International Conference on Mathematics, Computational Methods & Reactor Physics (M&C 2009)
Saratoga Springs, New York, May 3-7, 2009, on CD-ROM, American Nuclear Society, LaGrange Park, IL (2009)

USING A ZERO-VARIANCE SCHEME TO ACCELERATE THE FISSION SOURCE CONVERGENCE IN A MONTE CARLO CALCULATION

Stavros Christoforou and J. Eduard Hoogenboom
Delft University of Technology, Department of Applied Sciences
Mekelweg 15, 2629 JB, Delft, The Netherlands
s.christoforou@tudelft.nl; j.e.hoogenboom@tudelft.nl

ABSTRACT

We have used Boltzmann entropy in order to test whether a zero-variance based scheme can speed up the fission source convergence in a Monte Carlo calculation. It is shown that the choice of the initial source distribution significantly influences the evolution of the source, even leading to cases where the source does not converge at all throughout the calculation. The results from a loosely coupled system based on the NEA/OECD source convergence benchmarks indicate that, when using a biasing scheme such as the one we have developed, there can be significant improvement in the convergence, up to 3 times faster, which coupled with an figure of merit improvement of 1.5 leads to more efficient calculations.

Key Words: Monte Carlo, Adjoint biasing, Source convergence

1. INTRODUCTION

In a Monte Carlo calculation, the fission source distribution must be converged before one can reliably begin statistical sampling for k_{eff}. Usually, k_{eff} convergence is taken as an estimate of the source convergence in order to determine the number of inactive cycles in a simulation. However, k_{eff} being an integral quantity, can converge faster than the source. This is especially true in systems with dominance ratio close to 1, where k_{eff} can appear converged while the source distribution is clearly not.

In the last few years, several attempts have been made in order to reliably diagnose whether source convergence had been achieved in a simulation. An often used one suggests the use of the Shannon entropy of the source distribution [1]. The Shannon entropy S_{Sh} is defined as

$$S_{\mathrm{Sh}} = -\sum_{i=1}^{B} p_i \times \ln(p_i) \ , \tag{1}$$

where B is the number of spatial bins we divide our geometry into and $p_i = \frac{\text{\# of source particles in bin i}}{\text{total \# of source particles}}$. Convergence of the Shannon entropy of the source distribution is a much more reliable criterion for fission source convergence compared to k_{eff} and an easy one to test for.

A further development has been presented by Dumonteil et. al [2], who used the Boltzmann entropy S_{Bol} as a higher order correction to the Shannon entropy:

$$S_{\mathrm{Bol}} = S_{\mathrm{Sh}} - \frac{\ln\left((2N_s\pi)^{B-1}\prod_{i=1}^{B} p_i\right)}{2N_s} \ , \tag{2}$$

where N_s is the total number of particles. As seen in Eq. (2), the two expressions are equal for $N_s \to \infty$. In this paper, the concept of entropy will be used in order to determine whether a zero-variance-based scheme, developed by the authors for variance reduction purposes, can improve the convergence of the fission source distribution of the system.

2. THE ZERO-VARIANCE SCHEME FOR CRITICALITY CALCULATIONS

Our zero-variance scheme uses adjoint functions in order to bias the source, transition and collision kernels in a Monte Carlo simulation, which can lead to a result with zero variance [3]. The weight of the particles is corrected in order to compensate for the kernel biasing, so that an unbiased estimate of k_{eff} can be obtained.

Following the lifetime of a particle, we start with its generation at the source. Following the zero-variance scheme, particles should be selected by a distribution biased by the adjoint emission density $\chi^*(\mathbf{r}, E, \mathbf{\Omega})$, rather than the true source $S(\mathbf{r}, E, \mathbf{\Omega})$. Therefore, the biased source function is

$$\overline{S}(\mathbf{r}, E, \mathbf{\Omega}) = S(\mathbf{r}, E, \mathbf{\Omega}) \frac{\chi^*(\mathbf{r}, E, \mathbf{\Omega})}{R} \ , \tag{3}$$

