Monte Carlo Application of Arnoldi's Method for Acceleration of Eigenvalue and Fission Source Convergence

Jeremy Lloyd Conlin[*,1,2] and James Paul Holloway[2]

[1] *Los Alamos National Laboratory*
PO Box 1663, MS F663, Los Alamos, NM 87545
[2] *University of Michigan, College of Engineering*
Department of Nuclear Engineering and Radiological Sciences
1221 Beal Avenue, Ann Arbor, MI 48109-2106

Abstract —*This paper introduces the explicitly restarted Arnoldi's method for calculating eigenvalues and eigenvectors in a Monte Carlo criticality calculation. Arnoldi's method is described along with the power method. The power method has been used for decades for Monte Carlo criticality calculations despite the availability of other algorithms with better convergence properties. The Monte Carlo application of the transport-fission operator of the Boltzmann transport equation is defined and the Monte Carlo implementation of both Arnoldi's method and the power method are described. A brief discussion of eigenvalue and fission source convergence is given. Numerical simulations of 1-D slab geometries are presented, demonstrating the convergence of both the eigenvalue and fission source (as measured by the Shannon entropy) for both Arnoldi's method and the power method. The results show that Arnoldi's method does not need to discard iterations like the power method because both the eigenvalue and fission source appear to converge immediately, even for problems with high dominance ratios.*

I Introduction

For decades[1–4] the power method has been the algorithm of choice for estimating the fundamental eigenvalue and eigenvector of the transport-fission operator of the Boltzmann transport equation in Monte Carlo criticality calculations. The power method is a robust application of a Krylov subspace method. A Krylov subspace method is one where only the knowledge of the application of a linear operator to a vector is required to estimate the eigenvalues and eigenvectors of the linear operator; an explicit (e.g. matrix) form of the linear operator is not necessary. The Monte Carlo implementation of the power method for neutral particle transport applications is straightforward making it desirable for criticality applications.

Despite the robust and uncomplicated features of the power method, it suffers from slow convergence to the desired eigenvalue and fission source. The convergence rate is proportional the dominance ratio, the

---

[*]`jlconlin@lanl.gov`

ratio of the first higher-order eigenvalue to the fundamental eigenvalue,

$$DR = \lambda_1/\lambda_0.$$

For many particle transport calculations of interest, this ratio is close to unity causing slow convergence.

Much work[5] has been done during the last decade investigating the convergence of the both the eigenvalue and fission source in Monte Carlo criticality applications. Ueki and Brown[6,7] have shown the necessity of performing many discarded power iterations before the results from the power method iterations can be used for estimating the fundamental eigenvalue without biasing the results with bad eigenvalue estimates or a poorly converged fission source.

Arnoldi's method[8] is another Krylov subspace method commonly used to estimate eigenvalues and eigenvectors of linear operators for which an explicit form of the operator is not known or is expensive to calculate and/or store. Warsa et al.[9] have shown Arnoldi's method to provide a significant improvement over the power method in deterministic calculations, achieving cost savings of 1–2 orders of magnitude using Arnoldi's method over the power method. For Monte Carlo criticality calculations, we have demonstrated[10] the capability of Arnoldi's method in calculating multiple eigenvalues and eigenvectors for 1-D geometries.

In this paper we will describe Arnoldi's method can be used to accelerate the convergence of the eigenvalue and the fission source as compared to the power method in a Monte Carlo criticality application. In Section II, we review the power method (Section II.B) and introduce Arnoldi's method and our one-dimensional implementation will be described (Section II.C). In Section III, we briefly review how fission source convergence is quantified and describe how it is done with Arnoldi's method. In Section IV, we will compare Arnoldi's method to the power method for a few problems with high dominance ratios. In Section V we conclude and propose future research for Arnoldi's method.

II   MONTE CARLO EIGENVALUE CALCULATIONS WITH KRYLOV SUBSPACE METHODS

Krylov subspace methods[11] approximate solutions (i.e. eigenvalues and eigenvectors) from a subspace of the form

$$\mathcal{K}_m \;=\; \text{span}\left\{v, \mathbf{A}v, \mathbf{A}^2 v, \ldots, \mathbf{A}^{m-1}v\right\}, \tag{1}$$

where v is some function or vector and $\mathbf{A}$ is a linear operator. The basis vectors of the Krylov subspace,

$$\mathcal{K}_m \;=\; \text{span}\left\{v_1, v_2, \ldots, v_m\right\}, \tag{2}$$

are the results of applying the linear operator to v multiple times, $v_i = \mathbf{A}^{i-1}$. Note that to calculate the n-th basis vector, one doesn't need to apply $\mathbf{A}$ $(n-1)$ times; it is sufficient to apply $\mathbf{A}$ to the previously calculated basis vector. Knowing this, we can see that Krylov subspace is calculated iteratively with one basis vector calculated at each iteration.

Krylov subspace methods do not require an explicit form of the operator $\mathbf{A}$, but do rely on the existence of a procedure or algorithm to apply the linear operator, $\mathbf{A}$, to a vector, v.

The iterative process of creating a Krylov subspace by the application of the linear operator on a vector makes Krylov subspace techniques attractive to Monte Carlo particle transport calculations. To define the linear operator for Monte Carlo criticality applications we begin with the Boltzmann transport eigenvalue equation

$$\mathbf{\Omega} \cdot \nabla \psi(\mathbf{r}, \mathbf{\Omega}) + \Sigma_t \psi(\mathbf{r}, \mathbf{\Omega}) =$$

$$\frac{\Sigma_s}{4\pi} \int \psi(\mathbf{r}, \mathbf{\Omega}) \, d\Omega + \frac{1}{k} \frac{\nu \Sigma_f}{4\pi} \int \psi(\mathbf{r}, \mathbf{\Omega}) \, d\Omega, \quad (3)$$

and write it in operator form

$$(\mathbf{L} + \mathbf{C} - \mathbf{S})\psi = \frac{1}{k}\mathbf{F}\psi \tag{4a}$$

$$\mathbf{T}\psi = \frac{1}{k}\mathbf{F}\psi \tag{4b}$$

where $\mathbf{L}, \mathbf{C}$, and $\mathbf{S}$ are the leakage, collision, and scattering operators respectively and $\mathbf{F}$ is the fission operator. The operator $\mathbf{T}$ is the transport-collision operator. The left-hand side represents the neutron loss and scattering mechanisms and the right-hand side represents the neutron gain mechanism. We define

$$v \equiv \mathbf{F}\psi \tag{5}$$

as the Monte Carlo fission source and

$$\mathcal{A} \equiv \mathbf{F}\,\mathbf{T}^{-1} \tag{6}$$

as the transport-fission operator. We can manipulate (4b) to obtain

$$\mathcal{A}v = kv. \tag{7}$$

This is a standard eigenvalue problem with k and v the eigenvalue and eigenvector of the transport-fission operator, $\mathcal{A}$. The operation of $\mathcal{A}$ on any fission source can be evaluated by Monte Carlo, as described next.

A fission source, v, is a distribution of fission neutrons, or—in the case of the first iteration—a predefined list of locations where fission neutrons may be found. The Monte Carlo application of $\mathcal{A}$ to v is done by randomly picking fission neutrons from the fission source, transporting them through the medium and storing any new neutrons created by fission. These new fission neutrons make up a new fission source and are the result of the operation of $\mathcal{A}$ on v. It is important to note that any of the standard variance reduction techniques can be used to reduce the statistical uncertainty of the application of $\mathcal{A}$.

