper demonstrates the capability of this new method using OpenMC. The analysis will be performed on a Godiva sphere benchmark by propagating the uncertainty on two input parameters: the average neutron multiplicity $\overline{\nu}$ and ^{235}U density.

1 Introduction

Nuclear data describe the nuclear properties of atomic nuclei and their physical nuclear interactions. A comprehensive collection of nuclear reactions and decay data are needed for simulations of nuclear applications. Examples of nuclear data include basic nuclear properties, cross-sections, half-lives, decay modes, etc... Therefore, nuclear data are the basic information to understand all the physical processes underlying most of nuclear technologies. Applications of nuclear data include areas of nuclear science and technology such as fission energy, fusion energy, nuclear fuel cycles, waste management and decommissioning and accelerator driven systems. There are also non-energy applications such as medicine (diagnosis and therapy), production of radioisotopes for medical and industrial applications, dosimetry and radiation safety, nuclear security, materials analysis. Quantifying the uncertainty of nuclear data is of paramount importance when talking about nuclear simulations. Nuclear data

*e-mail: grego01@mit.edu

uncertainties have a significant impact on reactor physics, influencing key parameters such as the neutron multiplication factor k_{eff}, power distributions, and safety margins. Accurate uncertainty quantification helps in optimizing reactor designs, improving safety protocols, and ensuring economic competitiveness [1]. Uncertainty quantification is crucial for the development and deployment of advanced nuclear systems, such as Generation IV reactors and accelerator-driven systems. These technologies demand high precision in nuclear data to ensure their safety, efficiency, and sustainability. Moreover, understanding nuclear data uncertainties can guide experimental efforts to improve the quality of nuclear data and refine theoretical models. The most famous methods used to propagate these uncertainties from nuclear data are sensitivity analysis (with perturbation theory) and Total Monte Carlo (TMC). Perturbation theory relies on sensitivity profiles and covariance data to propagate uncertainties. This deterministic approach calculates the impact of variations in nuclear data on reactor parameters by linearizing the relationship between input uncertainties and output quantities. Although widely used, this method has limitations, including the assumption of linearity and the need for extensive covariance data. On the other hand, the TMC method is based on the variations of parameters for physical models to produce random nuclear data using batch statistics, in the case of Monte Carlo simulations. This stochastic approach does not require linearity assumptions and can capture the full range of possible outcomes by generating probability distributions for reactor parameters. A new algorithm called Embedded Monte Carlo (EMC) was recently developed in [2]. It is based on estimating the history moments to obtain the distribution of uncertainties of the simulation. Since EMC was never implemented for reactor physics problem, the goal of this paper is to compare this new algorithm to the well-known TMC and fast TMC methods on the Godiva benchmark problem used for nuclear data uncertainty propagation.

2 Background

2.1 Monte Carlo statistics

In Monte Carlo codes, there are two different types of statistics possible to implement: history and batch. With batch statistics, the mean value is based on the average of the means over a number of batches. If the number of histories stays constant for batch statistics, then the total mean is the same as history statistics. Additionally, the variance describes the same value even if this will result in fewer degrees of freedom. The reduction in degrees of freedom has a few effects. First, using a small number of batches can result in a significant bias. Indeed, there is a correlation between generations and histories. The main reason of this bias of variance is due to the assumption made that all particles and realizations (batches) are independent [3]. Nevertheless, in an eigenvalue computation, the source sites for each batch depend on the fission neutrons of the preceding batch resulting in a correlation between successive batches. The solution to this bias is to sum together all histories in a generation to eliminate the correlation between histories within a single generation. It is important to note that the general and more accurate formula for the variance of the mean is the following: [4]:

$$Var(\overline{x}) = \frac{1}{N^2} \cdot \left[\sum_{i=1}^{N} Var(x_i) + \sum_{i=1}^{N} \sum_{i \neq j} Cov(x_i, x_j) \right] \quad (1)$$

where the covariance of two different samples from the same population (thus μ being the mean for both samples) is defined as $Cov(x_i, x_j) = E[(x_i - \mu)(x_j - \mu)]$. If the assumption that all batches are independent, then the second term in Equation (1) vanishes and this quantity becomes the unbiased variance estimator of the mean computed in Monte Carlo codes.

3 Methodology

3.1 Total Monte-Carlo (TMC)

This method introduces an additional kind of uncertainty to the calculation results, called aleatoric uncertainty, which results from the randomness of the calculation procedure performed by the Monte Carlo sampling process. This output uncertainty $\sigma_{statistical}$ can be reduced or even eliminated by increasing the number of neutron histories or the number of batches in the simulation. The uncertainty methodology with Monte Carlo is based on the use of random samples of nuclear data libraries and performing separate calculations for each random sample. This approach doesn't use any approximation, no first order or linear approximation is required but due to its stochastic nature, it carries this inherent statistical uncertainty. The idea is to simulate a large number of calculations all alike but with different random nuclear data in each of them, the results consist of a probability distribution function from which different moments can be extracted. This method can be flexible in the use of covariance files but in [1], the TMC method generates random nuclear data from fundamental theoretical nuclear quantities with the help of a nuclear reaction code like TALYS [5]. For each random ENDF-6 file [1] and at the end of the n number of random samples, n different k_{eff} values with their statistical uncertainty are obtained. If we assume that the samples are uncorrelated, the total variance of the simulation σ^2_{total} can be separated in two parts [6]:

