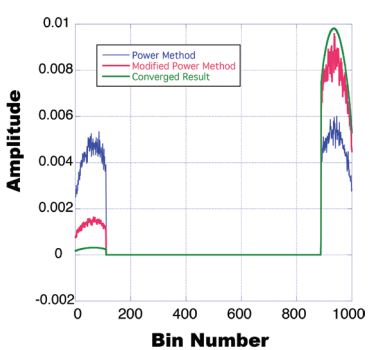
Monte Carlo Methods for Multiple Extremal Eigenvalues of Very Large Matrices

James E. Gubernatis, T-4; Thomas E. Booth, XCP-4

We developed a Monte Carlo method to estimate several of the largest or smallest eigenvalues and eigenvectors of a very large matrix. Our new method has three key components. One is a new procedure for estimating the eigenvalues, the second is a set of procedures for sampling a mixed sign function, and the third is an exact procedure for canceling signed random walkers over a region. Our new method for estimating the eigenvalues does not require a step explicitly orthogonalizing the eigenpairs. This is valuable because the matrices we can treat are sometimes so large that we cannot store the information necessary to do this. The sampling of mixed signed functions is accomplished by using random walkers of mixed sign and transferring the sampling task to the promotion of interference between them, so that on the average they represent the nodal structure of the eigenvector.

Fig. 1. The fission density amplitude convergence for a nuclear fuel slab problem. The left and right slabs, of slightly different widths, are the nuclear fuels that fission neutrons. In between are slabs that more strongly scatter and absorb the neutrons. The green curve is the result of running our new method or the standard method a very long time. The blue (standard) and red (new) show the state of the standard and new methods after 200 cycles of the simulation.



Finding the eigenvalues and eigenvectors of a matrix is a ubiquitous task in science and engineering. These eigenpairs may represent, for example, the acoustic modes of a vibrating system such as a violin. For many applications, ample well-developed software is readily available to solve the eigenvalue problem and works very well provided computer memory is sufficient to store the matrix. Typically, this software returns all the eigenvalues and eigenvectors. In other applications the matrix becomes too large to store. Fortunately, just the largest or smallest eigenvalue often provides adequate valuable information. For this type of problem other techniques exist that are sometimes useable deterministically, but are most often based on the Monte Carlo method. The Monte Carlo method becomes essential when the matrix is too big to store. This type of problem, for example, occurs in calculations of what is called the ground state energy of atoms, molecules, and solids. The ground state energy is the smallest eigenvalue of the Hamiltonian matrix representing the physical systems.

Knowing more than just one extremal eigenvalue is often desired and sometimes required. In designing a nuclear reactor, for example, the extremal eigenvalue is set by the details of the design, but the closeness of the second eigenvalue to it determines not only the stability of the design but also the efficiency of the Monte Carlo method simulating the design. Clearly, there are cases where it is better if one can find the second eigenvalue than if one cannot.

Determining more than one extremal pair of eigenvalues by Monte Carlo methods has been a difficult task. Monte Carlo is a method that draws "random samples" from probability distributions. By definition, probability distributions must be non-negative everywhere. The technique that the Monte Carlo method uses to compute the dominant extremal eigenvalue draws samples from the unknown eigenvector. Fortunately for many applications, the dominant eigenvectors have no negative components. The very nature of eigenvectors, however, requires all pairs to be mutually orthogonal. This means that some components of each subdominant eigenvector must be negative. Hence, direct extension of existing Monte Carlo methods to the computation of multiple extremal eigenpairs requires sampling from at least one eigenvector that is not a natural probability function.

We have succeeded in extending the most commonly used Monte Carlo method for extremal eigenvalue computations to the computations of multiple extremal eigenpairs [1-3]. By “multiple” we mean of the order of a half dozen or so. We have benchmarked our extension on problems drawn from classical statistical mechanics [1-3], nuclear engineering [4-7], and applied mathematics [8]. Besides providing extra desired or needed information about the problem, our method also accelerates the convergence to the dominant state.

Our extension has three key components. One is a new procedure for estimating the eigenvalues [4,5], the second is a set of procedures for properly sampling for a mixed-sign function for which the location of its zeroes is unknown [2,6-8], and the third is an exact procedure for canceling signed walkers over a region as opposed to at a point [7]. This last component is crucial for eigenvalue problems defined in the continuum, such as the nuclear criticality problem in nuclear engineering. Our new method for estimating the eigenvalues does not require a step at which we need to enforce the orthogonalization

For more information contact James E. Gubernatis at jg@lanl.gov.