This application of the transport-fission operator will be used in both the power method and Arnoldi's method. The only difference between the two applications is the way that fission neutrons are sampled from the fission source; in the power method, fission neutrons are sampled from discrete fission sites while in Arnoldi's method, fission neutrons are sampled from a discretized fission source. The specifics of each method are described in the following two sections.

The power method is a straightforward implementation of a Krylov subspace method. It begins with an estimate of the fundamental eigenvalue, k, and eigenvector v. At each iteration the linear operator, $\mathcal{A}$, is applied to a fission source, which is then normalized by the estimate of the eigenvalue,

$$v_{j+1} = \frac{1}{k_j}\mathcal{A}v_j. \tag{8}$$

Here v_j is the estimate of the fundamental eigenvector calculated in the j-th iteration of the power method. The fundamental eigenvalue is estimated in the $(j+1)$-th iteration as

$$k_{j+1} = k_j\frac{\int \mathcal{A}v_j}{\int v_j} = k_j\frac{\int v_{j+1}}{\int v_j}. \tag{9}$$

Using Eqs. (8)–(9) and knowing how to apply the transport-fission operator as discussed in the previous section, we can use the power method to estimate the fundamental eigenvalue and eigenvector of the transport-fission operator, $\mathcal{A}$. Note that the power method only stores the previous fission source and so discards information about the previously calculated basis vectors of the Krylov subspace.

The eigenvalue estimates calculated at the end of each power iteration are stored. The mean, $\overline{k}$, and

variance, σ^2, of these eigenvalue estimates can be estimated as

$$\overline{k} = \frac{1}{N} \sum_{j=1}^{N} k_j \tag{10}$$

$$\sigma^2 = \frac{1}{N} \frac{1}{N-1} \sum_{j=1}^{N} \left(k_j - \overline{k}\right)^2, \tag{11}$$

where N is the number of eigenvalue estimates. The standard deviation, σ, of the mean of the eigenvalue estimates is a measure of the statistical uncertainty in the mean eigenvalue estimate.

To examine the convergence rate of the power method we note that in Eq. (8), $v_j = \mathcal{A}^j v$. We can write v_j as a linear combination of the eigenvectors of $\mathcal{A}$

$$v_{j+1} = c_0 \mathcal{A}^{j+1} x_0 + \cdots + c_{n-1} \mathcal{A}^{j+1} x_{n-1} \tag{12}$$

$$\mathcal{A} v_j = c_0 \lambda_0^{j+1} x_0 + \cdots + c_{n-1} \lambda_{n-1}^{j+1} x_{n-1} \tag{13}$$

where $\lambda_0, \ldots, \lambda_{n-1}$ and $x_0, \ldots, x_{n-1}$ are the $n-1$ eigenvalues and eigenvectors of $\mathcal{A}$ with (λ_0, x_0) being the fundamental eigenpair; $c_0, \ldots, c_{n-1}$ are just some constants.

With a little bit of algebra we can rewrite Eq. (13) as

$$\mathcal{A} v_j = c_0 x_0 + \mathcal{O}(\lambda_1/\lambda_0)^j. \tag{14}$$

Eq. (14) is equivalent to Eq. (8) and shows that the power method converges no faster than the dominance ratio λ_1/λ_0.

II.C Arnoldi's Method

Arnoldi's method[8] iteratively builds a Krylov subspace by applying the linear operator to a vector in the same manner as in the power method. However, Arnoldi's method retains, orthogonalizes, and normalizes the basis vectors at each iteration. Arnoldi's method stores all the computed basis vectors, and uses the information so retained to improve the convergence compared to the power method. The orthonormalized basis vectors are called Arnoldi vectors.

Arnoldi's method begins like the power method with a vector which is then normalized

$$v_1 = \frac{v}{\|v\|_2}, \tag{15}$$

where

$$\|v\|_2 = \langle v, v \rangle^{1/2} \tag{16}$$

is just the Euclidean norm of the vector v. v_1 is the first basis vector of the Krylov subspace.

At the m-th iteration of Arnoldi's method the linear operator is applied to a vector, v_m, giving $\mathcal{A}v_m$. This result is then orthogonalized against all previously calculated Arnoldi vectors

$$\tilde{v}_{m+1} = \mathcal{A}v_m - \sum_{j=1}^{m} h_{jm} v_j, \tag{17}$$

where

$$h_{jm} = \langle \mathcal{A}v_m, v_j \rangle \tag{18}$$

is the inner product between $\mathcal{A}v_m$ and v_j,

$$\langle \mathcal{A}v_m, v_j \rangle = \int \left[\mathcal{A}v_m(x) \right] v_j(x) \, \mathrm{d}x. \tag{19}$$

The orthogonalization utilizing Eqs. (17) and (18) is a Gram-Schmidt orthogonalization. At the end of the iteration, $\tilde{v}_{m+1}$ is normalized,

$$v_{m+1} = \frac{\tilde{v}_{m+1}}{h_{m+1,m}}, \tag{20}$$

where

$$h_{m+1,m} = \|\tilde{v}_{m+1}\|_2. \tag{21}$$

The orthogonalization and normalization of the Arnoldi vectors involves taking the inner product between vectors. The results of these inner products (h_{jm} and $h_{m+1,m}$ in Eqs. (18) and (21) respectively) form an upper-Hessenberg matrix, $H_{m+1,m}$. (An upper-Hessenberg matrix is upper triangular but with the first subdiagonal non-zero.) The upper-Hessenberg matrix, $H_{m+1,m}$, is the projection of $\mathcal{A}$ onto the Krylov subspace.

We can write the m-th Arnoldi iteration in matrix form as

$$\mathcal{A}V_m = V_{m+1} H_{m+1,m}. \tag{22}$$

The elements of $H_{m+1,m}$ are the results of the inner product calculations and the columns of V are the orthonormalized Arnoldi vectors. If we separate the last column of V_{m+1} and the last row of $H_{m+1,m}$ we

can also write this as[12]

$$AV_m = V_m H_m + v_{m+1} h_{m+1,m} e_m^T \tag{23}$$

where v_{m+1} is the Arnoldi vector calculated during the m-th iteration, $h_{m+1,m}$ is the normalization factor for v_{m+1} and e_m is the m-th standard basis vector. Eq. (23) is called the Arnoldi factorization. During the m-th Arnoldi iteration, we add a row and a column to the Arnoldi factorization.

Using the Arnoldi factorization we can calculate estimates of the eigenvalues and eigenvectors of $\mathcal{A}$. The upper-Hessenberg matrix H_m is the projection of $\mathcal{A}$ onto the Krylov subspace. In practice, H_m is generated after a small number of iterations and thus is small and its eigenvalue and eigenvectors can easily be calculated with whatever method is desired, such as the standard QR method which we have used for all our calculations. The eigenpairs of H_m, (μ_i, x_i), can be used to find the Ritz pairs—approximate eigenpairs—of $\mathcal{A}$.

