ABSTRACT

QIONG, ZHANG. Development of SUBSPACE-Based Hybrid Monte Carlo-Deterministic Algorithms for Reactor Physics Calculations. (Under the direction of Hany Abdel-Khalik.)

This dissertation develops an innovative hybrid Monte-Carlo-Deterministic (MC-DT) method which places high premium on attaining high computational efficiency for reactor analysis applications. Over the past few decades, there have been a plethora of techniques proposed to enable the hybridization of MC and DT methods with great success primarily for shielding applications where one is often interested in estimating the flux at few given points. The basic idea is to employ a simplified deterministic model to get an estimate of the flux solution, which is subsequently employed to bias the MC particles. In one implementation, adjoint deterministic calculations are employed to set weight-windows to accelerate convergence of MC simulation. Some progress has been made recently for reactor analysis applications where one is interested in calculating the flux distribution everywhere in the reactor core, which is much more computationally demanding than shielding applications because of the huge increase in the number of responses required. We believe the efficiency of these methods however is still too low to enable using MC methods in routine analysis calculations where typically one needs to execute the flux solver in the order of 10^3-10^5 times. To be acceptable to nuclear practitioners, e.g. fuel vendors and utilities, the efficiency of hybrid MC-DT needs to method that of existing deterministic methods used for routine design calculations. This dissertation contributes a new hybrid method denoted hereinafter by the SUBSPACE method, which primarily focuses on improving the efficiency of hybrid methods for reactor analysis applications, whereby highly accurate estimates of the energy-dependent flux are required everywhere in the reactor core, including a detailed pin power

distribution for each fuel assembly. The SUBSPACE method achieves its higher computational efficiency by taking advantage of the correlations between the responses. These correlations are introduced by the physics of radiation transport. Research over the past ten years has shown that the effective degrees of freedom in reactor analysis problems are very few despite the high dimensionality of the associated models. The SUBSPACE method takes advantage of this situation by identifying a small number of degrees of freedom towards which the MC particles are biased in a similar manner to existing hybrid methods. Significant gains in computational efficiency have been demonstrated using this method. The dissertation derives the mathematical theory behind the SUBSPACE method and applies it to realistic reactor analysis models. Two different implementations of the SUBSPACE method are presented, the first one described above relies on an adjoint deterministic model to calculate weight-windows for MC particles biasing. The second one is referred to the Gaussian Process (GP) method. The reason for this name is that the responses correlations are captured based on the assumption that the responses can be treated as Gaussian processes, which is a reasonable assumption for radiation transport. The applicability of the SUBSPACE method is also demonstrated for different types of models, including k-eigenvalue core-wide models, assembly models used for cross-sections homogenization for subsequent core-wide calculations, and depletion calculations. Given the favorable results obtained here, we believe the applicability of the MC method for large scale reactor analysis could be realized over the near future.

Development of SUBSPACE-Based Hybrid Monte Carlo-Deterministic Algorithms for Reactor Physics Calculations

by
Qiong Zhang

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Nuclear Engineering

Raleigh, North Carolina

2012

APPROVED BY:

_____	_____
Dr. Hany Abdel-Khalik Committee Chair	Dr. Yousry Azmy
_____	_____
Dr. John Mattingly	Dr. Robert White

BIOGRAPHY

Qiong Zhang was born on December 18th, 1986 in a small town in Nei Mongol, China. She received her Bachelor of Science in Nuclear Engineering from Tsinghua University in China in July 2007. After graduation, she worked as a certified nuclear engineer in China Nuclear Power Engineering Co. until July 2009. Her work responsibilities focused on designing nuclear equipment for power plants in China. She also participated in designing and modeling the EPR coolant system in cooperation. In August 2009, she joined the Department of Nuclear Engineering, North Carolina State University as a research assistant. Since then, she has been working on the project proposed in this dissertation. In June 2011, she completed an 8-week internship in Oak Ridge National Laboratory working on the development of reliable performance metrics for Monte Carlo mesh tallies.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my advisor, Prof. Abdel-Khalik, for his support, guidance, and encouragement throughout the course of this work. I greatly appreciate his advice at every stage of this work and thank him for his invaluable guidance in both my academic and professional development. I am truly grateful to my committee members: Prof. Azmy, Prof. Mattingly, Prof. White, for their effort and time on my dissertation work.

I would like to thank Dr. John Wagner for many insightful and thought provoking discussions during my internship at ORNL and for providing the KADAVR sequence. In addition, I would like to acknowledge the computer resources provided by NCSU for Academic Computing.

Finally, I am deeply grateful to my parents for their support, encouragement, and faith in my abilities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The computer simulation of radiation transport within matter is essential to many nuclear engineering applications such as reactor physics design calculations, radiation detection modeling and shielding design. Fundamentally, there are two distinctive methods for modeling radiation transport: the Monte Carlo (MC) and the deterministic (DT) methods. While the two methods are both used to simulate the same type of problems, each method has its unique advantages and limitations.

The MC is well known for its ability of modeling basic physical phenomena of radiation interactions by making virtually no simplifying modeling assumptions and is considered the most accurate method for solving radiation transport problems. However, it is computationally very expensive and as of now cannot be applied effectively on a routine basis to complete reactor design calculations, where a significant number of executions of the computational models are often required. Meanwhile, the DT is superior in its computational efficiency enabled by various simplifying assumptions about the physics. The limitation of this method is that it is difficult to validate the quality of the assumptions made, and the results are therefore often benchmarked against MC calculations, considered to be the gold standard for radiation transport. Another limitation of DT methods is that the computational cost often goes super linearly with the number of cells in the numerical scheme employed. For this reason, it is not possible to use one detailed deterministic model to analyze the entire

reactor core even with leadership computers. Reactor physicists have therefore resorted to homogenization techniques (described below) to reduce the computational cost. These methods make additional assumptions which are typically refined over many years of experience.

The current commercial reactor design process exclusively utilizes deterministic radiation transport models in conjunction with homogenization techniques to render computationally efficient simulation of the core-wide behavior over the cycles of operation. The reactor design process is typically divided into three stages trading off details in the resolution of energy and space:

1. Infinite medium or 1D transport calculation with high energy resolution (either point-wise of many group representation, e.g. 100s-10000s groups). Cross section resonances are explicitly resolved and appropriately self-shielded. The result of this stage is a self-shielded cross section set comprised of 10s~100s groups.
2. The lattice or assembly level calculation solves the 2D transport equation on a 2D slice of a single assembly using reflective boundary conditions on all four edges which implies the reactor is infinitely loaded with identical assemblies. The flux solution is typically obtained using the Method of Collision Probabilities, the Method of Characteristic or unstructured S_N transport solvers.
3. The few-group cross section set is used in a core-wide, 3D, coarse mesh diffusion calculation.

The most attractive feature of MC is its ability to utilize continuous (also referred to as point-wise) energy-dependent cross sections; thereby removing the errors introduced by the multi-group method. In addition to that, MC does not suffer from truncation error present in deterministic models. Moreover, no assumptions about the boundary conditions of each assembly are made. The MC results however are statistical quantities, where the statistical fluctuations are described by a standard deviation. For a reliable MC simulation, the standard deviation on all quantities of interest must be made very small. Due to the central limit theorem, the standard deviation decreases as:

$$\sigma \propto \frac{1}{\sqrt{N}}, \tag{1}$$

where σ is the standard deviation and N is the number of executed histories. For example, a factor of 10 reduction in standard deviation requires a 100 times more particles to simulate, which increases the computational time/cost by the same amount. The standard analog MC simulation is therefore considered impractical without applying powerful variance reduction techniques. Variance reduction techniques minimize statistical errors e.g. by biasing the MC particles towards more important regions in phase space (the basis for importance sampling) or by eliminating the fluctuation in the particles weight.

Since the advent of the MC method, variance reduction techniques have been proposed and employed to reduce the statistical uncertainty in the response of interest with a fixed number of history.

The most widely used variance reduction techniques are Russian Roulette, implicit capture and exponential transformation. For its first couple of decades of application, typical MC problems comprised only a small number of responses such that variance reduction parameters that are instrumental for the above variance reduction methods could be easily devised by hand. However, one typical feature of reactor physics application is that responses are sought everywhere in phase-space and therefore the variance reduction must also be rendered everywhere, e.g., thermal and fast flux everywhere in the core and throughout cycles of depletion; this challenge is known as Global Variance Reduction (GVR).

Recognizing the advantages and deficiencies of both methods, hybrid MC-DT techniques have been proposed by researchers to speed up MC simulation. The essential idea of any hybrid technique is to obtain an inexpensive deterministic estimate of the adjoint or forward flux; then use the estimate to generate variance reduction parameters for GVR. Two typical examples of hybrid MC-DT techniques for GVR are: importance sampling and weight windows. Importance sampling concentrates on the regions in phase space that are most important for the response of interest; weight window controls the particle weight by rouletting and splitting based on the weight function that is derived from the deterministic estimate of the forward or adjoint flux. Common to all hybrid MC-DT methods, a deterministic estimate of the importance function is employed; the difference of hybrid MC-

DT methods manifests in what type of transport problem is solved (forward, adjoint, etc). Hybrid methods have been proven highly effective in speeding up MC computations; in

general all hybrid deterministic MC methods address one or several of the following three problems that cause poor MC performance in GVR:

1. Too few particles contribute to the response of interest leading to large relative variance;
2. The magnitude of particle weights that contribute to the response of interest fluctuate;
3. Particles with small weights are retained until collision and cause a waste of execution time.

To overcome the above challenges of current hybrid methods, we present in this work a novel hybrid MC-DT method denoted by the SUBSPACE method. The method is based on the adjoint function and therefore belongs to the same family as the FW-CADIS method developed by the SCALE group at ORNL [33]. For this reason, we employ the FW-CADIS method as a basis for comparison. Similar to FW-CADIS, the SUBSPACE method employs importance maps obtained from deterministic adjoint models to derive automatic weight-window biasing. In contrast to FW-CADIS, the SUBSPACE method does not calculate flux-based weighting of the adjoint source term.

Instead, it capitalizes on correlations between the responses of interest which are typically present in reactor physics models and is highly effective in reducing the execution time required for GVR. An extension of the SUBSPACE method, denoted by Gaussian Process (GP), is also presented in this work, which only requires a forward deterministic model when adjoint models are not available.

This dissertation is organized as follows: In Chapter 2, fundamental deterministic (forward and adjoint) and MC theories are reviewed to provide a basic theoretical background on DT and MC method; different variance reduction techniques are also included in this Chapter. A brief introduction on MC-DT hybrid methods is given and the literatures on various hybrid methods developed in the last couple of decades are summarized and reviewed. Chapter 3 presents the motivation, development and mathematical description of two proposed hybrid methods, that is the SUBSPACE and GP methods. In Chapter 4, the performance of proposed hybrid methods is demonstrated through a series of numerical experiments at both the assembly and core level. Comparisons between the performance of the SUBSPACE, GP, and FW-CADIS methods are well presented. Furthermore, the applicability of the SUBSPACE method in solving k-eigenvalue problems is shown for core-level problems. Chapter 5 focuses on the actual applications of the SUBSPACE method in reactor physics: cross section functionalization at assembly level and depletion. Conclusions for this work and suggestions for future research are given in Chapter 7.

CHAPTER 2

THEORY

2.1 Deterministic Theory

2.1.1 Forward Deterministic Theory

In the context of reactor analysis applications, the deterministic transport theory is concerned with the solution of the integro-differential form of the forward and adjoint linear Boltzmann transport equation that are originally derived by Ludwig Boltzmann to study the kinetics of gases[50]; it is hereinafter simply referred to as the transport equation. The two forms of the forward transport equation used within this work model the neutron flux in a host medium (1) for the case that the flux is maintained by an external distributed source (fixed source form) and (2) for the case that fission in the host medium maintains the flux without an external source (k-eigenvalue form). In its operator form, the fixed source form of the transport equation is given by the following expression [3]:

$$\underline{L}\psi\left(\vec{r},\hat{\Omega},E\right)=\underline{S}\psi\left(\vec{r},\hat{\Omega},E\right)+q(\vec{r}) \tag{2}$$

In Eq. (2), the dependent variable $\psi\left(\vec{r},\hat{\Omega},E\right)$ is the angular flux, i.e. the product of the neutron density and its speed, and $q(\vec{r})$ is the distributed source. $\psi\left(\vec{r},\hat{\Omega},E\right)$ depends on the six independent phase space variables: space $\vec{r}=(x,y,z)$, direction of movement of the neutrons and neutron energy E.

The operators $\underline{L}$ and $\underline{S}$ are the streaming-plus-collision and the scattering operator given by [3] respectively:

$$\begin{aligned} \underline{L}\psi &= \hat{\Omega}\cdot\nabla\psi + \Sigma_t\left(\vec{r},E\right)\psi\left(\vec{r},\hat{\Omega},E\right) \\ \underline{S}\psi &= \int_0^\infty dE' \int_{4\pi} d\hat{\Omega}'\Sigma_s\left(\vec{r},\hat{\Omega}',E'\to\hat{\Omega},E\right)\psi\left(\vec{r},\hat{\Omega}',E'\right) \end{aligned} \tag{3}$$

Instrumental in the definition of these two operators are the macroscopic total and double-differential scattering cross sections $\Sigma_t\left(\vec{r},E\right)$ and $\Sigma_s\left(\vec{r},\hat{\Omega}',E'\to\hat{\Omega},E\right)$, respectively. It is sufficient to say here that the macroscopic cross sections are input parameters characterizing the interaction probabilities of the neutrons with the host-medium; we refer to [3] for a comprehensive definition of macroscopic cross sections. The second form of the forward transport equation used within this work is the k-eigenvalue form:

$$\underline{L}\psi = \underline{S}\psi + \frac{1}{k}\underline{F}\psi \ , \tag{4}$$

where the fission operator $\underline{F}$ is defined as:

$$\underline{F}\psi = \frac{\chi(E)}{4\pi}\int_0^\infty dE' \int_{4\pi} d\hat{\Omega}'\nu\left(\vec{r},E'\right)\Sigma_f\left(\vec{r},\hat{\Omega}',E'\right)\psi\left(\vec{r},\hat{\Omega}',E'\right). \tag{5}$$

Describing the fission process necessitates introducing additional material properties: The fission cross section $\Sigma_f\left(\vec{r},E\right)$, the fission yield ν,

i.e. the average number of neutrons created per fission event, and the fission spectrum $\chi(E)$ which describes the initial energy distribution of neutrons born in a fission event. Finally the

eigenvalue or multiplication factor k is the ratio of the total neutrons gains in the system and the total neutron losses:

$$k = \frac{\text{Total neutron gains}}{\text{Total neutron losses}} \tag{6}$$

Depending on the value of the multiplication factor the system is classified as subcritical (k<1), critical (k=1) or supercritical (k>1).

In order to create a well-posed set of equations, boundary conditions have to be specified on the inflow faces:

$$\psi\left(\vec{r},\hat{\Omega},E\right) = \psi_{BC}\left(\vec{r},\hat{\Omega},E\right), \vec{r} \in \vec{r}_B \text{ and } \hat{\mathrm{n}}\cdot\hat{\Omega} < 0. \tag{7}$$

where ψ_B is a given function on the boundary and $\hat{n}$ is the outward normal vector on the domain boundary. Eq. (7) strictly only covers explicit boundary conditions; while in this work we frequently utilize reflective boundary conditions, where the inflow flux depends on the outflow flux in the "reflected" direction. Reflective boundary conditions can be described by the following equation:

$$\begin{aligned} &\psi\left(\vec{r},\hat{\Omega},E\right) = \psi\left(\vec{r},\hat{\Omega}',E\right), \vec{r} \in \vec{r}_B \text{ and } \hat{\mathrm{n}}\cdot\hat{\Omega} < 0, \\ &\hat{\mathrm{n}}\cdot\hat{\Omega} = -\hat{\mathrm{n}}\cdot\hat{\Omega}', \left(\hat{\Omega}'\times\hat{\Omega}\right)\cdot\hat{\mathrm{n}} = 0 \end{aligned} \tag{8}$$

2.1.2 Adjoint Deterministic Theory

For the derivation of the adjoint transport equation it is necessary to introduce the notion of an inner product over the phase space since the adjoint operators will be defined with respect to this inner product. Let the inner product of the two generic functions f and g be given by

$$\langle f, g \rangle = \int_0^\infty dE \int_{4\pi} d\hat{\Omega} \int_D dV \, f g, \tag{9}$$

where D denotes the spatial domain of the problem. Given a generic operator H then its adjoint is $H^\dagger$ defined by the following relationship:

$$\left\langle \psi^\dagger, H\psi \right\rangle = \left\langle H^\dagger \psi^\dagger, \psi \right\rangle, \tag{10}$$

where $\psi^\dagger$ denotes the adjoint function. The fixed source adjoint transport equation can then be formulated in terms of the adjoint counterparts of the transport-and-collision and scattering operators in a manner very similar to the forward transport equation:

$$\underline{L}^\dagger \psi^\dagger = \underline{S}^\dagger \psi^\dagger + q^\dagger(\vec{r}), \tag{11}$$

where the following definitions are used:

$$\underline{L}^\dagger \psi^\dagger = -\hat{\Omega} \cdot \nabla \psi + \Sigma_t \left(\vec{r}, E\right) \psi \left(\vec{r}, \hat{\Omega}, E\right) \tag{12}$$

$$\underline{S}^\dagger \psi^\dagger = \int_0^\infty dE' \int_{4\pi} d\hat{\Omega}' \Sigma_s \left(\vec{r}, \hat{\Omega}, E \to \hat{\Omega}', E'\right) \psi^\dagger \left(\vec{r}, \hat{\Omega}', E'\right). \tag{13}$$

The adjoint source is chosen for the specific purpose that the solution of the adjoint problem has for the user. In general if the user is interested in some system response (e.g. a reaction rate)

$$R = \left\langle \Sigma_R, \psi \right\rangle, \tag{14}$$

then the adjoint source is chosen to the gradient of the response with respect to the angular flux:

$$q^{\dagger} = \frac{\partial R}{\partial \psi}. \tag{15}$$

In contrast to the forward problem, adjoint boundary conditions need to be given on outflow boundaries:

$$\psi^{\dagger}\left(\vec{r}, \hat{\Omega}, E\right) = \psi^{\dagger}_{BC}\left(\vec{r}, \hat{\Omega}, E\right), \vec{r} \in \vec{r}_B \text{ and } \hat{\mathrm{n}} \cdot \hat{\Omega} > 0. \tag{16}$$

2.1.3 Discretization of the Transport Equation

Neither the forward nor the adjoint transport equations can be solved analytically for realistic reactor physics applications. Therefore, numerical methods need to be devised such that the transport equation could be discretized in all phase space variables to facilitate its solution on a digital computer. The discretization of the energy variable is performed via the multi-group formalism; the discretization of the angular space, for the purpose of this work, uses the S_N method and the discretization of the spatial variables is typically done using the method of step characteristics [49].

The multi-group formalism decomposes the range of energy into G bins numbered from highest to lowest energy group by $g = 1,\ldots,G$ such that the g^{th} bin has lower and upper energy boundaries E_{g+1} and E_g, respectively. Then the transport equation is integrated over the range of energy bin g and rearranging then gives its multi-group form. Obtaining the multi-group transport equation is only demonstrated for the forward k-eigenvalue problem since the multi-group forms of the forward and adjoint fixed source equations follow readily from the following development. Consider the multi-group k-eigenvalue equations in operator form:

$$\underline{L}_g \psi_g = \underline{S}_g \psi_g + \frac{1}{k} \underline{F}_g \psi_g, \text{for } g = 1,\ldots,G. \tag{17}$$

The group streaming-and-collision, scattering and fission operators, group fluxes and multi-group constants are going to be introduced after the following comments outlining some key features of the derivation of the multi-group transport equations:

1. The angular dependence of the scattering cross section is assumed to be azimuthally isotropic such that its functional dependence can be written in terms of the cosine in between the incoming and outgoing directions. Then the scattering cross sections is expanded into a truncated series of Legendre polynomials:

$$\Sigma_s\left(\vec{r}, E', \hat{\Omega}' \to E, \hat{\Omega}\right) = \Sigma_s\left(\vec{r}, E' \to E, \mu_0\right) \approx \sum_{l=0}^{L} \frac{2l+1}{4\pi} \Sigma_{sl}\left(\vec{r}, E' \to E\right) P_l\left(\mu_0\right). \tag{18}$$

   The order of this expansion is referred as the scattering order.

2. The straight-forward way of defining the group total collision cross sections would render it angularly dependent, a feature that is typically not supported by transport solvers. Therefore, some care is taken to define an isotropic group collision cross section.
3. Easier derivations assume energy separability which in practical applications is never satisfied. While the final form of the equations is much simpler its range of applicability is questionable.

Armed with the comments 1 through 3, the definitions pertinent for Eq. (18) of the operators and group angular and scalar fluxes is:

$$
\begin{aligned}
&\underline{L}_g\psi_g = \hat{\Omega}\cdot\nabla\psi_g + \Sigma_{t,g}\left(\vec{r}\right)\psi_g\left(\vec{r},\hat{\Omega}\right)\\
&\underline{S}_g\psi_g = \sum_{l=0}^{L}\sum_{m=0}^{M}\left(2-\delta_{m0}\right)Y_{lm}\left(\hat{\Omega}\right)\frac{(l-m)!}{(l+m)!}\frac{2l+1}{4\pi}\sum_{g'=1}^{G}\tilde{\Sigma}_{s\,lg'g}\phi_{lg'}^{m}\left(\vec{r}\right)\\
&\underline{F}_g\psi_g = \frac{\chi_g}{4\pi}\sum_{g'=1}^{G}\left(\nu\Sigma_f\right)_{g'}\phi_{g'}\left(\vec{r}\right)\\
&Y_{lm}\left(\hat{\Omega}\right):\text{Spherical Harmonic of order } l,m\\
&\psi_g\left(\vec{r},\hat{\Omega}\right) = \int_{E_{g+1}}^{E_g} dE\psi\left(\vec{r},\hat{\Omega},E\right)\\
&\phi_{lg}^{m}\left(\vec{r}\right) = \int_{4\pi} d\hat{\Omega}Y_{lm}\left(\hat{\Omega}\right)\psi_g\left(\vec{r},\hat{\Omega}\right)\\
&\phi_g\left(\vec{r}\right) = \phi_{0g}^{0}\left(\vec{r}\right)\text{ (Scalar group flux)}
\end{aligned}
\tag{19}
$$

The group constants used within the above equation Eq. (19) are given by the following expressions:

$$\Sigma_{t,g} = \frac{\int_{E_{g+1}}^{E_g} dE\, \Sigma_t(\vec{r},E)\phi(\vec{r},E)}{\phi_g(\vec{r})}$$
$$\left(\nu\Sigma_f\right)_g = \frac{\int_{E_{g+1}}^{E_g} dE\, \nu(\vec{r},E)\Sigma_f(\vec{r},E)\phi(\vec{r},E)}{\phi_g(\vec{r})} \tag{20}$$
$$\chi_g = \int_{E_{g+1}}^{E_g} dE\, \chi(\vec{r},E)$$

$$\tilde{\Sigma}_{slg'g} = \Sigma_{slg'g} + \delta_{gg'}\left(\frac{4\pi}{2l+1}\frac{(l+m)!}{(l-m)!}\Sigma_{tg} - \Sigma_{tlg}\right)$$
$$\Sigma_{tlg} = \frac{\int_{E_{g+1}}^{E_g} dE\, \Sigma_t(\vec{r},E)\phi_l^m(\vec{r},E)}{\phi_{lg}^m(\vec{r})} \tag{21}$$
$$\Sigma_{slg'g} = \frac{\int_{E_{g+1}}^{E_g} dE \int_{E_{g'+1}}^{E_{g'}} dE'\, \Sigma_{sl}(\vec{r},E'\to E)\phi_l^m(\vec{r},E')}{\phi_{lg}^m(\vec{r})}$$

It should be noted that the multi-group equations, aside from the scattering approximation, are exact. However, in the process of the derivation group constants were introduced that depend on the problem's solution itself such that the equations cannot be solved without knowing the solution beforehand. The path usually taken in deterministic transport theory is to guess a spectrum, compute the group constants and then solve the multi-group equations. It is this step that typically introduces a great deal of approximation into the deterministic solution process. In addition, we point out that the MC method does not need to make this approximation because it naturally supports continuous/pointwise cross sections sets.

The next step in the discretization sequence is to discretize angular variables. Within this work exclusively the S_N method is utilized for this purpose. For this method the multi-group equations are solved only along discrete rays $\hat{\Omega}_n, n = 1,\ldots,N$ and the scalar flux moments are then calculated using a quadrature rule that associates a distinct weight w_n for each discrete ordinate. Integrations over the angular variable are then approximated as follows:

$$\int_{4\pi} d\hat{\Omega}\,\bullet \rightarrow \sum_{n=1}^{N} w_n \bullet \tag{22}$$

After the discretization of the angular variable we obtained a set of NxG coupled PDEs that continuously depend in the spatial variable. Postponing the treatment of the coupling for later, the spatial variable is discretized by subdividing the spatial domain into non-overlapping cuboidal cells and applying standard methods to discretize the equations on each of these mesh cells. See [49] for various methods; within this work the Step Characteristic method is typically used because it guarantees positive cell average angular fluxes which is important when S_N solutions are used in subsequent MC computations.

As the multi-group S_N equations are coupled and a direct solution of the global system of equations is impossible due to its size a multi-level solution algorithm is typically applied. The legacy method present in most transport solvers is the inner-outer iteration scheme where fission and up-scattering sources are updated during the outer iteration, down-scattering sources are updated during a sweep through all energy groups that is contained within a single outer iteration and self-scattering sources are updated in the inner iterations which are performed for each group during the sweep through all groups.

The DENOVO code that is used as DT solver uses more advanced techniques, namely it replaces the inner iteration procedure by a GMRES solver which greatly increases efficiency. For a comprehensive description of the solution methods in DENOVO refer to [49].

2.2 Monte Carlo Transport Theory

MC methods are widely applied in computational simulations and especially useful for simulating complicated systems with multiple degrees of freedom. MC methods are a class of stochastic numerical analysis techniques based on the application of random sampling; that is, to evaluate random variables by using random numbers. Random variables are measured through random experiments. In contract to deterministic transport methods that solve an explicit transport equation to obtain the average behavior of particles quantified e.g. by the scalar flux, MC methods infer specific quantities of interest, e.g. the reaction rate in a sub-volume, by explicitly simulating many particles' life and averaging over the outcome. Naturally, deterministic methods provide a full set of information for the problem such as fast and thermal fluxes throughout the system, while MC only gives certain information requested by the user, for example, one specific tally in the phase space.

At its core, the MC technique consists of simulating a single particles life from its birth to its destruction. This basic unit of a MC simulation is typically referred to as a history. In general many histories have to be performed to obtain statistically meaningful results.

The life of a neutron from its birth to its destruction is governed by randomness: The spatial location where it is born, the initial direction of movement and the energy are all random

variables. Further, the distance to the next collision site, the collision type and its outcome are random processes as well. From this perspective, simulating a single particle history can be performed using the following basic steps:

1. Through random sampling pick a specific position where a particle is born as well as a specific energy and a specific travel direction;
2. Determine the distance to collision. That is, the distance the particle travels before its first interaction.
3. Determine the type and consequences of the interaction. If the interaction does not lead to an event where the particle is killed; that is, (1) the particle is absorbed or (2) the particle exits the phase space of interest, new energy and direction will be picked through next random sampling and the particle transport will be continued. In particular, if secondary particles are created, for example, if an elastic scattering event takes place, the transport of the secondary particles will be simulated through the same procedure after the primary particle is terminated.
4. If the interaction leads to an event that the particle is terminated as stated above: particle is absorbed or exits the volume of interest; then all the contribution obtained throughout this particle history to the quantities of interest will be recorded and is referred to as the score.

At the same time, a new source particle will be selected through random sampling and the same procedure will repeat until the total number of particle histories is successfully simulated.

1. Neutron Scatter, photon production
2. Fission, photon production
3. Neutron Capture
4. Neutron Leakage
5. Photon Scatter
6. Photon Leakage
7. Photon Capture

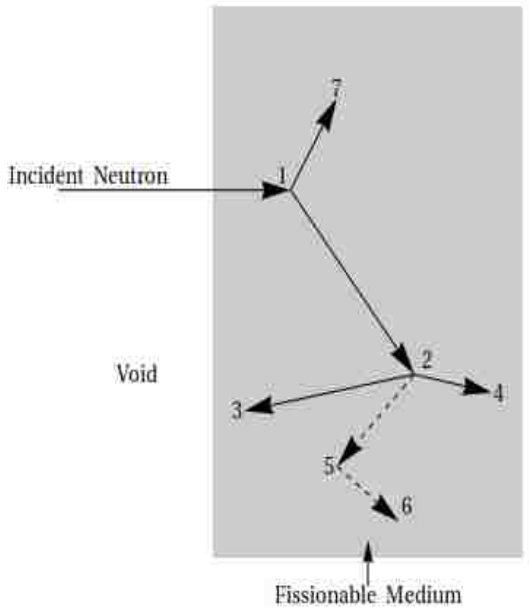


Fig. 2.1: Incident Neutron Events (referred from MCNP Manual)

Fig. 2.1 illustrates a possible sequence of events for a neutron incident on a slab of fissionable material. The first collision of the neutron within the slab takes place at event 1. The neutron is scattered after the collision in the direction towards 2 while a photon is simultaneously produced and stored for later discussion. At event 2, the incoming neutron is terminated by fission while two emerging neutrons and one photon are born.

