HYBRID MONTE CARLO–DETERMINISTIC METHODS FOR REACTOR ANALYSIS

FISSION REACTORS

KEYWORDS: *FW-CADIS, SUBSPACE approach, hybrid Monte Carlo–deterministic methods*

ZEYUN WU, QIONG ZHANG, and HANY ABDEL-KHALIK*

North Carolina State University, Raleigh, North Carolina 27695-7909

Received May 20, 2011
Accepted for Publication January 10, 2012

A new variant of a hybrid Monte Carlo–deterministic approach for simulating particle transport problems is presented and compared to the SCALE FW-CADIS approach. The new approach, denoted as the SUBSPACE approach, improves the selection of the importance maps in order to reduce the computational overhead required to achieve global variance reduction—that is, the uniform reduction of variance everywhere in the phase-space. The intended applications are reactor analysis problems where detailed responses for all fuel assemblies are required everywhere in the reactor core. Like FW-CADIS, the SUBSPACE approach utilizes importance maps obtained from deterministic adjoint models to derive automatic weight-window biasing. Unlike FW-CADIS, the SUBSPACE approach does not employ flux-based weighting of the adjoint source term. Instead, it utilizes pseudoresponses generated with random weights to help identify the correlations between the importance maps that could be used to reduce the computational time required for global variance reduction. Numerical experiments, serving as proof of principle, are presented to compare the SUBSPACE and FW-CADIS approaches in terms of the global reduction in standard deviation and the associated figures of merit for representative nuclear reactor assembly and core models.

I. INTRODUCTION

Over the course of the last five decades, stochastic Monte Carlo methods and deterministic methods have been developed separately for simulating particle transport problems. Recently, there has been growing interest in coupling Monte Carlo and deterministic methods by employing a hybrid approach to combine their benefits and overcome some of their individual deficiencies.[1–6] The main idea is to bias Monte Carlo sampling using an estimate of the solution obtained inexpensively from a simplified deterministic model. In the Monte Carlo community, this procedure represents a form of "variance reduction." The procedure has been successfully demonstrated to reduce the variance for a single response,[1–3] often representing a functional of the solution over a region in the phase-space, i.e., a detector's response in a given region. Global variance reduction, however, still remains an important challenge, especially when Monte Carlo methods are to be used for reactor analysis applications. Global variance reduction denotes problems where one seeks to reduce the variances for all responses evaluated everywhere in the phase-space, such as group fluxes, reaction rates density, and homogenized few-group cross sections. Several approaches have been proposed in the literature to uniformly distribute Monte Carlo particles throughout the domain. Among the most prominent approaches are the correction,[2] the variational variance reduction,[3] the forward-weighted consistent adjoint-driven importance sampling[4] (FW-CADIS), the linear tally combination,[5] and the coarse-mesh finite difference.[6] Our approach may be considered as a variant of the FW-CADIS or the linear tally combination approaches. We focus our comparison on the FW-CADIS approach since all numerical experiments were conducted using the Monaco with Automated Variance Reduction using Importance Calculations (MAVRIC) sequence[7] of the SCALE 6.0 package,[8] which is based on the FW-CADIS approach.

Hybrid approaches such as FW-CADIS employ an approximate adjoint function to assign importance values to various regions in the phase-space. The importance of a particle describes the contribution of this particle to the response of interest. In principle, if the adjoint map is known exactly and employed by a zero-variance biasing scheme, the response could be obtained with zero variance. Solving the adjoint problem exactly, however, is as

*E-mail: abdelkhalik@ncsu.edu

difficult as solving the forward problem. Hence, it is computationally sufficient to employ an approximation of the adjoint solution to bias the forward Monte Carlo solution via the use of weight windows. Given a weight window that is not zeroed anywhere in the phase-space, one is guaranteed to reach the exact solution in the limit.

When more than one response is needed (assume a total of I responses), the adjoint-based weight-window maps are expected to be different for different responses. This problem may be addressed in one of three ways: First, one could execute the Monte Carlo simulation I times in a brute-force manner, each corresponding to a different single-response weight-window map. Clearly, the brute-force approach will be computationally intractable as I is expected to be large for realistic problems; I represents the total number of responses evaluated everywhere in the phase-space. Second, one could form one weight-window map that captures the important features of each map. This could be done by creating a linear combination of the importance maps associated with the original single-response weight-window maps, with the weights reflecting the importance of each map. This is the approach adopted by FW-CADIS, where the weights are selected to be proportional to the inverse of the forward flux. In this approach, referred to hereinafter as FW-CADIS, a response evaluated at a region in the phase-space will have higher weight for its weight-window map if its corresponding flux is small. The logic behind this approach is that regions receiving fewer particles will have higher variances for their associated responses. Therefore, by sending more particles to these regions, the variances of their associated responses are expected to decrease.

While the idea behind the FW-CADIS approach is based on sound reasoning, it does not represent the only way to assign weights, which is expected to depend largely on the application of the model. Moreover, it does not assure that the associated computational burden will be less than that reached by the brute-force approach. The following numerical experiment demonstrates this situation for a model involving two responses with relatively independent weight-window maps: Consider a point detector model problem with only two responses as depicted in Fig. 1. Each response represents a point detector on the side of a concrete shield with the distributed source in the center. All dimensions are in centimeters, and both detectors are located 10 cm from the shielding surface. Shield 1 is twice as thick as shield 2 in order to render detector 1's response noticeably smaller than detector 2's response.

Given the weight-window maps for the two responses, the total computational time with the brute-force approach is equal to the sum of the times required to separately reduce the variance for each detector to the desired level, e.g., $\sigma = 4\%$. Employing the FW-CADIS approach, the total time required (represented by number of histories) is found to be considerably

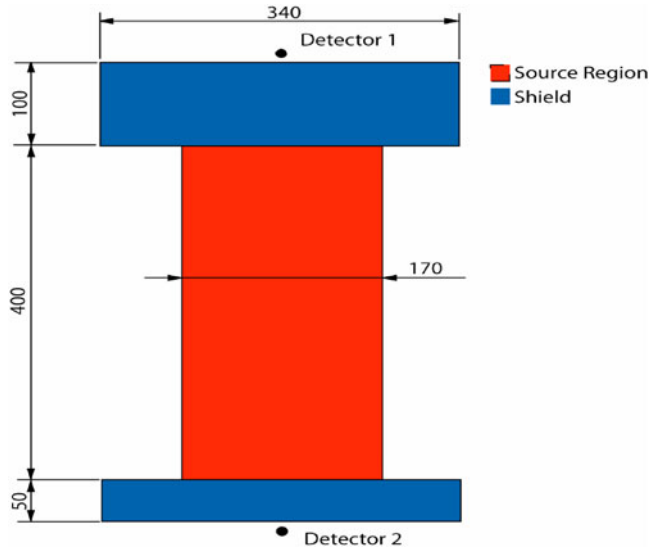

Fig. 1. Test case for FW-CADIS and brute-force approaches.

higher than for the brute-force approach. Figure 2 compares the number of histories required to reach the same level of variance for the two detectors. The first two cases are produced using the brute-force approach, and the last two cases are generated using the FW-CADIS approach.

