Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 192 (2003) 387–405
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Variational variance reduction for particle transport eigenvalue calculations using Monte Carlo adjoint simulation

Jeffery D. Densmore [a,*], Edward W. Larsen [b]

[a] Transport Methods Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[b] Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, MI 48109, USA

Received 7 March 2003; received in revised form 8 July 2003; accepted 19 July 2003

Abstract

The Variational Variance Reduction (VVR) method is an effective technique for increasing the efficiency of Monte Carlo simulations [Ann. Nucl. Energy 28 (2001) 457; Nucl. Sci. Eng., in press]. This method uses a variational functional, which employs first-order estimates of forward and adjoint fluxes, to yield a second-order estimate of a desired system characteristic – which, in this paper, is 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 significantly large values of N.

Keywords: Variance reduction; Monte Carlo; Criticality; Variational methods; Adjoint simulation

1. Introduction

The use of variational functionals can increase the efficiency of Monte Carlo simulations [1–14]. In this technique, a system characteristic, such as a detector response or an eigenvalue, is estimated using a variational functional 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,

* Corresponding author. Tel.: +1-505-665-9198.
E-mail addresses: jdd@lanl.gov (J.D. Densmore), edlarsen@engin.umich.edu (E.W. Larsen).

0021-9991/$ - see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/j.jcp.2003.07.007

(iii) estimating the forward solution using Monte Carlo, and (iv) evaluating the variational functional using the deterministic-adjoint and forward Monte Carlo information [3–11,14]. 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 [10,12,13].

Variational functionals gain their accuracy over direct functionals by the fact that first-order errors in the forward and adjoint flux estimates yield second-order errors in the estimate of the desired system characteristic. Direct functionals yield only first-order estimates of system characteristic when first-order estimates of the forward flux are used. Due to their increased accuracy, variational functionals have found many applications in nuclear engineering. Summaries of these applications are available from Becker [15], Lewins [16], Kaplan [17], and Stacey [18]. However, until recently, variational functionals have not been applied to Monte Carlo simulations.

If the statistical error in the Monte Carlo-calculated forward flux is $O(1/\sqrt{N})$, where N is the number of 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.

The VVR method differs from traditional variance reduction schemes in that it does not require the particle transport process to be simulated in an unphysical, or non-analog, manner. Instead, it interprets the information generated by the Monte Carlo simulation in a non-analog way. However, this does not preclude the combination of the VVR method with other variance reduction methods. In fact, if a particular variance reduction technique requires an estimate of the adjoint flux [19–21], this same adjoint estimate may be used in both the VVR method and the variance reduction scheme. This dual use of the adjoint flux can decrease the overhead cost of the adjoint calculation.

The VVR technique becomes more efficient as the accuracy of the adjoint flux estimate increases. In fact, for some applications, if one were to obtain the exact adjoint flux, the VVR method would yield the exact result regardless of the accuracy of the forward flux estimate. In this special case, the VVR method becomes a zero-variance method [22,23].

Allagi and Lewins [2], and Allagi, Lewins, and Parks [1] 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 did not 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 rest 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-diffusion 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. The work presented in this paper is extracted from the Ph.D. dissertation of the first author [10].

2. Problem description

We consider the following 3-D, monoenergetic criticality problem [24]:

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

where an estimate of the eigenvalue k is desired. Here, $\psi(\vec{r},\vec{\Omega})$ is the eigenfunction or forward angular flux, $\vec{r}$ is the spatial variable, and $\vec{\Omega}$ is the angular variable. The spatial domain is represented by V. In Eq. (1), the transport operator L is defined as

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

and the fission operator F is defined by

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

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

Corresponding to the above forward transport problem is the adjoint problem [24]

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

where $\psi^*(\vec{r},\vec{\Omega})$ is the adjoint eigenfunction or angular flux (or, the *importance function* [16]), and the eigenvalue k is identical to the eigenvalue of Eq. (1). In Eq. (4), the adjoint transport operator L^* is defined as

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

and, for the one-group problems considered in this paper, the fission operator F is self-adjoint. Vacuum or reflective boundary conditions for *outgoing* directions apply to Eq. (4). The specific boundary conditions are chosen to be consistent with the forward problem.

The forward and adjoint operators described above have the following properties [24]:

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

and

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

Here $f(\vec{r},\vec{\Omega})$ and $h(\vec{r},\vec{\Omega})$ are arbitrary functions, and we have used the inner-product notation

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

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:

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

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 is the ratio of the total neutron production rate to the total neutron loss rate; this is the standard definition of k [25]. 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,

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

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 neutron production in a given fission generation. The denominator $\langle 1, L\Psi\rangle$ is interpreted as the total neutron loss in a generation. Since fission is treated as a terminal process and the neutron loss is simply the total number of particles in a given generation, division of neutron production by neutron loss is simply a normalization by the total number of histories simulated.

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

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

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.

H has several useful properties. First, we note that

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

and

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

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. This can be seen from Eq. (1). Hence, in the limit of an infinite number of histories used to estimate the forward flux, 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 forward flux estimate. This can be seen from Eqs. (4), (6), and (7) (and assuming Ψ satisfies the forward boundary conditions exactly, but is otherwise arbitrary). In this case, 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 non-analog manner [23]. This zero-variance VVR method differs from other zero-variance methods in the fact that it requires no non-analog 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 k. To do this, we assume $\Psi = \psi + \delta\psi$ and $\Psi^* = \psi^* + \delta\psi^*$ and write Eq. (11) as

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

or

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

The zeroth-order terms on the right-hand side of Eq. (15) (containing no error terms in the numerator) are

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

The second-order term of Eq. (15) (containing only the product $\delta\psi\delta\psi^*$ in the numerator) is

$$\frac{\langle \delta\psi^*, F\delta\psi\rangle\langle \psi^*, L\psi\rangle - \langle \psi^*, F\psi\rangle\langle \delta\psi^*, L\delta\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle\langle \psi^*, L\psi\rangle} = \mathrm{O}(\delta\psi\delta\psi^*), \tag{17}$$

which cannot be simplified. Finally, the first-order term of Eq. (15) (containing only the error terms $\delta\psi$ and $\delta\psi^*$ separately in the numerator) is

