Global Monte Carlo Simulation with High Order Polynomial Expansions

FINAL REPORT

DE-FG07-04ID14607

William R. Martin, PI
James Paul Holloway, co-PI

Kaushik Banerjee
Jesse Cheatham
Jeremy Conlin

University of Michigan
Department of Nuclear Engineering and Radiological Sciences
Ann Arbor, MI 48109-2104

December 2007
Table of Contents

Abstract ........................................................................................................................................ 3
I. Introduction ............................................................................................................................... 4
II. Estimation of the Fission Source with FET ............................................................................ 5
   A. Fission source iteration ........................................................................................................ 5
   B. Application of FET to Estimate the Fission Source .......................................................... 6
   C. Controlling the Truncation and Statistical Errors with a Cost-Benefit Ratio .................... 7
   D. Formulation of FET with Piecewise Polynomial Basis Functions .................................... 8
   E. Computational Results with Piecewise Polynomial Basis Functions ............................. 11
   F. Adaptive Coarsening Method ............................................................................................17
   G. Comments ....................................................................................................................... 17
III. Using FET to Accelerate Fission Source Convergence ........................................................ 18
   A. Implementation in MCNP4C ............................................................................................. 18
   B. Numerical results ............................................................................................................. 19
   C. Alternative FET-like Methods for Fission Source Convergence .................................... 30
IV. Development of Monte Carlo-based Response Matrix Method ............................................. 38
   A. Background ...................................................................................................................... 38
   B. Alternative FET-based response matrix approach ........................................................... 39
V. References ............................................................................................................................. 41
VI. Other Grant Information ....................................................................................................... 43
   A. Graduate Students .......................................................................................................... 43
   B. Budget and Expenditures ............................................................................................... 43
   C. Publications Arising from the Grant ................................................................................ 43
Abstract

The functional expansion technique (FET) was recently developed for Monte Carlo simulation. The basic idea of the FET is to expand a Monte Carlo tally in terms of a high order expansion, the coefficients of which can be estimated via the usual random walk process in a conventional Monte Carlo code. If the expansion basis is chosen carefully, the lowest order coefficient is simply the conventional histogram tally, corresponding to a flat mode. This research project studied the applicability of using the FET to estimate the fission source, from which fission sites can be sampled for the next generation. The idea is that individual fission sites contribute to expansion modes that may span the geometry being considered, possibly increasing the communication across a loosely coupled system and thereby improving convergence over the conventional fission bank approach used in most production Monte Carlo codes. The project examined a number of basis functions, including global Legendre polynomials as well as “local” piecewise polynomials such as finite element hat functions and higher order versions. The global FET showed an improvement in convergence over the conventional fission bank approach. The local FET methods showed some advantages versus global polynomials in handling geometries with discontinuous material properties. The conventional finite element hat functions had the disadvantage that the expansion coefficients could not be estimated directly but had to be obtained by solving a linear system whose matrix elements were estimated. An alternative fission matrix-based response matrix algorithm was formulated. Studies were made of two alternative applications of the FET, one based on the kernel density estimator and one based on Arnoldi’s method of minimized iterations. Preliminary results for both methods indicate improvements in fission source convergence. These developments indicate that the FET has promise for speeding up Monte Carlo fission source convergence.
I. Introduction

The prediction of global neutron scalar flux and thermal power profiles throughout a reactor core is key for improved performance of nuclear reactors, from efficient utilization of fuel to improved burning and transmutation of nuclear waste isotopes and increased confidence regarding margins for safety. As nuclear reactor designs become more complex, there is an increasing need to perform 3-D "full-core" simulations that will provide accurate predictions of the neutron phase space density in space, energy, and angle, as well as time. This is inherently a 7-D problem including time dependence and has traditionally stressed computational methods and computing capacity since the early days of nuclear power. The Monte Carlo method can solve the Boltzmann transport equation “exactly” to the extent that the geometry and composition can be specified and the neutron interaction cross sections are known.

However, the caveat "in principle" is operative for full-core reactor analyses because the use of Monte Carlo to analyze a large, loosely-coupled system such as a modern light water reactor is problematical. Smith [1] estimated 20 billion histories would be needed to obtain a solution with 1% statistics on assembly-wise peak powers, and a large factor (100x) in this estimate was due to the fact that the dominance ratio (ratio of two largest eigenvalues) was very close to unity for loosely-coupled systems, necessitating a large number of generations (and a large number of particles per generation) to converge the solution to within the desired precision. This extreme computational time is compounded by the observation that over 6 billion tallies would be needed to characterize the solution, including depletion.

A recently developed methodology, the functional expansion technique (FET), utilizes high order polynomials to estimate quantities of interest in a Monte Carlo simulation, such as the scalar flux distribution within a lattice, and results indicate that the resultant high order approximations are substantially better estimates of the true distributions than histogram (binning) estimates used in conventional Monte Carlo codes.

As a result of this success with FET, we proposed a research project to use FET to accelerate Monte Carlo fission source convergence for loosely-coupled systems [2-7]. The idea here was to expand the fission source in terms of expansion functions (e.g., high order global polynomials, piece-wise polynomials, etc.) and then sample from this fission source during the fission source iterations. The hope was in improvement in convergence due to either the high order polynomial basis functions and/or the increased coupling of regions across the reactor as a result of the estimation and sampling approaches inherent in the FET. It was also proposed to imbed the FET into a response matrix methodology as an alternative approach to speed up fission source convergence. This proposed effort included the following key tasks:

- Estimation and sampling of the fission source with FET, including the use of piecewise expansion functions as well as global basis functions
- Using FET to accelerate fission source convergence within MCNP5
- Imbed FET into a Monte Carlo-based response matrix method

Work on each of these topics is discussed in the sections below.
II. Estimation of the fission source with FET

A. Fission source iteration

This section describes the methodology to use FET with global polynomials to (1) estimate the fission source distribution during a fission generation and (2) sample from this source distribution to obtain the fission neutron birth sites for the next generation.

Estimating the system eigenvalue from a Monte Carlo simulation is relatively simple because it is an integral quantity,

\[
\kappa_{\text{eff}} = \frac{\int \int_E v(x,E) \Sigma_f(x,E) \phi_k(x,E) \, dx \, dE}{\int \int_E S_k(x,E) \, dx \, dE},
\]

(1)

where \( \phi_k \) is the scalar flux distribution that results from an assumed (or predicted) source distribution \( S_k, \Sigma_f \) is the macroscopic fission cross section in the system, and \( v(x,E) \) is the average number of neutrons produced per fission event at \( x \) due to a neutron of energy \( E \). This quantity is typically estimated by either the collision or track length estimator:

\[
\hat{\kappa}_{\text{eff}}^C = \frac{1}{N} \sum_{i=1}^{N} \sum_{c=0}^{C_i} \sum_{j=1}^{J} \left[ \frac{v_j(E_{i,c}) \Sigma_{f,j}(x_{i,c},E_{i,c})}{\Sigma_{t,j}(x_{i,c},E_{i,c})} \right],
\]

(2)

\[
\hat{\kappa}_{\text{eff}}^{\text{TL}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C_i} \sum_{j=1}^{J} v_j(E_{i,c}) \Sigma_{f,j}(x_{i,c},E_{i,c}),
\]

(3)

where \( w_{i,c}, x_{i,c}, \) and \( E_{i,c} \) are the weight, position, and energy, respectively, of neutron \( i \) just prior to event \( c \), \( N \) is the total number of independent particle histories in a single neutron generation \( \ell \), and \( L \) is the total number of generations. The number of events due to history \( i \) is defined by \( C_i \), and is indexed by the variable \( c \) and \( d_{i,c} \) is the total distance traveled by particle \( i \) between events \( c-1 \) and \( c \). (An event is a collision for the collision estimator and either a collision or boundary crossing for the track length estimator.) The total number of fissile isotopes in a problem is defined as \( J \), with each individual isotope denoted by the identifier \( j \). The macroscopic fission and total cross sections for an isotope \( j \) are represented by \( \Sigma_{f,j} \) and \( \Sigma_{t,j} \), respectively. The average number of neutrons released per fission event in isotope \( j \) is denoted by \( v_j \). At the end of generation \( \ell \), Eqs. (2) and (3) gives an estimate for the effective multiplication factor \( \hat{\kappa}_{\text{eff}}^C \) for the system with a collision estimator or track length estimator, respectively.

