Nuclear Mathematical and Computational Sciences: A Century in Review, A Century Anew
Gatlinburg, Tennessee, April 6-11, 2003, on CD-ROM, American Nuclear Society, LaGrange Park, IL (2003)

VARIATIONAL VARIANCE REDUCTION FOR CRITICALITY CALCULATIONS USING MONTE CARLO ADJOINT FLUXES

Jeffery D. Densmore
Transport Methods Group CCS-4
Los Alamos National Laboratory
Los Alamos, New Mexico USA 87545
jdensmor@engin.umich.edu

Edward W. Larsen
Department of Nuclear Engineering and Radiological Sciences
University of Michigan
Ann Arbor, Michigan USA 48109
edlarsen@engin.umich.edu

ABSTRACT

The Variational Variance Reduction (VVR) method has been shown to be an effective technique for increasing the efficiency of Monte Carlo simulations. This method uses a variational functional, which employs first-order estimates of the forward and adjoint fluxes, to yield a second-order estimate of a desired system characteristic – which, in this paper, we take to be the criticality eigenvalue k. If Monte Carlo estimates of the forward *and* adjoint fluxes are used, each having global "first-order" errors of $O(1/\sqrt{N})$, where N is the number of histories used in the Monte Carlo simulation, then the statistical error in the VVR estimation of k will in principle be $O(1/N)$. In this paper, we develop this theoretical possibility and demonstrate with numerical examples that implementations of the VVR method for criticality problems can approximate $O(1/N)$ convergence for sufficiently large values of N.

Key Words: variance reduction, Monte Carlo, criticality, variational methods, adjoint simulation

1. INTRODUCTION

The use of variational functionals is an effective method of increasing the efficiency of Monte Carlo simulations [1–12]. In this technique a system characteristic, such as a detector response or an eigenvalue, is estimated using a variational rather than a direct functional. Direct functionals, which are employed in traditional Monte Carlo calculations, require estimates of the forward flux in the detector region only. Variational functionals, which are theoretically more accurate than direct functionals, require estimates of forward and adjoint fluxes over the entire phase space of the problem. Recently, variational functionals have been employed as a variance reduction device by (i) calculating the adjoint flux estimate using a relatively inexpensive deterministic method, (ii) representing the adjoint flux as a function in phase space, (iii) estimating the forward solution using Monte Carlo, and (iv) evaluating the variational functional using the deterministic adjoint and forward Monte Carlo flux information [3–11]. We have called this general procedure *Variational Variance Reduction* (VVR). In this paper, we examine the use of *Monte Carlo-generated adjoint flux estimates* in the VVR method, applied to 3-D, monoenergetic criticality problems [7, 12].

Variational functionals are more accurate than direct functionals because they combine *first-order* estimates of the forward and adjoint fluxes to yield a *second-order* estimate of the desired system characteristic. Direct functionals, used in standard Monte Carlo simulations, yield only a first-order estimate of the system characteristic from a first-order estimate of the forward flux. If the error in the Monte Carlo-calculated forward flux is $O(1/\sqrt{N})$, where N is the number histories in the Monte Carlo simulation, then the error in standard Monte Carlo estimates of the system characteristic, obtained using a direct functional, is also first-order, i.e. $O(1/\sqrt{N})$. This is a well-known, almost universal result: the statistical error in a Monte Carlo simulation is inversely proportional to the square root of the computing effort. However, if a variational functional is used with Monte Carlo estimates for both the forward and adjoint fluxes, each having errors of $O(1/\sqrt{N})$, then – in the absence of additional errors – the error in the variational estimate of the system characteristic will be second order, i.e. $O(1/N)$. Unfortunately, there are additional (truncation) errors, discussed below, which degrade this performance for large N to $O(1/\sqrt{N})$. Nevertheless, the variational functional is significantly more accurate, i.e. has significantly smaller statistical errors, than the standard functional.

The VVR method is more costly to implement per particle than traditional Monte Carlo, due to the extra work required to estimate the adjoint flux and evaluate the variational functional. However, our numerical results show that the accuracy of the variational functional outweighs this extra cost, yielding an overall more efficient Monte Carlo simulation.

Allagi and Lewins [1], and Allagi, Lewins, and Parks [2] have also examined generating both the forward and adjoint flux estimates via Monte Carlo. Although their results demonstrate the increase in efficiency when using variational functionals, they made no attempt to examine the increase in the convergence rate over other Monte Carlo methods. Also, their adjoint flux was estimated directly from the forward Monte Carlo simulation. This can lead to a bias in the results of the variational estimate [1]. In the present paper, we employ distinct Monte Carlo simulations to estimate the forward and adjoint fluxes.

We now outline the remainder of this paper. In Section 2 we outline the 3-D monoenergetic criticality problems of interest. In Section 3 we describe the variational functional employed in the VVR method, and the direct functional used in standard Monte Carlo calculations. In Section 4 we discuss the application of the variational functional in the VVR method, including the representation of the adjoint flux and the evaluation of the variational functional using forward and adjoint Monte Carlo simulations. In Section 5 we describe the behavior of the variational functional when Monte Carlo is used to estimate both the forward and adjoint fluxes. Here, we demonstrate theoretically that the error in the eigenvalue estimate converges as $O(1/N)$. In Section 6 we numerically compare (i) the VVR method using a Monte Carlo adjoint simulation, (ii) the VVR method using a deterministic adjoint calculation, and (iii) standard Monte Carlo. This comparison consists of examining the convergence rates of the above three methods for a class of criticality problems. We conclude with a brief discussion in Section 7.

2. PROBLEM DESCRIPTION

We consider the following three-dimensional, monoenergetic criticality problem:

$$L\psi = \frac{1}{k} F\psi \quad , \quad \mathbf{r} \in V \quad , \quad \mathbf{\Omega} \in 4\pi \quad , \tag{1}$$

where an estimate of the the eigenvalue k is desired. Here, $\psi(\mathbf{r}, \mathbf{\Omega})$ is the eigenfunction or forward angular flux, $\mathbf{r}$ is the spatial variable, and $\mathbf{\Omega}$ is the angular variable. The spatial domain is represented by V. In Eq.