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

which can be simplified using Eq. (12) as

$$\frac{\langle \psi^*, F\delta\psi\rangle + \langle \delta\psi^*, F\psi\rangle - k[\langle \delta\psi^*, L\psi\rangle + \langle \psi^*, L\delta\psi\rangle]}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle}. \tag{19}$$

Using the definition of the exact forward flux, Eq. (1), the second and third terms in the numerator of Eq. (19) cancel, leaving

$$\begin{aligned}\frac{\langle \psi^*, F\delta\psi\rangle - k\langle \psi^*, L\delta\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} &= \frac{\langle \psi^*, F\delta\psi\rangle - k\langle \psi^*, L\delta\psi\rangle + \langle \psi^*, F\psi\rangle - k\langle \psi^*, L\psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} \\ &= \frac{\langle \psi^*, F\Psi\rangle - k\langle \psi^*, L\Psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} = \frac{\langle F^*\psi^*, \Psi\rangle - k\langle L^*\psi^*, \Psi\rangle}{\langle \psi^* + \delta\psi^*, L(\psi + \delta\psi)\rangle} = 0,\end{aligned} \tag{20}$$

where we have used Eqs. (4), (6), (7), and (12), and assumed that the forward flux estimate satisfies the boundary conditions exactly. Thus, using Eqs. (16), (17), and (20), Eq. (15) yields the result

$$H[\psi + \delta\psi, \psi^* + \delta\psi^*] = k + \mathrm{O}(\delta\psi\delta\psi^*), \tag{21}$$

which states that first-order estimates of the forward and adjoint fluxes yield a second-order estimate of the eigenvalue k. Hence, the variational function 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 Monte Carlo information. This tally grid is represented by disjoint spatial cells V_i, each with a boundary ∂V_i and a unit outward normal vector $\vec{n}_i$. No constraints are placed on this grid, except that 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 $1 \leqslant i \leqslant I$. These spherical harmonic moments are defined by

$$\Phi^*_{nmi} = \frac{1}{\Delta V_i} \int_{V_i} \int_{4\pi} \overline{Y}_{nm}(\vec{\Omega}) \Psi^*(\vec{r}, \vec{\Omega}) \, \mathrm{d}\Omega \, \mathrm{d}^3 r \tag{22}$$

for $0 \leqslant n \leqslant L$ and $-n \leqslant m \leqslant n$. In Eq. (22), ΔV_i is the volume of tally cell i and L is the order of angular representation desired. In our simulations, we estimate Φ^*_{nmi} using a track-length estimator [24]. Then, using these spherical harmonic moments, we represent Ψ^* as the following histogram-in-space function in each tally cell i:

$$\Psi^*(\vec{r}, \vec{\Omega}) \approx \sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} Y_{nm}(\vec{\Omega}), \quad \vec{r} \in V_i, \ \ \vec{\Omega} \in 4\pi. \tag{23}$$

This representation of the adjoint flux preserves the first $(L+1)^2$-cell-averaged spherical harmonic moments (in this paper, we use $L \leqslant 3$). 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 [7–14]. More elaborate representations of the adjoint flux are possible, including representations that involve polynomial expansions of the angular *and* spatial variables [8–11,14]. 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 is dependent on forward flux information only. To evaluate this reduced functional, we require the following spherical harmonic moments in each tally cell $1 \leqslant i \leqslant I$,

$$\Phi_{nmi} = \int_{V_i} \int_{4\pi} Y_{nm}(\vec{\Omega}) \Psi(\vec{r}, \vec{\Omega}) \, \mathrm{d}\Omega \, \mathrm{d}^3 r, \tag{24}$$

and corresponding spherical harmonic moments over the surface of each tally cell,

$$\Gamma_{nmi} = \int_{\partial V_i} \int_{4\pi} (\vec{\Omega} \cdot \vec{n}_i) Y_{nm}(\vec{\Omega}) \Psi(\vec{r}, \vec{\Omega}) \, \mathrm{d}\Omega \, \mathrm{d}^2 r \tag{25}$$

for $0 \leqslant n \leqslant L$ and $-n \leqslant m \leqslant n$. In our work, we evaluate Eq. (24) using a track-length estimator, and Eq. (25) using a surface-crossing estimator [24].

We now demonstrate the evaluation of the reduced functional using Eq. (23) for Ψ^* and Eqs. (24) and (25) for Ψ. We begin by rewriting Eq. (11) as

$$H[\Psi, \Psi^*] = \frac{\langle \Psi^*, F\Psi \rangle}{\langle \Psi^*, T\Psi \rangle + \langle \Psi^*, C\Psi \rangle - \langle \Psi^*, S\Psi \rangle}, \tag{26}$$

where Ψ^* is the adjoint representation given by Eq. (23). In Eq. (26), the streaming operator T is defined by

$$T\psi = \vec{\Omega} \cdot \vec{\nabla}\psi(\vec{r}, \vec{\Omega}), \tag{27}$$

the collision operator C is defined by

$$C\psi = \Sigma_t(\vec{r})\psi(\vec{r}, \vec{\Omega}), \tag{28}$$

and the isotropic scattering operator S is defined by

$$S\psi = \frac{\Sigma_s(\vec{r})}{4\pi} \int_{4\pi} \psi(\vec{r}, \vec{\Omega}')\, d\Omega'. \tag{29}$$