In this paper, we propose a third approach to addressing the simultaneous reduction of variances for I responses, i.e., global variance reduction. It is denoted as the SUBSPACE approach, and it may be considered as a trade-off between the brute-force and the FW-CADIS approaches, where instead of evaluating all I weight-window maps (brute force) or a single weight-window map (FW-CADIS), only a small number r of pseudoresponse maps are evaluated, such that $1 < r \ll I$. The pseudoresponses are random linear combinations of the original I single responses. Implementing this approach into the FW-CADIS framework should be straightforward (detailed in Sec. III), as it will only require the execution of FW-CADIS r times with the results combined statistically to determine the responses' mean values and standard deviations. Moreover, the weights for the pseudoresponses are generated randomly using the SUBSPACE approach thus eliminating the need for an extra forward model execution to determine the flux-based weights as currently done by the FW-CADIS approach. Finally, given the independence of the r executions, the SUBSPACE approach allows for coarse-grained parallelization, thereby taking advantage of parallel computing environments.

This paper is organized as follows. Section II introduces the mathematical theory behind the proposed approach and a general algorithm for its implementation. A general algorithm is sought because we believe the idea behind the proposed approach can be applied to other

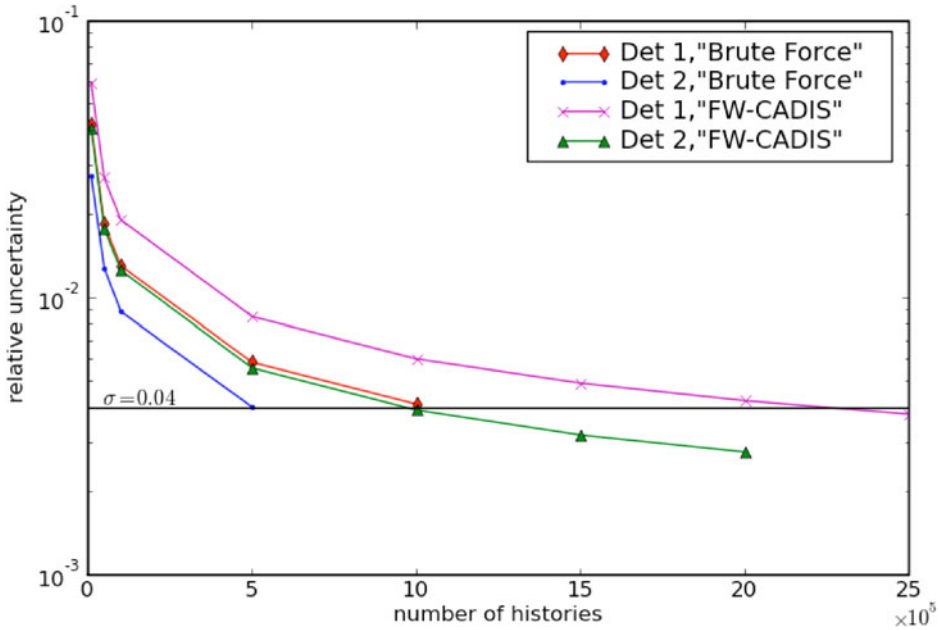

Fig. 2. Comparison of FW-CADIS and brute-force approaches.

variance reduction techniques that employ importance maps to bias Monte Carlo particles. For the sake of comparison with the FW-CADIS approach, Sec. III describes how the proposed approach could be directly implemented into the FW-CADIS framework. Numerical experiments and concluding remarks are presented in Secs. IV and V, respectively.

II. PROPOSED SUBSPACE APPROACH

The formation of the pseudoresponses and their associated weight-window maps is mathematically equivalent to projecting the single-response weight-window maps onto a subspace of smaller dimensions that captures their variability. This is possible because the I single-response weight-window maps will likely be correlated. This typically happens when responses represent distributions that are evaluated everywhere in the phase-space. For example, in a core model, the flux in a fuel assembly is expected to be highly correlated to the fluxes in the nearby assemblies. Thermal fluxes are expected to correlate more to nearby assemblies than fast fluxes because of the shorter mean free path. Identifying these correlations in an automated manner could be used to identify the minimum number of weight-window maps that are independent, denoted by r. We show that each of the r independent correlations represents a weight-window map that is associated with a pseudoresponse. By reducing the variances for the pseudoresponses, one can effectively reduce the variances for the original I responses. If $r \ll I$, computational savings could be achieved. Earlier work has shown that in reactor calculations, responses representing distributions such as group fluxes and reaction rates are highly correlated.[9,10]

The degree of correlations between the responses can be described by the singular values decline of the matrix containing the single-response importance maps. This may be described mathematically as follows: Let u_i be the response (i.e., tally) calculated at the i'th mesh cell. Mathematically, it may be described by the inner product of the forward flux solution $\vec{\psi}$ and a response function $\vec{\sigma}_i$ of the form

$$u_i = \langle \vec{\psi}, \vec{\sigma}_i \rangle \quad \text{and} \quad i = 1, \ldots, I \ , \tag{1}$$

where I is the total number of mesh cells, which is the same as the total number of tallies (i.e., responses). The importance map for u_i is obtained as the solution of an adjoint problem of the form

$$L^* \vec{\psi}_i^* = \frac{\partial u_i}{\partial \vec{\psi}} = \vec{\sigma}_i \ , \tag{2}$$

where

L^* = adjoint transport operator

$\vec{\psi}_i^*$ = importance map associated with response u_i.