The first of the two newly born fission neutrons is captured and terminated in event 3. The other fission neutron leaks out of the slab volume at event 4. The photon produced from the

fission collides at event 5 and leaks out at event 6. Back to the photon produced at event 1: it is now captured at event 7. So far the simulation history is completed for this incident neutron and the score has been obtained. The MC simulation proceeds by starting another incident neutron and repeats the basic steps that are comprised within a single history.

The transport simulation procedure described above is usually referred as an analog MC simulation. Meanwhile, non-analog MC refers to algorithms where distributions are sampled in a manner to increase the probability that a particle contributes to the quantity of interest and thus reduce the variance of the desired quantity or to control the particle's weight to ensure that scores do not vary too much in their individual contribution. Hereinafter, certain classes of very difficult problems that require days of computer time with analog MC simulation could be easily solved in hours or even minutes of computer time applying non-analog MC methods.

It is essential for non-analog MC that an unbiased estimate is preserved to ensure the correctness of the simulation. To this end the particle weight w, which indicates the total number of particles that are transported within a single history (think of a "particle package"), is adjusted. In non-analog MC simulations where variance reduction techniques are applied, the weight of the particle is adjusted following the conservation formula below:

$$w(\textit{biased pdf}) = w_0(\textit{unbiased pdf}) \quad , \tag{23}$$

where w_0 is the particle weight from analog MC simulation before the variance reduction techniques are applied. Equation (23) basically states that non-analog MC is considered to be

a correct transport simulation so far as the particle weights are conserved during the (biased) transport process.

2.2.1 Probability Theory and Statistical Uncertainties

The statistical uncertainty is a very important quantity in MC transport simulations. As the number of simulated histories is always finite, the estimates of the mean tally values are subject to statistical errors whose magnitude is quantified by the standard deviation. In addition the standard deviations can be used to infer if a tally is statistically well-behaved. If a tally is not statistically well behaved, the true confidence interval will not be reflected corrected by the associated uncertainty and thus the obtained results will be misleading. In general, for a well behaved tally, the statistical uncertainty will be proportional to $N^{-\frac{1}{2}}$, where N is the number of histories. For well-behaved tallies that means that increasing the number of histories decreases the standard deviation. However, for a poorly behaved tally, the statistical uncertainty may increase as the number of histories increases. By simulating particle histories, a range of scores $\{x_{1,\dots}x_{i,\dots}x_N\}$ will be generated to each particle history depending on the selected tally and applied variance reduction techniques. Define $f(x)$ as the probability density function of history score probability, the expected value of score x is then defined as true mean $\bar{x}$:

$$\bar{x} = \int x f(x) dx \tag{24}$$

The true underlying distribution function $f(x)$ is unknown and so is the true mean.

The MC simulation samples the underlying distribution function and obtains a range of scores $\{x_{1,\ldots}x_{i,\ldots}x_N\}$. The range of scores is then used to compute an estimate of the true mean: $\overline{x_e}$, hereinafter referred to as sample mean:

$$\overline{x_e} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{25}$$

where x_i is the score of ith particle history sampled from $f(x)$ as stated before, and N is the total number of particle histories. The sample mean $\overline{x_e}$ is the average value of $\{x_{1,\ldots}x_{i,\ldots}x_N\}$ for all particle histories in the MC problem. The relationship between the true mean $\bar{x}$ and the sample mean $\overline{x_e}$ is described in [4] by the Strong Law of Large Numbers: if $\bar{x}$ is finite, $\overline{x_e}$ tends to the limit of as the history N methods infinity.

Akin to the true mean, the true variance σ^2 (i.e: the square of the true standard deviation) can be computed from the underlying distribution function by taking the second central moment:

$$\sigma^2 = \int (x-\bar{x})^2 f(x)dx = \overline{x^2} - (\bar{x})^2 \quad . \tag{26}$$

The square root σ is defined as the standard deviation of the population of scores. In MC simulations, this quantity is usually estimated as the sample standard deviation s. The sample variance s^2 is denoted by:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x_e})^2 \tag{27}$$

The sample variance could also be calculated directly from probability density function $f(x)$ as shown in the [5] as

$$s^2 = \sum_{x=0}^{\infty} (x - \bar{x})^2 f(x) \approx \overline{x^2} - (\bar{x})^2 \tag{28}$$

The larger the data set, i.e. the more histories are run in an MC simulation, the more reliable are the sample mean and the sample variance, i.e. the more they are to be close to the true mean and variance, respectively. In general MC simulations, the sample mean $\overline{x_e}$ and the sample variance s^2 are two representative properties from the probability density function $f(x)$ that are of particular interest.

In the above discussion, it is implicitly assumed that the process of interest involves only a single random variable. In reality, processes involve a vector $\vec{x}$ of random variables which are distributed according to an underlying multi-varied probability distribution function $f(\vec{x})$.

In this case, the variance of the i[th] random variable is defined analogously to the single variable case:

$$\sigma_i^2 = \int (x_i - \bar{x}_i)^2 f(\vec{x}) d\vec{x} = \overline{x_i^2} - (\bar{x}_i)^2 \quad . \tag{29}$$

In addition to the single variable case, the interaction of the random variables is important for understanding the underlying process. Therefore, the concept of covariance is introduced to measure the interaction of two random variables:

$$\mathrm{cov}\left(x_i, x_j\right) = \int (x_i - \bar{x}_i)(x_j - \bar{x}_j) f(\vec{x}) d\vec{x}. \tag{30}$$

The variance is a special case of the covariance since $\mathrm{cov}(x_i, x_i) = \sigma_i^2$. For convenience the variances and covariances are typically collected in the covariance matrix $\underline{C}$:

$$\left(\underline{C}\right)_{i,j} = \mathrm{cov}\left(x_i, x_j\right). \tag{31}$$

The covariance matrix is a symmetric positive semi-definite matrix of dimension N [REF], where N is the length of the vector $\vec{x}$. It features the variances on the diagonal and the (true) covariances on the off-diagonals. In section 2.4, it will be shown that the covariances characterize the interdependence of the random variables (i.e. their correlation) such that the properties of the covariance matrix are at a high premium and matrix decomposition can provide valuable information of the system of interest.

The Central Limit Theorem of probability is a very importance theorem for the estimation of the reliability of statistical estimators [4]; it is used to define confidence intervals for the precision of obtained results:

$$\lim_{N\to\infty} \Pr[\bar{x} + \alpha\frac{\sigma}{\sqrt{N}} < \overline{x_e} < \bar{x} + \beta\frac{\sigma}{\sqrt{N}}] = \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-t^2/2}dt \tag{32}$$

Where α and β refer to arbitrary values and $\Pr$ stands for the probability. For large N approximation, it could be demonstrated [4] that

$$\Pr[\alpha < \frac{\overline{x_e} - \bar{x}}{\sigma\sqrt{N}} < \beta] \to \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-t^2/2}dt \tag{33}$$

From Central Limit Theorem of probability, it is observed that for large N approximation the sample mean $\overline{x_e}$ distribution could be considered as a normal distribution. This means that

the sample mean $\overline{x_e}$ distribution is normally distributed with the mean of true mean $\bar{x}$. Given that the sample variance s^2 is considered equivalent to the true variance σ^2 for a MC simulation with sufficient particle histories as discussed before, the following two statements on confidence intervals could be deducted from the standard table for the normal distribution:

$$\begin{aligned} &\overline{x_e} - s < \bar{x} < \overline{x_e} + s \quad \text{with 68\% confidence} \\ &\overline{x_e} - 2s < \bar{x} < \overline{x_e} + 2s \quad \text{with 95\% confidence} \end{aligned} \tag{34}$$

Within this work, it will be necessary to estimate the statistical uncertainty of derived important quantities that are not directly tallied. However, in general a functional form is known that describes the derived quantities in terms of directly tallied quantities.

Then the law of Error Propagation [5] could be used to estimate the statistical uncertainty associated with the derived quantities:

$$\begin{aligned} &f\left(x_1, x_2, \ldots, x_N\right): \text{known functional form} \\ &\sigma_f{}^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_i^2 \end{aligned} \tag{35}$$

where the functional form $f\left(x_1, x_2, \ldots, x_N\right)$ represents the derived quantity. The Error Propagation law is applicable to almost all circumstances in measurements.

It should be noted that the variables $x_1, x_2, \ldots, x_N$ are required to be truly independent to avoid the correlation effects.

As previously discussed, the estimated statistical uncertainty squared σ^2 should be proportional to $1/N$. Furthermore, the execution time T of an MC computation is asymptotically proportional to the number of simulated histories N , because each history basically follows the same algorithm. Therefore, the product of variance σ^2 and execution time T is expected to be asymptotically constant for every estimated quantity in MC simulation. The inverse of the product $\sigma^2 T$ is defined as Figure of Merit (FOM) of a single tally [6]:

$$FOM = \frac{1}{\sigma^2 T} \tag{36}$$

The single-tally FOM is a very important and useful statistic parameter to assess the statistical behavior of a tally bin in MC simulation. It is a tally reliability indicator to tell how well the tally behaves. The single-tally FOM is usually a fairly approximate constant as a function of the number of histories for each tally with the possible exception of early statistical fluctuations. Meanwhile, the single-tally FOM is also employed to compare MC simulation efficiencies and optimize variance reduction parameters. This is achieved by comparing the FOMs of several MC simulations with various variance reduction parameters and picking the parameters that associate with the highest FOM. However, for large and complicated systems, it will take multiple processors-days to perform such analysis even for short MC simulations.

The global FOM is thus introduced for comparing different methods to obtain effective variance reduction parameters [**7**]. As a direct extension of the single-tally FOM, a global FOM is defined for a set of tallies by replacing the single variance by a value representative of the distribution of variances. The single value could be the maximum variance or the mean variance of the tally set. In general, it is desirable to have a small spread of tally variances because the largest variance usually limits the confidence in the simulation's results. In [**7**], four different types of global FOMs are introduced:

$$FOM_1 = \frac{1}{\sigma_{\max}^2 T} \tag{37}$$

$$FOM_2 = \frac{1}{\bar{v} T} \tag{38}$$

$$FOM_3 = \frac{1}{(\bar{v} + \sigma_v) T} \tag{39}$$

$$FOM_4 = \frac{1}{[\bar{v} + (\kappa_v / 3)^{1/4} \sigma_v] T} \tag{40}$$

In above equations, $\sigma_{\max}$ stands for the maximum relative uncertainty in one single tally throughout the domain; T stands for the time; $\bar{v}$ stands for the mean of the variance, σ_v for the relative uncertainty of the variance and κ_v stands for Kurtosis, the peak factor.

The advantage for these global FOMs is that the added parameters are time-independent for large N thus preserving the asymptotic behavior for large N. These global FOMs could be employed to assess the efficiency of differing weight-window mesh resolutions as well as the employment of different variance reduction techniques in conjunction with global weight-

window maps [**7**]. In this work, the second form of global FOM is utilized to compare the efficiency of hybrid MC-DT methods in GVR.

2.2.2 Variance Reduction Techniques

The analog MC method mimics the physical process of particle transport: One could say that nature actually plays unbiased MC. However in practice, calculations are expected to produce accurate answers with reasonable efficiency. In MC simulations, variance reduction techniques, also named biasing techniques, are widely employed to decrease the statistical uncertainty σ of the desired results while trying to keep the computational time within a reasonable range and more importantly keep the estimates of the mean value unbiased. Two possible solutions to achieve the goal are: 1. Reduce the standard deviation s for a fixed number of histories; 2. Increase the total particle history N given a fixed amount of time. However, in the most recent history of MC, there has always been a trade-off between these two solutions for the reason that the decrease of s usually causes a corresponding increase of computation time per history; while the increase of total particle history N would lead to a higher standard deviation s [6]. Therefore, variance reduction techniques are applied to achieve a compromise to decrease s without increasing the total history N significantly. Generally, these techniques could be divided into two categories: 1. Techniques that produce particles work by decreasing s; 2. Techniques that destroy particles work by increasing the particle history N [6] [8].

Over the course of last few decades, many variance reduction techniques have been developed. It should be noted though that one technique cannot serve for all application purposes, and different techniques might also contradict with each other if used together. Below is a summary of several of the most widely applied variance reduction techniques:

Source Biasing

In many problems, a large amount of source particles are distributed over space, energy, angle and time. Among them certain particles tend to make more contributions to a specific result compared to other particles and thus are considered of high importance. Source biasing works well for such problems. Source biasing samples from a biased source distribution in phase space; this means the particles of high importance are much more likely to be selected than particles of low importance. The weight of the source particles will be consequently corrected to conserve the total weight of particles in the given interval as well as to preserve an unbiased estimate [6] [8].

Implicit Capture

In analog MC simulation, efforts are wasted on tracking interactions that lead to an absorption of the particle. For example, a neutron that has been tracked for multiple collisions could suddenly be terminated by absorption and thus no contribution is made to the calculation. Implicit capture, also known as survival biasing, is a variance reduction technique that avoids the problem of killing particles by absorption. To guarantee a fair game, the weight of the particle after the collision is multiplied by the non-absorption probability to

the weight before the collision [9]:

$$w_{after\ collison} = w_{\text{before } collison} \times \left(1 - \frac{\Sigma_a}{\Sigma_t}\right) \tag{41}$$

where Σ_a is the absorption cross section and Σ_t stands for the total cross section.

The advantage of implicit capture is that efficiency is increased as particles can be terminated only by transmission or reflection and thus will always be contributing to the calculations without the loss of absorption.

Exponential Transform

Exponential transform, or path stretching, is a technique that biases particles into the desired area by biasing the distance between collisions. When a particle is moving towards the detector, distances that are greater than one mean free path would be sampled. When a particle is moving away from the detector, distances that are shorter than one mean free path would be sampled. After the sampling, the total macroscopic cross section would be modified as [9]:

$$\Sigma_t^{\text{mod}ified} = \Sigma_t^{original}(1 - p\mu) \tag{42}$$

where p is an exponential transform parameter to vary the degree of biasing, and μ is the cosine of the angle between sampled direction and the original direction.

Thus in general, particles moving towards the detector will have their paths stretched and particles moving away from the detector will have their paths shrunk. Exponential transform

technique is particularly useful for deep-penetration transport problems where a source particle has a slight chance to method the detector.

Splitting with Russian Roulette

Splitting in MC simulation usually contains three categories: geometry splitting, energy splitting and time splitting [6] [9].

Geometry splitting is a variance reduction technique extensively employed in Monte Carlo codes. The essential idea of geometry splitting is to increase the number of particles moving towards a desired direction of interest while particles moving towards an uninterested direction are killed to avoid wasting further efforts. Geometry splitting divides the problem geometry into a set of cells and each cell is assigned with an importance I . If a particle of a weight w migrates from a cell of importance I_k to a cell with a higher importance I_j , the particle will be split into a number of $n = I_j / I_k$ identical particles each of the weight w/n .

On the other hand, if a particle of a weight w migrates from a cell of importance I_k to a cell with a lower importance I_j , Russian roulette will be played, and the particle will be killed with probability $p = 1-(I_j / I_k)$. If the particle is not killed, it will be followed further with probability I_j / I_k and weight $w \times I_k / I_j$. It should be noted although geometry splitting is generally considered a very reliable technique in MC simulations; it does not work in problems that have extreme angular dependence because no particles ever enter an important cell where the particles could be split in this extreme case [6].

Energy splitting works similarly to geometry splitting. It is performed on the energy space and independent of spatial cell. Particles could be split when they pass in between different energy ranges and Russian roulette is played correspondingly to reduce the particle number and computational time.

Time splitting is similar to the above two types of splitting, except a particle's time could only increase. Since particles are more important later in time for certain circumstances, it is more useful to split particles as time increases. One example is the case where a detector responds primarily to late time particles. If too many late time particles exist, the late time particles could be roulette to preserve reasonable calculation efficiency [6].

Weight Window

The weight window [6] is a representation of phase space splitting and Russian roulette technique implemented in the MCNP code. The phase space could stand for space, space-energy or space-time. Figure 2.2 below borrowed from MCNP manual provide a detailed presentation about how the weight window technique works.

A weight window is a window of suitable weights defined by the upper weight bound and the lower weight bound. In MCNP, these weight bounds are defined by the user. If a particle is between the upper and lower weight bounds, no action will be taken; if a particle is below the lower weight bound, Russian roulette will be played to decide if the particle be terminated or increase to a weight within the window; if a particle is above the upper weight bound, it will be split to multiple particles with weights within the window.

There are several importance differences between weight window technique and geometry split/roulette technique although they work in the similar way [6].

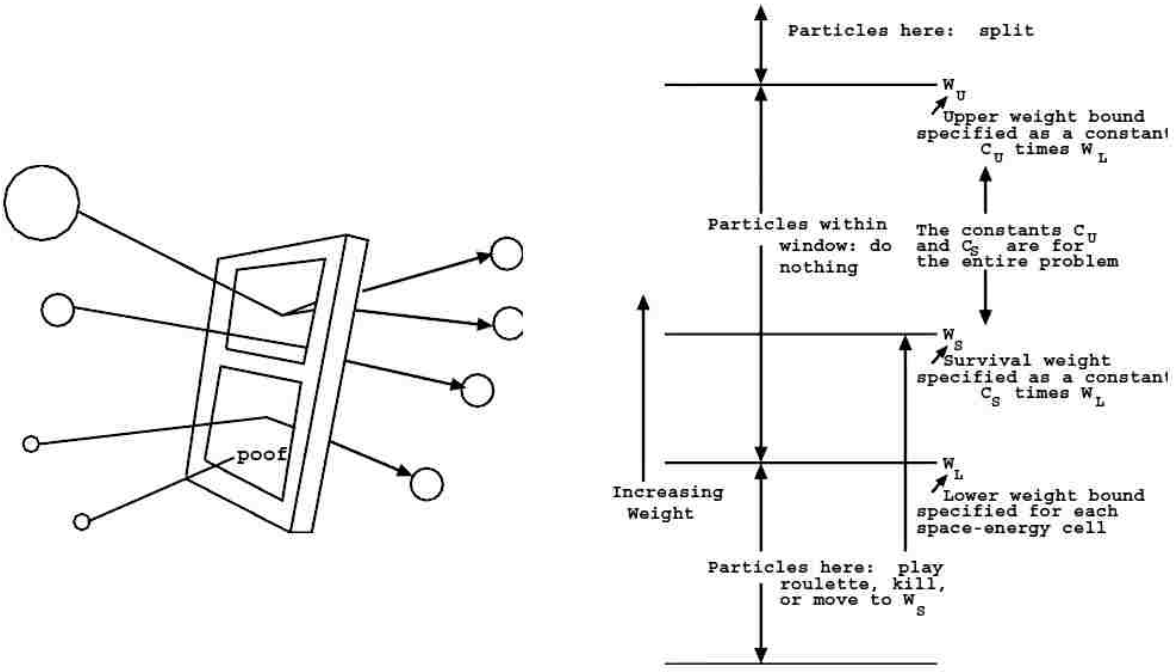


Fig. 2.2: Weight Window (referred from MCNP manual)

W_L , the lower weight bound.

W_S , the survival weight for particles playing roulette

W_U , the upper weight bound.

1. Geometry splitting is only space dependent while the weight window is space-energy dependent or space-time dependent.
2. Geometry splitting does not concern the particle weight while the particle weight is an important consideration in weight window for deciding the corresponding action.

3. The weight window could be applied at surfaces or collision sites or both. Meanwhile, splitting could be only employed at surfaces.
4. The weight window is able to control fluctuations of particle weight introduced by other variance reduction techniques while the geometry splitting is weight independent and preserve weight fluctuations.
5. The weight window could be turned off for selected energy or space regimes.

The weight window could be generated by the weight window generator. A weight window generator generates weight window importance functions automatically [6]. Dividing the phase space into a number of different phase space "cells" or regions, the importance of a cell is then defined as the expected score generated by a unit weight particle after entering the cell. Therefore, the cell's importance is estimated as:

$$\text{Importance} = \frac{\textit{total}\ \text{score due to the particles entering the cell}}{\textit{total}\ \text{weight entering the cell}} \tag{43}$$

The weight window technique requires an accurate estimate of the importance function. Otherwise, if a phase space is inappropriately sampled, unreliable importance estimate or even no importance estimate will be instead generated. It is also important to notice that if mesh cells are created too small, the weight window generator will also fail due to inadequately sampled subdivisions. In general, weight window is a powerful variance reduction technique that is superior to geometry splitting/roulette.

2.2.3 Variance Reduction based on the Adjoint Function

In MC particle transport problems, to calculate the response at a certain location is equivalent to computing the following integral:

$$R = \int_{V_d} \int_E \varphi(\vec{r}, E)\sigma_d(\vec{r}, E)dEdV \tag{44}$$

where φ is the forward scalar flux, σ_d is the response function in phase space over the volume V_d.

Using the fundamental property of the adjoint equation $\langle \psi^\dagger, H\psi \rangle = \langle H^\dagger \psi^\dagger, \psi \rangle$, the response R at a certain location could also be introduced as:

$$R = \int_{V_S} \int_E \varphi^+(\vec{r}, E)q(\vec{r}, E)dEdV \tag{45}$$

where φ^+ is the adjoint flux, q is the source density over the source volume V_S.

The adjoint function represents the particle importance with respect to the corresponding response. However, the evaluation of Eq. (45) requires knowledge of the adjoint function that is typically not available in a closed form. Instead, an approximation of the adjoint flux is employed to generate an importance map and a biased source [8]. Based on the adjoint function, Wagner introduces an alternative pdf $\hat{q}(\vec{r}, E)$ to replace the original pdf $q(\vec{r}, E)$ [8]:

$$R = \int_{V_S} \int_E \frac{\varphi^+(\vec{r}, E)q(\vec{r}, E)}{\hat{q}(\vec{r}, E)} \hat{q}(\vec{r}, E)dEdV, \tag{46}$$

where $\hat{q}(\vec{r},E)$ is given by:

$$\hat{q}(\vec{r},E)=\frac{\psi^{+}(\vec{r},E)q(\vec{r},E)}{\int_{V_S}\int_{E}\psi^{+}(\vec{r},E)q(\vec{r},E)dEdV} \tag{47}$$

This is a typical example of importance sampling. However, since the adjoint function is not exact, MC simulations for particle transport are necessary. As stated before, in source biasing, the weight of a source particle is corrected as

$$w(\vec{r},E)\hat{q}(\vec{r},E)=w_0 q(\vec{r},E) \tag{48}$$

where w_0 is the unbiased weight of a particle and usually set equal to 1. By substituting the equation (47) into equation (48), the following result could be obtained:

$$w(\vec{r},E)=\frac{\int_{V_S}\int_{E}\psi^{+}(\vec{r},E)q(\vec{r},E)dEdV}{\psi^{+}(\vec{r},E)}=\frac{R}{\psi^{+}(\vec{r},E)} \tag{49}$$

which shows an inverse relationship between the adjoint function and the particle weight. In this way, the target weight matches the particle's energy and position precisely. This relationship has been verified through computational analysis in [14] as well as been derived in [8].

In [8], it is demonstrated that the number of particles emerging in $(\vec{r},E)$ from an event in $(\vec{r}',E')$ is adjusted by the ratio of importance:

$$r=\frac{\psi^{+}(\vec{r},E)}{\psi^{+}(\vec{r}',E')} \tag{50}$$

If r greater than 1, particles are created by splitting; if r lower than 1, particles are destroyed through rouletting. Employing the ratio r, the following equation could be obtained as shown in [8]:

$$w(\vec{r},E) = w(\vec{r}',E')\frac{\psi^{+}(\vec{r},E)}{\psi^{+}(\vec{r}',E')} \tag{51}$$

Since the above relationships for the particle weights were derived from importance sampling in a consistent manner, it is referred to as the CADIS (Consistent Adjoint Driven Importance Sampling) method [8] and is implemented in SCALE 6.1 code sequence [24]. The CADIS method uses a discrete ordinates code to determine the adjoint particle flux. The obtained adjoint flux, regarded as the importance of particles, is employed in generating the biased source and corresponding weight window map [1]. The adjoint-flux-based importance map is required to be consistent with the source biasing to avoid insufficient survival particles that will cause a waste of computational time.

An extension of the CADIS method, forward-weighted CADIS (FW-CADIS) is considered a state-of-art variance technique [10]. It is developed by applying an inexpensive discrete ordinates code to perform a forward Sn calculation to estimate the expected tally responses. The adjoint sources that correspond to each tally are weighted inversely by the forward tally estimate. Then the standard CADIS is applied with an importance map and a biased source that are generated using the adjoint flux computed from the adjoint Sn calculation.

For example, in order to calculate a detector response function $\sigma_d(\vec{r},E)$ over a mesh tally, the adjoint source would be written as [10]:

$$q^{+}(\vec{r},E)=\frac{\sigma_d(\vec{r},E)}{\int\psi(\vec{r},E)\sigma_d(\vec{r},E)\ \mathrm{dE}}\ , \qquad 52)$$

where $\psi(\vec{r},E)$ is an estimate of the forward flux.

In the current implementation in SCALE 6.1, two options for forward weighting are available [24]:

1. For tallies where the entire group-wise flux is required with low relative uncertainties, the adjoint source should be weighted inversely by the forward flux $\varphi(\vec{r},E)$;

2. For tallies where only an energy-integrated quantity $\int\sigma_d(E)\varphi(\vec{r},E)dE$ is desired, the adjoint source should be weighted inversely by the energy-integrated quantity.

The table below from SCALE 6.1 MAVRIC Manual shows how the adjoint source is weighted by certain quantity to optimize the forward Monte Carlo simulation at multiple tally locations [24]:

Table 2.1: Adjoint Source Weighting in FW-CADIS

Calculation	Adjoint Source
Energy and spatial dependent flux $\psi(\vec{r},E)$	$q^{+}(\vec{r},E)=\frac{1}{\psi(\vec{r},E)}$
Spatial dependent total flux $\int\psi(\vec{r},E)dE$	$q^{+}(\vec{r},E)=\frac{1}{\int\psi(\vec{r},E)dE}$
Spatial dependent total dose rate $\int\psi(\vec{r},E)\sigma_d(\vec{r},E)dE$	$q^{+}(\vec{r},E)=\frac{\sigma_d(\vec{r},E)}{\int\psi(\vec{r},E)\sigma_d(\vec{r},E)dE}$

The overall goal of FW-CADIS is to achieve the particle density uniform over MC tallies via an importance map based on adjoint flux information. By obtaining more uniform particle densities, more uniform relative errors for the tallies will be realized. FW-CADIS has been proven to be a powerful GVR technique in many applications [10].

2.2.4 Correlation of Responses in MC Simulation

A correlation in statistics refers to inter-dependence of two (or in general more) random variables, i.e. how does one variable change if another one is changed or in the extreme of total correlation: does fixing one random variable totally determine another random variable. For example, the median income of a school district can be conjectured to be correlated to the average SAT scores within this district, because, in general, the richer the area the better the school. Correlations often (not always) imply an underlying causality but do not investigate or require it between the two correlated quantities. For example fact Z can cause both X and Y such that X and Y are correlated. However, X and Y can be completely unrelated in terms of a causal relationship.

The covariance introduced in section 2.1 is a measure of the linear dependence of a pair of random variables. However, its general use is limited because the magnitude of the covariance scales with the magnitude of the random variables and therefore general comparisons in-between different systems are difficult because of a possible difference in scaling. It is more convenient to measure the correlation of two random variables using a scaled/dimensionless quantity which gives rise to the correlation coefficient(s).

The most commonly used correlation coefficient was introduced by Pearson and is defined as [40]:

$$p_{i,j} = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_i \sigma_j} \tag{53}$$

where $\mathrm{cov}(x_i, x_j)$ stands for the covariance between variables x_i and x_j; σ_i and σ_j stands for the relative uncertainty of each variable.

An important property of Pearson's correlation coefficient is that it is only sensitive to linear correlations: A linearly uncorrelated, but non-linearly highly correlated phenomenon would seem to be uncorrelated when only Pearson's correlation coefficient is used. Values of the correlation coefficient range from -1 representing a perfectly negative linear correlation to +1 representing a perfectly positive correlation. In any of these limiting cases a random variable is totally determined by fixing another random variable by a relationship of the form:

$$x_i = \pm m x_j + c. \tag{54}$$

Totally uncorrelated variables exhibit a correlation coefficient of zero, but the converse, i.e. a zero correlation coefficient implies uncorrelated random variables, is not true because random variables can be correlated non-linearly. It is, however, true that a zero correlation coefficient implies that random variables are linearly uncorrelated.