where the expected detector response R is used for normalization. The transition kernel $T(\mathbf{r}' \to \mathbf{r}, E', \mathbf{\Omega}')$ is biased by the adjoint collision density $\psi^*(\mathbf{r}, E', \mathbf{\Omega}')$, since it represents the importance of particles entering a collision at $(\mathbf{r}, E', \mathbf{\Omega}')$. Similarly, we bias the collision kernel $C(\mathbf{r}, E' \to E, \mathbf{\Omega}' \to \mathbf{\Omega})$ by the adjoint emission density $\chi^*(\mathbf{r}, E, \mathbf{\Omega})$. The two biased kernels are then formulated as follows:

$$\overline{T}(\mathbf{r}' \to \mathbf{r}, E', \mathbf{\Omega}') = T(\mathbf{r}' \to \mathbf{r}, E', \mathbf{\Omega}') \frac{\psi^*(\mathbf{r}, E', \mathbf{\Omega}')}{\chi^*(\mathbf{r}', E', \mathbf{\Omega}')} \ , \tag{4}$$

and

$$\overline{C}(\mathbf{r}, E' \to E, \mathbf{\Omega}' \to \mathbf{\Omega}) = C(\mathbf{r}, E' \to E, \mathbf{\Omega}' \to \mathbf{\Omega}) \frac{\chi^*(\mathbf{r}, E, \mathbf{\Omega})}{\psi^*(\mathbf{r}, E', \mathbf{\Omega}')} \ . \tag{5}$$

Biasing the kernels by the adjoint functions as above, we reach a result with zero variance. However, it is obvious that the zero-variance scheme is mostly of theoretical value. Indeed, in practice, if one wishes to obtain the adjoint functions χ^* and ψ^* in order to bias the forward solution, one must solve the adjoint problem, thus making the forward calculation redundant. However, the scheme can still lead to a decrease in variance (never to zero, of course) when approximate adjoint functions are used. That way, obtaining a computationally cheap estimate of the adjoints could help reduce the variance of the expensive forward Monte Carlo simulation.

3. SOURCE AND KERNEL BIASING IN THE ZERO-VARIANCE SCHEME

Even if we restrict our problem to 1D geometry, the adjoint functions cannot be analytically calculated and used for any but the simplest of systems. Therefore, we use an adjoint deterministic calculation with a special source term in order to obtain the adjoints as discrete values at predetermined points in a grid. However, biasing the kernels requires that the adjoint function $\chi^*(x)$ is available at any point x (in a 1D geometry). In order to achieve that, we have to interpolate between the grid points where the adjoints are available. In this paper we make use of two methods of interpolation: constant adjoints throughout a bin, where a bin is defined as the space between grid points, and linear interpolation of adjoints between two

successive grid points. When using binwise constant adjoints, the adjoint is taken as the average value of the adjacent grid points.

The use of a deterministic code means that our adjoints are generated in groupwise form in energy, and discrete form in direction. We have kept those characteristics in the Monte Carlo simulation, in order to have a better comparison measure.

In the scheme, when using a multigroup treatment, the energy of neutrons generated by fission at a point $P(\mathbf{r}, g, \mathbf{\Omega})$ is selected from the fission spectrum χ_f biased by the direction averaged adjoint function $\chi^*(\mathbf{r}, g)$. In general terms:

$$\overline{S}(\mathbf{r}, g, \mathbf{\Omega}) = S(\mathbf{r}, g, \mathbf{\Omega}) \frac{\chi^*(\mathbf{r}, g, \mathbf{\Omega})}{\int S(\mathbf{r}, g, \mathbf{\Omega}) \chi^*(\mathbf{r}, g, \mathbf{\Omega}) dP} \ . \tag{6}$$

Since we want to have an unbiased estimate for k_{eff}, we need to weigh the particles appropriately. The correction weight is simply the ratio of the initial and modified PDFs:

$$W = \frac{S(\mathbf{r}, g, \mathbf{\Omega})}{\overline{S}(\mathbf{r}, g, \mathbf{\Omega})} = \frac{\int S(\mathbf{r}, g, \mathbf{\Omega}) \chi^*(\mathbf{r}, g, \mathbf{\Omega}) dP}{\chi^*(\mathbf{r}, g, \mathbf{\Omega})} \ . \tag{7}$$