$$\sigma^2_{total} = \sigma^2_{statistical} + \sigma^2_{nuclear-data} \tag{2}$$

where $\sigma^2_{statistical}$ corresponds to the statistical uncertainty derived from the number of histories used in the simulation and $\sigma^2_{nuclear-data}$ is due to the different random nuclear data files between calculations. It induces a spread in the distribution of the calculated reactor quantity, which can unequivocally be assigned to the spread of nuclear data (e.g. cross sections, angular distributions, nubar, double differential data, prompt fission neutron spectrum). The statistical uncertainty typically varies as $1/\sqrt{N}$ where N is the number of histories [7].

3.1.1 Fast TMC

One of the main drawback of Total Monte Carlo is the computation time required in comparison to a single calculation. Indeed, the minimization of the statistical uncertainty is making the uncertainty propagation inefficient compared to deterministic methods. The idea of fast TMC, defined in [9], is to obtain a Monte Carlo simulation with the same calculation time as a single run by reducing the number of histories for each run and using different seeds in each run. Uncertainties could then be provided for a Monte Carlo simulation within the same calculation time as a single simulation. The main differences between TMC and fast TMC are that the number of histories is reduced for each sample and that each run has a different random seed in order to make sure there is no correlation coming from the pseudo-random number generator. It is important to note that this reduced number of histories does not impact the number of random samples since increasing this number does not lead to a more precise answer.

3.2 Embedded Monte-Carlo (EMC)

The Embedded Monte Carlo method rests on the observation that the output quantity of interest Z is an expectation value of the form $Z(X) = E_Y[Y|X]$, where X is the input parameter

[1] The term "ENDF-6" refers to the format used for evaluated nuclear data files and not to the "ENDF-VI" library

and Y is the stochastic process being simulated (neutron random walk). This method samples different possible nuclear data values from their uncertainty distribution (e.g normal distribution), and is able to converge towards accurate uncertainty estimates despite running only a few non-converged nuclear simulations for each possible nuclear data sample (thereby reducing computational cost). Estimating the forward propagated moments and then approximate the distribution of output means using the principle of maximum entropy [2] explained in [10] . Instead of running a full MC simulation for each sample of the distribution to obtain directly the nuclear data uncertainties, the EMC method runs only one MC simulation where each batch has a unique random sample of the nuclear data. The idea is to compute history statistics that requires only one batch of neutrons. To fully grasp EMC, it is important to recall how all the estimators are computed. As in TMC, for each random sample M, N neutrons are simulated which correspond to N independent stochastic processes $Y_n|X_m$ (X_m corresponds to the mth random sample of parameter X, drawn from its uncertainty distribution). It is possible to obtain the mean and the variance for each sample using Equations (3) and (4) respectively (using the notation in [2]):

$$\widehat{Z}(X_m) = \frac{1}{N}\sum_{n=1}^{N} Y_n|X_m \tag{3}$$

$$\widehat{\Sigma_Y}(X_m) = \frac{1}{N-1}\sum_{n=1}^{N}\left(Y_n|X_m - \widehat{Z}(X_m)\right)^2 \tag{4}$$

where $\widehat{Z}(X_m)$ and $\widehat{\Sigma_Y}(X_m)$ are the mean and variance between the N stochastic processes respectively (also called inner loop). The inner loop of the EMC algorithm is on the neutrons Y_n and the outer loop is on the samples of the uncertainty distribution X_m. After running all M random samples, the following estimators are computed to find the nuclear data uncertainty for this simulation [2]:

$$\langle\widehat{Z}\rangle = \frac{1}{M}\sum_{m=1}^{M}\widehat{Z}(X_m) \tag{5}$$

$$\langle\sigma^2_{\widehat{Z}}\rangle = \frac{1}{M-1}\sum_{m=1}^{M}\left(\widehat{Z}(X_m) - \langle\widehat{Z}\rangle\right)^2 \tag{6}$$

$$\langle\widehat{\Sigma_Y}\rangle = \frac{1}{M}\sum_{m=1}^{M}\widehat{\Sigma_Y}(X_m) \tag{7}$$

$$\langle\widehat{\Delta^2 Z}\rangle = \langle\sigma^2_{\widehat{Z}}\rangle - \frac{\langle\widehat{\Sigma_Y}\rangle}{N} \tag{8}$$

where $\langle\widehat{Z}\rangle$ and $\langle\widehat{\Delta Z}\rangle$ are respectively the unbiased estimators of the mean and variance of the uncertainty distribution of the output quantity of interest, $Z(X) = E_Y[Y|X]$. $\langle\widehat{\sigma^2_Z}\rangle$ represents the total variance, mixing both inner loop (aleatoric uncertainty from stochastic process Y) and outer loop (epistemic uncertainty from input parameter X). $\langle\widehat{\Sigma_Y}\rangle$ illustrates the average over the different input parameters of each history variance. It is possible to see that Equation (6) is the same definition as σ^2_{total} and that Equation (7) is identical to $\sigma^2_{statistics}$. Therefore, Equation (8) defines the way EMC computes the nuclear data uncertainty which is very close to the TMC method. Nevertheless, the main difference between TMC and EMC that should be emphasized is that the mean of the statistical uncertainty in EMC is divided by the number of neutrons not by the number of batches as defined in TMC. This can be explained by the fact that the statistical uncertainty is based on the mean variance of the population. In MC codes such as OpenMC or MCNP, Equations (3) and (4) are computed based on the number

