Explicitly Restarted Arnoldi's Method for Monte Carlo Nuclear Criticality Calculations

Jeremy Lloyd Conlin

August 17, 2009

Acknowledgements

- James Hollowy
- Bill Martin, Ed Larsen, Martin Strauss
- Scott and Geri Conlin
- Annie, Brigham, Lily, Emma
- Trisha Conlin

Particle Transport

Boltzmann Transport Equation:

$$\Omega \cdot \nabla \psi(\mathbf{r}, \Omega) + \Sigma_t \psi(\mathbf{r}, \Omega) = \frac{\Sigma_s}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega + \frac{1}{k} \frac{\nu \Sigma_f}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega,$$

Particle Transport

Boltzmann Transport Equation:

$$\Omega \cdot \nabla \psi(\mathbf{r}, \Omega) + \Sigma_t \psi(\mathbf{r}, \Omega) = \frac{\Sigma_s}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega + \frac{1}{k} \frac{\nu \Sigma_f}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega,$$

Operator Form:

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

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

Particle Transport

Boltzmann Transport Equation:

$$\Omega \cdot \nabla \psi(\mathbf{r}, \Omega) + \Sigma_t \psi(\mathbf{r}, \Omega) = \frac{\Sigma_s}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega + \frac{1}{k} \frac{\nu \Sigma_f}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega,$$

Operator Form:

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

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

Define:

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

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

Particle Transport

Boltzmann Transport Equation:

$$\Omega \cdot \nabla \psi(\mathbf{r}, \Omega) + \Sigma_t \psi(\mathbf{r}, \Omega) = \frac{\Sigma_s}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega + \frac{1}{k} \frac{\nu \Sigma_f}{4\pi} \int \psi(\mathbf{r}, \Omega)\, d\Omega,$$

Operator Form:

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

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

Define:

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

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

$$\mathcal{A}v = kv$$

Krylov Subspace Methods

- Estimate eigenpairs from Krylov subspace:

$$\mathcal{K}_m(\mathcal{A}, v) \equiv \text{span}\left\{v, \mathcal{A}v, \mathcal{A}^2 v, \ldots, \mathcal{A}^{m-1} v\right\}$$

- Subspace constructed iteratively

$$v_i = \mathcal{A} v_{i-1}, \qquad \text{for } i = 1, 2, \ldots, m-1$$

- Krylov subspace

$$\mathcal{K}_m(\mathcal{A}, v) \equiv \text{span}\left\{v_0, v_1, v_2, \ldots, v_{m-1}\right\}$$

- Explicit form of $\mathcal{A}$ not required

Monte Carlo Application of Transport-Fission Operator $\mathcal{A}$

To apply $\mathcal{A}$ to v_{i-1}:

1. Sample neutron from v_{i-1}
2. Transport neutron
3. Record positions of fission neutrons
4. Repeat...

Fission Source:

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

$$v_i = \mathcal{A}v_{i-1}$$

Transport-fission Operator:

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

$$\mathcal{A} = \mathbf{F}\left(\mathbf{L} + \mathbf{C} - \mathbf{S}\right)^{-1}$$

Krylov Subspace Method: Power Method

- Straightforward Krylov subspace method

$$v_i = \frac{1}{k_{i-1}} \mathcal{A} v_{i-1}$$

$$k_i = k_{i-1} \frac{\int \mathcal{A} v_{i-1}}{\int v_{i-1}}$$

- k_i and v_i converge to fundamental eigenpair as i becomes large
- Eigenvalue convergence proportional to dominance ratio

$$\mathrm{DR} = \lambda_1 / \lambda_0$$

Monte Carlo Statistics

- Iterations before power method converges are discarded
  - *Inactive* iterations
- Iterations after convergence are *Active* iterations
- Calculate mean and variance of eigenvalue estimates

$$\overline{\lambda} \equiv \frac{1}{N} \sum_{n=1}^{N} \lambda_n$$

$$\sigma^2_{\overline{\lambda}} \equiv \frac{1}{N-1} \left(\frac{1}{N} \sum_{n=1}^{N} \left(\lambda_n - \overline{\lambda} \right)^2 \right)$$

- $\sigma_{\overline{\lambda}}$ is a measure of statistical uncertainty of $\overline{\lambda}$

Arnoldi's Method—Alternative to Power Method

Power method has been used for more than 50 years

- Straightforward implementation
- Fundamental eigenvalue and eigenvector
- Only one Krylov basis vector is stored and used
- Slow convergence

Arnoldi's Method—Alternative to Power Method

Power method has been used for more than 50 years

- Straightforward implementation
- Fundamental eigenvalue and eigenvector
- Only one Krylov basis vector is stored and used
- Slow convergence

Arnoldi's method:

- Application of $\mathcal{A}$ same as in Power method
- Multiple eigenpairs
- All Krylov basis vectors are stored and used.
- Faster convergence

Arnoldi's Method

Krylov Subspace

$$\mathcal{K}_m(\mathcal{A}, v) \equiv \text{span}\left\{v, \mathcal{A}v, \mathcal{A}^2 v, \ldots, \mathcal{A}^{m-1} v\right\}$$

$$\mathcal{K}_m(\mathcal{A}, v) \equiv \text{span}\left\{v_0, v_1, v_2, \ldots, v_{m-1}\right\}$$

- Krylov subspace built iteratively
- Vectors are orthogonalized and normalized
- Normalized basis vectors (v_i) are called Arnoldi vectors
- All Arnoldi vectors are stored and used

Arnoldi Method Iteration

Arnoldi Iteration

$$v_1 = \frac{v}{\|v\|_2}$$

$$\tilde{v}_2 = \mathcal{A} v_1$$

$$\tilde{v}_2 = \mathcal{A} v_1 - h_{1,1} v_1$$

$$v_2 = \frac{\tilde{v}_2}{h_{2,1}}$$

Arnoldi Method Iteration

Arnoldi Iteration

$$v_1 = \frac{v}{\|v\|_2}$$

$$\tilde{v}_2 = \mathcal{A} v_1$$

$$\tilde{v}_2 = \mathcal{A} v_1 - h_{1,1} v_1$$

$$v_2 = \frac{\tilde{v}_2}{h_{2,1}}$$

At m-th iteration

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

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

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

Arnoldi Factorization

$$\mathcal{A}V_m = V_m H_m + v_{m+1} h_{m+1,m} e_m^T$$

- Columns of V_m are Arnoldi vectors
- Elements of H_m are h_{jm}
- H_m is the projection of $\mathcal{A}$ onto Krylov subspace

- $H_m \in \mathbb{R}^{m \times m}$, upper-Hessenberg matrix
- m is small, easy to calculate eigenpair of H_m, (μ_i, x_i)

Finding Ritz Pairs from Arnoldi Factorization

- (μ_i, x_i) is an eigenpair of H_m

$$\begin{aligned}
\mathcal{A}V_m &= V_m H_m &&+ v_{m+1} h_{m+1,m} e_m^T \\
\mathcal{A}V_m x_i &= V_m \left(H_m x_i\right) &&+ v_{m+1} h_{m+1,m} e_m^T x_i \\
\mathcal{A}V_m x_i &= V_m \left(\mu_i x_i\right) &&+ v_{m+1} h_{m+1,m} e_m^T x_i \\
\mathcal{A}y_i &= \mu_i y_i &&+ v_{m+1} h_{m+1,m} e_m^T x_i
\end{aligned}$$

- $y_i = V_m x_i$
- (μ_i, y_i) is a *Ritz pair* or approximated eigenpair of $\mathcal{A}$
- Residual:

$$|r_i| = \left\|\mathcal{A}y_i - \mu y_i\right\| = \left|h_{m+1,m}\right| |e_m x_i| .$$

Explicitly Restarted Arnoldi's Method

- Begin with estimate of desired eigenvector
- Calculate eigenpairs of H_m after a fixed number of iterations
- Restart Arnoldi with desired eigenvectors as new starting vector
- Several iterations make up one Arnoldi *restart*

Monte Carlo Arnoldi's Method

- $\mathcal{A}$ is applied exactly as it is applied in power method
- Eigenvalues and eigenvectors estimated at end of every restart
- Mean and variance of estimates can be calculated
- Can do inactive restarts (but don't need to!)

- Orthogonalization will create negative Arnoldi vectors
  - Negative fission source
- Need to define inner product between two fission sources

Negative Sources

- To sample from negative fission source, normalize

$$\int \left|\nu(x)\right| \mathrm{d}x = q$$

$$p(x) = \frac{\left|\nu(x)\right|}{q}$$

- $p(x)\,\mathrm{d}x$ is probability of picking a point in $\mathrm{d}x$ about x
- Neutron is given weight

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

Spatial Discretization

- Orthonormalization requires inner product between fission sources:

$$h_{jk} = \langle v_j, v_k \rangle = \int v_j(x) v_k(x) \, \mathrm{d}x$$

- Discretize fission source

$$v_\Pi(x) = \sum_{b=1}^{B} a_b \Pi_b(x)$$

$$\Pi_b(x) = \begin{cases} \left(\frac{1}{\Delta x_b}\right)^{1/2}, & x_b \leq x < x_{b+1} \\ 0, & \text{otherwise} \end{cases}$$

Inner Product of Fission Sources

- Discretization:

$$v_\Pi(x) = \sum_{b=1}^{B} a_b \Pi_b(x)$$

- Arnoldi vector representation:

$$v_\Pi = \left[a_1, a_2, \ldots, a_B\right]^T$$

- Inner product of discretized fission sources:

$$h_{jk} = \langle v_\Pi^{(j)}, v_\Pi^{(k)} \rangle = \sum_{b=1}^{B} a_b^{(j)} a_b^{(k)}$$

Sampling from Discretized Source

- $\nu_\Pi(x)$ is a first order accurate approximation of fission source
- To sample from $\nu_\Pi(x)$, normalize

$$q = \sum_{b=1}^{B} \left| a_b \right|$$

$$p_b = \left| a_b \right| / q$$

- p_b is probability of sampling from bin b
- Bin b is sampled *uniformly* to determine neutron position

