M&C 2017 - International Conference on Mathematics & Computational Methods Applied to Nuclear Science & Engineering, Jeju, Korea, April 16-20, 2017, on USB (2017)

Neutron Clustering in Monte Carlo Iterated-Source Calculations

Thomas M. Sutton and Anudha Mittal

Naval Nuclear Laboratory, Schenectady, New York, USA, thomas.sutton@unnpp.gov

Abstract – *Monte Carlo neutron transport codes generally use the method of successive generations to converge the fission source distribution to—and then maintain it at—the fundamental mode. Recently, a phenomena called 'clustering' has been noted that produces fission distributions that are very far from the fundamental mode. In this paper, a mathematical model of clustering in Monte Carlo is developed. The model draws on previous work for continuous-time birth-death processes, as well as methods from the field of population genetics.*

I. INTRODUCTION

Monte Carlo neutron transport codes generally use some variation on the method of successive generations [1] for criticality calculations. In this method, the starting locations for the neutrons in one generation are chosen from the locations of fission events occurring in the previous generation. Since the distribution of fission events is not known a priori, calculations start with some initial 'source guess'—for example a uniform fission source distribution. This method is the Monte Carlo equivalent of the well-known power method from matrix algebra. Since the power method converges to the fundamental eigenvector, one might hope that as the Monte Carlo calculation proceeds from one generation to the next the distribution of fission locations will converge to the fundamental-mode fission distribution of the neutron transport equation.

Recently, however, a phenomena called 'clustering' has been noted that produces fission location distributions that are very far from the fundamental-mode distribution. This phenomenon has been studied for actual neutron populations [2] as well as for the simulated neutron populations that are the basis of Monte Carlo neutron transport calculations [3,4]. These studies have utilized a time-dependent diffusion approximation to model the evolution of the neutron population. While this may be appropriate under certain conditions for actual neutron populations, it is an approximation for Monte Carlo calculations. The reason is that iterated-fission-source Monte Carlo proceeds by a series of discrete fission generations rather than by a continuous evolution in time.

The motivation is to provide a higher-fidelity theory which may then form the basis of further research into such areas as higher-order moments of the entropy function for the purpose of detecting fission source convergence [4], methods for detecting the presence of clustering, and methods for minimizing or even eliminating clustering.

This paper will also address a phenomenon not discussed in the previous works on clustering. As will be explained in detail in a later section, one may view the evolution of the neutron population in terms of a genealogical tree in which a neutron may trace its lineage back through previous generations via the fission process to one of the neutrons in the initial source guess. We show that as the Monte Carlo calculation proceeds, the neutrons in successive generations trace their lineages to ever fewer neutrons from the initial source. Given a sufficient number of generations, at some point all of the neutrons in the calculation will be descended from the same initial source neutron. This phenomenon is analogous to a process in population genetics—referred to as *fixation*—in which all alleles of a gene except one disappear from a population of organisms. We will take advantage of mathematical techniques developed in that field.

In Section II we describe the Monte Carlo model that will be assumed for the subsequent analyses. In Section III we introduce some useful terminology from population genetics. In Section IV we provide an example of clustering that will make later discussion more concrete. In Section V we discuss fixation, and compare the predictions of a theoretical model to the results of Monte Carlo calculations. In Section VI we recast the previous work done using a continuous-time assumption to the generational form corresponding to a Monte Carlo calculation. We use the theory developed to compute the mean-squared separation of neutron absorption sites, and compare the theoretical result to that of Monte Carlo calculations. Finally, in Section VII we provide some concluding remarks.

II. THE MONTE CARLO MODEL

In this section we describe the model problem that will be used in the analysis which follows as well as the Monte Carlo algorithm used to solve it. Both the problem and the algorithm are very simple, which allows a theoretical analysis that would be impossible in a more complex setting.

The model problem consists of a homogeneous multiplying medium within a three-dimensional cube with sides of length L. Reflecting boundary conditions are applied to all six surfaces. We further assume one-speed neutrons and isotropic scattering.

The Monte Carlo algorithm is purely analog with no survival biasing or other variance reduction method. The number of neutrons per generation is constant and denoted by N. The starting locations for the neutrons in the initial

generation are assumed to be sampled uniformly within the problem space. Since for our simple model every neutron will be absorbed (there being no leakage), the neutrons starting each generation will produce exactly N absorption locations. The starting locations for a neutron in any subsequent generation will be randomly chosen from all of the absorption locations created in the previous generation. Since the medium is homogeneous, each absorption location will be chosen with equal probability. Using this algorithm, some absorption locations may be chosen as the starting location for multiple neutrons while others may not be chosen at all.

III. TERMINOLOGY

In what follows, we will use certain terms from the fields of genealogy and population genetics. First, we will consider all neutrons in a given generation that are born at the same absorption location to be 'siblings', and the neutron that produced the absorption to be their 'parent'. One may then envision a genealogical tree with which a neutron's lineage may be traced from the current generation back to one of the starting neutrons in the first generation—its 'original ancestor'. A 'family' is defined as those neutrons that have the same original ancestor. We will also use the term 'most-recent common ancestor' (MRCA) of a group of neutrons in a given generation. If the neutrons in the group are siblings, their MRCA is their parent. For a group of neutrons that are not all siblings but are in the same family, their MRCA is that ancestor that is shared by all neutrons in the group occurring in the generation least far removed from the current generation. Neutrons in different families, of course, have no common ancestor.

