Geometrically Convergent Learning
Algorithms for Global Solutions of Transport Problems

Jerome Spanier

Department of Mathematics
Claremont Graduate University
925 N. Dartmouth Ave.
Claremont, CA 91711, USA
Email: jerome.spanier@cgu.edu

Abstract. In 1996 Los Alamos National Laboratory initiated an ambitious five year research program aimed at achieving geometric convergence for Monte Carlo solutions of difficult neutron and photon transport problems. Claremont students, working with the author in Mathematics Clinic projects that same year and subsequently, have been partners in this undertaking. This paper summarizes progress made at Claremont over the two year period, with emphasis on recent advances.

The Claremont approach has been to maintain as much generality as possible, aiming ultimately at the Monte Carlo solution of quite general transport equations while using various model transport problems - both discrete and continuous - to establish feasibility. As far as we are aware, prior to this effort, only the discrete case had been seriously attacked by sequential sampling methods: by Halton beginning in 1962 [1] and subsequently by Kollman in his 1993 Stanford dissertation [2]. In work performed in Claremont, an adaptive importance sampling algorithm consistently outperformed a sequential correlated sampling algorithm based on Halton’s ideas for matrix problems. These findings are contrary to what Halton reported in 1962 and in subsequent papers.

These learning algorithms based on very different Monte Carlo strategies have recently been successfully extended to continuous problems. This paper outlines the methods and ideas employed, sketches the algorithms used and exhibits the geometric convergence obtained. A rationale for the results obtained so far and an indication of some of the remaining obstacles to achieving fully practical computation of global transport solutions by these means is also presented.

1 Introduction

The Los Alamos National Laboratory initiated, in 1996, an ambitious five year research program aimed at achieving dramatically accelerated convergence for its widely used Monte Carlo Neutron and Photon program MCNP. This initiative, in which the efforts of research groups both inside and outside the National Laboratories have been coordinated, sought the holy grail of Monte Carlo: a geometrically convergent algorithm for estimating the global
solution of very general transport problems with high precision algorithm on which such a solution rests clearly must be capable of it progresses.

Two quite different such algorithms have been developed for an algorithm making use of correlated sampling; and an algorithm of importance sampling, both applied sequentially, or adaptively. reports on recent progress made in Claremont by a team of gradudate (Carole Hayakawa, Rong Kong, Yongzeng Lai and Diana Verzi) w the author, with financial support from Los Alamos National Lab

In [3], [4], [5], some of the early progress was reported, much rected at the solution of matrix transport problems. We found that of matrices developed from discrete approximations to transport consistently superior convergence was observed for an algorithm for importance sampling compared with one based on correlated sam This is contrary to what Halton reported in [1]; our result might be explained by the fact that our algorithms were tested on matrix dis of continuous transport problems, rather than on matrices chosen However, geometric convergence was obtained for both the correplling and the importance sampling algorithms, applied sequenti

A method for discretizing continuous transport problems by low discrepancy sequences in the physical phase space was also and the resulting discrete problem was then solved effectively by the methods previously developed for the matrix case. This method is used to estimate the solution at arbitrary points of the phase making use of a variant of Nystrom’s method [6]. The combinato be quite effective, especially because a theoretical error bound is minimized when the discrepancy is least. This suggests that low-point sets should be used to define the discrete transport problem when the phase space is multidimensional 1 and the effective approach was reported in [5].

More recently, a direct attack on the continuous transport proble also yielded geometrically convergent results. This method is bas expansion of the global solution in an appropriately chosen basis then truncated to produce a finitely defined computer approxim able at any point of phase space. This more recent work will be the of the present paper.

2 Operator Transport Equations

It has been convenient and economical to formulate our transpo in terms of the operator transport equation

\[ \phi = \mathcal{K}\phi + \sigma \]

1 For modeling steady state transport in general media, the appro space is six dimensional.
where $\mathcal{K}$ is a linear operator from $B$ to $B$ (a Banach space). We assume that the norm of the operator $\mathcal{K}$ is less than one $(\| \mathcal{K} \| < 1)$ to guarantee that the equation (1) has exactly one solution (although the weaker assumption $\| \mathcal{K}^{n_0} \| < 1$ for some integer $n_0 \geq 1$ will suffice; this follows from rather elementary fixed point theorems in $B$).

