

STATIONARITY AND SOURCE CONVERGENCE DIAGNOSTICS IN MONTE CARLO CRITICALITY CALCULATION

Taro Ueki and Forrest B. Brown

Diagnostics Applications Group (X-5), Applied Physics Division
Los Alamos National Laboratory
PO Box 1663, MS F663, Los Alamos, NM 87545
ueki@lanl.gov; fbrown@lanl.gov

ABSTRACT

In Monte Carlo (MC) criticality calculations, source error propagation through the stationary cycles and source convergence in the settling (inactive) cycles are both dominated by the dominance ratio (DR) of fission kernels, i.e., the ratio of the second largest to largest eigenvalues. For symmetric two fissile component systems with DR close to unity, the extinction of fission source sites can occur in one of the components even when the initial source is symmetric and the number of histories per cycle is larger than one thousand. When such a system is made slightly asymmetric, the neutron effective multiplication factor (k_{eff}) at the inactive cycles does not reflect the convergence to stationary source distribution. To overcome this problem, relative entropy (Kullback Leibler distance) is applied to a slightly asymmetric two fissile component problem with a dominance ratio of 0.9925. Numerical results show that relative entropy is effective as a posterior diagnostic tool.

Key Words: Monte Carlo, criticality, stationarity, source convergence

1. INTRODUCTION

It has been argued in the MC criticality analysis community that source error propagation through stationary iteration cycles is governed by the DR of fission kernels. [1] Researchers in favor of the argument have also claimed for decades that when DR is about unity, the cycle to cycle correlation (autocorrelation) of MC source distribution is strong. Figure 1 shows the trajectories of the first order autoregressive process with an autocorrelation coefficient of 0.999 driven by independent standard normal noises. One can easily observe that the trajectories do not appear to be in equilibrium and the crossing of the true mean of zero often does not occur over a thousand cycles even though the process is stationary. A similarly unstable phenomenon of MC source has been actively investigated for loosely coupled fissile component systems [2,3,4]. Recent work about DR and autocorrelation in MC criticality calculations has shown that when DR is nearly unity, the autocorrelation of the tallies of k_{eff} is not strong and may be negligible while the autocorrelation of MC source distribution is strong. [5] Therefore, in terms of computing k_{eff} alone, the difficulty in MC criticality calculations associated with large values of DR is when to start the tallying of k_{eff} , i.e., how to diagnose the convergence of source distribution. To meet this challenge, relative entropy [6] is defined so as to be utilized as a posterior diagnostic tool and is applied to a slightly asymmetric two fissile component problem with a dominance ratio of 0.9925.

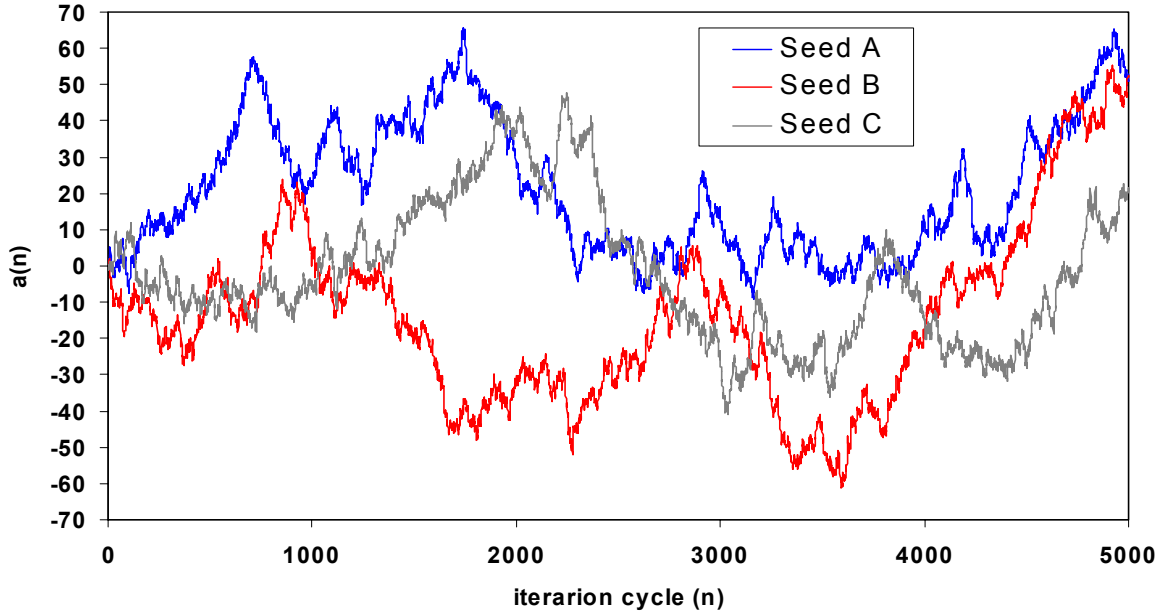


Figure 1: Trajectory of $a(n)=0.999*a(n-1)+b(n)$ for three independent initial random number seeds where $a(0)=0$, $b(n)$'s are independent and $b(n)\sim N(0,1)$ (standard normal)