IV. AN EXAMPLE OF CLUSTERING

In this section we provide a concrete example of clustering to motivate the analysis that follows. The model problem was run using the one-speed cross sections given in Table I for $L = 400\,\text{cm}$ and $N = 1{,}000$. The number of generations was 10,000. Figure 1 shows the locations of the source sites of the neutrons for the first generation in pink and the absorption sites for generations 1,000 and 10,000 in blue and black, respectively. As expected, the distribution of source sites in the first generation appears to be uniform. After 1,000 generations, however, the effect of clustering has become obvious with all 1,000 absorption sites being confined to a relatively small spatial region. After another 9,000 generations, the absorption sites are still tightly clustered not far from where they were after 1,000 generations. Despite the code being given the correct solution via the initial source guess, the solution has evolved to something that is quite different from a uniform distribution.

Table I. Macroscopic Cross Sections (cm^{-1})

Σ_T	Σ_S	Σ_A	Σ_F	$\nu\Sigma_F$
1.00	0.60	0.40	0.20	0.48

For the calculation illustrated in Fig. 1, it was found that beginning with generation 942 all of the neutrons belonged to the same family as defined in the previous section. The starting location of the original ancestor for the surviving family is indicated by the red point in Fig. 1. One can see that the center of mass (COM) of the cluster has not moved far from this point even after 10,000 generations.

In the remainder of this paper we will provide theoretical explanations of various aspects of this calculation. These include the conditions that lead to clustering, the size of the clusters, the cause of the extinction of families, and the number of generations it takes for all of the families except one to become extinct.

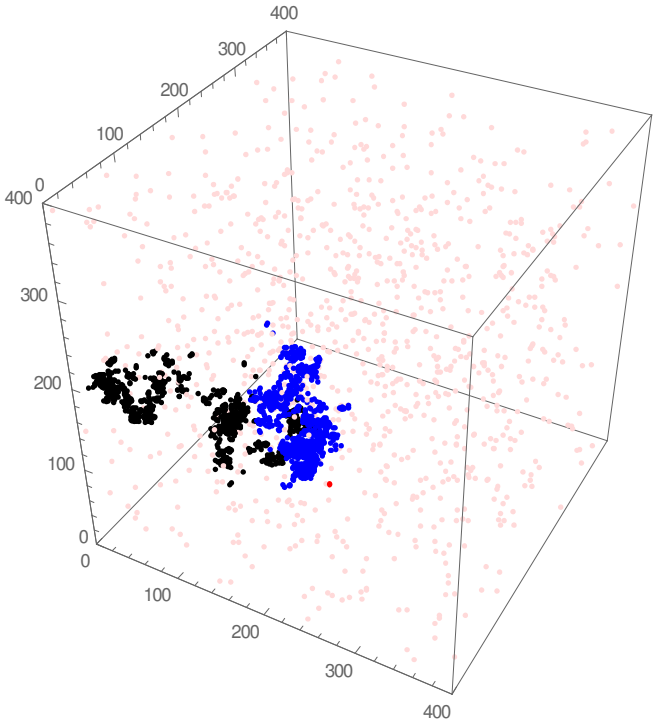

Fig. 1. Initial neutron source distribution (pink) and absorption distributions after 1,000 (blue) and 10,000 (black) generations. The red point marks the initial location of the original ancestor of all neutrons starting with generation 942.

V. EXTINCTION OF FAMILIES AND FIXATION

In the analysis of spatial clustering that follows in Section VI we will be concerned with pairs of neutrons within a generation, of which there are $N(N-1)/2$. For the initial generation, all of the pairs are uncorrelated since the starting location for each neutron is selected independently from that of the others. As the generations progress, spatial

correlation is introduced due to the possibility that some absorption locations may give rise to more than one neutron in the next generation. Since the probability that a given neutron will start from a given absorption location is $1/N$, the probability that any given absorption location is chosen as a starting location exactly k times in a generation is given by the binomial distribution

$$P_k = \frac{N!}{k!(N-k)!}\left(\frac{1}{N}\right)^k \left(1-\frac{1}{N}\right)^{N-k} . \quad (1)$$

The expected number of correlated neutron pairs produced per absorption location is $\nu_2/2$, where the second factorial moment is given by

$$\nu_2 = \sum_{k=0}^{N} P_k k(k-1) = \frac{N-1}{N} . \quad (2)$$

The total expected number of correlated pairs created in a generation with a MRCA in the previous generation is therefore $N\nu_2/2 = (N-1)/2$. The fraction of all pairs represented by these is

$$f = \frac{2}{N(N-1)} \frac{N\nu_2}{2} = \frac{1}{N} . \quad (3)$$

Since the total number of neutron pairs in a generation is fixed, the expected number of uncorrelated pairs and expected number of correlated pairs with a MRCA prior to the previous generation must each decrease by a corresponding factor of $1-f$ at each generation. At generation g, therefore, the expected fraction of uncorrelated pairs will be

$$\psi_{\mathrm{u}}(g) = \left(1-\frac{1}{N}\right)^{g-1} , \quad (4)$$

and the expected fraction of correlated pairs with a MRCA in generation g', where $1 \le g' \le g-1$, will be

$$\psi_{\mathrm{c}}(g'|g) = \frac{1}{N}\left(1-\frac{1}{N}\right)^{g-g'-1} . \quad (5)$$

From Eq. (4) one sees that the expected number of uncorrelated pairs is a monotonically decreasing function of the generation number. Since the members of an uncorrelated pair must come from different families, the implication is that the expected number of families remaining is also monotonically decreasing with increasing generation number.