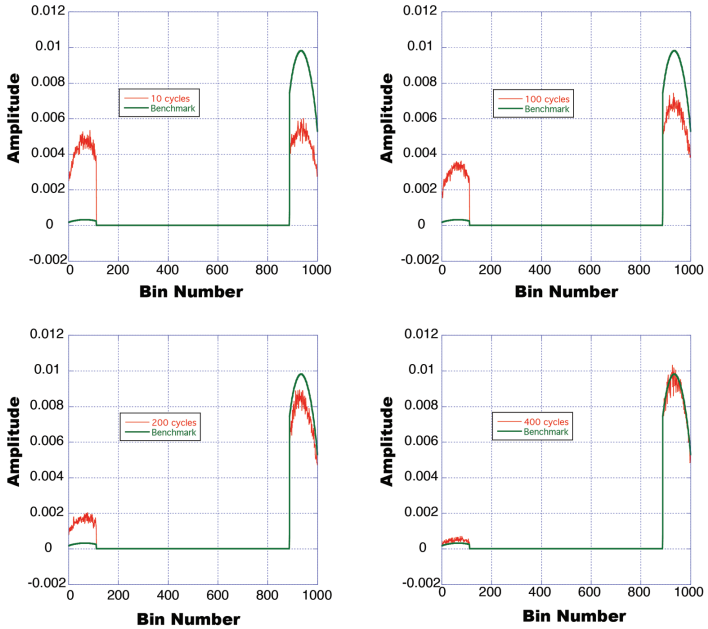


Fig. 2. The fission density amplitude convergence for a nuclear fuel slab problem. The slabs are the same as in Fig. 1, as is the green curve. The red curves show the state of the new method after 10, 100, 200, and 400 cycles of the simulation. Convergence occurred at approximately 450 cycles.

between eigenpairs explicitly. Not requiring this enforcement is valuable because the matrices we can treat are sometimes so large that we cannot store the information necessary to do this. The sampling of mixed-sign functions was accomplished by using random walkers of mixed sign, and transfers the sampling task to the promotion of interference between walkers of opposite signs so that on the average they represent the nodal structure of the eigenvector.

In Figs. 1 and 2 we show one of our benchmark calculations [6]. The generation of fission neutrons is being simulated in two slabs of nuclear fuel which are well, but not completely, shielded from each other. What makes this problem a difficult Monte Carlo simulation is the near equality of the dominant and subdominant eigenvalues, which causes the simulations to converge very slowly. In nuclear engineering parlance, the dominance ratio of the problem is nearly one. Our Monte Carlo method is an iterative one, meaning the iteration is cycled until it converges to sampling a stationary state. After the sampling becomes stationary, data is collected to compute estimates of important physical properties. In Fig. 1 we show how close the standard method is to convergence and our new one after 250 cycles. The exact solution has a specific asymmetry that our method has locked onto, but the standard method has not yet expressed. Figure 2 shows a sequence of partial convergences of the new method compared to the benchmark. For this problem we estimate that it takes 450 cycles to converge. We found that the standard method takes 5000 cycles.

Our first suite of benchmarks had dominant eigenvectors guaranteed to be positive so it was only the negativity of the subdominant state that was a concern. A next set of benchmarks was for problems where we knew the dominant state a priori. Here we created a modified algorithm using this information, thereby reducing the simulation to finding just the subdominant eigenpairs [8]. Curiously an application of this specialization is the Monte Carlo method itself. A measure of efficiency is the dominance ratio of the transition matrix defining the Monte Carlo steps [8]. By computing it we can compare the likely efficiency of different algorithms and, in some cases, by adjusting the parameters of a given algorithm we can improve its efficiency. Currently, we are studying problems where not even the dominant eigenvector is non-negative. This is a common occurrence in many quantum mechanical problems where it is called the sign problem. In our other work we have already demonstrated that in some cases we can solve this problem.

[1] Gubernatis, J.E., and T.E. Booth, *J Comput Phys* **227**, 8508 (2008).
[2] Booth, T.E., and J.E. Gubernatis, *Phys Rev E* **80,** 046704 (2009).
[3] Booth, T.E., and J.E Gubernatis, *Comput Phys Comm* **180,** 509 (2010).
[4] Booth, T.E., *Nucl Sci Eng* **143,** 291 (2003).
[5] Booth, T.E., *Nucl Sci Eng* **154,** 48 (2006).
[6] Booth, T.E., and J.E. Gubernatis, "Improved Criticality Convergence via a Modified Monte Carlo Power Iteration," *Int Conf Math Comput Meth Reactor Phys (M&C 2009)*, Saratoga Springs, New York; Los Alamos Report LA-UR-08-06461 (2009).
[7] Booth, T.E., and J.E. Gubernatis, *Nucl Sci Eng* **165,** 283 (2010).
[8] Rubenstein, B.M. et al., *Phys Rev E* **82,** 36701 (2010).

Funding Acknowledgments
LANL Laboratory Directed Research and Development - Exploratory Research Program