The adaptive methods for solving equation (1) iteratively in which we are interested can then be described by introducing a sequential dependence in both the source $\sigma$ and kernel $\mathcal{K}$ via the operator equation

$$\phi = \mathcal{K}^{(k)} \phi + \sigma^{(k)} \quad \sigma^{(0)} = \sigma, \quad \mathcal{K}^{(0)} = \mathcal{K}. \quad (2)$$

In this way, Halton's sequential correlated sampling [1], [7] results by selecting $\mathcal{K}^{(k)} = \mathcal{K}$ for all $k$ and by defining the "reduced source"

$$\sigma^{(k+1)} = \sigma^{(k)} + \mathcal{K} \hat{\psi}^{(k)} - \hat{\psi}^{(k)} \quad (3)$$

where $\hat{\psi}^{(k)}$ is an approximate solution to equation (2), while various more general biased and unbiased importance sampling sequential methods result by modifying both the source $\sigma$ and the kernel $\mathcal{K}$ suitably from one iteration to the next. Our interest in this paper is predominantly, but not exclusively, on this latter family of adaptive importance sampling methods, with a view to applying these directly to integral equations, taking advantage of the superior convergence rates observed in [5] for this method when compared with Halton's sequential correlated sampling method applied to matrix problems.

The more general operator formulation described above has two important applications, which otherwise have to be treated separately:

I. If $B = \mathbb{R}^k$ or $C^k$, regarded as $s$-dimensional real or complex vector spaces, $\mathcal{K} = A$ is an $s \times s$ real or complex matrix, $\phi = x$, $\sigma = a$ are $s$-vectors, then equation (1) reduces to the matrix equation

$$x = Ax + a$$

and may be interpreted as describing transport in an infinite homogeneous medium with $s$ discrete energy states available for each particle.

II. If $B = L^1(\Gamma)$, and $\mathcal{K}$ is an integral operator, equation (1) becomes the continuous transport equation

$$\psi(P) = \int_\Gamma K(P, P') \psi(P') dP' + S(P)$$

$$= \mathcal{K} \psi(P) + S(P) \quad (4)$$

where

$$\mathcal{K} \psi(P) = \int_\Gamma K(P, P') \psi(P') dP'.$$

This integral equation may be shown to be the appropriate equation to describe the directional, or vector, particle collision density, $\psi(P) = \Sigma_t(P) \phi(P)$,
where \( \Sigma_t(P) \) is the total macroscopic cross section\(^2\) at \( P \) and \( \phi(P) \) is the particle vector flux. In equation (4), \( P, P' \) denote generic points in phase space \( \Gamma \), \( S(P) \) is the density of first collisions, and \( K(I) \) is the transport kernel describing transitions from \( P' \) to \( P \).

In the context of equation (4), it is usual to apply Monte Carlo to the estimation of integrals of the form

\[
I = \int_{\Gamma} g(P)\psi(P)dP
\]

where \( g \) is some known bounded and (usually) nonnegative function which is frequently defined in terms of some interaction probability that satisfies equation (4). Our intention here, however, is to reconstruct the defined approximation \( \hat{\psi} \) to \( \psi \) from knowledge of a finite number of integrals of \( \psi \), each of which is of the form (5).

In contemplating the application of adaptive strategies to the solution of integral equations such as (4), we must, clearly, produce algorithms that seek only a finite number of unknown values. Since algorithms of primary interest to us will often require approximated solutions of the solution of equations such as (4) acting as an approximate function, there are only two obvious ways to obtain such information:

1. Replace equation (4) by a discrete approximation in the matrix equation and solve that equation adaptively. This was the view adopted in [5].