Using Eq. (1), one can compute the expected fraction of families that go extinct after the first generation as

$$P_0 = \left(1-\frac{1}{N}\right)^N , \quad (6)$$

which for large N approaches $e^{-1} = 0.367879$.

To extend the analysis of how the number of families varies as the Monte Carlo calculation proceeds beyond the first generation, we will resort to two methods from the field of population genetics: the Wright-Fisher (W-F) model [5,6] and coalescent theory [7,8]. The W-F model is an idealized model of the genetic evolution of a population of organisms. Coalescent theory provides a quantitative method for analyzing the number of generations between the current one and the generation containing the MRCA of the members of some subset of the population.

This W-F model uses the following assumptions:

- the population size is constant,
- generations do not overlap in time,
- the parent of an individual is chosen uniformly with replacement from all of the individuals in the preceding generation, and
- there is no selection, mutation, or recombination.

Because of these assumptions, the W-F model is a much idealized representation of most real biological populations. The W-F model for haploid organisms (those with a single set of chromosomes), however, is exact for the Monte Carlo algorithm considered here. This is because the number of neutrons per generation is constant; the fission generations are sequential; the probability of selecting any one of the absorption locations in the previous generation as the source location of each neutron in the subsequent generation is uniform; and there are no equivalent concepts to selection, mutation, and recombination.

Using the W-F model, two important properties of the evolution of a population have been proved. First, in the limit as the number of generations goes to infinity the population becomes *fixed*. In genetics this means that all alleles but one disappear from a population. In our case, this means that all but one family becomes extinct for a sufficiently large number of generations. Second, the probability at some point in time that a particular allele will be the one that becomes fixed is given by its fraction of the total allele population at that time. For our application this means that at generation g the probability that a family will be the one that survives is equal to that family's fraction of the total population at that generation. Since at the beginning of the first generation each family represents the

same fraction $1/N$ of the population, each family initially has an equal probability of being the one that survives.

A detailed discussion of coalescent theory is beyond the scope of this work, so the method will only be briefly outlined below. The reader is directed to Ref. [8] for a very readable presentation of the theory. Consider a subset *n* of the *N* neutrons in a generation. Denote the probability that none of the neutrons in the subset share a common parent by q_n. If these neutrons have no common parents, then they must have *n* distinct ancestors in the previous generation. The probability that these *n* distinct ancestors themselves have distinct ancestors is also q_n. The probability that the members of the subset share no common ancestors in the previous $t-1$ generations, but that two or more share a common ancestor *t* generations in the past is thus

$$Q_n(t) = q_n^{t-1}(1-q_n). \tag{7}$$

Using the terminology of coalescent theory, a 'sample' will refer to some subset of the neutrons at the 'current' generation. The theory is developed by following the lineages of the members of the sample backwards in time through the branches of the genealogical tree. The merging of two branches in some past generation is called a 'coalescence'. It is assumed that at any generation the number of branches on the sample's tree is much less than the population size *N*, so that the probability of more than one coalescence per generation is negligible. Thus, as one looks back in time from the current generation, the number of branches on the sample's tree is reduced by one at coalescences and is constant between coalescences. Eventually, the there are just two branches left. These coalesce at the MRCA of the sample.

For the W-F model, we have the exact expression

$$q_n = \prod_{k=1}^{n-1}\left(1-\frac{k}{n}\right). \tag{8}$$

For $n \ll N$, we have the approximate expressions

$$q_n \cong 1-\frac{n(n-1)}{2N} \tag{9}$$

and

$$Q_n(t) \cong \frac{n(n-1)}{2N} e^{-\frac{n(n-1)}{2N}t}. \tag{10}$$

The expected number of generations to the next (again, going backwards in time) coalescence when there are *i* branches on the tree may be obtained from Eq. (10) as

$$T_i = \frac{2N}{i(i-1)}. \tag{11}$$

For a sample of size *n*, the average number of generations between the current generation and the generation containing the MRCA is thus

$$T_{\mathrm{MRCA}}(n) = \sum_{i=2}^{n} T_i = 2N\left(1-\frac{1}{n}\right). \tag{12}$$

Fixation occurs when one of the original ancestors becomes a common ancestor of the entire population. We use Eq. (12) to obtain an approximate expression for the expected number of generations to fixation by setting the sample size to the population size, i.e.

$$T_{\mathrm{fix}} = T_{\mathrm{MRCA}}(N) = 2N\left(1-\frac{1}{N}\right). \tag{13}$$

There are three aspects of this last step that involve approximations that warrant explanation. First, consider the generation at which all neutrons first share a common ancestor. One may then trace the lineages of the neutrons back through the generations until the MRCA is found. In general, this MRCA will be found in some generation after the initial generation. Equation (13) gives the expected number of generations between the MRCA and fixation, but does not account for the generations between the first and the one containing the MRCA.