The previous paragraphs described the relatively straightforward methods for estimating the effective multiplication factor \( k_{\text{eff}} \). However, the spatial dependence of the fission source
distribution is also needed in order to determine the converged flux/power distribution throughout the core. The conventional Monte Carlo approach is to store the sites of individual fission events that are produced by particles during the current generation \( \ell \). At each collision event, the number of secondary neutrons due to fission is sampled, along with the energy of the emitted neutron. The locations and energies of these fission neutrons are stored in a list, commonly referred to as the \textit{fission bank}. After each fixed-source calculation has ended, the contents of the fission bank are used as the source points for the next generation. Thus, the actual shape of the fission source is never known since the fission bank is only used to store samples of the source distribution.

**B. Application of FET to Estimate the Fission Source**

We now describe the FET for determining the fission source distribution and sampling from it. We begin by writing the fission source distribution as a series expansion in Legendre polynomials,

\[
S_\ell (\tilde{x}) = \sum_{n=0}^{M} \frac{2n+1}{2} a_n P_n (\tilde{x}),
\]

(4)

where \( P_n (\tilde{x}) \) is the \( n^{th} \) Legendre polynomial, \( a_n \) are the expansion coefficients, \( M \) is the truncation order of the series, and \( \tilde{x} \) is a scaled spatial variable defined on the Legendre domain \( \tilde{x} \in [-1,1] \). Using the orthogonality property of the Legendre polynomials, it is possible to solve Eq. (4) for the expansion coefficients

\[
a_n = \int \chi S_\ell (\tilde{x}) P_n (\tilde{x}) d\tilde{x}.
\]

(5)

Finally, rewriting Equation (5) in terms of the flux distribution from the previous neutron generation, \( \phi_{\ell-1} \),

\[
a_n = \int \int \nu (\tilde{x}, E) \Sigma_f (\tilde{x}, E) \phi_{\ell-1} (\tilde{x}, E) P_n (\tilde{x}) dE d\tilde{x},
\]

(6)

gives the an expression for the expansion coefficient in terms of a convenient flux moment integral that is suitable for estimation by Monte Carlo.

Using our previous results for Monte Carlo FET estimators for quantities of the form shown in Eq. (6), it follows that the expansion coefficients can be estimated with either a collision or track length FET estimator as follows:

\[
\hat{a}_n^C = \frac{1}{N} \sum_{\ell=1}^{n} \sum_{c=0}^{C} w_{i,c} P_n (\tilde{x}_{i,c}) \sum_{j=1}^{J} \left[ \nu_j (E_{i,c}) \frac{\Sigma_{f,j} (x_{i,c}, E_{i,c})}{\Sigma_{i,j} (x_{i,c}, E_{i,c})} \right],
\]

(7)
\[ \hat{a}_n^{\text{TL}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} w_{i,c} d_{i,c} \int \hat{x}_{i,c} P_i(\hat{x}) \, d\hat{x} \left[ \sum_{j=1}^{J} V_j(E_{i,c}) \Sigma_{j,j}(x_{i,c}, E_{i,c}) \right]. \]  

(8)

Notice that the FET estimators for the fission source expansion coefficients given in Eqs. (7) and (8) are similar to the eigenvalue estimators given in Eq. (2) and (3). In fact, in the case where \( n = 0 \), the FET estimators reduce exactly to the eigenvalue estimators, giving

\[ \hat{a}_0^C = \hat{k}_\text{eff}^C, \]

\[ \hat{a}_0^{\text{TL}} = \hat{k}_\text{eff}^{\text{TL}}. \]  

(9)

During each neutron generation, either Eq. (7) or Eq. (8) can be used to estimate a set of Legendre moments for the spatial source distribution. These moments can be used in Eq. (4) to give a functional approximation for the spatial shape of the fission source. During the next generation, source points can be sampled directly from this functional approximation, eliminating the need for the fission bank.

\section*{C. Controlling the Truncation and Statistical Errors with a Cost-Benefit Ratio}

As with all Monte Carlo tallies, minimizing the statistical uncertainty in each of the Legendre moment estimates is very important for achieving an accurate result. Poorly converged moments can contaminate the functional source approximation with statistical noise, thus slowing the convergence rate of the source iteration. In addition to the statistical uncertainty, functional approximations produced with FET also contain truncation error due to approximating the true fission source by a low order series expansion. Therefore, any FET-based source convergence scheme must use a sufficiently high order expansion to ensure that the source distribution is well represented by the functional approximation.

In the FET, these two sources of error are inversely related to each other. Low order expansion coefficients are easier to estimate with Monte Carlo and, therefore, have smaller statistical uncertainties. However, using too few coefficients in a series expansion will result in large truncation error and a low resolution approximation. On the other hand, keeping too many poorly converged coefficients will result in a final approximation that is heavily contaminated by statistical noise. In order to maximize the effectiveness of an FET source convergence method, an optimal balance must be found between these two sources of error, such that the total error in the approximation is minimized. To help with finding this optimal balance, a simple cost-to-benefit ratio has been developed \[7,8\] to help users determine which coefficient estimates should be included in the series expansion approximation. The metric is defined as

\[ R_n^2 = \frac{2n + 1}{2} \frac{\sigma_{\hat{a}_n}}{\hat{a}_n}, \]  

(10)
where $\hat{\sigma}_{n\alpha}$ is the sample standard deviation for the coefficient estimate $\hat{a}_n$. The $R^2_n$ metric is simply the ratio of the increase in statistical error to the decrease in truncation error due to adding a nonzero coefficient $\hat{a}_n$ to the series expansion in Eq. (4).

This cost-to-benefit ratio provides a convenient test for determining how many expansion coefficients should be used for a given source approximation. Coefficients with values of $R^2_n >> 1$ should not be included in the approximation because they are not well converged and do not add any useful information to the result. Terms with $R^2_n$ values less than, or close to 1, should be included in the approximation because they provide valuable information about the shape of the true function.

The methods discussed above were implemented in a special version of MCNP4C as discussed in Section II.A below.

D. Formulation of FET with Piecewise Polynomial Basis Functions

This section describes the application of piecewise basis polynomial basis functions (e.g., finite element basis functions) with FET instead of global basis functions that have been discussed above. This section includes derivations of expressions that can be used to compute the uncertainty (standard deviations and covariance matrix) of the node values of a finite element based functional expansion tally [9-10].

This section describes the application of piecewise polynomial basis functions (e.g., finite element basis functions) with FET rather than global basis functions that have been discussed up till now. The reason for this is that typical reactor configurations have numerous internal boundaries where material properties are discontinuous, resulting in strongly varying neutron flux distributions, including scalar fluxes as well as angular fluxes. These cause Gibbs-like phenomena in the approximate solution near the discontinuities and degrade the convergence rate of the method to the true solution. This behavior is also seen with the FET when used to analyze problems with discontinuities, or “kinks”, in the desired solution.

**Derivation of Equations.** First approximate the true pdf $p(x)$ with an approximate pdf $\tilde{p}(x)$ that is an expansion in a finite element basis,

$$\tilde{p}(x) = \sum_{n=1}^{M} a_n A_n(x)$$

(11)

where $A_n(x)$ is zero at all points on a grid, except for point $x_n$, where it is unity. That is

$$A_n(x_m) = \delta_{nm}$$

(12)

It is assumed that the $A_n(x)$ are piecewise linear basis functions, which necessarily have a support spanning only three grid points in 1-D, but this does not affect the general considerations
presented here. The expansion coefficient \( a_n \) will represent the value of the approximation on the grid, i.e., \( \tilde{p}(x_n) = a_n, n = 1, 2, \ldots, M \).

To find the expansion coefficients, we require that the weighted residual vanish over the domain for all basis functions \( \Lambda_n(x) \),

\[
\int \left[ p(x) - \tilde{p}(x) \right] \Lambda_n(x) dx = 0, \quad n = 1, 2, \ldots, M
\]

which yields the following set of equations,

\[
\int p(x) \Lambda_n(x) dx = \sum_{m=1}^{M} a_m \int \Lambda_n \Lambda_m(x) dx, \quad n = 1, 2, \ldots, M
\]

Equation (14) can be represented as a matrix equation for \( a = \text{col}[a_1, a_2, \ldots, a_N] \),

\[
\Lambda a = c
\]

where \( \Lambda_{nm} = \int \Lambda_n(x) \Lambda_m(x) dx \)

and \( c_n = \int p(x) \Lambda_n(x) dx \)

We now sample a point \( x \) from the pdf \( p(x) \) using Monte Carlo, and estimate the integrals \( c_n \),

\[
\hat{c}_n = \frac{1}{N} \sum_{i=1}^{N} \int \Lambda_n(x_i) dx
\]

where \( N \) is the number of histories. We must also estimate the covariances, but this will be discussed later. With estimates for the elements of \( \hat{c} \) in hand we can now solve for the "estimated" expansion coefficients \( \hat{a} \),

\[
\hat{a} = \Lambda^T \hat{c}
\]

The quantity \( \hat{a} \) is an unbiased estimate of the true vector of coefficients \( \bar{a} = E[a] \).