To calculate the Ritz pairs of $\mathcal{A}$, we multiply the Arnoldi factorization on the right by an eigenvector of H_m, x_i

$$AV_m x_i = V_m (H_m x_i) + v_{m+1} h_{m+1,m} e_m^T x_i \tag{24}$$

$$AV_m x_i = V_m (\mu_i x_i) + v_{m+1} h_{m+1,m} e_m^T x_i \tag{25}$$

$$Ay_i = \mu_i y_i + v_{m+1} h_{m+1,m} e_m^T x_i, \tag{26}$$

where $y_i = V_m x_i$. We can see that the Ritz pair (μ_i, y_i) is an estimate of an eigenpair of $\mathcal{A}$. Note that we can calculate multiple Ritz pairs of $\mathcal{A}$ from multiple eigenpairs of H_m.

The term, $v_{m+1} h_{m+1,m} e_m^T x_i$ in Eq. (26) is the residual of the calculation; the absolute value of the residual can give an estimate of how well the Arnoldi process has converged. As the residual becomes increasingly small, the Ritz pair more nearly satisfies the eigenvector equation, and will converge to an exact eigenpair.

If the initial guess v is an eigenvector of $\mathcal{A}$ then the magnitude of the residual will be zero after one iteration. If the initial guess is a linear combination of n eigenvectors of $\mathcal{A}$ then the magnitude of the residual will become zero after n iterations. The closer our initial guess is to our desired eigenvectors, the faster Arnoldi's method will converge to the desired answer.

A Monte Carlo implementation of Arnoldi's method follows from the mathematical description given in the previous section with a definition of how to take the inner product of two fission sources. Additionally, a normalization of the Monte Carlo application of the transport-fission operator must be performed before orthogonalization of the Arnoldi vectors. Consideration must also be given to sampling from a potentially negative fission source.

II.D.1 Normalization of Fission Sources

In Arnoldi's method, the application of the transport-fission operator, $\mathcal{A}$, is applied in the same manner as it was in the power method; particles are sampled from a fission source, the particles are transported, and a new fission source is created from the induced fission neutrons produced during the transport of the source neutrons. When the application of $\mathcal{A}$ is complete, the new fission source is normalized by multiplying by the integral of the absolute value of the previous source

$$v'_{m+1}(x) = \mathcal{A}v_m(x)\frac{1}{N}\int |v_m(x)|\,\mathrm{d}x. \tag{27}$$

Here $v_m(x)$ is the source we sample from, $\mathcal{A}v_m(x)$ is the new source created after sampling from $v_m(x)$ and transporting. This normalization is similar to the normalization performed in the power method by multiplying the previous eigenvalue estimate as shown in Eq. (9).

Using $v'_{m+1}(x)$ as the result of applying the transport-fission operator, Arnoldi's method continues with the orthogonalization of the fission source

$$\tilde{v}_{m+1} = v'_{m+1}(x) - \sum_{j=1}^{m} h_{jm}v_j, \tag{28}$$

and the inner product of one Arnoldi vector to the new Arnoldi vector

$$h_{jm} = \langle v'_{m+1}, v_j \rangle. \tag{29}$$

The normalization of the fission source is unchanged from Eq. (20) in the Monte Carlo implementation with the above definitions.

II.D.2   Inner Product of Two Fission Sources

The inner product of two vectors is just the sum of element-wise multiplication of the two vectors. No simple definition exists for the inner product of two fission sources.

In order to define the inner product of two fission sources, we have taken the simple approach and discretized our fission source into B equally spaced spatial bins and represent the fission source in each bin with a first-order accurate spatial approximation

$$v_{\mathcal{L}}(x) = \sum_{b=1}^{B} \mathcal{L}_b(x), \tag{30}$$

where

$$\mathcal{L}_b(x) = \begin{cases} \alpha_b + \beta_b x, & x_b \leq x < x_{b+1} \\ 0, & \text{otherwise.} \end{cases} \tag{31}$$

Representing the fission source with $v_{\mathcal{L}}(x)$, the Arnoldi vector corresponding to a fission source is

$$v_{\mathcal{L}} \left[\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_B, \beta_B \right]^T, \tag{32}$$

and the inner product between two vectors is simply the sum of element-wise multiplication of two vectors as before

$$h_{jm} = \langle v_{\mathcal{L}}^{(m)}, v_{\mathcal{L}}^{(j)} \rangle \tag{33}$$

$$= \sum_{b=1}^{B} \alpha_b^{(m)} \alpha_b^{(j)} + \beta_b^{(m)} \beta_b^{(j)} \tag{34}$$

To determine the coefficients α_b and β_b for each bin we must evaluate two integrals, something that is is well suited for Monte Carlo methods. We first define the midpoint of bin b

$$x_{b,\text{mid}} = \frac{x_{b+1} + x_b}{2}. \tag{35}$$

Taking the zeroth and first spatial moments over the bin

$$\int_{x_b}^{x_{b+1}} \mathcal{L}_b(x)\, \mathrm{d}x = \frac{1}{2} \left(x_{b+1} - x_b \right) \left[2\alpha_b + \beta_b \left(x_{b+1} + x_b \right) \right] \tag{36a}$$

$$\int_{x_b}^{x_{b+1}} \left(x - x_{b,\text{mid}} \right) \mathcal{L}_b(x)\, \mathrm{d}x = \frac{\beta_b}{12} \left(x_{b+1} - x_b \right)^3, \tag{36b}$$

9

gives two equations for α_b and β_b. The left-hand side of equations 36 can be evaluated via Monte Carlo

$$\int_{x_b}^{x_{b+1}} \mathcal{L}_b(x)\,\mathrm{d}x = \frac{1}{N}\sum_{i=1}^{N}\omega_i \tag{37a}$$

$$\int_{x_b}^{x_{b+1}} (x_i - x_{b,\mathrm{mid}})\,\mathcal{L}_b(x)\,\mathrm{d}x = \frac{1}{N}\sum_{i=1}^{N}(x - x_{b,\mathrm{mid}})\,\omega_i \tag{37b}$$

where N is the number of source particles and ω_i and x_i are the weight and position of the particle that induces fission in bin b. By equating equation 36a with equation 37a and equation 36b with equation 37b we can obtain expressions for α_b and β_b respectively

$$\alpha_b = \frac{1}{x_{b+1} - x_b}\frac{1}{N}\sum_{i=1}^{N}\omega_i - \frac{\beta_b}{2}(x_{b+1} + x_b) \tag{38a}$$

$$\beta_b = \frac{12}{(x_{b+1} - x_b)^3}\frac{1}{N}\sum_{i=1}^{N}(x_i - x_{b,\mathrm{mid}})\,\omega_i. \tag{38b}$$

For Monte Carlo particle transport this means every time a fission occurs in bin b the tallies ω_i and $(x_i - x_{b,\mathrm{mid}})\,\omega_i$ are recorded.

The discretization of the fission source is a simple way of providing a way to calculate the inner product between fission sources. The discretization of the fission source can introduce a bias into the result and so care must be taken with how the geometry is discretized. Using a first-order accurate approximation to the fission source allows us to compute the inner product between two fission sources. Using a first-order accurate approximation greatly improves the estimate of the eigenvectors of $\mathcal{A}$ as well as reducing the statistical uncertainty in the estimated eigenvalues as reported by Conlin and Holloway [13]. Some ideas have been proposed [14] that may eliminate the need to discretize the fission source in order to define the inner product, however this has not yet been fully investigated.