(1), the transport operator L is defined as

$$L\psi = \mathbf{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}) + \Sigma_t(\mathbf{r})\psi(\mathbf{r}, \mathbf{\Omega}) - \frac{\Sigma_s(\mathbf{r})}{4\pi} \int_{4\pi} \psi(\mathbf{r}, \mathbf{\Omega}')d\Omega' \quad , \tag{2}$$

and the fission operator F is defined by

$$F\psi = \frac{\nu\Sigma_f(\mathbf{r})}{4\pi} \int_{4\pi} \psi(\mathbf{r}, \mathbf{\Omega}')d\Omega' \quad . \tag{3}$$

In Eq. (2), we have assumed isotropic scattering, but our work easily generalizes to include anisotropic scattering. To complete the problem description, boundary conditions must assign the angular flux for incoming directions on the outer boundary of the system. For criticality problems, we consider only vacuum or reflective boundary conditions.

Corresponding to the above forward transport problem is the adjoint problem

$$L^*\psi^* = \frac{1}{k}F\psi^* \quad , \quad \mathbf{r} \in V \quad , \quad \mathbf{\Omega} \in 4\pi \quad , \tag{4}$$

where $\psi^*(\mathbf{r}, \mathbf{\Omega})$ is the adjoint eigenfunction or angular flux (or, the *importance function* [14]), and the eigenvalue k is identical to the eigenvalue of Eq. (1). In Eq. (4), the adjoint transport operator L^* is defined as

$$L^*\psi^* = -\mathbf{\Omega} \cdot \nabla\psi^*(\mathbf{r}, \mathbf{\Omega}) + \Sigma_t(\mathbf{r})\psi^*(\mathbf{r}, \mathbf{\Omega}) - \frac{\Sigma_s(\mathbf{r})}{4\pi} \int_{4\pi} \psi^*(\mathbf{r}, \mathbf{\Omega}')d\Omega' \quad , \tag{5}$$

and, for the one-group problems considered in this paper, the fission operator F is self adjoint. Vacuum or reflective boundary conditions apply to Eq. (4) for outgoing directions.

The forward and adjoint operators described above have the following properties [15]:

$$\langle f, Lh \rangle = \langle L^*f, h \rangle \quad , \tag{6}$$

and

$$\langle f, Fh \rangle = \langle Ff, h \rangle \quad . \tag{7}$$

Here f and h are arbitrary functions, and we have used the inner-product notation

$$\langle f, h \rangle = \int_V \int_{4\pi} f(\mathbf{r}, \mathbf{\Omega})h(\mathbf{r}, \mathbf{\Omega})d\Omega d^3r \quad , \tag{8}$$

to denote integration over the entire phase space of the problem. The application of Eq. (6) requires that h and f respectively satisfy the boundary conditions of Eqs. (1) and (4) exactly. In this paper, we consider forward fluxes calculated by Monte Carlo only, which exactly satisfy vacuum or reflective boundary conditions.

3. DIRECT AND VARIATIONAL FUNCTIONALS

We now present two functionals that, given estimates of the forward flux Ψ and possibly the adjoint flux Ψ^*, estimate the eigenvalue k. The first is the *direct functional*, employed by standard Monte Carlo simulations:

$$G[\Psi] = \frac{\langle 1, F\Psi \rangle}{\langle 1, L\Psi \rangle} \quad . \tag{9}$$

This functional can be obtained simply by integrating Eq. (1) over phase space. Thus, if G is evaluated using the exact forward flux, then Eq. (9) will yield the exact eigenvalue. G has the standard interpretation of being the ratio of the total neutron production rate to the total neutron loss rate [15]. Also, if $\Psi = \psi + \delta\psi$ is a "first-order" estimate of the exact forward flux, then G yields a first-order estimate of the eigenvalue,

$$G[\psi + \delta\psi] = k + O(\delta\psi) \quad . \tag{10}$$

When employing the direct functional G in a Monte Carlo simulation, the numerator $\langle 1, F\Psi \rangle$ is often evaluated by using a track-length or collision estimator to calculate the total fission rate in a given fission generation. The denominator $\langle 1, L\Psi \rangle$ is interpreted as the total loss rate. Since fission is treated as a terminal process and the loss rate is simply the total number of particles in a given generation, division of the fission rate by the loss rate is simply a normalization by the total number of histories simulated.

Corresponding to the above direct functional is the *variational functional*

$$H[\Psi, \Psi^*] = \frac{\langle \Psi^*, F\Psi \rangle}{\langle \Psi^*, L\Psi \rangle} \quad , \tag{11}$$

which requires estimates of the forward and adjoint fluxes throughout the entire phase space of the problem. H can be derived by taking the inner product [defined by Eq. (8)] of Eq. (1) with Ψ^*. This functional has the interpretation of being the importance-weighted fission rate divided by the importance-weighted loss rate.

The variational functional H has several useful properties. First, we note that

$$H[\psi, \Psi^*] = k \quad , \tag{12}$$

and

$$H[\Psi, \psi^*] = k \quad . \tag{13}$$

Eq. (12) implies that the variational functional yields the exact eigenvalue, regardless of the adjoint estimate, if the exact forward flux is used to evaluate the variational functional. Hence, in the limit of an infinite number of histories, the VVR method will calculate k exactly. Eq. (13) implies that if the exact adjoint flux is used, the variational functional yields the exact eigenvalue regardless of the forward flux estimate (as long as Ψ satisfies the exact boundary conditions). Now the VVR method becomes a form of a zero-variance method in which the Monte Carlo simulation is performed in an analog manner, but the resulting forward information is combined with the adjoint solution in a nonanalog manner [13]. This zero-variance VVR method differs from other zero-variance methods in the fact that it requires no nonanalog biasing.

We now demonstrate that when first-order estimates of the forward and adjoint fluxes are introduced into H, one obtains a second-order estimate of the eigenvalue. To do this, we assume $\Psi = \psi + \delta\psi$ and $\Psi^* = \psi^* + \delta\psi^*$ and write Eq. (11) as

$$\begin{aligned} H[\psi + \delta\psi, \psi^* + \delta\psi^*] &= \frac{\langle \psi^* + \delta\psi^*, F(\psi + \delta\psi) \rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi) \rangle} \\ &= k + \frac{\langle \psi^* + \delta\psi^*, F(\psi + \delta\psi) \rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi) \rangle} - \frac{\langle \psi^*, F\psi \rangle}{\langle \psi^*, L\psi \rangle} \\ &= k + \frac{\langle \psi^* + \delta\psi^*, F(\psi + \delta\psi) \rangle \langle \psi^*, L\psi \rangle - \langle \psi^*, F\psi \rangle \langle \psi^* + \delta\psi^*, L(\psi + \delta\psi) \rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi) \rangle \langle \psi^*, L\psi \rangle} \quad . \end{aligned} \tag{14}$$