Numerical Results

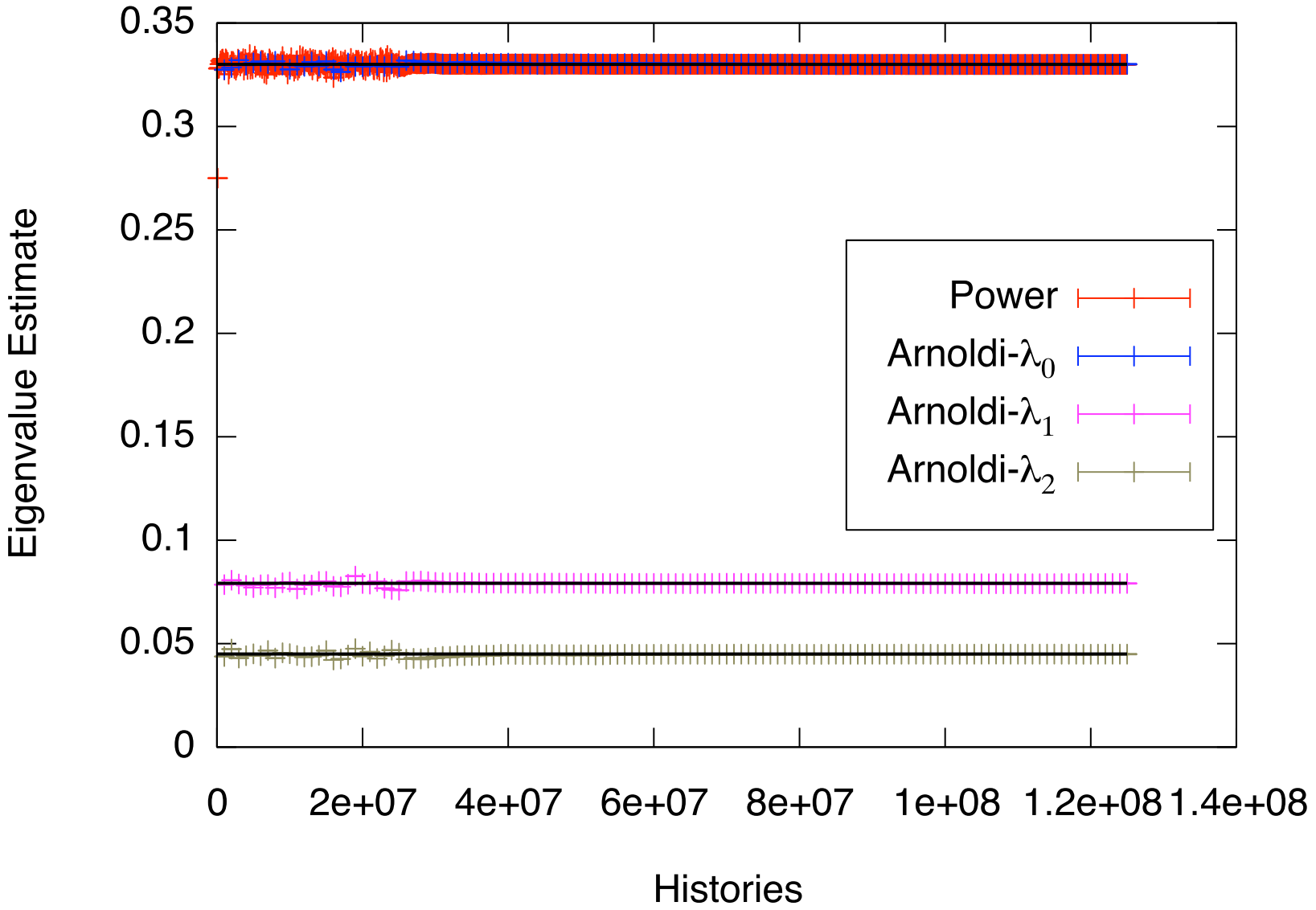
- Homogeneous slab
- $\nu\Sigma_f = 1.0$, $\Sigma_a = 0.2$, $\Sigma_s = 0.8$, $\Sigma_t = 1.0$
- 0.2, 2.0, 20 mfp thick
- Geometry and cross sections match published results

- Parameters:
  - 1×10^5 particles per iteration
  - 25 inactive, 100 active Arnoldi restarts, 10 iterations per restart
  - 250 inactive, 1000 active power iterations
  - 50 spatial bins (0.2 mfp), 75 spatial bins (2.0, 20 mfp)

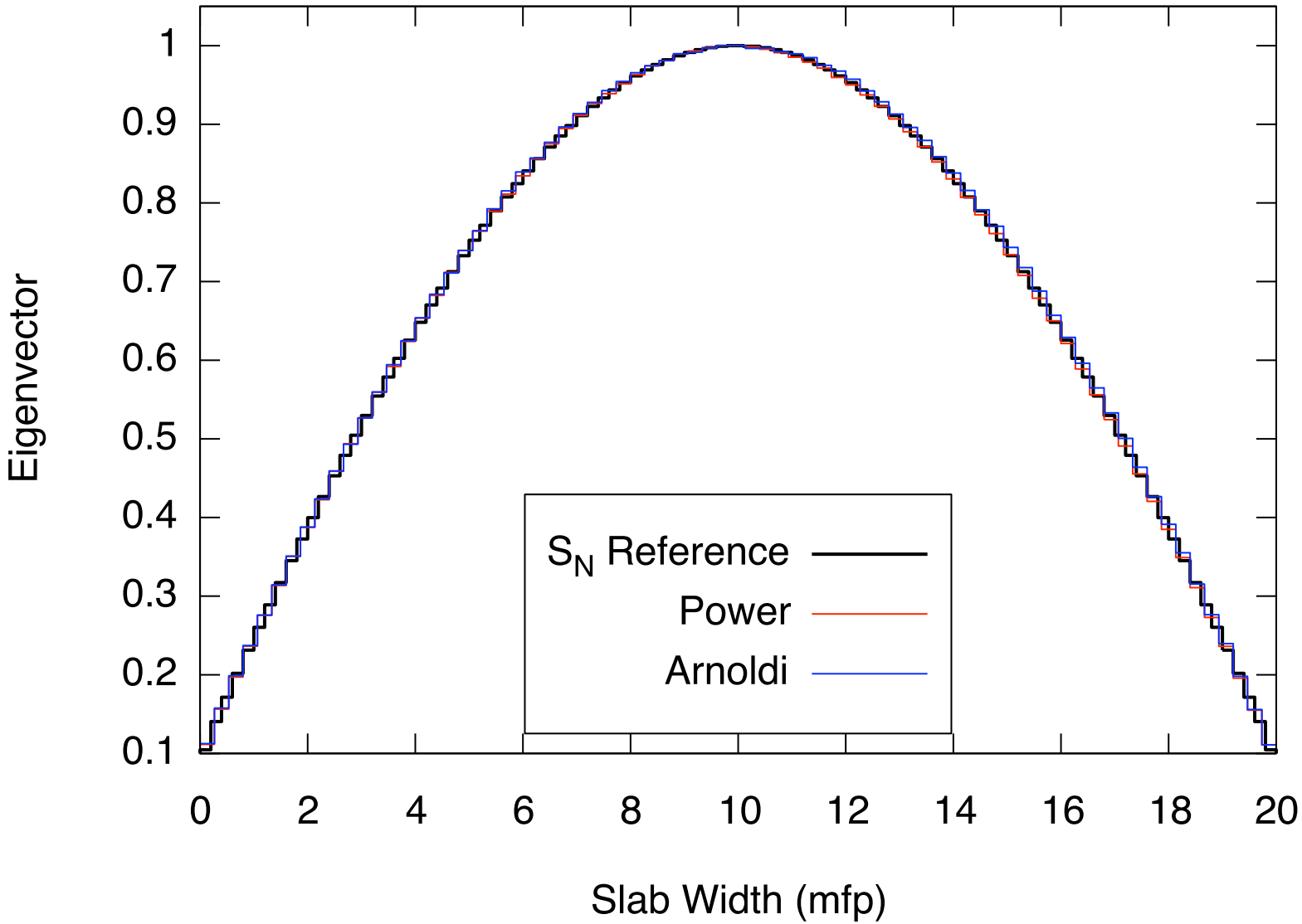
Numerical Results

Width (DR)	Method	Eigenvalue	Standard Deviation	Reference
0.2 (0.2400)	Power Arnoldi	0.329979 0.33008	6.3×10^{-5} 1.8×10^{-4}	0.330000
2.0 (0.4015)	Power Arnoldi	2.09593 2.09652	2.7×10^{-4} 6.9×10^{-4}	2.09599
20 (0.9079)	Power Arnoldi	4.82734 4.8290	6.3×10^{-4} 1.5×10^{-3}	4.82780

Eigenvalue Convergence



Fundamental Eigenvector



Multiple Eigenvectors

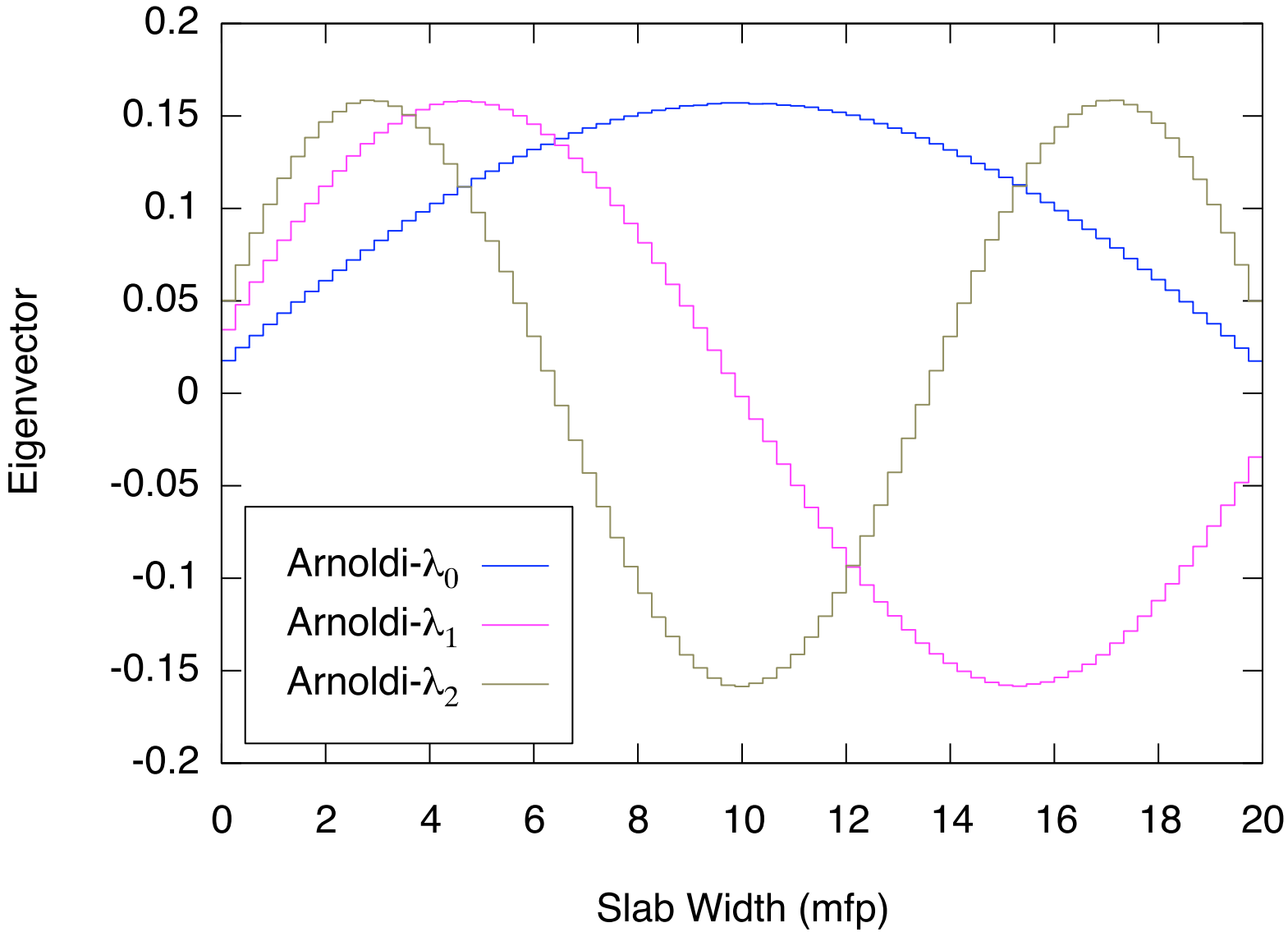


Figure of Merit

Width (mfp)	Method	Standard Deviation	FOM	Time (sec)
0.2	Power	6.3×10^{-5}	1.7×10^{6}	149.0
	Arnoldi	1.8×10^{-4}	3.3×10^{5}	95.3
2.0	Power	2.7×10^{-4}	5.5×10^{4}	258.1
	Arnoldi	6.9×10^{-4}	9.8×10^{3}	212.5
20	Power	6.3×10^{-4}	5.4×10^{3}	463.0
	Arnoldi	1.5×10^{-3}	1.1×10^{3}	378.5

Figure of Merit

Figure of Merit

$$\text{FOM} \equiv \frac{1}{\sigma^2 T}$$

Figure of Merit

Figure of Merit

$$\text{FOM} \equiv \frac{1}{\sigma^2 T}$$

Variance

$$\sigma^2 \equiv \frac{1}{N-1}\left(\frac{1}{N}\sum_{n=1}^{N}(x_n - \overline{x})^2\right)$$