In Fig. 2.3, Pearson's correlation coefficients are listed for four data sets featuring very different linear correlations. It could be seen from the plot that the upper left data set has the highest correlation of a correlation coefficient 0.99; while the lower right data set has the

lowest correlation of a negative correlation coefficient -0.2. The most correlated data set $p = 0.99$ has the best linear fit while the least correlated data set $p = -0.2$ has the worst linear fit.

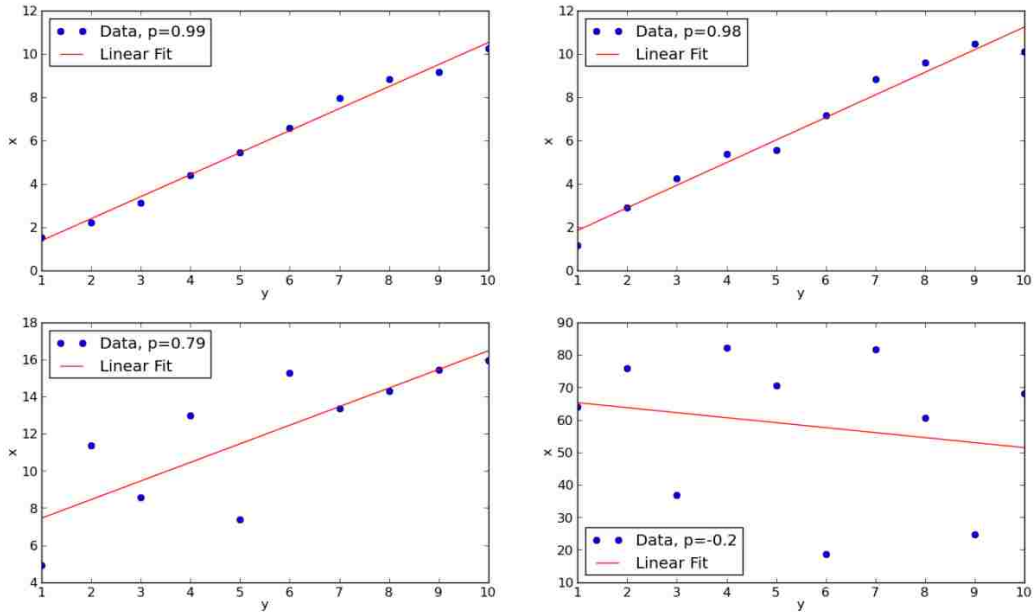
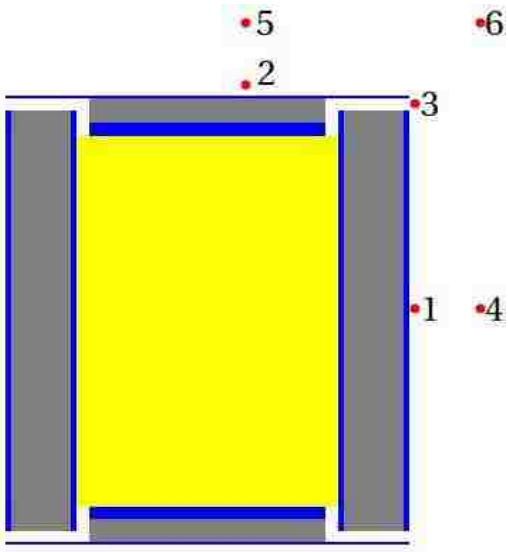


Fig. 2.3: Pearson's Correlation Coefficients

In reactor physics, correlations also exist between responses of different tallies in the regime. This means the more correlated two tallies are, the more similar their response will be. 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. Earlier work has shown that in reactor calculations responses representing distributions such as group fluxes and reaction rates are highly correlated [30].

A full-size cylindrical cask model from SCALE 6.1 manual [24] is employed to show such correlations between responses. Neutron flux responses are detected at 6 different detector locations as shown in the Fig. 2.4. Detector 1 and 4, Detector 2 and 5, Detector 3 and 6 are located very close to each other respectively. The adjoint flux profile for each detector is plotted separately in the Fig. 2.5. It is quite obvious that the detector 1 has the most similar adjoint flux profile to detector 4. The same similarity also exists in between detector 2 and 5, detector 3 and 6.

In MC simulation particularly for challenging reactor physics problems, identifying these correlations in an automated manner could be used to identify the minimum number of weight-window maps that are independent of each other. Motivated by this fact, a new hybrid GVR technique named the SUBSPACE method, is developed by taking advantage of the correlations between responses and their associated weight-window maps to improve the efficiency of MC calculation. The detailed implementation of the SUBSPACE method is introduced in Chapter 4.



Detector locations: (coordinates in cm)

1: 180, 0, 0; midplane, 10cm from surface

2: 0, 0, 295.6; on axis, 10cm from surface

3: 180, 0, 267.9; center of gap, 10cm r_{max}

4: 270, 0, 0; midplane, 100 cm from surface

5:0, 0, 385.6; on axis, 100cm from surface

6:270, 0, 385.6; 100cm from each surface

Fig. 2.4: Cask Geometry and Detector Locations from SCALE 6.1 Manual

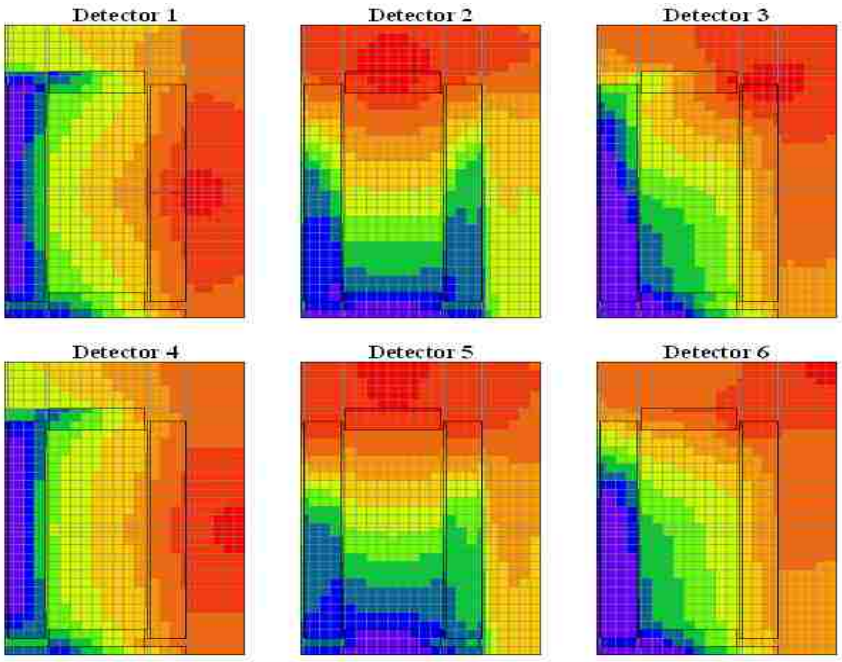


Fig. 2.5: Adjoint Flux Profiles of Detector 1~6

2.3 Monte-Carlo-Deterministic Hybrid Method

Over the course of the last five decades, there has been a growing interest in coupling Monte Carlo and deterministic methods employing a hybrid method to combine their benefits and overcome some of their individual deficiencies [1],[2],[16],[19],[21],[33]. 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". Among the variance reduction techniques, splitting/roulette and implicit capture have been the most widely applied techniques for reducing the variance of the Monte Carlo calculations [17]. These techniques, together with many others, are available in most standard Monte Carlo codes including MCNP [6], TRIPOLI [22], MORSE [26], and MCBEND [42].

These techniques have been successfully demonstrated to reduce the variance for a single response [1], [13], [16] often representing a functional of the solution over a region in the phase space, i.e., a detector's response in a given region. Examples are the TRIPOLI MC code [22] and more recently the Lift (local importance function transform) Method developed by University of Michigan [43]. The TRIPOLI MC code employs an importance function to obtain a number of advanced biasing schemes, for example, exponential biasing, quota sampling and collision biasing. The importance function is generated on a user-defined mesh via several means, including a method based on graph theory and a 2-D discrete ordinates adjoint solver [44].

Lift Method employs an approximation of the exact adjoint solution to approximate the zero-variance method for the source-detector problem, and therefore overcomes the difficulty of searching the exact solution of the adjoint transport problem. This approximation uses a deterministic adjoint calculation to obtain localized biasing parameters for source biasing, collision biasing and path length biasing [43].

When these methods prove to be powerful, nevertheless, they are very sensitive to the accuracy of the importance function, require additional user input and are statistically instable compared to the splitting/roulette methods alone [31]. Furthermore, once the problem is expanded from a detector to a large region, Global Variance Reduction (GVR) becomes the primary goal, especially when Monte Carlo methods are to be used for reactor analysis applications. In this case, the above methods are no longer effective for the purpose.

As stated before, GVR 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. Therefore a uniform distribution of MC particles throughout the domain is crucial. Responding to this challenge, a number of methods have been developed in the field. In general, these methods could be divided into two categories: i) Methods that employ a deterministic forward flux solution as the basis of a

global MC calculation; ii) Methods that employ the concept of adjoint function (usually deterministic) to bias MC calculation and obtain a uniform global variance reduction.

One of the most widely applied variance reduction methods that utilize deterministic forward flux solution is the weight window generator developed by Booth and Hendricks [2]. Weight window generator, originally known as forward-adjoint generator, is an importance estimation technique. The weight window technique is implemented in the standard version of MCNP. It applies splitting and roulette to obtain variance reduction. A forward MC calculation is performed, and the particle importance is calculated as the ratio of the total score entering a cell to the total weight entering a cell.

Solomon and Booth later developed the Talley Linear Combination technique by implementing a linear tally combination in MCNP [21]. In conjunction with MCNP's weight window generator, this technique optimizes a linear combination of tallies and is shown to be equivalent to increasing the adjoint source strengths of detectors by means of the linear multiplier. Solomon and Booth have successfully extended this technique to generate global importance in conjunction with MCNP's mesh tally capabilities.

Larsen's group has also developed several methods, based on forward flux estimate, to obtain lower uncertainties everywhere in the problem. Cooper and Larsen's weight window method [18] employs a forward deterministic solution to set a space-angle-energy- dependent weight window to obtain a uniform distribution of Monte Carlo particles throughout the system. While the method is shown significantly more efficient than implicit capture, however, since the forward flux does not represent the expected uniform response throughout the phase space, the method is not effective when applied to a large realistic problem.

Becker and Larsen proposed a hybrid technique that calculates an inexpensive deterministic global estimate of the forward flux then uses MC to estimate the multiplicative correctons [13]. Since the correcton flux has much less spatial variation compared to the actual physical flux, a more uniform distribution of correctons is obtained instead of the distribution of the original angular flux. Similar to Cooper's method, the correcton method employs an inexpensive deterministic global estimate of the forward flux. Different from Cooper's method [18], the correcton method distributes correctons instead of angular flux particles in MC simulations and therefore shows more efficiency in obtaining a uniform global variable reduction. However, the correcton method is more complicated to implement than Cooper's method, and has not been widely incorporated in computational codes up to date.

While the above methods, for example the weight window generator, prove useful in the iterative process of obtaining variance reduction parameters [20, 31], they share the same fundamental problem that an insufficient number of particles pass through the space-energy interval and cause the lack of accuracy in estimating the importance, which means, either no importance estimate or an unreliable importance estimate for this interval.

Wagner states:"Current forward MC importance generators are restricted by their statistical nature and are of limited use in multi-dimensional deep-penetration problems."[8]

Responding to the difficulties related to statistical importance estimation, people turned to the possible methods that utilize the concept of the adjoint function based on the fact that a deterministic means to generate accurate importance would significantly accelerate the MC calculation for large and complicated problems. The adjoint function, as the solution to the

adjoint Boltzmann transport equation [50], has been long recognized for its physical significance as a measure of the particle importance to desired responses. Kalos[11] and Coveyou [15] proposed using the adjoint function as an importance function to advantageously bias the MC calculation in their early work. Realizing this important physical interpretation of the adjoint function, several works in the field have employed this fact for biasing MC calculations.

One of the earliest literatures in this area by Tang [46], [47] employs two-dimensional (2-D) discrete ordinates adjoint functions to bias multigroup MC calculations. This method later is implemented in the SAS4 sequence [48] of the SCALE code package (SCALE, 2001) for spent fuel cask dose calculations using 1-D Sn adjoint functions.

AVATAR method [20] developed by Los Alamos National Laboratory employs the adjoint function to construct space- energy- and angular-dependent weight windows for MCNP. The adjoint function is calculated by a 3-D Sn code THREEDANT [45].

Since AVATAR is automated, it requires little user intervention to generate the weight window and thus human efficiency is significantly increased. However, the difficulty associated with the automated process still remains, that is the user must provide the weight window values and corresponding spatial and energy grid.

Barrett and Larsen developed the Variational Variance Reduction method [16]. This variational method uses a variational functional that employs first-order estimates of forward and adjoint fluxes that are of low accuracy to yield a more accurate second-order estimate of

a desired system characteristic, for example, the criticality eigenvalue. Variational Variance Reduction method is different from traditional variance reduction methods given that the particle weight in the MC transport process is never modified. However, it has been utilized together with other variance reduction methods that indeed modify the particle weight [1]. The Variational Variance Reduction method is currently under development to solve energy-dependent multi-dimensional problems.

Motivated by the expensive computational expense of performing MC calculations particularly in reactor physics, Wagner and Haghighat proposed the CADIS (Consistent Adjoint Driven Importance Sampling) method [1]. The CADIS method calculates consistent source biasing and weight window parameters based on the space- and energy- dependent deterministic adjoint function. The equations for obtaining the biasing parameters are consistently derived with the adjoint function as the weighting function [8].

Since the CADIS method eliminates the manual effort of selecting variance reduction parameters; it significantly saves user's preparation time for large MC calculations.

The CADIS method has been implemented into codes ADVANTG (based on MCNP) and the MAVRIC sequence of SCALE 6.1. It has been demonstrated very effective in optimizing the calculation of a response for a single point or a small region and thus extremely powerful for variance reduction of deep-penetration problem with a distributed source. However, CADIS method has not been successful in optimizing global distributions due to the inefficiency in specifying the adjoint sources throughout the phase space and therefore fails to obtain a uniform global variance reduction [10].

To overcome this difficulty, Wagner and Peplow developed the FW-CADIS (forward-weighted Consistent Adjoint Driven Importance Sampling) method as a direct extension of the CADIS method. FW-CADIS method employs two approximate calculations: one forward and one adjoint, to generate consistent source biasing and weight window parameters for MC simulations. The forward calculation is used to define appropriate adjoint sources that are later used in a deterministic adjoint calculation to generate the adjoint function. The generated adjoint function is employed to assign importance values to various regions in the phase space and thus uniform responses are achieved throughout the system. FW-CADIS method has been incorporated in the MAVRIC sequence of SCALE 6.1. It has been proven that FW-CADIS method is very promising in optimizing global distributions.

While the idea behind the FW-CADIS method 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 a less computational burden compared to other methods. In this dissertation, we propose a new method to address the simultaneous reduction of response variances everywhere in the phase space. The new method, denoted by the SUBSPACE method, optimizes the selection of the weight windows for reactor analysis problems where detailed properties of all fuel assemblies are required everywhere in the reactor core. Like the FW-CADIS, the SUBSPACE method utilizes importance maps obtained from deterministic adjoint models to derive automatic weight-window biasing. In contrast to FW-CADIS, the SUBSPACE method identifies the correlations between weight window maps to minimize the computational time required for global variance reduction, i.e.,

when the solution is required everywhere in the phase space. The correlations are employed to reduce the number of maps required to achieve the same level of variance reduction that would be obtained with single-response maps. Preliminary numerical experiments are discussed in the following Chapters to serve compare the SUBSPACE and FW-CADIS methods in terms of the reduction in standard deviation of spatially distributed responses.

CHAPTER 3

PROPOSED HYBRID METHODS

3.1 SUBSPACE Method

3.1.1 Motivation

As stated in the previous chapters, hybrid methods 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 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 method 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 method adopted by FW-CADIS, where the weights are selected to be proportional to the inverse of the forward flux. In this method, 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 method 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.

Now, while the FW-CADIS method proves useful, it does not guarantee the associated computational burden will be less than that reached by the brute force method. 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 two responses only as depicted in Fig. 3.1. Each response represents a point detector on the side of a concrete shield with distributed source in the center. All dimensions are in centimeter, both detectors are located 10 cm from the shielding surface. Shield 1 is twice as thick as shield 2 in order to render its 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 method 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 method, the total time required (represented by number of histories) is found to be considerably higher than the brute force method.

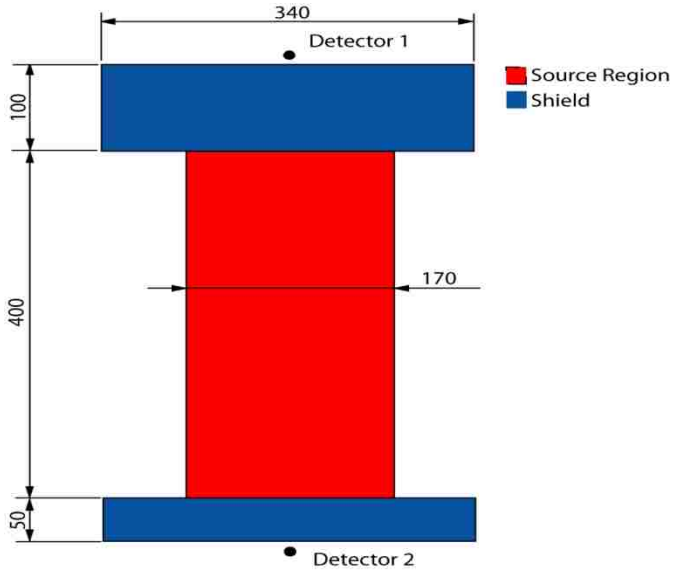


Fig. 3.1: Test Case for FW-CADIS and Brute Force Methods

Fig. 3.2 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 method, and the last two are generated using the FW-CADIS method.

In this work, we propose a third method to addressing the simultaneous reduction of variances for I responses, i.e. global variance reduction. It is denoted by the SUBSPACE method and it may be considered as a trade-off between the brute force and the FW-CADIS methods, 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 pseudo response maps are evaluated, such that $1 < r \ll I$.

The pseudo responses are random linear combinations of the original *I* single-responses. Implementing this method into the FW-CADIS framework should be straightforward, 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 pseudo responses are generated randomly using the SUBSPACE method thus eliminating the need for an extra forward model execution to determine the flux-based weights as currently done by the FW-CADIS method. Finally, given the independence of the *r* executions, the SUBSPACE method allows for coarse-grained parallelization, thereby taking advantage of parallel computing environments.

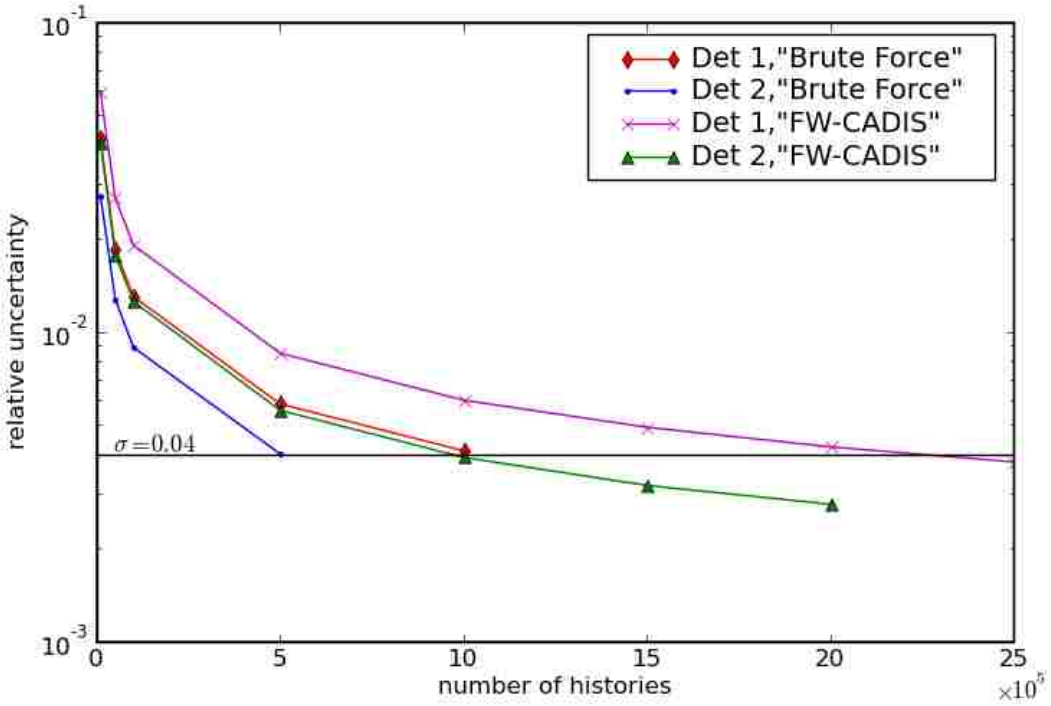


Fig. 3.2: Comparison between Brute Force and FW-CADIS Method

3.1.2 Proposed SUBSPACE Method

The formation of the pseudo responses and their associated weight-window maps is mathematically equivalent to projecting the single-response weight-window maps onto a subspace of smaller dimension which 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 pseudo response. By reducing the variances for the pseudo responses, 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 [27, 31].

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:

$$u_i = \left\langle \vec{\psi}, \vec{\sigma}_i \right\rangle, \text{ and } i = 1,...,I \tag{55}$$

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

$$L^* \vec{\psi}_i^* = \frac{\partial u_i}{\partial \vec{\psi}} = \vec{\sigma}_i \tag{56}$$

where L^* is the adjoint transport operator and $\vec{\psi}_i^*$ is the importance map associated with response u_i . In MAVRIC, the deterministic code DENOVO [49] is used to obtain an approximate solution $\vec{\psi}_i^*$ of the adjoint problem in Eq. (56) on a coarse grid in phase space whose grid points may be indexed by $j = 1,..,J$. The $\vec{\psi}_i^*$ may be written as:

$$\vec{\psi}_i^* = \left[\psi_{i,1}^* \quad ... \quad \psi_{i,J}^* \right]^T \tag{57}$$

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:

$$\mathbf{\Psi} = \begin{bmatrix} \psi_{1,1}^* & \cdots & \psi_{1,J}^* \\ \vdots & & \vdots \\ \psi_{I,1}^* & \cdots & \psi_{I,J}^* \end{bmatrix} = \begin{bmatrix} \vec{\psi}_1^{*T} \\ \vdots \\ \vec{\psi}_I^{*T} \end{bmatrix} \tag{58}$$

In the numerical experiments section, 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 $\mathbf{\Psi}$ [28]. 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 ten 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 [29].

This means that only operations of the form $\mathbf{\Psi}^T\vec{\eta}_j$ (where $\vec{\eta}_j \in \mathbb{R}^I$ is a randomly generated vector) are required to determine the decline in the singular values which 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 pseudo responses, representing random linear combination of the original *I* responses. Let $\tilde{u}_j$ be the j^{th} pseudo response defined by:

$$\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} u_i \qquad \text{for } j = 1,...,r \tag{59}$$

Using the definition for u_i and $\vec{\psi}_i^*$ from Eq. (55) and Eq. (56), one can write:

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

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

Using the linearity of the transport operator, one can write an expression for the importance map for the pseudo response u_i as:

$$\tilde{\vec{\psi}}_j^* = \sum_{i=1}^{I} \eta_{i,j} \vec{\psi}_i^* \qquad \text{for } j = 1, ..., r \tag{60}$$

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

$$\tilde{\vec{\psi}}_j^* = \mathbf{\Psi}^T \vec{\eta}_j , \tag{61}$$

where $\vec{\eta}_j^T = \begin{bmatrix} \eta_{1,j} & \dots & \eta_{I,j} \end{bmatrix}$.

If the $\left\{ \vec{\eta}_j \right\}_{j=1}^{r}$ randomly generated, one can show that the vectors $\left\{ \tilde{\vec{\psi}}_j^* \right\}_{j=1}^{r}$ are independent and span a subspace of size *r* which belongs to the range of the matrix $\mathbf{\Psi}^T$ [27, 29].

Let $\mathbb{C}$ represent the subspace generated by the vectors $\left\{ \tilde{\vec{\psi}}_j^* \right\}_{j=1}^{r}$, and $\mathbb{C}^{\perp}$ as the orthogonal subspace. Now, split each of the importance vectors into two components, one that lives in the subspace $\mathbb{C}$, $\left\{ {}^{\mathbb{C}}\vec{\psi}_i^* \right\}_{i=1}^{I}$, and the other in the subspace $\mathbb{C}^{\perp}$, $\left\{ {}^{\mathbb{C}^{\perp}}\vec{\psi}_i^* \right\}_{i=1}^{I}$. Using elements from random matrix theory [29], 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 method is expected to happen with a small estimate for the rank *r*.

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

Requirements:

- A general methodology that employs an importance map $\vec{\psi}_i^*$ to bias Monte Carlo particles towards a given response u_i.
- The capability to calculate an importance map $\tilde{\vec{\psi}}_j^*$ for a pseudo response defined as a random linear combination of the original *I* responses as defined in Eq. (60).

Objective:

- Identify *r* pseudo response, 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 SVD of the matrix containing the importance maps for the *r* pseudo responses: $\left[\tilde{\vec{\psi}}_1 \quad \tilde{\vec{\psi}}_2 \quad ... \quad \tilde{\vec{\psi}}_r\right]$. If the singular values do not significantly decline, increase the estimate for *r*.

b) PARALLEL DO $j = 1, ..., r$

1. Generate a random vector $\vec{\eta}_j \in \mathbb{R}^I$
2. Form a pseudo response $\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} u_i$
3. Calculate the importance map $\vec{\psi}_i^*$ associated with $\tilde{u}_j$
4. Bias Monte Carlo particles based on the $\vec{\psi}_i^*$
5. Tally the original *I* responses until 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 [30]:

$$u_i^{\mu} = u_i^{\sigma} \sum_{j=1}^{r} \frac{u_{i,j}^{\mu}}{\left(u_{i,j}^{\sigma}\right)^2} \quad \text{and} \quad \frac{1}{u_i^{\sigma}} = \sum_{j=1}^{r} \frac{1}{\left(u_{i,j}^{\sigma}\right)^2} \tag{62}$$

End Result:

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

This algorithm is composed of three steps. Step (a) requires an estimate of the rank *r*. In the following sections, we will 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 method with a special choice for the pseudo response. 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 hand side of the adjoint equation. Next section provides more details on this step for incorporating the SUBSPACE method into the FW-CADIS framework. Step (c) combines the results from the *r* executions, each with *N* independent histories, assume that they are statistically independent [30].

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.

3.1.3 FW-CADIS-Based Implementation

In this section, we discuss the implementation of the SUBSPACE method 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. Second, an adjoint model is executed. The inverse of the flux estimated by the forward model is employed to design the right hand 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 (see Eq. (56)) **[**10**].** 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 method, the user-defined weights are selected randomly and are assigned via the input file. This is equivalent to setting a pseudo response as a random linear combination of all responses as required by Eq. (59). The resulting adjoint solution satisfies Eq. (60) which is equivalent to multiplying the SRI matrix with a random vector.

Next, as discussed earlier, the SUBSPACE method does not require the flux-based weights, thus eliminating the needs for the extra forward model execution. Fortunately, this is also possible, as FW-CADIS framework is flexible enough to provide the user the option to bypass forward flux weighting. Therefore to implement the SUBSPACE method, one needs to 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 statistically combine them as given by Eq. (62). This implementation strategy has been adopted in our work and is employed in the following work to analyze both a core model and an assembly model.

3.2 GAUSSIAN PROCESS METHOD

3.2.1 Motivation

The use of variance reduction methods has proved beneficial in accelerating the convergence of Monte Carlo simulation. While several methods have been proposed to achieve this goal in recent years, hybrid methods that employ deterministic models to bias Monte Carlo particles have received the most attention from the nuclear community's researchers. This section presents a new hybrid method based on the assumption that radiation transport may

be treated as a Gaussian Process [57]. Results obtained using one relatively small dimensional problem (assembly model) indicate that the performance of the Gaussian Process method is comparable to that of the SUBSPACE method. Moreover, it is believed that researching different global variance methods will help provide insight into their mechanics and the hybridization potential needed to combine their benefits.

We focus in this dissertation work on methods that employ the adjoint model with a source term representing the derivative of the given response with respect to the state variable (i.e., flux)[1]. While this methods results in accelerating the convergence for the given response, it may grossly estimate the variance of other responses. A brute force method of repeating the above procedure for different responses could be attempted but quickly proves to be computationally intractable, especially when responses are desired everywhere in the phase

[1] Note that in FW-CADIS terminology, the derivative of response with respect to flux is called the response function; both terms represent the right hand side of the adjoint equation (i.e. the source term).

space[2]. To combat that, other set of techniques, referred to as Global Variance Reduction (GVR) techniques, have been proposed. GVR techniques attempt to distribute particles equally over the phase space to achieve uniform variance reduction [62].