The particle direction is then selected from the isotropic distribution biased by the directional adjoint function, in a way similar to collision biasing, shown later. Again, after biasing we need to alter the statistical weight of the particle. In this case, it needs to be set inversely proportional to the adjoint function for the selected energy group and direction. This requires appropriate normalization of the weights of all particles in a new batch. In practice, when using a multigroup treatment, the initial weight is set equal to

$$W_S = \frac{\sum_{g''} \chi_{fg''} \chi^*_{g''}(\mathbf{r})}{\chi^*_g(\mathbf{r}, \mathbf{\Omega})} \ . \tag{8}$$

For biasing the transition kernel to select a new collision site, the scheme requires the adjoint function ψ^* for particles entering a collision at P as the biasing function. The new collision site is therefore selected from the biased transition kernel

$$\overline{T}(\mathbf{r}' \rightarrow \mathbf{r}, g, \mathbf{\Omega}) = \frac{T(\mathbf{r}' \rightarrow \mathbf{r}, g, \mathbf{\Omega}) \psi^*(\mathbf{r}, g, \mathbf{\Omega})}{\int T(\mathbf{r}' \rightarrow \mathbf{r}'', g, \mathbf{\Omega}) \psi^*(\mathbf{r}'', g, \mathbf{\Omega}) dV''} \ . \tag{9}$$

We can convert the $\chi^*(x)$ from our deterministic output to the required $\psi^*(x)$ using the following equation:

$$\psi^*(\mathbf{r}, g, \mathbf{\Omega}) = \frac{\nu \Sigma_{fg}}{\Sigma_{tg}} + \sum_{g'} \frac{\Sigma_{g \rightarrow g'}}{\Sigma_{tg}} \sum_{N_{\text{dir}}} w_m \chi^*(\mathbf{r}, g', \mathbf{\Omega}) \ . \tag{10}$$

In practice, the new path length needs to be selected from the normalized probability

$$\xi = \frac{\int_0^s T(\mathbf{r} \rightarrow \mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega}) \psi^*(\mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega}_m) ds'}{\int_0^{s_{max}} T(\mathbf{r} \rightarrow \mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega}) \psi^*(\mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega}_m) ds'} \ . \tag{11}$$

where ξ is a random number, uniformly distributed between 0 and 1, $\mathbf{\Omega}_m$ is the direction cosine and g is the current energy group of the particle. A weight factor is then applied, in order to keep the final result unbiased:

$$W_T = \frac{\int_0^{s_{max}} T(\mathbf{r} \rightarrow \mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega})\psi^*(\mathbf{r} + s'\mathbf{\Omega}, g, \mathbf{\Omega}_m)ds'}{\psi^*(\mathbf{r}, g, \Omega)} \ . \tag{12}$$

In order to apply the biasing scheme to the collision kernel, we have to bias it by $\chi^*(x)$. The biasing is done in two steps - initially, we select the new energy group g after the collision with probability

$$\overline{p}(g|\mathbf{r}, g') = \frac{\Sigma_s(\mathbf{r}, g' \rightarrow g)\chi^*(\mathbf{r}, g)}{\sum_{g''} \Sigma_s(\mathbf{r}, g' \rightarrow g'')\chi^*(\mathbf{r}, g'')} \ . \tag{13}$$

After that, we need to select an outgoing direction, with probability

$$\overline{p}(\mathbf{\Omega}|\mathbf{r}, g', g) = \frac{\Sigma_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega}|g', g)\chi^*(\mathbf{r}, g, \mathbf{\Omega})}{\int \Sigma_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega}''|g', g)\chi^*(\mathbf{r}, g, \mathbf{\Omega}'')d\mathbf{\Omega}''} \ . \tag{14}$$

The (combined) weight factor to be applied for the scattering biasing will then be

$$W_C = \frac{\sum_{g''} \Sigma_s(\mathbf{r}, g' \rightarrow g'', \mathbf{\Omega}' \rightarrow \mathbf{\Omega})\chi^*(\mathbf{r}, g'', \mathbf{\Omega})}{\chi^*(\mathbf{r}, g, \mathbf{\Omega}_m)} \ . \tag{15}$$

4. TEST SETUP: A LOOSELY COUPLED SYSTEM

In order to test the effect of biasing on source convergence, we examined the case of a loosely coupled system. The particular case comes from the OECD/NEA benchmarks on source convergence, and more specifically benchmark 3 [4]. It is composed of a one-dimensional slab (infinite in the y- and z- directions), as shown in Fig. 1.