The zero-th order terms on the right side of of Eq. (14) (containing no error terms in the numerator) can be written as

$$k + \frac{\langle \psi^*, F\psi\rangle\langle \psi^*, L\psi\rangle - \langle \psi^*, F\psi\rangle\langle \psi^*, L\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle\langle \psi^*, L\psi\rangle} = k \quad . \tag{15}$$

The second-order term of Eq. (14) (containing only the product $\delta\psi\delta\psi^*$ in the numerator) is given by

$$\frac{\langle \delta\psi^*, F\delta\psi\rangle\langle \psi^*, L\psi\rangle - \langle \psi^*, F\psi\rangle\langle \delta\psi^*, L\delta\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle\langle \psi^*, L\psi\rangle} = O(\delta\psi\delta\psi^*) \quad , \tag{16}$$

which cannot be simplified. Finally, the first-order term of Eq. (14) (containing only the error terms $\delta\psi$ and $\delta\psi^*$ separately in the numerator) is

$$\frac{[\langle \psi^*, F\delta\psi\rangle + \langle \delta\psi^*, F\psi\rangle]\langle \psi^*, L\psi\rangle - \langle \psi^*, F\psi\rangle[\langle \delta\psi^*, L\psi\rangle + \langle \psi^*, L\delta\psi\rangle]}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle\langle \psi^*, L\psi\rangle} \quad , \tag{17}$$

which can be simplified using Eq. (12) as

$$\frac{\langle \psi^*, F\delta\psi\rangle + \langle \delta\psi^*, F\psi\rangle - k[\langle \delta\psi^*, L\psi\rangle + \langle \psi^*, L\delta\psi\rangle]}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} \quad . \tag{18}$$

Using the definition of the exact forward flux, Eq. (1), the second and third terms in the numerator of Eq. (18) cancel, leaving

$$\begin{aligned} \frac{\langle \psi^*, F\delta\psi\rangle - k\langle \psi^*, L\delta\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} &= \frac{\langle \psi^*, F\delta\psi\rangle - k\langle \psi^*, L\delta\psi\rangle + \langle \psi^*, F\psi\rangle - k\langle \psi^*, L\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} \\ &= \frac{\langle \psi^*, F\Psi\rangle - k\langle \psi^*, L\Psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} = \frac{\langle F^*\psi^*, \Psi\rangle - k\langle L^*\psi^*, \Psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} = 0 \quad , \end{aligned} \tag{19}$$

where we have used Eqs. (4), (6), (7), and (12), and assumed the forward flux estimate satisfies the boundary conditions. Then, using Eqs. (15), (16), and (19), Eq. (14) yields the result

$$H[\psi + \delta\psi, \psi^* + \delta\psi^*] = k + O(\delta\psi\delta\psi^*) \quad , \tag{20}$$

which states that first-order estimates in the forward and adjoint fluxes yield a second-order estimate in the eigenvalue k. Thus, the variational functional H has the capacity to estimate k more accurately than the direct functional G.

4. IMPLEMENTATION

We now describe the implementation of the variational functional into a variance reduction method for criticality problems. To begin, we subdivide the problem domain V into a "tally" grid of I cells for the purpose of collecting the forward Monte Carlo information. This tally grid is represented by cells V_i, each with a boundary ∂V_i and a unit outward normal vector $\mathbf{n}_i$. No constraints are placed on this tally grid, other than that the material properties (cross-sections) are constant within each cell.

Next, we generate an estimate of the adjoint flux Ψ^* and use the resulting information to calculate cell-averaged spherical harmonic moments in each tally cell $1 \le i \le I$. These spherical harmonic moments are defined by

$$\phi^*_{nmi} = \frac{1}{\Delta V_i}\int_{V_i}\int_{4\pi} Y^*_{nm}(\mathbf{\Omega})\Psi^*(\mathbf{r}, \mathbf{\Omega})d\Omega d^3r \quad , \tag{21}$$

for $0 \leq n \leq N$ and $-n \leq m \leq n$. In Eq. (21), ΔV_i is the volume of tally cell i, and N is the order of angular representation desired. In our simulations, we estimate ϕ^*_{nmi} using a track length estimator. Then, using these spherical harmonic moments, we represent Ψ^* as the following histogram-in-space function in each tally cell i:

$$\Psi^*(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} \phi^*_{nmi} Y_{nm}(\mathbf{\Omega}) \quad , \quad \mathbf{r} \in V_i \quad , \quad \mathbf{\Omega} \in 4\pi \quad . \tag{22}$$

This representation of the adjoint flux preserves the first N cell-average spherical harmonic moments. (In this paper, we use $N \leq 3$.) The resulting low-order angular representation is accurate in diffusive systems, where the angular flux depends weakly on the angular variable. Thus, the resulting VVR method should perform very well for problems involving large nuclear reactor cores dominated by scattering [7–12]. More elaborate representations of the adjoint flux are possible, including representations that involve polynomial expansions of the angular *and* spatial variables [7, 9–11]. For simplicity, we consider only a piecewise-constant spatial representation of the adjoint flux estimate in this paper.

The above representation of the adjoint flux estimate can be introduced into H to obtain a *reduced functional*, which requires knowledge of low-order angular moments of the forward flux in each tally cell and on each tally cell boundary. To do this, we first write H as

$$H[\Psi, \Psi^*] = \frac{\langle \Psi^*, F\Psi \rangle}{\langle \Psi^*, T\Psi \rangle + \langle \Psi^*, C\Psi \rangle - \langle \Psi^*, S\Psi \rangle} \quad , \tag{23}$$

where the streaming operator T is defined by

$$T\psi = \mathbf{\Omega} \cdot \nabla \psi(\mathbf{r}, \mathbf{\Omega}) \quad , \tag{24}$$

the collision operator C is defined by

$$C\psi = \Sigma_t(\mathbf{r}) \psi(\mathbf{r}, \mathbf{\Omega}) \quad , \tag{25}$$

and the scattering operator S is defined by

$$S\psi = \frac{\Sigma_s(\mathbf{r})}{4\pi} \int_{4\pi} \psi(\mathbf{r}, \mathbf{\Omega}') d\Omega' \quad . \tag{26}$$