3.2.2 Mathematical Description of GP Method

A Gaussian Process (GP) is one whose random fluctuation described by a normal distribution. In our context, this means that the statistical fluctuations of all responses resulting from the Monte Carlo sampling process could be described by normal distributions. Let $\vec{r} = \begin{bmatrix} r_1 & r_2 & \dots & r_n \end{bmatrix}^T \in \mathbb{R}^n$ represent a vector of the *n* responses of interest representing *n* random Gaussian processes. Consider N realizations of these random processes denoted by: $\left\{ r_i^k \right\}_{i=1}^n, k = 1, 2, ..., N$. The covariance between the two responses r_i and r_j is given by:

$$\operatorname{cov}(r_i, r_j) = \lim_{N \to \infty} \frac{1}{N-1} \sum_{k=1}^{N} \left(r_i^k - \bar{r}_i \right) \left(r_j^k - \bar{r}_j \right) \tag{63}$$

where $\bar{r}_i$ is the estimated mean of the response r_i, i.e.,

$$\bar{r}_i = \frac{1}{N} \sum_{k=1}^{N} r_i^k$$

[2] The term 'brute force' is not to be confused with its more common usage that implies unbiased execution of Monte Carlo. In our context, 'brute force' implies the direct application of the single response variance reduction method for each individual response.

The covariance information between all pairs of *n* responses may be compactly represented by a symmetric covariance matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ such that: $\mathbf{C}_{ij} = \text{cov}(r_i, r_j)$. This matrix may be rewritten using singular value decomposition as follows [28]:

$$\mathbf{C} = \mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^T = \sum_{i=1}^{r} \sigma_i \vec{w}_i \vec{w}_i^T \ , \tag{64}$$

where $\mathbf{W} = \left[\vec{w}_1 \quad \vec{w}_2 \quad .... \quad \vec{w}_r\right] \in \mathbb{R}^{n \times r}$ is a matrix of *r* orthonormal singular vectors, $\boldsymbol{\Sigma} = \text{diag}\left\{\sigma_1, \sigma_2, ..., \sigma_r\right\}$ a diagonal matrix of *r* nonzero singular values, and *r* the rank of the covariance matrix [62]. The power of this decomposition lies in the following property: if one defines pseudo responses of the form: $\xi_i = \vec{w}_i^T \vec{r}$, one can show that the *r* pseudo responses, unlike the original responses, are uncorrelated, i.e.

$$\text{cov}(\xi_i, \xi_j) = \delta_{ij} \ , \tag{65}$$

where δ_{ij} is the Kronecker delta function, such that:

$$\delta_{ij} = 0 \text{ for } i \neq j \text{ and } \delta_{ij} = 1 \text{ for } i = j \ . \tag{66}$$

Note that the pseudo responses are merely linear combinations of the original responses with the weights determined by the singular vectors of the covariance matrix [36]. If the covariance matrix is numerically ill-conditioned, as observed in earlier work, the effective number of pseudo responses may become much smaller than the number of original responses. One can take advantage of this situation by designing weight windows to bias MC particles towards the pseudo responses rather than the original responses.

To provide a basis for comparison, we contrast the new method to both the FW-CADIS and SUBSPACE method [62]. In all three methods, the weight windows-based biasing is determined by solving an adjoint equation of the form (demonstrated for a single response):

$$\mathbf{L}^{*}\left(\phi_i^{*}\right)=\frac{\partial r_i}{\partial \phi} \tag{67}$$

where $\mathbf{L}^{*}$ is the adjoint transport operator with appropriate boundary conditions, $\vec{\phi}$ is the forward flux, and $\vec{\phi}_i^{*}$ is the adjoint flux corresponding to the i^{th} response r_i. If $\vec{\phi}_i^{*}$ is employed to determine weight-windows, one can reduce the variance for only the response r_i and runs the risk of grossly estimating the variances for all other responses. To achieve GVR, the FW-CADIS method formulates a single pseudo response of the form:

$$\xi^{\text{FW-CADIS}}=\sum_{i=1}^{n}\frac{1}{\phi_i}r_i \tag{68}$$

In this method, more MC particles are sent to regions in the phase space where the flux is low. The rationale behind this is that regions with lower flux receive fewer MC particles and their associated responses are therefore expected to have higher variances. To promote uniform variances across the phase space, more particles are encouraged to go to lower flux regions.

The SUBSPACE method argues that the single-response adjoint fluxes are highly correlated, which may be exploited to reduce the number of adjoint calculations required to achieve global variance reduction. This is done by formulating *k* pseudo responses of the form:

$$\xi_j^{\text{Subspace}} = \sum_{i=1}^{n} \alpha_{i,j} r_i \ , \ \ j = 1,2,...,k \tag{69}$$

where *k* is the effective rank of the matrix comprised of all *n* single-response adjoint fluxes, and $\alpha_{i,j}$ are randomly generated weights.

Finally, the GP method formulates *r* pseudo responses of the form:

$$\xi_j^{\text{GP}} = \vec{w}_j^T \vec{r} \ , \ \ j = 1,2,...,r \tag{70}$$

where *r* is the rank of the unknown covariance matrix.

The central requirement for the GP method is the estimation of the covariance matrix $\mathbf{C}$. Since the focus of current work is to demonstrate proof of principle, the matrix $\mathbf{C}$ is estimated after MC calculations are completed. In reality, it can be generated in a couple of different manners. One way is to employ first a FW-CADIS or a SUBSPACE method with few initial histories to provide an initial estimate for $\mathbf{C}$ which could be updated dynamically as more histories are tracked. It could also be estimated deterministically noting that the primary random process in radiation transport is the interaction between a particle and a target nucleus. Since the process of interaction is characterized by cross-sections, one can sample cross-sections using a deterministic model to estimate the covariance matrix [62].

Assume that *l* model executions are employed to build an estimator for the covariance matrix $\mathbf{C}$. In the j^{th} execution, one records the *n* responses denoted by a vector $\vec{r}^{\,j}$ and $j = 1,2,...,l$. After l executions, the estimator for covariance matrix may be constructed as follows:

$$\mathbf{C} = \frac{1}{l-1}\sum_{i=1}^{l}\left(\vec{r}^{\,j} - \bar{\vec{r}}\right)\left(\vec{r}^{\,j} - \bar{\vec{r}}\right) \tag{71}$$

where $\bar{\vec{r}}$ is a vector of estimated means of all responses. An SVD of the matrix **C** is then recovered as in Eq. (64).

Note that in the FW-CADIS method, only one adjoint and one forward flux solution are required. The SUBSPACE method and the GP methods require no forward flux estimates, but they require respectively, *r* and *k* adjoints, where *r* is the effective rank of the matrix comprising all single-responses adjoint, and *k* the effective rank of the covariance matrix described above. These additional adjoint runs have to be considered when comparing figure of merits [62].

CHAPTER 4

NUMERICAL EXPERIMENT

4.1 Fixed Source Simulation

4.1.1 SUBSPACE Method Performance

This section is divided up into three subsections. The first subsection employs a BWR assembly model and the second a PWR core model; both compare the performance of the FW-CADIS and SUBSPACE methods with an estimated rank of r = 10 as discussed in the third subsection.

4.1.1.1 Assembly Model

The first experiment is based on a two-dimensional MARVIC model for a 7x7 BWR assembly; a cross-section of the assembly is shown in Fig.4.1. The assembly model represents the southeast assembly of a typical 2x2 BWR control cell which 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 7x7 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. $I = 49$.

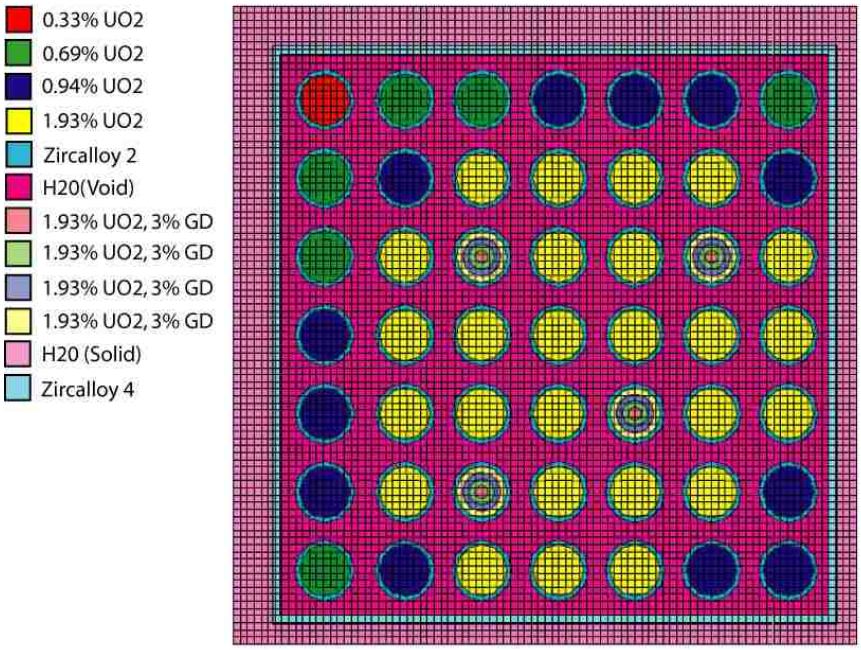


Fig. 4.1: BWR Assembly Model

Table 4.1: 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

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 is given in Table 4.1.

Since the publicly available version of the MAVRIC sequence does not have an eigenvalue solver and is currently limited to source-driven problems only, a fixed source subcritical configuration is analyzed. 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 0.33, 0.69, 0.94, and 1.93% U-235 and 3% gadolinium are employed. The 27 neutron and 19 photon energy group libraries from SCALE [25] are employed for the analysis of the BWR model. For the flux and reaction rates 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 methods are employed to analyze this model with equal number of histories. For each response type, e.g. flux, fission reaction rate, etc., two vectors each of 49 relative uncertainties (one for each fuel pin cell) are obtained from both the FW-CADIS and the SUBSPACE methods. Figs. 4.2 through 4.5 show the reduction of the relative reduction of uncertainty for the flux and reaction rates distribution at the 49 fuel pin sites. Let $\sigma_i^{FW-CADIS}$ and $\sigma_i^{SUBSPACE}$ 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. The reduction in the relative standard deviation resulting from the use of the SUBSPACE method is defined by:

$$\Delta\sigma_i = \frac{\sigma_i^{FW-CADIS} - \sigma_i^{SUBSPACE}}{\sigma_i^{FWCADIS}} * 100\% \tag{72}$$

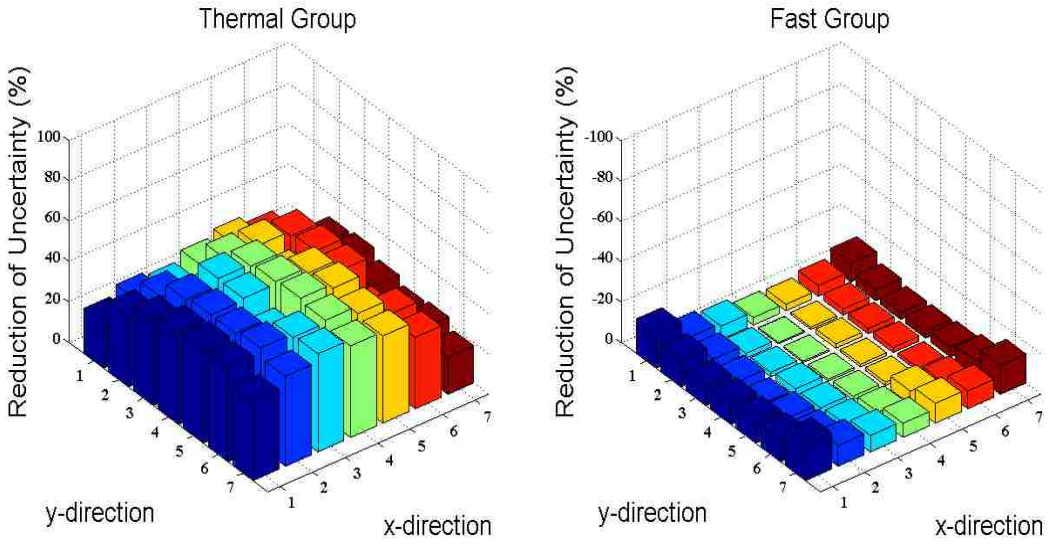


Fig. 4.2: Thermal and Fast Fluxes

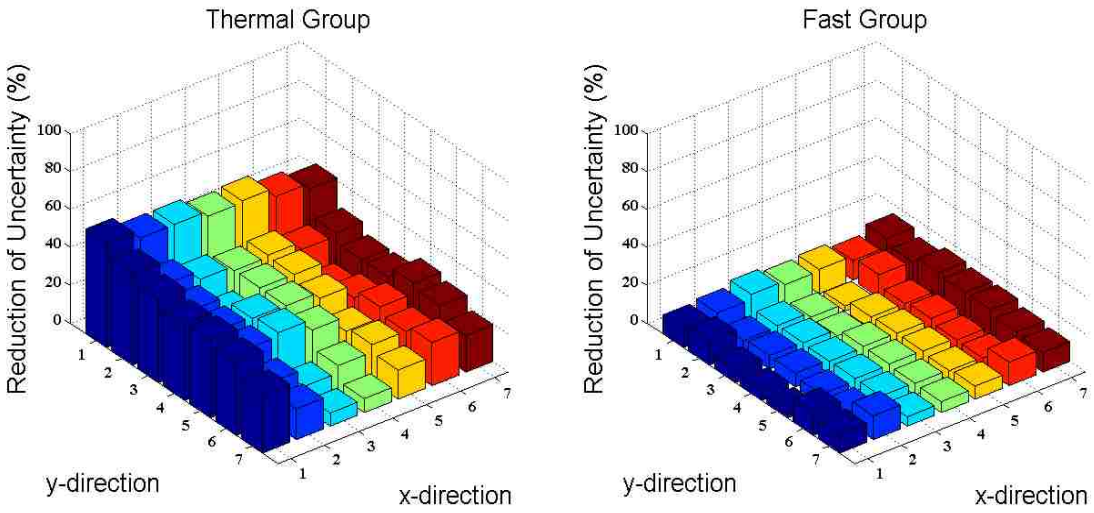


Fig. 4.3: Thermal and Fast Fission Rates

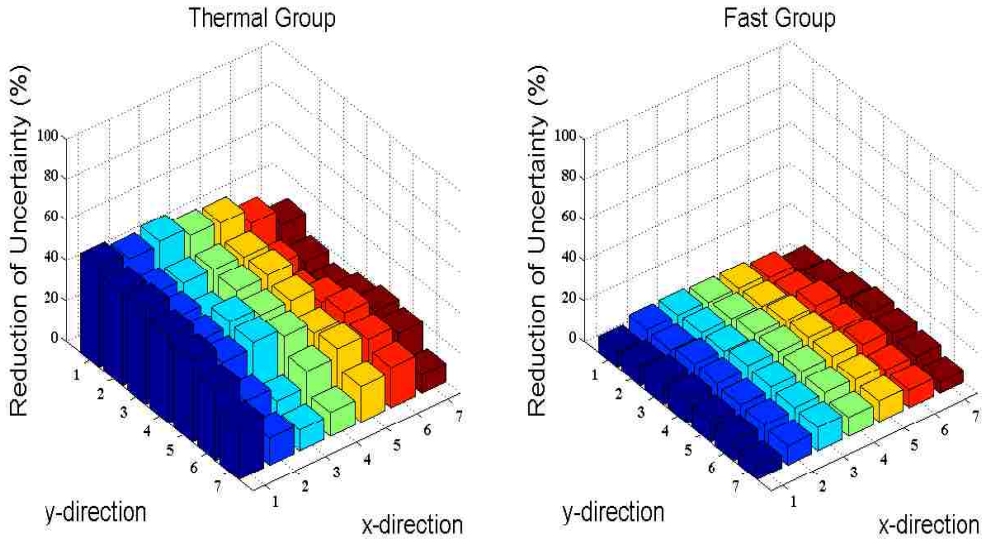


Fig. 4.4: Thermal and Fast Inelastic Scattering Rates

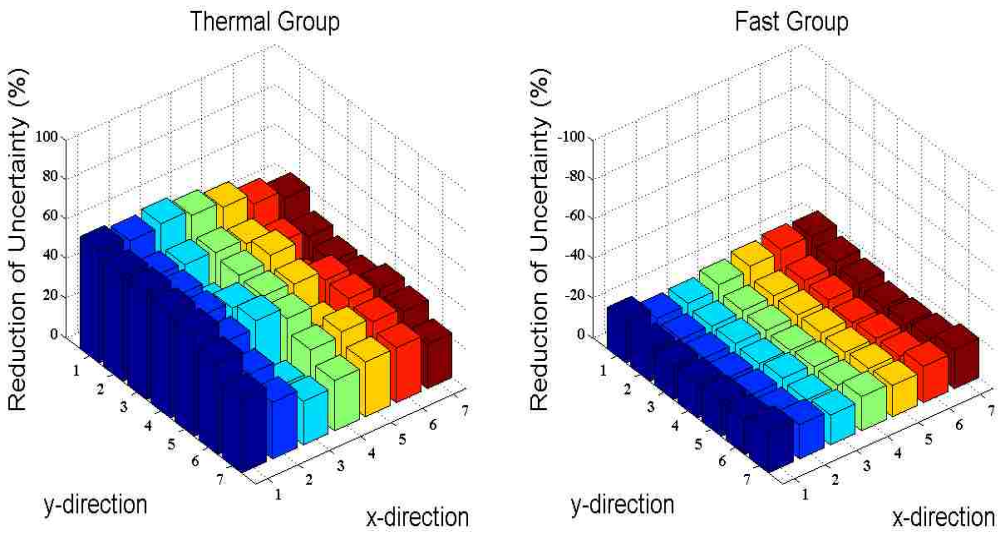


Fig. 4.5: Thermal and Fast Elastic Scattering Rates

The results show that the SUBSPACE method produces lower variances for the thermal flux than FW-CADIS. This is due to the fact that the mean free path for the neutron in a thermal reactor is short which correlates each pin's responses to its nearest neighboring pins. Fast neutrons however have longer mean free path and they are able to visit the entire assembly from their birth to their death, hence the assembly features are more smeared for fast neutrons than they are for thermal neutrons. Since the thermal flux has more significant meaning compared to the fast flux in reactor physics, therefore, in the following work, the thermal flux is always employed as an important quantity to compare the performance of FW-CADIS and SUBSPACE method. Table 4.2 provides a more detailed comparison between the two methods:

Table 4.2: Performance Metrics for Hybrid Methods in Assembly Model

Method	Relative Standard Variance		Number of MC Particles	DT Execution Time (sec)	MC Execution Time (min)	FOM
	AVG	STD				
FW-CADIS	0.2037	0.0454	1.00E+07	31.48	469.88	0.0513
SUBSPACE	0.1182	0.0271	1.00E+07	173.36	581.42	0.1231

The 2nd column in the table labeled AVG is the average value of the relative standard deviations for all the *I* responses. Thermal flux is picked as the desired response for analysis. The STD column represents the standard deviation of the responses' relative standard deviations (this is more commonly known as the square root of the variance of variances); it provides a measure of the spread of the standard deviations throughout the phase space. A large value for STD indicates that the responses variances are not uniformly reduced; therefore a small value would be considered more favorable. Mathematically, each of these metrics are defined as follows for a given method *k*, where *k* = FW-CADIS and SUBSPACE.

$$\mathrm{AVG}^k = \frac{1}{I}\sum_{i=1}^{I}\sigma_i^k \tag{73}$$

$$\mathrm{STD}^k = \sqrt{\frac{1}{I-1}\sum_{i=1}^{I}\left(\sigma_i^k - \mathrm{AVG}^k\right)^2} \tag{74}$$

The number of Monte Carlo particles employed by the FW-CADIS method is 10^7 particles. The same number was used by the SUBSPACE method but distributed over 10 different runs (corresponding to an estimated rank of $r=10$) each with 10^6 particles. The DT execution time for the FW-CADIS method is composed of one adjoint and one forward run.

The SUBSPACE DT time is composed of 10 adjoint runs. The MC time is the time spent by the MONACO code[3]. Finally, the FOM (figure or merit) is based on the following formula:

$$\mathrm{FOM}^k = \frac{1}{\left(\mathrm{AVG}^k\right)^2 T}, \tag{75}$$

where T is the total time including both the DT and MC times. Notice that the DT time is negligible compared to the MC time which is to be expected since all DT calculations are based on source driven models. These assembly results show a 2.5 speed up 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 method does reduce the variances in a uniform manner like the FW-CADIS method.

[3] We noticed that the time spent by MONACO is always a bit higher when using the SUBSPCE method 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 method 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 that the same responses means and standard deviations were obtained as with the standard MAVRIC sequence, the time required was also higher like the SUBSPACE method. This implies that the MONACO code requires an 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 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.

4.1.1.2 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 [34] and [35]. The full core model consists of 193 fuel assemblies (blue regions in Fig. 4.6) laid out in a 17×17 grid scheme and surrounded by light water (red regions in Fig. 4.6).

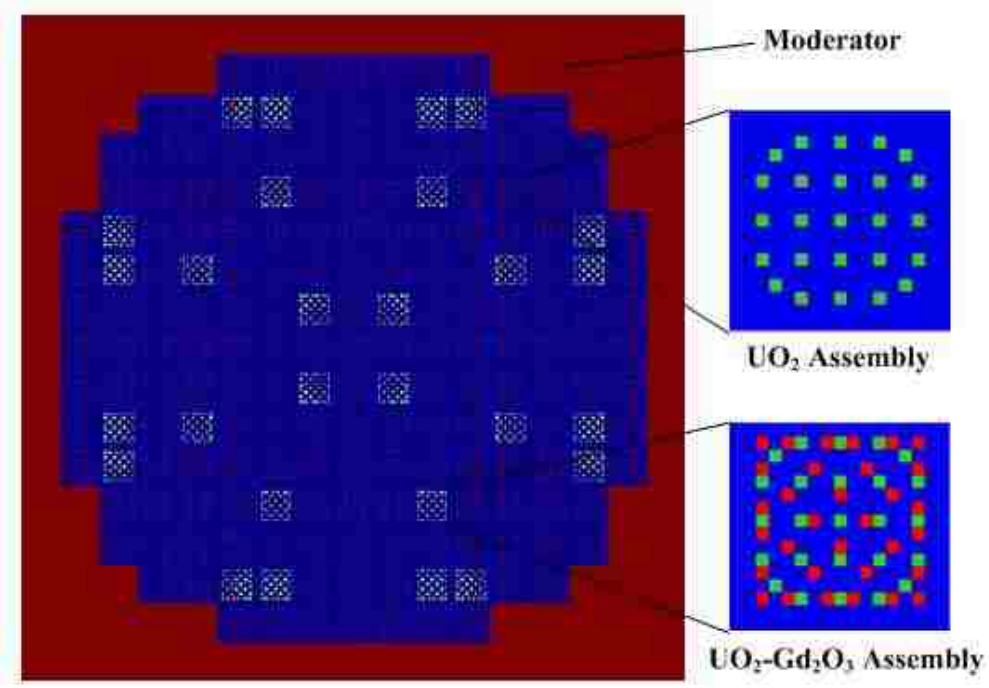


Fig. 4.6: PWR Full Core Model

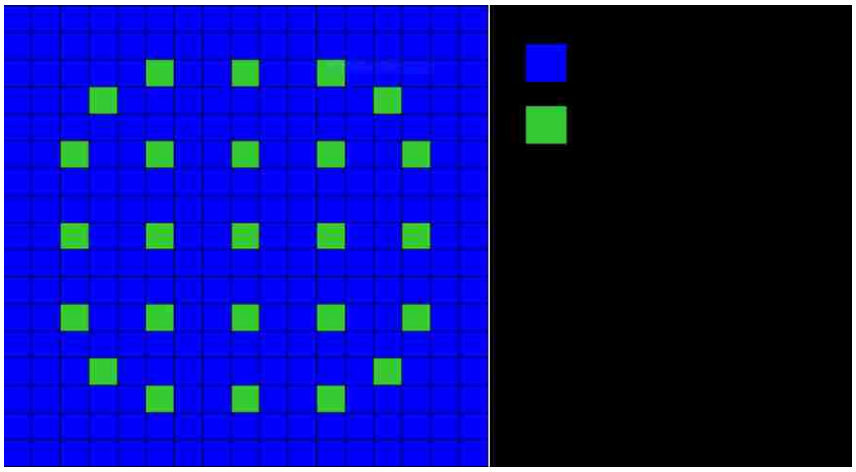


Fig. 4.7: UO2 Fuel Assembly

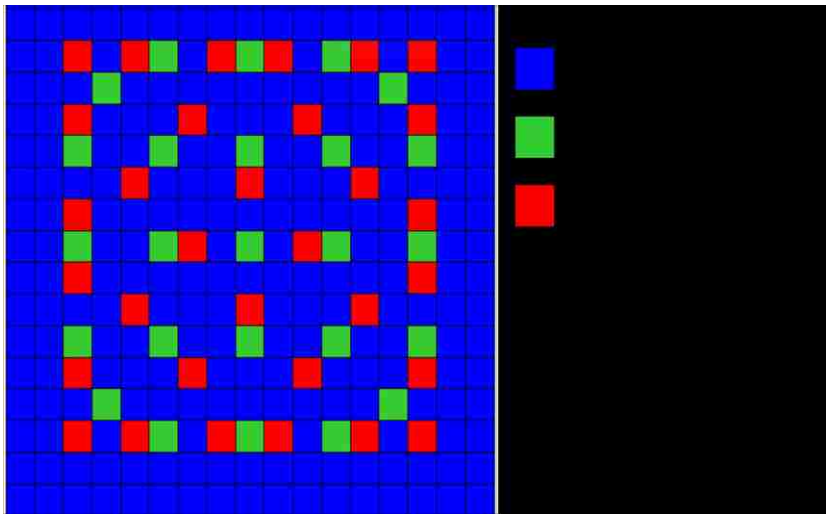


Fig. 4.8: UO2-Gd2O3 Fuel Assembly

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. 4.6 and the loading pattern of each assembly type is plotted in

Fig. 4.7 and Fig. 4.8. 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.

The reduction of relative uncertainties is associated with the assembly-homogenized thermal and fast fluxes and obtained by comparing the FW-CADIS and SUBSPACE methods. As in the previous section, it is calculated as:

$$\Delta\sigma_i = \frac{\tilde{\sigma}_i^{FWCADIS} - \tilde{\sigma}_i^{Subspace}}{\tilde{\sigma}_i^{FWCADIS}} * 100\%$$

where $\tilde{\sigma}_i^{FWCADIS}$ and $\tilde{\sigma}_i^{Subspace}$ are, respectively, the standard deviations for the i th assembly. The spatial dependence of the reduction of uncertainty across the whole core is plotted respectively for the thermal and fast fluxes in Fig. 4.9 and Fig. 4.10.

The result shows that the SUBSPACE method gives a better performance than the FW-CADIS method for thermal flux in all 289 assemblies with the maximum reduction of uncertainty up to 80%. For fast flux, FW-CADIS method performs better and the maximum reduction of uncertainty is up to 20%, specifically for assemblies adjacent to the core's boundaries. Meanwhile, SUBSPACE method shows better results for the rest of the assemblies. This is very similar to the previously discussed BWR assembly model following the same reasoning that thermal neutrons feature shorter mean free paths that consequently restrict corresponding thermal responses to neighboring assemblies only. In contrast, fast neutrons are of longer mean free paths and would hence be more likely to travel throughout the whole core range. Fig. 4.9 shows that for both thermal and fast groups, reasonably uniform uncertainty reductions are obtained.

Between the boundary assemblies and the interior assemblies, the levels of uncertainty reduction are quite different. This is possibly because the boundary assemblies are moderator assemblies while the interior assemblies are fuel assemblies.

In a similar manner to Table 4.2, the core model results of the thermal flux response are presented in Table 4.3. Notice that amount of reduction in terms of FOM is doubled as the size of the problem is increased. Although this may not sound intuitive at a first glance, but from previous experience, we have notice that as the size of the model is increased the potential for reduction via responses correlations is also increased.