2. THEORETICAL BACKGROUND

Let $F(\vec{r}' \rightarrow \vec{r})$ be the expected number of the first generation descendant particles per unit volume at \vec{r} resulting from a particle born at \vec{r}' . In the case of a position independent energy spectrum, $F(\vec{r}' \rightarrow \vec{r})$ is the fission kernel defined by the product of energy and angular spectrums, an inverse transport operator and a fission operator, with the last operator defined as $\int \int v\Sigma_f(\vec{r}, E)\psi(\vec{r}, E, \vec{\Omega})d\Omega dE$ for the operand ψ and the fissile descendant generation cross section $v\Sigma_f$. The eigenfunctions and eigenvalues of F are denoted by S_j and k_j :

$$S_j(\vec{r}) = \frac{1}{k_j} \int S_j(\vec{r}')F(\vec{r}' \rightarrow \vec{r})dr' \tag{1}$$

where k_j are ordered as $k_0 > |k_1| > |k_2| > \dots$. As in previous work [1], the eigenvalue k_j 's are assumed to be discrete. Note that k_{eff} is the largest eigenvalue k_0 and S_0 is called the fundamental mode eigenfunction and assumed to be normalized to k_0 :

$$\int S_0(\vec{r})dr = k_0. \tag{2}$$

The normalization condition (2) cannot generally be assumed for S_j , $j \geq 1$ because in symmetric problems eigenfunctions may integrate to zero for some of the non-fundamental modes. In order

to simplify later derivations, the following normalization scheme is imposed on the nonfundamental mode eigenfunctions:

$$S_j(\vec{r}) \leftarrow \frac{k_j S_j(\vec{r})}{\int S_j(\vec{r}) dr} \text{ when } \int S_j(\vec{r}) dr \neq 0, \quad (3)$$

i.e., the whole domain integral of $S_j(\vec{r})$ is normalized to the corresponding eigenvalue as far as $\int S_j(\vec{r}) dr \neq 0$ and no specification is made otherwise. The source (distribution of fission sites) after simulating the m -th stationary cycle in a MC criticality calculation is written as

$$\hat{S}^{(m)}(\vec{r}) = NS(\vec{r}) + \sqrt{N} \hat{e}^{(m)}(\vec{r}), \quad m \geq 0 \quad (4)$$

where $\hat{e}^{(m)}(\vec{r})$ is the fluctuating component of the stationary source, N the number of particle histories per cycle, the hats indicate a realization of stochastic quantities, and $S(\vec{r})$ is the expected value (ensemble average) of $\hat{S}^{(m)}(\vec{r})/N$:

$$S(\vec{r}) = \frac{1}{N} E[\hat{S}^{(m)}(\vec{r})]. \quad (5)$$

Note that Eq. (5) implies $E[\hat{e}^{(m)}(\vec{r})] = 0$. The scaling by N and \sqrt{N} in (4) is based on the random nature of individual particle tracking processes. In other words, the relative fluctuation of particle population $\int [\hat{S}^{(m)}(\vec{r}) - E[\hat{S}^{(m)}(\vec{r})]] dr / \int E[\hat{S}^{(m)}(\vec{r})] dr$ can be scaled by the inverse of the square root of population when the population is sufficiently large. In addition, the number of particle histories is assumed to be fixed throughout cycles. The bias of $S(\vec{r})$ is of order $1/N$ for discretized [7] and continuous [8] models:

$$S(\vec{r}) - S_0(\vec{r}) = O(1/N). \quad (6)$$

The expected value (ensemble average) of the normalized source conditional on $\hat{S}^{(m-1)}(\vec{r})$ is written as

$$\frac{N \hat{S}^{(m-1)}(\vec{r})}{\int \hat{S}^{(m-1)}(\vec{r}') dr'}.$$

Then, the random noise component $\hat{e}^{(m)}(\vec{r})$ resulting from the starter selection and subsequent particle tracking can be introduced as

$$\sqrt{N} \hat{e}^{(m)}(\vec{r}) \equiv \hat{S}^{(m)}(\vec{r}) - \frac{N \int F(\vec{r}' \rightarrow \vec{r}) \hat{S}^{(m-1)}(\vec{r}') dr'}{\int \hat{S}^{(m-1)}(\vec{r}'') dr''}. \quad (7)$$

Recent work [5] has shown that Eqs. (4) and (7) yield

$$E[\hat{\varepsilon}^{(p)} \hat{\varepsilon}^{(q)}] = 0, \quad p > q, \quad (8)$$

and

$$\hat{e}^{(m)}(\vec{r}) = \int A(\vec{r}' \rightarrow \vec{r}) \hat{e}^{(m-1)}(\vec{r}') dr' + \hat{\varepsilon}^{(m)}(\vec{r}) + O(N^{-1/2}), \quad (9)$$

where A is defined as

$$A(\vec{r}' \rightarrow \vec{r}) \equiv \frac{1}{k} F(\vec{r}' \rightarrow \vec{r}) - \frac{1}{k^2} \int F(\vec{q} \rightarrow \vec{r}) S(\vec{q}) dq \quad (10)$$

with

$$\int S(\vec{r}) dr = k. \quad (11)$$

Eqs. (1), (2), (6) and (11) enable one to rewrite Eq. (9) as

$$\hat{e}^{(m)}(\vec{r}) = \int A_0(\vec{r}' \rightarrow \vec{r}) \hat{e}^{(m-1)}(\vec{r}') dr' + \hat{\varepsilon}^{(m)}(\vec{r}) + O(N^{-1/2}) \quad (12)$$

where A_0 is defined as

$$A_0(\vec{r}' \rightarrow \vec{r}) \equiv \frac{1}{k_0} [F(\vec{r}' \rightarrow \vec{r}) - S_0(\vec{r})]. \quad (13)$$