From Eq. (2), Eqs. (27)–(29) compose the transport operator L:

$$L\psi = T\psi + C\psi - S\psi. \tag{30}$$

We now evaluate Eq. (26) term by term. First, the fission term is

$$\langle \Psi^*, F\Psi \rangle = \int_V \int_{4\pi} \Psi^*(\vec{r}, \vec{\Omega}) \left[\frac{\nu\Sigma_f(\vec{r})}{4\pi} \int_{4\pi} \Psi(\vec{r}, \vec{\Omega}')\, d\Omega' \right] d\Omega\, d^3r. \tag{31}$$

Introducing Eq. (23), Eq. (31) can be written as

$$\langle \Psi^*, F\Psi \rangle = \sum_{i=1}^{I} \int_{V_i} \int_{4\pi} \left[\sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} Y_{nm}(\vec{\Omega}) \right] \left[\frac{\nu\Sigma_{f_i}}{4\pi} \int_{4\pi} \Psi(\vec{r}, \vec{\Omega}')\, d\Omega' \right] d\Omega\, d^3r, \tag{32}$$

where $\nu\Sigma_{f_i}$ is ν times the fission cross-section in tally cell i. Using the identity

$$1 = \sqrt{4\pi} Y_{00}(\vec{\Omega}) = \sqrt{4\pi} \overline{Y}^*_{00}(\vec{\Omega}), \tag{33}$$

and the orthogonality property of the spherical harmonic functions, Eq. (32) yields

$$\langle \Psi^*, F\Psi \rangle = \frac{1}{4\pi} \sum_{i=1}^{I} \sqrt{4\pi} \Phi^*_{00i} \nu\Sigma_{f_i} \int_{V_i} \int_{4\pi} \Psi(\vec{r}, \vec{\Omega}')\, d\Omega'\, d^3r. \tag{34}$$

Finally, using Eqs. (24) and (33), Eq. (34) can be rewritten as

$$\langle \Psi^*, F\Psi \rangle = \sum_{i=1}^{I} \nu\Sigma_{f_i} \Phi^*_{00i} \Phi_{00i}. \tag{35}$$

Next, we evaluate the (isotropic) scattering term

$$\langle \Psi^*, S\Psi \rangle = \int_V \int_{4\pi} \Psi^*(\vec{r}, \vec{\Omega}) \left[\frac{\Sigma_s(\vec{r})}{4\pi} \int_{4\pi} \Psi(\vec{r}, \vec{\Omega}')\, d\Omega' \right] d\Omega\, d^3r. \tag{36}$$

Since Eq. (36) is similar to Eq. (31), we can evaluate the scattering term in the same manner as the fission term:

$$\langle \Psi^*, S\Psi \rangle = \sum_{i=1}^{I} \Sigma_{s_i} \Phi^*_{00i} \Phi_{00i}. \tag{37}$$

In Eq. (37), Σ_{s_i} is the scattering cross-section in tally cell i. If scattering is anisotropic, the summation extends over all spherical harmonic moments present in Ψ^*.

The collision term is written as

$$\langle \Psi^*, C\Psi \rangle = \int_V \int_{4\pi} \Psi^*(\vec{r},\vec{\Omega})\Sigma_t(\vec{r})\Psi(\vec{r},\vec{\Omega})\,d\Omega\,d^3r. \tag{38}$$

Employing Eq. (23), Eq. (38) becomes

$$\langle \Psi^*, C\Psi \rangle = \sum_{i=1}^{I} \int_{V_i} \int_{4\pi} \left[\sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} Y_{nm}(\vec{\Omega}) \right] \Sigma_{t_i} \Psi(\vec{r},\vec{\Omega})\,d\Omega\,d^3r, \tag{39}$$

where Σ_{t_i} is the total cross-section in tally cell i. Applying Eq. (24), Eq. (39) yields

$$\langle \Psi^*, C\Psi \rangle = \sum_{i=1}^{I} \sum_{n=0}^{L} \sum_{m=-n}^{n} \Sigma_{t_i} \Phi^*_{nmi} \Phi_{nmi}. \tag{40}$$

Finally, the streaming term is written as

$$\langle \Psi^*, T\Psi \rangle = \int_V \int_{4\pi} \Psi^*(\vec{r},\vec{\Omega})\vec{\Omega}\cdot\vec{\nabla}\Psi(\vec{r},\vec{\Omega})\,d\Omega\,d^3r. \tag{41}$$

Using the adjoint representation given by Eq. (23), Eq. (41) yields

$$\langle \Psi^*, T\Psi \rangle = \sum_{i=1}^{I} \int_{V_i} \int_{4\pi} \left[\sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} Y_{nm}(\vec{\Omega}) \right] \vec{\Omega}\cdot\vec{\nabla}\Psi(\vec{r},\vec{\Omega})\,d\Omega\,d^3r. \tag{42}$$

Next, we apply the divergence theorem [26], given by

$$\int_{V_i} \vec{\nabla}\cdot\vec{F}(\vec{r})\,d^3r = \int_{\partial V_i} \vec{n}_i \cdot \vec{F}(\vec{r})\,d^2r, \tag{43}$$

where $\vec{F}(\vec{r})$ is arbitrary vector-valued function. Eq. (42) becomes

$$\langle \Psi^*, T\Psi \rangle = \sum_{i=1}^{I} \int_{\partial V_i} \int_{4\pi} \left[\sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} Y_{nm}(\vec{\Omega}) \right] (\vec{\Omega}\cdot\vec{n}_i)\Psi(\vec{r},\vec{\Omega})\,d\Omega\,d^2r. \tag{44}$$

Then, employing Eq. (25), Eq. (44) can be written as

$$\langle \Psi^*, T\Psi \rangle = \sum_{i=1}^{I} \sum_{n=0}^{L} \sum_{m=-n}^{n} \Phi^*_{nmi} \Gamma_{nmi}. \tag{45}$$

Thus, using Eqs. (35), (37), (40), and (45), Eq. (26) can be completely evaluated using Monte Carlo estimates of ψ^* and ψ obtained by adjoint and forward calculations. In this implementation of the VVR method, the adjoint information is provided by Eq. (22), and the forward information is provided by Eqs. (24) and (25).