Second, we have violated the assumption that the sample size is much smaller than the population size. However, using Eq. (11) we see that the expected number of generations required to coalesce from two ancestors to one is $T_2 = N$. Thus, for half of the expected number of generations required for fixation there are only two branches on the tree. For one-sixth of the number of generations there are only three branches, for one-twelfth the number of generations only four branches, etc. Thus, the approximation is valid for the terms in Eq. (13) that contribute the most to T_{fix}.

The final point is that fixation occurs at the first generation for which the entire population shares a common ancestor, but in the derivations leading to Eq. (13) the number of generations to the first coalescence (going backwards) is assumed to be drawn from a distribution. However, the probability that all *N* neutrons in a generation have distinct ancestors in the previous generation is P_1^N, where P_1 is computed using Eq. (1). For large values of *N*, P_1^N is an extremely small number so that the number of generations to the first coalescence is almost certainly one. Despite these approximations, as we will see Eq. (13) seems to yield a reasonably accurate result.

Figure 2 illustrates the extinction of families and fixation. The red line shows the average number of families as a function of generation for an ensemble of 200 calculations identical to the problem discussed in Sec. IV. The green line is the number of families versus generation for a representative member of the ensemble (run 1 of 200, the same run that produced the results shown in Fig. 1). From Eq. (6), the expected fraction of families to become extinct during the first generation is 0.367695. The observed fraction from the ensemble of Monte Carlo calculations is 0.366895.

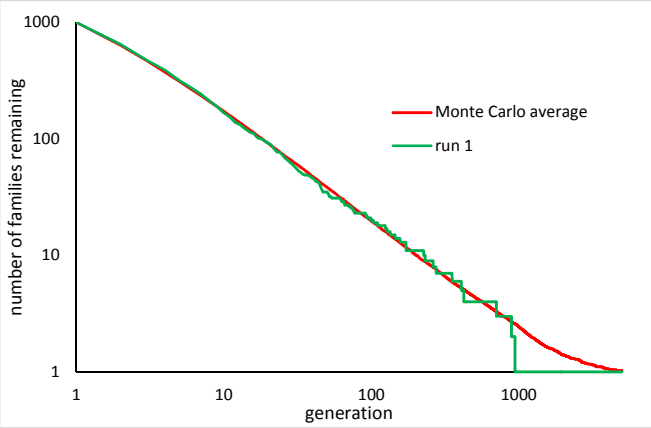

Fig. 2. Number of families as a function of generation. The red line is an average over 200 independent calculations. The green line is from a single representative calculation.

Figure 3 shows a binned representation of the probability density for T_{fix} obtained from an ensemble of 400 calculations using the same model as before. The bin width is 250 generations. The average value is 2,008.88, which compares well with the value of 1,998 obtained using Eq. (13). As can be seen from the plot, there is a wide spread in the distribution of T_{fix}. For large N, the variance is given by

$$\sigma^2_{T_{\text{fix}}} = 4N^2\left(\frac{\pi^2}{3}-3\right), \tag{14}$$

which yields a standard deviation of 1076.79, which also compares well with the value of 1091.10 obtained using the simulation data.

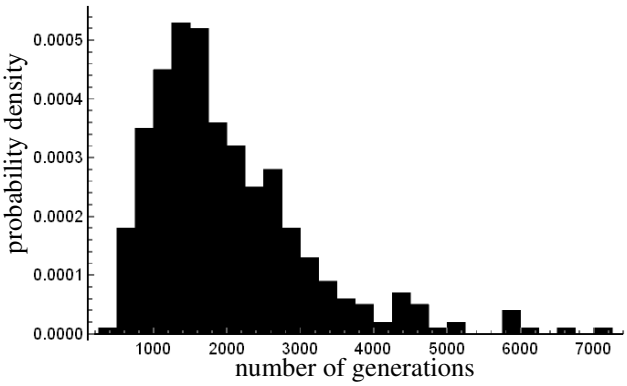

Fig. 3. Probability density of the number of generations to fixation.

VI. SPATIAL CLUSTERING IN MONTE CARLO CALCULATIONS

1. The Two-particle Distribution Function

In our examination of spatial clustering we will use the two-particle distribution function $h_2(\mathbf{x},\mathbf{y},g)$, which is the joint probability density for one neutron being absorbed at **x** and a different neutron being absorbed at **y** in generation g. For the homogeneous problem considered here, the absorption probability density per neutron is uniform and equal to L^{-3}. In the first generation all of the neutrons are uncorrelated, hence

$$h_2(\mathbf{x},\mathbf{y},1) = \frac{N(N-1)}{L^6}. \tag{15}$$

In subsequent generations the contribution to the two-particle distribution function due to uncorrelated neutrons decreases proportionally to the fraction of uncorrelated pairs, i.e.

$$h_2^{\text{unc}}(\mathbf{x},\mathbf{y},g) = \frac{N(N-1)}{L^6}\psi_{\text{u}}(g). \tag{16}$$