[2] The probability distribution which best represents the current state is the one with the largest entropy

of batches of the simulation and not the number of particle histories. The aim of EMC is to be able to get correct statistical estimates in order to run M samples with only a single batch of N neutrons for each sample [11].

3.3 OpenMC

OpenMC is a community-developed Monte Carlo neutron and photon transport simulation code. It is capable of performing fixed source, k-eigenvalue, and subcritical multiplication calculations on models built using either a constructive solid geometry or CAD representation. OpenMC supports both continuous energy and multi-group transport. This open-source software simulates neutral particles (presently neutrons and photons) moving stochastically through an arbitrarily defined model that represents a real-world experimental setup. In the case of a nuclear reactor model, neutrons are especially important because they induce fission in isotopes of uranium and other elements. Knowing the behavior of neutrons allows one to determine how often and where fission occurs. The amount of energy released is then directly proportional to the fission reaction rate since most heat is produced by fission. By simulating many neutrons (millions or billions), it is possible to determine the average behavior of these neutrons (or the behavior of the energy produced, or any other quantity one is interested in) very accurately [12].

3.4 Uncertainty propagation workflow

In the present work, the uncertainty propagation scheme can be analysed as three major steps:

1. Gaussian perturbation on the input parameters which correspond to the average neutron multiplicity $\overline{\nu}$ and ^{235}U density.

2. Monte Carlo simulations performed using 32 CPU cores (16 cores per socket with 64-bit CPU architecture) with OpenMC working in fixed source and eigenvalue modes for each random sample using a Godiva sphere benchmark.

3. Analysis of the output parameters k_{eff} and fission reaction rates depending on the simulation mode used in order to extract the moments of the distributions in order to compute the nuclear data uncertainty.

Concerning ^{235}U density, the modifications were made directly in the input file and not further steps were needed in order to propagate the uncertainty. Nevertheless, for the average neutron multiplicity, a set of random HDF5 file was pre-generated for U235 in order to modify only the *cross_sections.xml* file before each simulation. Each new U235 file contained a dataset representing the total number of neutrons produced per fission event for different energy range, which was perturbed using values drawn from a normal distribution.

4 Verification and results

In order to test this new method and compare it with TMC, the well-known critical Godiva benchmark is used to propagate uncertainties on the average neutron multiplicity $\overline{\nu}$ and the density of uranium $\rho_{^{235}\mathrm{U}}$ [13]. All the simulations are made using the nuclear data library ENDF/B-VII.1. HDF5 files are created by first processing source ENDF files into ACE files and then using the *openmc.data* Python module to convert ACE data into HDF5. This benchmark is a bare spherical mass of highly enriched uranium (HEU) with well-documented uranium material properties and critical dimensions, simplifying its implementation in OpenMC [14]. Table 1 presents the isotopic composition of the Godiva sphere used in the simulations.

Table 1: Material composition [15]

Isotope	Density [nuclei/b-cm]
^{235}U	0.045000
^{238}U	0.002498
^{234}U	0.000492

4.1 Eigenvalue mode

In order to avoid modifying OpenMC to compute history statistics for EMC moments, each k_{eff} value is used to compute the variance between each sample to obtain the total uncertainty and a reference value is used for the statistical uncertainty. The underlying assumption is that the statistical variance of each sample is in the same order of magnitude, implying that averaging these values yields a result equivalent to using a single reference value. In this sense, it is possible to compute the nuclear data uncertainty by taking the difference between the total uncertainty (variance of all the random samples) and the statistical uncertainty. This is very useful due to the fact that EMC only simulates one batch of neutrons for each random sample. Before reducing the number of neutrons per simulation, it is important to compute these reference values (mean and variance) for the comparison between fast TMC and EMC. Table 2 presents the uncertainties for the fully converged simulation using TMC brute force method. The final k_{eff} value and statistical error estimation are given by OpenMC from a combination of three individual k_{eff} estimators. This technique produces the maximum likelihood estimate for the combined average k_{eff}. For ^{235}U density, with an initial standard deviation of 0.03, the reference is obtained using $5 \cdot 10^2$ random samples and $5 \cdot 10^7$ neutrons, distributed over $5 \cdot 10^3$ active cycles and 10^3 inactive cycles, requiring 360 min of simulation time. Similarly, for $\overline{\nu}$, with an initial standard deviation of 0.05, the reference is also derived with the same number of random samples, neutrons and active cycles, but requiring 40 minutes more of simulation time. In Table 2, it is important to note that the difference between the mean value and the reference value is higher for the ^{235}U density case than for $\overline{\nu}$. This can be explain by the absence of a linear relationship between the ^{235}U density and k_{eff}. Indeed, with a mean of 0.045 and a standard deviation of 0.03 means that 68% of the data falls within one standard deviation from the mean: [0.015, 0.075]. Higher densities than 0.045 have minimal impact on k_{eff}; for instance, with a density of 0.076, k_{eff} is 1.011. Conversely, densities below 0.015 are extremely unlikely but significantly affect k_{eff}; for example, at a density of 0.006, k_{eff} drops to 0.863. For this particular case of ^{235}U density, 33 out of 500 random samples were below 0.01, substantially lowering the mean k_{eff} from the reference value.