2. Represent the required approximate solution \( \hat{\psi} \) as a finite combination of appropriately chosen basis functions \( B_1, B_2, \ldots, B_M \):

\[
\hat{\psi}(P) = \sum_{i=1}^{M} a_i B_i(P)
\]

and determine the coefficients \( a_i \) of combination adaptively by Monte Carlo simulation. This is the perspective of the present paper.

3 Correlated and Importance Sampling

A closer inspection of equations (2) and (3) reveals that the methodological correlated sampling substracts an approximate solution from the equation at each stage, and uses the random walks generated by \( M \) in the next stage to identify an additive correction to the previous solution. The random walks in stage \( k \) originate in states that are defined as the average distance travelled between collisions in an infinite homogeneous medium of material with cross section \( \Sigma_t(P) \) has dimensions inverse length.\(^2\)
“reduced source” \( \sigma^{(k)} \), but successive collision states for each random walk in every stage are then generated using the original transport kernel. As the reduced source \( \sigma^{(k)} \) tends to zero, so does the correction \( \hat{g}^{(k)} \) derived from the reduced source, and the accuracy of the globally defined solution improves geometrically provided sufficiently many random walks are processed in each stage (see \[8\] for a result that assures such convergence in a probabilistic sense).

When importance sampling is applied adaptively to obtain the geometric convergence, the algorithm is quite different from the one just described. Information used to improve Monte Carlo estimates adaptively is described in terms of an importance function. The theory of importance sampling is well known (see, e.g., Section 6 of \[5\], also \[9\]). Here we will only summarize in two theorems the needed features of this theory of zero variance importance sampling applied to transport problems.

To this end, consider the problem of estimating the integral \((5)\), where \( \psi \) satisfies equation \((4)\). Equivalently, one can estimate \( I \) as

\[
I = \int_{\mathcal{R}} S(P) \psi^*(P) dP \tag{7}
\]

where

\[
\psi^*(P) = \int_{\mathcal{R}} K^*(P, P') \psi^*(P') dP' + g(P) \tag{8}
\]

is an integral equation dual to \((4)\). In these equations, the adjoint kernel \( K^* \) is defined as usual by \( K^*(P, P') = K(P', P) \). Then the theory (see \[5\] or \[9\]) tells us that \( \psi^* \) is an exact importance function for the problem defined by equations \((5)\) and \((4)\) while \( \psi \) is an exact (dual) importance function for the problem defined by equations \((7)\) and \((8)\). From this latter observation we expect that information about \( \psi \) should be useful in improving estimates of \( I \) as represented by \((7)\), a fact that we will use in applying adaptive importance sampling methods.

Referring back to equation \((6)\), we notice that if the basis functions \( B_i \) are chosen to be both orthogonal and normalized, then each coefficient \( a_i \) in the representation \((6)\) takes the form of equation \((5)\) where \( \psi \) satisfies equation \((4)\) and where the known function \( g \) appearing in \((3)\) has been identified as the \( i \)th basis function, \( B_i(P) \). Because of the duality mentioned above, we then have a choice between two equivalent methods of applying importance sampling to this problem. The first of these methods would represent each \( a_i \) as

\[
a_i = \int_{\mathcal{R}} B_i(P) \psi(P) dP
\]

and use information about the adjoint solution \( \psi^* \) as an importance function. However, in this framework, the adjoint equation \((8)\) becomes

\[
\psi^*_i(P) = \int_{\mathcal{R}} K^*(P, P') \psi^*_i(P') dP' + B_i(P). \tag{9}
\]
That is, a different importance function $\psi_i^*$ would be needed for $i$. On the other hand, if one estimates $a_i$ as equation (7), then source becomes $B_i$ and we seem to need to solve a different problem index $i$ again, except that in this case a different source is required. However, it is possible, in simulating equation (9), to make use of sampling methods, sampling initial points of random walks in such a way that each random walk can be used to estimate many different $a_i$'s, as can be found in [10]. This is the scheme we used since it is more efficient than having to identify and use as many different functions as there are indices $i$.