In MAVRIC, the deterministic code[7] DENOVO is used to obtain an approximate solution $\vec{\psi}_i^*$ of the adjoint problem in Eq. (2) on a coarse grid in phase-space whose grid points may be indexed by $j = 1, \ldots, J$. The $\vec{\psi}_i^*$ may be written as

$$\vec{\psi}_i^* = [\psi_{i,1}^* \ldots \psi_{i,J}^*]^T \ , \tag{3}$$

where $\psi_{i,j}^*$ describes the importance values of the particles entering the phase-space at point j, which eventually contributes to the response u_i. The importance maps corresponding to all responses may be assembled in a matrix $\mathbf{\Psi}$ of the form. We denote this matrix by single-responses-importance (SRI) matrix:

$$\mathbf{\Psi} = \begin{bmatrix} \psi_{1,1}^* & \cdots & \psi_{1,J}^* \\ \vdots & & \vdots \\ \psi_{I,1}^* & \cdots & \psi_{I,J}^* \end{bmatrix} = \begin{bmatrix} \vec{\psi}_1^{*T} \\ \vdots \\ \vec{\psi}_I^{*T} \end{bmatrix} \ . \tag{4}$$

In the numerical experiments in Sec. IV we show that a typical SRI matrix of interest to reactor analysis problems exhibits a significant decline in its singular value spectrum. This implies that a great deal of correlation exists between the importance maps of its different responses. Research in linear algebra has shown that one could take advantage of this behavior by approximating the matrix with other matrices of much smaller dimensions. The dimension of the smaller matrices is determined by the effective rank of the matrix $\mathbf{\Psi}$ (Ref. 11). Linear algebra is replete with matrix decomposition methods that could be used to determine the rank of a matrix whose elements are explicitly available, which is not the case in our problem. Explicit evaluation of the matrix implies evaluation of the adjoint model for all possible responses, which is overwhelming for routine design calculations. In the past 10 years, research in the applied linear algebra community has shown that great insight into the singular value spectrum could be obtained via simple matrix-vector products operations employing random vectors.[12] This means that only operations of the form $\mathbf{\Psi}^T\vec{\eta}_j$ (where $\vec{\eta}_j \in \mathbf{R}^I$ is a randomly generated vector) are required to determine the decline in the singular values that could be used to determine an effective rank for the matrix. We show next how these randomized matrix-vector products could be easily generated as pseudoresponses, representing random linear combination of the original I responses. Let $\tilde{u}_j$ be the j'th pseudoresponse defined by

$$\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} u_i \quad \text{for } j = 1, \ldots, r \ . \tag{5}$$

Using the definitions for u_i and $\vec{\psi}_i^*$ from Eqs. (1) and (2), one can write

$$\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} \langle \vec{\psi}, \vec{\sigma}_i \rangle = \left\langle \vec{\psi}, \sum_{i=1}^{I} \eta_{i,j} \vec{\sigma}_i \right\rangle$$

and

$$L^* \tilde{\vec{\psi}}_j^* = \frac{\partial \tilde{u}_j}{\partial \vec{\psi}} = \sum_{i=1}^{I} \eta_{i,j} \vec{\sigma}_i \ .$$

Using the linearity of the transport operator, one can write an expression for the importance map for the pseudoresponse $\tilde{u}_j$ as

$$\tilde{\vec{\psi}}_j^* = \sum_{i=1}^{I} \eta_{i,j} \vec{\psi}_i^* \quad \text{for } j = 1, \ldots, r \ . \tag{6}$$

Note that $\tilde{\vec{\psi}}_j^*$ is a linear combination of all I importance maps, which can be rewritten using linear algebra notations as

$$\tilde{\vec{\psi}}_j^* = \mathbf{\Psi}^T \vec{\eta}_j \ , \tag{7}$$

where $\vec{\eta}_j^T = [\eta_{1,j} \ \ldots \ \eta_{I,j}]$. If the $\{\vec{\eta}_j\}_{j=1}^r$ are randomly generated, one can show that the vectors $\{\tilde{\vec{\psi}}_j^*\}_{j=1}^r$ are independent and span a subspace of size r, which belongs to the range of the matrix $\mathbf{\Psi}^T$ (Refs. 9 and 12). Let $\mathbb{C}$ represent the subspace generated by the vectors $\{\tilde{\vec{\psi}}_j^*\}_{j=1}^r$, and $\mathbb{C}^\perp$ as the orthogonal subspace. Now, split each of the importance vectors into two components: one that lives in the subspace $\mathbb{C}$, $\{{}^{\mathbb{C}}\vec{\psi}_i^*\}_{i=1}^I$, and the other that lives in the subspace $\mathbb{C}^\perp$, $\{{}^{\mathbb{C}^\perp}\vec{\psi}_i^*\}_{i=1}^I$. Using elements from random matrix theory,[12] one can show that as r is increased, the components in the subspace $\mathbb{C}^\perp$ continue to shrink, and the components in the subspace $\mathbb{C}$ continue to increase. More importantly, for most realistic problems, the components living in the $\mathbb{C}^\perp$ subspace are significantly reduced with a small estimate for the rank r; this is because the major decline of the singular values is expected to happen over the first few dimensions associated with the highest singular values. Therefore, most of the acceleration rendered by the proposed SUBSPACE approach is expected to happen with a small estimate for the rank r.

The general algorithm to implement this approach may be described as follows:

Requirements:

- a general methodology that employs an importance map $\vec{\psi}_i^*$ to bias Monte Carlo particles toward a given response u_i
- the capability to calculate an importance map $\tilde{\vec{\psi}}_j^*$ for a pseudoresponse defined as a random linear combination of the original I responses as defined in Eq. (6).

Objective:

- Identify r pseudoresponses and employ them to reduce variance for all I responses.

Algorithm:

(a) Estimate the rank r. If no prior knowledge about the rank is available, pick a small value, e.g., $5 < r < 20$, and execute step (b). Calculate the singular value decomposition of the matrix containing the importance maps for the r pseudo-

responses: $[\tilde{\tilde{\psi}}_1 \ \tilde{\tilde{\psi}}_2 \ \ldots \ \tilde{\tilde{\psi}}_r]$. If the singular values do not significantly decline, increase the estimate for r.

(b) PARALLEL DO $j = 1, \ldots, r$

1. Generate a random vector $\tilde{\eta}_j \in \mathbb{R}^I$
2. Form a pseudoresponse $\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} u_i$
3. Calculate the importance map $\tilde{\tilde{\psi}}_j^*$ associated with $\tilde{u}_j$
4. Bias Monte Carlo particles based on the $\tilde{\tilde{\psi}}_j^*$
5. Tally the original I responses until the number of histories is exhausted
6. Record the responses $u_{i,j}^{\mu}$ and their standard deviations $u_{i,j}^{\sigma}$

END DO

(c) COMBINE the responses and their standard deviations from the r runs as follows[13]:

$$u_i^{\mu} = u_i^{\sigma} \sum_{j=1}^{r} \frac{u_{i,j}^{\mu}}{(u_{i,j}^{\sigma})^2} \quad \text{and} \quad \frac{1}{u_i^{\sigma}} = \sum_{j=1}^{r} \frac{1}{(u_{i,j}^{\sigma})^2} \ . \tag{8}$$

End Result:

- The u_i^{μ} and u_i^{σ} are the mean and the standard deviation, respectively, for the i'th response calculated by the SUBSPACE approach.