4.1.1 Density of ^{235}U

In a standard normal distribution, there is a non-zero probability of sampling negative values, which are not realistic for physical densities. By truncating the distribution at zero, these non-physical values are eliminated, and although a bias on the moments of the normal distribution is possible, it is likely to be negligible [8]. A truncated normal distribution that only samples positive value is used for ^{235}U density. For fast TMC, the total number of neutrons was divided by the number of random samples reducing the simulations to 10^5 neutrons (divided in 10^3 active cycles with 10^3 inactive cycles) as defined in [9]. Concerning EMC, $5 \cdot 10^2$ random samples were used with 10^3 inactive cycles and one active cycles with 10^5 neutrons. Only the mean and total uncertainty are computed for each value of $\overline{\nu}$ and the reference statistical error is used. Fig. 1a presents the neutron multiplication factor distribution and three normal distribution with moments computed by different methods: Fast TMC with the blue solid line, EMC with the orange dashed line, and a reference computed with TMC given by Table 2 with the red dotted line. All the detailed results of Section (4.1) for fast TMC and EMC can be found in Table 2. It is possible to analyse that the mean values for k_{eff} are identical between TMC and EMC. Nevertheless, the variance computed by EMC is much closer to the reference than fast TMC. This can be explained by the fact that fast TMC computes the nuclear data uncertainty based on the statistical uncertainty of each random sample. Indeed, this method performs relatively well provided that σ^2_{stat} is no greater than 5% of the total uncertainty σ^2_{tot} [9]. In this case, the ratio $\frac{\sigma_{stat}}{\sigma_{tot}}$ is equal to 0.2 which means that the statistical uncertainty is too high to get a good quantification of the nuclear data uncertainty. However, EMC is able to bypass this statistical threshold to provide a better estimate when propagating the uncertainty on ^{235}U density in eigenvalue mode.

Table 2: Nuclear data uncertainties using reference TMC, Fast TMC and EMC in eigenvalue mode

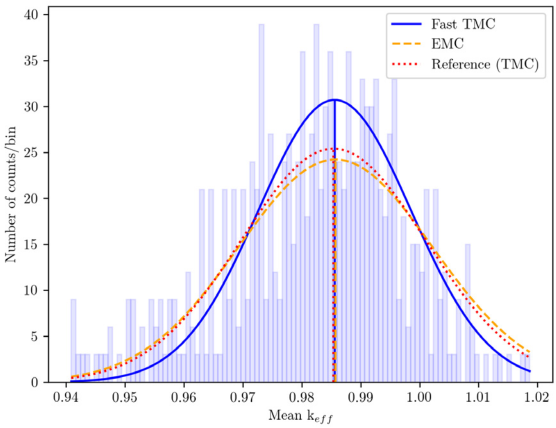
Method	Parameter	Mean k_{eff}	σ^2_{stat}	σ^2_{tot}	σ^2_{ND}
TMC	$\rho_{^{235}\mathrm{U}}$	0.985	$9.524 \cdot 10^{-5}$	$2.560 \cdot 10^{-3}$	$2.465 \cdot 10^{-3}$
	$\bar{\nu}$	1.000	$9.654 \cdot 10^{-5}$	$2.508 \cdot 10^{-3}$	$2.411 \cdot 10^{-3}$
	Both	0.986	$9.545 \cdot 10^{-5}$	$6.703 \cdot 10^{-3}$	$6.608 \cdot 10^{-3}$
Fast TMC	$\rho_{^{235}\mathrm{U}}$	0.986	$4.303 \cdot 10^{-4}$	$2.117 \cdot 10^{-3}$	$1.687 \cdot 10^{-3}$
	$\bar{\nu}$	0.999	$9.848 \cdot 10^{-4}$	$2.513 \cdot 10^{-3}$	$1.529 \cdot 10^{-3}$
	Both	0.986	$21.81 \cdot 10^{-4}$	$5.133 \cdot 10^{-3}$	$2.952 \cdot 10^{-3}$
EMC	$\rho_{^{235}\mathrm{U}}$	0.986	N/A	$2.808 \cdot 10^{-3}$	$2.712 \cdot 10^{-3}$
	$\bar{\nu}$	1.001	N/A	$2.623 \cdot 10^{-3}$	$2.526 \cdot 10^{-3}$
	Both	0.986	N/A	$5.359 \cdot 10^{-3}$	$5.263 \cdot 10^{-3}$