The contribution to the two-particle distribution function at generation g due to neutron pairs with a MRCA in a previous generation g' is given by

$$N(N-1)\psi_{\text{c}}(g'|g)\frac{\int d\mathbf{r}G(\mathbf{x},\mathbf{r},g-g')G(\mathbf{y},\mathbf{r},g-g')}{L^3}. \tag{17}$$

Here, $N(N-1)\psi_{\text{c}}(g'|g)$ is the number of correlated pairs in generation g with a MRCA in generation g'. The Green's function $G(\mathbf{x},\mathbf{r},g-g')$ is the probability per unit volume that a neutron will be absorbed at **x** in generation g

given that its ancestor in generation g' was absorbed at **r**. The integral divided by the volume in Eq. (17) is thus the joint probability that a pair of neutrons with a uniformly-distributed MRCA in generation g' will be absorbed in unit volumes about **x** and **y** in generation g. Combining Eqs. (16) and (17), the two-particle distribution function is given by

$$h_2(\mathbf{x},\mathbf{y},g) = \frac{N(N-1)}{L^6}\psi_u(g) + \frac{N(N-1)}{L^3}\sum_{g'=1}^{g-1}\psi_c(g'|g) \times \int d\mathbf{r} G(\mathbf{x},\mathbf{r},g-g')G(\mathbf{y},\mathbf{r},g-g'), \quad (18)$$

Following Meyer, *et al.* [9] and Zoia, *et al.*, [10], we use the Markov property of the Green's function, i.e.

$$\int d\mathbf{r} G(\mathbf{x},\mathbf{r},g-g')G(\mathbf{y},\mathbf{r},g-g') = G(\mathbf{x},\mathbf{y},2(g-g')), \quad (19)$$

thus

$$h_2(\mathbf{x},\mathbf{y},g) = \frac{N(N-1)}{L^3}\left[\frac{1}{L^3}\psi_u(g) + \sum_{g'=1}^{g-1}\psi_c(g'|g)G(\mathbf{x},\mathbf{y},2(g-g'))\right]. \quad (20)$$

2. The Diffusion Approximation

To make further progress we need an explicit expression for the Green's function, which we will now obtain using a diffusion approximation. In analogy to the method employed by Meyer, *et al.* [9] for the continuous-time case, the generalized trajectory for a neutron in generation g is defined as its own trajectory plus the trajectories of all of its ancestor neutrons starting from the source location of its original ancestor. Such a generalized trajectory may be envisioned as a sequence of isotropic scattering events for a pseudo-particle diffusing through a purely scattering medium. For the continuous time case, the corresponding Green's function is given by

$$\Phi(\mathbf{x},\mathbf{y},t) = (4\pi Dt)^{-3/2} e^{-\frac{|\mathbf{x}-\mathbf{y}|^2}{4Dt}}, \quad (21)$$

where D is a diffusion coefficient. The mean-squared distance traveled by a diffusing particle in time t is $\langle r^2(t)\rangle = 6Dt$, thus Eq. (21) may be rewritten as

$$\Phi(\mathbf{x},\mathbf{y},t) = \left(\frac{3}{2\pi\langle r^2(t)\rangle}\right)^{3/2} e^{-\frac{3|\mathbf{x}-\mathbf{y}|^2}{2\langle r^2(t)\rangle}}. \quad (22)$$

Defining $\langle r_1^2\rangle$ as the mean-squared distance between a neutron's birth and absorption in a single generation, the Green's function for the pseudo-particles may be written as

$$G(\mathbf{x},\mathbf{y},g) = \left(\frac{3}{2\pi\langle r_1^2\rangle g}\right)^{3/2} e^{-\frac{3|\mathbf{x}-\mathbf{y}|^2}{2\langle r_1^2\rangle g}}. \quad (23)$$

To illustrate the plausibility of this diffusion approximation, we consider the case of a point source of pseudo-particles at the origin in an infinite medium. 1,000,000 pseudo-particles were simulated for 100 generations using the cross section data from Table 1. The simulation started all pseudo-particles from the origin, and tracked them using the standard analog Monte Carlo transport algorithm with the exception that at the beginning of each generation exactly one neutron was born at each absorption location. Figure 3 shows the mean-squared distance from the origin versus generation. Using Eq. (23), the mean-squared distance from the origin is given by $\langle r^2(g)\rangle = \langle r_1^2\rangle g$. The linearity of the simulation data indicates that the diffusion approximation is valid for this case. The slope of the line yields $\langle r_1^2\rangle = 4.998\ \text{cm}^2$.

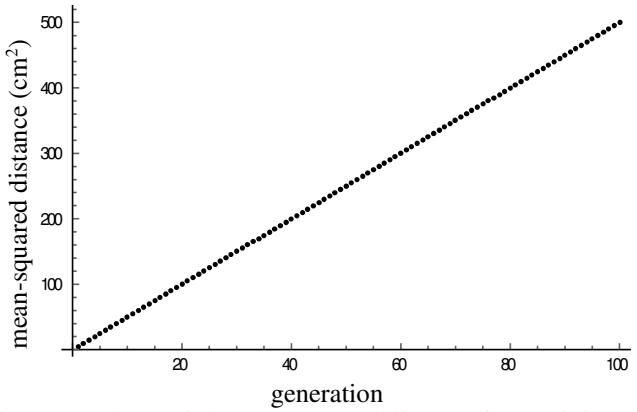

Fig. 3. Pseudo-particle mean-squared distance from origin.