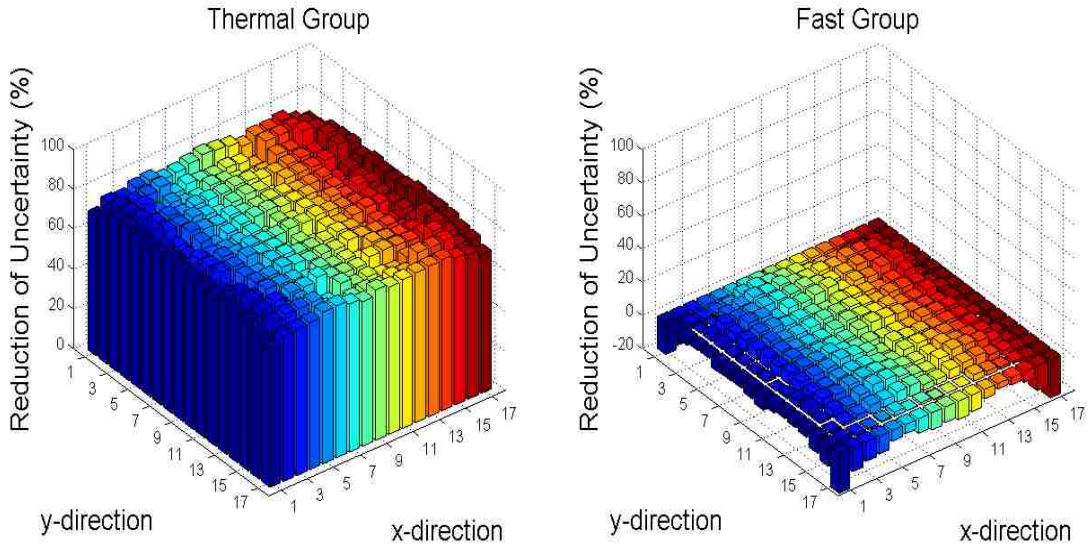


Fig. 4.9: Thermal and Fast Scalar Fluxes for Full Core

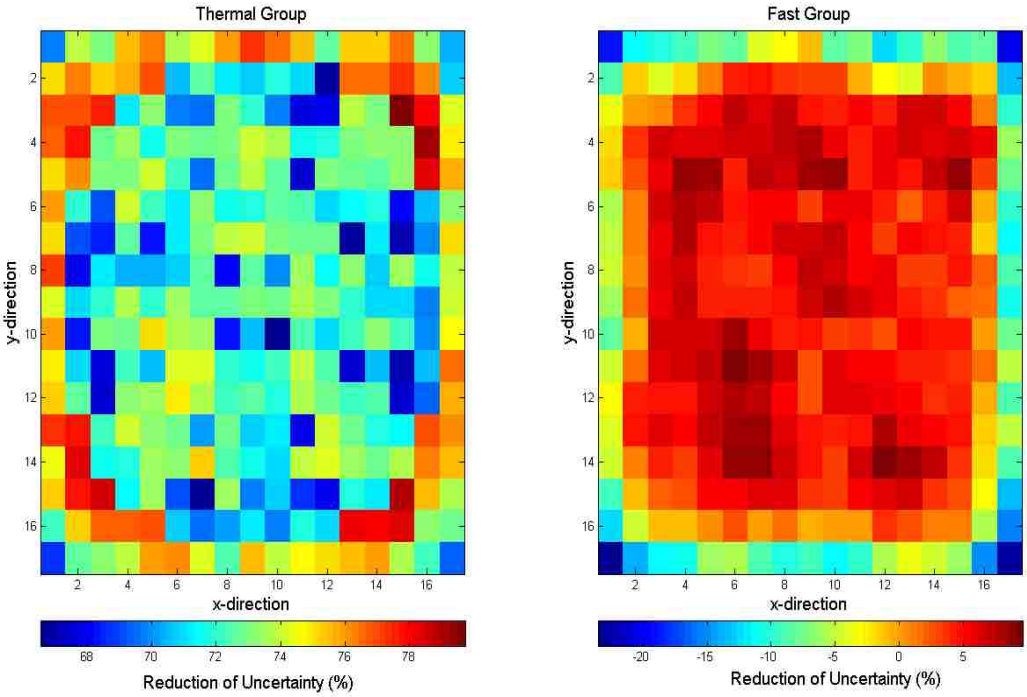


Fig. 4.10: Thermal and Fast Scalar Fluxes for Full Core

Table 4.3: Performance Metrics for Hybrid Methods in Core Model

Method	Relative Standard Variance		Number of MC Particles	DT Execution Time (sec)	MC Execution Time (min)	FOM
	AVG	STD				
FW-CADIS	0.904	0.133	1.0E+07	1.85	74.23	0.0165
SUBSPACE	0.338	0.031	1.0E+07	9.59	97.45	0.897

4.1.1.3 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 method 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 [36]. Notice that the singular values plotted in Fig. 4.11 fall down by three orders of magnitude by the time the tenth singular value is reached. After that, the singular values continue to fall down but at a much smaller 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 models results, previously completed with r = 10, are repeated with different estimates for the rank. Fig. 4.12 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.

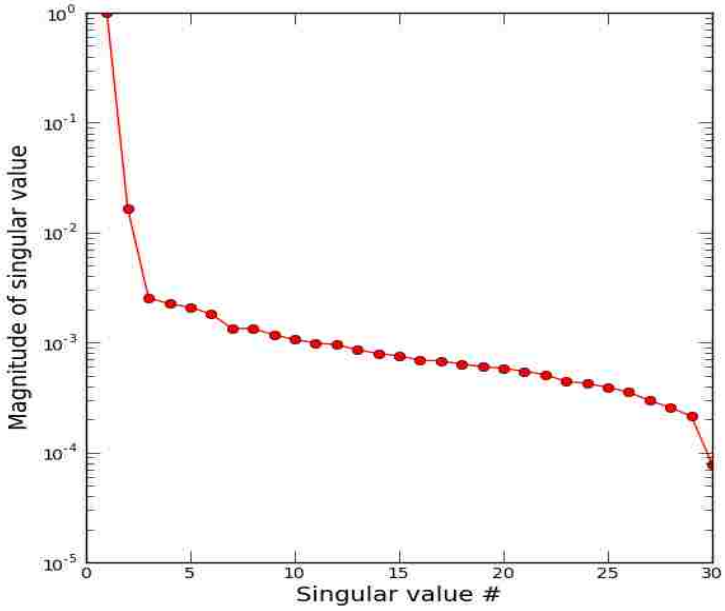


Fig. 4.11: Singular Values of the SRI Matrix

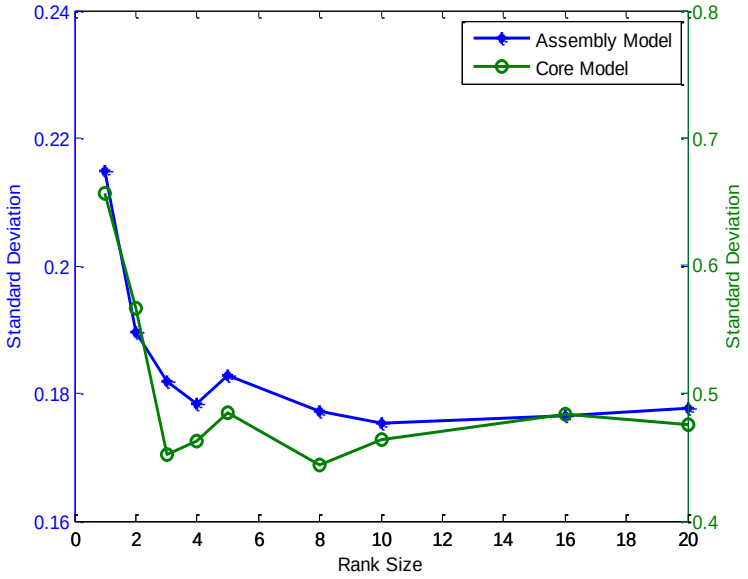


Fig. 4.12: Variance Reduction Sensitivity to Rank Estimate.

4.1.2 GP Method Performance

4.1.2.1 Assembly Model

The 7x7 BWR assembly model introduced in the previous chapter is employed to compare the GVR performance of the SUBSPACE and the GP methods as a preliminary case study. The responses employed for the sake of comparison are the spatial thermal fluxes and various reaction rates densities at the 49 array locations. The run time for all adjoints and forward deterministic calculations were in the order of seconds. Therefore, the Monte Carlo model is found to dominate the total execution time which was in the order of hours. Hence the comparisons were based on the reduction in standard deviations only.

Figure 4.13 compares the performance of the SUBSPACE and FW-CADIS methods in terms of the relative reduction in standard deviation defined by:

$$\varepsilon = \frac{\sigma_i^{\text{Subspace}} - \sigma_i^{\text{FW-CADIS}}}{\sigma_i^{\text{FW-CADIS}}} \times 100$$

Fig. 4.13 shows that the SUBSPACE method consistently reduces the standard deviation everywhere in the assembly model by approximately 40-50% as compared to the FW-CADIS method. If the same standard deviation is to be reached by both methods, a 50% reduction translates into 4 times reduction in the number of histories required by the MC's law of large numbers.

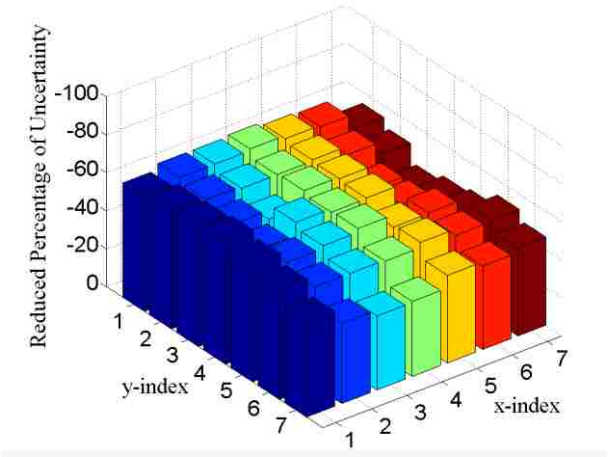


Fig. 4.13: Standard Deviation Comparison for Thermal Fission Reaction Rate Density

In a similar manner, Fig. 4.14 compares the performance of the SUBSPACE and GP methods by plotting the quantity:

$$\varepsilon = \frac{\sigma_i^{\text{GP}} - \sigma_i^{\text{Subspace}}}{\sigma_i^{\text{Subspace}}} \times 100$$

Results indicate that both the SUBSPACE and GP methods are close in performance.

In our case study, the run time for all adjoints and forward deterministic calculations were in the order of seconds. The Monte Carlo model is found to dominate the total execution time which was in the order of hours. Hence for this proof of principle study, comparisons were based on the reduction in standard deviations only.

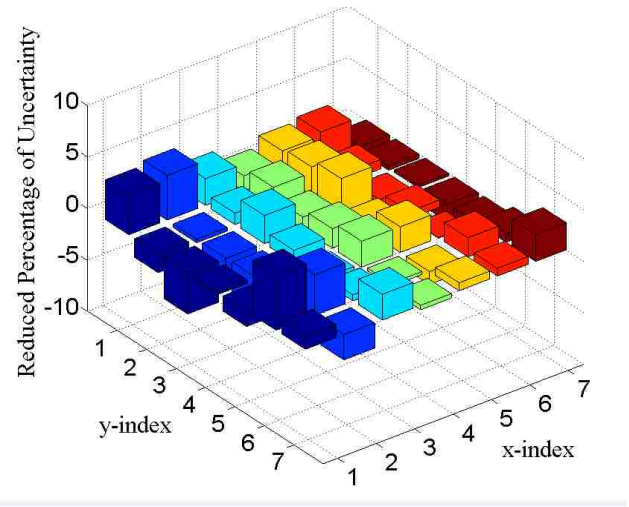


Fig. 4.14: Standard Deviation Comparison for Thermal Flux

4.1.2.2 Core Model

For the purpose of comparison between SUBSPACE and GP methods, the responses of the spatial thermal and fast scalar fluxes at the 289 assembly locations in the full core are employed. Figs. 4.15 and 4.16 compare the performance of the SUBSPACE and GP methods by plotting the quantity:

$$\varepsilon = \frac{\sigma_i^{\mathrm{GP}} - \sigma_i^{\mathrm{Subspace}}}{\sigma_i^{\mathrm{Subspace}}} \times 100$$

as similarly applied for the BWR assembly model. σ_i^{GP} and $\sigma_i^{\mathrm{Subspace}}$ are the standard deviations for the i th assembly tally in the full core, respectively.

Fig. 4.15 shows how the reduction of uncertainty varies for each assembly in the full core. Fig. 4.16 presents more information on the uniformity of the reduction throughout the core. It could be seen from Fig. 4.15 that in interior assemblies the reduction of uncertainty varies

from 0 up to 20% for both thermal group and fast group, while in boundary assemblies, the reduction of uncertainty changes from 0 down to -40% for thermal and fast group. It implies that GP method performs slightly better for boundary assemblies and SUBSPACE method has more advantages for interior assemblies. From Fig. 4.16, it is shown that for both boundary assemblies and interior assemblies, the reduction of uncertainty is of decent uniformity in each energy group. Generally, results have well indicated that the SUBSPACE and GP methods are comparatively close in performance for both thermal group and fast group in the PWR full core model. This is similar to the case study result of the BWR assembly model.

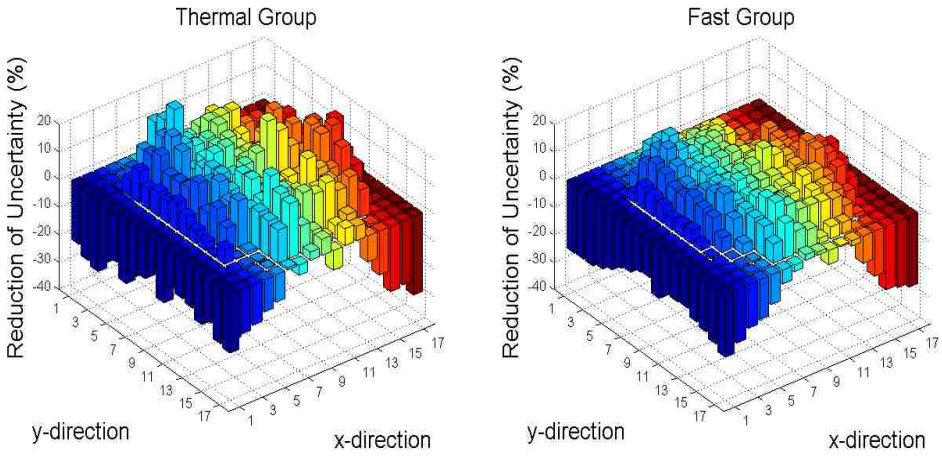


Fig. 4.15: Thermal and Fast Scalar Fluxes for Full Core

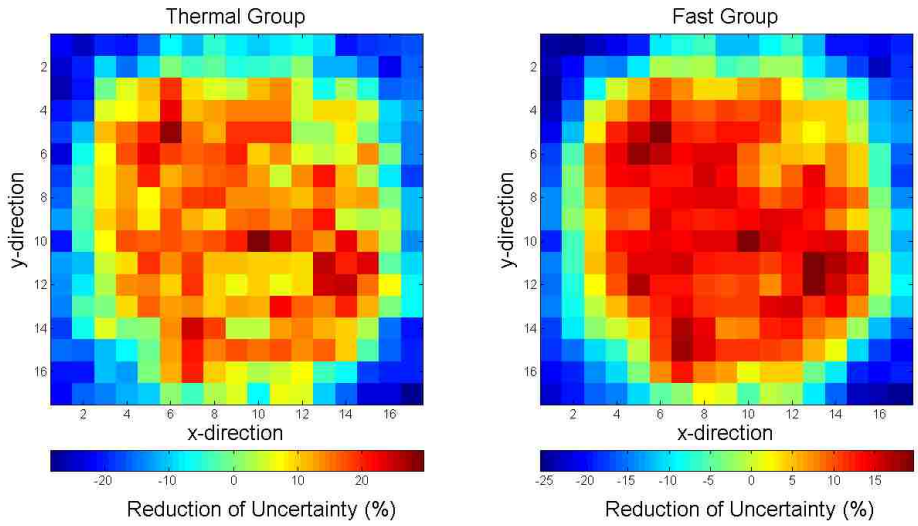


Fig. 4.16: Thermal and Fast Scalar Fluxes for Full Core

4.2. SUBSPACE Method Applied in Eigenvalue Calculations

4.2.1 Monte Carlo in Criticality Calculations

While the Monte Carlo method has been widely applied and demonstrated to be efficient in radiation transport, its realm of applicability for reactor analysis is typically limited, for example, to benchmark deterministic results at certain limited state-points for certain limited responses of interest. This is due to the restriction of the prohibitive computational requirements for obtaining a fully converged system-wide solution. Among all the associated challenges, the most significant challenge is that in actual reactor analysis where millions of different material, tally regions resolving regions of different fuel irradiation, and temperature distributions are involved, the estimates from the MC simulation result in a slow, non-uniform convergence. Concerning these difficulties, it is important to enable direct MC

application for core-level reactor analyses. In [52], it states that the current research has focused on two areas: (1) Hybrid Monte Carlo-deterministic methods for obtaining high-precision fluxes throughout the phase space in k-eigenvalue problems. (2) Efficient Monte Carlo domain-decomposition algorithm to solve the problem with multiple processors for massively parallel systems. In this work, we focus on the first area as mentioned above.

Monte Carlo methods have been applied in computing the fundamental mode eigenfunction of critical systems since the 1950s [53~56]. Many MC codes use the standard power method for computing Keff and solving k-eigenvalue problems such as MCNP [6] and KENO in the SCALE6.1 code package [25]. In these MC simulations, each iteration cycle corresponds to one single fission generation [51]. Power iteration is performed till the convergence is obtained for both Keff and the source distribution. After the convergence, responses of interest .e.g. flux and reaction rates are accumulated. While this power iteration calculation is becoming routine in standard MC codes, three essential limitations restrict the fidelity of the calculation: (1)The initial source guess tends to influence the obtained results, therefore it is important to ensure sufficient initial cycles are discarded prior to the tally calculation. (2) In each cycle, it must be guaranteed that there are sufficient number of neutron particles, otherwise the bias in Keff and tallies of desired response will contaminate the result; (3) Bias in the uncertainties on Keff and responses of interest must be taken into account [51].

To address the above limitations in eigenvalue calculation, a trial run is usually performed to determine the number of cycles to be discarded before tally calculations. For long production runs, a minimum 5000 neutrons per cycle should be applied to prevent bias in

Keff and tallies of response of interest. Furthermore, symmetric geometry is preferred as taking advantage of symmetry will reduce the bias in uncertainties as well as accelerating the convergence. For example, a quarter core model with reflective boundaries proves more efficient in reducing the problem dominance ratio than a full core model [51].

In this chapter, the SUBSPACE method is extended and applied to enable global variance reduction for reactor analysis which require precise responses to be evaluated everywhere in the phase space by performing the k-eigenvalue simulation. A numerical experiment based on a quarter-core PWR model is discussed to compare the SUBSPACE and FW-CADIS methods in terms of the reduction in standard deviation of spatially distributed responses. As stated in Chapter 3, the SUBSPACE method belongs to the same adjoint-based family as FW-CADIS method. It identifies the correlations between weight window maps to minimize the computational time required for global variance reduction. The correlations are employed to reduce the number of maps required to achieve the same level of variance reduction that would be obtained with single-response maps.

4.2.2 Implementation of the SUBSPACE Method

In its standard form, the k-eigenvalue transport equation is written as:

$$[\Omega\cdot\nabla+\sum\nolimits_T(\vec{r},E)]\psi(\vec{r},E,\Omega)=\iint\psi(\vec{r},E',\Omega')\sum\nolimits_S(\vec{r},E'\to E,\Omega\cdot\Omega')d\Omega'dE' + \frac{1}{k_{eff}}\frac{\chi(E)}{4\pi}\iint\nu\sum\nolimits_F(\vec{r},E')\psi(\vec{r},E',\Omega')d\Omega'dE'$$

It could be simplified as:

$$\mathrm{L}\vec{\psi} = \frac{1}{k_{eff}} F\vec{\psi} \ ,$$

which is the problem usually solved in standard MC codes.

Weight window based Monte-Carlo deterministic hybrid methods employ an approximate adjoint function tailored to the responses of interest:

$$u = \left\langle \vec{\psi}, \vec{\sigma} \right\rangle \tag{76}$$

to assign a permissible range of weights to various regions in the phase space: If the particle weight is below the range it is rouletted and if it is above the range it is split. In global variance reduction responses are desired everywhere, i.e.

$$u_i = \left\langle \vec{\psi}, \vec{\sigma}_i \right\rangle, \text{ and } i = 1, ..., I \tag{77}$$

such that a variance reduction scheme needs to develop an adjoint function accommodating all responses. The FW-CADIS method, introduced in previous chapters, has been successfully extended to address the k-eigenvalue problems [52]. This is realized by: 1. A forward eigenvalue deterministic calculation is performed to obtain a precise estimate of the space-energy-dependent fluxes; 2. A subsequent fixed-source deterministic adjoint calculation is conducted where the adjoint sources are weighted by the fluxes from step 1. As stated before, with the FW-CADIS method, the adjoint sources (one for each response) are weighted by the inverse of the forward quantity being sought, i.e. the following forward eigenvalue problem and the corresponding adjoint problem are solved:

$$L\vec{\psi} = \frac{1}{k_{eff}} F\vec{\psi} \Rightarrow u_i = \langle \vec{\psi}, \vec{\sigma}_i \rangle \tag{78}$$

$$L^*\vec{\psi}^* = \sum_{i=1}^{I} \frac{1}{u_i} \frac{\partial u_i}{\partial \vec{\psi}} = \sum_{i=1}^{I} \frac{\vec{\sigma}_i}{u_i}, \tag{79}$$

where the standard transport theory notations from previous chapters are used.

Meanwhile, the SUBSPACE method has also been extended to address the eigenvalue problems. In contrast to the FW-CADIS, the SUBSPACE method generates adjoint functions for pseudo-responses that are linear combinations of the original responses:

$$\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} u_i \tag{80}$$

The linear combinations identify the minimum number of pseudo single-responses required to reach the same level of variance reduction that would be achieved if all single-response weight window maps are employed to reduce variances for all responses. Using the definition for $\tilde{u}_j$ from Eq. (80), one can write:

$$\tilde{u}_j = \sum_{i=1}^{I} \eta_{i,j} \langle \vec{\psi}, \vec{\sigma}_i \rangle = \left\langle \vec{\psi}, \sum_{i=1}^{I} \eta_{i,j} \vec{\sigma}_i \right\rangle \tag{81}$$

$$L^*\tilde{\vec{\psi}}_j^* = \frac{\partial \tilde{u}_j}{\partial \vec{\psi}} = \sum_{i=1}^{I} \eta_{i,j} \vec{\sigma}_i \tag{82}$$

The adjoint flux map utilized for the SUBSPACE method is generated using the three-dimensional discrete ordinates code Denovo [49].

Similar to the FW-CADIS method, the optimization objective of the SUBSPACE method is the group-wise fluxes in fissionable regions instead of the eigenvalue[52]. Different from the

FW-CADIS method, instead of solving a forward eigenvalue problem, the SUBSPACE method solves a single adjoint fixed-source problem with the adjoint source being determined as derivatives of the pseudo responses with respect to the adjoint function. The adjoint function is then employed to compute the upper and lower weight window bounds and subsequently the weight window map is written in a format suitable for MCNP [6]. The Monte-Carlo computations are all performed using MCNP.

4.2.3 PWR Quarter Core Model

A three-dimensional quarter core PWR model is employed to compare the performance of the FW-CADIS and the SUBSPACE methods. The PWR quarter core model features a generic three-dimensional layout. The x-y-z dimensions are 204.25x204.25x335.28cm.

The model consists of $48\frac{1}{4}$ 17x17 fuel assemblies, with 264 fuel rods per assembly and 3% uniform fuel enrichment. The adjoint fixed-source Denovo calculations use an S4 level symmetric quadrature and an 461 × 461 × 10 spatial grid resolving the unit-cells. The 27 neutron and 19 photon energy group libraries included in the SCALE package are employed for generating the adjoint flux maps. For the flux and reaction rates responses, the first 14 neutron groups (10.678eV<E<20MeV) define the fast group and the last 13 groups (E<3.059eV) are thermal. A cross sectional view of the model is presented in Fig.4.17, where green represents moderator and reflector (water) and red the fuel pins.

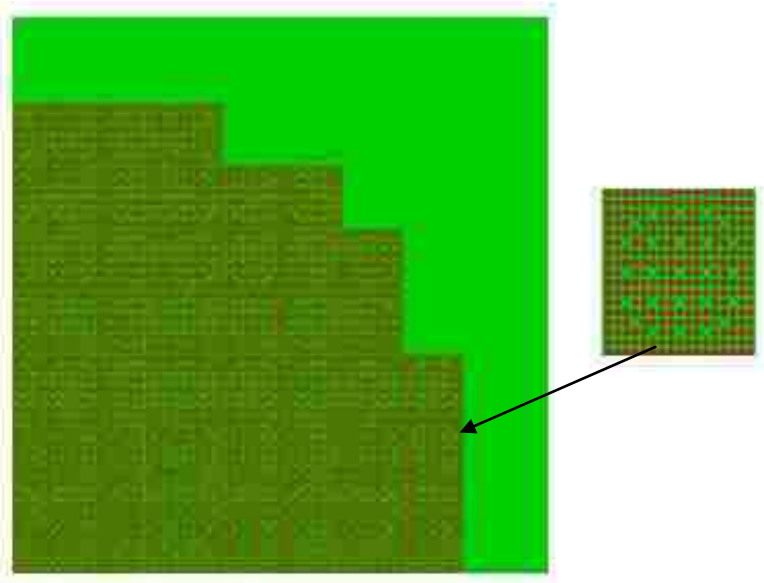


Fig. 4.17: A Cross Section View of the 3-D PWR Quarter Core Model

Continuous-energy MCNP5 simulations are conducted using 50,000 histories/cycle, 2500 active cycles with 500 inactive cycles starting from an initially uniform fission source. FW-CADIS and the SUBSPACE methods are applied respectively and a thermal flux energy binfrom 0.15 to 0.275ev is selected to compare the performance of two methods. The obtained results of the relative uncertainty associated with the thermal flux, for each method, are clearly illustrated along the z-dimension from -150.876cm to 150.876cm for each layer (1^{st} ~10^{th}) in the model as shown in Figs. 4.18~4.27. The colorbar identifies the percentage of relative uncertainty.