Eqs. (8) and (12) are the functional version of the first order autoregressive process when N is sufficiently large. Introducing the operator notation of A_0 as

$$\hat{e}^{(m)}(\vec{r}) = A_0 \hat{e}^{(m-1)}(\vec{r}) + \hat{\varepsilon}^{(m)}(\vec{r}) + O(N^{-1/2}),$$

Def. (13) and Eqs. (1)-(3) yield

$$A_0 S_0(\vec{r}) = 0, \quad (14)$$

and

$$A_0^i S_j(\vec{r}) = \begin{cases} \left(\frac{k_j}{k_0}\right)^i [S_j(\vec{r}) - S_0(\vec{r})] & \text{when } \int S_j(\vec{r}) dr \neq 0 \\ \left(\frac{k_j}{k_0}\right)^i S_j(\vec{r}) & \text{when } \int S_j(\vec{r}) dr = 0 \end{cases}, \quad i \geq 1, j \geq 1. \quad (15)$$

Eqs. (14) and (15) imply that the cycle to cycle error propagation is governed by the ratio of the non-fundamental to fundamental mode eigenvalues if the completeness assumption of S_j 's is valid. Therefore, when DR is close to unity and the ratio of the second to fundamental mode eigenvalues is much smaller than unity, the fluctuation of MC source may show the behavior observed in the trajectory of the first order autoregressive process $a(n+1) = \phi a(n) + b(n)$ with $\phi \approx 1$.

Previous work in statistics showed that the discretized version of Eq. (12) without the $O(N^{-1/2})$ term was equivalent to the vector autoregressive moving average process of order p and $p-1$ (ARMA($p, p-1$)). [9] Here, ARMA models are a generalization of autoregressive processes and ARMA(1,0) is the first order autoregressive process. Thus, the fluctuation of MC source could in principle be explained in the framework of ARMA processes. However, we do not delve into this aspect of model identification. In the next section, we examine whether or not the duration of no crossing of true mean (ensemble average) lasts over a thousand cycles for the MC source of a two fissile component problem with a DR larger than 0.999 as observed in the first order autoregressive process in Figure 1.

3. ONE GROUP TWO FISSILE COMPONENT PROBLEMS

This section shows numerical results of symmetric and slightly asymmetric two fissile component problems having DR larger than 0.99. The problem specification is as follows:

Problem 1

- 5 region slab, with void boundary conditions on both sides and one-group isotropic cross sections,
- the regions are (left to right) 1.0, 1.0, 5.0, 1.0, and 1.0 cm thickness,
- the materials are (left to right) 2 (fuel), 1 (scatterer), 3 (absorber), 1, and 2,
- material 1 (scatterer)
 $\Sigma_{total} = 1.0 \text{ cm}^{-1}$, $\Sigma_{scattering} = 0.8 \text{ cm}^{-1}$, $\Sigma_{capture} = 0.2 \text{ cm}^{-1}$,
- material 2 (fuel)
 $\Sigma_{total} = 1.0 \text{ cm}^{-1}$, $\Sigma_{scattering} = 0.8 \text{ cm}^{-1}$, $\Sigma_{capture} = 0.1 \text{ cm}^{-1}$, $\Sigma_{fission} = 0.1 \text{ cm}^{-1}$, $\nu = 3.0$,
- material 3 (absorber)
 $\Sigma_{total} = 1.0 \text{ cm}^{-1}$, $\Sigma_{scattering} = 0.1 \text{ cm}^{-1}$, $\Sigma_{capture} = 0.9 \text{ cm}^{-1}$.

Problem 2

- Same as problem 1 except 1.01 cm thickness for the rightmost slab.

The eigenvalues of Problems 1 and 2 were computed by the Green's function methods [10] and the ratio of the first five non-fundamental to fundamental mode eigenvalues are shown in Table I. The DR values therein were also obtained by analyzing the spectral radius of the outer iterations in a discontinuous finite element discrete ordinate computation [11], where the initial flux in

Problem 1 is a random guess in order to avoid the suppression by a symmetric initial flux guess. Figure 2 shows the cycle progression of the MC source at the right fuel in Problem 1 for 2000 and 200000 histories per cycle computations. The extinction of fission source sites is observed in the 2000 histories per cycle computation even though the initial source is correct in the resolution associated with two fuel bins. The fluctuation of MC source significantly decreases in the 200000 histories per cycle computation. The duration of no crossing of the mean value of 0.5 is often more than a thousands cycles in both the computations. This may be troublesome because active cycles for MC tallying would have to be at least 10000 cycles for problems with DR larger than 0.999. If a middle plane reflective boundary condition is employed, the DR of Problem 1 is reduced to 0.305 and the difficulty in the computation disappears. Moreover, it has been proved for symmetric two fissile components systems that the MC source fluctuation does not affect k_{eff} tallies in the first order approximation. [2] Therefore, real difficulty exists in slightly asymmetric problems even though the DR may become smaller than that of the symmetric counterpart.