4.1.2 Average neutron multiplicity

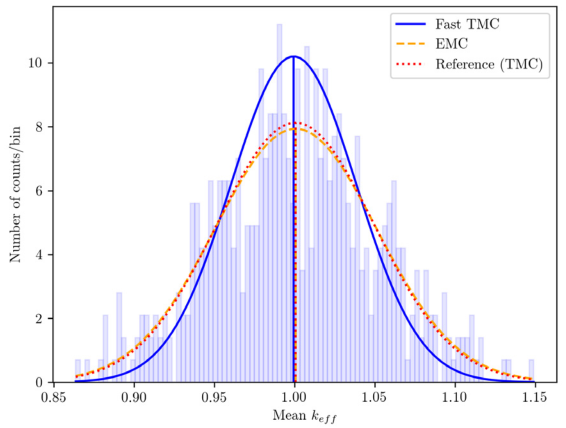
For the other parameter, the same operating mode is used. For EMC, $5 \cdot 10^2$ random samples were used with 10^3 inactive cycles and one active cycles with 10^5 neutrons whereas for fast TMC, the same number of random samples and inactive cycles is used with 10^3 active cycles and 10^2 neutrons following the same general principle as in Section (4.1.1). Only the mean and total uncertainty are computed for each value of $\bar{\nu}$ and the reference statistical error is used. Figure 1b presents the output distribution of the effective neutron multiplication factor k_{eff} when varying the average neutron multiplicity. Firstly, it is important to note that the histogram of k_{eff} distribution is not converged due to the low number of neutrons used in the simulation. In this case, it is possible to observe the same pattern for the nuclear data uncertainty computed based on fast TMC. Indeed, the ratio $\frac{\sigma_{stat}}{\sigma_{tot}}$ is approximately 0.39 which explains this variance difference. As in the previous case, EMC is able to offer a more reliable estimate for the average neutron multiplicity in eigenvalue mode with a low number of neutrons. It is possible to see that inactive cycles are mandatory in order for the fission source to converge but that moments can be computed with only one active cycle. A separate study on the assumption made on the statistical uncertainty for eigenvalue mode will be performed later.

4.1.3 Combined case

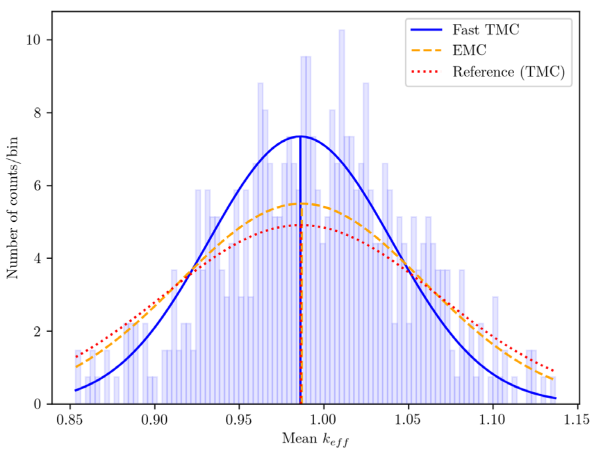
Another interesting case to study is the comparison between fast TMC and EMC when both parameters are perturbed. For fast TMC, $5 \cdot 10^2$ random samples were used with 10^3 inactive



(a) ^{235}U density case



(b) Average neutron multiplicity case



(c) Combined case

Figure 1: k_{eff} distribution for Godiva sphere using eigenvalue mode for ^{235}U density, average neutron multiplicity and both input parameters perturbed using $5 \cdot 10^2$ random samples and respectively $5 \cdot 10^7$, 10^5 and 10^5 neutron histories for TMC, fast TMC and EMC

cycles and 10^3 active cycles with 10^2 neutrons whereas for EMC, the same number of random samples and inactive cycles is used with one active cycles and 10^2 neutrons following the same general principle as in Section (4.1.1). Figure 1c illustrates the output distribution of the effective neutron multiplication factor when both parameters are perturbed. It is possible to observe a similar behaviour to the two previous cases with the histogram obtained from fast TMC not fully converged. Moreover, EMC delivers a more accurate estimate of the nuclear data uncertainty than fast TMC when reducing the number of neutron histories and varying the average neutron multiplicity and ^{235}U density. Table 2 shows that modifying ^{235}U density has much more impact on k_{eff} than varying the average neutron multiplicity.

4.2 Fixed source mode

In order to compute history statistics for fixed source simulation, the number of batches corresponds to the number of particles simulated in OpenMC. Table 3 presents the uncertainties for the reference fixed source simulation computed with a fully converged TMC algorithm. For ^{235}U density, with an initial standard deviation of 0.03, the reference is obtained using 10^3 random samples and $5 \cdot 10^6$ neutrons, distributed over $5 \cdot 10^2$ cycles, requiring 1125 minutes of simulation time. Similarly, for nubar, with an initial standard deviation of 0.001, the reference is derived using the same number of samples, neutrons, and cycles, but requires 1340 minutes of simulation time. It is important to note that the initial standard deviations differ between the two parameters, leading to dissimilar results. Furthermore, the impact of these parameters on the fission reaction rate varies across different cells of the Godiva sphere. Two cells, with diameters of 1 cm and 8 cm respectively, were selected to validate and verify the simulation results, as well as to capture the heterogeneities in material properties, neutron flux, and reaction rates, and their impact on the overall behavior of the system. For both cases, an average of the fission reaction rate in each cell is made between collision and tracklength estimators. Collision estimators directly tally interactions at collision sites, offering detailed insights into local reaction rates. Tracklength estimators, on the other hand, score contributions at every track for the particle rather than every collision. Each time a neutron makes a track inside the region, the track length is recorded. Using both methods enhances the accuracy and reliability of the results by allowing cross-verification and better versatility, ultimately improving the robustness of the nuclear data analysis.