II.D.3   Negative Fission Sources

During the orthogonalization and normalization steps of Arnoldi's method, the fission source will inevitably have regions that are negative. This occurs because we are orthogonalizing multiple basis vectors, so only one of them can be single-signed. This requires that the Monte Carlo procedure sample a multi-signed source and work with negative weight particles.

To sample from a distribution, it must be everywhere positive and integrate to 1. If $p(x)$ is a probability distribution function, the quantity $p(x)\,\mathrm{d}x$ is the probability of choosing a point in $\mathrm{d}x$ about x. For a fission

source $v(x)$ which may have negative regions it is first normalized such that

$$\int |v(x)|\, \mathrm{d}x = q \tag{39a}$$

$$p(x) = \frac{|v(x)|}{q}. \tag{39b}$$

With this normalization, the quantity $p(x)\,\mathrm{d}x$ is the probability of choosing a point in $\mathrm{d}x$ about x. A neutron position x_s is sampled from $p(x)$ and is given an initial weight of

$$\omega = \frac{v(x_s)}{|v(x_s)|}, \tag{40}$$

or, alternatively

$$\omega = \begin{cases} 1, & v(x_s) > 0 \\ -1, & v(x_s) < 1. \end{cases} \tag{41}$$

Neutrons sampled where $v(x_s) < 0$ reduce the tally where they score; neutrons sampled where $v(x_s) > 0$ contribute positively to the tally where they score. A neutron is never sampled where $v(x_s) = 0$ because there the probability of choosing this point is exactly zero.

Once a particle has been sampled, Monte Carlo transport proceeds as usual, with the neutron scoring $\omega\,(\nu\Sigma_f/\Sigma_T)$ in the proper bin at each collision. If non-analog Monte Carlo is being done, particle weight is reduced at each collision and Russian Roulette is played if the absolute value of the weight becomes small. Giving a neutron a signed weight does not interfere with any of the standard variance reduction or tallying techniques.

II.D.4 Explicitly Restarted Arnoldi's Method

As the Arnoldi process proceeds, the estimate of the fundamental eigenvalue and eigenvector improves with each iteration. After a few iterations we have a better estimate of the fundamental eigenpair than what we started with. However with each additional iteration the new Arnoldi vector must be orthogonalized against one more vector increasing the computational expense. In addition, one more Arnoldi vector must be stored, increasing the memory requirements.

To reduce the computational and memory requirements Arnoldi's method can be restarted[15] after a set number of iterations. This is known as explicitly restarted Arnoldi's method. Arnoldi's method is restarted with a linear combination of the best estimate of the desired eigenvectors as the starting vector for the next set of iterations. Each set of iterations will be referred to as an Arnoldi *restart* and each restart consists of

the same number of Arnoldi *iterations*. At the end of every restart, estimates of the desired Ritz pairs of $\mathcal{A}$ are calculated and stored; the mean and variance can be calculated just like in the power method.

III Eigenvalue and Fission Source Convergence

Before a Monte Carlo criticality code can begin to calculate the mean and variance of the eigenvalue estimates, the eigenvalue estimates must be converged. If the estimates are not converged, the mean will include poor estimates of the eigenvalue. These poor estimates affect the may cause the mean eigenvalue estimate to be biased. For this reason, all Monte Carlo codes[16] will discard some iterations before the eigenvalue estimates are stored. However it is often difficult to determine convergence and Monte Carlo codes do not determine this internally; the number of iterations to discard must be input by the user.

In addition to the eigenvalue convergence, the fission source must also be converged. Ueki and Brown[7] have demonstrated that the apparent convergence of the eigenvalue estimates does not imply that the fission source has converged. If a fission source has not converged, then the modeled system is not being adequately sampled and the estimated eigenvalue and eigenvector of the transport-fission operator may be incorrect.

Ueki and Brown[7] have come up with a way to quantify the convergence of the fission source, thus assisting the Monte Carlo user in identifying when their simulation has converged. They have used the Shannon entropy

$$H\left(S(x)\right) \equiv -\sum_{b=1}^{B} S_b \log\left(S_b\right) \tag{42}$$

as a measure of the convergence of the fission source. Here H is the Shannon entropy and S_b is number of fission neutrons in spatial bin b of the fission source v,

$$S_b = \int_{x_b}^{x_{b+1}} v(x)\,\mathrm{d}x. \tag{43}$$

As the calculation proceeds, the Shannon entropy will converge to some value; the actual value of the Shannon entropy is not important (as we will see later). What is important is that it converges to some value.

Calculating the Shannon entropy requires discretizing the fission source and counting the number of neutrons in each spatial bin. This does not affect the transport of the source since the discretization is only used for calculating the Shannon entropy. In Arnoldi's method, we calculate the Shannon entropy from the Ritz vectors of $\mathcal{A}$, and not from the fission source or Arnoldi vector. Using a first-order accurate spatial

approximation as described in Section II.D.2 we can define

$$S_b = \int_{x_b}^{x_{b+1}} |\mathcal{L}_b(x)| \, \mathrm{d}x. \tag{44}$$

With this definition of the Shannon entropy, the convergence of the fission source can be quantified. The value to which the Shannon entropy converges is not critical; what is critical is that the Shannon entropy converges to some value. In a Monte Carlo application, a converged value (i.e. Shannon entropy, eigenvalue estimates) is one where the individual values are randomly scattered about the mean of the individual values.

IV NUMERICAL CALCULATIONS

To demonstrate the differences in convergence rates between Arnoldi's method and the power method, we have performed a few 1-D Monte Carlo simulations; one homogeneous slab and two heterogeneous. In all of these simulations, the number of histories tracked and the number of iterations is the same between the power method and Arnoldi's method. However, the number of eigenvalue estimates and Shannon entropy measurements are much less for Arnoldi's method than for the power method; Arnoldi's method estimates an eigenvalue at the end of a restart (several iterations) while the power method estimates an eigenvalue at every iteration. In these simulations, we don't differentiate between active or inactive iterations since we are only interested in the convergence. The eigenvalues plotted are the eigenvalue estimates and *not* the mean of the eigenvalue estimates.

The geometries and materials for these examples were picked from published results[7,17] so that the estimated eigenvalue from the power method and Arnoldi's method could be compared to known results.

IV.A Homogeneous Slab

The first simulation we show here is a homogeneous 1-D slab with vacuum boundary conditions with $\nu\Sigma_f = 1.0$, $\Sigma_a = 0.2$, $\Sigma_s = 0.8$, and $\Sigma_t = 1.0$. This is the same material used in published results[17], but we have set our slab to be 50 mfp thick. A reference fundamental eigenvalue was calculated using an S_N code with 32 quadrature set and 200 equally spaced regions. The fundamental eigenvalue was calculated as $\lambda_0 = 0.997520$. Arnoldi's method had 25 iterations per restart and both Arnoldi and the power methods track 5×10^5 particles per iteration.

The convergence of the eigenvalue and Shannon entropy are shown in Figure 1; the eigenvalue estimates are shown on the bottom and the Shannon entropy is shown on top. The dashed black line shows either the

reference eigenvalue (0.997520 or the Shannon entropy at the end of the simulation for the power method, or the mean Shannon entropy for Arnoldi's method. The results from the power method are shown in Figure 1a and in Figure 1b for Arnoldi's method.

[Figure 1 about here.]