A vacuum boundary condition is imposed on both sides of the slab. The slab is divided in 3 sections: Two slabs containing a Uranyl solution with the same material concentrations, separated by a water slab. The width of the left Uranyl slab and the water slab was set to 20 cm, while the width of the right Uranyl slab was left variable, in order to test both symmetric and asymmetric cases. In this setup, k_{eff} converges faster than the fission source, leading to incorrect results when an insufficient number of initial cycles has been discarded. In addition, when the system is symmetric, there is a periodic shift in the distribution of the source between the two Uranyl slabs, making convergence even harder to predict.

Two cases from the benchmark were selected: Case 5 represents a fully symmetric, 20 cm - 20 cm - 20 cm configuration for the three slabs, while Case 6 introduces 10% asymmetry between the two fissile slabs, resulting in a 20 cm - 20 cm - 18 cm configuration. Using the compositions given in the benchmark specification, a reduced set of 3-group zone-mixed cross-sections for the 3 zones was prepared through the use of the SCALE code system [6]. For simplicity, only the P0 scattering matrix was used. The XSDRN code [5] was then used in order to generate the adjoint functions, which were passed to an in-house Monte Carlo code for biasing. A mesh of 0.5 cm was used. In addition, XSDRN was used in order to calculate the value of k_{eff} for the system, to be used as a reference point for our calculations.

In accordance to the benchmark specifications, 550 active generations or cycles were simulated, with another 50 generations skipped before tallying began. An initial number of 1000 neutrons was used, and

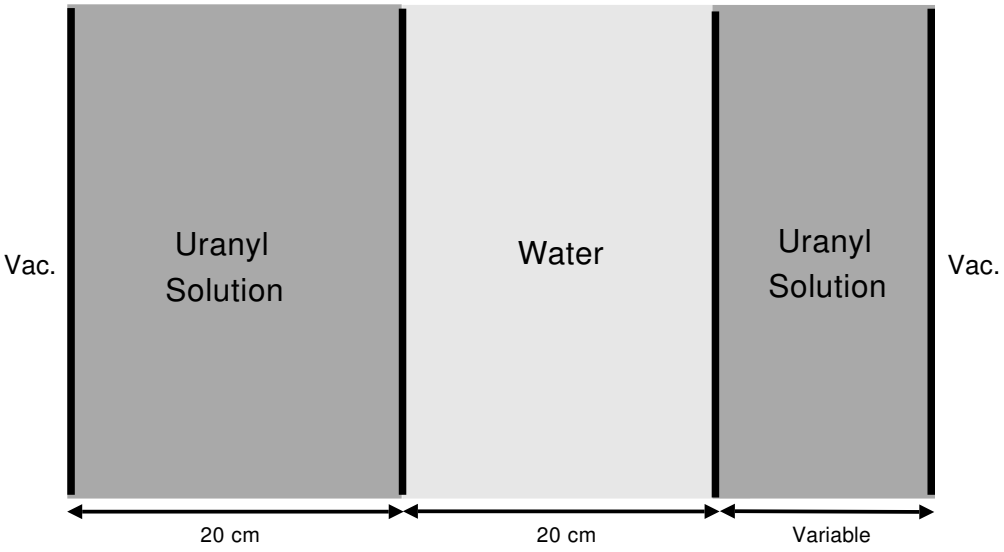

Figure 1. Geometry of the Loosely Coupled System

the number remained approximately constant throughout each calculation via population control. The Russian roulette threshold for implicit capture was set to 0.05 for all calculations, while a uniform initial source distribution was assumed.

5. RESULTS AND CONCLUSIONS

Initially, both systems were calculated for the 3 different variance reduction techniques: implicit capture only (no adjoint biasing), adjoint biasing using binwise constant adjoints and adjoint biasing using linearly interpolated adjoints. The k_{eff} and its error were then calculated and compared to the deterministic calculation. All cases fell within 3σ of the deterministic result, which indicates that the zero-variance scheme provides an unbiased result for k_{eff}.

Table I. Comparison of the different variance reduction techniques for the symmetric (case 5) and asymmetric (case 6) cases examined.

