Real Variance Estimation in iDTMC method using Spectral Analysis Method and Autoregressive (1) Model

Jaehyeong Jang[1] and Yonghee Kim[1*]

[1]*Department of Nuclear & Quantum Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea*
[*]*Corresponding author: yongheekim@kaist.ac.kr*

***Keywords :** iDTMC Method, Autoregressive (1) Model, Spectral Analysis Method, Real Variance

1. Introduction

The Monte Carlo (MC) method is a highly accurate approach for reactor analysis, directly simulating neutron behavior with minimal assumptions. However, this accuracy comes with significant computational costs due to the need to simulate many neutron histories and cycles.

The Improved Deterministic Truncation of Monte Carlo (iDTMC) method accelerates the MC method by integrating stochastic and deterministic techniques, enhancing efficiency and reducing computational demands[1].

Accurate uncertainty estimation remains crucial, as the inherent cycle correlation in MC simulations, especially in the iDTMC method, can lead to underestimated variance if not properly addressed. This study introduces two new methods for estimating the real variance of the iDTMC method using Spectral Analysis Method and Autoregressive (1) models.

2. Methodology

2.1. The iDTMC Method

The iDTMC method employs p-CMFD to accelerate the convergence of the FSD during the inactive cycles and uses p-FMFD to generate pin-level reactor solutions during the active cycles.

The p-CMFD method solves the one-group neutron balance equation, Eq. (1):

$$\sum_s \frac{A_s}{V_i}(J_{s1} - J_{s0}) + \Sigma_a^i \varphi_i = \frac{1}{k_{eff}} \nu\Sigma_f^i \varphi_i \quad (1)$$

where A is the surface area, V_i is the volume of node i, s denotes the surface index, ϕ and J are the flux and current, respectively, Σ represents cross section and the effective multiplication factor is denoted as k_{eff}. The current across a specific surface is determined by the sum of partial currents. The primary difference between p-CMFD and p-FMFD lies in the node size: p-CMFD considers subassemblies as nodes, whereas p-FMFD considers individual pins within the subassemblies as nodes.

The overall iDTMC methodology is illustrated in Fig. 1. After a few skip cycles, p-CMFD is applied during inactive cycles to accelerate the convergence of the fission source distribution (FSD). Once converged, the MC solution is truncated using p-FMFD, which is decoupled from MC to avoid instability. The parameters for p-FMFD during active cycles are accumulated from the inactive cycles.

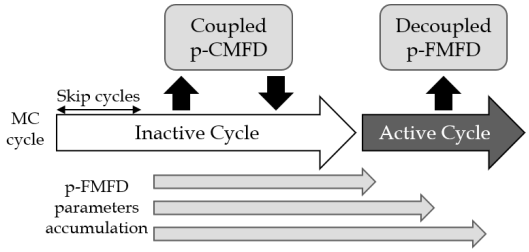

Figure 1. Schematic of the iDMTC method

2.2. Real variance estimation using correlated sampling

To estimate the real variance, a previously developed method using correlated sampling for accumulated FMFD parameters is employed [2].

For a desired number of samples, Uniform Random Numbers (URN) are generated using Latin Hypercube Sampling (LHS) to ensure a well-distributed set of pseudo-random numbers. These URN samples are then transformed into normally distributed random numbers.

The normal samples are then correlated according to the correlation matrix of the accumulated parameters:

$$C = \begin{bmatrix} 1 & \rho(\Sigma_t, \Sigma_a) & \rho(\Sigma_t, \nu\Sigma_f) \\ \rho(\Sigma_a, \nu\Sigma_f) & 1 & \rho(\Sigma_a, \nu\Sigma_f) \\ \rho(\Sigma_t, \nu\Sigma_f) & \rho(\Sigma_t, \Sigma_a) & 1 \end{bmatrix} \quad (2)$$

while Σ_t, Σ_a and $\nu\Sigma_f$ are the accumulated total, absorption and ν-fission cross sections, respectively, and $\rho(X, Y)$ represents the correlation coefficient between variables X and Y, as calculated by Eq. (3).