Table I. Ratio of non-fundamental to fundamental mode eigenvalues

	Problem 1	Problem 2
First (DR)	0.99957	0.99250
Second	0.30465	0.30563
Third	0.30464	0.30243
Fourth	0.16774	0.16827
Fifth	0.16774	0.16652

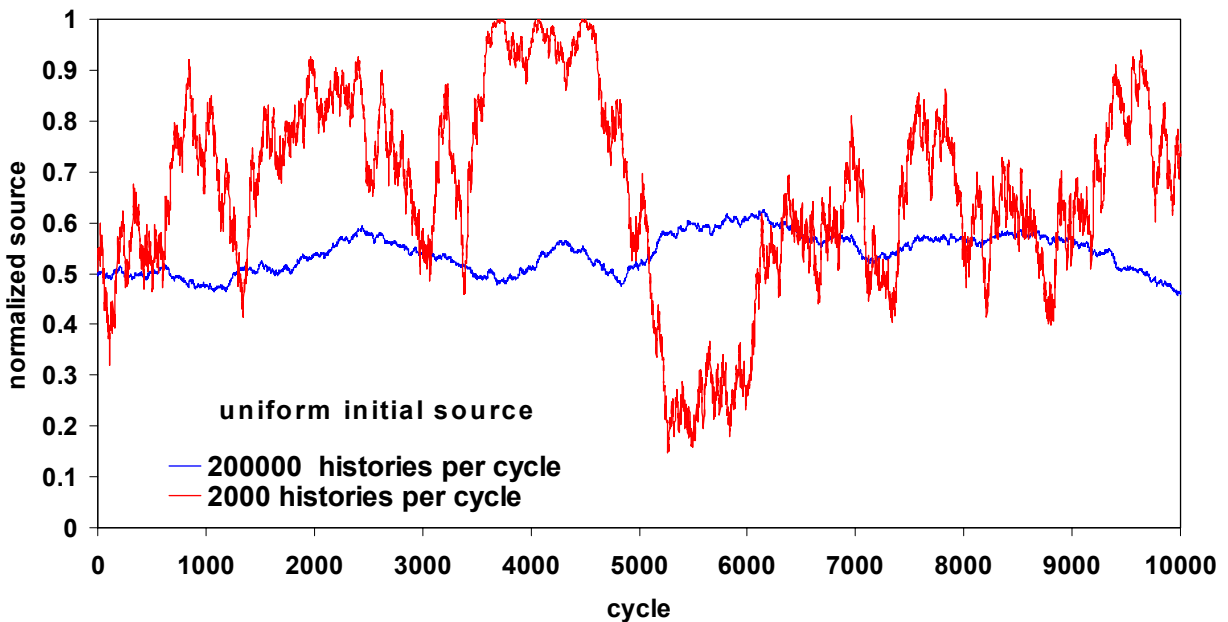


Figure 2: Normalized source at right fuel in Problem 1

Figure 3 shows the cycle progression of k_{eff} and MC source at right fuel in Problem 2 divided by their respective true mean. Here, the fundamental mode eigenvalue 0.427425 computed by the Green's function method is used as the true stationary mean of k_{eff} , and the average of MC source over 16 replicas of 1250 inactive and 1250 active cycles with 10000 histories per cycle, 0.9744 ± 0.0013 (right + left = 1), is used as the true mean of the MC source at the right fuel. One can observe that the duration of no crossing of the true mean value often becomes more than eighty cycles for the MC source; 379-463, 482-571, 726-815 and 905-1000 cycles. However, a more striking phenomenon is that k_{eff} does not reflect the source convergence process at all. Figure 4 shows the MC confidence interval coverage rate of the fundamental mode eigenvalue of Problem 2 computed by the Green's function method [10]. Figure 5 shows the ensemble average of one σ (68%) confidence intervals of k_{eff} where the end points of the confidence intervals are averaged over 2000 replicas and thus the interval lengths are not multiplied by the factor $1/\sqrt{2000}$. The comparison of Figures 3 with 4 and 5 clearly indicates the necessity of the diagnostic method of the convergence of MC source distribution. On the other hand, (2), (3), and (15) yield

$$\int A_0^i S_j(\vec{r}) dr = \begin{cases} k_0 \left(\frac{k_j}{k_0}\right)^i \left(\frac{k_j}{k_0} - 1\right) & \text{when } \int S_j(\vec{r}) dr \neq 0 \\ 0 & \text{when } \int S_j(\vec{r}) dr = 0 \end{cases}, i \geq 1, j \geq 1. \quad (16)$$

Due to the factor $k_j/k_0 - 1$, the effect of DR does not strongly appear in the autocorrelation of k_{eff} 's when DR is close to unity. [5] Therefore, in terms of computing k_{eff} alone, effective MC source convergence diagnostics would indeed improve confidence interval estimation for problems with DR close to unity.