Spread

$$s \equiv \frac{1}{N}\sum_{n=1}^{N}(x_n - \overline{x})^2$$

Figure of Merit

Figure of Merit

$$\text{FOM} \equiv \frac{1}{\sigma^2 T}$$

Variance

$$\sigma^2 \equiv \frac{1}{N-1}\left(\frac{1}{N}\sum_{n=1}^{N}(x_n - \overline{x})^2\right)$$

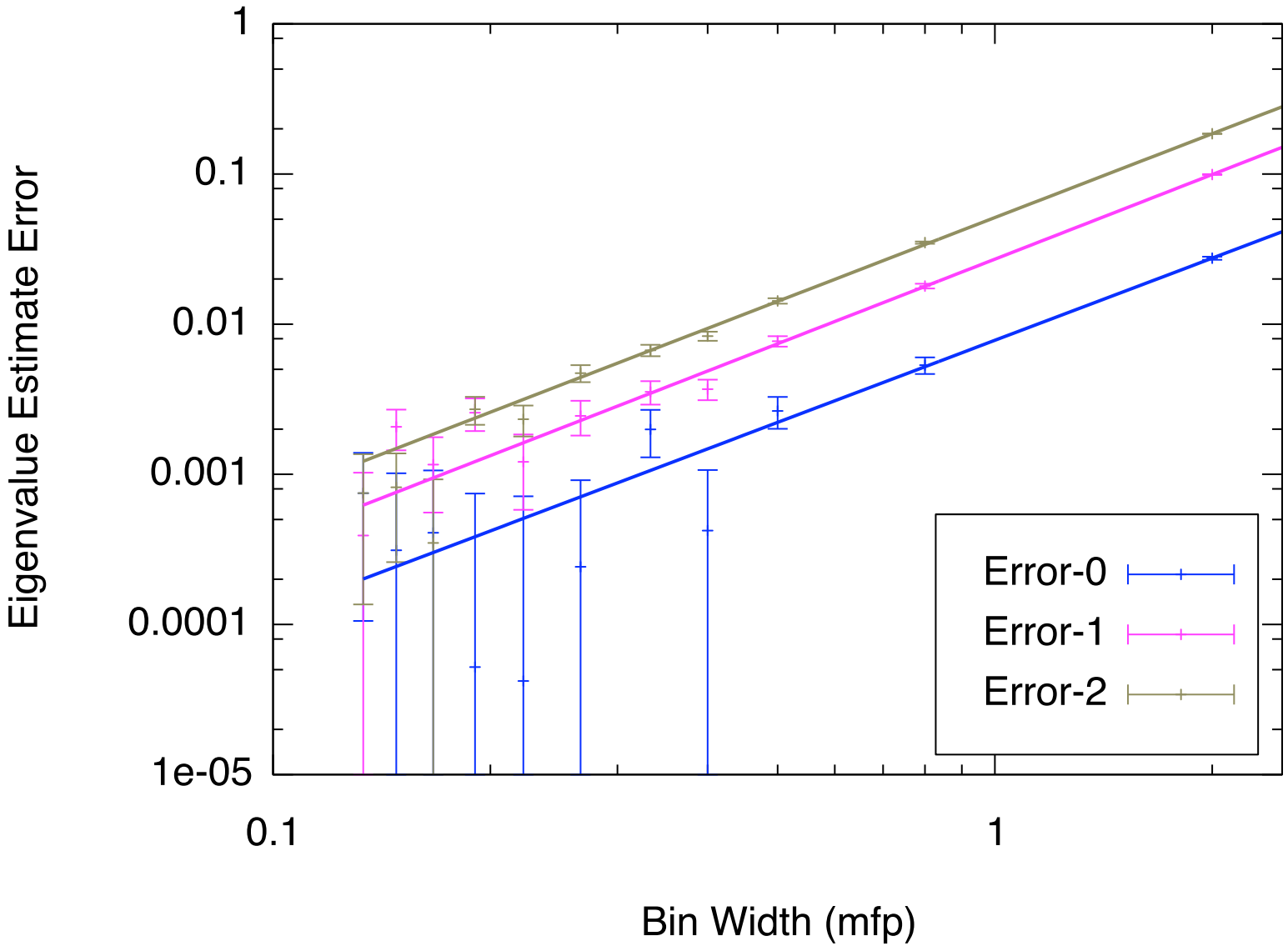
Spread

$$s \equiv \frac{1}{N}\sum_{n=1}^{N}(x_n - \overline{x})^2$$

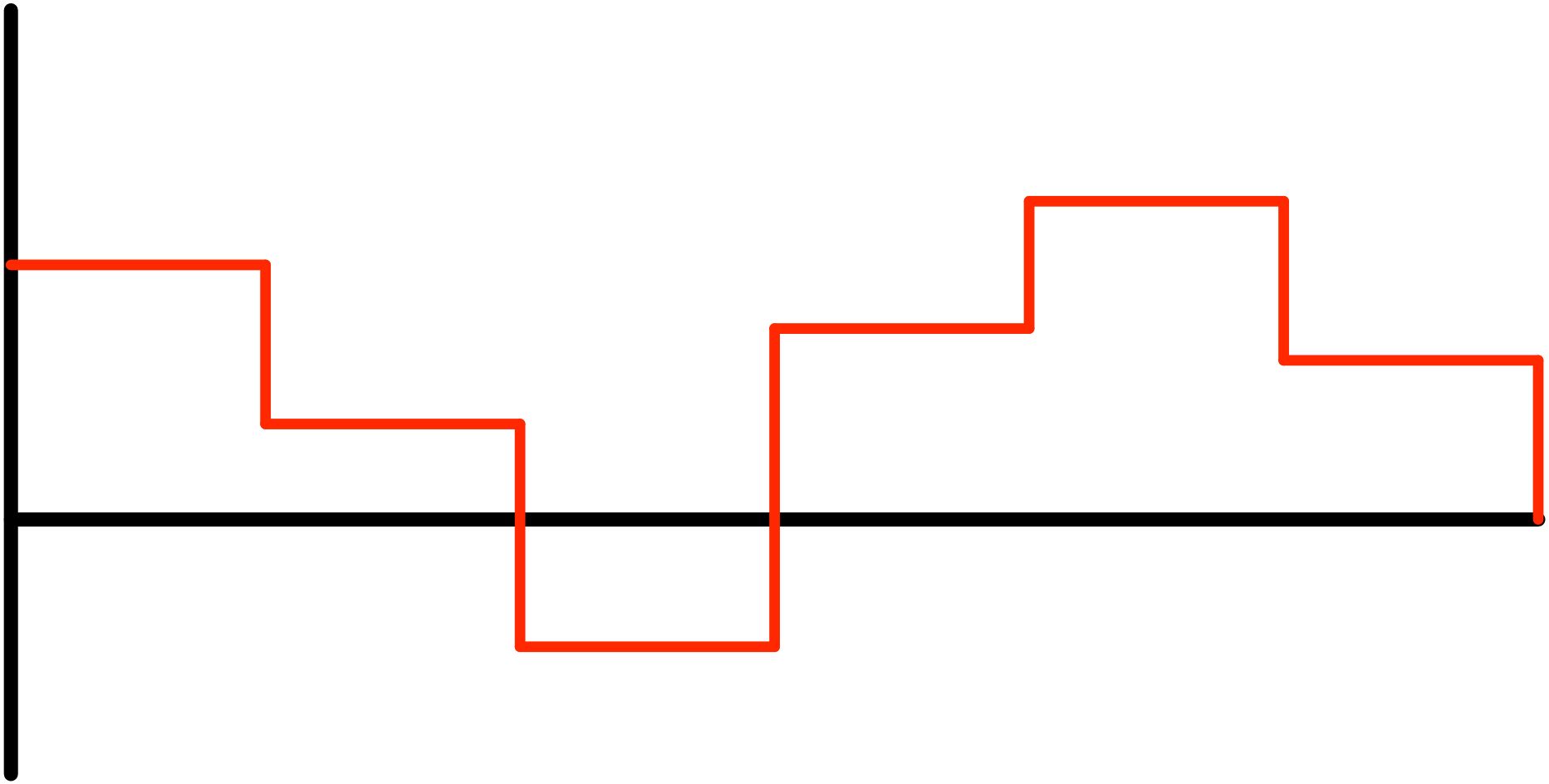
	0.2 mfp	2.0 mfp	20 mfp
Power	0.0020	0.0084	0.0201
Arnoldi	0.0018	0.0069	0.0153