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[XY] - E[X]E[Y]}{\sigma_X \sigma_Y} \quad (3)$$

Cholesky decomposition is applied to incorporate the correlations into the normally distributed random samples, transforming the positive semi-definite matrix into a lower triangular matrix L as follows:

$$C = LL^T \quad (4)$$

Assuming that the normal samples are initially uncorrelated, the correlation matrix of the normal sample matrix X is represented as:

$$E[XX^T] = I_{n\times n} \quad (5)$$

Multiplying L by the samples results in a new correlation matrix, as shown in Eq. (6), matching the original correlation matrix we constructed:

$$C' = E[LXX^TL^T] = LE[XX^T]L^T = LL^T = C \quad (6)$$

The correlated random samples are transformed back into uniformly distributed random numbers (URNs). These correlated URNs are then used to sample FMFD parameters from the discrete accumulated CDF. Since the samples may not be perfectly correlated, some iteration is required. Finally, these perturbed samples are used to estimate the real variance using first-order perturbation theory, optimizing computing time.

Fig. 2 illustrates the entire procedure for real variance estimation using correlated sampling. This estimation method is shown to be effective for Light Water Reactor (LWR) type problems [2, 5].

Given that the iDTMC method accelerates FSD convergence and FMFD parameters are accumulated during the inactive cycles, there is likely a strong cycle-wise correlation. To address this, we propose using an Autoregressive (AR) model and the Spectral Analysis Method (SAM) to improve real variance estimation. The first method uses an AR (1) model to estimate cycle-wise correlation coefficients, while the second reproduces cycle-wise correlated parameters and estimates variance with SAM

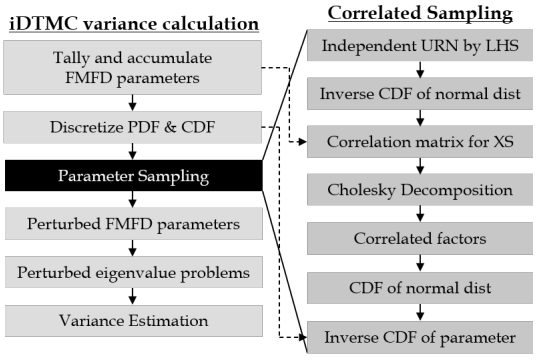

Figure 2. Real variance estimation using correlated sampling

2.3. Autoregressive (1) model

Previous studies suggest that treating consecutive Monte Carlo (MC) samples as a subset of an Autoregressive (AR) model is an effective analytical approach [3]. A first-order AR (AR (1)) model, with a mean μ, can be represented as follows:

$$X_t - \mu = \phi(X_{t-1} - \mu) + Z_t(0, \sigma^2) \quad (7)$$

Here, X_t represents the sample tallied from the t[th] cycle for the total T active cycles, ϕ is the correlation coefficient, and Z_t is the Gaussian white noise with a mean of 0 and a variance of σ^2.

All parameters in the AR (1) model can be estimated using Maximum Likelihood Estimator (MLE). The log-likelihood function of AR (1) model, denoted as $logL(\phi, \mu, \sigma^2)$, is expressed as:

$$logL(\phi,\mu,\sigma^2) = -\frac{1}{2}\log\left(\frac{2\pi\sigma^2}{1-\phi^2}\right) - \frac{(X_1-\mu)^2}{2\,\sigma^2/(1-\phi^2)} - \frac{T-1}{2}\log(2\pi\sigma^2) - \sum_{t=2}^{T}\left\{\frac{[(X_t-\mu)-\phi(X_{t-1}-\mu)]^2}{2\sigma^2}\right\} \quad (8)$$

By differentiating the log-likelihood function with respect to μ, σ^2 and ϕ and setting the derivatives to zero, we can estimate $\hat{\phi}$. With $\hat{\phi}$ in hand, we can then estimate the other parameters $\hat{\mu}$ and $\hat{\sigma}^2$.