**Covariance Matrix.** Our next task is to compute the uncertainty in each of the coefficients that results from the stochastic estimates of the values \( c_n \). Note that the \( c_n \) terms are not independently estimated so we need to introduce their covariance matrix.
\[ C = E\left[(e - \bar{e})(e - \bar{e})^T\right] \]  

(20)

where \( \bar{e} = E[e] \) and \( E \) denotes expected value. The diagonal elements of this matrix are the variances of each of the \( c_n \) terms, while the off-diagonal elements are of the form

\[ C_{nm} = E\left[(c_n - \bar{e}_n)(c_m - \bar{e}_m)\right] \]  

(21)

and can be estimated using

\[ \hat{C}_{nm} = \frac{1}{N} \sum_{i=1}^{N} (A_n(x_i) - \hat{c}_n)(A_m(x_i) - \hat{c}_m) \]  

(22)

We can now compute the covariance matrix \( A \) of the \( a \) vector, starting from the definition

\[ A = E\left[(a - \bar{a})(a - \bar{a})^T\right] \]  

(23)

and then inserting Eqs. (6) and (8),

\[ A = E\left[(A^{-1}c - A^{-1}\bar{e})(A^{-1}c - A^{-1}\bar{e})^T\right] \]
\[ = E\left[(A^{-1}(c - \bar{e})(c - \bar{e})^T A^{-1}] \right. \]
\[ = A^{-1}E\left[(c - \bar{e})(c - \bar{e})^T A^{-1}\right] \]

(24)

so finally

\[ A = A^{-1}C A^{-1T} \]  

(25)

Here \( A^{-1T} \) denotes the transpose of the inverse matrix \( A^{-1} \). Thus, once we know the covariance matrix \( C \) we can compute the covariance matrix \( A = A^{-1}C A^{-1T} \). Note that \( A \) contains the variances of each of the \( a_n \) as its diagonal elements. It should also be noted that \( C \) contains non-diagonal elements that may be either positive or negative, hence treating the \( c_n \) as independent can result in an over-estimate or an under-estimate of the error in the \( a_n \) terms.

We can compute the variances in the random variables \( c_n \) and in the random variables \( a_n \) using the covariance matrices just described. But what we really want is the variance in the estimates of the samples means \( \hat{c}_n \) and \( \hat{a}_n \), for each \( n \). For \( \hat{c}_n \) we appeal directly to the Central Limit Theorem, which gives us directly that \( \hat{c}_n \) is normally distributed about the true mean value \( \bar{c}_n \) with variance
\[ \sigma^2(\hat{c}_n) = \frac{1}{N} \sigma^2(c_n) \approx \frac{c_{nn}}{N} \]  \hspace{1cm} (26)

Through the linear relationship between \(a\) and \(c\) it is clear that each \(\hat{a}_n\) is also an average of \(N\) independent samples of identically distributed random variables, and it was the covariance matrix of these that we computed in Eq. (25). Therefore, the Central Limit Theorem still applies, and we have \(\hat{a}_n\) is normally distributed about the true mean value \(\bar{a}_n\) with variance,

\[ \sigma^2(\hat{a}_n) = \frac{1}{N} \sigma^2(a_n) \approx \frac{a_{nn}}{N} \]  \hspace{1cm} (27)

E. Computational Results using FET with Piecewise Polynomial Basis Functions

The conventional Monte Carlo algorithm utilizes “histogram” tallies for all output quantities, where the score for a cell is simply the sum of the individual scores from each neutron trajectory that contributes to that region, either by colliding in and/or traversing the region. This can be viewed as a piecewise constant expansion of the tally. Since we are trying to develop a FET-based tally methodology that improves on the conventional histogram tally, we decided to look more carefully at local basis functions that span conventional histogram regions, and compare the accuracy and efficiency of local FET versus histogram tallies. As a result, the following schemes were investigated:

- Conventional tallies – histogram tally within each cell
- Local Legendre tallies – local Legendre expansion within each cell
- Continuous finite element tallies – linear, cell-edged finite element basis functions
- Modified continuous hat tallies – linear from cell-center to cell-center

Numerical results for these schemes will now be given. For all of these simulations, the same cells are used for any of the above schemes; the primary differences are (1) the choice of basis functions, (2) the domain of support for the basis functions, (3) the definitions of the expansion coefficients, (4) the location of the expansion coefficients, and (5) whether or not a linear system of equations has to be solved for the expansion coefficients. The last item may be important because the system of equations will be as large as the number of tally cells, which can result in an increase in computational time to solve this system for the expansion coefficients. Table 1 summarizes these properties for each of the above methods that we investigated.

Local Legendre functional expansion tallies use Legendre polynomial expansions within each cell of a spatial mesh. As noted above, local Legendre polynomial expansions suffer from fewer scores as the mesh size is decreased. While this is a well-understood effect with using histogram tallies, it creates more significant problems for the local Legendre expansion tally because the shape of the distribution has to be estimated as well as its average value. As the mesh is refined, the statistical fluctuations in the estimates of these derivatives can have a deleterious effect on the resultant tally. The top plot in Figure 1 illustrates this effect.
Table 1. Properties of Different Tally Methods

<table>
<thead>
<tr>
<th>Tally Method</th>
<th>Basis Functions</th>
<th>Expansion Coefficients</th>
<th>Solve System of Equations?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Domain of support</td>
<td>Continuous across cell edge?</td>
</tr>
<tr>
<td>Histogram (conventional)</td>
<td>Piecewise constants</td>
<td>Cell</td>
<td>No</td>
</tr>
<tr>
<td>Local Legendre</td>
<td>Continuous linear within cell</td>
<td>Cell</td>
<td>No</td>
</tr>
<tr>
<td>Linear FEM</td>
<td>Piecewise linear (hats)</td>
<td>Cell and nearest neighbors</td>
<td>Yes</td>
</tr>
<tr>
<td>Modified FEM</td>
<td>Centered hats</td>
<td>Cell and nearest neighbors</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 1. Legendre Expansion with Large Bins on a Truncated Gaussian
(The errors given in the title are the L2 errors for the differences between the actual and approximate distributions.)
The results in Figure 1 were obtained by sampling from a truncated Gaussian and using the resultant samples to estimate the original distribution using the methods listed in Table 1. This plot shows the conventional histogram tally along with the local Legendre tally for a linear functional expansion within each cell. The slope is not very well estimated near the peak of the distribution where the slope is close to zero and passes from positive to negative.

If there are not enough scores to resolve a linear shape, then the first moment coefficient (i.e., the slope) only adds noise to the tally. To alleviate this problem, a significance test (see Section II.C) is used based on the ratio of the increase in statistical error to the decrease in truncation error due to the addition of the first moment to the tally:

\[ R_1^2 = \frac{\sigma_{a_1}^2 \cdot 3}{\Delta x} \]  

In Eq. (28), \( a_1 \) is the estimated expansion coefficient for the first Legendre moment and \( \sigma_{a_1}^2 \) is the estimated variance in \( a_1 \). Basically, \( R_1^2 \) is the relative variance in the first moment times a normalization factor that accounts for a local linear Legendre expansion over the mesh \( \Delta x \). The first Legendre moment expansion coefficient is set to zero if \( R_1^2 > \frac{1}{1} \). Otherwise the coefficient is assumed to add statistically significant information about the distribution within the tally mesh and is retained. The bottom plot in Figure 1 shows the effect of using the \( R_1^2 \) significance test; namely, the statistically insignificant slope in the central mesh cell did not pass the significance test and was eliminated. As can be seen, the resultant tally yields a more accurate fit visually as well as quantitatively.

In Figure 1, only the central bin’s first moment failed the significance test and was set to zero. Figure 2 shows the truncated Gaussian case with a finer mesh and a larger number of samples. In this case, several of the first moment expansion coefficients did not pass the significance test. It appears that the significance criterion \( R_1^2 > \frac{1}{1} \) was too strict in this case and eliminated first moments that appear to be acceptable, at least by visual inspection. This behavior can be changed by finding an appropriate value of \( R \) for the significance test, i.e., where \( R_1^2 > R \), and this will be a topic of investigation over the next few months.

Other linear FET techniques have also been examined. The linear continuous finite element expansions actually proved to be effective when expanding smooth, continuous functions. However, if an insufficient number of bins are used in the finite element method to characterize the problem, then the function that is to be expanded can appear to be kinked, i.e. discontinuous in the first derivative. Since the finite element expansion is generated by solving a linear system of equations, the appearances of kinks in the function being expanded are anticipated to give fitting errors beyond the mesh in which the kink originated.