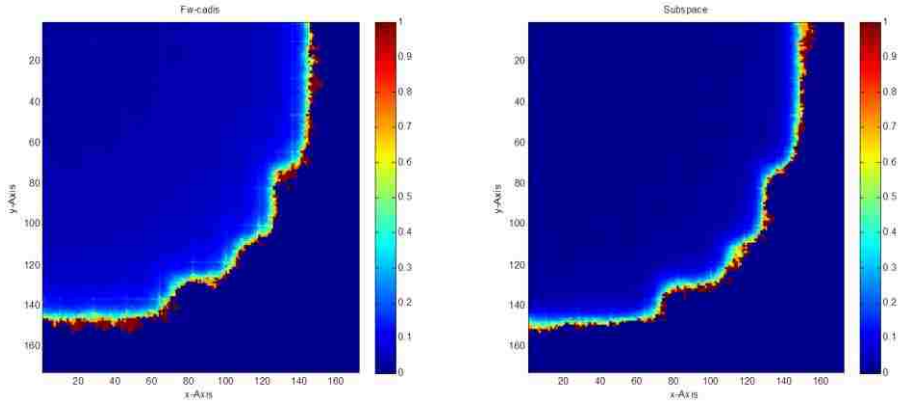


Fig. 4.18: Relative Uncertainty Distribution of Thermal Flux for 1st layer

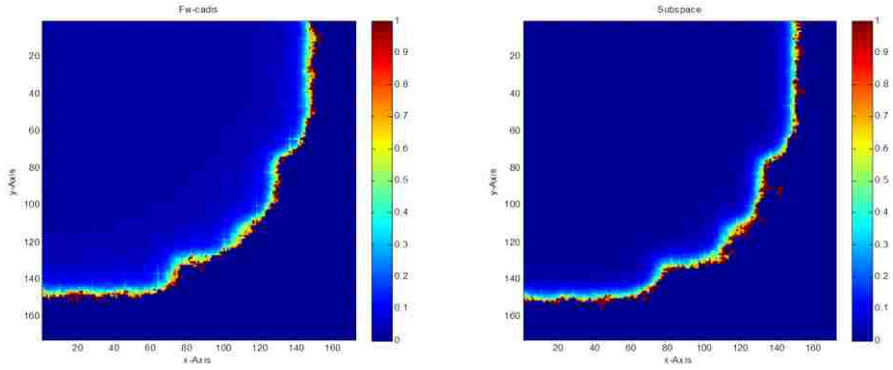


Fig. 4.19: Relative Uncertainty Distribution of Thermal Flux for 2nd layer

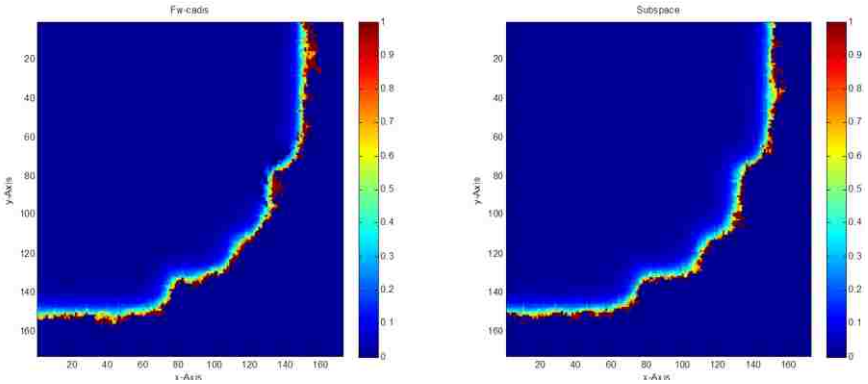


Fig. 4.20: Relative Uncertainty Distribution of Thermal Flux for 3rd layer

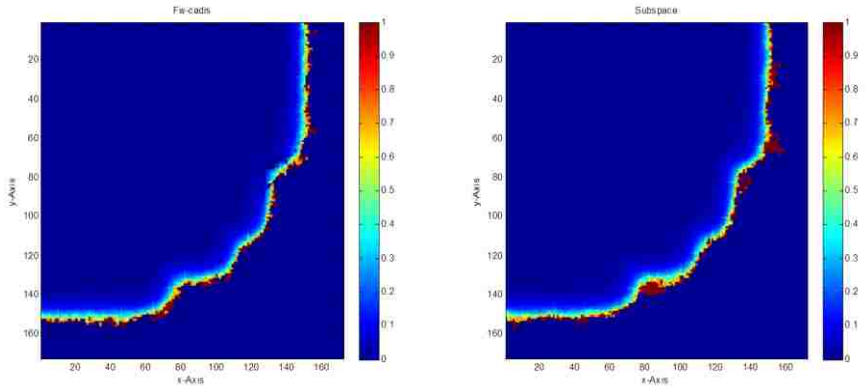


Fig. 4.21: Relative Uncertainty Distribution of Thermal Flux for 4thlayer

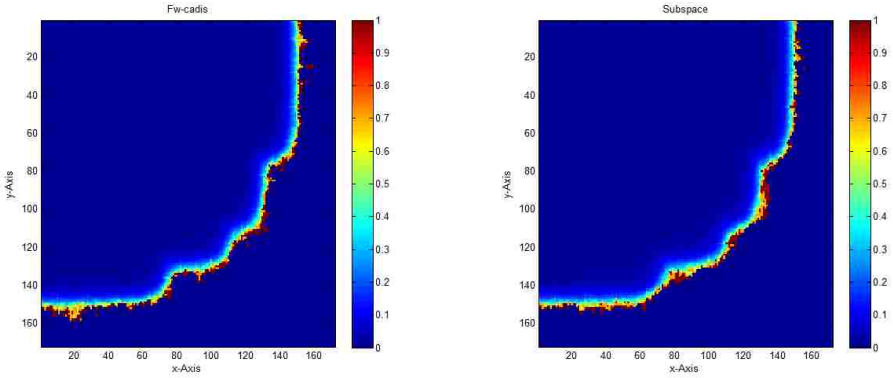


Fig. 4.22: Relative Uncertainty Distribution of Thermal Flux for 5th layer

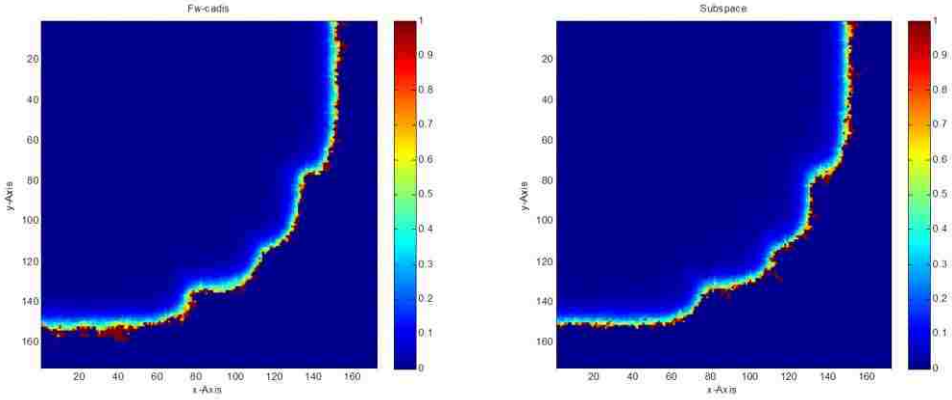


Fig. 4.23: Relative Uncertainty Distribution of Thermal Flux for 6th layer

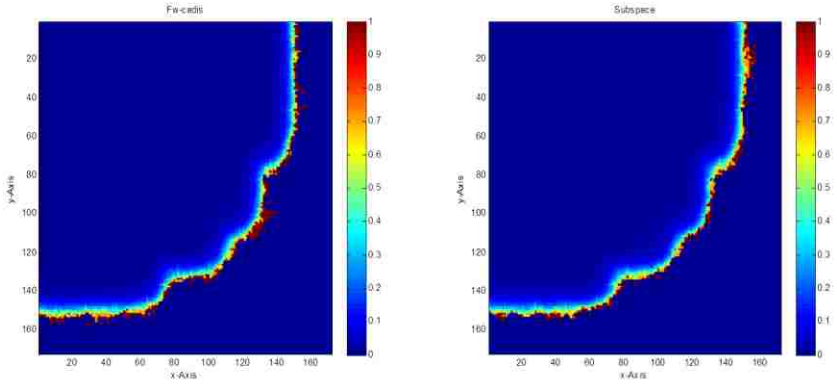


Fig. 4.24: Relative Uncertainty Distribution of Thermal Flux for 7th layer

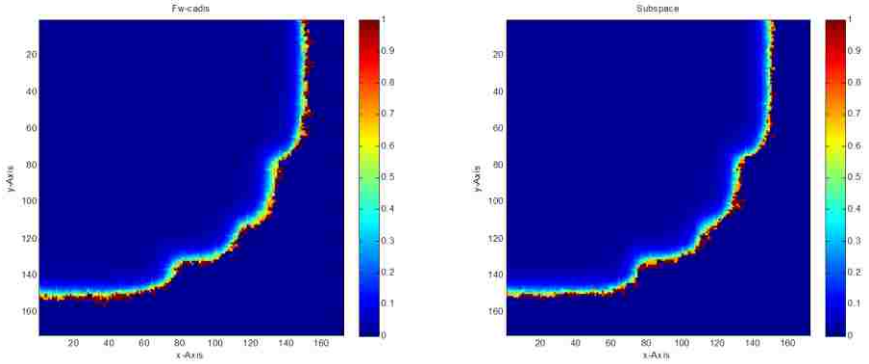


Fig. 4.25: Relative Uncertainty Distribution of Thermal Flux for 8th layer

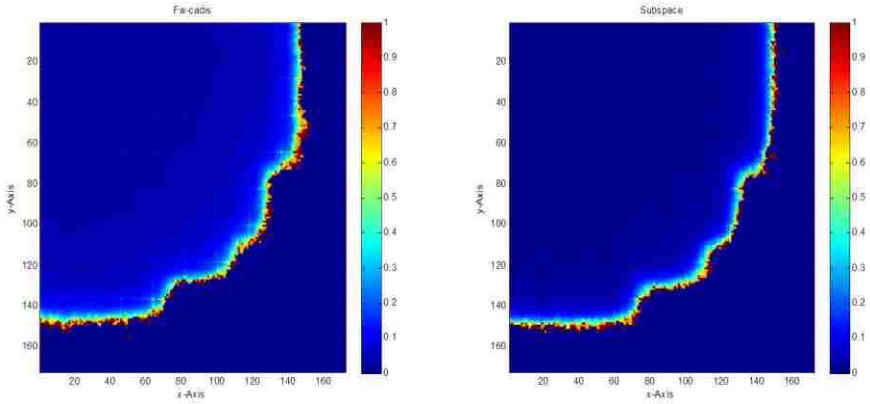


Fig. 4.26: Relative Uncertainty Distribution of Thermal Flux for 9th layer

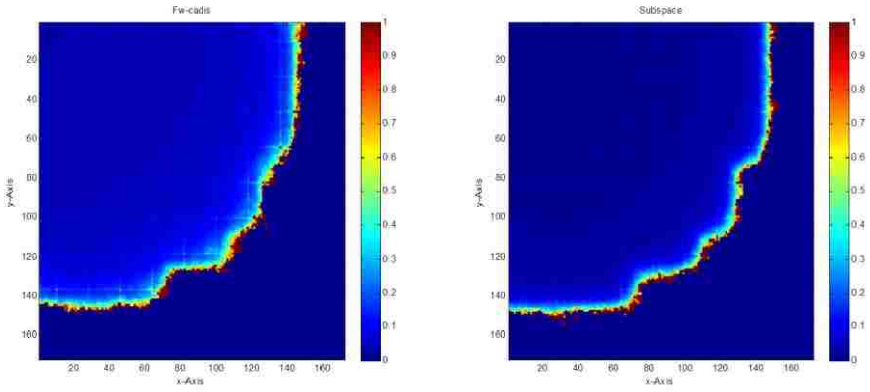


Fig. 4.27: Relative Uncertainty Distribution of Thermal Flux for 10th layer

The distribution of relative uncertainties for thermal flux in the given energy bin is plotted in the Fig. 4.28 for the fifth layer of the model. Only the cells within the reactor core (i.e. excluding the reflector) are taken into account. It is clearly shown that both the FW-CADIS and the SUBSPACE method obtain the same number of cells (9000~10000) with relative uncertainty between 0~0.2%.

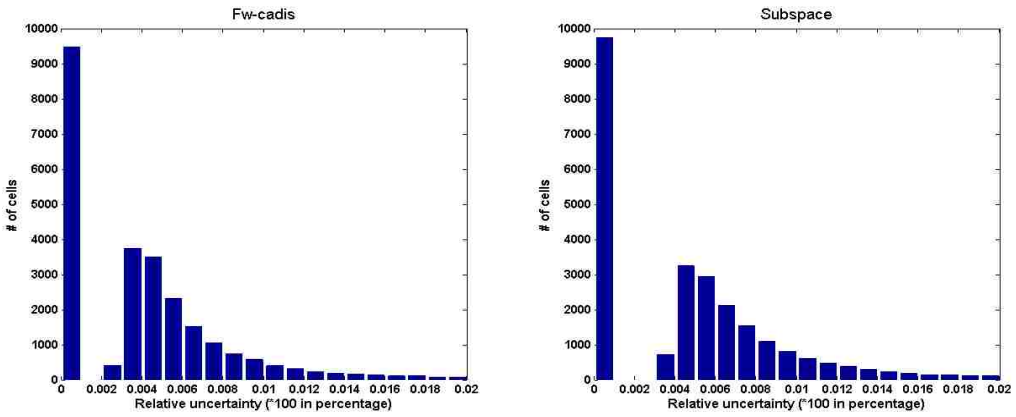


Fig. 4.28: Histogram of Relative Uncertainty Distribution

To make the comparison more straight-forward, the standard deviation and the mean value of the distribution of variance obtained with FW-CADIS and SUBSPACE method respectively are calculated as shown in Table 4.4. For reference, the analog results are also provided. When applying the SUBSPACE method, the obtained mean value of variance is 0.0181; while the mean value of variance obtained by performing FW-CADIS method is 0.0346, which is 91% higher. The SUBSPACE method also generates a standard deviation of the variance distribution which is 48% lower compared to the same quantity generated by FW-CADIS. Based on these two quantities, it is safe to claim the SUBSPACE method obtains responses of interest of higher accuracy compared to the FW-CADIS method.

Table 4.4: Standard Deviation and Mean Value of the Variance Distribution

	Stdev of variance	Mean of variance
FW-CADIS	0.1503	0.0346
SUBSPACE	0.1017	0.0181
Analog	0.2145	0.0588

In order to compare the efficiency, the global FOMs of the two methods need to be calculated. The execution time of each method is listed in Table 4.5. For the Denovo computation, since the SUBSPACE-based cases run in parallel on multiple processors, the total time is taken as the time associated with the longest run among the 30 parallel jobs.

This follows from the fact that the user cannot proceed with his analysis unless all parallel jobs finished such that the longest running job limits the user's ability to accelerate the analysis procedure. This is based on the definition of parallel computing [8]. It is worth mentioning that the Denovo time consumed by applying the SUBSPACE method is about 1/30 of the Denovo time by applying the FW-CADIS method. This could be explained by two reasons: 1.The parallelization associated with applying the SUBSPACE method enhances the computational efficiency; 2. FW-CADIS performs a forward eigenvalue calculation as the basis of solving the adjoint problem. However, in realistic and complicated core-level systems with millions of cells, the forward eigenvalue calculation generally takes a significant amount of time and therefore the total computational efficiency deteriorates.

The global FOM is calculated as the inverse product of the mean value of variance distribution and total execution time, which is the sum of the Denovo execution time and the MCNP execution time:

$$FOM = \frac{1}{v(t_{Denovo} + t_{MCNP})}$$

For the analog run, the total time equals to the MCNP execution time. The global FOM when applying the SUBSPACE method is 0.0649 while the global FOM of the FW-CADIS method is 0.0071. The analog MCNP results in a global FOM value of 0.000698.

Therefore, the SUBSPACE method shows a speed-up of 9 over the FW-CADIS method and a speed-up of 93 over the analog.

Table 4.5: Execution Time and Global FOM

Execution Time				FOM
	Denovo	MCNP	Total	
SUBSPACE	90.83min	761.76min	852.59min	0.0649
FW-CADIS	2883.05min	1208.91min	4091.96min	0.0071
Analog	N/A	24364.05min	24364.05min	0.000698

CHAPTER 5

CROSS SECTION FUNCTIONALIZATION

5.1 Cross Section Functionalization

First a brief overview is given on the standard reactor physics methodology and short description is provided of the GVR methodology employed in this work.

The routine execution of a detailed model for the entire reactor core which resolves all the fuel pins details in all the fuel assemblies is still considered challenging even with access to leadership computing resources. To overcome this challenge, reactor physicists devised a multi-level homogenization method to help reduce the dimensionality of the problem. In this method, the reactor core domain is divided up into smaller regions, often chosen to represent full or parts of a fuel assembly taken at different axial levels. Lattice physics (or assembly) calculations are used to analyze these regions in more detail, often done with many energy groups and fine spatial and angular mesh. The flux solution from lattice physics calculations is used to generate cross-sections that are homogenized over coarser energy groups and spatially over each region in a manner that preserves reaction rates over the various regions. The few-group cross-sections are then used in core-wide simulation where the geometrical, energy, and spatial details of the regions are now smeared which reduces the effective dimensionality of the core-wide problem.

Given the dependence of the flux solution from lattice physics calculations on the isotopic concentration, the fuel temperature, the coolant temperature and/or voiding, presence of poison in the coolant, amount of control rod insertion, etc., the few-group cross-sections must

be generated at a matrix of different conditions to enable interpolation of the correct value for core-wide simulation. This is a formidable task as for typical LWR models the number of these conditions is overwhelmingly large.

Take for example a BWR model: One typically has in the order of 30 lattice designs, each depleted using lattice physics calculations to end of life with about 50 depletion steps. This is often repeated with 3 different voiding histories, e.g. no voiding, medium voiding, and high voiding. This is important as the increased voiding affects the spectrum and subsequently the depletion characteristics. At each depletion step, about 5 different branch calculations are executed. In each branch calculation, one parameter, e.g. fuel temperature increase or decrease, control rod insertion, etc., is changed and another flux solution is obtained. The total number of flux solutions for a typical BWR is 30 x 50 x 3 x 5 = 22500. If each flux solution takes in the order of few seconds, which is possible with highly customized commercial codes, these calculations can be completed over a short period of time.

With MC models however, unless one has a reasonably fast convergence scheme, the use of MC would be infeasible for routine reactor physics calculations.

To address this challenge, variance reduction techniques have been developed to accelerate Monte Carlo convergence. The idea is that if one has an approximate idea about the solution, one can use that knowledge to bias MC particles. For adjoint-based variance reduction techniques, which represent our current interest, a simplified deterministic model is used to calculate an adjoint flux for the response of interest, say a detector response placed somewhere in the reactor core. Given that the adjoint flux can be shown mathematically to

describe the importance of particles at different points in the phase space, one can design weight window maps based on the adjoint information to bias MC particles. This is done by splitting particles that are important and playing Russian roulette with particles that are less important. The idea has been successfully demonstrated in the FW-CADIS methodology, which generalizes the idea of variance reduction to problems with global responses, i.e., that is when responses are desired everywhere in the phase space. This is done by employing an additional deterministic forward flux solution to assign more weight to regions with low flux and less weight to regions with high flux which renders a uniform variance reduction over all responses of interest. An assembly model represents such an example where the flux is required everywhere in the assembly to properly homogenize the cross-sections.

The SUBSPACE method, as introduced in previous chapters, is applied to perform GVR with three primary advantages over existing FW-CADIS methodology. First, the forward flux solution is not required, which results in considerable time savings especially for eigenvalue problem with dominance ratio close to unity. Second, via the use of the pseudo responses, representing random linear combinations of the original responses, the number of MC

particles required to reach a given level of variance reduction is significantly reduced. Finally, the method allows one to split the total number of MC particles over multiple trains of MC simulation, which improves the efficiency of the methods by allowing one to take advantage of parallel computing environment.

In previous chapters, the figure of merit (FOM) of the SUBSPACE method is calculated, which was found to be in the range of 2-10 times faster than the FW-CADIS method. The

lower range for the gain is for source-driven problems with small dimensionality in terms of the responses and the complexity of the geometry and energy details. The gain increases as the dimensionality of the model increases and reaches its maximum particularly for core-wide eigenvalue problems.

5.2 Implementation of the SUBSPACE Method

The goal of this exercise is to use Monte-Carlo methods to generate homogenized few-group cross sections thus replacing the standard deterministic lattice codes used today. The responses of interest in this task are homogenized and collapsed cross sections given by the general form:

$$\sum_{x,g} = < \phi_g, \sigma_{x,g} > / \phi_g \ ,$$

where x denotes the reaction type (x=fission, capture, scattering). However, using this response to directly bias particles via the adjoint methodology would be difficult.

Following the methodology, the derivative of the response with respect to the flux is used to construct the adjoint source. However, the derivative of the homogenized cross section with respect to the flux can be negative which might lead to negative adjoint sources. Negative sources are, however, are bound to lead to problems with the biasing procedure. Therefore, we tailor an adjoint source for the all reaction rates:

$$R_x = \langle \phi, \sigma_x \rangle$$

which homogenized cross sections are desired for. In addition we add the flux because it appears in the denominator of the cross section thus bears significance for its accuracy. The global aspect of the GVR problem now does not only lie in the spatial extent of the problem but also to simultaneously optimize various reactions rates and the flux.

The proposed GVR method: the SUBSPACE method, is based on a mathematical method that takes advantage of the correlation between the various responses. Mathematically, this is realized by generating adjoint functions for pseudo-responses which are random linear combinations of the original responses:

$$\tilde{u}_j = \sum_{i=1}^{I} w_{i,j} u_i \tag{83}$$

The linear combinations identify the minimum number of pseudo responses required to reach the same level of variance reduction that would be achieved if all single-response weight window maps are employed to reduce variances for all responses one at a time.

The pseudo response would correspondingly be described as:

$$\tilde{u}_x = \sum_g \sum_x w_{x,g} < \phi_g, \sigma_{x,g} > \tag{84}$$

where x now stands for x=fission, capture, scattering and flux (note: sigma_flux = 1). The adjoint source is equal to:

$$L^* \tilde{\phi}_x^* = \frac{\partial \tilde{u}_x}{\partial \phi} = \sum_g \sum_x w_{x,g} \sigma_{x,g} \tag{85}$$

Therefore, given the desired problem-specific cross section as the response of interest, the pseudo response is constructed as a linear combination of weighted original cross section responses from given libraries. For the SUBSPACE method, the weighting factors are randomly sampled and the algorithm is applied to obtain the optimized number of correlations between responses. For the FW-CADIS method, the weighting factors are obtained from a forward calculation and calculated as the inverse of the response, in this case: the reaction rate.

5.3 PWR Assembly Model

The current reactor physics analysis comprises a stage in which the transport equation is solved for a single assembly. The goal of this stage is to obtain collapsed and homogenized cross sections sets for a subsequent core-wide coarse mesh diffusion calculation. In order to compare the performance of Monte-Carlo and deterministic calculations, macroscopic cross-sections are homogenized over the entire assembly for a simplified PWR assembly model. In this work, the deterministic calculation is done by NEWT (New ESC-based Weighting Transport code) from SCALE 6.1 code package by ORNL [60].

The NEWT computer code is a multi-group discrete-ordinates radiation transport code with flexible meshing capabilities that enables two-dimensional neutron transport calculations using complex geometric models. In particular, NEWT proves useful in eigenvalue and source calculations as well as in preparing collapsed, flux-weighted cross sections in AMPX working library format. Compared to Denovo, NEWT allows more flexibility in defining boundary conditions. Furthermore, NEWT has also been incorporated into the TRITON

sequence of SCALE6.1 [61] serving as a flux-solver for depletion calculations. When used in depletion mode under TRITON, which is a two-dimensional transport and depletion module for characterization of spent nuclear fuel, NEWT could be used to generate lattice-physics cross sections for subsequent core calculations. The Monte-Carlo calculations are performed by MCNP5. Both the analog and the GVR (global variance reduction) technique based Monte-Carlo simulations are performed. The GVR technique applied in this work is the SUBSPACE method.

The assembly model employed is a PWR UO2 fuel assembly featuring 17x17 pin cells as shown in Fig. 5.1. Red stands for moderator, green stands for UO2 fuel and blue stands for UO2-Gd2O3 fuel. The assembly pitch is 21.505cm. The fuel rod pitch is 1.265cm and the thickness of cladding is 0.064cm. The outer diameter of the pellet is 0.824cm. The inner and the outer diameter of the cladding are 0.824cm and 0.952cm respectively. The gap in-between fuel and cladding is ignored. The average U235 enrichment is 6.2 wt%. The assembly is composed of UO2 and UO2-Gd2O3 (Gd) fuel rods. The assembly model is simulated under the hot condition with the pellet temperature at 900 K and the moderator temperature at 600 K [35].

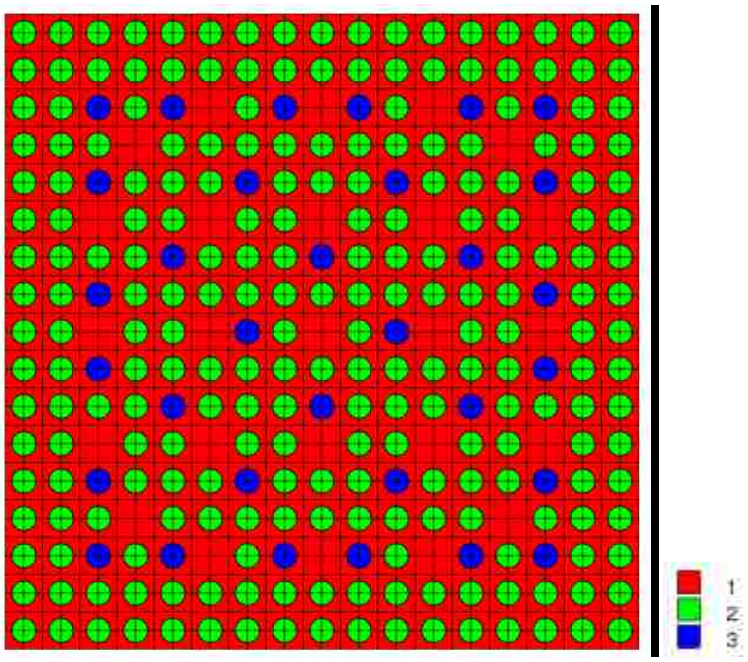


Fig. 5.1: PWR Assembly Model

Collapsed (two-group format) and homogenized fission, capture and total cross sections are prepared using (1) NEWT (2) analog MCNP and (3) SUBSPACE MCNP in a two-group format. The thermal group ranges from 0 through 0.625eV and the fast group comprises the energy range above 0.625eV.

For the analog Monte-Carlo and the deterministic calculation, the obtained cross section data values are very close to each other as shown in Table 5.1. The difference in percent of the obtained cross section values range from 0.6423% (total cross section for the fast group) to 6.3818% (fission cross section for the thermal group) as listed in Table 5.2.

For the analog Monte-Carlo and the SUBSPACE-based Monte-Carlo calculation, the obtained cross section data values are nearly identical to each other (Table 5.1).

The discrepancy of cross section data presented in Table 5.2 varies from 0.0031% (fission cross section for the fast group) to 0.0330% (capture cross section for the thermal group). These results demonstrate that a much stronger resemblance exists between the analog Monte-Carlo and the variance technique based Monte-Carlo compared to the analog Monte-Carlo and the deterministic calculations. This could be explained by the error naturally associated with the deterministic calculation. The analog Monte-Carlo simulation completes 8000 cycles, 20000 histories per cycle in 1657.8 minutes while the SUBSPACE-based Monte-Carlo simulation completes the same number of histories in 178.3 minutes. Given that the uncertainties associated to the cross section values in both cases are below 0.1% and could be safely considered at the same level, the SUBSPACE method shows a speed-up of 9.3. It should be noticed that the reason why both the analog and the biased Monte-Carlo simulation obtain the same level of uncertainty is because in this work, the response of interest is one single response (homogenized cross section) and consequently every particle counts towards the desired response. Therefore the analog Monte-Carlo performs alright in this case. In previous chapters, it has been shown that in a more complicated quarter-core model, given the thermal flux everywhere as the response of interest, the SUBSPACE method could achieve a speed-up that is close to 100 over the analog Monte Carlo.

Table 5.1: Cross Section Data

Analog_MCNP XSection			
	Total	Fission	Capture
Fast group	0.3337	0.0039	0.0067
Thermal group	0.8300	0.0891	0.0665
SUBSPACE_MCNP XSection			
	Total	Fission	Capture
Fast group	0.3337	0.0039	0.0067
Thermal group	0.8299	0.0890	0.0665
NEWT XSection			
	Total	Fission	Capture
Fast group	0.3316	0.0038	0.0068
Thermal group	0.8224	0.0834	0.0645

Table 5.2: Difference compared to Analog in percentage (%)

SUBSPACE_MCNP XSection			
	Total	Fission	Capture
Fast group	0.0034	0.0031	0.0088
Thermal group	0.0081	0.0176	0.0330
NEWT XSection			
	Total	Fission	Capture
Fast group	0.6423	2.5519	2.2611
Thermal group	0.9115	6.3818	2.9869

The PWR assembly is now depleted and the numerical experiment previously performed for the undepleted assembly is repeated. The depletion process comprises a five-cycle depletion case executed by the TRITON sequence in the SCALE code package [61]. The TRITON control module can be used to provide automated, problem-dependent cross-section processing followed by calculation of the neutron multiplication factor for a 2-D configuration using NEWT. Used in conjunction with NEWT, TRITON could perform 2-D lattice calculations for non-traditional lattice designs such as hexagonal arrays. Moreover, TRITON is able to provide reliable transport modeling accuracy such as to predict the burnup of nuclear materials in configurations that have a strong spatial dependence on the neutron flux and other physics parameters characterizing the system.

During the depletion study Triton alternates calls to NEWT and ORIGEN to compute the fluxes everywhere in the assembly and to advance time, respectively, using a predictor-corrector algorithm. Given these abilities, TRITON is chosen as the primary code to perform the depletion simulations in this work.

The depletion study is initialized with the assembly operating at a power level of 21.220 MW/MTHM for a period of 100 days. This is followed by four identical cycles of operation each of a 100-day period. For the depletion purpose, the 289 pin cells are divided into 9 groups based on different material as shown below. The macroscopic cross sections are tallied for each group of pin cells separately.