We recall ([9], Sections 2.5, 2.6, 2.7) that the random variational approach to estimating the integral $I$ may be defined as a transport problem:

A. **Terminal estimator**

$$\xi_T(\alpha) = \frac{g(P_k)}{\mu(P_k)}$$

where $p(P_k) = 1 - \kappa[P_k]$ = probability of termination in state $k$.

B. **Collision estimator**

$$\xi_C(\alpha) = \sum_{j=1}^{k} g(P_j);$$

C. **Track length estimator**

$$\xi_{TL}(\alpha) = \sum_{j=1}^{k} d_j g_j$$

where $d_j =$ length of $j$th track, $g_j =$ value of $g$ (assumed to be unity) on $j$th track. Here, $\alpha = (P_1, ..., P_k)$ is a typical random walk terminated as a result of absorption there.

A generalization of the track length estimator to include the case that the function $g$ varies along the $j$th track has proved to be very useful. This estimator, which is treated in [11], is called the generalized track length estimator and is defined by

D. **Generalized track length estimator**

$$\xi_{GTL}(\alpha) = \sum_{j=1}^{k} \int_{T_j} g(P) dP$$

where the integral is taken along the $j$th track.

A very important theoretical tool in achieving error reduction in Monte Carlo applications is the existence of zero variance sampling procedures. As we have seen, the zero variance scheme involves knowledge of the
of the dual problem; if available, this information could be used to produce the value $I$ without any random sampling merely by utilizing equation (7). Even though not realizable in practice, the zero variance theory nevertheless plays a crucial role in the design of adaptive procedures for steadily lowering sampling errors. The following result guarantees that a zero variance terminal estimator can be produced based upon the knowledge of an importance function $\psi^*$.

Theorem 1 Assuming that the importance function $\psi^*(P)$ satisfying

$$\psi^*(P) = \int K^*(P, Q)\psi^*(Q)dQ + g(P)$$

is known, where the kernel $K$ satisfies $K(P, Q) \geq 0$ for all $P, Q$ and where $K^*(P, Q) \equiv K(Q, P)$, the choices

$$\hat{S}(P) = \frac{\psi^*(P)S(P)}{\int \psi^*(P)S(P)dP}; \quad \hat{K}(P, Q) = \frac{K(P, Q)\psi^*(P)}{\psi^*(Q)}; \quad \hat{\beta}(Q) = \frac{\beta(Q)}{\psi^*(Q)}$$

(10)

for generating, respectively, the initial states, transitions from state $Q$ to state $P$, and termination at $Q$, produce $\text{Var}[\hat{\xi}_T] = 0$, where the random variable $\hat{\xi}_T$ is defined by

$$\hat{\xi}_T = \frac{g(P_n)K(P_n, P_{n-1}) \cdots K(P_2, P_1)S(P_1)}{\hat{\beta}(P_n)\hat{K}(P_n, P_{n-1}) \cdots \hat{K}(P_2, P_1)\hat{S}(P_1)}$$

(11)

The proof of this result, which may be found in [9], is really just a computation that confirms that the value of the transformed terminal estimator $\hat{\xi}_T$ equals $I$ for every random walk $\alpha$ generated according to (10).

A dual importance sampling theory can be applied to the adjoint problem formulation:

Estimate

$$I = \int \psi^*(P)S(P)dP$$

where

$$\psi^*(P) = \int K^*(P, Q)\psi^*(Q)dQ + g(P).$$

We make one additional assumption:

$$\int K^*(P, Q)dP = \int K(Q, P)dP = 1 - p^*(Q)$$

and $0 \leq p^* \leq 1$. Since $K^* \neq K$ in general, this condition cannot otherwise be assured.