Using the same truncated Gaussian distribution as in Figures 1 and 2, the finite element expansion shown in Figure 3 has errors similar to that shown in Figure 1 with the first Legendre
moment expansion. However, when the number of bins and particles is increased, the finite element FET seems to work extremely well as shown in Figure 4.

Figure 4 demonstrates that the finite element expansions are slightly better than the corresponding Legendre expansion (without the cost-benefit test) when using the same bins and number of histories. While the added benefit of continuity is a plus, the linear finite elements add computational complexity because a global linear system of equations must be solved as a post processing step. For a large number of tally cells, this could be a limiting factor.

To maintain continuity but remove the coupled system of equations, two other linear expansion ideas are examined. The first is an approximate collocation scheme with linear interpolation between the collocation points, in which the delta function is approximated by a hat function. The Monte Carlo is then used to estimate the integral of the sampled flux against this approximate delta function. The magnitude of one score in this “hat” tally is weighted by the magnitude of the linear line evaluated at the position of the score. In one-D a given score will contribute to two collocation point tallies that bracket the score. The advantage of this method is the increase in the region that contributes to a tally score, essentially doubling the scoring region for a uniform bin size. Although there is coupling between adjacent cells, there is no need to solve a coupled system of equations for the expansion coefficients.
The tally created by this scoring scheme is normalized by the total number of samples multiplied by the number of meshes $N$ divided by 2 for normalization. A straight line is drawn between the
results of each bin’s normalized tally and constitutes the functional expansion for this tally. This method is called the Modified Hat (MH) expansion.

Another very simple linear interpolation method is to draw a linear line between the centers of neighboring histograms; this is a more reasonable scheme to compare histogram tallies to the piecewise linear tallies. To effectively match the histogram interpolation with the MH expansion, the histograms at the edge of the entire problem are cut in half while the interior histograms keep their original width. This effectively allows the interpolation points of the MH, histogram and the finite element methods to be directly compared in Figures 5 and 6.

All of the methods should appear highly correlated in Figure 5 since they all are using the exact same samples to calculate their distribution. While the results vary from run to run, the MH on average has the lowest error followed by the histogram interpolation and finally the finite element method as shown in Table 2. With more particles run, the finite element method has the smallest error on average, then the histogram method, then finally the MH method as shown in Table 2. Figure 6 plots the results for the same case as Figure 5 but with more particles.

<table>
<thead>
<tr>
<th>Number of Particles</th>
<th>Finite Element Error</th>
<th>Modified Hat Error</th>
<th>Average Histogram Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.063</td>
<td>0.049</td>
<td>0.054</td>
</tr>
<tr>
<td>10000</td>
<td>0.020</td>
<td>0.030</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Figure 5. Histograms, Modified Hats, and Finite Elements on a Truncated Gaussian with 1000 Particles
F. Adaptive coarsening method

This method is an extension of the discontinuous Legendre FET to allow adaptive coarsening of the bins depending on the problem data. The idea is to use the uncertainty filter discussed above and if the first moment of a given cell has a large uncertainty ($R_1 > 1$), merge the cell with its neighboring cell and expand the solution in a discontinuous Legendre basis for the larger cell. This is straightforward to do because it only requires the Legendre moments over the original cells, not the sampled data. Figure 7 shows a comparison of a local Legendre (discontinuous Legendre) FET over a mesh that is too fine, hence yielding poor statistics, compared to a coarsened mesh using the above method. These results are preliminary but are promising enough to warrant a slight change in direction for this research project. We plan to direct effort to this method in the coming months and move away from the use of finite element basis functions.

![Figure 6. Histograms, Modified Hats, and Finite Element with 10,000 Particles](image)

G. Comments

However, there is a drawback to the use of local basis functions for the FET: a neutron will contribute to the score for a given expansion coefficient if it moves (or collides) in a region of local support for that basis function. For global basis functions, every neutron trajectory contributes to every expansion coefficient, which results in a reduced variance compared to the use of local basis functions. So there is a tradeoff between truncation error and statistical error, as previously reported [7-8]. In general, a larger domain of support for the basis function translates to more “counts” hence decreased variance, but for a realistic flux a larger basis of support translates into a larger truncation error.
III. Using FET to Accelerate Fission Source Convergence

A. Implementation in MCNP4C

In order to test the use of functional expansion tallies for eigenvalue calculations, the FET estimators given in Eqs. (7) and (8) were implemented in a modified version of MCNP4C [11]. The set of Legendre polynomials was chosen as the expansion basis set. The initial version of the code is designed to estimate the first 20 spatial Legendre moments of the fission source. In this test code, both the collision and track length FET estimators operate concurrently with the original fission bank method. This means that each neutron generation produces three separate approximations for the spatial shape of the next fission source in our test code. The birth sites for the following generation can then be sampled from any of the three source approximations.

For the FET source approximations, particle birth locations are sampled directly from the appropriate functional approximation using simple rejection sampling. To determine the functional approximation, the code considers all 20 spatial source moments estimated during the previous generation and rejects (sets to zero) those moments that have a cost-to-benefit ratio \((R^2)\) greater than a user defined threshold. This filtering helps to reduce statistical noise in the functional approximation. The remaining coefficients are then normalized by the zeroth coefficient and used in the Legendre series expansion to give a normalized polynomial source.
approximation. For rejection sampling, sample points along the polynomial can be quickly and efficiently evaluated using the Legendre recursion relationships.

At birth, the initial direction of a source neutron is sampled isotropically in angle. The initial energy of the source neutron is sampled from the Maxwell fission energy spectrum with a mean temperature of 1.2895 MeV

\[ P(E) = 0.77059 e^{-\left(\frac{E}{0.77059}\right)} \sqrt{E} . \]  

(29)

Note that the distribution in Eq. (29) does not take into account the energy of the neutron that causes the fission event. Unlike the fission bank method, the 1-D FET source convergence method does not retain any information about the parent neutrons for the source particles. This is because the source neutrons in the FET method are not directly associated with a single parent neutron. Rather they are sampled independently from the approximate source distribution created from all of the neutrons in the previous generation. This independent sampling strategy means that an unlimited number of source locations can be generated, without concern about intra-generation correlation between the samples.

For the fission bank source approximation, particle birth locations are taken directly from the list of stored fission sites during the previous generation. The internals of the fission bank source iteration scheme remained unchanged from the original distribution version of MCNP4C, with two minor exceptions. First, the source particle energies stored in the fission bank were not used. Instead, the initial energy of each source particle was sampled from the Maxwell fission spectrum given in Eq. (29), as was done with the FET approach. This resampling in energy allows a fair comparison between the FET and fission bank methods. Also, MCNP4C was modified to give a histogram edit for the contents of the fission bank after each generation. This change was made to allow easier visualization of the fission bank source shape.

B. Numerical results

For numerical verification of the FET source convergence method, the modified version of MCNP was run on two one-dimensional homogeneous benchmark problems, a fast reactor problem characteristic of a tightly coupled system and a thermal reactor problem characteristic of a loosely coupled system.

**Fast reactor test problem.** A homogeneous fast-spectrum critical assembly was modeled as a 100 cm bare slab of pure uranium with a density of 10.97 g/cc. The isotopic composition of the uranium was 6.5% U\textsuperscript{235} and 93.5% U\textsuperscript{238}. In order to characterize this system a reference criticality calculation was performed using an unmodified version of MCNP4C. The reference calculation used 500 generations of 10,000 neutron histories. The first 50 generations of the simulation were discarded to avoid biasing the results with the initial (poor) source distribution. The simulation produced an eigenvalue estimate of 0.96674 with a relative standard deviation of 0.00021.

When compared with the diameter of the reactor, the mean free path of neutrons in the system is relatively large. This means that neutrons are able to spread from one side of the reactor to the other in only a few generations. Systems such as these are typically referred to as tightly coupled systems. In practice, source iteration schemes are very efficient for tightly coupled systems because local perturbations within the fission source shape are quickly dissipated, allowing rapid
convergence of the global solution. Thus, it should be anticipated that both the fission bank and FET source convergence methods should provide fast convergence to the critical fission source distribution.

To compare the FET and fission bank source convergence methods, two separate k-code calculations were run. Each simulation used 3,000 source generations with 10,000 neutron histories per generation. Because the initial convergence of the fission source is of primary interest, no source generations were discarded. During each generation the collision FET was used to estimate the first 20 spatial Legendre expansion coefficients of the fission source. These coefficients were filtered using a cost-to-benefit ratio threshold of 0.95. Any coefficients with an $R^2_n$ value greater than 0.95 were set to zero. For sampling source points for the following generation, one simulation used sites stored in the fission bank, while the second simulation sampled directly from the FET functional approximation.