VR Technique	Case 5		Case 6	
	k_{eff}	σ_{relative}	k_{eff}	σ_{relative}
Implicit Capture	1.0491	6.2100×10^{-4}	1.0471	6.7709×10^{-4}
Binwise constant	1.0489	4.2250×10^{-4}	1.0481	4.0409×10^{-4}
Linear interpolation	1.0489	3.4384×10^{-4}	1.0473	3.6378×10^{-4}
Deterministic result	1.0486		1.0475	

In loosely coupled systems, particles have a low probability of moving between fissile volumes. This means that the initial guess for the fission source should play an important role in the evolution of the source and the k_{eff} eigenvalue. To test whether this happens, we ran each calculation using two different source configurations: One where the source is uniformly distributed over the fissile regions, and one where the source is concentrated in one point, namely at $x = 5.0$ cm.

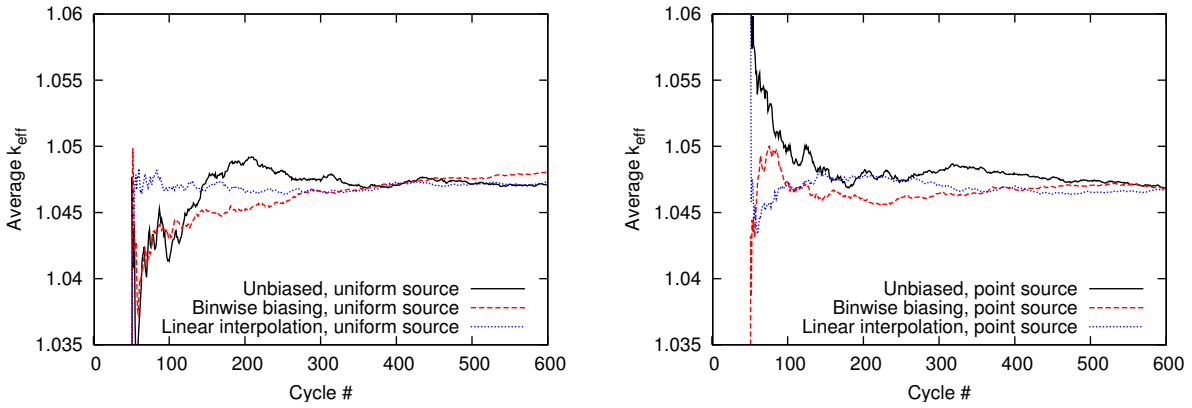

Figure 2. k_{eff} for the asymmetric system using uniform (left) and point (right) initial source. The calculation of the mean starts after the first 50 inactive cycles.

We can see the evolution of the k-eigenvalue for the two different cases in the asymmetric system in Fig. 2. In all cases, k_{eff} has converged to the expected value by the end of the calculation. As we can see, biasing using linearly interpolated adjoints helps convergence significantly, something that can be explained by the next plot.

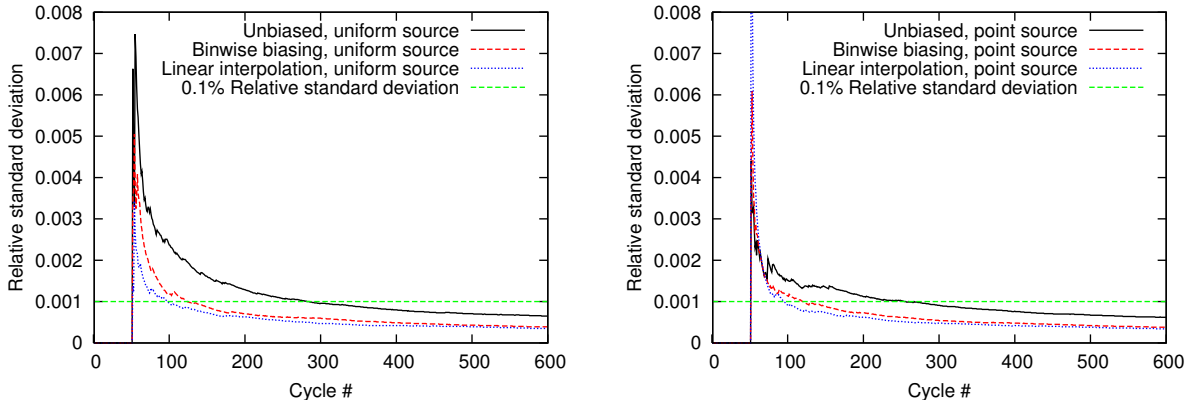

Figure 3. Relative standard deviation for the asymmetric system using uniform (left) and point (right) initial source