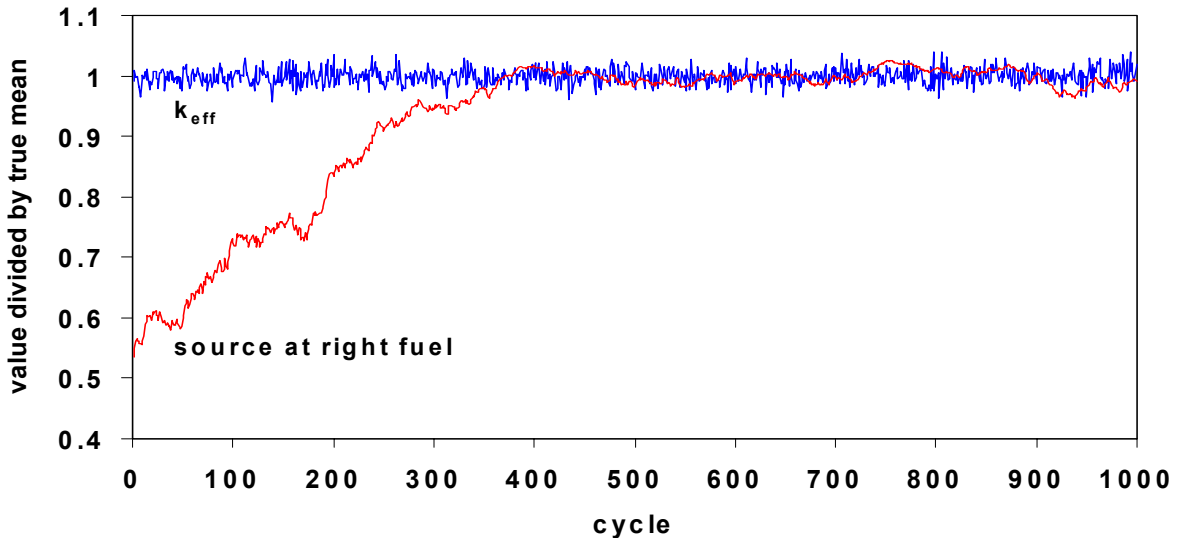


Figure 3: Source at right fuel and k_{eff} divided by their respective true mean in Problem 2 (5000 histories per cycle; uniform initial source)

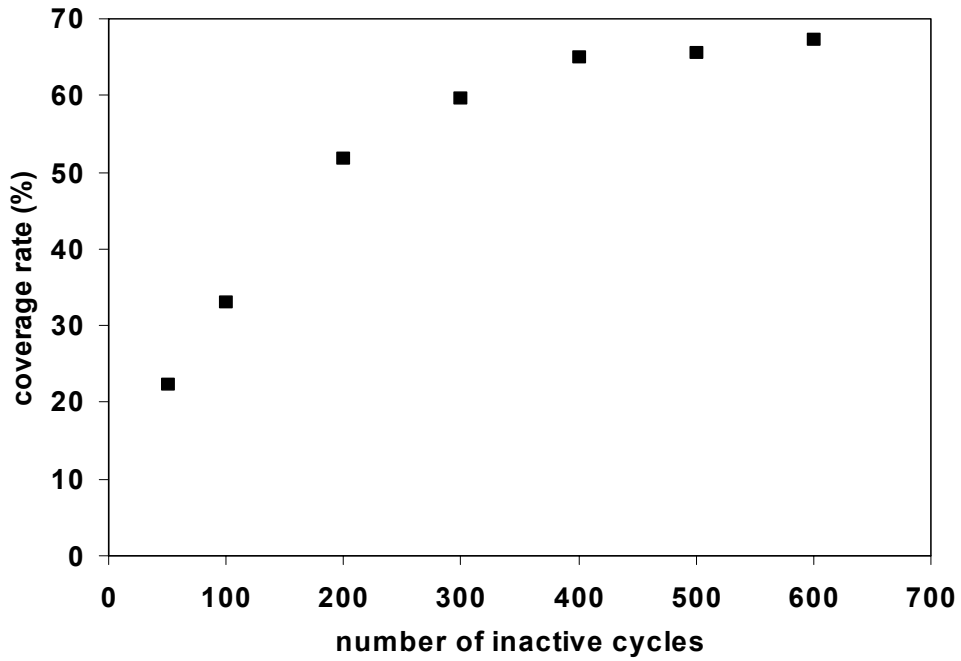


Figure 4: One σ (68%) confidence interval coverage rate of the fundamental mode eigenvalue computed by Green's function method over 2000 replicas of 200 active cycle simulations with 5000 histories per cycle (Problem 2)