Spread of eigenvalue estimates

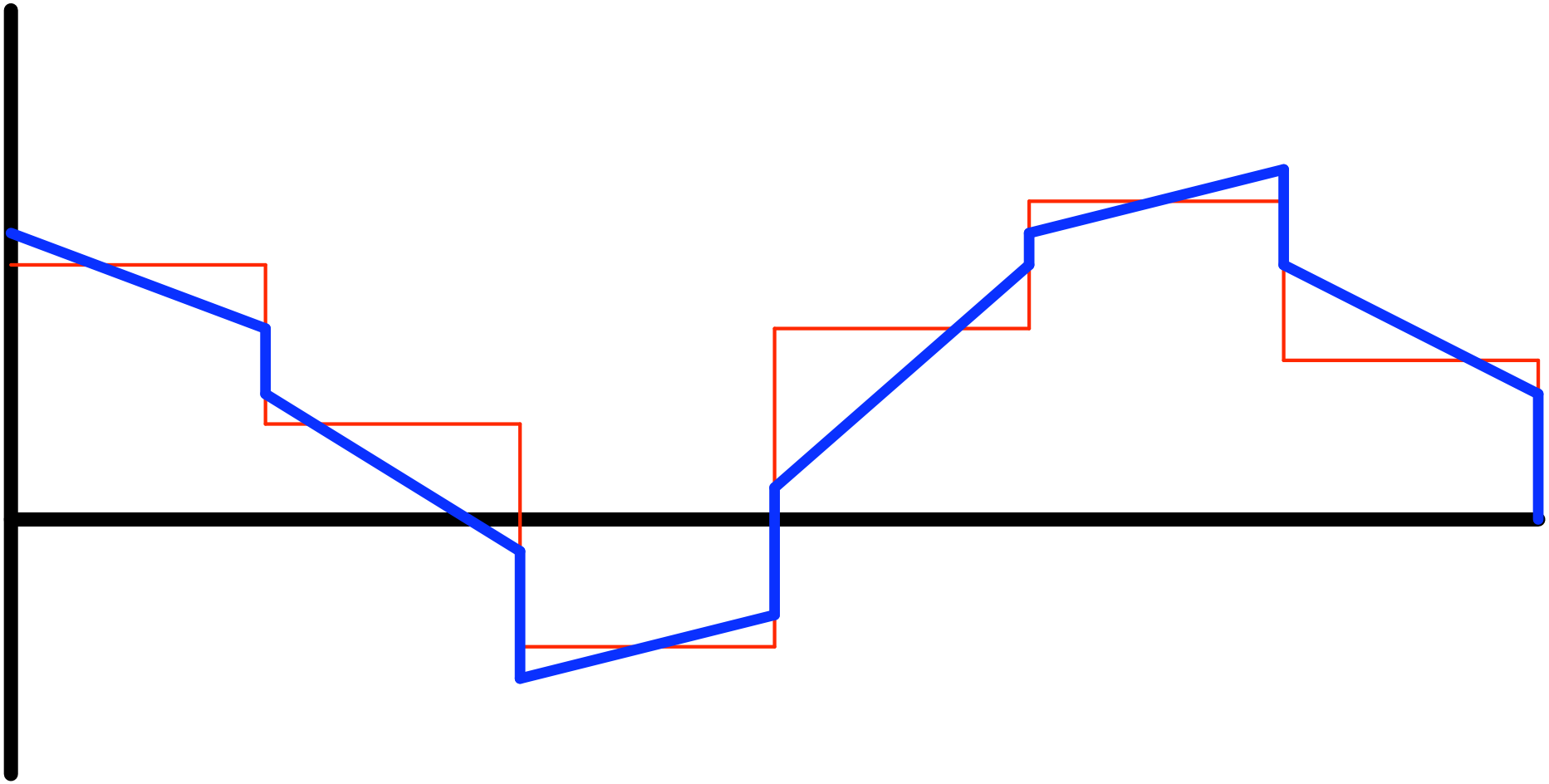
Discretization Error



Spatial Approximations



Spatial Approximations



Spatial Approximations

First Order Accurate

$$v_{\Pi}(x) = \sum_{b=1}^{B} a_b \Pi_b(x)$$

$$\Pi_b(x) = \left(\frac{1}{\Delta x_b} \right)^{1/2}$$

$$v_{\Pi} = \left[a_1, a_2, \ldots, a_B \right]^T$$

Second Order Accurate

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

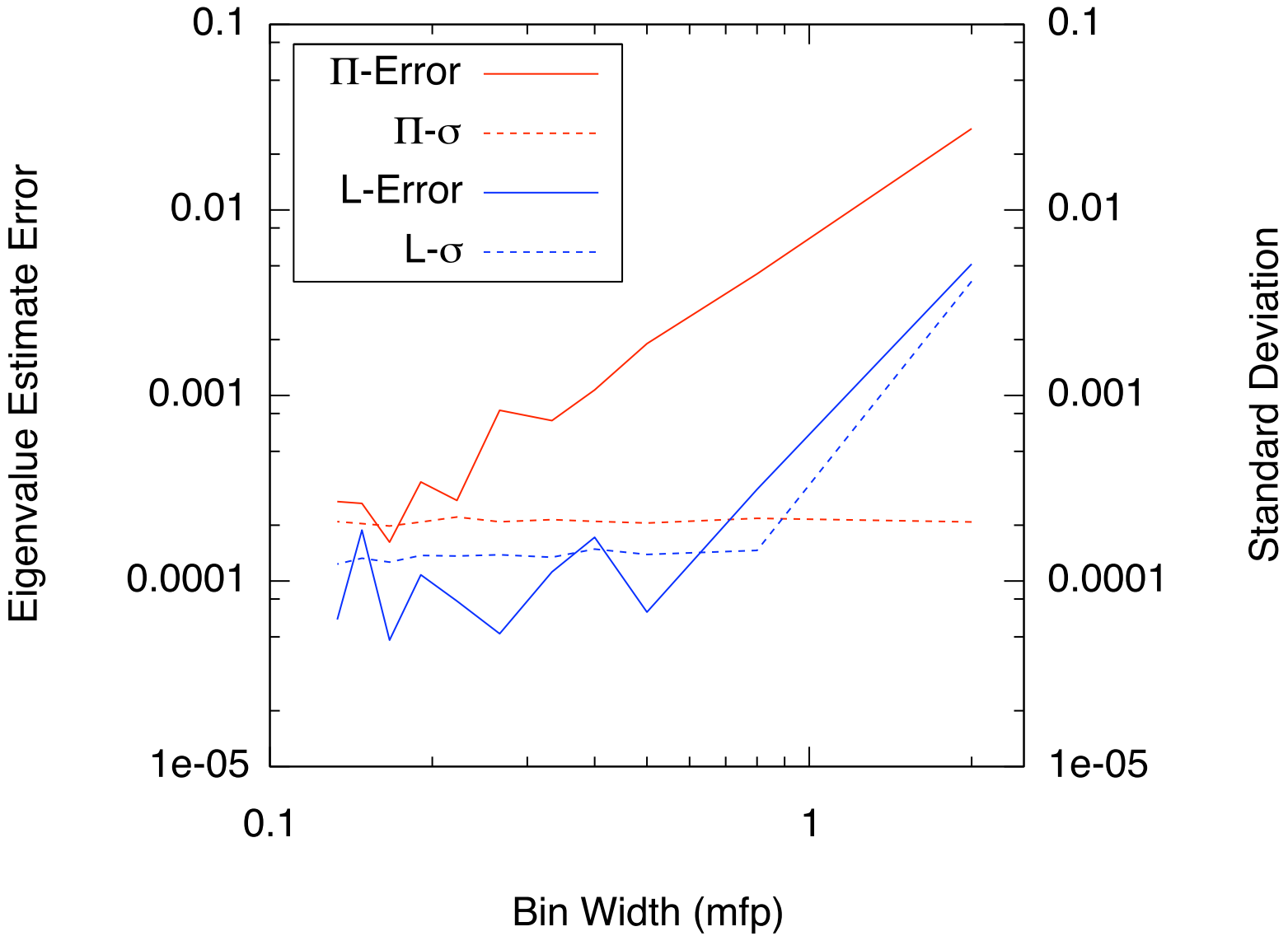
$$\mathcal{L}_b(x, \alpha_b, \beta_b) = \alpha_b + \beta_b x$$

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

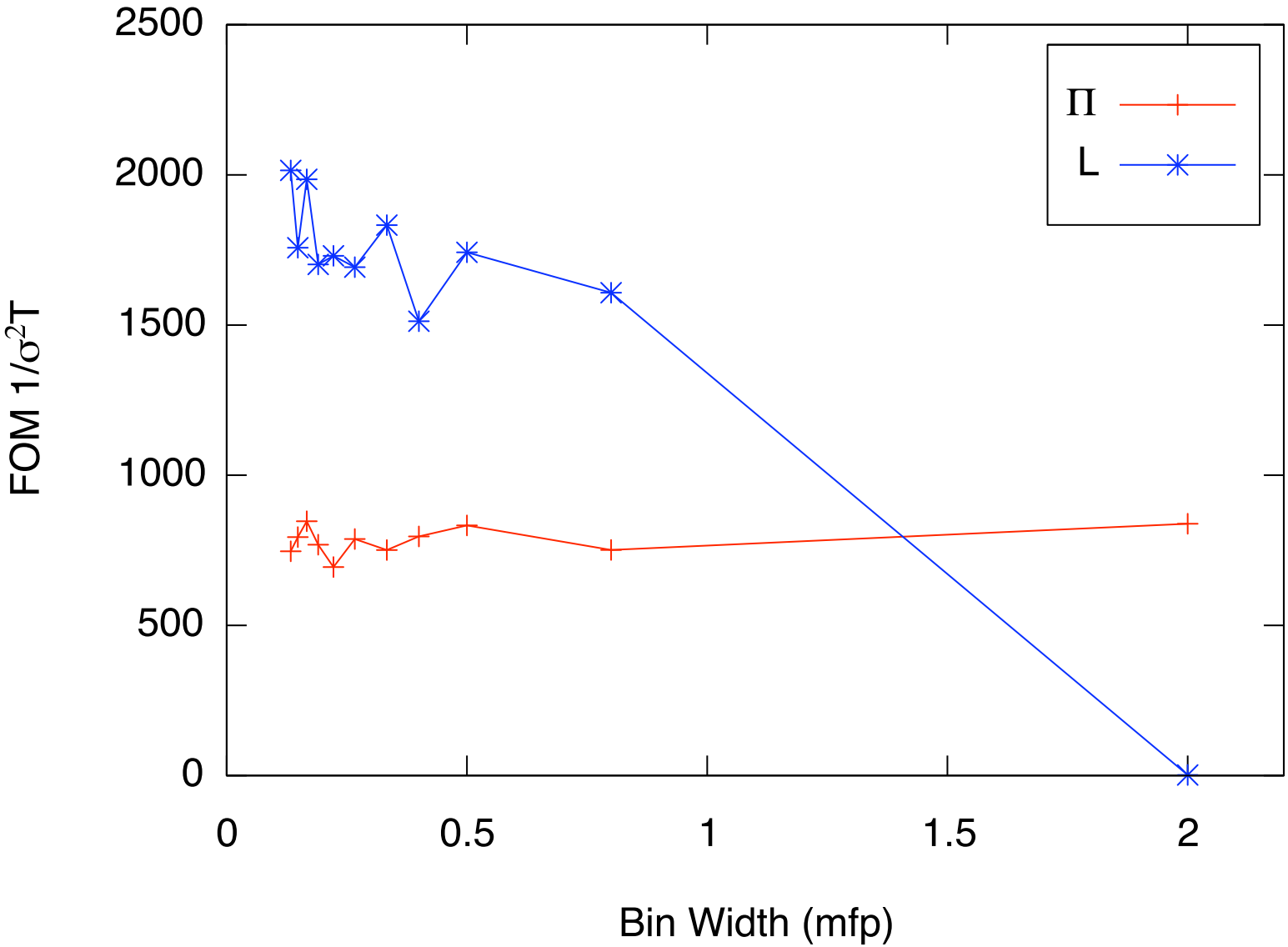
Numerical Results

- 20 mfp homogeneous thick
- $\nu\Sigma_f = 1.0$, $\Sigma_a = 0.2$, $\Sigma_s = 0.8$, $\Sigma_t = 1.0$
- 1×10^6 particles per iteration
- 50 inactive, 100 active Arnoldi restarts, 10 iterations per restart
- First and second order accurate spatial approximations
- 10–150 spatial bins