Table 3: Reference data using TMC for fixed source simulations

Cell	Parameter	$RR_{fission}$	σ^2_{stat}	σ^2_{tot}	σ^2_{ND}
1	$\rho_{^{235}\mathrm{U}}$	0.644	$6.073 \cdot 10^{-7}$	$3.047 \cdot 10^{-2}$	$3.047 \cdot 10^{-2}$
	$\overline{\nu}$	0.668	$7.916 \cdot 10^{-7}$	$1.394 \cdot 10^{-2}$	$1.394 \cdot 10^{-2}$
7	$\rho_{^{235}\mathrm{U}}$	2.042	$119.2 \cdot 10^{-7}$	$76.08 \cdot 10^{-2}$	$76.07 \cdot 10^{-2}$
	$\overline{\nu}$	2.170	$167.2 \cdot 10^{-7}$	$37.67 \cdot 10^{-2}$	$37.66 \cdot 10^{-2}$

4.2.1 Density of ^{235}U

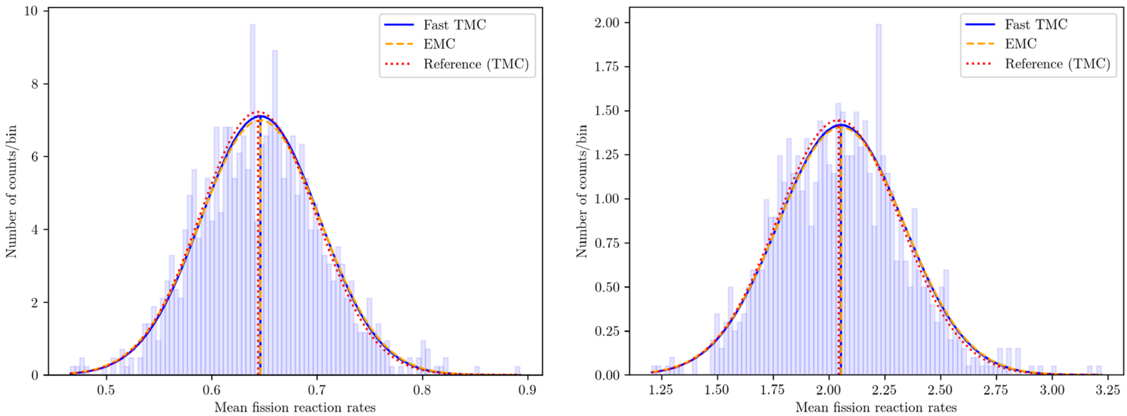
For fixed source simulations, the same truncation methodology of the initial normal distribution is used to produce random samples for ^{235}U density. For fast TMC, the number of neutrons is divided by the number of random samples (M=10^3) which means that $5 \cdot 10^3$ neutrons (divided in 50 cycles) were simulated in 30 minutes. For EMC, the number of batches is set to $5 \cdot 10^3$ with one neutron per batch in order to get history statistics with a simulation time of 35 minutes. Figures 2a and 2b present the fission reaction rate distribution for inner and outer cells with three normal distribution with moments computed by different methods: Fast TMC with the blue solid line, EMC with the orange dashed line, and a reference computed with TMC given by Table 3 with the red dotted line. For both cells, it is possible to see that fast TMC and EMC are in agreement with the reference mean and variance. The differences in the mean, the statistical uncertainty and the nuclear data uncertainty between fast TMC and EMC are lower than 10^{-3}, 10^{-6} and 10^{-3} respectively. It is possible to conclude that EMC is able to reproduce the same results as fast TMC using history statistics in fixed source simulations. One limitation in the application of EMC is that most of the MC codes are using batch statistics in order to compute the statistical uncertainty of the simulation. To fully embed the propagation of uncertainties and run only one batch of neutrons, modifications in OpenMC source code and python API should be made to obtain history statistics in fixed source mode. All the detailed results (mean, variance, etc.) for fast TMC and EMC can be found in Table 4.