2.4. Spectral Analysis Method

The Spectral Analysis Method (SAM) is a technique used for analyzing noise signals and stationary time series. SAM can estimate real variance while inherently accounting for cycle-wise correlations [4].

Assuming that $\{X_t\}$ is a stationary Gaussian process the autocovariance function, which describes the dependency between two cycles, is defined as:

$$\gamma_h = Cov(X_t, X_{t+h}) = E(X_t, X_{t+h}) - \mu^2 \quad (9)$$

where h is the lag between two cycles, and E denotes expectation. The sample autocovariance $\hat{\gamma}_h$, unbiased estimate for γ_h, is given by:

$$\hat{\gamma}_h = \frac{1}{T-h}\sum_{t=1}^{T-h}(X_t - \bar{X})(X_{t+h} - \bar{X}) \quad (10)$$

where $\bar{X}$ is sample mean, calculated as $\bar{X} = \sum_{t=1}^{T}\frac{X_t}{T}$. The real variance, can be expressed in terms of the autocovariance function:

$$var(\bar{X}) = \frac{1}{T^2}var\left(\sum_{t=1}^{T}X_t\right) = \frac{1}{T^2}\left\{\sum_{i=1}^{T}\sum_{j=1}^{T}\gamma_{(i-j)}\right\} \quad (11)$$

For large T, this simplifies to:

$$var\left(\sqrt{T}\bar{X}\right) \rightarrow \sum_{h=-\infty}^{h=\infty}\gamma_h \quad (12)$$

Since any autocovariance structure corresponds to a particular frequency pattern, we can establish the spectral density function $f(\omega)$ using the autocovariance function:

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_h e^{ih\omega}, \omega \in (-\pi, \pi) \quad (13)$$

where $i = \sqrt{-1}$. By evaluating $f(\omega)$ at $\omega = 0$:

$$f(0) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_h \quad (14)$$

Finally, combining Eq. (12) with Eq. (14), the real variance can be represented using the spectral density function for large T:

$$var(\bar{X}) \approx \frac{2\pi f(0)}{T} \quad (15)$$

Since the spectral density function cannot be determined directly, it must be estimated. The periodogram, which captures the frequency pattern of the samples, serves as a sample spectral density. The periodogram $I(\omega_k)$ is defined as:

$$I(\omega_k) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t e^{-it\omega_k} \right|^2 \quad (16)$$

for frequency $\omega_k = \frac{2\pi T}{k}$. Given that the zero-frequency value of the periodogram is biased and not a consistent estimator of $f(0)$, we use a local average of periodogram ordinates near the zero frequency:

$$\hat{f}(0) = \frac{1}{M} \sum_{m=1}^{M} I(\omega_m) \quad (17)$$

where M is a positive integer that satisfies $M \leq T$. Combining Eq. (15) and Eq. (17), we estimate the real variance using the periodogram:

$$var(\bar{X}) = \frac{2\pi}{MT} \sum_{k=1}^{M} I(\omega_k) \quad (18)$$

Using a Taylor series analysis, the bias and variance of our zero-frequency spectral density function estimation are [5]:

$$E\left[\hat{f}(0)\right] - f(0) \approx \frac{\pi f'(0)(M+1)}{T} \quad (19)$$

$$var\left(\hat{f}(0)\right) \approx \frac{1}{M^2} \sum_{m=1}^{M} \hat{f}(\omega_m) \quad (20)$$

By increasing M, the variance of the estimation can be reduced, but the bias will increase, indicating a trade-off. Moreover, many active cycles are crucial for minimizing the bias of the estimation.

2.5. Two new methods to estimate real variance

One new approach modifies the original method to create cycle-wise uncorrelated but parameter-wise correlated samples. Using an AR (1) model, the cycle-wise correlation is removed by estimating the correlation coefficient and generating uncorrelated parameters. For Gaussian-distributed samples, multiplying by the Cholesky matrix achieves parameter-wise correlation without needing iterations. The schematic of this method is illustrated in Fig. 3.