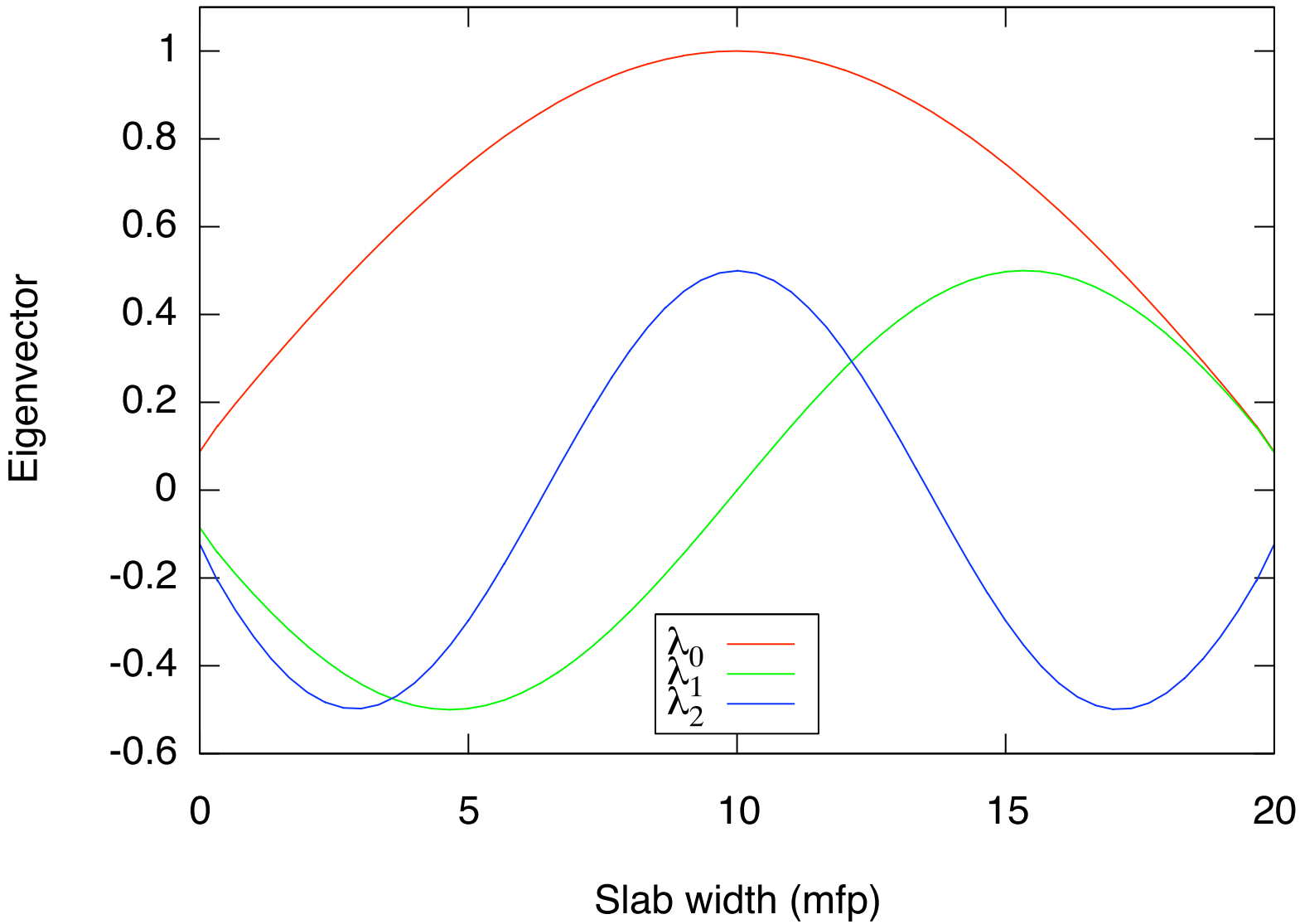
Numerical Results



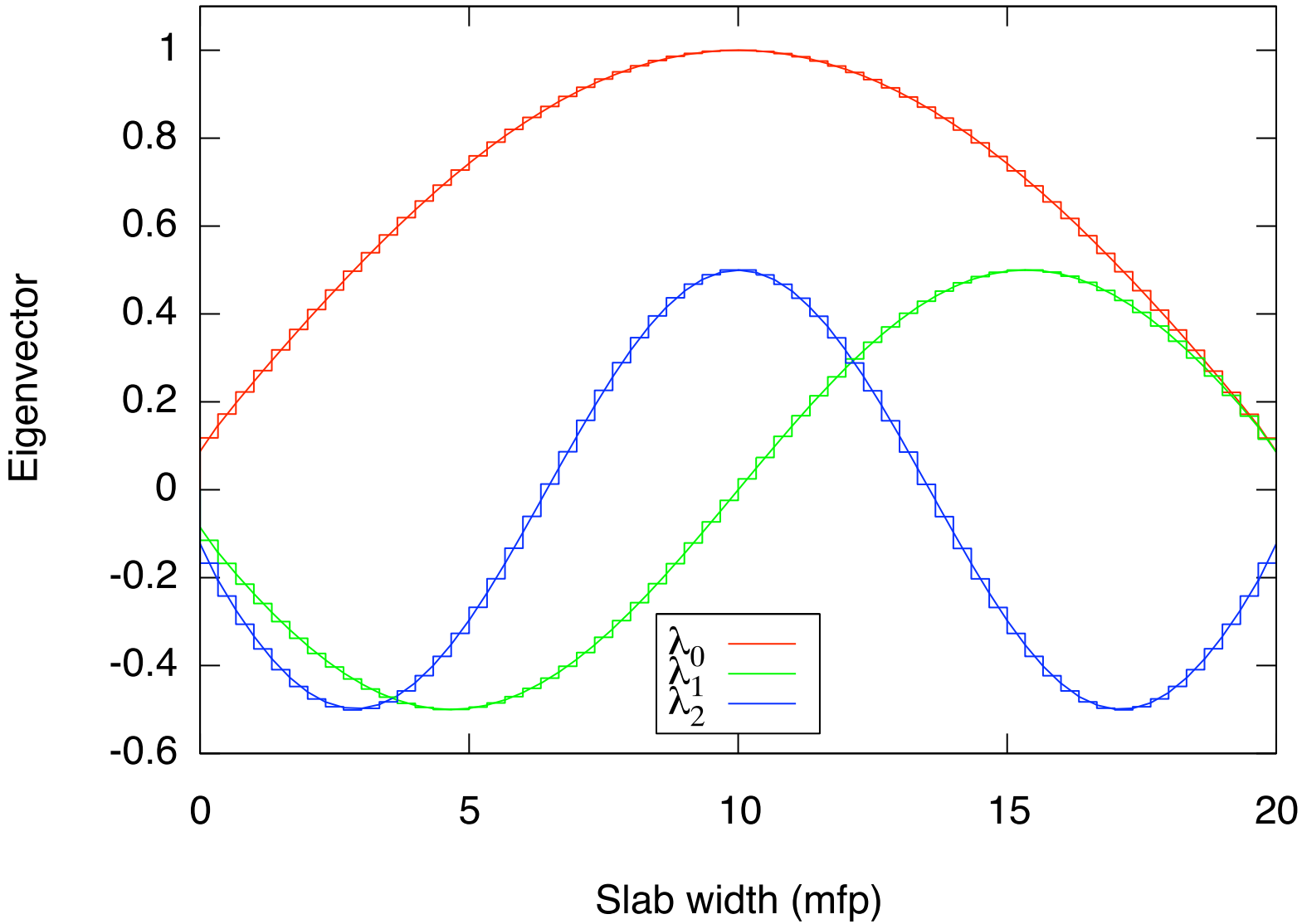
Spatial Discretization—FOM



Spatial Discretization—Eigenvectors



Spatial Discretization—Eigenvectors



Spatial Discretization—Conclusions

- Spatially discretizing can cause an error in the eigenvalue estimate
- Second order accurate approximation:
  - Reduces error
  - Reduces standard deviation
  - Increases figure of merit
  - Improves eigenvector estimate

Relaxed Arnoldi

Residual

$$|r_m| = \|\mathcal{A}y - \mu y\| = \left|h_{m+1,m}\right| |e_m x_i| \,.$$

Relaxed Arnoldi

Residual

$$|r_m| = \left\|\mathcal{A}y - \mu y\right\| = \left|h_{m+1,m}\right| |e_m x_i| \,.$$

Bouras and Frayssé:

- Can *relax* application of $\mathcal{A}$
- Maintain convergence of Arnoldi's method
- Save computational expense

Relaxing Monte Carlo Arnoldi

Relax Monte Carlo = Run fewer particles

- As residual decreases:
  - Relaxation *increases*
  - Number of particles tracked *decreases*
- Relaxation controlled by parameter η

Number of particles tracked

$$N_k = f\, N_0$$

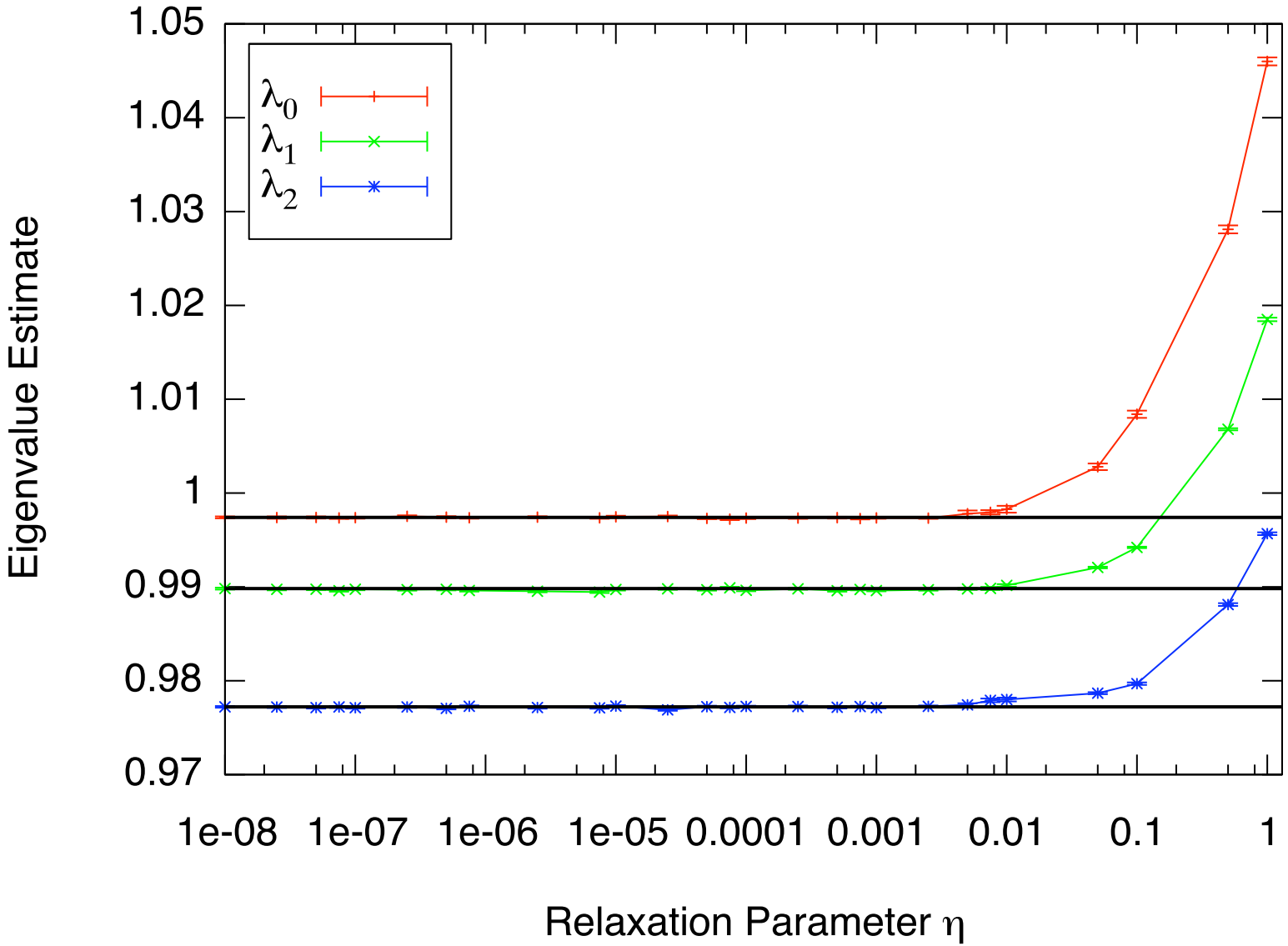
$$f = \begin{cases} \left(\left\| r_{k-1} \right\| / \eta \right)^2, & \left\| r_{k-1} \right\| < \eta \\ 1, & \left\| r_{k-1} \right\| \geq \eta \end{cases}$$