Following Zoia, *et al.* [10], de Mulatier, *et al.* [2], and Nowak, *et al.* [4], an eigenfunction expansion [11] is applied to Eq. (23) yielding

$$G(\mathbf{x},\mathbf{y},g) = \left[\frac{1}{L^3} + \sum_{ijk}\varphi_i(x_1)\varphi_i^\dagger(y_1)\varphi_j(x_2)\varphi_j^\dagger(y_2) \times \varphi_k(x_3)\varphi_k^\dagger(y_3)e^{-\frac{\langle r_1^2\rangle}{6}\left(\frac{\pi}{L}\right)^2(i^2+j^2+k^2)}\right], \quad (24)$$

where

$$\varphi_k(x) = \cos\left(\frac{k\pi x}{L}\right), \tag{25}$$

$$\varphi_k^{\dagger}(x) = \frac{2}{L}\cos\left(\frac{k\pi x}{L}\right), \tag{26}$$

and x_q and y_q for $q = 1, 2, 3$ are the three Cartesian components of the **x** and **y** vectors. The summation is over all combinations *i*, *j*, and *k* except $i = j = k = 0$ since that term is given explicitly as L^{-3} . Substitution of Eq. (24) into Eq. (20) produces

$$\begin{aligned} h_2(\mathbf{x},\mathbf{y},g) = \frac{N(N-1)}{L^3}\Bigg[\frac{1}{L^3} \\ +\sum_{g'=1}^{g-1}\psi_c(g'|g)\sum_{ijk}\varphi_i(x_1)\varphi_i^{\dagger}(y_1) \\ \times\varphi_j(x_2)\varphi_j^{\dagger}(y_2)\varphi_k(x_3)\varphi_k^{\dagger}(y_3) \\ \times e^{-\frac{\langle r_1^2\rangle}{3}\left(\frac{\pi}{L}\right)^2\left(i^2+j^2+k^2\right)}\Bigg]. \end{aligned} \tag{27}$$

3. The Mean-Squared Distance Between Pairs

The expected mean-squared distance between pairs of absorption locations is defined using the two-particle distribution function as

$$\left\langle r_p^2(g)\right\rangle \equiv \frac{\int d\mathbf{x}\int d\mathbf{y} h_2(\mathbf{x},\mathbf{y},g)|\mathbf{x}-\mathbf{y}|^2}{\int d\mathbf{x}\int d\mathbf{y} h_2(\mathbf{x},\mathbf{y},g)}. \tag{28}$$

In the absence of correlation the distribution of absorption locations would be uniform, in which case $\left\langle r_{p,\,unc}^2\right\rangle = L^2/2$. Significant deviation from this value would be an indicator of clustering.

Using Eq. (27) for the pair correlation function, we obtain

$$\begin{aligned} \left\langle r_p^2(g)\right\rangle = \frac{L^2}{2}\Bigg[1 - \frac{96}{(N-1)\pi^4}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^4} \\ \times\frac{N\left(1-e^{-(2k-1)\frac{\langle r_1^2\rangle}{3}\left(\frac{\pi}{L}\right)^2(g-1)}\left(\frac{N-1}{N}\right)^g\right)-1}{1+N\left(e^{(2k-1)\frac{\langle r_1^2\rangle}{3}\left(\frac{\pi}{L}\right)^2}-1\right)}\Bigg]. \end{aligned} \tag{29}$$

As *g* becomes large, the mean-squared distance between pairs approaches an asymptotic value given by

$$\begin{aligned} &\left\langle r_p^2(\infty)\right\rangle \\ &= \frac{L^2}{2}\left[1 - \frac{96}{\pi^4}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^4}\frac{1}{1+N\left(e^{(2k-1)\frac{\langle r_1^2\rangle}{3}\left(\frac{\pi}{L}\right)^2}-1\right)}\right]. \end{aligned} \tag{30}$$

For $\langle r_1^2\rangle / L^2 \ll 1$, the above expression may be simplified by replacing the exponential by its series expansion and retaining only the first-order term. Even though this approximation is not valid for terms corresponding to large values of *k*, these terms contribute negligibly to the summation. The resulting expression is

$$\left\langle r_p^2(\infty)\right\rangle = 2N\left\langle r_1^2\right\rangle\left[1 - \sqrt{\frac{4N\langle r_1^2\rangle}{3L^2}}\tanh\left(\sqrt{\frac{3L^2}{4N\langle r_1^2\rangle}}\right)\right], \tag{31}$$

which is identical in form to equivalent expressions obtained for the case of continuous time variation [2,4].