This algorithm is composed of three steps. Step (a) requires an estimate of the rank r. In Sec. IV, we show that the variance reduction is not very sensitive to the choice of the rank estimate. For interested readers, we include in the Appendix an elaborate algorithm that describes how the rank for the matrix $\mathbf{\Psi}$ could be determined exactly using r_{ex} adjoint model executions, where r_{ex} is the exact rank. Step (b) represents an execution of an existing variance reduction approach with a special choice for the pseudoresponse. Since the importance function is often calculated using an adjoint model, this should be fairly easy to implement for most codes via simple manipulation of the right side of the adjoint equation. Section III provides more details on this step for incorporating the SUBSPACE approach into the FW-CADIS framework. Step (c) combines the results from the r executions, each with N independent histories, under the assumption that they are statistically independent.[13] The expression for the mean value implies that simulations with high variance will have little impact on the unbiased estimate for the mean value. The formula for the variance implies that the overall variance is reduced as more simulations are executed, which is consistent with the law of Monte Carlo sampling.

III. FW-CADIS–BASED IMPLEMENTATION

In this section, we discuss the implementation of the SUBSPACE approach into the FW-CADIS framework. FW-CADIS proceeds in two steps: First, a forward deterministic model is executed to calculate an estimate for the flux everywhere in the phase-space, and second, an adjoint model is executed. The inverse of the flux estimated by the forward model is employed to design the right side of the adjoint model, referred to as the adjoint source, which can be expressed as the derivative of the response with respect to the flux[4] [see Eq. (2)]. This results in giving more weight to regions in the phase-space where the flux is low and less weight to regions with high flux. In addition to flux-based weights, FW-CADIS allows for user-defined weights for each region in the phase-space. To implement the SUBSPACE approach, the user-defined weights are selected randomly and are assigned via the input file. This is equivalent to setting a pseudoresponse as a random linear combination of all responses as required by Eq. (5). The resulting adjoint solution satisfies Eq. (6), which is equivalent to multiplying the SRI matrix with a random vector.

Next, as discussed earlier, the SUBSPACE approach does not require the flux-based weights, thus eliminating the need for the extra forward model execution. Fortunately, this is also possible as the FW-CADIS framework is flexible enough to provide the user the option to bypass forward flux weighting. Therefore, to implement the SUBSPACE approach, one must specify an estimate for the rank r. Next, the FW-CADIS sequence is executed in parallel r times with the forward flux weighting bypassed and random weights assigned to the adjoint source; both of these could be specified via the input cards to FW-CADIS. After the r FW-CADIS executions are completed, a script is needed to read the responses and their standard deviations and to statistically combine them as given by Eq. (8). This implementation strategy has been adopted in our work and is employed in Sec. IV to analyze both a core model and an assembly model.

IV. NUMERICAL EXPERIMENTS

This section is divided into three subsections. Section IV.A employs a boiling water reactor (BWR) assembly model, and Sec. IV.B employs a pressurized water reactor (PWR) core model. Both compare the performance of the FW-CADIS and SUBSPACE approaches with an estimated rank of $r = 10$. Section IV.C investigates the impact of the rank estimate on the variance reduction results.

IV.A. Assembly Model

The first experiment is based on a two-dimensional MAVRIC model for a 7×7 BWR assembly; a cross section of the assembly is shown in Fig. 3. The assembly

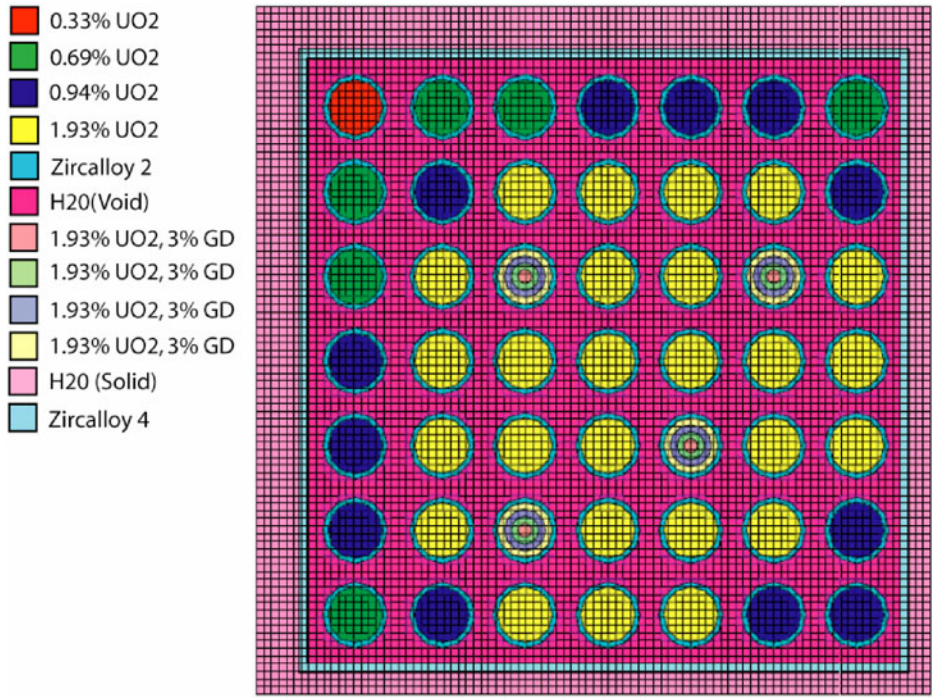

Fig. 3. A 7 × 7 BWR model.

model represents the southeast assembly of a typical 2 × 2 BWR control cell that contains four assemblies and a cruciform control blade. Within the scope of this preliminary work, the control blades are not modeled. The BWR assembly contains 49 fuel rods of different compositions in a regular 7 × 7 fuel rod array. A single mesh tally is defined over the 49 square regions each comprising a fuel rod and its neighboring moderator within what is commonly referred to as a fuel pin cell. The thermal flux tallied over each fuel pin cell is employed as a response giving rise to a total of 49 responses, i.e., $I = 49$.

Employing the MAVRIC procedure, the mesh tallies, the corresponding responses, and their weight windows are defined over each fuel pin cell. The whole array is surrounded by a layer of zirconium and an outer layer of water. Essential technical data of the model problem are given in Table I.