We can see that the Shannon entropy is converged in both Arnoldi's method and the power method and the eigenvalue estimates have converged to the reference solution. Arnoldi's method has converged in both the eigenvalue and Shannon entropy immediately (1 restart, or 25 iterations) while the power method requires approximately 200 iterations for convergence. From this data we can say that the power method must discard the first 200 iterations so as not to pollute the result with bad eigenvalue estimates. Arnoldi's method need not discard any iterations. We can see that the standard deviation of the eigenvalue estimates is much smaller for Arnoldi's method than for the power method. The implications of this are not yet fully realized.

IV.B   Heterogeneous Slab

We now show a simulation with a heterogeneous, 1-D slab with vacuum boundary conditions. This problem was first proposed by Kornreich and Parsons [18] and also used by Ueki and Brown [7] to demonstrate the difficulty of eigenvalue and fission source convergence. The heterogeneous slab has five regions of widths 1.0, 1.0, 5.0, 1.0, and 1.0 cm; the full width of the slab is 9.0 cm. The materials are of three types, fissile, scatterer, and absorber. The cross sections for the material types are (all in units of cm^{-3})

Fission $\Sigma_t = 1.0$, $\Sigma_s = 0.8$, $\Sigma_\gamma = 0.1$,
$\Sigma_f = 0.1$, $\nu = 3.0$;

Scatter $\Sigma_t = 1.0$, $\Sigma_s = 0.8$, $\Sigma_a = 0.2$;

Absorber $\Sigma_t = 1.0$, $\Sigma_s = 0.1$, $\Sigma_a = 0.9$.

The materials are setup with the thick slab in the center made of absorber material surrounded by thin slabs of scatterer material; the outer slabs are the fissile material and are surrounded by a vacuum. A graphical representation of this geometry is shown in Figure 2.

We demonstrate the convergence of the fundamental eigenvalue and the fission source for Arnoldi's method and the power method for two variations of this geometry. The first is an asymmetric variation of the geometry described above; the right most slab has a width of 1.01 cm. The other uses the symmetric geometry exactly as described. The main difference between these two scenarios is a change in the dominance

ratio. The dominance ratio is 0.999566 for the symmetric problem and 0.992504 for the asymmetric problem. Because the dominance ratio is larger for the symmetric problem, we expect to see a slower convergence.

In both the symmetric and asymmetric geometries, the initial fission source is solely in the left-most bin. This makes fission source convergence particularly difficult because neutrons that enter the absorber region have a low probability of traversing the region to cause a fission in the fissile region on the right of the slab. This difficulty is the same for both Arnoldi's method and the power method so the comparison between the methods is fair. For these simulations, 1×10^6 particles are tracked in each iteration with Arnoldi's method using 10 iterations per restart.

[Figure 2 about here.]

The results for the asymmetric problem are shown first in Figure 3; the power method results are shown in Figure 3a and the results from Arnoldi's method are shown in Figure 3b. As in the homogeneous results the dashed line shows the reference eigenvalue[18] of 0.427425 or the Shannon entropy at the end of the simulation for the power method and the mean Shannon entropy for Arnoldi's method.

[Figure 3 about here.]

The convergence of the eigenvalue using the power method requires 800–1000 iterations while the Shannon entropy doesn't converge for approximately 1250 iterations. In comparison, the eigenvalue estimates and Shannon entropy are immediately converged for Arnoldi's method.

The results for the symmetric problem (shown in Figure 4) demonstrate a slightly different behavior for the power method. The power method convergence is shown in Figure 4a and Arnoldi's method convergence is shown in Figure 4b. The eigenvalue appears to have converged immediately for both the power method and Arnoldi's method, in addition the Shannon entropy has converged immediately for Arnoldi's method as in the asymmetric problem. However, the Shannon entropy convergence for the power method takes much longer to converge. We ran the power method for 2000 more iterations than was run for Arnoldi's method and it appears that the Shannon entropy may be approaching convergence after that many iterations, but it is difficult to say whether it has converged.

[Figure 4 about here.]

V   Conclusions

Previous papers[10,13,19] have shown that a Monte Carlo implementation Arnoldi's method can accurately calculate the fundamental eigenvalue as well as the first two higher-order eigenmodes of 1-D slab geometries.

In this paper we have shown that the eigenvalue and fission source (as measured by the Shannon entropy) converges immediately for Arnoldi's method. In comparison—as was seen by Ueki and Brown [7]—the power method may need several hundred iterations, or more, before either the eigenvalue or fission source has converged.

In all the examples in this paper we see that the noise in the Shannon entropy is much smaller for the power method than for Arnoldi's method. In the power method we see a very systematic trend towards a converged value of the Shannon entropy with little noise superimposed on that trend. Arnoldi's method, on the other hand, has no trend towards a converged Shannon entropy (it is converged right away), but the noise in the Shannon entropy is greater than that for the power method. The full implication of this is not fully understood, but may indicate that the correlation between Arnoldi restarts is less than the correlation between power method iterations. This stands to reason as there are many Arnoldi iterations within each restart causing the fission source in the last iteration to be less coupled to the initial fission source for the iteration.

Arnoldi's method can save considerable computational expense because it does not require inactive Monte Carlo iterations in order to converge the fission source and the fundamental eigenvalue. Because of it's superior convergence capabilities, Arnoldi's method may be useful in a Monte Carlo criticality code in converging the eigenvalue and the fission source with the code switching to the traditional power method after convergence is finished.

We have not demonstrated using Arnoldi's method for 3-D or energy dependant calculations, but the mathematical description of the method fully supports multi-dimensional calculations as long as the inner product between two fission sources has been defined. For this paper we have discretized the fission source in order to calculate the inner product between two fission sources. As discussed in Section II.D.2, the fission source may not have to be discretized. Further work on this remains to be done.

REFERENCES

1. W. Goad and R. Johnston. A monte carlo method for criticality problems. *Nuclear Science and Engineering*, 5(6):371–375, 1959.

2. E. L. Kaplan. Monte carlo methods for equilibrium solutions in neutron multiplication. Technical Report UCRL-5275-T, Livermore National Laboratory, 1958.

3. J. Lieberoth. A monte carlo technique to solve the static eigenvalue problem of the boltzmann transport equation. *Nukleonik*, 11(5):213–&, 1968.

4. M. R. Mendelson. Monte carlo criticality calculations for thermal reactors. *Nuclear Science and Engineering*, 32(3):319–&, 1968.

5. Forrest B. Brown. A review of monte carlo criticality calculations–convergence, bias, statistics. In *International Conference on Mathematics, Computational Methods & Reactor Physics (M&C 2009)*, Saratoga Springs, NY, May 3-7 2009. American Nuclear Society.

6. Taro Ueki and Forrest B. Brown. Informatics approach to stationarity diagnostics of the monte carlo fission source distribution. In *American Nuclear Society 2003 Winter Meeting*. American Nuclear Society, American Nuclear Society, November 16-20 2003.

7. Taro Ueki and Forrest B. Brown. Stationarity modeling and informatics-based diagnostics in monte carlo criticality calculation. *Nuclear Science and Engineering*, 149(1):38–50, January 2005.

8. W. E. Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics*, 9(1):17–29, 1951. ISSN 0033-569X.

9. J. S. Warsa, T. A. Wareing, J. E. Morel, J. M. McGhee, and R. B. Lehoucq. Krylov subspace iterations for deterministic k-eigenvalue calculations. *Nuclear Science and Engineering*, 147(1):26–42, May 2004. ISSN 0029-5639.