It is instructive to examine the behavior of Eq. (31) for various limiting cases. For an infinite medium $(L \to \infty)$, the term in brackets goes to unity and we have $\left\langle r_p^2(\infty)\right\rangle \simeq 2N\left\langle r_1^2\right\rangle$. To understand this result, note that for large *g* the average number of generations separating two correlated neutrons from their MRCA is equal to the number of neutrons in a generation, i.e.

$$\lim_{g\to\infty}\sum_{g'=1}^{g-1}\psi_c(g'|g)(g-g') = N \ . \tag{32}$$

Over the course of N generations, two neutrons born at the same location in an infinite medium will thus have an expected squared separation of $2N\left\langle r_1^2\right\rangle$. For small but finite values of $N\left\langle r_1^2\right\rangle/L^2$ the bracketed term in Eq. (30) is positive and less than unity, so that $\left\langle r_p^2(\infty)\right\rangle$ is smaller than for an infinite medium.

For the case of $N\to\infty$, we obtain $\left\langle r_p^2(\infty)\right\rangle\to\left\langle r_{p,\,unc}^2\right\rangle$, i.e. the result for an uncorrelated system of particles. For sufficiently large but finite N, Eq. (31) may be approximated as

$$\left\langle r_p^2(\infty)\right\rangle\simeq\left\langle r_{p,\,unc}^2\right\rangle\left[1-\frac{3L^2}{10N\left\langle r_1^2\right\rangle}\right], \tag{33}$$

from which we see that correlation reduces $\left\langle r_p^2(\infty)\right\rangle$ from the value for an uncorrelated system. This behavior has previously been noted for the continuous-time case [4].

Figure 4 shows the behavior of the mean-squared distance between pairs for our model problem as the calculation proceeds through the generations. The black line is the theoretical prediction determined from Eq. (29) using $\left\langle r_1^2\right\rangle=4.998\ \text{cm}^2$ and truncating the summation at $k=20$. Initially the mean-squared distance between pairs is the value corresponding to a uniform distribution, i.e. $L^2/2=80{,}000\ \text{cm}^2$. As the calculation proceeds, however, the effect of clustering is evident as the expected value decreases to the asymptotic value of $\left\langle r_p^2(\infty)\right\rangle=7957\ \text{cm}^2$ determined using Eq. (30). This is 20% less than the expected value of the mean-squared distance between pairs for an infinite medium ($2N\left\langle r_1^2\right\rangle=9{,}994\ \text{cm}^2$). Also shown is a plot of $\left\langle r_p^2(g)\right\rangle$ obtained by averaging over an ensemble of 200 independent Monte Carlo calculations (red line), and a plot of a single representative member of the ensemble (run 1 of 200, green line). Despite the use of the diffusion approximation in the theoretical model, agreement with the average of the Monte Carlo results is quite good. Note the precipitous drop in the plot for run 1 at about the point where fixation occurs (generation 942). From this point on all the neutrons are in a single tight cluster.

Figure 5 shows the cluster from Fig. 1 at generation 10,000 along with the location of the original ancestor of all of the neutrons in the cluster (red point) and the COM of the population (magenta point). The COM is given by

$$\mathbf{r}_{\text{COM}}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{r}_i\ , \tag{34}$$

where $\mathbf{r}_i$ is the location of the i^{th} neutron absorption location. Since the mean-squared distance between points uniformly distributed within a sphere of radius R is $6R^2/5$. we may use $\rho\equiv\sqrt{\left\langle r_p^2(\infty)\right\rangle}$ as a characteristic length associated with the radial extent of the cluster. Figure 5 shows a superimposed sphere of radius ρ centered on the COM. By taking the ratio of the volume of the sphere to that of the 400 cm cube in which it is located, we see that after the cluster has formed all of the neutrons are found in less than 5% of the volume of the problem.

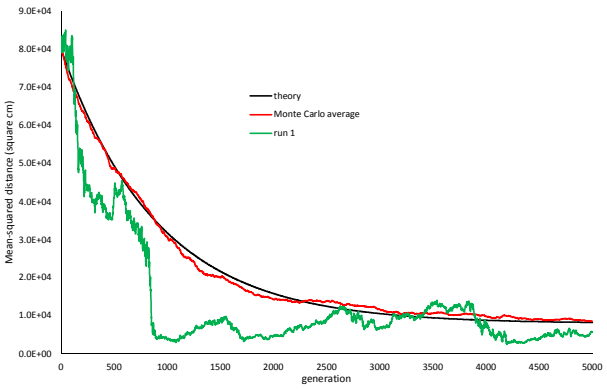

Fig. 4. Mean-squared distance between pairs for N = 1000.

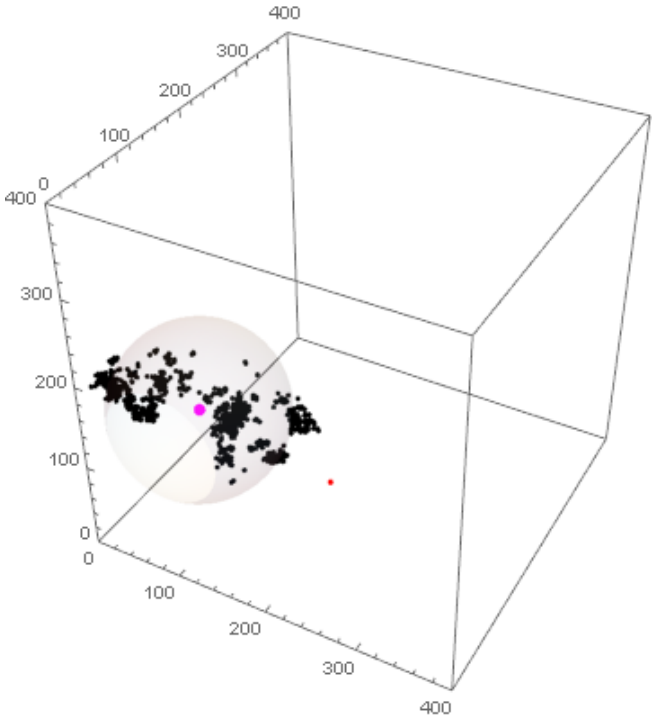

Fig. 5. Absorption distribution after 10,000 generations (black), a sphere of radius ρ centered at the COM (magenta), and the location of the original ancestor (red).