Introducing Eqs. (22) into (23), we may express the various terms in the reduced functional as:

$$\langle \Psi^*, F\Psi \rangle = \sum_{i=1}^{I} \nu \Sigma_{fi} \phi^*_{00i} \phi_{00i} \quad ; \tag{27}$$

$$\langle \Psi^*, S\Psi \rangle = \sum_{i=1}^{I} \Sigma_{si} \phi^*_{00i} \phi_{00i} \quad ; \tag{28}$$

$$\langle \Psi^*, C\Psi \rangle = \sum_{i=1}^{I} \sum_{n=0}^{N} \sum_{m=-n}^{n} \Sigma_{ti} \phi^*_{nmi} \phi_{nmi} \quad ; \tag{29}$$

$$\langle \Psi^*, T\Psi \rangle = \sum_{i=1}^{I} \sum_{n=0}^{N} \sum_{m=-n}^{n} \phi^*_{nmi} \Gamma_{nmi} \quad . \tag{30}$$

In these equations, ϕ_{nmi} and Γ_{nmi} are defined by:

$$\phi_{nmi} = \int_{V_i} \int_{4\pi} Y_{nm}(\mathbf{\Omega}) \Psi(\mathbf{r}, \mathbf{\Omega}) d\Omega d^3 r \quad , \tag{31}$$

and

$$\Gamma_{nmi} = \int_{\partial V_i} \int_{4\pi} (\mathbf{\Omega} \cdot \mathbf{n}_i) Y_{nm}(\mathbf{\Omega}) \Psi(\mathbf{r}, \mathbf{\Omega}) d\Omega d^2 r \quad , \tag{32}$$

for $0 \leq n \leq N$ and $-n \leq m \leq n$. These space-angle moments of Ψ can easily be estimated by standard Monte Carlo techniques. In this paper, we estimate ϕ_{nmi} by a track length estimator and Γ_{nmi} by a surface crossing estimator. We emphasize that in the construction of the reduced functional, we have approximated the adjoint flux by Eq. (22), but we have introduced no such approximation to the forward flux. Hence, if the exact forward flux is introduced into the reduced functional, the exact eigenvalue k is obtained.

Using Eqs. (22) and (27)-(30), Eq. (23) can be completely evaluated with both adjoint and forward Monte Carlo calculations. In this implementation of the VVR method, the adjoint information is provided by Eq. (21), while the forward information is provided by Eqs. (31) and (32). We again emphasize that the resulting expression is an unbiased estimator for k.

When using the above VVR method in performing a criticality calculation, generations of neutrons are followed as in traditional Monte Carlo. Neutrons in a given generation are tracked, and the resulting information is used to provide an estimate of k and source locations for the next generation [15]. However, a "full generation" for this VVR method consists of both a generation of forward particles and a generation of adjoint particles. At the end of each such "full" generation, we introduce estimates of ϕ_{nmi}, Γ_{nmi}, and ϕ^*_{nmi} into Eqs. (27)-(30) to evaluate H [Eq. (23)]. Thus, a variationally-calculated estimate of k is provided by each generation. These generation-wise estimates are then used to calculate a mean and a sample variance, as in traditional Monte Carlo.

If a deterministic method is used to estimate the adjoint flux, then the constants ϕ^*_{nmi} are determined from this calculation prior to the beginning of the Monte Carlo simulation. The subsequent Monte Carlo simulation proceeds as described in the above paragraph, with the difference that only forward Monte calculations are performed.

5. ERROR ANALYSIS OF VARIATIONAL ESTIMATES

Now we discuss the performance of the VVR method described above. From Eq. (20), for each individual fission generation, the error in the variationally-estimated eigenvalue satisfies

$$\text{Error} = O(\delta\psi \delta\psi^*) \quad . \tag{33}$$

If the error in the Monte Carlo-estimated forward flux is $O(1/\sqrt{N})$, where N is the number of histories simulated to estimate the forward flux (i.e. the number of forward neutrons simulated per fission generation), Eq. (33) gives

$$\text{Error} = O(\frac{\delta\psi^*}{\sqrt{N}}) \quad , \tag{34}$$

where $\delta\psi^*$ is the error in the adjoint flux estimate. For a fixed estimate of the adjoint flux, Eq. (34) is characteristic of most Monte Carlo simulations, the error being inversely proportional to the square root of the computational effort. However, the standard deviation is now proportional to $\delta\psi^*$, the error in the

adjoint flux estimate. If $\delta\psi^* = 0$, the variance is zero, and one has a zero-variance method. (As with all other known zero-variance methods, to achieve this one must know the adjoint flux exactly.) If $\delta\psi^*$ is small, the variance in the estimation of k is small. In previous implementations of the VVR method [3–11], we have estimated the adjoint flux by a deterministic calculation, in which case the adjoint flux estimate is fixed and $\delta\psi^*$ is "small," the degree of smallness depending on the accuracy of the technique used to perform the adjoint calculation. As predicted by Eq. (34), we observed convergence of the estimate of k with $O(1/\sqrt{N})$ error, but with a much smaller variance than observed with the standard Monte Carlo method for estimating k.

However, if the adjoint estimate is *not fixed*, but instead the forward *and* the adjoint flux estimates are obtained using Monte Carlo – each having error $O(1/\sqrt{N})$ (N is now the number of forward and adjoint histories per fission generation; the total number of histories simulated is $2N$) – the error in Eq. (33) becomes:

$$\text{Error} = O(\frac{1}{N}) \quad . \tag{35}$$

Thus, in the presence of only statistical errors, i.e. $\delta\psi = O(1/\sqrt{N})$ and $\delta\psi^* = O(1/\sqrt{N})$, the error in the estimate of k at the end of each fission generation decreases more rapidly with respect to computational effort when Monte Carlo is employed to generate both the forward and adjoint flux estimates.
Unfortunately, the increased convergence rate predicted by Eq. (35) will not continue indefinitely. Due to the bilinear form of the variational functional, it is necessary to represent the adjoint flux estimate as a function in phase space [Eq. (22)], and this representation inevitably contains a truncation error. (The estimate of the forward flux is treated continuously in angle and space.) The truncation error present in the representation of the adjoint flux will dominate the error in the Monte Carlo adjoint simulation for N large, in which case the error will converge as $O(1/\sqrt{N})$. However, for N not large, a convergence rate faster than $O(1/\sqrt{N})$ will occur.