TABLE I

BWR Model Specifications

Assembly pitch (cm)	15.24
Fuel pitch (cm)	1.8745
Fuel rod diameter (cm)	1.2116
Cladding thickness (cm)	0.1092
Canning thickness (cm)	0.2032
Material temperature (K)	552.833

In contrast to usual practice in reactor physics calculation where k-eigenvalue computations are performed, a fixed-source subcritical configuration is analyzed instead because of the current limitations of the MAVRIC sequence—MAVRIC is the sequence in the SCALE package that embodies the FW-CADIS approach. The publicly available version of the MAVRIC sequence does not have an eigenvalue solver and is currently limited to source-driven problems only. To overcome this limitation, a NEWT model is employed to approximate the fission source, which is subsequently reduced by adjusting fuel enrichment to render a subcritical system. An isotropic fixed source distributed uniformly throughout the fuel pins is employed to find a physical flux solution. The fuel enrichment is adjusted rendering a k-effective value of 0.88. Five different fuel enrichments of 0.33%, 0.69%, 0.94%, and 1.93% ^{235}U and 3% gadolinium are employed. The 27 neutron and 19 photon energy group libraries from SCALE are employed for the analysis of the BWR model. For the flux and the reaction rate responses, the first 14 neutron groups (10.678 eV $< E <$ 20 MeV) define the fast group, and the last 13 groups ($E <$ 3.059 eV) are thermal.

Both the FW-CADIS and the SUBSPACE approaches were employed to analyze this model with equal number of histories. Let $\sigma_i^{\text{FW-CADIS}}$ and $\sigma_i^{\text{SUBSPACE}}$ denote the relative standard deviations for the i'th response—the relative standard deviation is the ratio of the response's absolute standard deviation to the response's mean value. Figure 4 shows the reduction in the relative standard deviation for thermal flux distribution in the assembly.

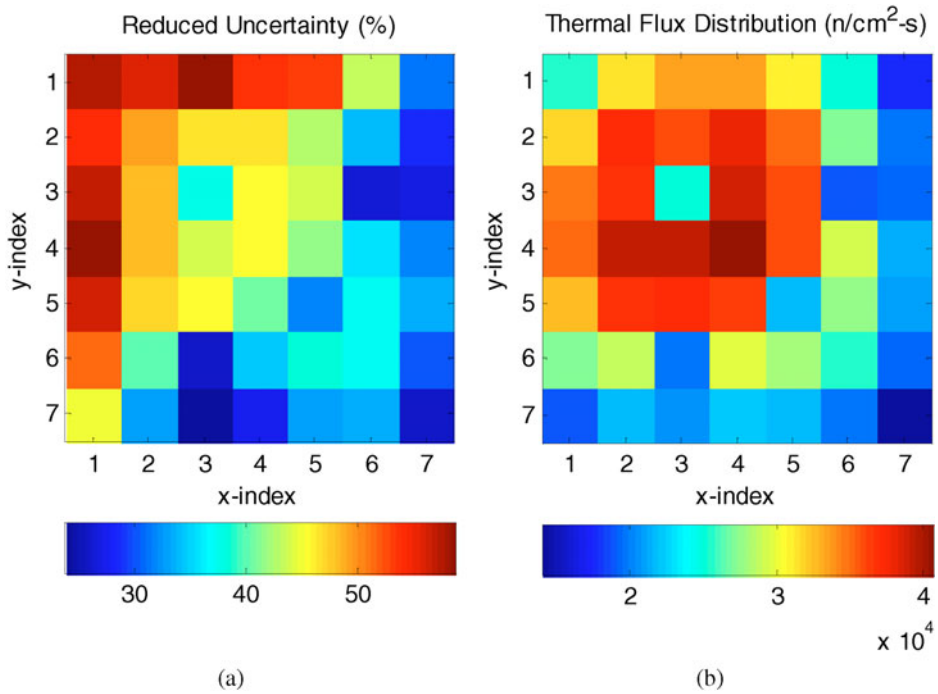

Fig. 4. Variance reduction results for BWR 7 × 7 assembly model; (a) standard deviation reduction and (b) mean thermal flux distribution.

The reduction in the relative standard deviation resulting from the use of the SUBSPACE approach is defined by

$$\Delta\sigma_i = \frac{\sigma_i^{\text{FW-CADIS}} - \sigma_i^{\text{SUBSPACE}}}{\sigma_i^{\text{FW-CADIS}}} * 100\% \ . \tag{9}$$

For reference, the mean of the thermal flux distribution is also shown in Fig. 4, where each square represents the thermal flux tallied over one pin cell.

Table II provides a more detailed comparison between the two approaches. The column in Table II labeled AVG is the average value of the relative standard deviations for all I responses. The column labeled STD represents the standard deviation of the responses' relative standard deviations (this is more commonly known as the square root of the variance of variances); this value provides a measure of the spread of the standard deviations throughout the phase-space. A large value for STD indicates that the response variances are not uniformly reduced; therefore, a small value would be considered more favorable. Mathematically, each of these metrics is defined as follows for a given approach k, where k = FW-CADIS and SUBSPACE:

$$\text{AVG}^k = \frac{1}{I}\sum_{i=1}^{I} \sigma_i^k$$

TABLE II

Performance Metrics for FW-CADIS and SUBSPACE Approaches in an Assembly Model

Approach	Relative Standard Variance		Number of Monte Carlo Particles	Deterministic Execution Time (s)	Monte Carlo Execution Time (min)	FOM
	AVG[a]	STD[b]				
FW-CADIS	0.2037	0.0454	1.00E+07[c]	31.48	469.88	0.0513
SUBSPACE	0.1182	0.0271	1.00E+07	173.36	581.42	0.1231

[a]AVG = average.
[b]STD = standard.
[c]Read as 1.00×10^7.

and

$$\mathrm{STD}^k = \sqrt{\frac{1}{I-1}\sum_{i=1}^{I}(\sigma_i^k - \mathrm{AVG}^k)^2} \quad .$$

The number of Monte Carlo particles employed by the FW-CADIS approach is 10^7 particles. The same number was used by the SUBSPACE approach but distributed over ten different runs (corresponding to an estimated rank of $r = 10$) each with 10^6 particles. The deterministic execution time for the FW-CADIS approach is composed of one adjoint and one forward run. The SUBSPACE deterministic time is composed of ten adjoint runs. The Monte Carlo time is the time spent by the MONACO code.[a] Finally, the figure of merit (FOM) is based on the following formula:

$$\mathrm{FOM}^k = \frac{1}{(\mathrm{AVG}^k)^2 T} \quad ,$$

where T is the total time including both the deterministic and the Monte Carlo times. Notice that the deterministic time is negligible compared to the Monte Carlo time, which is to be expected since all deterministic calculations are based on source-driven models. These assembly results show a 2.5 speedup factor over FW-CADIS results. Moreover, notice that the STD metric is reduced by the same amount as the AVG metric, implying that the SUBSPACE approach does reduce the variances in a uniform manner like in the FW-CADIS methodology.