Numerical Results

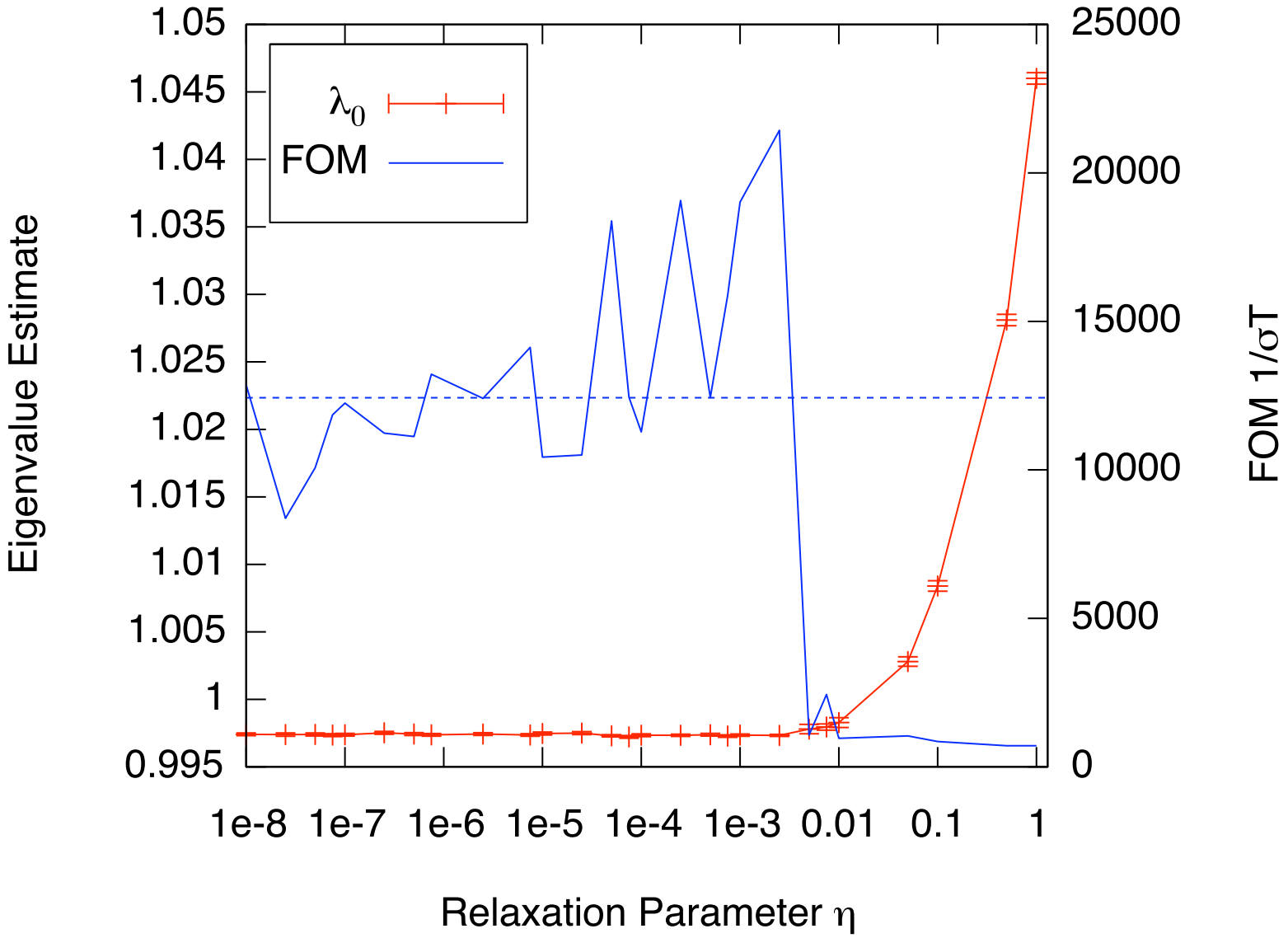
- $\nu\Sigma_f = 1.0$, $\Sigma_a = 0.2$, $\Sigma_s = 0.8$, $\Sigma_t = 1.0$
- 50 mfp thick slab (DR = 0.9924)
- 100 inactive restarts

- η ranges from 1E8–1.0
- Total number of particles is the same for all values of η

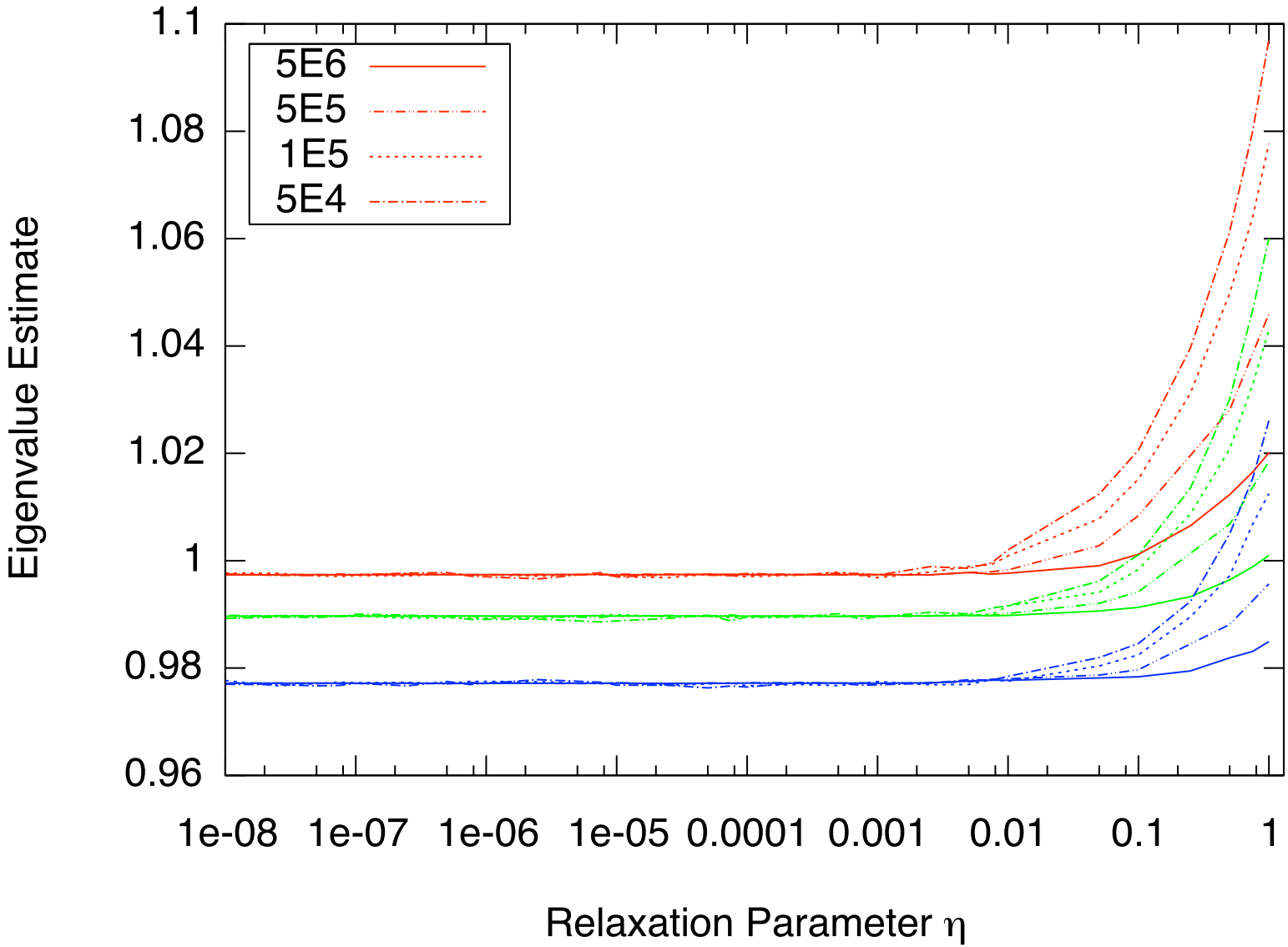
Numerical Results



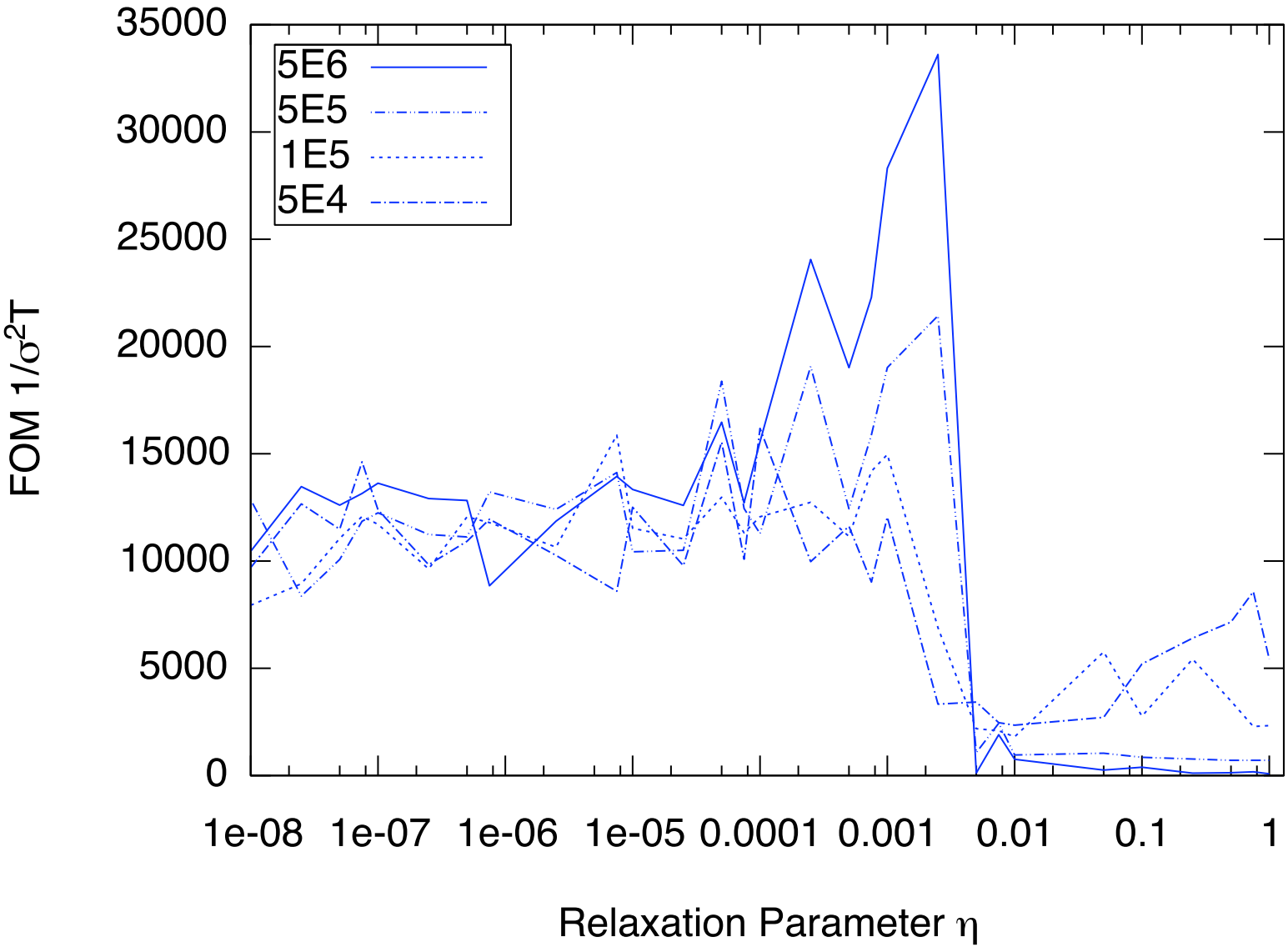
Numerical Results



Numerical Results



Numerical Results



Convergence in Monte Carlo Criticality Calculations

- Eigenvalue estimate converges proportional to dominance ratio λ_1/λ_0
- Fission source converges more slowly
- Shannon entropy