To illustrate the effect of the truncation error in Eq. (22), we write the error in the forward flux estimate as

$$\delta\psi = \frac{A}{\sqrt{N}} \quad , \tag{36}$$

and the error in the adjoint flux estimate as

$$\delta\psi^* = \frac{A^*}{\sqrt{N}} + B^* \quad . \tag{37}$$

Here, A and A^* are constants related to the variance in the forward and adjoint Monte Carlo simulations, respectively, and B^* is a measure of the truncation error present in the adjoint flux representation, Eq. (22). By Eq. (33), the error in the VVR estimate of k satisfies

$$\text{Error} \cong \left(\frac{A}{\sqrt{N}}\right)\left(\frac{A^*}{\sqrt{N}} + B^*\right) \quad . \tag{38}$$

If the truncation error is much smaller than the statistical error (if $A^*/\sqrt{N} \gg B^*$), then the error is given by:

$$\text{Error} \approx \frac{AA^*}{N} \quad , \quad N \ll \left(\frac{A^*}{B^*}\right)^2 \quad . \tag{39}$$

If the truncation error dominates the statistical error, (if $A^*/\sqrt{N} \ll B^*$), then the error is $O(1/\sqrt{N})$:

$$\text{Error} \approx \frac{AB^*}{\sqrt{N}} \quad , \quad \left(\frac{A^*}{B^*}\right)^2 \ll N \quad . \tag{40}$$

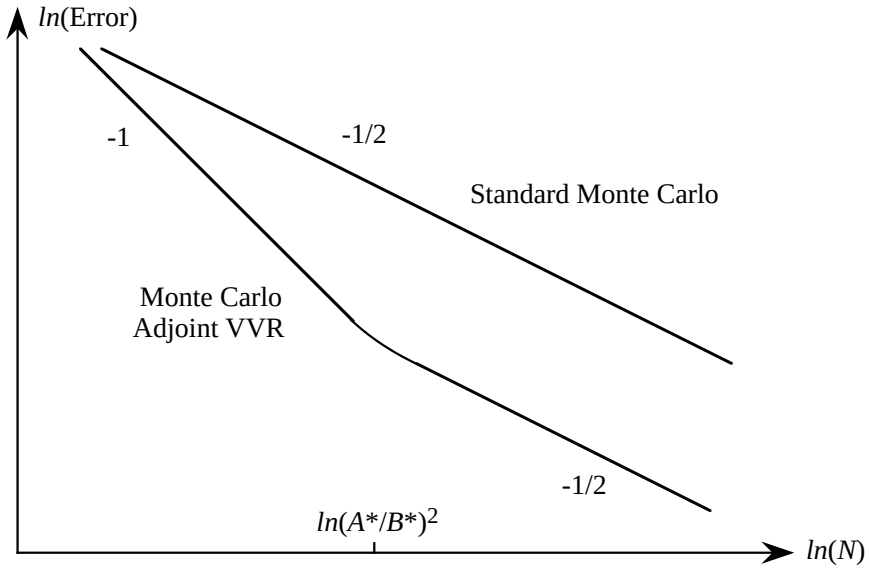

Figure 1. Convergence of Monte Carlo Adjoint VVR

The transition point between these two behaviors of the error in the VVR-estimated eigenvalue,

$$N \cong \left(\frac{A^*}{B^*}\right)^2 \quad , \tag{41}$$

is large if the truncation error B^* is small. Thus, if the spatial and angular representation of the adjoint flux estimate is sufficiently accurate, the VVR method employing both forward and adjoint Monte Carlo simulations will converge as $O(1/N)$ for a significant number of histories N.

In Figure 1 we sketch (i) the behavior of standard Monte Carlo k-eigenvalue calculations, and (ii) the predictions of the theory outlined above for the Monte Carlo adjoint VVR method. The slopes of the lines are indicated explicitly in the figure. Our actual numerical simulations, discussed in the next section, bear a closer resemblance to the left half of Figure 1 than the right half. However, there is little doubt that for large N, the experimentally-observed VVR error must eventually decline as $N^{-1/2}$.

We now relate the error predictions described above to the variance observed in a Monte Carlo simulation. We consider a criticality calculation that employs M fission generations and N histories (forward or adjoint) per fission generation. Each generation produces an estimate of the exact eigenvalue k of the form

$$k_m = k + e_m \quad , \tag{42}$$

where k_m is the eigenvalue estimate of the m^{th} fission generation and e_m represents the statistical error in the eigenvalue estimate. The expected value of the generation-wise eigenvalue estimates is given by

$$\overline{k} = E[k_m] = k + E[e_m] \quad , \tag{43}$$

which is the exact eigenvalue if the expected value of the statistical error is zero (i.e. the simulation is unbiased). The variance in the generation-wise eigenvalue estimates is

$$\sigma^2 = E[(k - k_m)^2] = E[e_m^2] \quad . \tag{44}$$

Table I. Monoenergetic Cross-Sections

Material	Σ_t	Σ_s	$\nu\Sigma_{fg}$
Fuel	0.444	0.344	0.12
Moderator	0.267	0.262	-

If the statistical error is given by Eq. (34), the standard deviation of the Monte Carlo simulation is given by

$$\sigma = O(\frac{1}{\sqrt{N}}) \quad . \tag{45}$$

This corresponds to a standard Monte Carlo simulation or the VVR method employing a deterministically-calculated adjoint flux. However, if the VVR method is used with Monte Carlo-calculated forward and adjoint fluxes, then the statistical error in the eigenvalue estimate is given by Eq. (35), and the standard deviation in the Monte Carlo simulation is

$$\sigma = O(\frac{1}{N}) \quad . \tag{46}$$

In practice, the standard deviation of the mean is usually reported in the results of a Monte Carlo simulation,

$$\sigma_{\overline{k}} = \frac{\sigma}{\sqrt{M}} \quad . \tag{47}$$

In this case, the statistical uncertainty in the Monte Carlo adjoint VVR method decreases as $O(1/N)$ with respect to the number of histories simulated per fission generation, but only as $O(1/\sqrt{M})$ with respect to the number of fission generations employed. In the numerical simulations described below, we fix the total number of fission generations, and observe the decrease in statistical uncertainty as the number of histories simulated per fission generation increases. For a fixed amount of work, this strategy should minimize the overall statistical errors in the Monte Carlo adjoint VVR method.

6. NUMERICAL RESULTS

We now present the results from two three-dimensional criticality problems of the form of Eq. (1). In these problems, the VVR method with analog tracking is compared to a traditional Monte Carlo calculation using a track-length estimate of the direct functional and survival biasing. (Due to the extra computational expense per history, the VVR method becomes less efficient when used in combination with survival biasing. Hence, we do not present results using the VVR method with survival biasing.)