We note that in the evaluation of the variational functional H, the adjoint flux estimate Ψ^* is approximated by Eq. (23) – which contains spatial and angular truncation errors – but the forward flux estimate Ψ is not represented in a way that contains truncation errors. Specifically, the approximation for Ψ^* is introduced into H [Eq. (11) or (26)], yielding a reduced functional that only requires moments of forward flux estimate [Eqs. (24) and (25)] to be numerically evaluated. It is not necessary to approximate Ψ in a manner

similar to the way Ψ^* is approximated in Eq. (23). Therefore, if the exact forward flux ψ is used to evaluate Eqs. (24) and (25), the exact eigenvalue k will be calculated. Hence, the VVR method will yield the exact eigenvalue in the limit of an infinite number of histories being used in the forward Monte Carlo simulation, irrespective of the error in the adjoint estimate. However, as in standard Monte Carlo criticality calculations, VVR results using a finite number of Monte Carlo histories will contain bias [27,28].

When using the above VVR method to perform 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 [24]. 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. (35), (37), (40), and (45). Thus, a variationally calculated estimate of k is provided by each generation. We use these generation-wise estimates to calculate a mean and a sample variance, as in traditional Monte Carlo.

If a deterministic method is used to estimate the adjoint flux, the constants Φ^*_{nmi} are determined from this calculation prior to the beginning of the forward Monte Carlo simulation. The subsequent Monte Carlo simulation proceeds as described in the above paragraph, with the difference that only forward Monte Carlo calculations are performed. In previous work, we have shown that this deterministic-adjoint VVR method is a viable way of computing k [8–10,14]. In the numerical simulations discussed below, we compare this deterministic-adjoint VVR method to the Monte Carlo-adjoint VVR method described above.

5. Error analysis of variational estimates

We now discuss the performance of the VVR method described above. From Eq. (21), for each individual fission generation, the error in the variationally estimated eigenvalue satisfies

$$\text{Error} = \mathrm{O}(\delta\psi\delta\psi^*). \tag{46}$$

If the error in the Monte Carlo-estimated forward flux is $\mathrm{O}(1/\sqrt{N})$, 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. (46) gives

$$\text{Error} = \mathrm{O}\left(\frac{\delta\psi^*}{\sqrt{N}}\right), \tag{47}$$

where $\delta\psi^*$ is the error in the adjoint flux estimate. For a fixed estimate of the adjoint flux, Eq. (47) is characteristic of most Monte Carlo simulations, the error being inversely proportional to the square root of the computational effort. However, the error is also 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 [3–11,14], 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. (47), we observed convergence of the estimate of k with $\mathrm{O}(1/\sqrt{N})$ error, but with a 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* adjoint flux estimates are obtained using Monte Carlo – each having error $\mathrm{O}(1/\sqrt{N})$ (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. (46) becomes:

$$\text{Error} = \mathrm{O}\left(\frac{1}{N}\right). \tag{48}$$

Thus, in the presence of only statistical errors, i.e., $\delta\psi = \mathrm{O}(1/\sqrt{N})$ and $\delta\psi^* = \mathrm{O}(1/\sqrt{N})$, 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. (48) 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. (23)], 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 $\mathrm{O}(1/\sqrt{N})$. However, for N not large, a convergence rate faster than $\mathrm{O}(1/\sqrt{N})$ will occur.

To illustrate the effect of the truncation error in Eq. (23), we write the error in the forward flux estimate as

$$\delta\psi = \frac{A}{\sqrt{N}} \tag{49}$$

and the error in the adjoint flux estimate as

$$\delta\psi^* = \frac{A^*}{\sqrt{N}} + B^*. \tag{50}$$

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. (23). By Eq. (46), the error in the VVR estimate of k satisfies

$$\text{Error} \cong \left(\frac{A}{\sqrt{N}}\right)\left(\frac{A^*}{\sqrt{N}} + B^*\right). \tag{51}$$

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

$$\text{Error} \approx \frac{AA^*}{N}, \quad N \ll \left(\frac{A^*}{B^*}\right)^2. \tag{52}$$

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

$$\text{Error} \approx \frac{AB^*}{\sqrt{N}}, \quad \left(\frac{A^*}{B^*}\right)^2 \ll N. \tag{53}$$

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

$$N \cong \left(\frac{A^*}{B^*}\right)^2, \tag{54}$$

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 $\mathrm{O}(1/N)$ for a significant number of histories N.

In Fig. 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

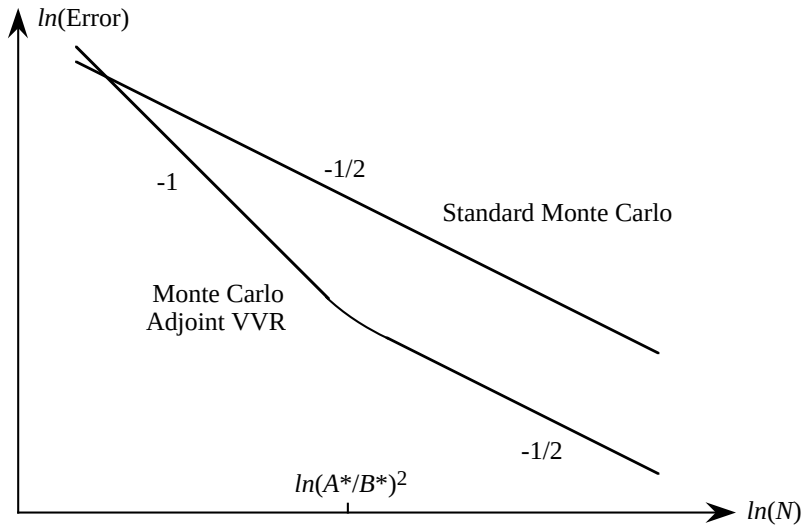

Fig. 1. Convergence of Monte Carlo adjoint VVR.

closer resemblance to the left half of Fig. 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 also note that, for N small, the statistical error of the Monte Carlo-adjoint VVR method may be larger than that of standard Monte Carlo simulation. This behavior, which is observed in our numerical results, is caused by a poorly estimated adjoint eigenfunction when N is small.