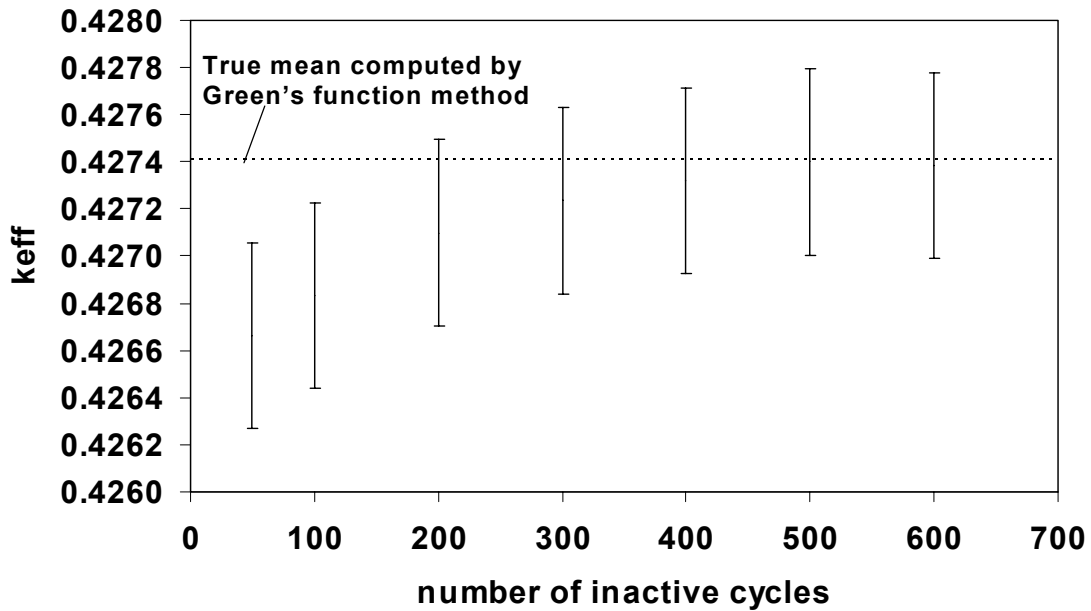


Figure 5: Ensemble average of one σ (68%) confidence intervals of k_{eff} over 2000 replicas of 200 active cycle simulations with 5000 histories per cycle (Problem 2)

4. POSTERIOR SOURCE CONVERGENCE DIAGNOSIS

In this section, we define the relative entropy of MC source distribution so as to be utilized as the posterior diagnosis of source convergence. [5] The relative entropy (Kullback Leibler distance) of the normalized binned sources S^B and T^B is defined as [6]

$$D(S^B \parallel T^B) = \sum_{i=1}^B S^B(i) \ln \left(\frac{S^B(i)}{T^B(i)} \right),$$

where B is the number of spatial bins and i the bin number. $D(S^B \parallel T^B)$ is a statistical distance between S^B and T^B in a sense that $D(S^B \parallel T^B)$ is nonnegative and becomes zero only when $S^B(i) = T^B(i)$ for all bins. In addition, $D(S^B \parallel T^B)$ satisfies the pair convexity [6]

$$D(\lambda S_1^B + (1-\lambda)S_2^B \parallel \lambda T_1^B + (1-\lambda)T_2^B) \leq \lambda D(S_1^B \parallel T_1^B) + (1-\lambda)D(S_2^B \parallel T_2^B), 0 \leq \lambda \leq 1,$$

where S_1^B , S_2^B , T_1^B and T_2^B are all binned sources normalized to unity. Then, the S^B convexity follows by setting $T_1^B = T_2^B = T^B$:

$$D(\lambda S_1^B + (1-\lambda)S_2^B \parallel T^B) \leq \lambda D(S_1^B \parallel T^B) + (1-\lambda)D(S_2^B \parallel T^B), 0 \leq \lambda \leq 1,$$

i.e., $D(S^B \parallel T^B)$ is convex when viewed as a function of S^B . The formal meaning of $D(S^B \parallel T^B)$ in information theory is the penalty of the data compression limit (minimum of expected prefix-free codeword length) of the random bins generated from S^B under the false assumption that the distribution generating the bins is T^B [6] The data compression limit of the random bins generated from S^B under the presence of the knowledge of S^B is nearly equal to Shannon entropy of S^B , i.e., the randomness of the distribution S^B . [6] They are formerly stated as

$$\begin{aligned} & \left(\text{Shannon entropy of } S^B \right) \\ & \leq \left(\begin{array}{l} \text{Data compression limit of the random bins generated from } S^B \\ \text{under the presence of the knowledge of } S^B \end{array} \right) \\ & \leq \left(\text{Shannon entropy of } S^B \right) + 1 \\ & \left(\text{Shannon entropy of } S^B \right) + D(S^B \parallel T^B) \\ & \leq \left(\begin{array}{l} \text{Data compression limit of the random bins generated from } S^B \text{ under} \\ \text{the false assumption that the distribution generating the bins is } T^B \end{array} \right) \\ & \leq \left(\text{Shannon entropy of } S^B \right) + D(S^B \parallel T^B) + 1 \end{aligned}$$

These inequalities involve the concept of the penalty resulting from the hypothesis that T^B is stationary MC source distribution under the condition that one has observed S^B . On the other hand, the S^B convexity is a desired characteristic for measuring the distance to a fixed reference

distribution. Therefore, $D(S^B||T^B)$ can be interpreted as a measure of the inefficiency of assuming that the true distribution in a MC simulation is T^B when the observed distribution is S^B . One can then utilize the relative entropy in the following manner. After all cycles are simulated:

- 1) Compute $T^B(i)$ by averaging MC source over the second half of active cycles.
- 2) Plot $D(S^B||T^B)$ for each source S^B through all cycles starting at the initial cycle.
- 3) Check whether $D(S^B||T^B)$ crosses the average of $D(S^B||T^B)$ over the second half of the active cycles before the first active cycle.