IV.B. Core Model

The second numerical experiment employs a prototypical PWR full-core model. It is designed as a slight variation to the benchmark problems presented in Refs. 14 and 15. The full-core model consists of 193 fuel assemblies (blue regions in Fig. 5) laid out in a 17 × 17 grid scheme and surrounded by light water (red regions in Fig. 5) (colors online only).

The cubic volume of the whole active core is 365.6 × 365.6 × 335.3 cm^3. The cubic volume of each assembly is 21.505 × 21.505 × 335.28 cm^3. Two types of fuel assemblies are modeled (blue regions): a UO_2 fuel assembly and a UO_2-Gd_2O_3 fuel assembly. The loading pattern of the full core is shown in Fig. 5. Each assembly consists of a 17 × 17 grid of pin cells with each pin cell measuring 1.265 × 1.265 cm^2 in the X-Y plane.

Figure 6 shows the relative reduction in the thermal flux distribution throughout the core resulting from using the SUBSPACE approach. The mean thermal flux is also displayed for reference as was done in Fig. 4

In a similar manner to Table II, the core model results are listed in Table III. Notice that the amount of the reduction in terms of FOM is doubled as the size of the problem is increased. Although this may not sound intuitive at first glance, from previous experience we have noticed that as the size of the model is increased the potential for reduction via response correlations is also increased.

IV.C. Rank Estimate

In this section we demonstrate the sensitivity of the variance reduction results to the rank estimate. As discussed earlier, one could employ a rigorous approach to estimate the exact rank of the matrix $\mathbf{\Psi}$ such as the range-finding algorithm described in the Appendix. However, in most applications employing Monte Carlo models, a small estimate of the rank should be sufficient. This is because the very first few singular values of the matrix $\mathbf{\Psi}$ display a significant decline with the rate of decline decreasing with increased rank. To illustrate this, the algorithm in the Appendix is employed to estimate the first 30 singular values of the matrix $\mathbf{\Psi}$. This could be achieved by executing the algorithm with different user-defined tolerance.[16]

Notice that the singular values plotted in Fig. 7 fall by three orders of magnitude by the time the tenth singular value is reached. After that, the singular values continue to fall but at a much lower rate. Given that the statistical uncertainties for the responses are expected to be in the 0.1% to 1% range, only the initial reduction in the singular values should be sufficient to estimate the rank.

To analyze the impact of the rank estimate on the variance reduction results, the assembly and core model results, previously completed with $r = 10$, are repeated with different estimates for the rank. Figure 8 plots the standard deviation for one of the responses as a function of the estimated rank. Results show that the initial decline in the standard deviation occurs over the first few singular values, which is consistent with the shape of the singular values. After that, the reduction in the standard deviation is negligible.

[a] We noticed that the time spent by MONACO is always slightly higher when using the SUBSPACE approach, which implies the weights for the adjoint source are supplied by the user via the input file rather than evaluated directly by the code. To understand this, a weight-window map was generated using the same approach employed by FW-CADIS, i.e., based on the inverse of the forward flux, and was then manually fed into the MONACO code. We noticed that although the same response means and standard deviations were obtained as with the standard MAVRIC sequence, the time required was also higher like in the SUBSPACE approach. This implies that the MONACO code requires additional time likely when reading the weight windows from an input buffer. This is a minor issue and can likely be handled by experienced code developers. For the sake of the current work, the higher times recorded by MONACO are employed in all FOM results, so slightly better results should be expected upon resolution of this issue.

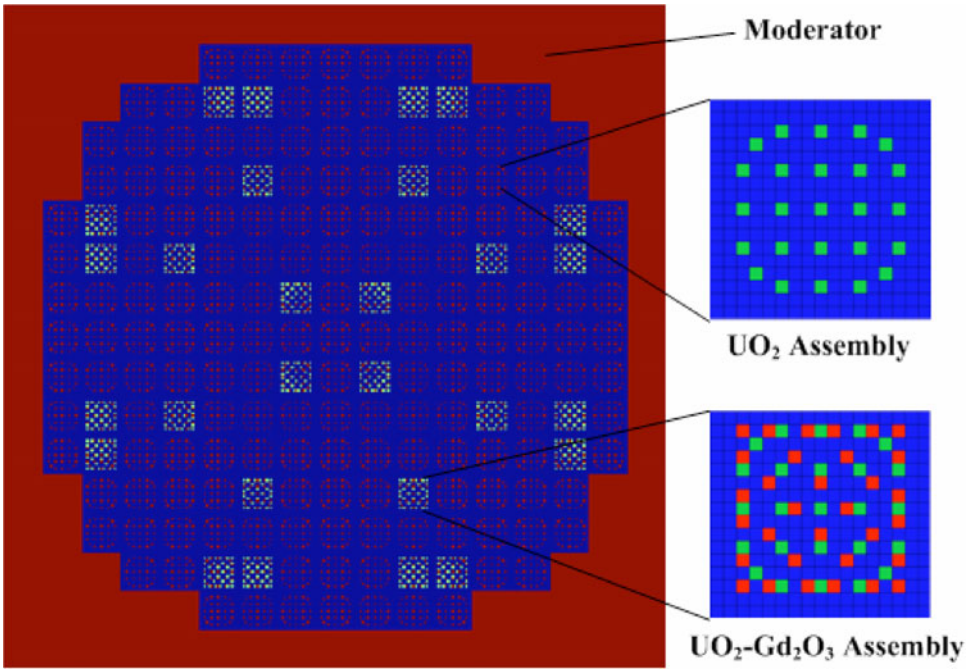

Fig. 5. PWR core model.