VII. CONCLUSION

In this paper we have developed a mathematical model of clustering in Monte Carlo iterated-fission-source calculations. This model explicitly accounts for the sequential-generation nature of the Monte Carlo source iteration procedure, whereas earlier works in this area used a continuous-time model. A novel aspect of the current work is the analysis of the decline in the number of 'lineages' remaining as the calculation proceeds through the generations.

It must be stressed that the Monte Carlo algorithm used and the problem studied are not typical of real-world applications. This was intentional. The simplicity of the algorithm and problem allowed for theoretical analyses that would not have been possible otherwise. Furthermore, the particular model parameters used were chosen to emphasize the effect of clustering. While clustering may indeed be an issue for real-world problems, only for very pathological situations would it be likely to manifest itself so dramatically as in this paper.

From the theoretical results presented in Sec. VI, it is clear that the error in the mean-squared-distance between pairs due to clustering decreases with increasing number of neutrons per generation. To demonstrate this, the model problem was rerun using 250 generations of 40,000 neutrons each. The total number of neutrons run is thus the same as for the calculation discussed in Sec. IV. The mean-squared distance between pairs versus generation is plotted in Fig. 6. Instead of decreasing as was the case with the calculation illustrated in Fig. 4, the value fluctuated about the correct value of 80,000. Although it has not been proven here, one might expect that Monte Carlo results in general will suffer less from the effects of clustering for larger generations. Thus, if one is considering whether to run calculations with many smaller generations or fewer larger generations, one should choose the latter.

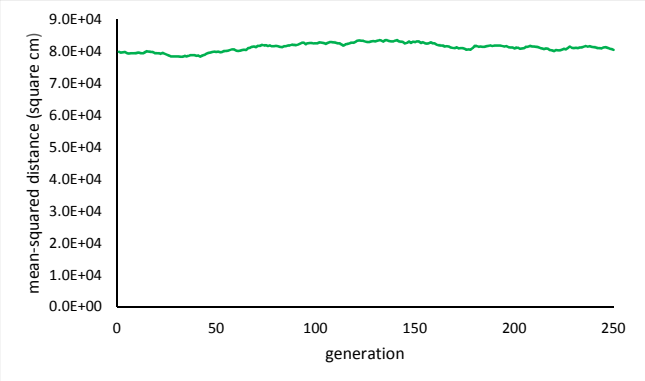

Fig. 6. Mean-squared distance between pairs for N = 40,000.

ACKNOWLEDGMENT

The authors would like to thank Andrea Zoia of CEA for several helpful discussions.

REFERENCES

1. I. LUX and L. KOBLINGER, *Monte Carlo Particle Transport Methods: Neutron and Photon Calculations*, CRC Press, Boca Raton, Florida (1991).
2. C. DE MULATIER, *et al*., "The Critical Catastrophe Revisited," *J. Stat. Mech*., P08021, doi:10.1088/1742-5468/2015/08/P08021, (2015).
3. E. DOMONTEIL, *et al*., "Particle Clustering in Monte Carlo Criticality Simulations," *Ann. Nucl. Energy*, **63**, 612, http://dx.doi.org/10.1016/j.anucene.2013.09.008, (2014).
4. M. NOWAK, *et al*., "Monte Carlo Power Iteration: Entropy and Spatial Correlations," *Ann. Nucl. Energy*, **94**, 856, http://dx.doi.org/10.1016/j.anucene.2016.05.002, (2016).
5. S. WRIGHT, "Evolution in Mendelian Populations," *Genetics*, **16**, 97, (1931).
6. R. DURRETT, *Probability Models for DNA Sequence Evolution*, Springer, New York (2008).
7. J. F. C. KINGMAN, "The Coalescent," *Stochast. Proc. Appl.*, **13**, 235, (1982).
8. R. R. HUDSON, "Gene Genealogies and the Coalescent Process," *Oxford Surveys in Evolutionary Biology*, **7**, 1, (1990).
9. M. MEYER, *et al*., "Clustering of Independently Diffusing Individuals by Birth and Death Processes," *Phys. Rev. E*, **54**, 5567, (1996).
10. A. ZOIA, *et al*., "Clustering of Branching Brownian Motions in Confined Geometries," *Phys. Rev. E*, **90**, 042118, (2014).
11. D. S. GREBENKOV and B.-T. NGUYEN, "Geometrical Structure of Laplacian Eigenfunctions," *SIAM Rev.*, **55**, 601, (2013).