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

$$k_m = k + e_m, \tag{55}$$

where k_m is the eigenvalue estimate of the mth 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

$$\bar{k} = E[k_m] = k + E[e_m], \tag{56}$$

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

$$\sigma^2 = E[(k - k_m)^2] = E\left[e_m^2\right]. \tag{57}$$

If the statistical error is given by Eq. (47), the standard deviation of the Monte Carlo simulation is given by

$$\sigma = \mathrm{O}\left(\frac{1}{\sqrt{N}}\right). \tag{58}$$

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. (48), and the standard deviation in the Monte Carlo simulation is

$$\sigma = \mathrm{O}\left(\frac{1}{N}\right). \tag{59}$$

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

$$\sigma_{\bar{k}} = \frac{\sigma}{\sqrt{M}}. \tag{60}$$

In this case, the statistical uncertainty in the Monte Carlo-adjoint VVR method decreases as $\mathrm{O}(1/N)$ with respect to the number of histories simulated per fission generation, but only as $\mathrm{O}(1/\sqrt{M})$ 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.

Eqs. (58)–(60) also give some insight in how the computational effort should be allocated in each of the different Monte Carlo techniques. When solving criticality problems, the user must determine: (i) the total number of neutrons to simulate in a single fission generation, N and (ii) the total number of fission generations to simulate, M. It is of interest if there is a combination of N and M, for a fixed computational expense, that optimizes the statistical error, $\sigma_{\bar{k}}$.

We first consider a standard Monte Carlo simulation, or the VVR method employing a deterministic-adjoint flux estimate. In this case, the total work is $\mathrm{O}(NM)$. From Eqs. (58) and (60), the statistical error at the end of the simulation is

$$\sigma_{\bar{k}} = \mathrm{O}\left(\frac{1}{\sqrt{\text{work}}}\right). \tag{61}$$

Thus, in the case of standard Monte Carlo or deterministic-adjoint VVR, it does not matter how the work is divided. One can have more particles per generation and fewer generations, or vice versa (as long as M and N individually do not become too small). The resulting error in k will not be sensitive to the individual values of M and N, for a fixed amount of work.

In the case of Monte Carlo adjoint VVR, the total work is $\mathrm{O}(2NM)$ (there are N particles in each forward and adjoint fission generation). From Eqs. (59) and (60), the statistical error at the end of the simulations is

$$\sigma_{\bar{k}} = \mathrm{O}\left(\frac{\sqrt{M}}{\text{work}}\right). \tag{62}$$

In this case, the error is minimized if M is minimized – or at least held constant. Thus, for a fixed amount of work, more computational effort should be allocated to simulating more particles per fission generation. This estimate should be valid until N and computational effort become large.

6. Numerical results

We now present the results from two 3-D 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. However, survival biasing does improve the efficiency of standard Monte Carlo, for the problems presented here.)

In comparing the VVR method to traditional Monte Carlo, we examine two 3-D criticality problems. Each problem consists of a 25 cm × 25 cm × 25 cm system, 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 × 50 cm × 50 cm symmetric system with only one-eighth of the problem being modeled.) The first problem is an array of 5 cm × 5 cm × 20 cm fuel regions in moderator material, with the fuel array surrounded by 5 cm of moderator material. A diagram of this system through the $z = 0$ and $y = 0$ planes is depicted in Fig. 2. The cross-sections used in this problem are given in Table 1, in units of cm^{-1}.

We also examine a more difficult problem with a voided moderator channel. A plot of the voided-channel system through the $z = 0$ and $y = 0$ planes is depicted in Fig. 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 system.

We simulated these two problems using a standard 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 scheme included a P_1 representation of the adjoint flux using a deterministic-diffusion adjoint calculation (P_1-D), and a P_L representation of the adjoint flux using a Monte Carlo adjoint calculation, where $0 \leqslant L \leqslant 3$ (P_L-MC). In the Monte Carlo-adjoint VVR methods, we used the same number of particles in the forward and adjoint calculations. These simulations were performed using a 2.5-cm tally grid, while the diffusion calculations were performed on 1.25-cm fine grid for spatial discretization. We note that although the Monte Carlo-adjoint VVR simulations employed twice as many particles per generation as the standard Monte Carlo simulations

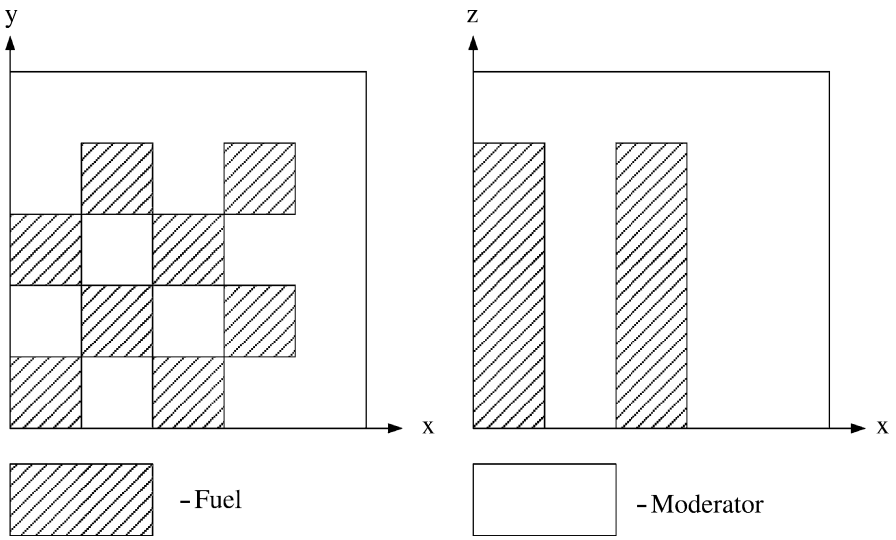

Fig. 2. Problem 1 geometry.