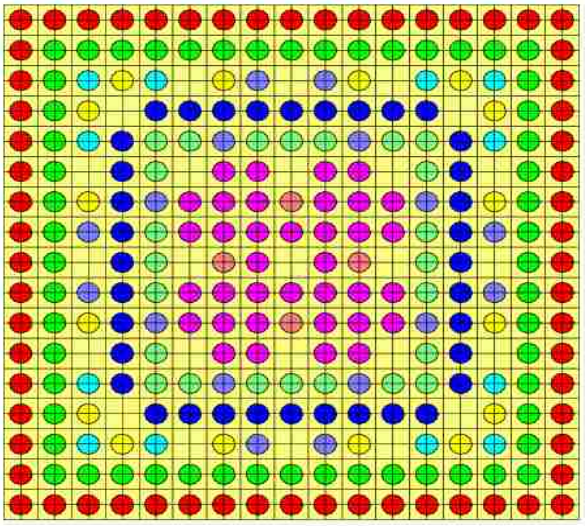


Fig. 5.2: Depletion Pattern of the PWR Assembly Model

The changes of the level of nuclides (Plutonium and Uranium for example) throughout the five-cycle-depletion are also plotted below to show how the isotropic compositions evolve:

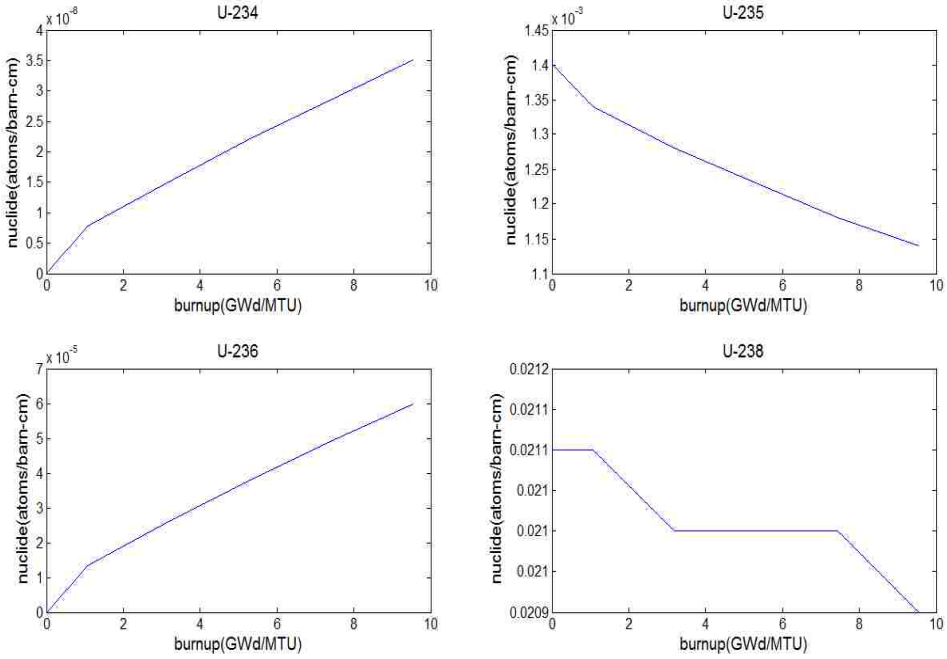


Fig. 5.3: Isotopes of Uranium through the five-cycle-depletion

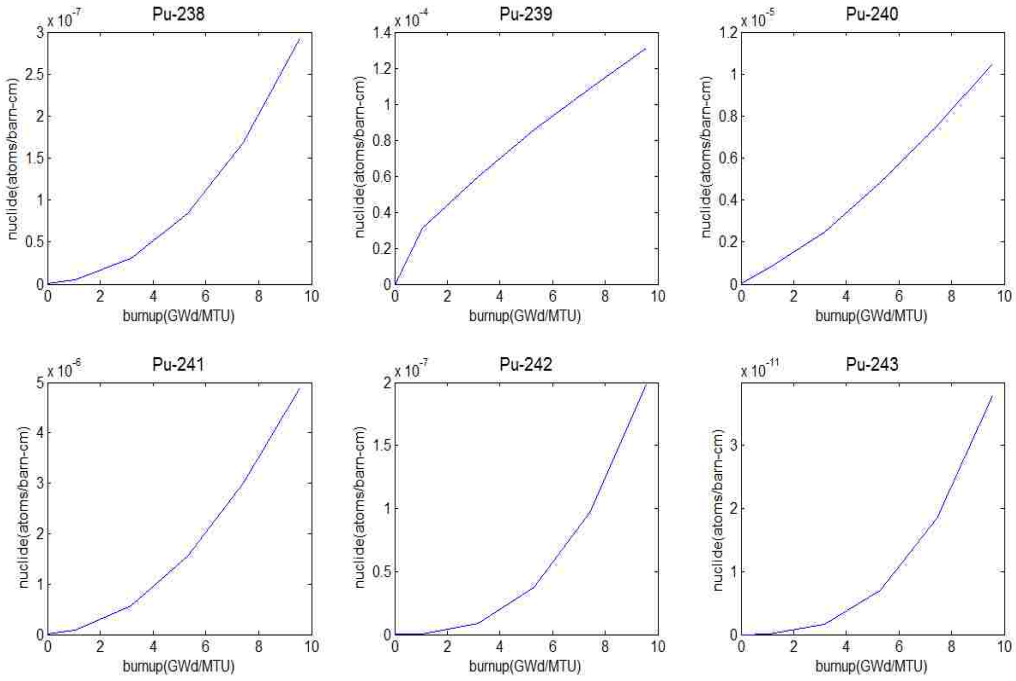


Fig. 5.4: Isotopes of Plutonium through the five-cycle-depletion

Based on the Triton depletion results, the number densities of each nuclide are computed at each separate depletion cycle for the subsequent Monte Carlo calculation. The Monte Carlo calculations are performed using MCNP5; we tally group-wise homogenized macroscopic cross sections for each separate depletion cycle. Both analog and GVR (global variance reduction) accelerated Monte-Carlo simulations are performed. The GVR technique applied in this work is the SUBSPACE method introduced in Chapter 4, which is an adjoint-based technique that utilizes pseudo responses generated with random weights to help identify the correlations between the importance maps and thus reduce the computational time required for global variance reduction.

Utilizing the SUBSPACE method, two types of weight window are employed to compare the efficiency: 1. A general weight window that is generated for the undepleted, homogenized assembly; 2. The optimized weight window that is generated for each depletion cycle specifically. For case 1 the same general weight window is applied for all the 5 depletion cycles, while for case 2 five different optimized weight windows are applied to each corresponding depletion cycle separately.

For both the analog and the biased MCNP runs, 5200 cycles, 20000 histories per cycle are performed. Fission, absorption, scattering and total cross section data are obtained for each depletion cycle, each material group and each energy group (thermal: 0~0.625eV and fast: 0.625ev~20MeV). Comparing the analog and the biased computations, the obtained results are close enough to be safely considered identical. Taken fission cross section for example, as presented from Table 5.3 to 5.6, the maximum difference between the analog results and

the results employing the general weight window is less than 1.5% for the thermal group and less than 0.05% for the fast group. The same proximity is observed between results obtained by employing the general weight window and results obtained by the optimized weight windows. Therefore, it could be claimed that analog, general weight window and optimized weight window calculations generate the same sets of cross section data.

Table 5.3: Difference compared to General Weight Window for Fast Group

	Depletion Cycle 1(%)	Depletion Cycle 2(%)	Depletion Cycle 3(%)	Depletion Cycle 4(%)	Depletion Cycle 5(%)
Material 1	0.0213	0.0096	0.0073	0.0074	0.0425
Material 2	0.0280	0.0024	0.0048	0.0024	0.0025
Material 3	0.0068	0.0162	0.0234	0.0119	0.0288
Material 4	0.0299	0.0281	0.0071	0.0096	0.0122
Material 5	0.0046	0.0564	0.0119	0.0218	0.0123
Material 6	0.0113	0.0092	0.0186	0.0259	0.0286
Material 7	0.0537	0.0640	0.0283	0.0211	0.0210
Material 8	0.0036	0.0214	0.0178	0.0000	0.0035
Material 9	0.0072	0.0143	0.0107	0.0142	0.0035
Material 10	0.0213	0.0096	0.0073	0.0074	0.0425

Table 5.4: Difference compared to General Weight Window for Thermal Group

	Depletion Cycle 1(%)	Depletion Cycle 2(%)	Depletion Cycle 3(%)	Depletion Cycle 4(%)	Depletion Cycle 5(%)
Material 1	1.1806	1.1435	0.9210	0.4037	1.0300
Material 2	1.1711	0.0637	0.2474	1.3985	0.7313
Material 3	1.0180	0.1162	1.0399	0.9323	0.3871
Material 4	0.2414	1.2394	0.6126	1.3196	1.4935
Material 5	0.3778	0.2827	1.6882	1.1380	0.1592
Material 6	0.5834	1.1277	1.1778	1.2481	0.1442
Material 7	0.1601	1.0984	0.9221	1.3220	1.4429
Material 8	0.3760	0.9207	1.4785	1.4647	0.7187
Material 9	0.4252	0.5062	0.6405	0.5822	0.1021
Material 10	1.1806	1.1435	0.9210	0.4037	1.0300

Table 5.5: Difference compared to Optimized Weight Window for Fast Group

	Depletion Cycle 1(%)	Depletion Cycle 2(%)	Depletion Cycle 3(%)	Depletion Cycle 4(%)	Depletion Cycle 5(%)
Material 1	0.028	0.012	0.012	0.037	0.005
Material 2	0.037	0.012	0.005	0.015	0.005
Material 3	0.014	0.014	0.002	0.009	0.017
Material 4	0.037	0.023	0.007	0.019	0.022
Material 5	0.000	0.012	0.029	0.036	0.012
Material 6	0.018	0.014	0.021	0.021	0.031
Material 7	0.014	0.039	0.014	0.014	0.025
Material 8	0.000	0.004	0.039	0.007	0.004
Material 9	0.000	0.029	0.036	0.018	0.011
Material 10	0.028	0.012	0.012	0.037	0.005

Table 5.6: Difference compared to Optimized Weight Window for Thermal Group

	Depletion Cycle (%)	Depletion Cycle 2(%)	Depletion Cycle 3(%)	Depletion Cycle 4(%)	Depletion Cycle 5(%)
Material 1	0.168	0.721	0.757	0.863	0.922
Material 2	1.310	0.656	0.058	1.112	1.047
Material 3	0.250	0.792	0.036	1.249	0.359
Material 4	0.563	1.354	0.611	1.157	0.676
Material 5	0.145	0.676	0.316	0.611	0.959
Material 6	0.010	0.066	1.155	0.859	0.781
Material 7	1.324	0.545	1.153	0.558	1.009
Material 8	0.447	0.702	0.746	1.244	1.670
Material 9	0.103	0.959	1.426	1.272	1.403
Material 10	0.168	0.721	0.757	0.863	0.922

In Table 5.7 and Table 5.8, the execution time for each method is presented in minutes. The data from Table 5.3 and Table 5.4 suggest a speedup between 7~8 is observed compared with respect to the analog runs. It is also observed that the speedups obtained when applying general weight windows and when applying optimized weight windows are very close. Therefore, we can conclude that two types of weight window are comparable. Since it is way more expensive to generate five optimized weight windows than generating one weight window, the general weight window significantly improves the efficiency and could be applied in more complicated reactor calculations.

Table 5.7: Execution Time Applying Single Weight Window

	Depletion Cycle 1	Depletion Cycle 2	Depletion Cycle 3	Depletion Cycle 4	Depletion Cycle 5
analog	19437	19366	19817	19612	19530
General ww	2337	2717	2445	2336	2681
speedup	8.31	7.12	8.10	8.39	7.28

Table 5.8: Execution Time Applying Multiple Weight Windows

	Depletion Cycle 1	Depletion Cycle 2	Depletion Cycle 3	Depletion Cycle 4	Depletion Cycle 5
analog	19437	19366	19817	19612	19530
Optimized ww	2112	2419	2319	2301	2412
speedup	9.20	8.00	8.54	8.52	8.09

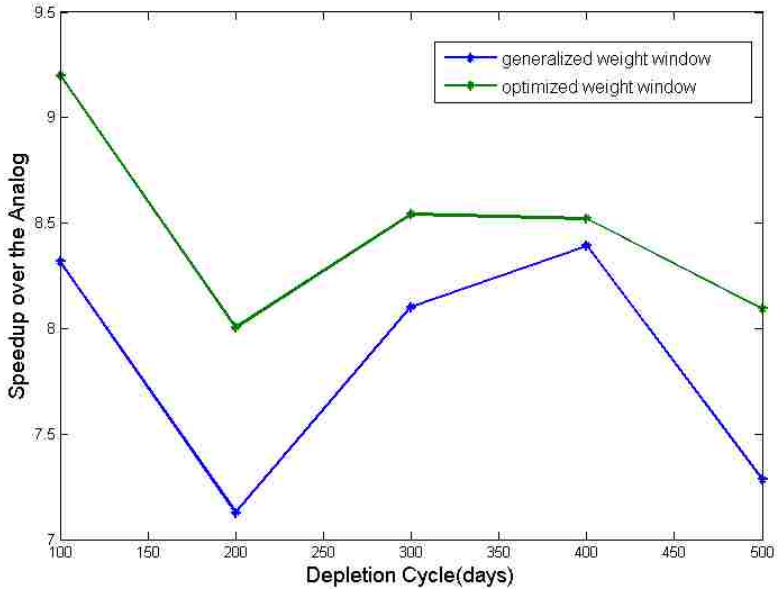


Fig. 5.5: Speedup of Different Weight Windows through the five-cycle-depletion

5.4 Cross Section Functionalization on BWR Assembly Model

A peach-bottom 7*7 BWR assembly model is employed and rebuilt for our purposes. The assembly model represents the southeast assembly of a typical 2x2 BWR control cell which contains four assemblies and a cruciform control blade that is not modeled within the scope of this work. The BWR assembly contains 49 fuel rods in a regular 7x7 fuel rod array. Each fuel pin is assigned with a unique fuel composition. The moderator around the fuel pin is separated into unit cells. The 49 unit cells are tallied and homogenized independently as shown in the Fig. 5.6:

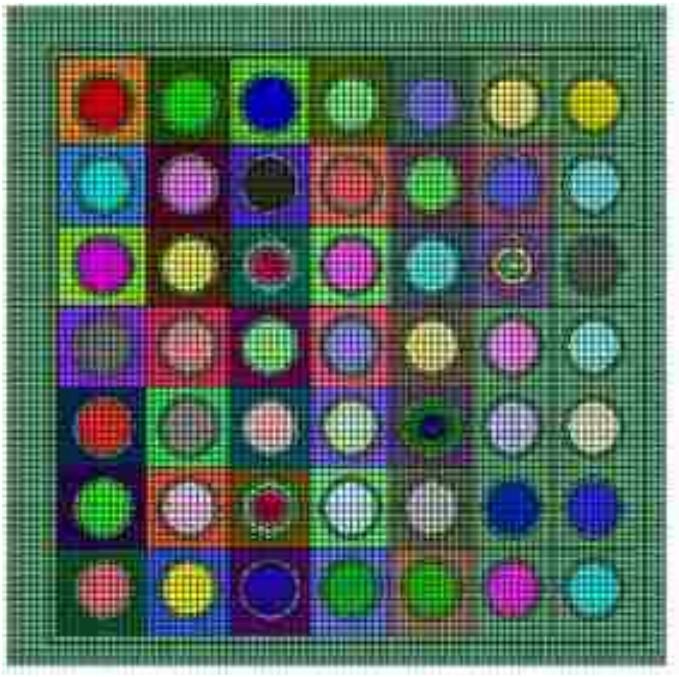


Fig. 5.6: BWR Assembly Model

Table 5.9: BWR Model Specification

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

The BWR model is implemented in both MCNP and NEWT computer codes. Both codes obtain the identical Keff proving consistency of the two models. NEWT is a multigroup, discrete-ordinates radiation transport code that could be used to prepare collapsed weighted cross sections and perform fixed-source and eigenvalue calculations. Compared to the Denovo code from the previously used MAVRIC sequence, NEWT allows the user to define the boundary conditions more accurately.

The 44-group SCALE library is employed and collapsed into thermal group and fast group. The thermal group ranges from 0 through 0.625eV and the fast group comprises the energy range above 0.625eV. Collapsed cross sections for each energy group are obtained from the 44-group library. An importance map is created based on a NEWT adjoint fixed-source calculation where the sources are constructed from the numerical values of the collapsed cross sections from 44-group library. The importance map is consequently used in the MCNP

calculation and for each material in the assembly the desired responses: functionalized cross sections are obtained.

The material representing unit cell 1 in the assembly serves as an example to compare the performance of different hybrid methods. The analog Monte Carlo simulation completes 2000 active cycles, 20000 histories per cycle in 351.02 minutes. When applying the SUBSPACE method, the same number of histories is completed in 131.14 minutes and applying FW-CADIS it is completed in 99.72 minutes.

The relative uncertainties of the obtained numerical results are shown in Table 5.10 for thermal and fast energy groups. The SUBSPACE method obtains an average uncertainty level that is 2~3 times lower compared to the analog for the fast group and 4~5 times lower for the thermal group. Meanwhile, the SUBSPACE method also shows a better performance in reducing uncertainty compared to the FW-CADIS method.

The results of the global FOM are shown in Table 5.11. The global FOM is calculated as:

$$FOM_{GVR} = \frac{1}{\bar{v}(t_{NEWT} + t_{MCNP})}$$

$$FOM_{analog} = \frac{1}{\bar{v}t_{MCNP}}$$

It is seen that for thermal group, SUBSPACE method obtains a speedup between 32~38 over the analog and for the fast group a speedup between 9~16 over the analog.

Table 5.10: Relative Uncertainty of Homogenized Cross Sections

	Fast Group			Thermal Group		
	Analog	SUBSPACE	FW-CADIS	Analog	SUBSPACE	FW-CADIS
Fission	0.103%	0.050%	0.072%	0.106%	0.028%	0.042%
Capture	0.206%	0.085%	0.117%	0.106%	0.028%	0.042%
Scattering	0.078%	0.042%	0.057%	0.099%	0.028%	0.042%

Table 5.11: Global FOM of Homogenized Cross Sections

	Fast Group					
	Analog	SUBSPACE	Speed-up	Analog	FW-CADIS	Speed-up
Fission	2.68E+03	3.05E+04	11.35	2.68E+03	1.92E+04	7.18
Capture	6.70E+02	1.04E+04	15.58	6.70E+02	7.31E+03	10.92
Scattering	4.67E+03	4.23E+04	9.07	4.67E+03	3.13E+04	6.71
	Thermal Group					
	Analog	SUBSPACE	Speed-up	Analog	FW-CADIS	Speed-up
Fission	2.52E+03	9.53E+04	37.81	2.52E+03	5.57E+04	22.10
Capture	2.52E+03	9.53E+04	37.81	2.52E+03	5.57E+04	22.10
Scattering	2.90E+03	9.53E+04	32.79	2.90E+03	5.57E+04	19.16

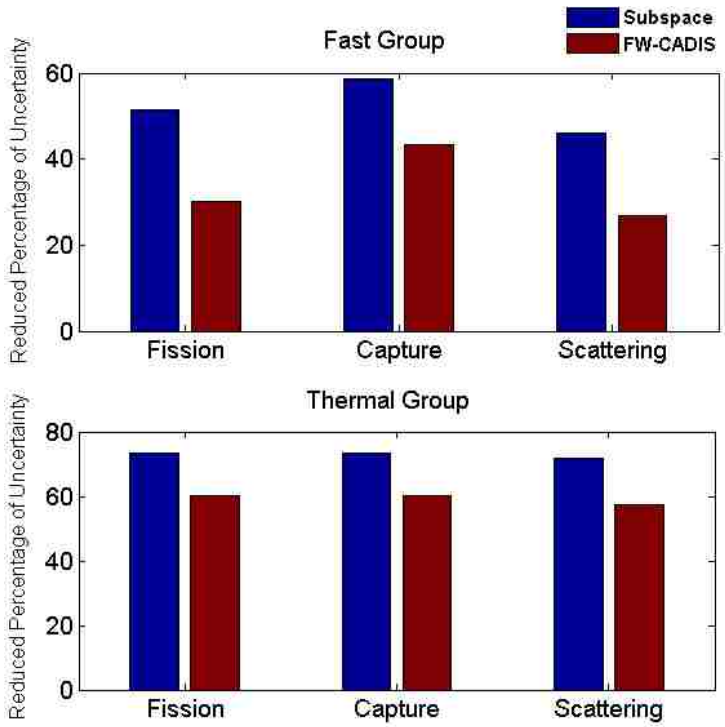


Fig. 5.7: Reduced Percentage of Relative Uncertainty in GVR Calculations

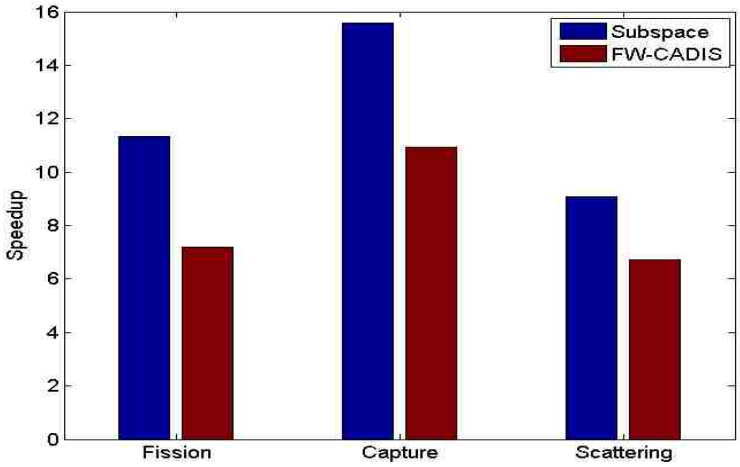


Fig. 5.8: GVR Calculation Speedup for Fast Group

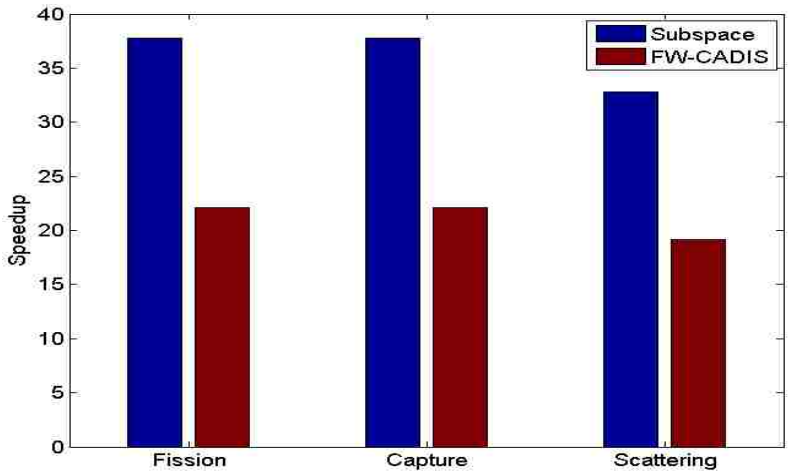


Fig. 5.9: GVR Calculation Speedup for Thermal Group

5.5 Depletion Study on BWR Assembly Model

The 7x7 BWR assembly model described in Section 5.4 is employed in this section for demonstrating the feasibility of the Subspace method in a depletion study. The depletion calculation is conducted by TRITON. The 44 group energy library from SCALE, as introduced in Section 5.4, is employed and collapsed into thermal 0~0.625eV and fast 0.625ev~20MeV groups. The 49 fuel pins are divided into 8 groups based on the different composition for depletion purpose as shown in Fig. 5.10. The assembly model is simulated under hot condition with a pellet temperature of 900 K and a moderator temperature at 600 K. A constant power level of 45.220 MW/MTU is maintained during the depletion which is split over 5 depletion cycles, each spanning a 100 days period. Throughout the 5 depletion steps, the computed multiplication factor decreases from 1.08682 to 0.92166. The Burnup level increases from 1.13 GWd/MTU to 19.2 GWd/MTU as shown in Fig. 5.11. Fig. 5.12 serves as an example to show how the level of uranium isotopes evolves throughout depletion.

The Monte Carlo calculations are performed using MCNP5. The SUBSPACE method is implemented for representing GVR methodology. Since the assembly level homogenized cross section is chosen as the response of interest, the pseudo response is constructed as a linear combination of weighted original cross section responses from the SCALE library. The analog Monte Carlo is performed independently for comparison. For each depletion cycle, 20000 histories/cycle and 2000 active cycles are completed in MCNP simulation.

As demonstrated in Section 5.4, a single weight window in Monte Carlo simulation proves as accurate as multiple specified weight windows for all the depletion cycles. Therefore to guarantee the maximum efficiency, in this work a single averaged weight window is constructed for all the depletion cycles employing the SUBSPACE method.

The number densities of nuclides are obtained for each depletion cycle from the TRITON execution and different depletion scenarios are built. For each depletion scenario, an adjoint fixed source problem, where the SUBSPACE pseudo response is constructed by linearly combining original cross section data from the library, is solved and the corresponding importance map is obtained. Multiple importance maps (3~5 per depletion cycle) are generated employing the SUBSPACE method to represent the complete depletion process.

All the importance maps are then linearly combined into one single importance map, based on which an "average" weight window is constructed for the following Monte Carlo simulations. To compare the performance of the SUBSPACE method versus the analog, the figure of merit for group flux and reaction rates (fission, capture, scattering) at each depletion level are calculated as:

$$FOM_{Subspace} = \frac{1}{v(t_{NEWT} + t_{MCNP})}$$

$$FOM_{analog} = \frac{1}{vt_{MCNP}}$$

For the SUBSPACE method, the total time is the sum of the deterministic calculation time from the adjoint fixed-source run by NEWT and the Monte Carlo calculation time from the MCNP simulation. For the analog, the total time is the Monte Carlo calculation time by MCNP. The final execution times of SUBSPACE and analog are listed in Table 5.13.

The FOM results are shown in Table 5.14 to 5.17. The speedups of FOM obtained by employing the SUBSPACE method are plotted in Fig. 5.13 to Fig. 5.16 to show how the FOM speedup evolves throughout the depletion cycles.

For both flux and reaction rates including fission, capture and scattering, the FOMs obtained when applying the SUBSPACE method gain a speedup that is between 40~50 for the thermal group and between 10~20 for the fast group compared to the analog FOMs.

In Fig. 5.13 to 5.16, it is shown that the speedups distribute evenly through the complete depletion process instead of showing an explicit increasing or decreasing trend. This well demonstrates the consistency of the performance of the single average weight window for all the depletion scenarios.