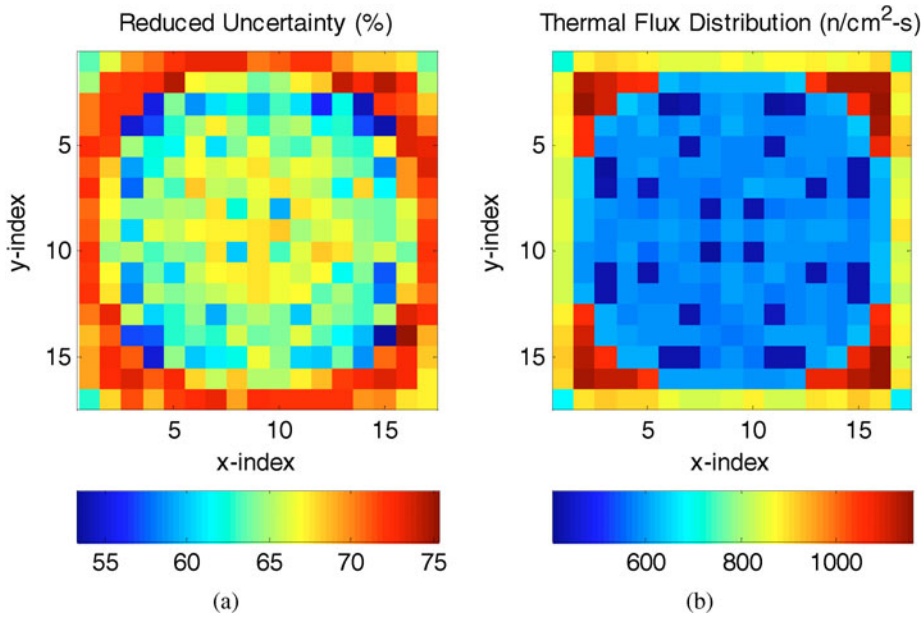

Fig. 6. Variance reduction results for prototypical PWR core model; (a) standard deviation reduction and (b) mean thermal flux distribution.

V. CONCLUSION

A new variant of a hybrid Monte Carlo–deterministic method has been presented. In this new approach, the correlations between the various single-response adjoint-based weight-window maps are determined via a SUBSPACE approach. The correlations describe a set of pseudoresponses whose number is much smaller than the number of original responses. By biasing the Monte Carlo particles toward the pseudoresponses, noticeable computational savings could be achieved. The SUBSPACE approach has been compared to the

TABLE III

Performance Metrics for FW-CADIS and SUBSPACE Approaches in a Core Model

Approach	Relative Standard Variance		Number of Monte Carlo Particles	Deterministic Execution Time (s)	Monte Carlo Execution Time (min)	FOM
	AVG[a]	STD[b]				
FW-CADIS	0.904	0.133	1.0E+07[c]	1.85	74.23	0.0165
SUBSPACE	0.338	0.031	1.0E+07	9.59	97.45	0.0897

[a]AVG = average.
[b]STD = standard.
[c]Read as 1.0×10^7.

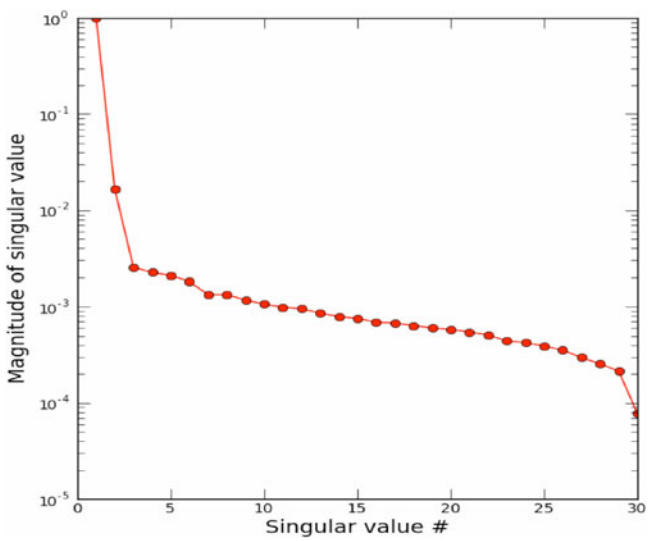

Fig. 7. Singular values of the SRI matrix.

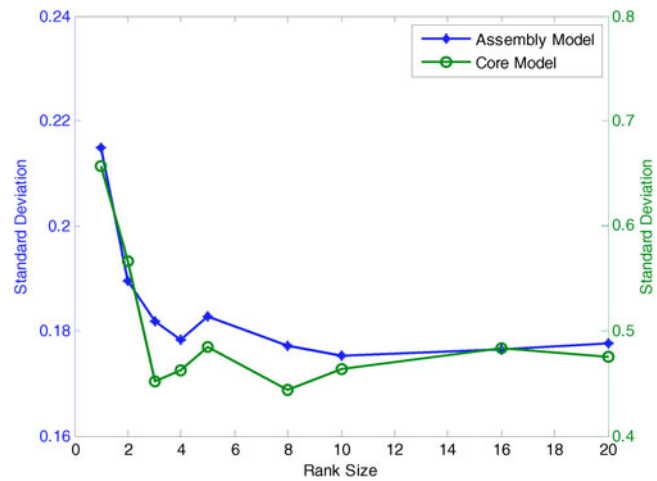

Fig. 8. Variance reduction sensitivity to rank estimate.

FW-CADIS approach implemented in the MAVRIC sequence of the SCALE code system, and the approach has been tested on a typical BWR assembly model and a prototypical PWR core model. The numerical experiments have demonstrated the mechanics of the proposed approach and showed a performance improvement over the FW-CADIS methodology as measured by the FOM. Moreover, the implementation of the SUBSPACE approach is straightforward. It requires executing the FW-CADIS methodology a number of times (the number is equal to the estimated rank of the SRI matrix) with the forward flux weighting bypassed. It was also noticed that with more complicated geometries, one can see more correlations between the weight windows, and hence, more computational savings could be achieved. Future work will focus on applying the SUBSPACE approach to an eigenvalue problem. Moreover, the computational savings can render possible the generation of sensitivity information of responses with respect to cross sections and the propagation of cross-section uncertainties through Monte Carlo–based models.

APPENDIX

Let $\mathbf{\Psi} \in \mathbb{R}^{I \times J}$ represent a matrix as defined in the paper. Let ε be a user-defined error tolerance. The matrix $\mathbf{\Psi}$ could be decomposed into two matrices, i.e., $\mathbf{\Psi} = \mathbf{\Psi}_r + \mathbf{\Psi}_{I-r}$, where $\mathbf{\Psi}_r = \mathbf{\Psi}\mathbf{Q}\mathbf{Q}^T$ has rank r and $\mathbf{\Psi}_{I-r} = \mathbf{\Psi}(\mathbf{I} - \mathbf{Q}\mathbf{Q}^T)$ has rank $I - r$ such that $\mathbf{Q} \in \mathbb{R}^{J \times r}$ is a matrix with orthonormal columns. With special requirements on the choice of the matrix $\mathbf{Q}$, one can use $\mathbf{\Psi}_r$ to approximate $\mathbf{\Psi}$ by upper-bounding the error resulting from $\mathbf{\Psi}_{I-r}$. In particular, one can prove with high probability that

$$\|\mathbf{\Psi}(\mathbf{I} - \mathbf{Q}\mathbf{Q}^T)\| \leq \varepsilon \ .$$

The rank r would then be defined as the minimum integer that satisfies the above criterion for a given user-defined tolerance. In most engineering problems, the tolerance could be selected to match the precision of the calculations. In deterministic calculations, the tolerance could be matched to the truncation errors induced by the numerical scheme employed. In probabilistic calculations, a much higher tolerance should be employed since in

most practical situations, the statistical uncertainties rendered by the models are much higher than the truncation errors. Typical examples are $\varepsilon = 10^{-8}$ for deterministic calculations and $\varepsilon = 10^{-3}$ for probabilistic calculations.