Figure 6 shows the posterior computation of relative entropy in Problem 2 following 1)-3) for the 200 inactive and 200 active cycle simulations with 20000 histories per cycle. One can easily observe that the crossing of the mean of relative entropy over the second half of active cycles does not occur before the first active cycle. Figure 7 shows the same posterior computation of relative entropy in Problem 2 for the 500 inactive and 1000 active cycle simulations with 5000 histories per cycle. A remarkable characteristic in Figure 7 is convergence performance difference for the different seeds. MC source convergence is observed to occur on average after about 350 cycles. Table II shows the cycle at the first crossing of the mean of relative entropy over the second half of active cycles for the simulations in Figure 7. The results in Table II are consistent with the trend in Figures 4 and 5.

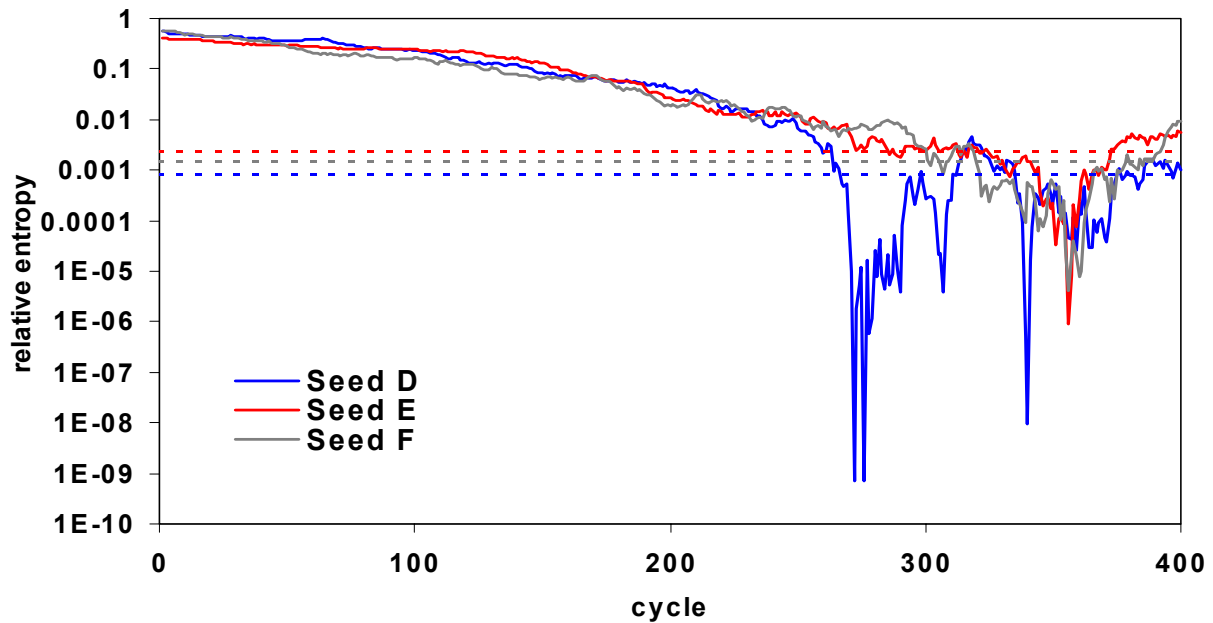


Figure 6: Posterior computation of relative entropy for 200 inactive and 200 active cycle simulations with 20000 histories per cycle in Problem 2 (dotted lines are the mean level over 301-400 cycles)