In comparing the VVR method to traditional Monte Carlo, we examine two three-dimensional criticality problems. Each problem consists of a 25 cm $\times$ 25 cm $\times$ 25 cm cubic reactor, with reflective boundaries on the $x = 0$, $y = 0$, and $z = 0$ surfaces and vacuum conditions on all other boundaries. (This corresponds to a 50 cm $\times$ 50 cm $\times$ 50 cm symmetric reactor with only one-eighth of the reactor being modeled). The first problem is an array of 5 cm $\times$ 5 cm $\times$ 20 cm fuel assemblies in moderator material, with the entire reactor surrounded by 5 cm of moderator material. A diagram of this reactor through the $z = 0$ and $y = 0$ planes is given in Figure 2. The cross-sections used in this problem are given in Table I, in units of cm^{-1}.

We also examine a more difficult reactor with a voided moderator channel. A plot of this voided-channel reactor through the $z = 0$ and $y = 10$ planes is given in Figure 3. The voided channel is simulated by

dividing the moderator cross-sections by a factor of 100, and the void is continued through the reflector to the exterior surface of the reactor.

The two problems described above were simulated using a traditional Monte Carlo calculation employing survival biasing (SB), along with several VVR schemes. These simulations employed 10,000, 50,000, 100,000, 500,000, and 1,000,000 histories per fission generation, along with 225 total fission generations and 25 "skipped" or "inactive" generations. The VVR schemes included P_1 and P_3 representations using a Monte Carlo adjoint (P_1-MC and P_3-MC), and a P_1 representation of the adjoint using a deterministic diffusion calculation (P_1-D). In the Monte Carlo-adjoint VVR methods, the same number of particles were used in the forward and adjoint calculations. These simulations were performed using a 2.5 cm tally grid, while the diffusion calculations were performed on a 1.25 cm grid for spatial discretization.

The statistical error (one standard deviation) in k as a function of computer time for the first problem is given in Figure 4. Since computer time is proportional to the number of histories simulated, the slope of the error should be $-1/2$ if the error is given by Eq. (34), and -1 if the error is given by Eq. (35). From Figure 4, both the traditional Monte Carlo error and the diffusion-estimated adjoint VVR error have the expected slope of -1/2. The P_1-MC and P_3-MC VVR methods converge at a much steeper rate, with a slope close to -1. The efficiency of these methods is less than that of the P_1-D VVR method, due to the extra expense of the Monte Carlo adjoint simulation. However, as more particles are simulated, the efficiency of the P_1-MC and P_3-MC methods should eventually surpass that of the P_1-D method. The P_3-MC method is less efficient than the P_1-MC method, due to the increased number of tallies collected in both the forward and adjoint calculations. However, the P_3-MC method converges slightly faster than the P_1-MC method, and it is likely that for very large computing times, the P_3-MC method will become more efficient.

Next, we examine the more difficult voided-channel problem. The statistical error (one standard deviation) in k as a function of computer time at the end of the Monte Carlo simulation is given in Figure 5. Again, we see that the traditional Monte Carlo simulation and the P_1-D VVR method converge as $1/\sqrt{N}$, while the P_1-MC and P_3-MC errors decrease at a faster rate. For this problem, the P_1-MC method outperforms the P_1-D method for large number of histories per generation because of the greater accuracy of the adjoint Monte Carlo calculation (the diffusion adjoint is not accurate in the void region). As before, the P_3-MC VVR method does not perform as well as the P_1-MC method, due to the cost of the extra adjoint and forward tallies required. However, as more particles are simulated and the accuracy of the adjoint representation increases, the P_3-MC method should eventually outperform the P_1-MC method.

7. CONCLUSIONS

We have examined the use of Monte Carlo-generated forward and adjoint flux estimates for k-eigenvalue problems employing the Variational Variance Reduction technique. In this method, forward and adjoint flux estimates are used to evaluate a variational functional that is theoretically more accurate than the direct functional employed in traditional Monte Carlo calculations. When forward and adjoint Monte Carlo simulations are used in the VVR method, the statistical error will converge at an $O(1/N)$ rate, provided the truncation error in the representation of the adjoint flux estimate is small compared to the statistical error. For N very large, the truncation error in the representation of the adjoint flux estimate will be large compared to the statistical error, and the statistical error in k will converge at the usual $O(1/\sqrt{N})$ rate.

In two example 3-D k-eigenvalue problems, we observed an increased convergence rate in the estimate of k when the VVR method is employed with Monte Carlo adjoint simulation. For reasonable numbers of