4.2.2 Average neutron multiplicity

Since the reference is computed using 10^3 random samples of $\overline{\nu}$, the number of neutrons is reduced to $5 \cdot 10^3$ in order to compute the first two moments of the output distribution using fast TMC and EMC methods. For EMC, $5 \cdot 10^3$ batches are utilized, each with a single neutron per batch. Conversely, the fast TMC method involves 50 batches, each containing 100 neutrons. Using Table 4, it is possible to analyze that the mean fission reaction rates $RR_{fission}$ are identical for both methods, indicating that history statistics is equivalent to batch statistics for the mean values. The primary difference between these two methods arises from the statistical uncertainty σ^2_{stat} which is varying by 10^{-3} between EMC and fast TMC. Fig. 3a and 3b show the fission reaction rate distributions and three log-normal distributions with moments computed by different methods: Fast TMC with the blue solid line, EMC with the orange dashed line, and a reference computed with TMC given by Table 4 with the red dotted line. Indeed, it is important to notice that both fission reaction rate distributions have negative skewness suggesting that a log-normal distribution is more suitable to reconstruct the histogram of output. Moreover, higher moments should be computed in order to correctly estimate the distribution's tail behaviour and for assessing the likelihood of extreme values. Indeed, EMC is able to compute the skewness (third moment) and the kurtosis (fourth moment) of the output distribution in order to obtain correct and unbiased estimators of the output distribution as defined in [2]. Furthermore, it is possible to see that even if the fast TMC histogram of $RR_{fission}$ is not fully converged, EMC and fast TMC can compute the moments of the distributions with high precision compared to the reference with errors smaller than 10^{-3}, 10^{-5} and 10^{-3} for the mean values, σ^2_{stat} and σ^2_{ND} respectively. This implies that reducing the number of neutrons per simulation does not affect the statistical accuracy of the results, and using history statistics is equivalent to using batch statistics in fixed source mode for both cases.

Table 4: Nuclear data uncertainties using TMC and EMC with low number of neutrons in fixed source mode

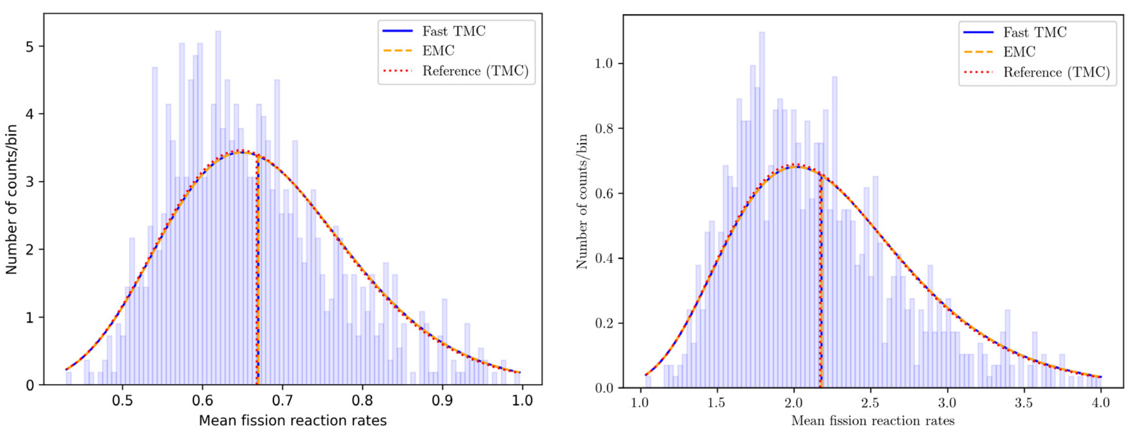
Method	Cell	Parameter	$RR_{fission}$	σ^2_{stat}	σ^2_{tot}	σ^2_{ND}
Fast TMC	1	$\rho_{^{235}\mathrm{U}}$	0.646	$6.197 \cdot 10^{-4}$	$3.770 \cdot 10^{-2}$	$3.151 \cdot 10^{-2}$
		$\overline{\nu}$	0.670	$8.158 \cdot 10^{-4}$	$1.502 \cdot 10^{-2}$	$1.420 \cdot 10^{-2}$
	7	$\rho_{^{235}\mathrm{U}}$	2.054	$121.8 \cdot 10^{-4}$	$91.31 \cdot 10^{-2}$	$79.13 \cdot 10^{-2}$
		$\overline{\nu}$	2.179	$174.1 \cdot 10^{-4}$	$40.37 \cdot 10^{-2}$	$38.63 \cdot 10^{-2}$
EMC	1	$\rho_{^{235}\mathrm{U}}$	0.646	$6.180 \cdot 10^{-4}$	$3.870 \cdot 10^{-2}$	$3.252 \cdot 10^{-2}$
		$\overline{\nu}$	0.670	$8.069 \cdot 10^{-4}$	$1.502 \cdot 10^{-2}$	$1.421 \cdot 10^{-2}$
	7	$\rho_{^{235}\mathrm{U}}$	2.054	$121.4 \cdot 10^{-4}$	$92.75 \cdot 10^{-2}$	$80.61 \cdot 10^{-2}$
		$\overline{\nu}$	2.179	$171.4 \cdot 10^{-4}$	$40.37 \cdot 10^{-2}$	$38.65 \cdot 10^{-2}$



(a) Cell N°1

(b) Cell N°7

Figure 2: Fission reaction rate distribution using fixed source mode for ^{235}U density using 10^3 random samples and respectively $5 \cdot 10^6$, $5{\cdot}10^3$ and $5 \cdot 10^3$ neutron histories for TMC, fast TMC and EMC