10. Jeremy Lloyd Conlin and James Paul Holloway. Arnoldi's method of minimized iterations for monte carlo criticality calculations. In *International Conference on Reactor Physics, Nuclear Power: A Sustainable Resource*, Interlaken, Switzerland, September 14-19 2008.

11. Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Manchester University Press, 1992.

12. David S. Watkins. *Fundamentals of Matrix Computations*. Wiley-Interscience, New York, second edition, 2002.

13. Jeremy Lloyd Conlin and James Paul Holloway. Second order approximations for spatial discretization in arnoldi's method for monte carlo criticality calculations. In *Trans. Am. Nucl. Soc.*, volume 101, pages 441–443, Washington, D.C., November 15–19 2009. American Nuclear Society.

14. Thomas E. Booth and James E. Gubernatis. Exact regional monte carlo weight cancellation for second eigenfunction calculations. *Nuclear Science and Engineering*, 165(3):283–291, July 2010.

15. Y. Saad. Variations on arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra and its Applications*, 34(DEC):269–295, 1980. ISSN 0024-3795.

16. X-5 Monte Carlo Team. Mcnp—a general n-particle transport code, version 5—volume i:overview and theory. In *Trans. Am. Nucl. Soc.*, volume 87, page 273. American Nuclear Society, November 2002.

17. J. A. Rathkopf and W. R. Martin. The finite element response matrix method for the solution of the neuron transport equation. *Progress in Nuclear Energy*, 18(1-2):237–250, 1986. ISSN 0149-1970.

18. D. E. Kornreich and D. K. Parsons. Semi-analytic results and insights into difficult source convergence problems. In *Trans. Am. Nucl. Soc.*, volume 87, pages 152–154, 2002.

19. Jeremy Lloyd Conlin and James Paul Holloway. Relaxation scheme for monte carlo explicitly restarted arnoldi's method for criticality calculations. In *International Conference on Mathematics, Computational Methods & Reactor Physics (M&C 2009)*, Saratoga Springs, NY, May 3-7 2009. American Nuclear Society.
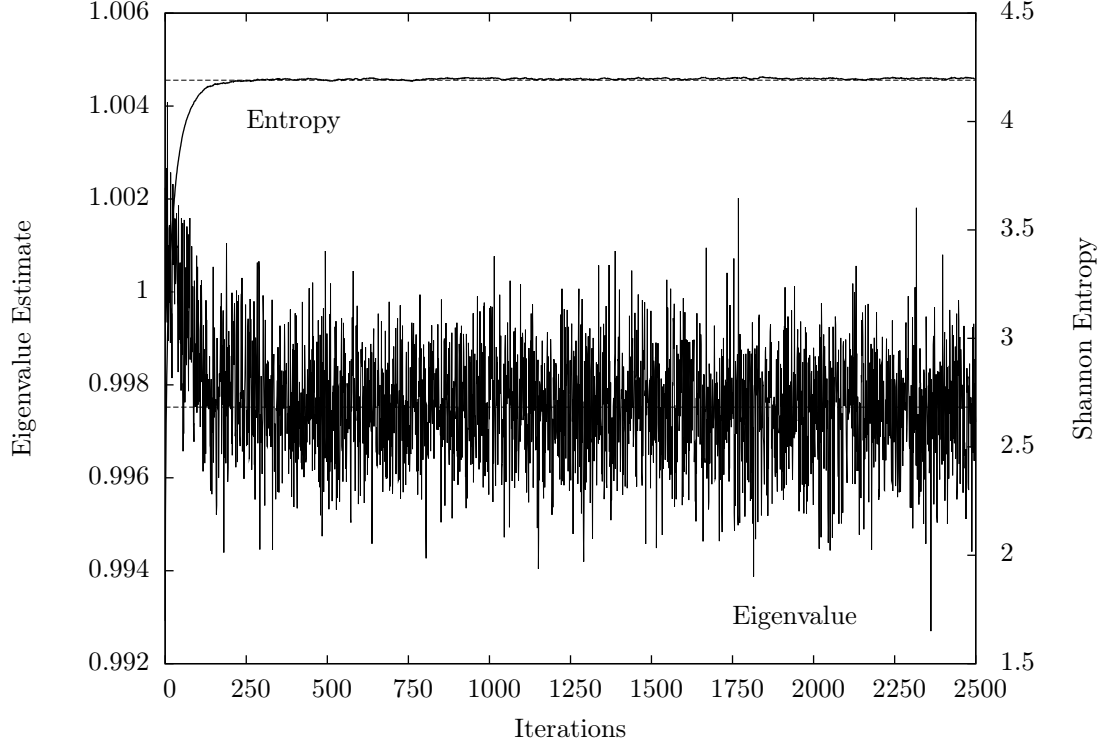
(a) Power method.
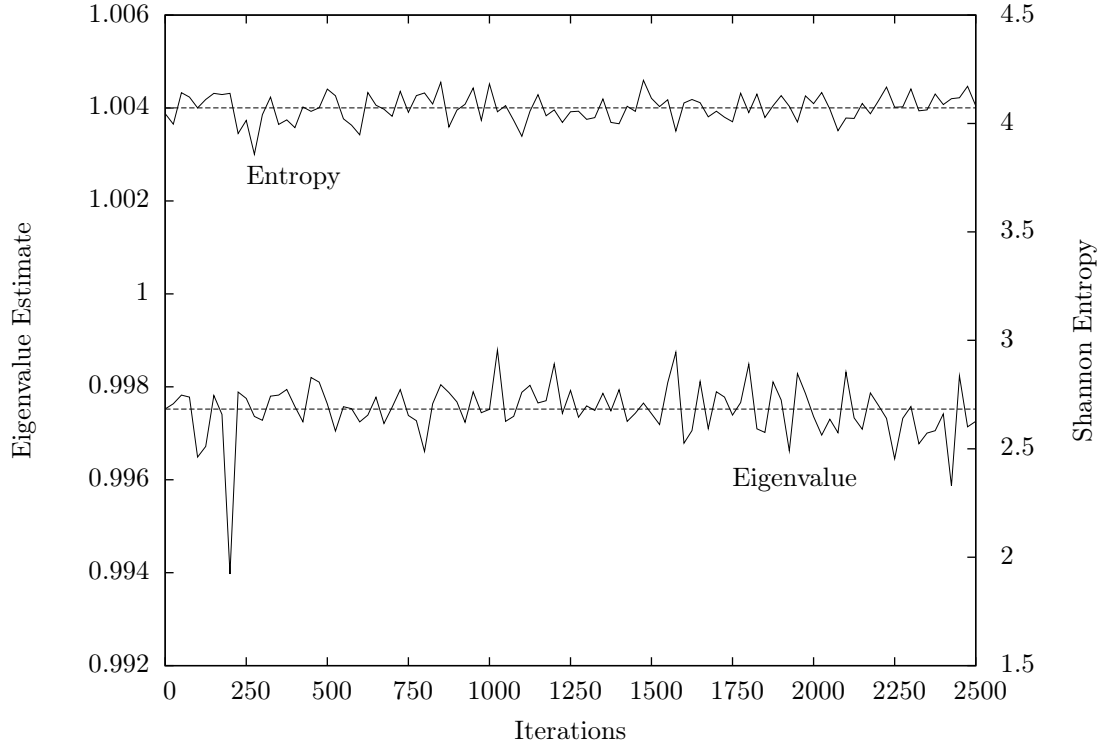
(b) Arnoldi's method.