For both sampling methods, as the fission source converges, the Legendre coefficients will each converge to a constant value. These constants are the source moments for the converged source distribution. As long as any of the expansion coefficients are showing systematic changes between generations the fission source is not converged and additional generations will be required in the source iteration process.

Figure 8 shows a plot of the first through the fourth Legendre coefficients for the fission bank and FET source convergence methods as a function of the source generation number. Results are shown for both the fission bank (red) and collision FET (blue) source sampling routines. Both sets of results were generated in independent Monte Carlo criticality calculations using 3000 generations and 10,000 histories per generation. Legendre moments with a cost-to-benefit ratio greater than 0.95 were considered poorly converged and filtered out of the functional approximation. As expected, the coefficients for both convergence methods appear to converge immediately to the same values and remain nearly constant over all of the generations. The first and third coefficients appear randomly distributed about zero, while the second and fourth coefficients appear randomly distributed about -0.18 and 0.007, respectively. The empty gaps that appear in the first, third and fourth moments are a result of the $R^2_n$ filtering used to reduce statistical uncertainty in the functional approximation. Data points that would normally appear within these gaps contain large statistical uncertainties and therefore were filtered out and set to zero. The width of these gaps can be reduced by either raising the $R^2_n$ filter threshold or increasing the number of histories per generation.
Based on the coefficient plots, both the fission bank and FET convergence methods produce nearly identical results for this benchmark problem. This is confirmed by a visual comparison between the FET fission source approximation and a histogram representation of the fission bank data taken after cycle number 3000, shown in Figure 9. This comparison highlights the excellent agreement between the fission source shapes produced by the two methods. These results verify that the new FET based method is implemented correctly and that it can be used to accurately estimate the converged fission source distribution for this problem.
Figure 9. Comparison of spatial source distribution produced by the fission bank and collision FET source sampling routines for the fast-spectrum critical assembly benchmark.

**Thermal reactor test problem.** The second series of benchmark tests were conducted on a homogenous, one-dimensional, thermal-spectrum critical assembly. The assembly was modeled as a 100 cm bare slab of a homogenized water and uranium mixture with a density of 10.97 g/cc. The ratio of hydrogen to uranium atoms in the mixture was 900:1. The isotopic composition of the uranium was 40% $^{235}\text{U}$ and 60% $^{238}\text{U}$. The system was again characterized by a reference criticality calculation using an unmodified version of MCNP4C. As with the previous benchmark, the reference calculation used 500 generations of 10,000 neutron histories, with the first 50 cycles being discarded. The estimated eigenvalue from the simulation was found to be 0.96531 with a relative standard deviation of 0.00010, and the mean free path of neutrons in the simulation was 0.113 cm.

The mean free path of neutrons in the water/uranium assembly is much smaller than the mean free path in the pure uranium assembly analyzed above. The smaller mean free path means that it will take neutrons in the system many generations to spread between local regions in the assembly. A system like this is commonly referred to as being loosely coupled. In general, source iteration techniques will converge slowly for loosely coupled systems. In fact, the convergence rate depends heavily on how loosely coupled the system under consideration is. The problem with loosely coupled systems is that separate local regions of a single global system often start to converge independently of one another. For example, during the early source generations, an extremely wide reactor may behave neutronically like two, or more, smaller reactors placed side by side. Only after many generations (typically thousands) do these local sources begin to coalesce into the global solution. Unfortunately, for large scale problems, such as commercial reactor cores, running a sufficiently large number of neutron generations to achieve source convergence with current Monte Carlo codes is computationally very expensive. It is hoped that the FET source convergence method can help to accelerate the initial fission source convergence rate for loosely coupled systems, thus reducing the overall computational cost.
To test this hypothesis, a series of Monte Carlo eigenvalue calculations were conducted on the water/uranium system. Each calculation used 1000 generations with either 10-, 50-, or 100,000 neutron histories per generation. Both collision and track length tallies were used to estimate sets of 20 Legendre expansion coefficients of the fission source. These coefficients were filtered using a cost-to-benefit ratio threshold of 0.6. Source points for the next generation were either taken directly from the fission bank or sampled from the FET source approximations. The combination of the three source sampling routines and the three different generation sizes gave a total of nine independent criticality simulations used for this study.

As with the fast-spectrum system, the convergence of the fission source was assessed by examining the convergence of the individual Legendre moments as a function of the generation number. Figures 10 through 12 shows the first through sixth expansion coefficients for all three source convergence methods. For all three figures, results are shown for the conventional fission bank (red), collision FET (blue), and track length FET (green) source sampling routines. All results were generated with independent Monte Carlo criticality calculations using 3000 generations and 10,000 histories per generation. Legendre moments with a cost-to-benefit ratio greater than 0.95 were considered poorly converged and filtered out of the functional approximation.

The convergence behavior of the expansion coefficients in the thermal system is strikingly different than that observed in the fast spectrum benchmark. In the fast system the moments appeared to change by a large margin between generations, but they always appeared to be approximately randomly distributed about a mean value. In the thermal system the values of individual moments change relatively little between generations. Furthermore, in the thermal system the moments are obviously not converged during the early cycles. Rather, each moment shows a steady trend from generation to generation. This trending clearly illustrates the convergence of the source shape. Statistical noise in the source shape of the thermal system appears as slow drifts over time in each of the moments, instead of large oscillations about a central value. These slow drifts make it difficult to assess when, and where a particular moment has converged.

For this example it is useful to compare the numerical results with an analytical approximation to the source shape, in order to verify that the methods are converging to the correct shape. Because this benchmark problem is a bare, homogenous, critical slab it is reasonable to expect that diffusion theory should apply and the fission source should be roughly a cosine shape over much of the slab. For comparative purposes the first 6 exact Legendre moments for a cosine over the slab are listed in Table 3. We note that the continuous energy Monte Carlo solutions should not be expected to converge to exactly the cosine shape. However, the fact that the dominant second order expansion coefficient appears to converge to a value close to the simple one-group diffusion approximation provides an indication that the methods are working as predicted.
10,000 Histories per Generation

$\alpha_1$

$\alpha_2$

$\alpha_3$

$\alpha_4$

$\alpha_5$

$\alpha_6$

Figure 10. Low order Legendre moments of the fission source by generation for the thermal-spectrum critical assembly for 10,000 neutrons/generation.
50,000 Histories per Generation

Figure 11. Low order Legendre moments of the fission source by generation for the thermal-spectrum critical assembly for 50,000 neutrons/generation.
100,000 Histories per Generation

In the simulation with 10,000 histories per generation (Figure 10), both FET source sampling routines clearly outperform the fission bank. The fission bank solution has a large 1st moment,
which indicates that the distribution is tilted towards one side of the assembly. The FET source method, by contrast, is able to quickly converge the odd moments to zero, which is physically correct for a symmetric system. In the simulation with 50,000 histories per generation (Figure 11), the FET appears to show faster convergence than the fission bank for the second coefficient (the dominant term). In fact, it appears as though the FET estimates for $a_2$ reach convergence near generation 400, several hundred cycles before the fission bank estimates. At 100,000 histories per generation (Figure 12) all three source sampling routines appear to show roughly the same performance. However, even with this large number of histories per generation the FET sampling methods seem to do slightly better on the odd moments. For all of these test cases, there appears to be no significant difference between the collision FET and track length FET source convergence methods.

Table 3. Exact Legendre expansion coefficients for a cosine fission source distribution.

<table>
<thead>
<tr>
<th>Legendre Order (n)</th>
<th>Expansion Coefficient ($a_n$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-2.60</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0.00903</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>-0.000161</td>
</tr>
</tbody>
</table>

Note that, in some cases, the results appear to converge, but may suddenly drift away from convergence (i.e. get a large first moment component) and then drift slowly back towards the original converged value. This behavior is observed for both the FET and the fission bank methods. The magnitude of this drift is, in some cases, surprisingly large. The causes of these drifts are not well understood. One possible explanation is that the drifts are due to a temporary false convergence towards a higher (non-stable) eigenmode. It is clear that analysis of the individual Legendre moments can provide valuable information about the convergence of the fission source shape.

For smaller numbers of histories per generation, the fission bank has more noise in the source shape than the FET based method. This difference is illustrated by a comparison of the functional approximation and a histogram representation of the fission bank data, shown in Figure 13. The blue line shows the 20th order Legendre approximation to the source shape. The red bars show a histogram representation of fission bank. Results used 3000 neutron generations with 10,000 histories per generation. Data shown are for the fission source produced after generation 3000.