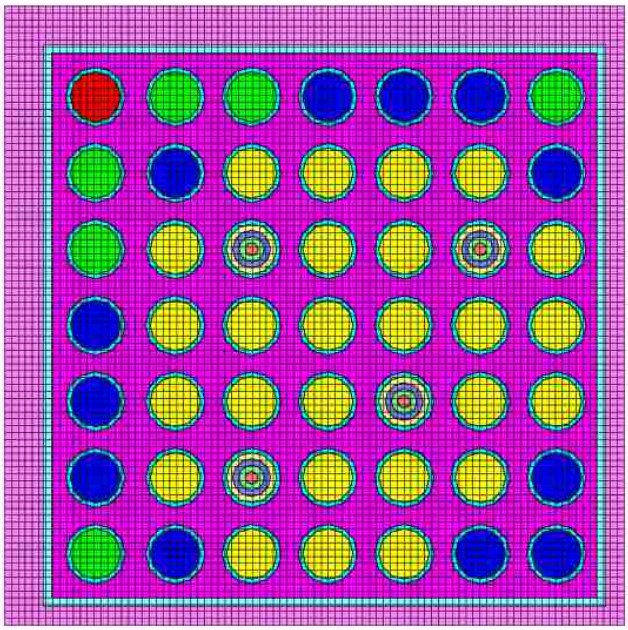


Fig. 5.10: Depletion Pattern of BWR Assembly Model

Table 5.12: Burnup through Depletion Cycles

Depletion Cycle	Depletion Days	Burnup (GWd/MTU)
1	25	1.13
2	125	5.65
3	225	10.2
4	325	14.7
5	425	19.2

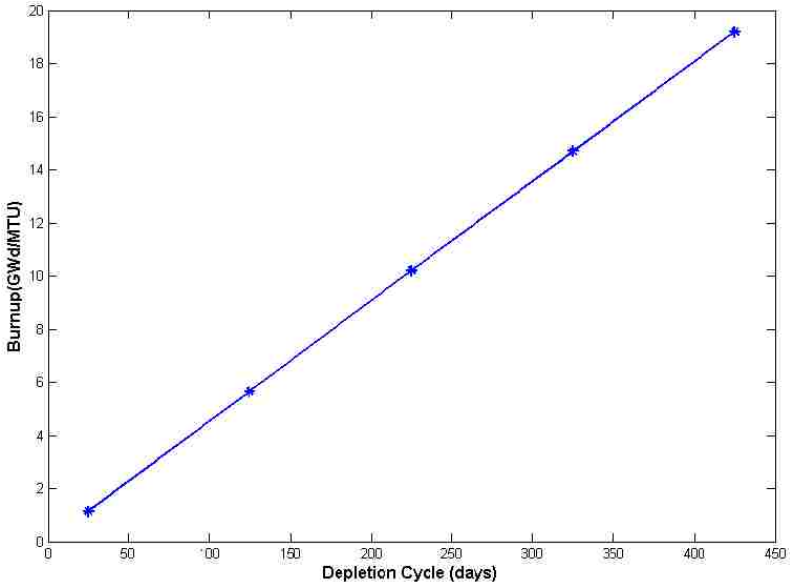


Fig. 5.11: Burnup through Depletion Cycles

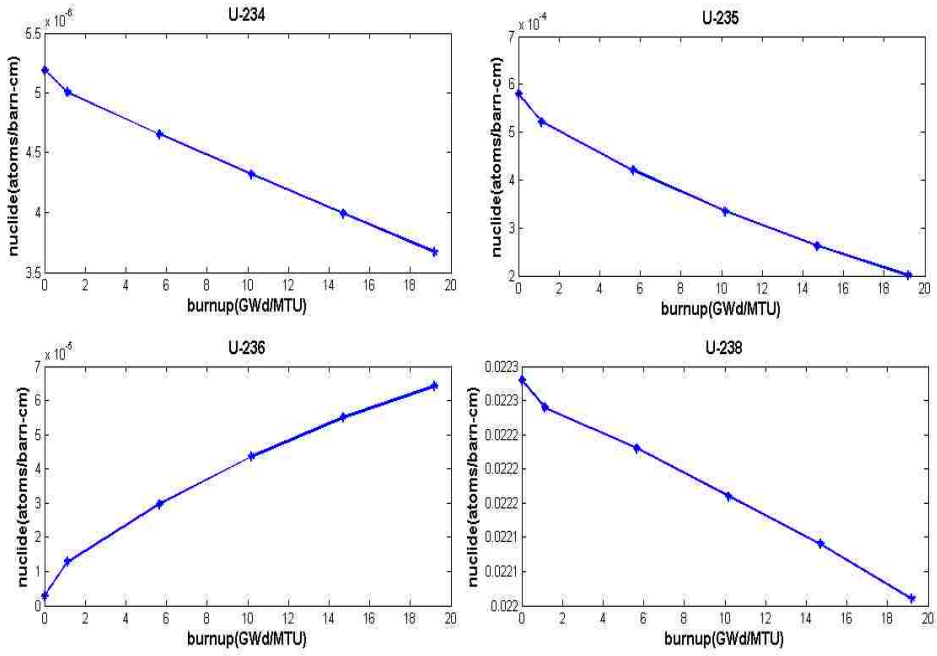


Fig. 5.12: Isotopes of Uranium through Depletion Cycles

Table 5.13: Total Execution Time of Depletion Calculations

Depletion Cycle	Total Execution Time (mins)	
	Analog	SUBSPACE
1	402.65	120.37
2	421.65	135.23
3	441.83	146.08
4	472.17	165.20
5	521.18	178.95

Table 5.14: The FOM Comparison of Flux

Depletion Cycle	FOM				SUBSPACE Speedup	
	Analog		SUBSPACE		Fast	Thermal
	Fast	Thermal	Fast	Thermal		
No depletion	9934.2	5068.5	92304.0	207684.0	9.3	41.0
1	9934.2	5068.5	82161.1	207684.0	8.3	41.0
2	6287.0	6287.0	76059.9	207684.0	12.1	33.0
3	5883.0	5883.0	67257.2	207684.0	11.4	35.3
4	5329.8	5329.8	62092.1	207684.0	11.7	39.0
5	5115.8	5115.8	59307.5	207684.0	11.6	40.6

Table 5.15: The FOM Comparison of Fission Rate

Depletion Cycle	FOM				SUBSPACE Speedup	
	Analog		SUBSPACE		Fast	Thermal
	Fast	Thermal	Fast	Thermal		
No depletion	3066.1	3880.5	33229.4	207684.0	10.8	53.5
1	3066.1	5068.5	29578.0	207684.0	9.6	41.0
2	2794.2	4619.0	27381.6	207684.0	9.8	45.0
3	2614.7	4322.2	16814.3	207684.0	6.4	48.1
4	1918.7	5329.8	15523.0	207684.0	8.1	39.0
5	1841.7	5115.8	14826.9	207684.0	8.1	40.6

Table 5.16: The FOM Comparison of Capture Rate

Depletion Cycle	FOM				SUBSPACE Speedup	
	Analog		SUBSPACE		Fast	Thermal
	Fast	Thermal	Fast	Thermal		
No depletion	620.9	5068.5	12980.2	207684.0	20.9	41.0
1	563.2	5068.5	11553.9	207684.0	20.5	41.0
2	513.2	4619.0	10695.9	207684.0	20.8	45.0
3	437.6	5883.0	9458.0	207684.0	21.6	35.3
4	396.4	5329.8	6899.1	207684.0	17.4	39.0
5	380.5	5115.8	6589.7	207684.0	17.3	40.6

Table 5.17: The FOM Comparison of Scattering Rate

Depletion Cycle	FOM				SUBSPACE Speedup	
	Analog		SUBSPACE		Fast	Thermal
	Fast	Thermal	Fast	Thermal		
No depletion	6898.7	5068.5	92304.0	207684.0	13.4	41.0
1	6898.7	5068.5	82161.1	207684.0	11.9	41.0
2	6287.0	6287.0	76059.9	207684.0	12.1	33.0
3	5883.0	5883.0	67257.2	207684.0	11.4	35.3
4	5329.8	5329.8	62092.1	207684.0	11.7	39.0
5	5115.8	5115.8	59307.5	207684.0	11.6	40.6

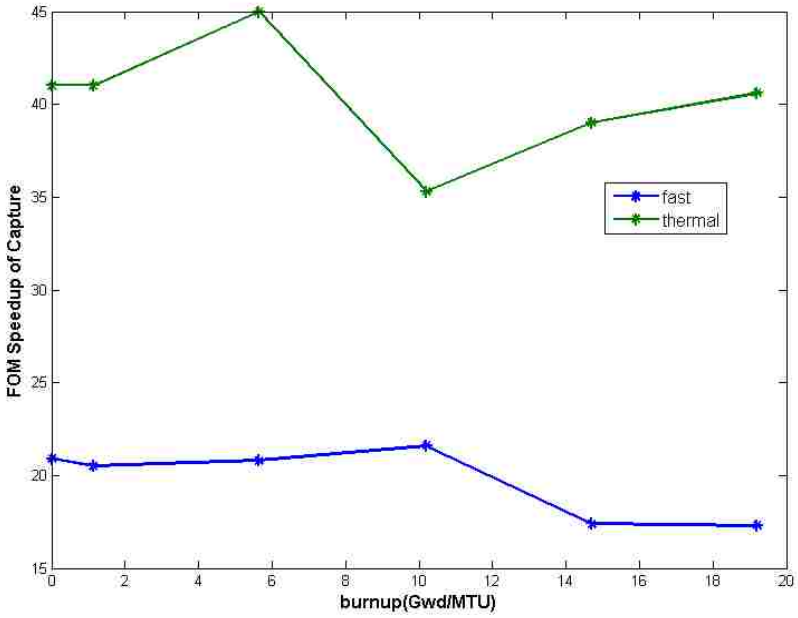


Fig. 5.13: FOM Speedup of Capture through Depletion

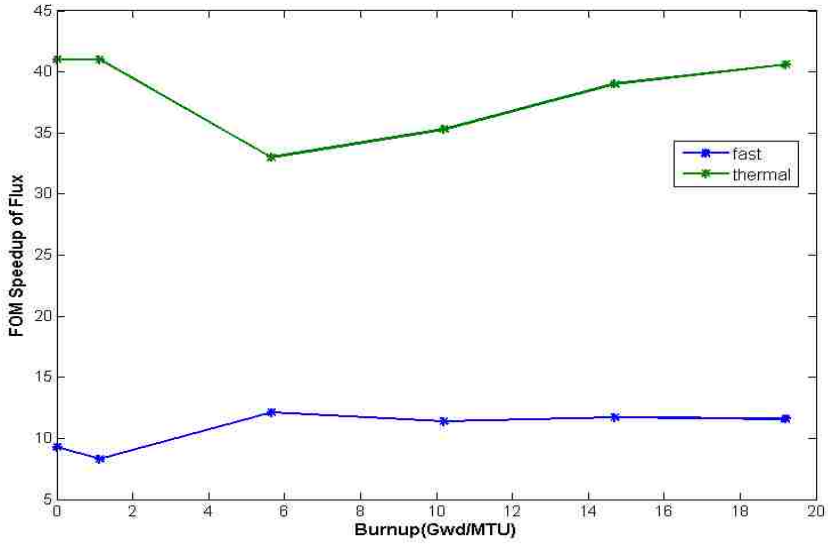


Fig. 5.14: FOM Speedup of Flux through Depletion

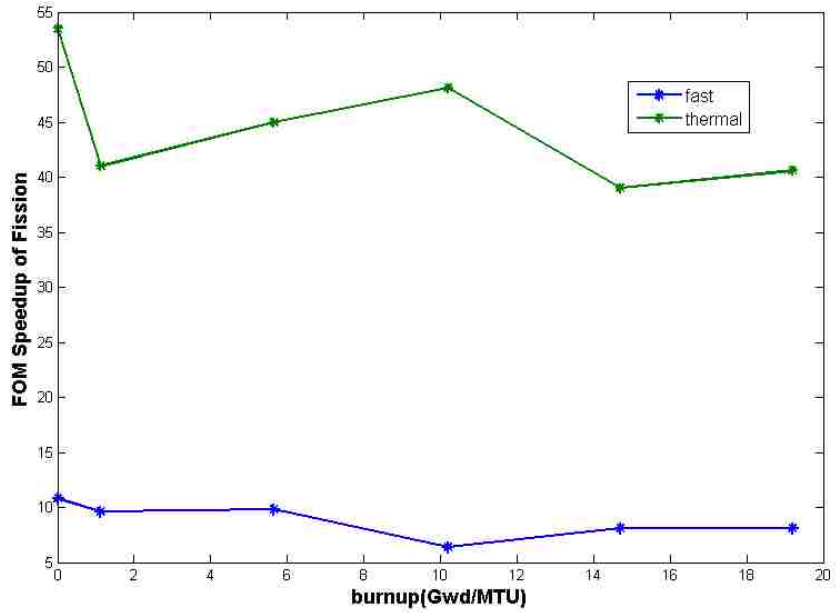


Fig. 5.15: FOM Speedup of Fission through Depletion

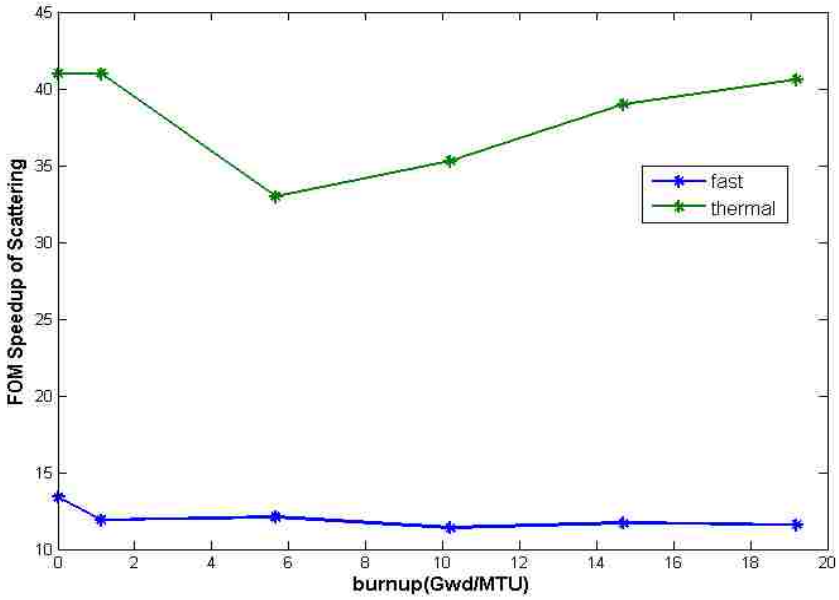


Fig. 5.16: FOM Speedup of Scattering through Depletion

CHAPTER 6

CONCLUSION

6.1 Summary and Conclusions

The ultimate purpose of this work is to develop a mathematically-justified, computationally-efficient, and massively-parallelized framework for elucidating the coupling between Monte Carlo and deterministic models. Up to date, the following objectives have been successfully accomplished:

a) Accelerate the convergence of Monte Carlo calculations via enhanced biasing methods;

b) Enhancing the accuracy and efficiency of coupled Monte Carlo-deterministic calculations for reactor analysis;

c) Determination of energy-collapsed cross-sections from Monte Carlo solutions for deterministic methods;

In this work, two new variants of hybrid Monte Carlo-deterministic GVR techniques are presented: The SUBSPACE and Gaussian Process (GP). In SUBSPACE method, the correlations between the various single-response adjoint-based weight-window maps are identified and determined. The correlations describe a set of pseudo responses whose number is much smaller than the number of original responses.

By biasing the Monte Carlo particles towards the pseudo responses, noticeable computational savings could be achieved. Meanwhile, as an extension of the SUBSPACE method, GP

method takes advantage of the correlations that exist between the variances for the various responses by treating them as Gaussian processes.

Numerical experiments are conducted to assess the performances of the proposed hybrid methods. Preliminary fixed source simulations are performed on both assembly level and core level. The SUBSPACE method is compared to the FW-CADIS method implemented in MAVRIC sequence of the SCALE code system and showed a performance that is more favorable. The same comparison is conducted with the GP method. The obtained results indicate that the performance of the GP method and the SUBSPACE method is comparable. Furthermore, the SUBSPACE method is extended to address k-eigenvalue calculations for reactor analysis. The method is tested and compared to the FW-CADIS method on a 3-D quarter-core model, showing significant computational savings and a speed-up up to 10 times over the FW-CADIS. It is expected with more complicated geometries and more generalization of the pseudo responses construction, one can see more correlations between the weight windows, and hence more computational saving could be achieved.

The proposed hybrid methods are specifically developed for the purpose of accelerating actual reactor physics calculation. Therefore, the applicability of the SUBSPACE method is examined for cross section functionalization and depletion on assembly level.

Significant speedups are obtained comparing to the analog Monte Carlo simulations under different core conditions.

6.2 Topics for Future Research

In future, these computational savings could render the generation of sensitivity information of responses with respect to cross-sections and the propagation of cross-sections uncertainties through Monte Carlo-based models. It is important to note here that while hybrid methods can improve the convergence of MC simulation, a single order of magnitude speedup over analog methods is not expected to render MC models competitive with deterministic methods for routine reactor physics calculations. Therefore, investigations on the use of the GPT-free methodology are recommended to reduce the computational cost required to generate all the depletion and branch cases. Based on recent results of applying GPT-free to a realistic assembly models, another two orders of magnitude speed up could be expected, since the GPT-free methodology allows one to directly calculate the change in the few-group cross-sections due to changes in core parameters without having to re-execute the MC model.

The following topics are recommended for future research:

Sensitivity Analysis

Generate a SUBSPACE-based sensitivity matrix representing the first order derivatives of the multi-response weight-window map with respect to all input data variations. This can be achieved via an ESM-based matrix-free method, involving both the forward and adjoint DT models, with the matrix representing the unknown sensitivity matrix and the vector representing random input data and output responses variations. The resulting SUBSPACE

sensitivity matrix can then be used to accelerate not only reference MC calculations, but also the associated sensitivity analysis calculations;

Uncertainty Quantification

Propagate uncertainties of basic input data by perturbing them along SUBSPACEs consistent with their a prior uncertainty information; SUBSPACEs are to be obtained via matrix revealing decompositions;

Monte Carlo Inverse Analysis

With both prior input data uncertainty and determined sensitivity information, complete model inversion to adjust nuclear data to account for the observed discrepancies between predictions and measurements of core attributes.

REFERENCES

1. A. HAGHIGHAT and J.C. WAGNER, "Monte Carlo Variance Reduction with Deterministic Importance Functions," Progress in Nuclear Energy, 42, 1 (2003).

2. T.E. BOOTH, “Automatic Importance Estimation in Forward Monte Carlo Calculations,” Trans. Am. Nucl. Soc., 41,308 (1982).

3. J.J. DUDERSTADT and L.J. HAMILTON, “Nuclear Reactor Analysis”, Wiley; 1 edition (1976).

4. L.L. CARTER and E.D. CASHWELL, “Particle Transport Simulation with the Monte Carlo Method”, ERDA Critical Review Series, TID-26607 (1975).

5. G.F. KNOLL, “Radiation Detection and Measurement,” John Wiley and Sons (2010).

6. J.F. BRIESMEISTER, Editor, “MCNP – A General Monte Carlo N-Particle Transport Code, Version 4A,” LA-12625, Los Alamos National Laboratory (1993).

7. B.C. KIEDROWSKI, A. IBRAHIM, “Evaluating the Efficiency of Estimating Numerous Monte Carlo Tallies,” 2011 ANS Annual Meeting, June 26-30, Hollywood, FL (2011).

8. J.C. WAGNER, Acceleration of Monte Carlo Shielding Calculations with an Automated Variance Reduction Technique and Parallel Processing, Ph.D. Dissertation, Pennsylvania State University. University Park, Pennsylvania (1997).

9. S.A. DUPREE, S.K. FRALEY, “A Monte Carlo Primer”, Springer; 1edition (2001).

10. 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).

11. M.H. KALOS et al, “Monte Carlo Methods in Reactor Computations,” in Computing Methods in Reactor Physics, H. GREENSPAN et al, Gordon and Breach Publishers, New York (1968).

12. M.H. KALOS and P.A. WHITLOCK, Monte Carlo Methods – Volume 1: Basics, John Wiley & Sons, New York (1986).

13. 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, 2 (2007).

14. S.N. CRAMER and J.S. TANG, “Variance Reduction Methods Applied to Deep-Penetration Monte Carlo Problems,” ORNL/TM-9643, Oak Ridge National Laboratory (1986).

15. R.R. COVEYOU, V.R. CAIN, and K.J. YOST, “Adjoint and Importance in Monte Carlo Application,” Nucl, Sci, Eng., 27,219 (1967).

16. J. D. DENSMORE and E. W. LARSEN, "Variational variance reduction for particle transport eigenvalue calculations using Monte Carlo adjoint simulation," Journal of Computational Physics, 192, 2 (2003).

17. I. LUX, L. KOBLINGER, “Monte Carlo Particle Transport Methods: Neutron and Photon Calculations,” CRC Press, Boca Raton, FL (1990).

18. M.A. COOPER, “An Automated Variance Reduction Method for Global Monte Carlo Neutral Particle Transport Problems,” Ph.D. Thesis, University of Michigan, Department of Nuclear Engineering and Radiological Sciences (1999).

19. M.A. COOPER and E.W. LARSEN. “Automated Weight Windows for Global Monte Carlo Particle Transport Calculations”. Nucl, Sci, Eng., 137:1, (2001).

20. K.A. VAN RIPER et al, “AVATAR—Automated Variance Reduction in Monte Carlo Calculations,” Proc. Joint Int. Conf. Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs, New York, October 5-9, 1997, Vol.1, p.661,American Nuclear Society (1997).

21. 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. Adv. Math. Comp., Saratoga Springs, NY, May (2009).

22. J.P. BOTH, H. DERRIENNIC, B. MORILLON, and J.C. NIMAL, “A Survey of TRIPOLI-4,” Proc. of the Eighth Int. Conf. on Radiation Shielding, Arlington. TX1, 373 (1994).

23. G.I. BELL and S. GLASSTONE, Nuclear Reactor Theory, Van Nostrand and Reinhold, New York (1967).

24. D.E. PEPLOW, “MAVRIC: MONACO with automated variance reduction using importance calculations”, ORNL/TM-2005/39, UT-Battelle, LLC, Oak Ridge National Laboratory, January (2009).

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

26. M.B EMMETT, “The MORSE Monte Carlo Radiation Transport Code System,” ORNL-4972. Oak Ridge National Laboratory (1975).

27. H. S. ABDEL-KHALIK, “Adaptive Core Simulation,” PhD Dissertation, Department of Nuclear Engineering, North Carolina State University, December (2004).

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

29. N. HALKO, P.G. MARTINSSON, and J. TROPP, “Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions.” SIAM Review, 53, 2 (2011).

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

31. M. A. JESSEE, H. S. ABDEL-KHALIK, and P. J. TURINSKY, “Cross-Section Uncertainty Propagation and Adjustment Algorithms for BWR Core Simulation,” Proceedings of the Advances in Nuclear Fuel Management IV, Hilton Head, April (2009).

32. T.E. BOOTH, “Analytic Monte Carlo Score Distributions for Future Statistical Confidence Interval Studies”, Nucl, Sci, Eng., 112, 159-169(1992)

33. J.C. WAGNER, D.E. PEPLOW, S.W. MOSHER, T.M. EVANS, “Review of Hybrid (Deterministic/Monte Carlo) Radiation Transport Methods, Codes, and Applications at Oak Ridge National Laboratory,” Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010, Hitotsubashi Memorial Hall, Tokyo, Japan, October 17-21 (2010).

34. A. YAMAMOTO, T. IKEHARA, T. ITO, and E. SAJI, “Benchmark Problem Suite for Reactor Physics Study of LWR Next Generation Fuels,” Journal of Nuclear Science and Technology, 39, 8 (2002).

35. S. DOUGLASS, F. RAHNEMA, and J. MARGULIES, “A Stylized three dimensional PWR whole-core benchmark problem with Gadolinium,” Annuals of Nuclear Energy, 37, 10 (2010).

36. J. HITE, Y. BANG, C. WANG, and H. S. ABDEL-KHALIK, “Heuristic Method for ESM-based Reduced Order Modeling,” Trans. Am. Nucl. Soc., 105, 475 (2011).

37. F.B. BROWN, “A Review of Monte Carlo Criticality Calculations-Convergence, Bias, Statistics,” Monte Carlo Codes, X-3-MCC, LANL, M&C 2009, May 3-7, Saratoga Springs, NY, May (2009).

38. D.E. PEPLOW, T.M. EVANS, J.C. WAGNER, “Simultaneous Optimization of Tallies in Difficult Shielding Problems,” Nucl_Technol_168_p785_792 (2008).

39. B.C. KIEDROWSKI, C.J. SOLOMON, "Statistical Assessment of Numerous Monte Carlo Tallies," Proc. Int. Conf. Math. Comp. Meth. and Appl. Nucl. Sci. Eng. M&C Rio De Janeiro, Brazil (2011).

40. D.S. RAHNEMA, F. MARGULIES, "A Stylized Three Dimensional PWR Whole-Core Benchmark Problem with Gadolinium." Ann. Nucl. Eng.37 (10), 1384–1403 (2010).

41. J.L. RODGERS, W.A. NICEWANDER, "Thirteen ways to look at the correlation efficient," The American Statistician, 42(1):59-66 (1988).

42. S. CHUCAS, I. CURL, T. SHUTTLEWORTH, and G. MORRELL, "Preparing the Monte Carlo Code MCBEND for the 21st Century," Proc. of the Eighth Int. Conf. on Radiation Shielding, Arlington, TX 1,381 (1994).

43. S.A. TURNER, "Automatic Variance Reduction for Monte Carlo Simulations via the Local Importance Function Transform," LA-13119-T, Los Alamos National Laboratory (1996).

44. F.X. GIFFARD, R. JACQMIN, J.C. NIMAL, and Y. PENELIAU, "Variance Reduction in 3-D Continuous-Energy Monte Carlo Simulations Using Importance Maps Generated by a Deterministic Code," Proceedings of the Mathematics and Computation, Reactor Physics and Environmental Analysis in Nuclear Applications, p. 273, Madrid, Spain, September 27-30 (1999).

45. R.E. ALCOUFFE, R.S. BAKER, F.W. BRINKLEY, D.R. MARR, R.D. O'Dell, and W.F. WALTERS, "DANTSYS: A Diffusion Accelerated Neutral Particle Transport Code System," Los Alamos National Laboratory, LA-12969-M (1995).

46. J.S. TANG, P.N. STEVENS, and T.J. HOFFMAN, Methods of Monte Carlo Biasing Using Two-Dimensional Discrete Ordinates Adjoint Flux, Oak Ridge National Laboratory, ORNL/TM-5414 (1976).

47. J.S. TANG, and T.J. HOFFMAN, Monte Carlo Shielding Analyses Using an Automated Biasing Procedure, Nucl. Sci. Eng. 99, p. 329 (1988).

48. J.S. TANG, SAS4: A Monte Carlo Cask Shielding Analysis Module Using an Automated Biasing Procedure, Sec. S4 of SCALE: A Modular Code System for Performing Standardized Computer Analyses for Licensing Evaluations, NUREG/CR-0200, Rev. 6 (ORNL/NUREG/CSD-2N2/R6) (1998).

49. T.M. EVANS, A.S. STAFFORD, R.N. SLAYBAUGH, K.T. CLARNO, "Denovo: A New Three Dimensional Parallel Discrete Ordinates Code in SCALE," Nuclear Technology / Volume 171 / Number 2 / Pages 171-200 (2010).

50. G.I. BELL and S. GLASSTONE, Nuclear Reactor Theory, Van Nostrand and Reinhold, New York (1967).

51. F.B. BROWN, A Review of Monte Carlo Criticality Calculations-Convergence, Bias, Statistics, Monte Carlo Codes, X-3-MCC, LANL, M&C, May 3-7, Saratoga (2009).

52. J.C. WAGNER and S.W. MOSHER, Forward-Weighted CADIS Method for Variance Reduction of Monte Carlo Reactor Analysis, ORNL, Trans. Am. Nucl. Soc. 103, 342-344 (2010).

53. E.L. KAPLAN, "Monte Carlo Methods for Equilibrium Solutions in Neutron Multiplication", UCRL-5275-T, Livermore National Laboratory (1958).

54. W. GOAD and R. JOHNSTON, "A Monte Carlo Method for Criticality Problems", Nucl. Sci. Eng.5, 371-375 (1959).

55. J. LIEBEROTH, "A Monte Carlo Technique to Solve the Static Eigenvalue Problem of the Boltzmann Transport Equation," Nukleonik 11,213 (1968).

56. M. R. MENDELSON, "Monte Carlo Criticality Calculations for Thermal Reactors," Nucl. Sci Eng. 32, 319-331 (1968).

57. M.C. KENNEDY, O'HAGAN, Bayesian calibration of computer model. J. of the Royal Statistical Society B. 63 (3), 425–464 (2011).

58. Q. ZHANG, H.S. ABDEL-KHALIK, 2011b. "Adjoint-based global variance reduction method for reactor analysis problems," Proc. Int. Conf. Math. Comp. Meth. and Appl. Nucl. Sci. Eng. Rio De Janeiro, Brazil (2011).

59. J.C. WAGNER, E.D. BLAKEMAN, D.E. PEPLOW, “Forward-weighted CADIS method for variance reduction of Monte Carlo calculations of distributions and multiple localized quantities”, Proc. Int. Conf. Adv. Math. Comp. and Reac. Phy. Saratoga Spring, NY (2009).

60. M.D. DEHART, "NEWT: A new transport algorithm for two-dimensional discrete ordinates analysis in non-orthogonal geometries", ORNL/TM-2005/39, UT-Battelle, LLC, Oak Ridge National Laboratory, January (2009).

61. M.D. DEHART, "TRITON: A two-dimensional transport and depletion module for characterization of spent nuclear fuel”, ORNL/TM-2005/39, UT-Battelle, LLC, Oak Ridge National Laboratory, January (2009).

62. Q. ZHANG, H.S. ABDEL-KHALIK, “Gaussian Process Approach for Global Variance Reduction”, Transactions of the American Nuclear Society, Vol 104, Hollywood, FL, June (2011).

63. Q. ZHANG, H.S. ABDEL-KHALIK, “Global Variance Reduction for Monte Carlo Reactor Physics Calculations”, Accepted to ANS Winter Meeting, Nov. 11-15 (2012).

64. Q. ZHANG, H.S. ABDEL-KHALIK, “Extending the SUBSPACE Hybrid Method for Eigenvalue Problems”, Accepted to ANS Winter Meeting, Nov. 11-15 (2012).

APPENDIX

APPENDIX

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

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

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 [11]:

1. Pick a small integer *s*, e.g. $s = 10$ is suitable for most practical calculations.
2. Generate *s* random vectors $\left\{\vec{\mu}_j\right\}_{j=1}^{s}$.
3. Calculate: $\left\{\vec{w}_j = \mathbf{\Psi}^T \vec{\mu}_j\right\}_{j=1}^{s}$.
4. Given a user-defined tolerance ε.
5. Given an estimate of the rank r_0, generate r_0 random vectors $\left\{\vec{\eta}_j\right\}_{j=1}^{r_0}$
6. Calculate: $\left\{\vec{z}_j = \mathbf{\Psi}^T \vec{\eta}_j\right\}_{j=1}^{r_0}$.
7. Form an orthonormal matrix $\mathbf{Q} \in \mathbb{R}^{J \times r_0}$ such that: $\mathrm{R}(\mathbf{Q}) = \mathrm{span}\left\{\vec{z}_1, ..., \vec{z}_{r_0}\right\}$. This could be done via a Gram-Schmidt orthogonalization procedure.
8. Calculate: $\varepsilon_1 = \max_{j=1,...,s}\left\{\theta_j = \left\|\left(\mathbf{I} - \mathbf{Q}\mathbf{Q}^T\right)\vec{w}_j\right\|\right\}_{j=1}^{s}$.
9. If $10\sqrt{\frac{2}{\pi}}\varepsilon_1 > \varepsilon$, 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. Notice 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.