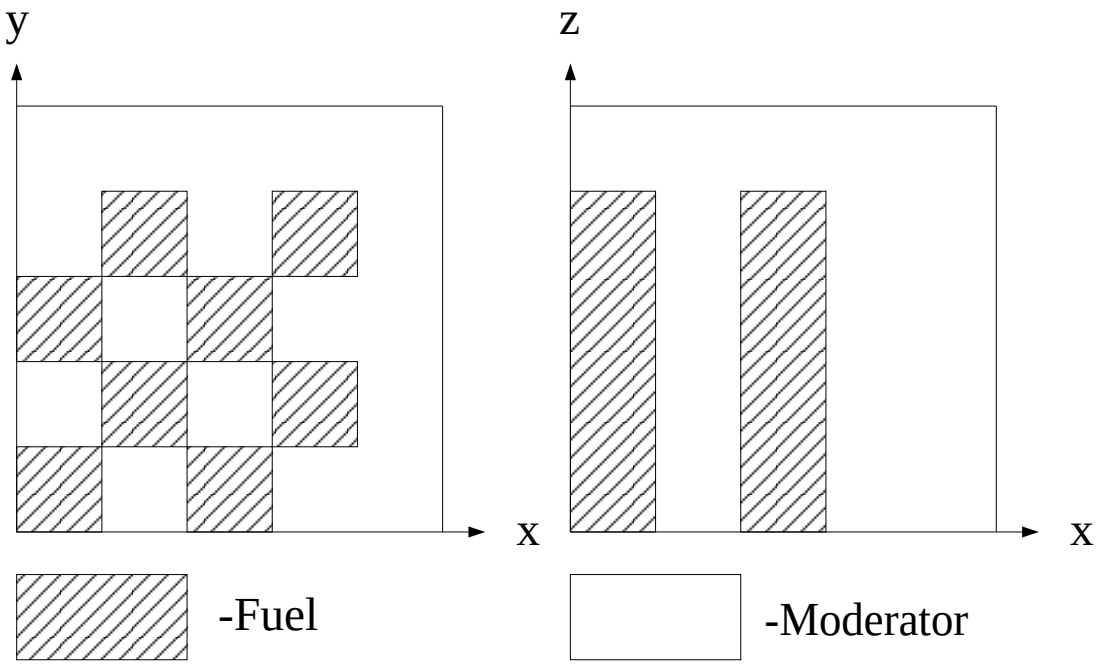

Figure 2. Problem 1 Reactor

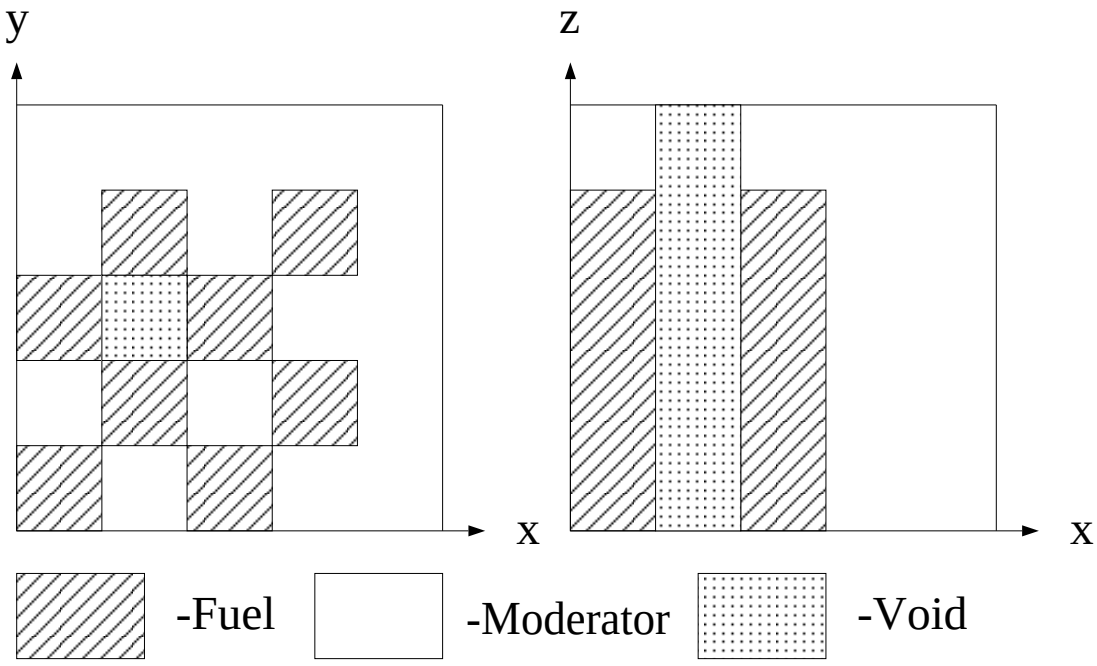

Figure 3. Problem 2 Reactor

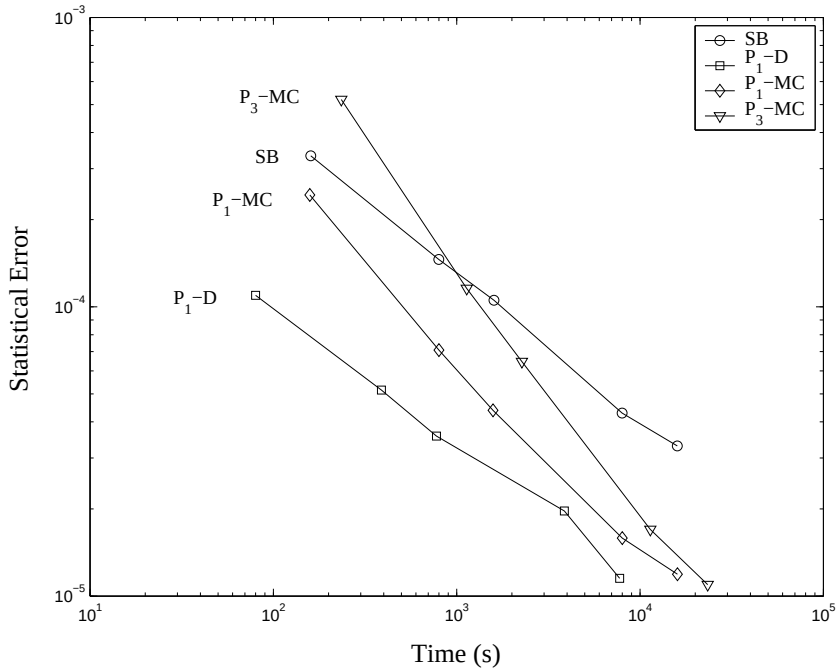

Figure 4. Problem 1 Statistical Error

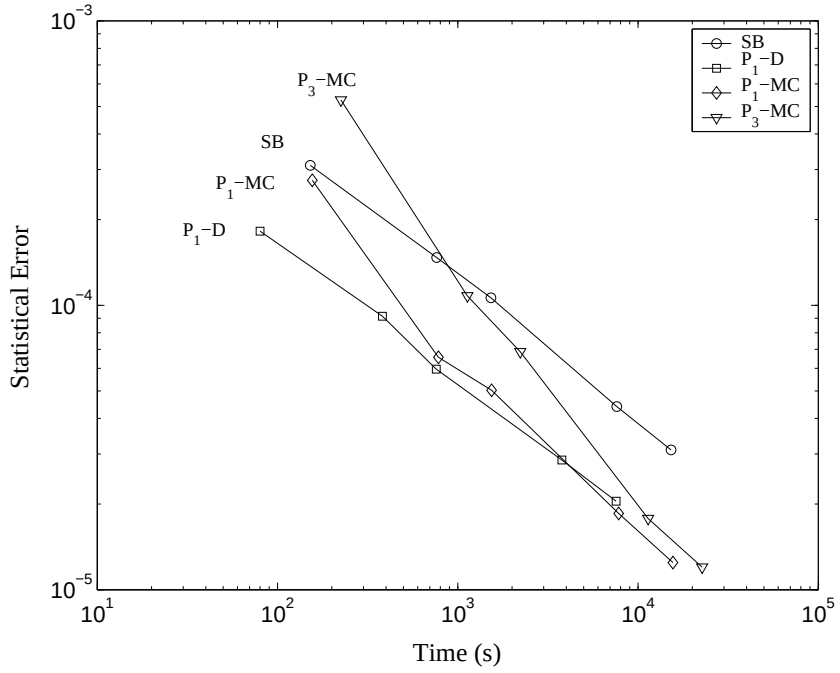

Figure 5. Problem 2 Statistical Error

histories, the Monte Carlo adjoint VVR method is more efficient than the traditional Monte Carlo simulation, but less efficient than the VVR method employing a deterministic diffusion calculation. As the number of histories increases (and if the problem is such that the diffusion calculation produces a less-than-accurate solution) the Monte Carlo adjoint VVR method can outperform other VVR methods that use a deterministic adjoint.