Table 1
Monoenergetic cross-sections

Material	Σ_t	Σ_s	$\nu\Sigma_f$
Fuel	0.444	0.344	0.12
Moderator	0.267	0.262	–

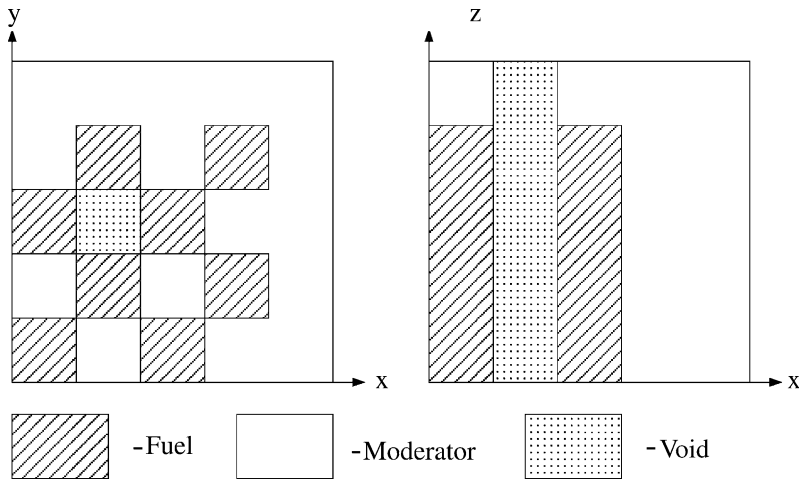

Fig. 3. Problem 2 geometry.

employed survival biasing, whereas all VVR methods used analog tracking. Thus, the SB method will be more costly per generation than the P_1-D and P_L-MC methods.

The statistical error (one standard deviation) in k as a function of computer time for the first problem is given in Fig. 4. Since computer time is proportional to the number of histories simulated, the slope of the error will be $-1/2$ if the error is given by Eq. (47), and -1 if the error is given by Eq. (48). From Fig. 4, the

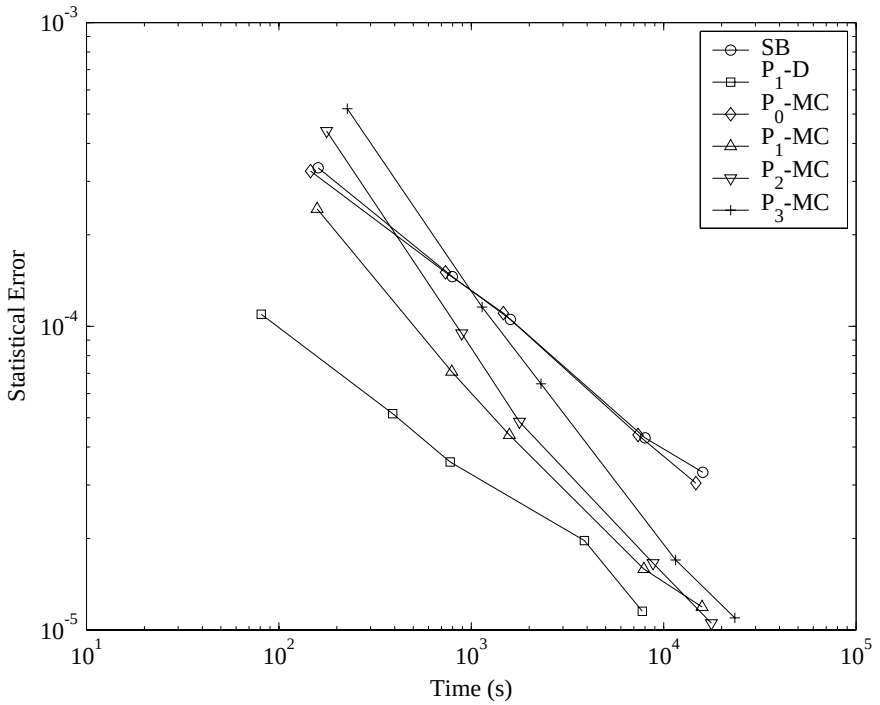

Fig. 4. Problem 1 statistical error.

standard Monte Carlo error and the diffusion-estimated-adjoint VVR error both have the expected slope of $-1/2$. The P_0-MC VVR method performs no more efficiently than standard Monte Carlo. This illustrates the need for higher angular resolution when employing the VVR method. The P_1-MC, P_2-MC, and P_3-MC VVR methods converge at a much steeper rate, with a slope close to -1. The efficiency of these method 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, P_2-MC, and P_3-MC VVR methods should eventually surpass that of the P_1-D method. We also note that, for most numbers of histories per generation, increasing the order of angular representation in the Monte Carlo-adjoint VVR method past $L = 1$ decreases the efficiency of the method. This is due to the increased number of forward and adjoint tallies required by this higher-order angular representation. However, these higher-order VVR methods converge at a greater rate than the P_1-MC method, and for large computer times, the efficiency of the P_2-MC, and P_3-MC methods should surpass that of the P_1-MC method.