In addition to the accuracy of k_{eff}, in a criticality calculation, especially in safety calculations, the

statistical error of the calculation must usually fall below a certain threshold. We therefore set as a target of the calculation to achieve a relative standard deviation of 0.1%. The real advantage of the biasing scheme is now apparent in Fig. 3. Using biasing, the required standard deviation is achieved using fewer than 50 active cycles, while the unbiased calculation requires an extra 200 cycles before convergence is achieved. Coupled with the faster convergence of k_{eff}, one can have the expected result after only 100 total cycles, rather than 300 or more required for the unbiased calculation. Here, one can also observe a slight

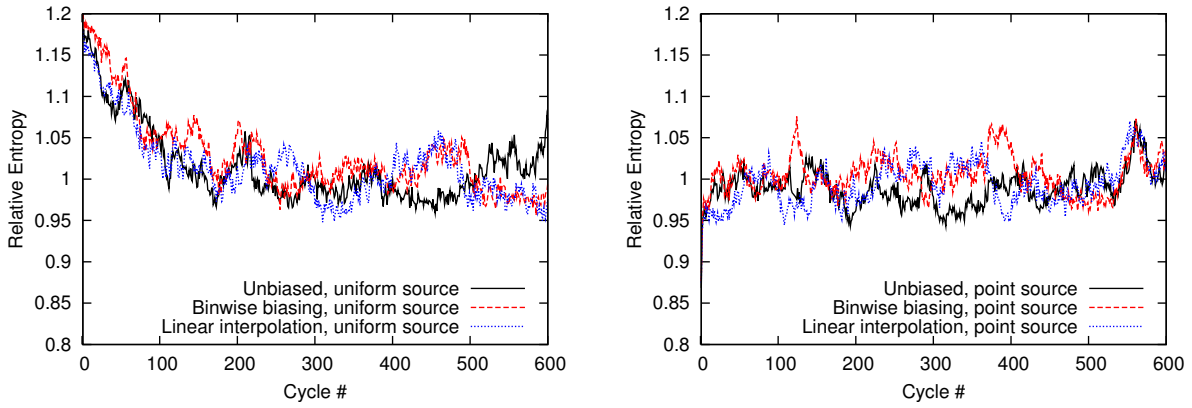

Figure 4. Progressive relative entropy for the asymmetric system using uniform (left) and point (right) initial source

improvement when going from a binwise constant approximation of the adjoints to a linearly interpolated approximation. As the normalization constants are now closer to the analytically calculated ones, particle weight is better controlled and therefore the standard deviation of a cycle is lower. This leads to a smaller variation between cycles, and a smaller standard deviation for the whole calculation.

However, as was discussed in the introduction, in such systems the source can converge slower than the k-eigenvalue. In Fig. 4, we can see the progressive relative entropy of the asymmetric system for the duration of the calculation. Here, progressive relative entropy is defined as the entropy of progressive cycles when compared to the mean of the final 200 cycles of the calculation. In our system, although the number of particles per cycle was low, the Boltzmann and Shannon relative entropies were indistinguishable from each other, so the plots represent both.

Surprisingly, all three methods showed very similar behaviour. The source converges almost at the same time scale as the k-eigenvalue, which means that the same number of cycles can be safely discarded in all cases. In this case, biasing does not seem to hold an advantage over the unbiased case when regarding source evolution. However, as we have already seen, the reduction in variance is noticeable and leads to a more efficient calculation, even taking into account that the biasing cases take twice as much in CPU time, compared to the unbiased one.

We will not make a discussion of the absolute efficiency of the scheme here, since the CPU time cost of the biasing can vary in orders of magnitude, depending on the implementation of the biasing technique in the computer code used. Some factors that affect the CPU time cost are the storage of biased probability tables, the root-finding method used during transition biasing and its tolerance, the detail of the adjoint grid and the Russian roulette threshold used for implicit capture. We will only mention here that in our implementation

on a simple Monte Carlo code, the efficiency improvement when using biasing was of the order of 50%.