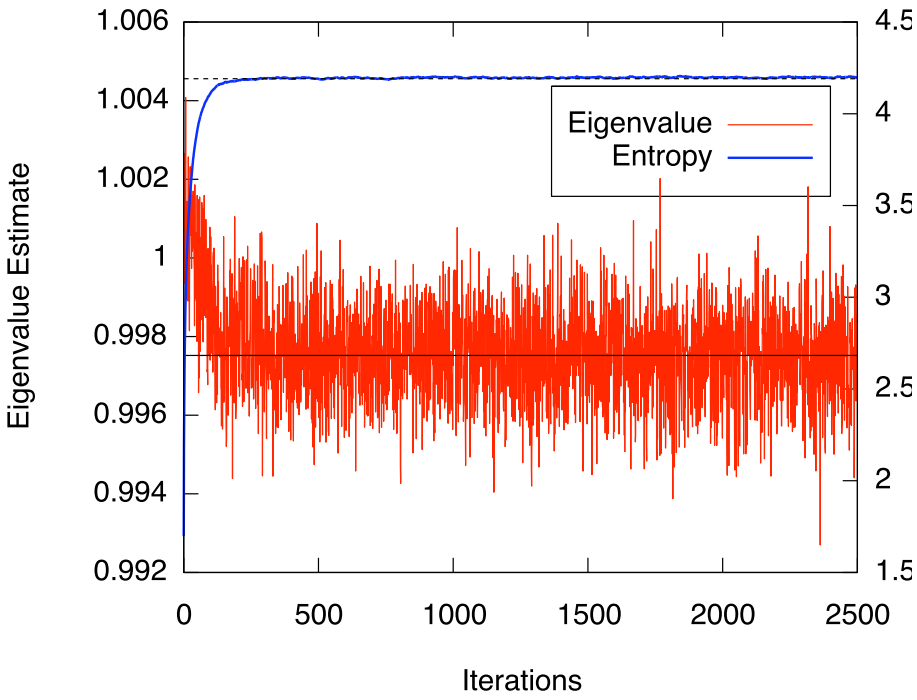
$$H\left(S(x)\right) \equiv -\sum_{b=1}^{B} S_b \log\left(S_b\right)$$

$$S_b = \int_{x_b}^{x_{b+1}} \left|v(x)\right| \mathrm{d}x$$

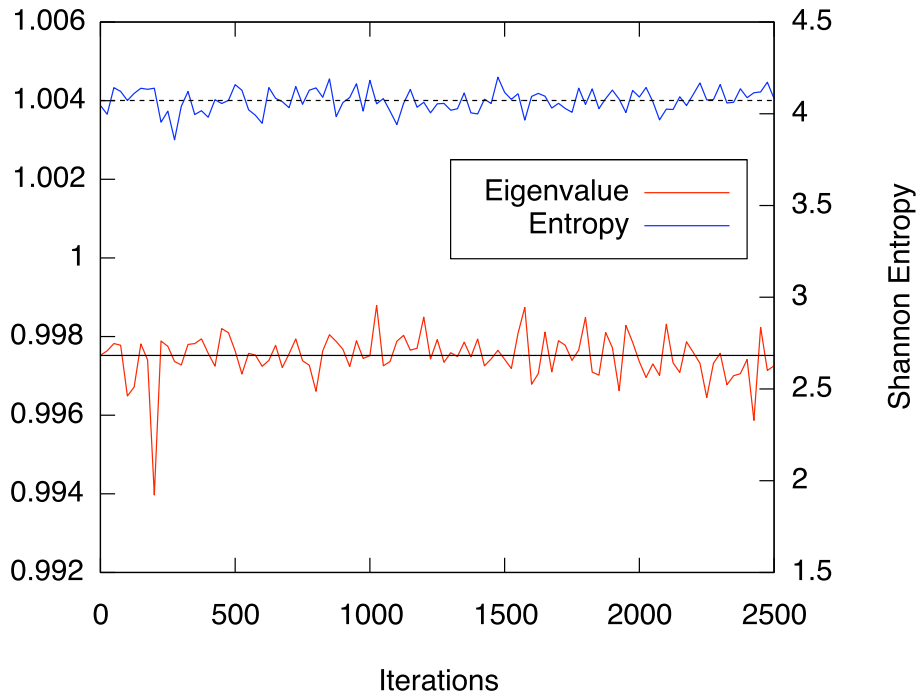
Homogeneous Geometry

- 50 mfp thick
- $\nu\Sigma_f = 1.0$, $\Sigma_a = 0.2$, $\Sigma_s = 0.8$, $\Sigma_t = 1.0$
- 5×10^5 particles per iteration
- 75 spatial bins (first order accurate)
- Source solely in left most bin

Homogeneous Geometry



Power Method

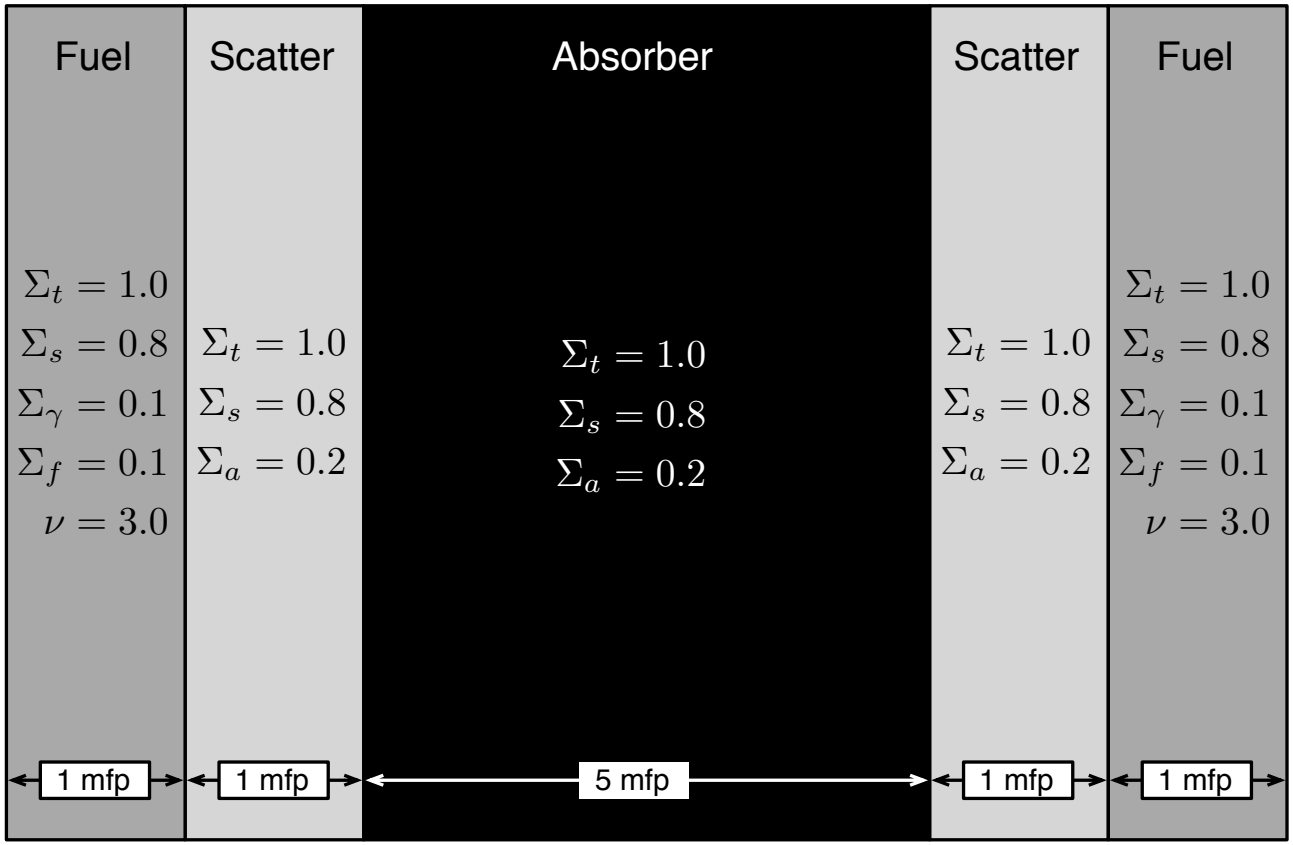


Arnoldi's Method

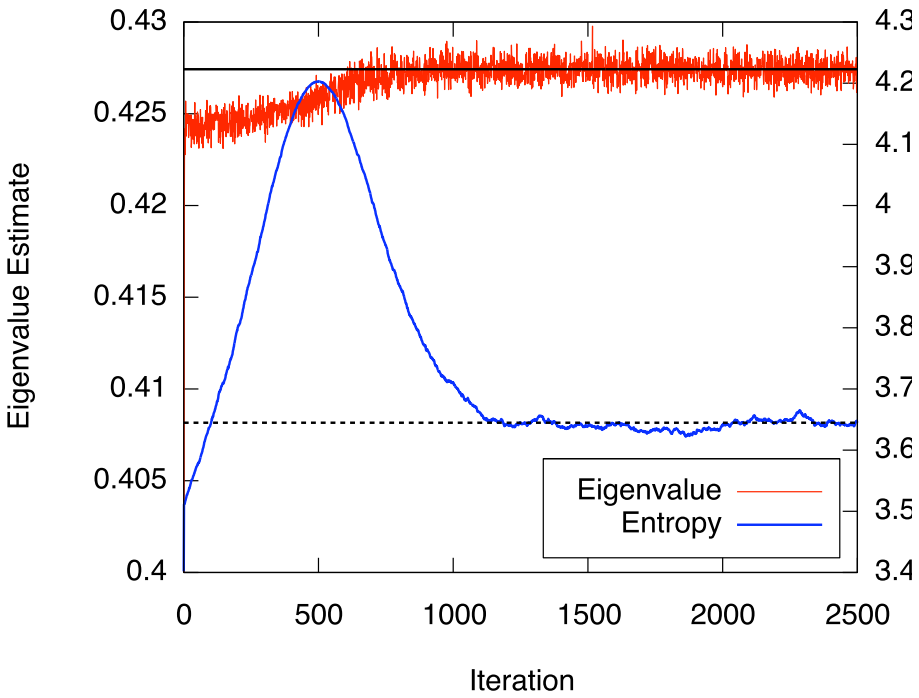
Heterogeneous Geometry

Parameters:

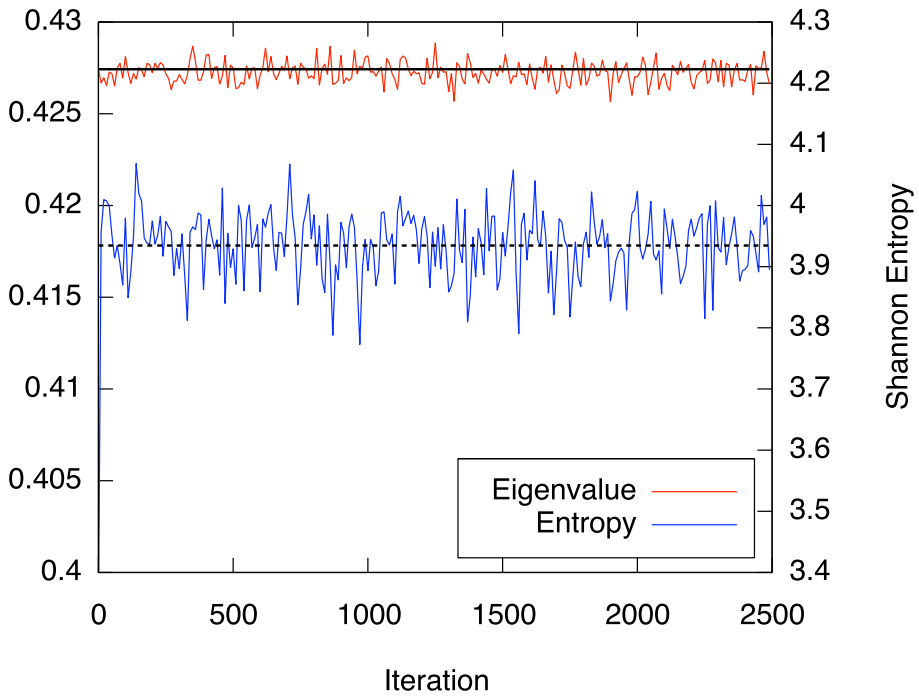
- 1×10^5 particles per iteration
- 10 iterations per Arnoldi restart
- 300 spatial bins (first order accurate)
- Source solely in left most bin
- Symmetric
- Asymmetric