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 Fig. 5. As with the first problem, the P_0-MC VVR method has about the same efficiency as standard Monte Carlo. Again, we see that the SB and P_1-D methods converge as $1/\sqrt{N}$, while the P_1-MC, P_2-MC, and P_3-MC errors decrease at a faster rate. For this problem, the P_1-MC and P_2-MC methods outperform the P_1-D method for a larger 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 other Monte Carlo-adjoint VVR methods, due to the cost of the required extra adjoint and forward tallies. However, as more particles are simulated and the accuracy of the

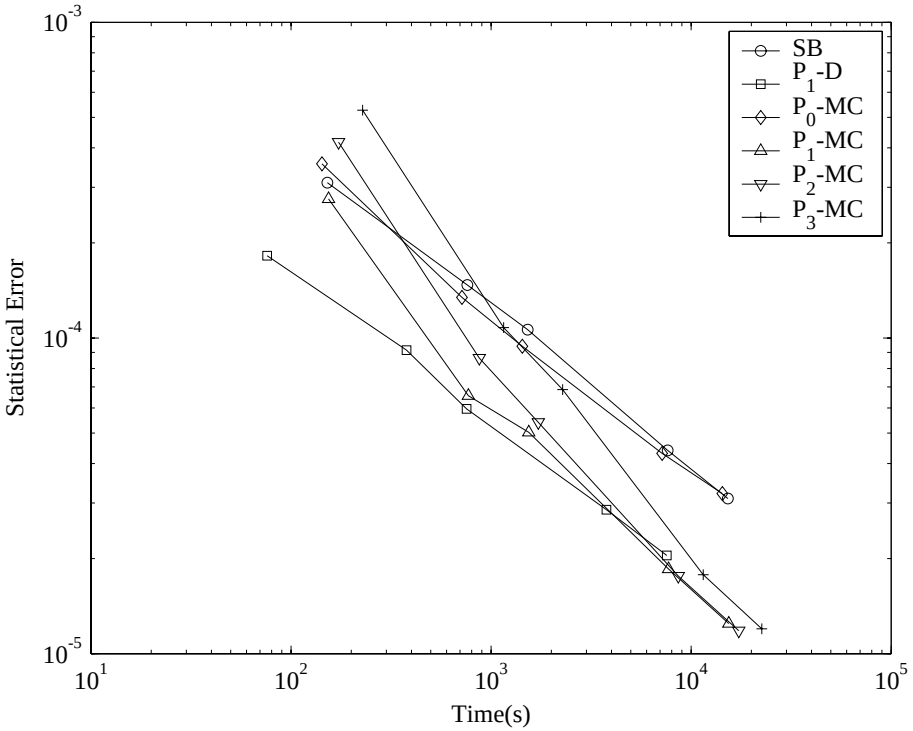

Fig. 5. Problem 2 statistical error.

Table 2
Simulation results

Simulation method	Problem 1 k	Problem 2 k
SB	0.960903 ± 0.000033	0.955561 ± 0.000031
P_1-D	0.960917 ± 0.000012	0.955502 ± 0.000020
P_0-MC	0.961022 ± 0.000030	0.955447 ± 0.000032
P_1-MC	0.960906 ± 0.000012	0.955500 ± 0.000012
P_2-MC	0.960921 ± 0.000011	0.955531 ± 0.000012
P_3-MC	0.960917 ± 0.000011	0.955515 ± 0.000012
Diffusion	0.9488	0.8967

adjoint flux estimate increases, the P_3-MC method should eventually outperform the P_1-MC and P_2-MC methods.

The resulting final eigenvalue estimates for both problems, along with their statistical errors, are given in Table 2 for each simulation method. For reference, we also include the eigenvalue estimated by the deterministic-diffusion adjoint calculation. We note that all VVR methods produced an eigenvalue estimate that is much more accurate than the diffusion-estimated eigenvalue. Also, most VVR-calculated eigenvalues were within a few standard deviations of the reference SB eigenvalue. The one exception is the eigenvalue calculated by the P_0-MC VVR method. For both problems, this simulation method produced an eigenvalue that was more than a few standard deviations away from the SB eigenvalue. It is not clear if this difference is due to the poor accuracy of the P_0 representation of the adjoint flux, or some not-understood bias due to too few histories being simulated in each fission generation.

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 histories, the Monte Carlo adjoint VVR method is more efficient than 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 non-analog simulation of the forward Monte Carlo particles, and the VVR numerical results reported above are in fact based on forward analog Monte Carlo simulations. 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 non-analog 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 matter whether these estimates come from analog Monte Carlo calculations, non-analog Monte Carlo

calculations, or (for the adjoint flux) deterministic calculations. Simulations not shown in this paper indicate that with survival biasing, the VVR method becomes less efficient because the longer times required to process particles histories outweigh the gain in accuracy [8–10]. However, there may exist classes of problems for which the use of non-analog techniques with VVR is advantageous, and if this is the case, the non-analog techniques can and should be used.

Also, our choice of a 2.5-cm "tally" grid and a 1.25-cm "fine" grid in our numerical simulations was motivated by our intuitive desire to obtain: (i) an accurate diffusion solution at minimal cost and (ii) an accurate representation of Ψ^* [Eq. (23)] at minimal cost. We did not perform extensive numerical experiments to determine optimal tally and fine grids. In general, we believe that (i) the truly optimal choice of tally and fine grids will be problem dependent and (ii) the behavior of the VVR method will not be seriously impacted by reasonable variations (from optimal) in the choice of these grids. Nevertheless, the choice of tally and fine grids is a potentially important issue in practical applications and deserves further study.

In this paper, we showed that the statistical error in the Monte Carlo-adjoint VVR method converges as $O(1/N)$, where N is the number of histories per fission generation – not the number of fission generations simulated. This fact may require users to the change the way they apply Monte Carlo simulation to criticality problems. The standard approach to solving criticality problems is to fix the number of histories per generation and then to simulate fission generations until a desired level of statistical error is reached. In the Monte Carlo-adjoint VVR method, a user should minimize the total number of fission generations, and redirect computational effort to simulating more histories in each fission generation [see Eq. (62)]. The Monte Carlo-adjoint VVR method can proceed in the standard manner, simulating fission generations until the desired statistical error is reached. However, it may be difficult to determine a priori how many histories to simulate in each generation in order to take full advantage of the $O(1/N)$ convergence of this VVR method. This is another issue we have not discussed, but which may be significant in practical applications.