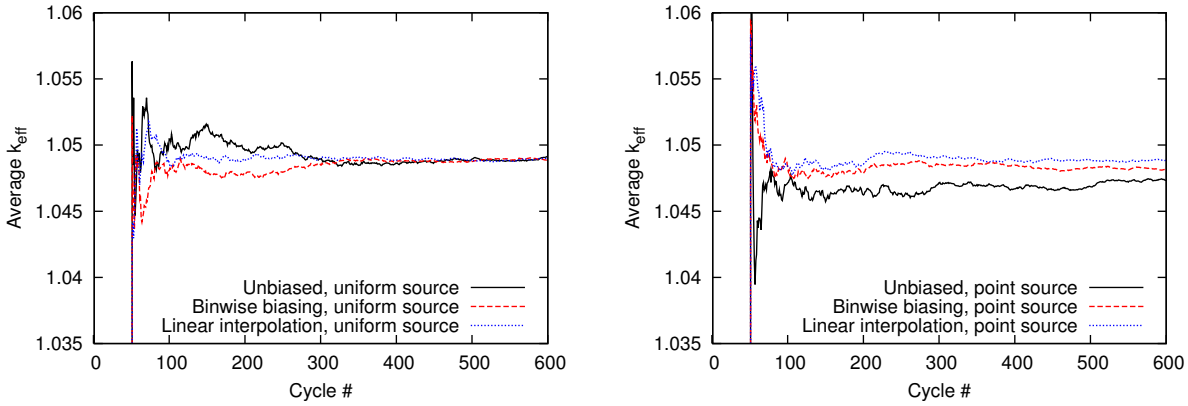

Figure 5. k_{eff} **for the symmetric system using uniform (left) and point (right) initial source.**

Moving to the symmetric system (case 5 in the benchmark), one can expect a more difficult convergence of the source distribution. When using a uniform initial source, k_{eff} converges to the expected eigenvalue in all cases. On the other hand, when using a point source in the left fissile slab, most of the particles periodically travel between the slabs, resulting in underprediction of the k_{eff}, since the fraction of fission in one slab can become very low during a cycle. Meanwhile, when biasing is used, particles are directed to both slabs by their higher importance when compared to the water slab, therefore showing similar behaviour between the uniform and point source configurations. This can be seen in Fig. 5.

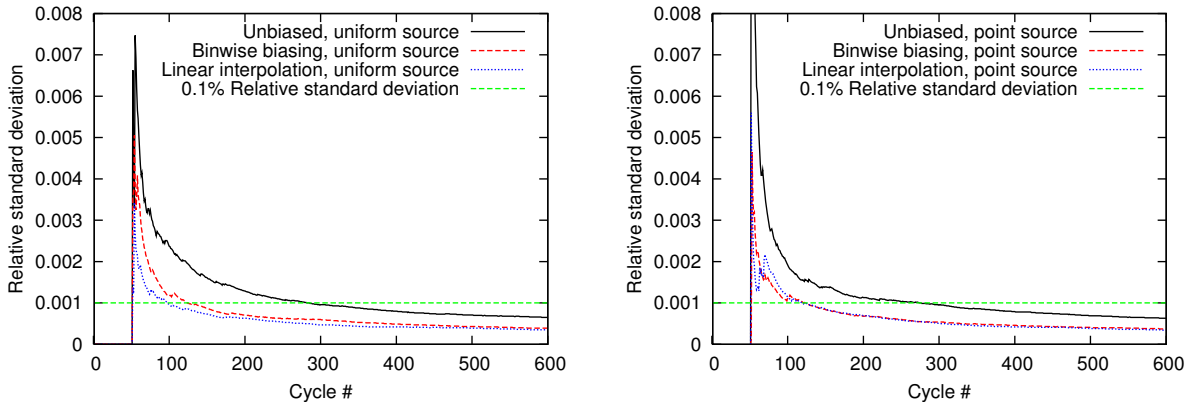

Figure 6. Relative standard deviation for the symmetric system using uniform (left) and point (right) initial source

In terms of the relative standard deviation, similar results to the asymmetric case can be observed in Fig. 6. Again, the target relative standard deviation is reached between 2 and 3 times faster when using biasing, again with the linear interpolation technique being slightly more efficient in reducing the error, compared to the binwise constant one.