Figure 1: Convergence of eigenvalue estimate and Shannon entropy for 50 mfp thick, homogeneous slab. The dashed line is the S_N eigenvalue solution for the fundamental eigenvalue ($\lambda_0 = 0.997520$) or the Shannon entropy.
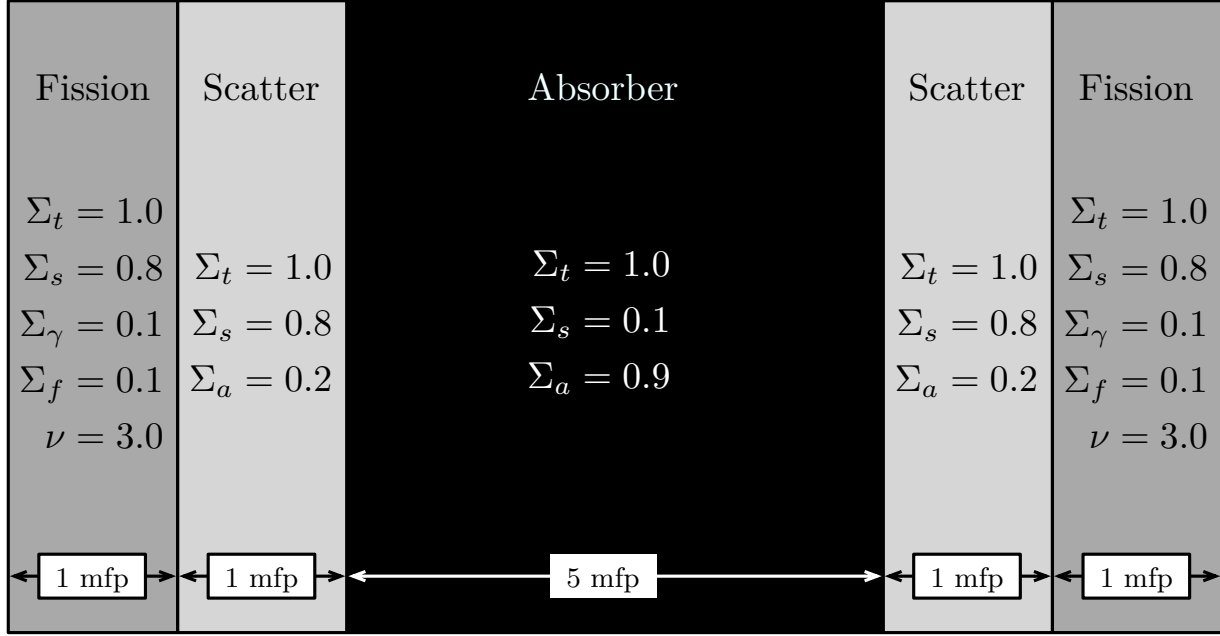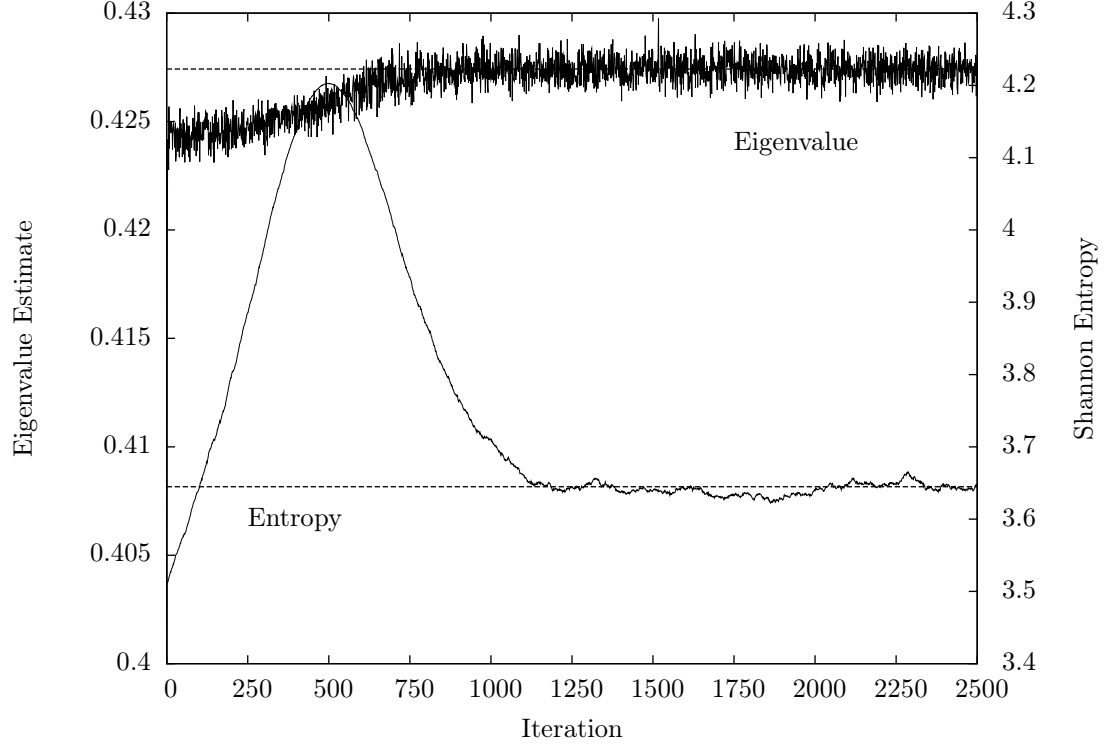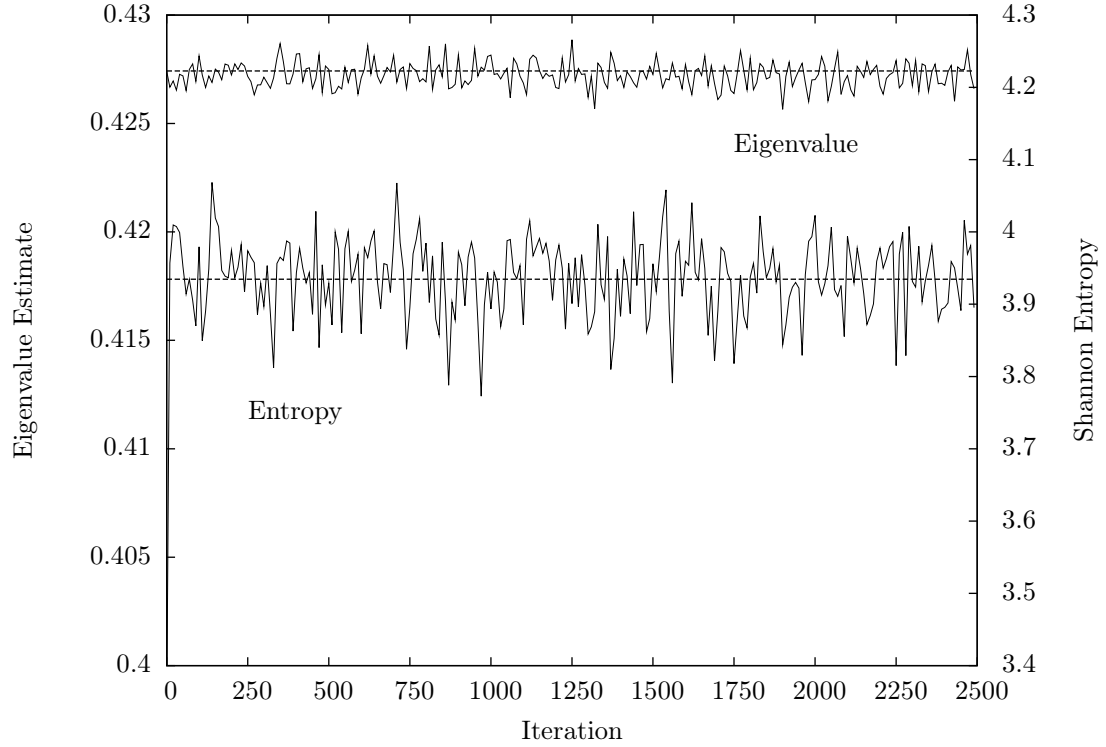
19