The second method combines the AR (1) model with the Spectral Analysis Method (SAM) to estimate real variance. Since the iDTMC method uses a limited number of active cycles, the resulting small sample size can lead to significant bias and variance in SAM estimates. To mitigate this, we use the AR (1) model to generate additional samples, enabling SAM to be applied with reduced bias and variance. The schematic of this method is shown in Fig. 4.

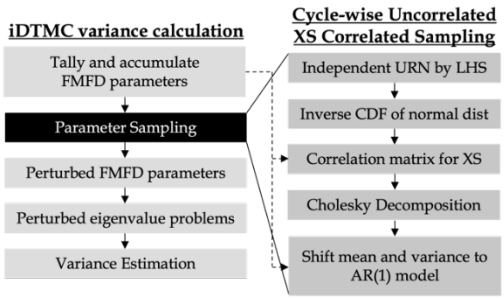

Figure 3. Real variance estimation using AR (1) model

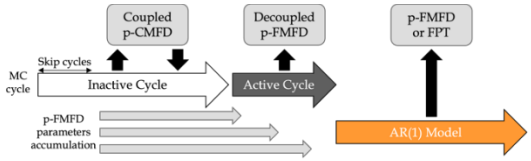

Figure 4. Real variance estimation using AR (1) model and SAM

3. Numerical Results

For real variance estimation, we solved the same SMR problem with identical histories and cycles as in [2]. All the calculation has been done with 240 Intel® Xeon® Gold 6148 cores parallel calculation.

The estimation result using AR (1) model is shown in Fig. 5, alongside real, apparent, and current methods. Real variance was determined as 15.29 pcm using a 45-batch calculation. The computing time for the current method was 82.20 minutes, while the new method required 73.88 minutes. By the 40th cycle, the estimated real standard deviation was 8.39 pcm, compared to 15.29 pcm for the actual value and 12.68 pcm for the current method. Fig. 6 presents a heat map of the correlation coefficients at the center surface, using the AR (1) model.

Results using the AR (1) model combined with the SAM are also shown in Fig. 5. For each active cycle, 100 parameters were generated, with a computing time of 78.41 minutes. At the 40th cycle, the estimated real

standard deviation was 9.81 pcm (M=1), 9.73 pcm (M=2), and 8.08 pcm (M=3). The periodogram for cycles 37 to 40 is depicted in Fig. 7.

We also tested with parameters from the AR (1) model without multiplying Cholesky matrix. The results, shown in Fig. 5 as 'AR (1)* model', had a computing time of 74.76 minutes and estimation of 17.13 pcm at 40th cycle.

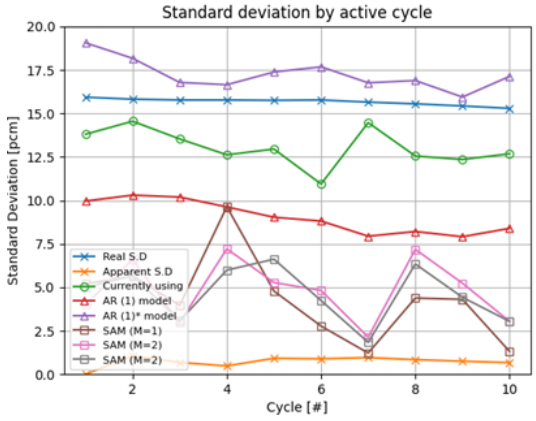

Figure 5. Real (blue), apparent(orange), currently using(green) and AR (1) model (red), AR (1)* model(purple), SAM with M=1(brown), M=2(pink) and M=3(gray) estimated standard deviation by active cycle