(a) Cell N°1

(b) Cell N°7

Figure 3: Fission reaction rate distribution using fixed source mode for the average neutron multiplicity using 10^3 random samples and respectively $5 \cdot 10^6$, $5{\cdot}10^3$ and $5 \cdot 10^3$ neutron histories for TMC, fast TMC and EMC

5 Conclusions and future work

To advance future nuclear technologies, ranging from space applications to medical diagnostics and carbon-neutral power generation, it's crucial to have highly accurate nuclear simulations backed by reliable confidence intervals. The foundation of our understanding in nuclear physics is built on nuclear data, which is derived from experiments that inherently include some level of uncertainty. Therefore, accurately accounting for the uncertainties in nuclear data within these simulations is crucial to ensure the accuracy and safety of nuclear systems. This aspect has even become a bottleneck in the development of certain innovative reactor technologies, such as the design of SMR, GEN-IV reaction or fusion systems. The Total Monte Carlo and its fast derivative approaches are advanced methods for nuclear data uncertainty quantification. By generating a comprehensive set of nuclear data files through random sampling, it enables a more thorough exploration of uncertainties and their impact on nuclear system performance, thereby improving reliability and safety. This study focuses on the comparison between the novel method Embedded Monte Carlo to fast TMC. Regarding the eigenvalue mode, it can be concluded that EMC provides more accurate estimates compared to fast TMC, thanks to assumptions made about statistical uncertainty. In terms of the fixed source mode, EMC is found to be equivalent to fast TMC when following the general principle. Nevertheless, further investigations, including convergence analysis and comparison with other methods, are recommended to determine the minimum number of neutrons needed to achieve an accurate estimation of output distribution moments for EMC. This paper is only the first step of a long road to demonstrate the role of Embedded Monte Carlo in the world of nuclear data uncertainty quantification. In any case, EMC creates the path to a new uncertainty propagation method that relies on an better understanding of Monte Carlo statistics and can be available to every user in the world due to the open-source nature of OpenMC.

Funding

This research is being performed using funding received from the DOE Office of Nuclear Energy's Nuclear Energy University Programs.

References

[1] A. J. Koning and D. Rochman, Annals of Nuclear Energy **35**, 2024–2030 (2008)

[2] Pablo Ducru (2021), *Nuclear Computations under Uncertainty : New methods to infer and propagate nuclear data uncertainty across Monte Carlo simulations* [PhD thesis, Massachusetts Institute of Technology]

[3] Picard Richard Roy, Anthony J. Zukaitis, and Robert Arthur Forster Iii, *Evaluating Equivalent Monte Carlo Calculations*, Vol. Tech. Rep. LA-UR-18-29274, Los Alamos National Laboratory, NM, USA, October 2018

[4] B. R. Herman (2014), *Monte Carlo and Thermal Hydraulic Coupling using Low-Order Nonlinear Diffusion Acceleration* [PhD thesis, Massachusetts Institute of Technology]

[5] A.J. Koning, M.C. Duijvestijn and S. Hilaire, *Talys-1.0*, Proceedings of the International Conference on Nuclear Data for Science and Technology, 22–27 April 2007, Nice, France, EDP Sciences, pp. 211–214

[6] D. Rochman, A. J. Koning and S.C. van der Marck, Annals of Nuclear Energy **36**, 810-831 (2009)

[7] D. Rochman, A. J. Koning, and C.M. Sciolla, Annals of Nuclear Energy **38**, 942–952 (2011)

[8] T. Loux and O. Davy, The American Statistian **75**, 294-299 (2021)

[9] D. Rochman, W. Zwermann, S.C. van der Marck, A.J. Koning, H. Sjostrand, P. Helgesson and B. Krzykacz-Hausmann, Nuclear Science and Engineering **177**, 337–349 (2014)

[10] Sung Y. Park and K. Anil, Journal of Econometrics, **150**, 219–230, (2009)

[11] J. A. Kulesza, T. R. Adams, J. C. Armstrong, S. R. Bolding, F. B. Brown, J. S. Bull, T. P. Burke, A. R. Clark, R. A. Forster III, J. F. Giron, T. S. Grieve, C. J. Josey, R. L. Martz,

G. W. McKinney, E. J. Pearson, M. E. Rising, C. J. Solomon Jr., S. Swaminarayan, T. J. Trahan, S. C. Wilson, A. J. Zukaitis, *MCNP Code Version 6.3.0 Theory and User Manual*, Los Alamos National Laboratory Tech. Rep. LA-UR-22-30006, NM, USA (2022)

[12] P. K. Romano, N. E. Horelik, B. R. Herman, A. G. Nelson, B. Forget and K. Smith, Annals of Nuclear Energy **82**, pp. 90–97 (2015)

[13] R.E. Patterson and G.A Newby, Nuclear Science and Engineerng **1**, 112-115 (1956)

[14] V. Patel, J. Navarro, W. Windes and P. Tsvekov, Annals of Nuclear Energy **165**, 108629 (2022)

[15] International Handbook of Evaluated Criticality Safety Benchmark Experiments, NEA/NSC/DOC(95)03, OECD-NEA (2012)