To a large extent this reduction in noise is due to the filtering of moments with high statistical uncertainty. In order to study the effect of different filtering thresholds on the source convergence a series of three additional criticality calculations were conducted using the collision FET source sampling method. Each of these additional calculations used 1000 generations with 10,000 neutron histories per generation. The filtering threshold for each of the three simulations was set to 0.3, 0.6 and 0.95, respectively. Plots of the first through sixth expansion coefficients are shown in Figure 14. All results were generated in independent Monte
Carlo criticality calculations using collision FET source sampling with 1000 generations and 100,000 histories per generation.

![Fission Source Distribution](image)

Figure 13. Comparison of spatial source distribution produced by the fission bank and collision FET source sampling routines for the thermal-spectrum critical assembly.

With the largest threshold, 0.95, the FET allows a large amount of fluctuation in the first and third expansion coefficients, but generally does a good job in quickly converging the important second and fourth moments. Reducing the threshold to 0.6 eliminates the fluctuation in the odd moments, but maintains the rapid convergence of the even terms. In the final simulation, the threshold was further reduced to 0.3. The results show that this threshold is set too low, and as a result, the FET filtered out important $a_6$ coefficients during the early generations. Without these initial $a_6$ terms, the fission source actually takes longer to converge than with the other filter threshold values.

These results indicate that the FET based method is a viable alternative to the traditional fission bank approach and offers several advantages over the traditional fission bank approach. The data can be filtered to eliminate some of the statistical noise present in the results. Also, the new method is not limited to a fixed number of source points per generation like the fission bank. This property of the FET based method means that it is possible to use a variable number of neutron histories per generation. In fact, one promising idea is to allow each generation to run until a number of expansion coefficients have passed some user defined convergence test, before starting the next generation.

Aside from accelerating the source convergence, tallying the expansion moments of the source distribution may also provide a new and interesting method for studying source convergence behavior.
Figure 14. Filtered low order Legendre moments of the fission source by generation for the thermal-spectrum critical assembly. Results are shown for the cost-to-benefit ratio filter threshold values of 0.95 (red), 0.6 (blue), and 0.3 (green).
C. Alternative FET-like Methods for Fission Source Convergence

Since the primary emphasis of this research grant was to accelerate fission source convergence with FET, two alternative methodologies have been studied with promising results. The first method below, the kernel density estimator, is the PhD thesis research for Kaushik Banerjee. The second method, implementing Arnoldi’s method with FET, is the PhD thesis research of Jeremy Conlin. These methods are discussed below.

1. Kernel Density Estimator

The implementation of the kernel density estimator (KDE) for fission source iterations can be viewed as a dynamic FET scheme because the samples (e.g., fission sites) are used implicitly to construct an expansion of the fission source in terms of a prescribed basis set, perhaps a Gaussian or a polynomial with finite support, for example. This section describes the KDE method as applied to Monte Carlo fission source convergence and presents preliminary results.

The Kernel Density Estimator (KDE), a nonparametric density estimator, is studied and applied to Monte Carlo eigenvalue calculations for reactor analyses. The Monte Carlo method is an important tool for performing reactor calculations because of its ability to handle complex geometry and complex physics. One major drawback of Monte Carlo is slow fission source convergence rates for large, loosely coupled systems. In this work, KDE is used to estimate the shape of the fission source at the end of each neutron generation and samples from this estimated source distribution are used as the starting particles for the next generation.

Previous applications of the KDE method to nuclear engineering have been to resample electron trajectories as part of a response matrix approach [13] and to resample photon trajectories in the phase space resulting from the patient-independent portion of a radiation transport calculation as part of a cancer treatment plan [13]. Both of these applications involve resampling of electron or photon trajectories close to previously sampled trajectories and the application of KDE to fission source iterations is a natural extension.

Our preliminary results for a large, loosely coupled system in 1D slab geometry are promising: the converged fission source distribution is satisfactory and there is a substantial increase in fission source convergence as measured by the relative source entropy [14,15]. We also note that the traditional sampling of the fission source introduces correlation among the source points between neutron generations [11] and it is hoped that the KDE method will reduce this correlation by decreasing the propagation of the correlated source particles from generation to generation. Finally, KDE can be used to represent Monte Carlo tallies instead of histograms, resulting in a smooth estimation of the tally distribution.

**Methodology.** Consider \( n \) real observations \( X_1, \ldots, X_n \) whose underlying density (\( f(x) \)) is to be estimated. The kernel estimator with kernel \( k \) for univariate data is defined by

\[
\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right)
\]  

(30)
where \( h \) (\( h \to 0 \) as \( n \to \infty \)) is the band width, also called the smoothing parameter. Usually, but not always, \( k \) will be a symmetric probability density function (pdf) about zero, the normal density for instance. The basic properties of \( \hat{f}(x) \) are well known from the literature [16]:

\[
\text{Bias} = E\left\{\hat{f}(x)\right\} - f(x) = O(h^2) \tag{31}
\]

and

\[
\text{var}\hat{f}(x) = \frac{1}{n} \left[ \int \frac{1}{h^2} k\left(\frac{x-y}{h}\right)^2 f(y)\,dy - \left\{ \frac{1}{h} \int k\left(\frac{x-y}{h}\right)f(y)\,dy \right\}^2 \right] \tag{32}
\]

The estimator \( \hat{f}(x) \) itself is a pdf, normalized to 1, as the kernel function \( \frac{1}{h} k\left(\frac{x-X_i}{h}\right) \) is normalized to 1. It is straightforward to sample from the estimator \( \hat{f}(x) \). Realizations \( Y \) from \( \hat{f}(x) \) can be generated as follows [16]:

1. Choose \( I \) uniformly with replacement from \( \{1, \ldots, n\} \) by \( I = \xi n + 1 \), where \( \xi \) is the random number between 0 and 1
2. Generate sample \( \varepsilon \) from the kernel \( k \) (e.g., Box-Muller if \( k \) is Gaussian).
3. Set \( Y = X_I + h\varepsilon \)

Therefore, the KDE method is a variation on FET in the sense that the samples are expanded in terms of basis functions but the coefficients of the expansion are not explicitly calculated this is all done on the fly.

The most widely used way of estimating the global accuracy of \( \hat{f} \) as an estimator of \( f \) is the mean integrated square error (abbreviated as MISE), which is defined by [16]

\[
\text{MISE}(\hat{f}) = E\int \left(\hat{f}(x) - f(x)\right)^2 \, dx \tag{33}
\]

The global band width \( (h) \) is selected from the point of view of minimizing the approximate mean integrated square error.

However the performance of \( \hat{f}(x) \) is poor near the boundaries due to the boundary effect that occurs in nonparametric curve estimation problems. This is due to the fact that the bias of \( \hat{f}(x) \) is of \( O(h) \) instead of \( O(h^2) \) at boundary points. Whenever data points are near boundaries the associated kernel will not integrate to unity over the problem domain because it overlaps the
boundary. Simply truncating the estimator at the boundaries and then renormalizing \( \hat{f}(x) \) to unity will not solve the problem as noted by Silverman [16] and also tested by the authors. To remove the boundary effect we have examined several methods that have been proposed [17]:

- **Reflection** – reflect the sampled point back into the computational domain.
- **Boundary kernel** – define a separate kernel for data points within a small distance (e.g., \( h \)) of the boundary.
- **Transformation** – transform the sampled domain within a short distance of the boundary to minimize the boundary bias.

Reflection boundary correction works well if the first derivative of the density function is zero at the boundary. However, we found that it does not work well in general. We have found that transformation seems to work best for the fission source application. Karunamuni and Alberts [17] present a transformation that preserves an \( O(h^2) \) boundary bias for a single boundary at \( x=0 \). We have extended this work to a slab of width \( a \) with two boundaries at \( x=0 \) and \( x=a \), resulting in the following estimator:

\[
\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} \left[ k\left( \frac{x-g_1(X_i)}{h} \right) + k\left( \frac{x+g_1(X_i)}{h} \right) \right], \quad 0 \leq x \leq h
\]

\[
= \frac{1}{nh} \sum_{i=1}^{n} k\left( \frac{x-X_i}{h} \right), \quad h < x < a - h
\]

\[
= \frac{1}{nh} \sum_{i=1}^{n} \left[ k\left( \frac{x-g_2(X_i)}{h} \right) + k\left( \frac{x-2a+g_2(X_i)}{h} \right) \right], \quad a - h \leq x \leq a
\]  

(34)

where \( h \) is the bandwidth \( k \) is the kernel function with support \([-1,1]\) and \( g_1 \) and \( g_2 \) are the boundary transformations. It can be shown that the above estimator has a boundary bias of \( O(h^2) \). The transformations are given by

\[
g_i(y) = y + 0.5d_0k'_0 y^2 + 0.5\left( d_0k'_0 \right)^2 y^3
\]

(35)

where

\[
d_0 = \frac{f^{(1)}(0)}{f(0)}
\]

(36)

\[
k'_0 = \frac{2\int_{c_0}^{1} (t-c_0)k(t)dt}{c_0 + 2\int_{c_0}^{1} (t-c_0)k(t)dt}
\]

(37)
and
\[ g_2(y) = (2a - y) + 0.5d_a k'_{c_a}(a - y)^2 + 0.5\left(d_a k'_{c_a}\right)^2 (a - y)^3 \] (38)

where
\[ d_a = \frac{f^{(l)}(a)}{f(a)} \] (39)

\[ k'_{c_a} = -\frac{2\int_{c_a}^{c}(t-c_a)k(t)dt}{c_a + 2\int_{-1}^{c_a}(t-c_a)k(t)dt} \] (40)

and \( c_0 = \frac{x}{h}, \ c_a = \frac{x-a}{h} \). The logarithmic derivatives \( d_0 \) and \( d_a \) are estimated by conventional FET. We have found that the estimation of these terms is important for obtaining good results near the boundaries.