Theorem 2 Assuming that the dual importance function $\psi(P)$ satisfying

$$\psi(P) = \int K(P, Q)\psi(Q)dQ + S(P),$$
is known, where the kernel $K$ satisfies $K(P, Q) \geq 0$ for all $P, Q$
\[
\hat{S}(P) = \int_{\mathcal{P}} \frac{\psi(P)g(P)}{\psi(Q)g(Q)P} \; dP ; \quad \hat{K}(P, Q) = \frac{K(Q, P)\psi(P)}{\psi(Q)} ; \quad \hat{p}
\]
for generating, respectively, the initial states, transitions from state $\hat{P}$, and termination at $\hat{Q}$, produce $\text{Var}[\xi^*_T] = 0$, where the rare event $\xi^*_T$ is defined by
\[
\xi^*_T = \frac{S(P_n) K(P_{n-1}, P_n) \cdots K(P_1, P_2) g(P_1)}{\hat{p}(P_n) \hat{K}(P_n, P_{n-1}) \cdots \hat{K}(P_2, P_1) \hat{S}(P_1)}
\]

The proof of this dual result is also just a computation.

4 Steady-State Transport

For the sake of illustration, we apply our ideas to the steady-state equation. We will subsequently specialize to treat the case of planar transport.

The general steady-state transport equation can be written
\[
\mathbf{\Omega} \cdot \nabla \phi + \sigma \phi = \int \int \sigma' f \phi' d\Omega' dE' + Q
\]
\[
= \int q
\]

where
\[
\mathbf{\Omega} = \text{unit direction vector},
\phi = \phi(r, \mathbf{\Omega}, E) = \text{angular flux}
\sigma = \sigma(r, E) = \text{total macroscopic cross section}
\sigma' f = \sigma(r, E') f(r; \Omega', E' \rightarrow \mathbf{\Omega}, E) \text{ describes scattering and}
\int \int f(r; \Omega', E' \rightarrow \mathbf{\Omega}, E) d\Omega' dE' = c(r, E')
\]
\[
= \text{mean number of secondaries per primary particle}
\phi' = \phi(r, \Omega', E')
Q = Q(r, \mathbf{\Omega}, E) = (\text{internal}) \text{ source}
\]
\[
q = q(r, \mathbf{\Omega}, E)
\]
\[
= \int \int \sigma(r, E') f(r; \Omega', E' \rightarrow \mathbf{\Omega}, E) \phi(r, \Omega', E') d\Omega' dE' + Q(r, t)
\]

We define
\[
\varphi(r, E) \equiv \int_{4\pi} \phi(r, \mathbf{\Omega}, E) d\Omega = \text{scalar flux}.
\]

From Equation (14) it is easy to convert to an integral equation for the angular flux by using an integrating factor based on the observable spatial gradient term is a directional derivative. The result is
\[
\phi(r, \mathbf{\Omega}, E) = \int_0^\infty \exp \left[ - \int_0^{s'} \sigma(r - s'\mathbf{\Omega}, E) ds' \right] q(r - s'\mathbf{\Omega}, \mathbf{\Omega}, E)
\]
If scattering is now assumed to be isotropic,

\[ f(r; \Omega', E' \rightarrow \Omega, E) = \frac{1}{4\pi} f(r; E' \rightarrow E) \]  

(19)

we can write

\[ q(r, \Omega, E) = \frac{1}{4\pi} \int \sigma(r; E' \rightarrow E) \phi(r, E') dE' + Q(r, \Omega, E) \]

(20)

where

\[ \sigma(r; E' \rightarrow E) \equiv \sigma(r, E') f(r; E' \rightarrow E) \].

(21)

From equation (20) we see that if the scalar flux \( \phi(r, E') \) is known, then \( \phi(r, \Omega, E) \) can be determined by substituting (20) into (18) and carrying out the integration. We will make use of this observation shortly.