The VVR techniques described in this paper do not require a nonanalog simulation of the forward Monte Carlo particles. Variance reduction occurs because of the processing of the forward and adjoint flux information in the variational functional. However, it is possible to run either the forward or the adjoint Monte Carlo simulations using nonanalog techniques, e.g. survival biasing. The VVR method requires only a good estimate of the adjoint flux and an unbiased estimate of the forward flux; it does not care whether these estimates come from analog Monte Carlo calculations, nonanalog Monte Carlo calculations, or (for the adjoint flux) deterministic calculations. Simulations not shown in this paper indicate that with survival biasing, the VVR methods become less efficient because the longer times required to process particle histories do not outweigh the gain in accuracy [7, 9, 10]. Thus, the VVR simulations presented in this paper used only analog methods to process the histories of Monte Carlo particles. However, there may exist classes of problems for which the use of nonanalog techniques with VVR is advantageous, and if this is the case, then nonanalog techniques can be used.

Our numerical results suggest that the use of a deterministic adjoint in the VVR method is generally more efficient than the use of a Monte Carlo adjoint. However, there are several practical issues that this judgment does not take into account:

1. First, the Monte Carlo adjoint VVR method is generally more efficient than the standard Monte Carlo method for estimating k. Thus – without comparisons to the deterministic adjoint VVR method – the Monte Carlo adjoint VVR method represents generally a gain in efficiency.

2. Developing a Monte Carlo-only VVR code may be easier and require less programming effort than developing a pair of codes (one deterministic, one Monte Carlo) that must be run in tandem, or one dual-purpose code containing both deterministic and Monte Carlo methods.

3. A Monte Carlo-only VVR code would probably be more user-friendly. For example, it would not require the user to understand the accuracy-limitations of a deterministic method that should be weighed in selecting a spatial discretization. A code user would not need to be knowledgeable about both deterministic and Monte Carlo methods.

4. For especially difficult problems, such as ones containing voids and streaming regions – for which deterministic methods are prone to error – it may be advantageous to use a Monte Carlo adjoint in which one has confidence that the estimations will improve with computing time.

5. If estimations of k are required at the level of accuracy where the Monte Carlo adjoint VVR and deterministic adjoint VVR errors generally cross (see Figures 4 and 5), then there is no computational gain in employing a deterministic adjoint.

In any case, we hope that this paper makes a strong argument that for many criticality problems it is more computationally-advantageous to use the VVR method to estimate k than the standard method – whether one uses a Monte Carlo estimate *or* a deterministic estimate for the adjoint flux.

Finally, although this paper considers only monoenergetic, isotropic, scattering criticality problems, the adjoint Monte Carlo VVR method can be extended without conceptual difficulty to source-detector problems, or to more realistic energy-dependent (criticality or source-detector) problems with anisotropic scattering. These generalizations will be the subject of future work.

ACKNOWLEDGMENTS

The work of the first author (JDD) was performed under appointment to the U.S. Department of Energy Nuclear Engineering and Health Physics Fellowship Program, sponsored by DOE's Office of Nuclear Energy, Science, and Technology.

REFERENCES

[1] M.O. Allagi and J.D. Lewins, "Real and Virtual Sampling in Variational Processing of Stochastic Simulation in Neutron Transport: The One-Dimensional Rod," *Ann. Nucl. Energy*, **25**, pp. 1521-1542 (1998).

[2] M.O. Allagi, J.D. Lewins, and G.T. Parks, "Variationally Processed Monte Carlo Transport Theory," *Ann. Nucl. Energy*, **25**, pp. 1055-1068 (1998).

[3] C.L. Barrett, *A Variationally-Based Variance Reduction Method for Monte Carlo Particle Transport Problems*, Ph.D. thesis, University of Michigan (1999).

[4] C.L. Barrett and E.W. Larsen, "A Variationally-Enhanced Monte Carlo Algorithm for Neutron Shielding Problems," *Trans. Am. Nucl. Soc.*, **79**, 166 (1998).

[5] C.L. Barrett and E.W. Larsen, "A Variational Variance Reduction Method for Monte Carlo Shielding and Eigenvalue Problems," *Proceedings of the International Conference on Mathematics and Computation, Reactor Physics, and Environmental Analysis*, Madrid, Spain, September 27-30, Vol. 2, 1371 (1999).

[6] C.L. Barrett and E.W. Larsen, "A Variationally-Based Variance Reduction Method for Monte Carlo Neutron Transport Calculations," *Ann. Nucl. Energy*, **28**, pp. 457-475 (2001).

[7] J.D. Densmore, *Variational Variance Reduction for Monte Carlo Reactor Analysis*, Ph.D. thesis, University of Michigan (2002).

[8] J.D. Densmore and E.W. Larsen, "A New Variational Variance Reduction Method for Monte Carlo Source-Detector Problems," *Trans. Am. Nucl. Soc.*, **83**, pp. 336-338 (2000).

[9] J.D. Densmore and E.W. Larsen, "Variational Variance Reduction for Monte Carlo Criticality Calculations," *Trans. Am. Nucl. Soc.*, **84**, pp. 177-178 (2001).

[10] J.D. Densmore and E.W. Larsen, "Variational Variance Reduction for Monte Carlo Criticality Problems," *Proceedings of the ANS International Meeting on Mathematical Methods for Nuclear Applications*, Salt Lake City, Utah, September 9-13 (2001).

[11] J.D. Densmore and E.W. Larsen, "Variational Variance Reduction for Monte Carlo Eigenfunction Problems," *Trans. Am. Nucl. Soc.*, **87**, 277 (2002).

[12] J.D. Densmore and E.W. Larsen, "Use of Monte Carlo Adjoint Simulations in VVR Criticality Calculations," *Trans. Am. Nucl. Soc.*, **87**, 278 (2002).

[13] E.W. Larsen and J.D. Densmore, "New Zero-Variance Methods for Monte Carlo Criticality Calculations," *Trans. Am. Nucl. Soc.*, **84**, pp. 174-177 (2001).

[14] J.D. Lewins, *Importance: The Adjoint Function*, Pergamon Press, Oxford, United Kingdom (1965).

[15] E.E. Lewis and W.F. Miller, Jr., *Computational Methods of Neutron Transport*, American Nuclear Society, La Grange Park, Illinois (1993).