Figure 2: Diagram of heterogeneous slab geometry. The symmetric simulation is shown in this figure. For the asymmetric, heterogeneous simulation, the right-most region will have a width of 1.01 mfp.
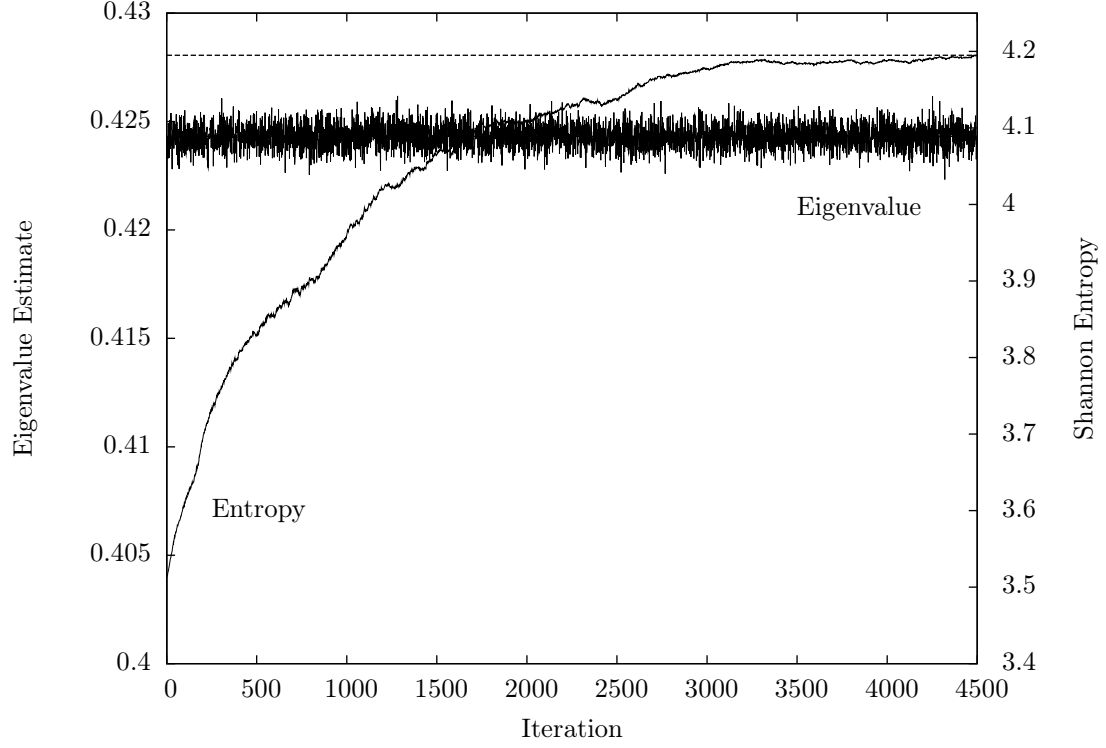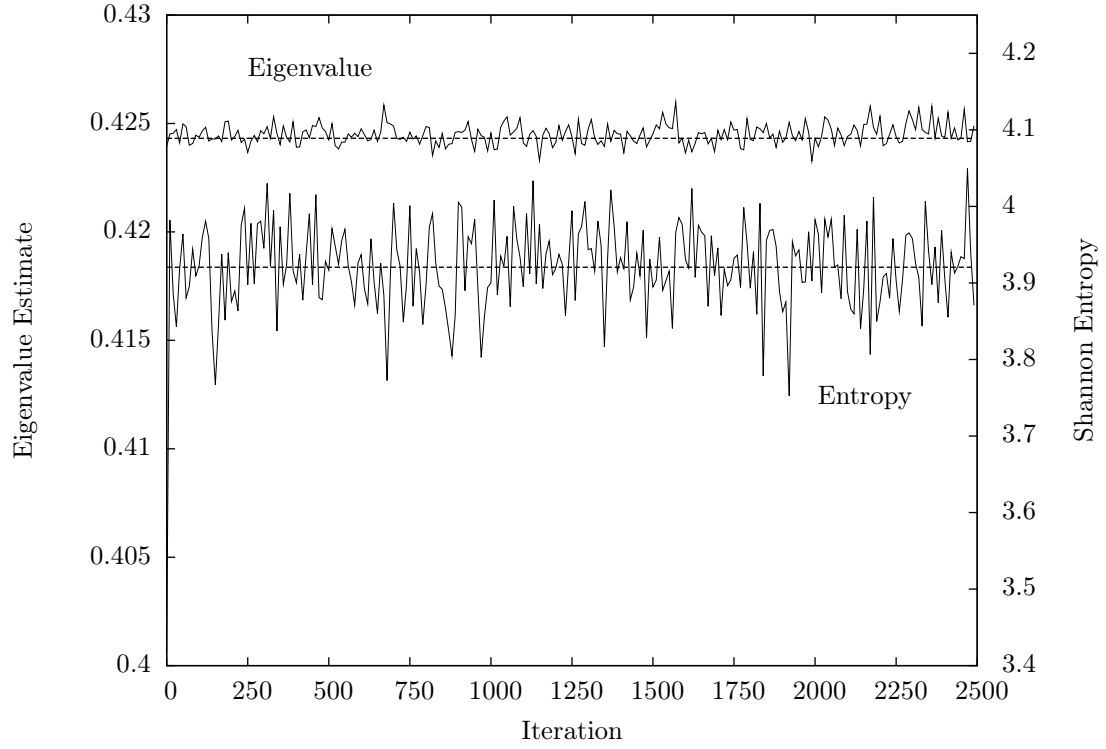
(a) Power method.

(b) Arnoldi's method.

Figure 3: Convergence of eigenvalue estimate and Shannon entropy for asymmetric geometry. The dashed line is the fundamental eigenvalue ($\lambda_0 = 0.427425\,18$), or the Shannon entropy at the end of the simulation for the power method and mean Shannon entropy for Arnoldi's method.

21

(a) Power Method

(b) Arnoldi's Method

Figure 4: Convergence of eigenvalue estimate and Shannon entropy for symmetric geometry. The dashed line is the fundamental eigenvalue ($\lambda_0 = 0.424316^{18}$), or the Shannon entropy at the end of the simulation for the power method and mean Shannon entropy for Arnoldi's method.