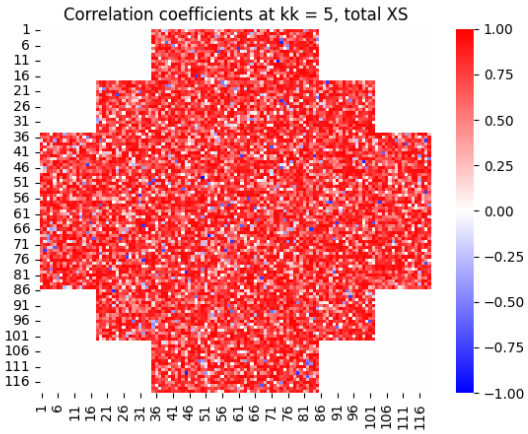

Figure 6. Heat map of total cross-section correlation coefficients calculated from AR (1)

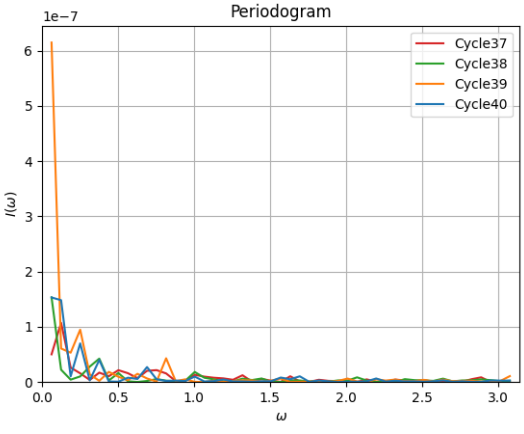

Figure 7. Periodogram from cycle 37(red), 38(green), 39(orange) and 40(blue)

4. Conclusions

Based on the numerical results, we identified significant cycle-wise correlations across all nodes. To address this, we eliminated these correlations, estimated using accumulated parameters and the AR (1) model, when sampling parameters. Additionally, we attempted to replicate cycle-wise correlated parameters using the AR (1) model and estimate the real variance with the SAM. Among the methods tested, the AR (1) model without multiplying the Cholesky matrix produced the most accurate results.

When estimating real variance using SAM with parameters adjusted by the Cholesky matrix, the periodogram revealed a lack of specific cycle-wise correlation, suggesting that the correlation matrix is influenced more by cycle-wise correlations than by parameter's one. Consequently, we developed a new correlation matrix that excludes cycle-wise correlation influences, leading to improved variance estimation using the AR (1) model compared to the original one.

These results suggest that the currently used correlation matrix may not effectively account for parameter-wise correlations and is overly influenced by cycle-wise correlations. Furthermore, the AR (1) model, when used without the Cholesky matrix, may offer a more accurate real variance estimation. Since the normal distribution does not require iterative processes to align with the original correlation, the computational time is reduced compared to the current method. Further research is necessary to confirm the reliability of AR (1) model. Additionally, exploring the correlation of other parameters, such as spatial correlations or correction factors in p-FMFD calculations, is needed for a more comprehensive understanding.

ACKNOWLEDGEMENT

This work was supported by the Korea Energy Technology Evaluation and Planning grant RS-2023-00304393 funded by the Korean government.

REFERENCES

[1] Inhyung Kim, Development of a Deterministic Truncation of Monte Carlo Solution for a Pin-Resolved Nuclear Reactor Analysis, 2021. KAIST, Ph.D. Dissertation.
[2] Inyup Kim and Yonghee Kim, Real variance estimation in iDTMC-based depletion analysis, Nuclear Engineering and Technology, Vol. 55, No. 11, pp. 4228-4237, 2023.
[3] O. JACQUET et al., Eigenvalue Uncertainty Evaluation in MC Calculations, Using Time Series Methodologies, Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications, p. 703, Springer (2001).
[4] J. P. Brockwell and A. R. Davis, Time Series: Theory and Methods, 2nd ed., Springer, New York, 1991.
[5] L. Jin and K. Banerjee, Variance Estimation in Monte Carlo Eigenvalue Simulations Using Spectral Analysis Method, Nuclear Science and Engineering, Vol. 191, No. 3, pp. 248-261, 2018.