**Numerical results.** The first problem is a simple example to demonstrate that KDE can reproduce a known pdf, in this case a clipped exponential on the interval \([0, 2]\). To depict this graphically, 100,000 data points were drawn from \( f(x) \) and Eq. (34) was used to evaluate \( \hat{f}(x) \) at 200 equally spaced points between 0 and 2. These points are plotted in Figure 15. To illustrate the importance of the boundary correction terms, the 200 points were evaluated without the boundary correction terms (using the same set of 100,000 data points) and are plotted in Figure 15. It is clear that the general boundary correction algorithm contained in Eq. (34) is necessary to avoid poor estimation of the pdf near the edges of the distribution.

![Figure 15. Using KDE with boundary correction for Example 1.](image-url)
Figure 15 also illustrates the potential for KDE to plot tallies, since the tally points can be treated as if they were un-normalized samples from a pdf and then plotted as we have done in Figure 15.

The second example is to predict the criticality of a large slab. Monte Carlo criticality calculations with KDE and the conventional fission source method are carried out for a 100 mfp (mean free path) wide multiplying slab. With KDE, the starting neutrons for each cycle are sampled from the estimated source distribution of the previous cycle using Eq. (34). Both simulations used 100,000 particles per batch and 3000 batches. Figure 16 presents the comparison between the conventional and KDE methods. The binned sources for both cases are calculated by averaging the source distributions over the final 1000 cycles. This averaged source is also used for the relative entropy calculation [14].

It is clear from Figure 16 that the KDE method has improved source convergence. Based on our parametric studies of 1D versus 2D for simple problems, we believe the advantage of KDE over conventional source iteration will increase for multi-D problems.

The use of KDE to implement fission source iteration in criticality problems has been shown to be successful in simple 1D geometries. KDE results in substantially faster source convergence than conventional fission source iteration for simple 1D system. A boundary correction method was developed and implemented and yields accurate fission source distributions near the boundaries. The results show the feasibility of the KDE approach and we are now extending this methodology to multi-D geometry and expect to have results soon.
2. Arnoldi’s Method with FET.

Krylov Subspace Methods. The power method is the method of choice for Monte Carlo criticality calculations. The power method generates a Krylov subspace [18] by applying a linear operator, $A$, to a starting vector, $v$, repeatedly. The result of these operations is a set of vectors, defining a Krylov subspace, $K_m = \text{span}\{v, Av, A^2v, \ldots, A^{m-1}v\}$.

After each iteration, the power method calculates a new estimate of the dominant eigenvalue. The associated eigenvector estimate is $A^2v$ and approaches the true eigenvector as $k$ gets larger. The variance of the average of the calculated eigenvalues is a measure of the statistical uncertainty in the eigenvalue estimate.

The linear operator $A$ does not have to be known explicitly to be able to use the power method, rather we only need to know how to apply the linear operator $A$ to the vector $v$. In Monte Carlo particle transport the vector $v$ is a fission source. We can apply the operator $A$ by transporting the fission source and generating a new fission source from the resulting fission sites.

Arnoldi’s method. Arnoldi’s method of Minimized Iterations [19] is another method used for calculating eigenvalues. Like the power method, Arnoldi’s method only requires knowing how to apply the operator $A$ to the vector $v$. Arnoldi’s method uses the same Krylov subspace as the power method but builds an orthonormal basis for it. The process of orthogonalizing and normalizing the basis vectors creates an upper Hessenberg matrix, $H$, of size equal to the number of basis vectors; $H$ is the orthogonal projection of $A$ onto the Krylov subspace. The Ritz pairs of $H$ are easily found because the size of $H$ is small; generally much smaller than $A$. The Ritz pairs of $H$ approximate the eigenpairs of $A$.

Arnoldi’s method requires all the basis vectors be stored in order to orthogonalize them. The memory requirements for orthogonalization of Arnoldi’s method increases with the size of the Krylov subspace. A way to reduce these memory requirements is to restart Arnoldi’s method after a predetermined number of iterations (size of Krylov subspace). At the end of an Arnoldi restart, a Ritz pair is calculated. The Ritz vector is used as the starting vector of a new Arnoldi calculation. This process can be repeated indefinitely. An Arnoldi restart can be treated similar to a power method iteration. A Ritz pair is calculated at the end of each restart, the variance of the average of these eigenvalues is the statistical uncertainty.

Recently Warsa et al. [20] have shown that Arnoldi’s method provides a significant decrease in the overall computational expense as compared to the power method for a deterministic $S_N$ transport code. This effect was even more dramatic for problems with poor dominance ratios. In this paper we show that Arnoldi’s method can be used in a Monte Carlo criticality application.

Implementation of Arnoldi’s method. The orthogonalization of the basis vectors in Arnoldi’s method requires taking the inner product of two vectors. In our studies, we have discretized the fission source into spatial bins and so have represented the infinite-dimensional fission source as a B-dimensional vector.
\[
q(x) = \sum_{i=1}^{B} a_i f_i(x) 
\]

where \( B \) is the number of spatial bins and

\[
f_i(x) = \begin{cases} 
1 & , \ x_i \leq x \leq x_{i+1} \\
0 & , \ \text{otherwise}
\end{cases}
\]

Representing the fission source vector in this manner, taking the inner product of two fission sources is simply the dot product of two vectors of coefficients \( \{a_i\}_{i=1}^{B} \). This introduces a discretization error from which the power method does not suffer when calculating the eigenvalue of \( A \). The power method creates a similar discretization error when binning tally results.

Operating \( A \) on a vector of coefficients \( \{a_i\}_{i=1}^{B} \) simply requires sampling the piecewise constant source function \( q(x) \) in Eq. (41), transporting these neutrons until they cause another fission and tallying the resulting fission neutrons over the bins.

The source function \( q_{n-1}(x) \) at Arnoldi iteration \( n-1 \) will generally be negative because it has been orthogonalized against previous fission sources. We define the normalized source distribution

\[
P(x) = \frac{|q_{n-1}(x)|}{\int|q_{n-1}(x)|dx}
\]

and sample source locations \( x_i \) from \( P(x) \) and give the corresponding particle an initial weight \( w = q_{n-1}(x)/|q_{n-1}(x)| \). This ensures sampling each bin with correct probability and preserving the correct sign for the particle weight. Monte Carlo transport then proceeds as usual with the neutron scoring \( w(\nu \Sigma_t / \Sigma_f) \) in the proper spatial bin at each collision. The resulting discrete fission source \( \hat{q}_n(x) \) is normalized after each iteration by the number of particles thrown per iteration, \( N \),

\[
q_n(x) = \frac{\int|q_{n-1}(x)|dx}{N} \hat{q}_n(x)
\]

Here the integral of the fission source from the previous iteration, \( \int|q_{n-1}(x)|dx \), is the strength of the previous source including both positive and negative sources. This scheme is equivalent to sampling a total of \( N \) histories from the positive and negative sources, with the ratio of the number of histories from the positive to the number of histories from the negative source equal to the ratio of the positive to the negative source strengths.
Numerical Results. We have performed some preliminary calculations using Arnoldi’s method to calculate the dominant eigenvalue and eigenvector for a 1-D, homogeneous slab of width equal to 20 mfp ($\Sigma_s = 0.5, \nu \Sigma_\epsilon = 0.5$) surrounded by vacuum. We used 5000 histories per iteration, 50 Arnoldi restarts, and five Arnoldi iterations per restart. The power method also used 5000 histories per iteration and 250 iterations; both methods had the same total number of active histories. Both methods discarded the first 10 iterations (2 Arnoldi restarts) to allow for source convergence. In Arnoldi’s method, we discretized the space into 50 equally spaced bins.