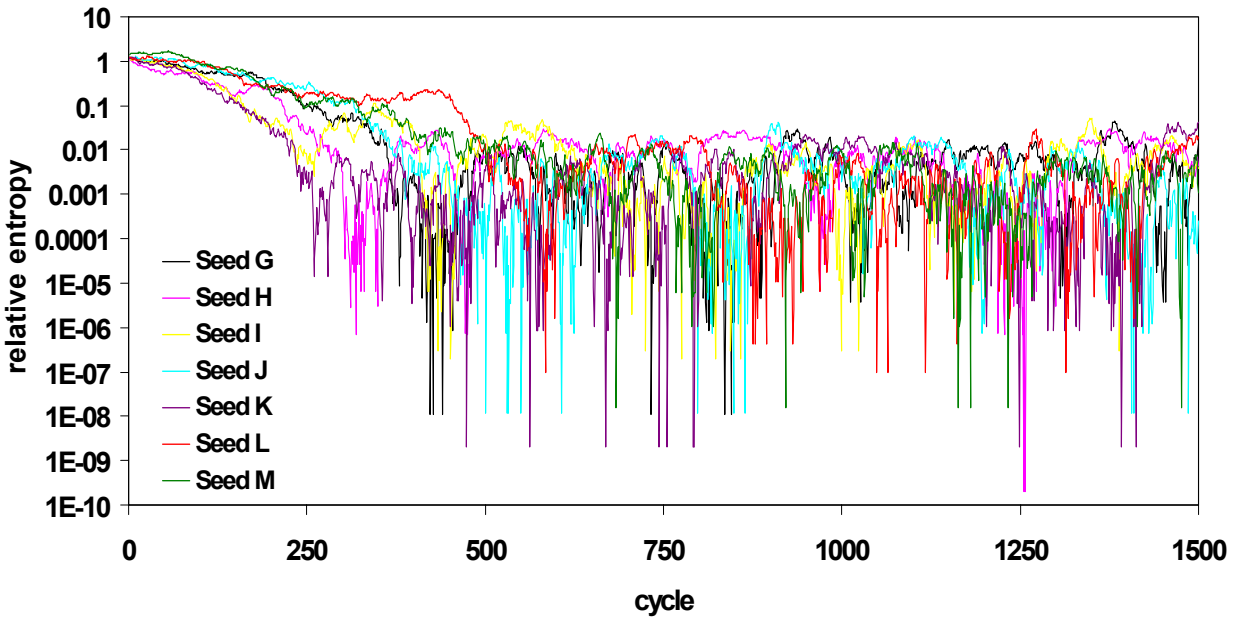


Figure 7: Posterior computation of relative entropy for 500 inactive and 1000 active cycle simulations with 5000 histories per cycle in Problem 2

Table II: Cycle at first crossing of the mean of relative entropy over the second half of active cycles for the simulations in Figure 7.

Seed	G	H	I	J	K	K	M
Cycle at first crossing	364	283	235	386	235	505	540

5. CONCLUSIONS

We have discussed similarities between the first order autoregressive process and the MC source in two fissile component problems with DR close to unity both numerically and mathematically. The possibility of modeling stationary MC source as ARMA(p,p-1) processes was pointed out based on mathematical expressions derived and previous work on vector ARMA processes [9]. We have successfully applied the relative entropy of MC source distribution to the convergence diagnosis of a slightly asymmetric two fissile component problem with DR close unity. The diagnostic proposed can be implemented for any number of space bins. Moreover, it is the direct evaluation of the penalty associated with the stationarity hypothesis of MC source distribution and does not need the additional knowledge, parameter estimation, or parametric assumption of a problem simulated, which is often required in convergence diagnostic methods available in operations research and statistics [12,13]. The only restriction in the usage of relative entropy is a posterior diagnostic nature.

ACKNOWLEDGMENTS

We thank Drew E. Kornreich and D. Kent Parsons for computing the dominance ratio and higher mode eigenvalues of the two fissile component problems studied in this work.

REFERENCES

1. D.B. Macmillan, "Monte Carlo Confidence Limits for Iterated Source Calculations," *Nuclear Science and Engineering*, **50**, pp.73-75 (1973).
2. T. Yamamoto, T. Nakamura and Y. Miyoshi, "Fission Source Convergence of Monte Carlo Criticality Calculations in Weakly Coupled Fissile Arrays," *Journal of Nuclear Science and Technology*, **37**, pp.41-52 (2000).
3. F. B. Brown, R. C. Little, A. V. Sood, D. K. Parsons, and T. A. Wareing, "MCNP Calculations for the OECD/NEA Source Convergence Benchmarks for Criticality Safety Analysis," LA-UR-01-5181 (2001).
4. R. N. Blomquist and E. M. Gelbard, "Adaptive Implementations of the Monte Carlo Power Method," *Nuclear Science and Engineering*, **141**, pp.85-100 (2002).
5. T. Ueki, F. B. Brown, D.K. Parsons and D.E. Kornreich, "Autocorrelation and Dominance Ratio in Monte Carlo Criticality Calculations," LA-UR-02-5700 (2002).
6. T. M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, USA (1991).
7. R.J. Brissenden and A.R. Garlick, "Biases in the Estimation of Keff and its Error by Monte Carlo Methods," *Annals of Nuclear Energy*, **13**, 2, 63 (1986).
8. T.M. Sutton and F.B. Brown, "Analysis of the Monte Carlo Eigenvalue Bias," (unpublished).
9. H. Akaike, "Markovian Representation of Stochastic Processes and Its Application to the Analysis of Autoregressive Moving Average Processes," *Annals of the Institute of Statistical Mathematics*, **26**, pp.363-387 (1974).
10. D.E. Kornreich and B.D. Ganapol, "The Green's Function Method for Nuclear Engineering Applications," *Nuclear Science and Engineering*, **126**, 293 (1997).
11. T.A. Wareing, J.M. McGhee, J.E. Morel and S.D. Pautz, "Discontinuous Finite Element S_N Methods on Three-Dimensional Unstructured Grids," *Nuclear Science and Engineering*, **138**, 256 (2001).
12. L.W. Schruben, "Detecting Initialization Bias in Simulation Output," *Operations Research*, **30**, 569 (1982).
13. M.K. Cowles and B.P. Carlin, "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review," *Journal of the American Statistical Association*, Vol.91, No.434, 883 (1996).