Asymmetric Heterogeneous Geometry

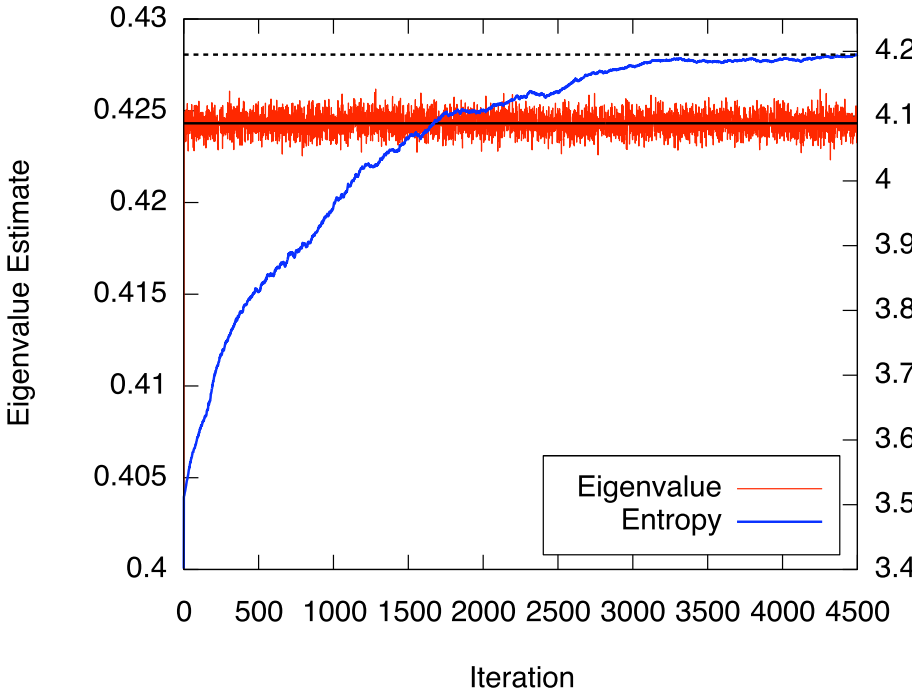


Power Method

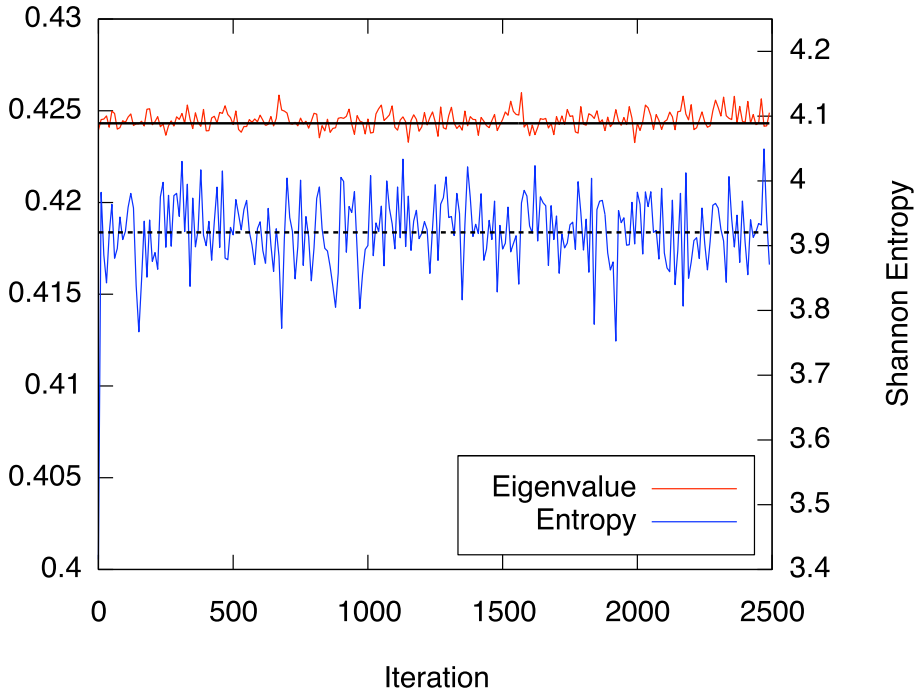


Arnoldi's Method

Symmetric Heterogeneous Geometry



Power Method



Arnoldi's Method

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations
- Arnoldi's method can estimate multiple eigenvalues

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations
- Arnoldi's method can estimate multiple eigenvalues
- Fission source is discretized
  - Discretization causes error in eigenvalue estimate
  - Second order accurate approximation reduces error and standard deviation

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations
- Arnoldi's method can estimate multiple eigenvalues
- Fission source is discretized
  - Discretization causes error in eigenvalue estimate
  - Second order accurate approximation reduces error and standard deviation
- Relaxing Arnoldi's method can increase the figure of merit

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations
- Arnoldi's method can estimate multiple eigenvalues
- Fission source is discretized
  - Discretization causes error in eigenvalue estimate
  - Second order accurate approximation reduces error and standard deviation
- Relaxing Arnoldi's method can increase the figure of merit
  - But you must be careful

Conclusions

- Arnoldi's method can be used in Monte Carlo Criticality Calculations
- Arnoldi's method can estimate multiple eigenvalues
- Fission source is discretized
  - Discretization causes error in eigenvalue estimate
  - Second order accurate approximation reduces error and standard deviation
- Relaxing Arnoldi's method can increase the figure of merit
  - But you must be careful
- Arnoldi's method converges in just a few iterations
  - Eigenvalue estimate
  - Fission source

Future Work

- Multi-dimensional and real-world problems
- Implicit restarts
- Calculating eigenvalue estimates at every iteration
- Condensing Arnoldi's method

Implicitly Restarted Arnoldi's Method

Danny Sorensen:

- Arnoldi's method can be *implicitly* restarted
- Implicit restarts trade iterations of the QR algorithm for Arnoldi iterations—much faster
- More robust restarts
- Mathematically equivalent to explicit restarts

Calculating Eigenvalue Estimates at Every Iteration

- $\sigma^2 \propto 1/N$
- Can increase N in Arnoldi's method by estimating eigenvalue after every iteration

Calculating Eigenvalue Estimates at Every Iteration

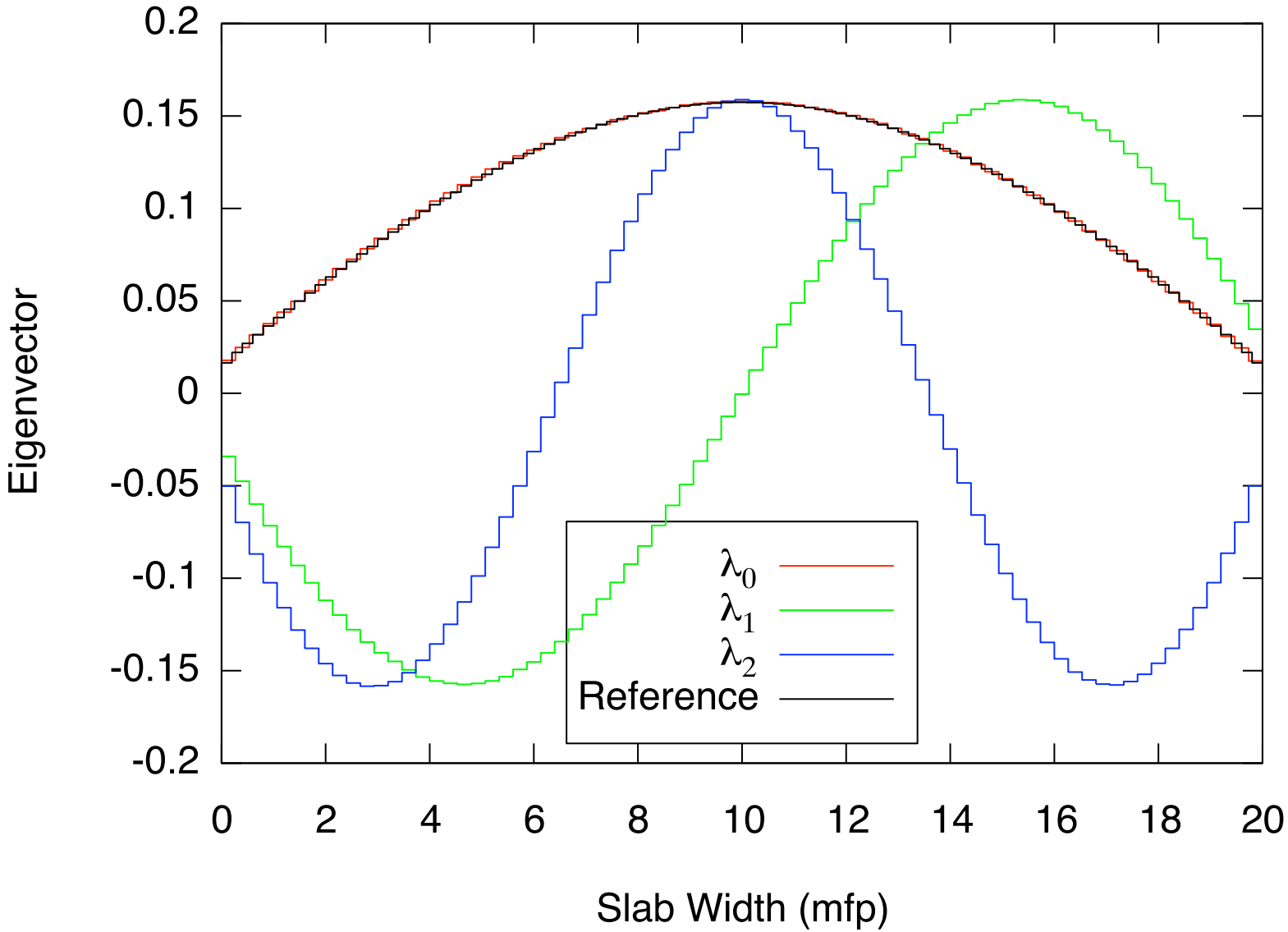
- $\sigma^2 \propto 1/N$
- Can increase N in Arnoldi's method by estimating eigenvalue after every iteration

	λ_0	σ	FOM
Arnoldi*	4.82806	6.3×10^{-4}	6.6×10^{3}
Power	4.82734	6.3×10^{-4}	5.4×10^{3}
Arnoldi	4.8290	1.5×10^{-3}	1.1×10^{3}

Condensing Arnoldi's Method

- Arnoldi's method does not need many iterations to converge
- Why not condense all particles tracked to just a few iterations?

Condensing Arnoldi's Method



Questions?

Extra Slides

Spatial Approximations

First Order Accurate

$$v_{\Pi}(x) = \sum_{b=1}^{B} a_b \Pi_b(x)$$

$$\Pi_b(x) = \left(\frac{1}{\Delta x_b} \right)^{1/2}$$

Second Order Accurate

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

$$\mathcal{L}_b(x, \alpha_b, \beta_b) = \alpha_b + \beta_b x$$

Spatial Approximations

First Order Accurate

$$v_\Pi = \left[a_1, a_2, \ldots, a_B\right]^T$$

$$h_{jk} = \langle v_\Pi^{(j)}, v_\Pi^{(k)} \rangle = \sum_{b=1}^{B} a_b^{(j)} a_b^{(k)}$$

Second Order Accurate

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

$$h_{jk} = \langle v_\Pi^{(j)}, v_\Pi^{(k)} \rangle = \sum_{b=1}^{B} \left(\alpha_b^{(j)} \alpha_b^{(k)} + \beta_b^{(j)} \beta_b^{(k)} \right)$$

Sampling from $v_{\mathcal{L}}(x)$

- To sample from $v_{\mathcal{L}}(x)$, normalize

$$q_b = \int \left|\mathcal{L}_b(x, \alpha_b, \beta_b)\right| \mathrm{d}x$$

$$Q = \sum_{b=1}^{B} q_b$$

$$p_b = q_b / Q$$

$$p = \frac{1}{Q}\left[p_1, \ldots, p_B\right]$$

- p_b is probability of sampling from bin b
- Bin b is sampled to determine neutron position from PDF

$$p_b(x) = \frac{1}{q_b}\left|\mathcal{L}_b(x, \alpha_b, \beta_b)\right|$$