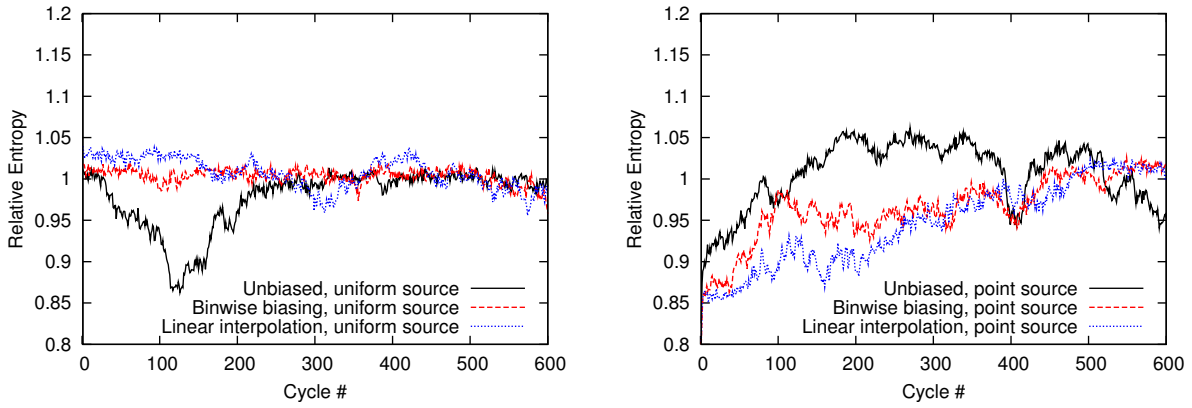

Figure 7. Progressive relative entropy for the symmetric system using uniform (left) and point (right) initial source

In Fig. 7 we can see the progressive relative entropy results for the symmetric system. Here, the results clearly show the importance of choosing a representative initial source distribution for the calculation. In the point source case, the entropy does not converge in any of the three cases, although the two biased cases show a slow convergence towards the mean. On the contrary, when starting with a uniform source, the biased cases show excellent results for the source entropy. The unbiased case, on the other hand, shows significant fluctuations even when uniform source is used, forcing a larger number of cycles to be simulated before one can consider it converged. Since the behaviour of symmetric, loosely coupled systems can change significantly with only small perturbations in the system, we ran the symmetric case 10 times with different initial random number seeds and compared the resulting entropies. In all but one case, the behaviour of the biased cases matched this of the plot shown above, while the unbiased case showed a dip similar to the one shown here, but in random stages of the source evolution.

We can therefore state with confidence that the convergence of the source in systems that are difficult to predict, such as symmetric systems, is assisted by biasing with the adjoints, using the variance reduction scheme we have developed. It is very important to correctly set the initial source distribution, since it can prevent the source from converging, even when biasing is used. However, even in simpler cases, where the source converges rapidly, the scheme can assist with the k_{eff} convergence by reducing the variance of the system, allowing for fewer generations to be simulated in order to reach the target error for a calculation.

REFERENCES

[1] T. Ueki and F. B. Brown, “Stationarity modeling and informatics-based diagnostics in Monte Carlo criticality calculations”, *Nuclear Science & Engineering*, **149**, pp.38-50 (2005).

[2] E. Dumonteil et al., “Source convergence diagnostics using Boltzmann entropy criterion application to different OECD/NEA criticality benchmarks with the 3-D Monte Carlo code Tripoli-4”, *Proceedings of PHYSOR 2006*, Vancouver, Canada, September 10-14 (2006).

[3] S. Christoforou and J. E. Hoogenboom, “A zero-variance-based scheme for variance reduction in Monte Carlo criticality”, *Proceedings of PHYSOR 2006*, Vancouver, Canada, September 10-14 (2006).

[4] Nuclear Energy Agency NSC Working Party on Nuclear Criticality Safety (WPNCS) homepage, http://www.nea.fr/html/science/wpncs/convergence/index.html.

[5] Greene, N. M. and Craven Jr, C. W. , "XSDRN, A Discrete Ordinates Spectral Averaging Code", ORNL-TM-2500, Oak Ridge National Laboratory (1969)

[6] "SCALE, A Modular Code System for Performing Standardized Computer Analyses for Licensing Evaluations", ORNL-TM-2005/39, Version 5.1, Vols. I-III, Oak Ridge National Laboratory (2006)