The eigenvalue convergence of the active cycles is shown in Figure 17 for both methods. The errorbars indicate one standard deviation from the estimated mean eigenvalue. The true eigenvalue (calculated with an $S_N$ code and compared to published results [21]) is marked as $\lambda_0$. In Figure 18, the estimated eigenvector is shown for both methods as well as an $S_N$ method for comparison. Errorbars have been left off this figure due to space limitations, but are on the order of 0.005 for both methods.

![Eigenvalue convergence of Arnoldi’s method compared to power method.](image)

Figure 17. Eigenvalue convergence of Arnoldi’s method compared to power method. Arnoldi’s method uses 5000 histories per iteration, 5 iterations per restart, and 50 active restarts – 2 restarts were discarded. The power method also used 5000 histories per iteration with 250 active iterations – 10 iterations were discarded. Only the active iterations are shown. $\lambda_0$ is the correct eigenvalue.
Conclusion. We have presented preliminary results showing that Arnoldi’s method can be used instead of the traditional power method to accurately calculate the dominant eigenvalue and eigenvector in a Monte Carlo particle transport calculation. Further studies are required to fully understand the effect of changing the parameters of Arnoldi’s method (e.g. number of iterations per restart, starting vector, spatial discretization). Based on the work of Warsa et al. [20] we can expect that Arnoldi’s method will be better able to converge the eigenvector than the power method for problems with dominance ratio close to 1.

Figure 18. Fission source distribution from power method and Arnoldi’s method with $S_N$ results for comparison. Arnoldi’s method uses 5000 histories per iteration, 5 iterations per restart, and 50 active restarts – 2 restarts were discarded. The power method also used 5000 histories per iteration with 250 active iterations – 10 iterations were discarded.

IV. Development of Monte Carlo-based Response Matrix Method

A. Background
This section is based on the original proposal for this grant and is included here for completeness. The conventional response matrix equations, expressed in terms of FET, may be expressed as follows:
\[ J^{'*} = R_{S-V} J^{'*} + \frac{1}{k} R_{S-V} \phi \]
\[ \dot{J}^*' = H \dot{J}^*' \]
\[ \dot{\phi} = R_{V-V} \dot{J}^*' + \frac{1}{k} R_{V-V} \dot{\phi} \]

where \( \phi \) is the vector of expansion coefficients representing the global solution as an expansion in an appropriate basis:

\[ \phi(r, \Omega, E, t) = \sum_{n=1}^{M} a_{n} \psi_{n}(r, \Omega, E, t) \]

The matrices appearing in Eq. (45) include a surface-to-surface response matrix \( R_{S-S} \), a volume-to-surface matrix \( R_{S-V} \), a surface-to-volume response matrix \( R_{V-S} \), a volume-to-volume response matrix \( R_{V-V} \), and a “permutation” matrix \( H \) that simply re-orders incoming quantities as outgoing quantities depending on the region being considered. The response matrices are readily evaluated by Monte Carlo if the basis functions are orthogonal and the lowest order basis function (n=0) mode is flat, such as with Legendre polynomials, because the response of the system due to an arbitrary mode \( \psi_{n} \) can be estimated with Monte Carlo by using superposition of the given mode and the fundamental mode, which guarantees a positive pdf to sample from. In essence, the Monte Carlo method is determining the “Green’s function” response for a unit input which is either an arbitrary volumetric basis function \( \psi_{n}(r, \Omega, E, t) \) for \( R_{S-V} \) and \( R_{V-V} \) or an arbitrary surface basis function for \( R_{V-S} \) and \( R_{S-S} \). Since Monte Carlo is being used for the determination of the response matrices, the resultant response matrix accounts for any complex physics or geometric detail, as long as it can be resolved by the “local” Monte Carlo problem.

B. Alternative FET-based response matrix approach

**Change in approach.** The fission matrix method based on FET that is developed below. The original intent was to develop a response matrix approach based on using FET to estimate the fission source terms, however this effort was stopped about a year ago as a result of recently identified work performed by researchers at Georgia Institute of Technology [22,23]. Basically, they have developed a global response matrix method that is based on using orthogonal function expansions on the surfaces of internal regions, thus computing exactly the effective “Green’s function” noted in the paragraph above, i.e., the region response due to a unit surface basis function input. Moreover, they have used Monte Carlo to evaluate the region responses, essentially identical to what we were proposing to do with FET as part of this grant. They used the methodology developed by us [2,3] to carry out the local response calculations and improved on it by using an innovative iteration strategy that does not require the use of volume-to-surface and volume-to-volume response matrices, resulting in a huge savings in computational effort and complexity. Thus, they have developed and published a key component of the methodology we were hoping to develop for the FET-based response matrix portion of this grant. As a result, the
decision was made to redirect our efforts towards an alternative response matrix approach rather than continue down the path described in the proposal for this grant.

**Alternative formulation using the fission matrix.** An alternative approach that utilizes FET to accelerate fission source convergence is a modification of the well-known fission matrix method. The fission matrix method may be described succinctly beginning with the usual transport equation in operator form:

$$A \psi(r, \hat{\Omega}, E) = \frac{1}{k} M \psi(r, \hat{\Omega}, E)$$

(47)

where $\psi(r, \hat{\Omega}, E)$ is the angular flux, $A$ is the transport-collision operator, and $M$ is the angular fission source operator:

$$A = \hat{\Omega} \cdot \nabla + \Sigma_s(r, E) - \int_0^\infty (\int_{4\pi} d\hat{\Omega}' \Sigma_s(r, E' \rightarrow E, \hat{\Omega}' \rightarrow \hat{\Omega}))$$

(48)

$$M = \frac{1}{4\pi} \chi(r, E) \int_0^\infty (\int_{4\pi} d\hat{\Omega}' \nu \Sigma_f(r, E'))$$

(49)

Now formally invert Eq. (47) for $\psi(r, \hat{\Omega}, E)$ and integrate over angle to obtain the scalar flux:

$$\phi = \frac{1}{k} A^T F \phi$$

(50)

where $F$ is defined:

$$F = \chi(r, E) \int_0^\infty dE' \nu \Sigma_f(r, E')$$

(51)

Now define the fission source $Q(r, E) = F \phi(r, E)$, to obtain

$$Q = \frac{1}{k} FA^{-1} Q$$

(52)

Now define $G = FA^{-1}$ and rearrange Eq. (52) to find:

$$GQ = kQ$$

(53)

Note the multiplication factor is the largest eigenvalue of the operator $G$. Now define the fission source vector $Q$ with elements $Q_i =$ number of fission neutrons emitted in region $i$. The discrete form of $G$ is the fission matrix $G$ with matrix elements $G_{ij} :$
The fission matrix has been used to estimate the overall multiplication factor and to accelerate fission source convergence. The drawback with the fission matrix is its sensitivity to fluctuations, or noise, in the Monte Carlo solution [24]. This can be seen by inspection of Eq. (54), where grid refinement to obtain higher resolution will be accompanied by poorer statistics as the number of “counts” in the smaller regions will decrease. However, this is the same conundrum faced by the histogram tally and motivated our development of FET. Instead of using region-wise tallies for the fission neutron emission rates, let us use mode-wise tallies:

\[
G_{uv} = \frac{\text{fission neutrons emitted in mode } u}{\text{fission neutrons starting in mode } v}
\] (55)

Note that the matrix \( G \) is a FET-based response matrix for fission. The use of a FET-based fission matrix may improve the statistics because a count anywhere contributes to all the modes. One interesting aspect of this formulation of the fission matrix is that the matrix elements \( G_{uv} \) can be computed readily using the methodology developed for the response functions using FET [2].

No further progress has been made on this alternative fission matrix-based response matrix method. Effort by the project participants was focused on the alternative FET methods described in Section III.C, namely the kernel density estimator and the Arnoldi method.

V. References


VI. Other Grant Information

A. Graduate Students

The following graduate students were supported by the grant:


B. Budget and Expenditures

The approved budget and actual expenditures are included in the following table:

<table>
<thead>
<tr>
<th>Item</th>
<th>DOE Grant</th>
<th>UM Costsharing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved budget</td>
<td>$299,705</td>
<td>$84,000</td>
<td>$383,705</td>
</tr>
<tr>
<td>Actual expenditures</td>
<td>$299,412</td>
<td>$83,918</td>
<td>$383,330</td>
</tr>
</tbody>
</table>

C. Publications Arising from the Grant