The following algorithm could be employed to determine the rank[12]:

1. Pick a small integer s; e.g., $s = 10$ is suitable for most practical calculations.

2. Generate s random vectors $\{\vec{\mu}_j\}_{j=1}^{s}$.

3. Calculate: $\{\vec{w}_j = \mathbf{\Psi}^T \vec{\mu}_j\}_{j=1}^{s}$.

4. Given a user-defined tolerance ε.

5. Given an estimate of the rank r_0, generate r_0 random vectors $\{\vec{\eta}_j\}_{j=1}^{r_0}$.

6. Calculate: $\{\vec{z}_j = \mathbf{\Psi}^T \vec{\eta}_j\}_{j=1}^{r_0}$.

7. Form an orthonormal matrix $\mathbf{Q} \in \mathbb{R}^{J \times r_0}$ such that $\mathrm{R}(\mathbf{Q}) = \mathrm{span}\{\vec{z}_1, \ldots, \vec{z}_{r_0}\}$. This could be done via a Gram-Schmidt orthogonalization procedure.

8. Calculate: $\varepsilon_1 = \max_{j=1,\ldots,s} \{\theta_j = \|(\mathbf{I} - \mathbf{Q}\mathbf{Q}^T)\vec{w}_j\|\}_{j=1}^{s}$.

9. If $10\sqrt{2/\pi}\,\varepsilon_1 > \varepsilon$, then the rank r_0 does not satisfy the tolerance ε; increase r_0 and return to step 6 until the exact rank is identified r_{ex}.

Note that the only requirement for this algorithm is the evaluation of the matrix-vector product, which requires the execution of the adjoint model. This algorithm requires r_{ex} adjoint model evaluations. Note that all steps required to identify the rank are simple vector manipulations, which could be done in a script outside the code; the only requirement is the access to the importance map calculated by the adjoint model.

ACKNOWLEDGMENTS

This work has been supported by a Department of Energy Nuclear Energy University Program grant titled "Development of Subspace-Based Hybrid Monte Carlo-Deterministic Algorithms for Reactor Physics Calculations."

The authors would like to thank A. Sood of Los Alamos National Laboratory for discussions and providing feedback on the proposed approach. The authors would also like to thank J. Wagner for hosting the second author as an intern at the Oak Ridge National Laboratory during the summer of 2011.

REFERENCES

1. A. HAGHIGHAT and J. C. WAGNER, "Monte Carlo Variance Reduction with Deterministic Importance Functions," *Prog. Nucl. Energy*, **42**, 1 (2003).

2. T. L. BECKER, A. B. WOLLABER, and E. W. LARSEN, "A Hybrid Monte Carlo–Deterministic Method for Global Particle Transport Calculations," *Nucl. Sci. Eng.*, **155**, 155 (2007).

3. J. D. DENSMORE and E. W. LARSEN, "Variational Variance Reduction for Particle Transport Eigenvalue Calculations Using Monte Carlo Adjoint Simulation," *J. Comput. Phys.*, **192**, 2 (2003).

4. J. C. WAGNER, E. D. BLAKEMAN, and D. E. PEPLOW, "Forward-Weighted CADIS Method for Global Variance Reduction," *Trans. Am. Nucl. Soc.*, **97**, 630 (2007).

5. C. J. SOLOMON, A. SOOD, and T. E. BOOTH, "A Weighted Adjoint Source for Weight-Window Generation by Means of a Linear Tally Combination," *Proc. 2009 Int. Conf. Mathematics, Computational Methods & Reactor Physics (M&C 2009)*, Saratoga Springs, New York, May 3–7, 2009, American Nuclear Society (2009) (CD-ROM).

6. M.-J. LEE, H. G. JOO, D. LEE, and K. SMITH, "A Feasibility Study of CMFD Acceleration in Monte Carlo Eigenvalue Calculation," *Trans. Korean Nuclear Society Autumn Meeting* (2009) (CD-ROM).

7. D. E. PEPLOW, "MAVRIC: MONACO with Automated Variance Reduction Using Importance Calculations," ORNL/TM-2005/39, Oak Ridge National Laboratory (Jan. 2009).

8. "SCALE: A Modular Code System for Performing Standardized Computer Analyses for Licensing Evaluations," ORNL/TM-2005/39, Version 6, Vols. I–III (2009).

9. H. S. ABDEL-KHALIK, "Adaptive Core Simulation," PhD Dissertation, Department of Nuclear Engineering, North Carolina State University (Dec. 2004).

10. M. A. JESSEE, H. S. ABDEL-KHALIK, and P. J. TURINSKY, "Cross-Section Uncertainty Propagation and Adjustment Algorithms for BWR Core Simulation," *Proc. Advances in Nuclear Fuel Management IV*, Hilton Head Island, North Carolina, April 12–15, 2009, American Nuclear Society (2009) (CD-ROM).

11. G. H. GOLUB and C. F. VAN LOAN, *Matrix Computations*, 3rd ed., John Hopkins (1996).

12. N. HALKO, P. G. MARTINSSON, and J. TROPP, "Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions," *SIAM Rev.*, **53**, 2 (2011).

13. A. PAPOULIS and S. PILLAI, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill (2002).

14. A. YAMAMOTO, T. IKEHARA, T. ITO, and E. SAJI, "Benchmark Problem Suite for Reactor Physics Study of LWR Next Generation Fuels," *J. Nucl. Sci. Technol.*, **39**, 8 (2002).

15. S. DOUGLASS, F. RAHNEMA, and J. MARGULIES, "A Stylized Three Dimensional PWR Whole-Core Benchmark Problem with Gadolinium," *Ann. Nucl. Energy*, **37**, 10 (2010).

16. J. HITE, Y. BANG, C. WANG, and H. S. ABDEL-KHALIK, "Heuristic Approach for ESM-based Reduced Order Modeling," *Trans. Am. Nucl. Soc.*, **105**, 475 (2011).