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 a 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) The Monte Carlo-adjoint VVR method is more easily extended to continuous-energy problems than the deterministic-adjoint VVR method. A Monte Carlo simulation does not require a multigroup energy discretization, as a deterministic method does. The ability to directly use continuous energy cross-sections also gives a Monte Carlo adjoint simulation an accuracy advantage over a deterministic-adjoint calculation.
(6) If estimation of k is required at the level of accuracy where the Monte Carlo-adjoint VVR and deterministic-adjoint VVR errors generally cross (see Figs. 4 and 5), then there is no computational gain in employing a deterministic adjoint.

In any case, we hope that this paper makes a solid 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 subject of future work.

Acknowledgements

The work of the first author (J.D.D.) 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, J.D. Lewins, G.T. Parks, Variationally processed Monte Carlo transport theory, Ann. Nucl. Energy 25 (1998) 1055.
[2] M.O. Allagi, J.D. Lewins, Real and virtual sampling in variational processing of stochastic simulation in neutron transport: the one-dimensional rod, Ann. Nucl. Energy 25 (1998) 1521.
[3] C.L. Barrett, E.W. Larsen, A variationally-enhanced Monte Carlo algorithm for neutron shielding problems, Trans. Am. Nucl. Soc. 79 (1998) 166.
[4] C.L. Barrett, A variationally-based variance reduction method for Monte Carlo particle transport problems, PhD Dissertation, University of Michigan, 1999.
[5] C.L. Barrett, E.W. Larsen, A variational variance reduction method for Monte Carlo shielding and eigenvalue problems, in: Proceedings of the International Conference on Mathematics and Computation, Reactor Physics, and Environmental Analysis, Madrid, Spain, September 27–30, 1999, vol. 2, p. 1371.
[6] C.L. Barrett, E.W. Larsen, A variationally-based variance reduction method for Monte Carlo neutron transport calculations, Ann. Nucl. Energy 28 (2001) 457.
[7] J.D. Densmore, E.W. Larsen, A new variational variance reduction method for Monte Carlo source-detector problems, Trans. Am. Nucl. Soc. 83 (2000) 336.
[8] J.D. Densmore, E.W. Larsen, Variational variance reduction for Monte Carlo criticality calculations, Trans. Am. Nucl. Soc. 84 (2001) 177.
[9] J.D. Densmore, E.W. Larsen, Variational variance reduction for Monte Carlo criticality problems, in: Proceedings of the ANS International Meeting of Mathematical Methods for Nuclear Applications, Salt Lake City, UT, September 9–13, 2001.
[10] J.D. Densmore, Variational variance reduction for Monte Carlo reactor analysis, PhD Dissertation, University of Michigan, 2002.
[11] J.D. Densmore, E.W. Larsen, Variational variance reduction for Monte Carlo eigenfunction problems, Trans. Am. Nucl. Soc. 87 (2002) 277.
[12] J.D. Densmore, E.W. Larsen, Use of Monte Carlo adjoint simulation in VVR criticality calculations, Trans. Am. Nucl. Soc. 87 (2002) 278.
[13] J.D. Densmore, E.W. Larsen, Variational variance reduction for criticality calculations using Monte Carlo adjoint fluxes, in: Proceedings of the Mathematics and Computations 2003: A Century In Review – A Century Anew, Gatlinburg, TN, April 6–10, 2003.
[14] J.D. Densmore, E.W. Larsen, Variational variance reduction for Monte Carlo eigenvalue and eigenfunction problems, Nucl. Sci. Eng., in press.
[15] M. Becker, The Principles and Applications of Variational Methods, The MIT Press, Cambridge, MA, 1964.
[16] J.D. Lewins, Importance: The Adjoint Function, Pergamon Press, Oxford, UK, 1965.
[17] S. Kaplan, Variational methods in nuclear engineering, Adv. Nucl. Sci. Tech. 5 (1969) 185.
[18] W.M. Stacey Jr., Variational Methods in Nuclear Reactor Physics, Academic Press, New York, 1974.
[19] S.A. Turner, E.W. Larsen, Automatic variance reduction for three-dimensional Monte Carlo simulations by the local importance function transform – I: analysis, Nucl. Sci. Eng. 127 (1997) 22.
[20] S.A. Turner, E.W. Larsen, Automatic variance reduction for three-dimensional Monte Carlo simulations by the local importance function transform – II: numerical results, Nucl. Sci. Eng. 127 (1997) 36.

[21] K.A. Van Riper, T.J. Urbatsch, P.D. Soran, D.K. Parsons, J.E. Morel, G.W. McKinney, S.R. Lee, L.A. Crotzer, F.W. Brinkley, T.E. Booth, J.W. Anderson, R.E. Alcouffe, AVATAR – Automatic variance reduction in Monte Carlo calculations, in: Proceedings 1997 Joint International Conference on Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs, New York, October 6–10, 1997, vol. 1, p. 661.
[22] I. Lux, L. Koblinger, Monte Carlo Particle Transport Methods: Neutron and Photon Calculations, CRC Press, Boca Raton, FL, 1991.
[23] E.W. Larsen, J.D. Densmore, New zero-variance methods for Monte Carlo criticality calculations, Trans. Am. Nucl. Soc. 84 (2001) 174.
[24] E.E. Lewis, W.F. Miller Jr., Computational Methods of Neutron Transport, American Nuclear Society, La Grange Park, IL, 1993.
[25] J.J. Duderstadt, L.J. Hamilton, Nuclear Reactor Analysis, John Wiley & Sons, New York, 1976.
[26] J.E. Marsden, A.J. Tromba, Vector Calculus, W.H. Freeman & Company, New York, 1996.
[27] R.J. Brissenden, A.R. Garlick, Biases in the estimation of k_{eff} and its error by Monte Carlo methods, Ann. Nucl. Energy 13 (1986) 63.
[28] E.M. Gelbard, A.G. Gu, Biases in Monte Carlo eigenvalue calculations, Nucl. Sci. Eng. 117 (1994) 1.