5 Plane Geometry

We now assume no energy dependence, no internal source \( Q \equiv 0 \), isotropic scattering, and that \( \phi \) depends only on \( x, \mu \). The integro-differential equation (14) then becomes

\[ \mu \frac{\partial \phi}{\partial x} + \sigma \phi = \frac{\epsilon}{2} \int_{-1}^{1} \phi(x, \mu') d\mu' \]

\[ = \frac{\epsilon}{2} \phi(x) \]

(22)

or, equivalently (provided \( \mu \neq 0 \))

\[ \phi(x, \mu) = \begin{cases} 
\exp(-\sigma x/\mu) \psi(0, \mu) + \int_{0}^{x} \exp\left[\left(\sigma/\mu\right)(s - x)\right] \left[\frac{\epsilon}{2} \phi(s)\right] ds, & \mu > 0, \\
\exp(\sigma(T - x)/\mu) \psi(T, \mu) + \int_{x}^{T} \exp\left[\left(\sigma/\mu\right)(s - x)\right] \left[\frac{\epsilon}{2} \phi(s)\right] ds, & \mu < 0
\end{cases} \]

(23)

We applied Theorem 2 adaptively to a series of model transport problems in planar geometry, using an approximation to the (dual) importance function \( \phi \) from stage \( k \) to improve this approximation in stage \( k + 1 \). The algorithm used, and details of this computation, are described in [10].

The first problems, reported in [10], involved monodirectional transport (i.e., problems in which only a single, fixed scattering direction \( \mu_0 \) is allowed, so the flux \( \phi \) is a function of a single spatial variable \( x \)). Thus we wrote

\[ \phi(x) \approx \sum_{i=0}^{n-1} a_i P_i\left(\frac{2x}{T} - 1\right), \quad 0 \leq x \leq T \]

where \( P_i \) denotes the \( i \)th Legendre polynomial and we estimated the coefficients \( a_i \) adaptively, using the ideas described above. Excellent results were obtained, and geometric convergence was achieved making possible very accurate global solutions. Even so, we observed that the convergence characteristics were degraded as the fixed scattering direction approached \( \theta_0 = \pi/2 \).
\( (\mu_0 = 0) \). The difficulties were traced to the need to use very large
of Legendre polynomials to represent the solution adequately ev-
the problem [12].

When we next attempted to use a double Legendre expansion
both the \( x \) and the \( \mu \) variation of the solution \( \phi(x, \mu) \) of equation
serious difficulties were encountered. Some of these difficulties are
in Figure 1, which represents the converged solution after 40 se-
related sampling stages, each consisting of \( W = 100,000 \) random
this computation, nine Legendre components were used for each
total of 81 in all. It is clear that the waviness near \( \mu = 0 \) is an
the use of a polynomial approximation. This waviness lessened b
plainly observable as we increased the number of Legendre compo-
for each variable. In addition to this waviness, the fact that the
well-behaved considered as a function of \( \mu \) at boundaries and/or
or at values of \( x \) that correspond to source locations, means that
Legendre polynomials (in \( \mu \)) will only be able to capture this si-
behavior if the number of terms is actually infinite. This suggests that
be desirable to formulate a method that avoids the Legendre ex-
\( \mu \) completely. Such a method is available, using the derivations
in the preceding section. There we showed that the angular flux
constructed from knowledge of the scalar flux alone by making
transport equation itself.

To test the effectiveness of this idea, assume that \( \phi(0, \mu) \) is give
the scalar flux is represented by

\[
\varphi(s) \approx \sum_{i=0}^{n-1} a_i P_i \left( \frac{2s}{T} - 1 \right), \quad 0 \leq s \leq T
\]

where \( a_i \) are spatial Legendre coefficients determined from a previ-
te stage, for example. Substituting (24) into (23) and carrying o-
tegration produces the next stage approximation to \( \phi(x, \mu) \) to use
reduce the source (in a sequential correlated sampling algorithm) or
an approximate (dual) importance function (to use in an adap-
tance sampling algorithm). The potential advantage of this techni-
the exponential factor \( \exp(\pm \sigma x / \mu) \) present in the solution (23) is
analytically rather than by means of a second expansion in \( \mu \). For
a modest extension of this idea works for anisotropic scattering as
omit the details of this extension.

The representation portrayed in Figure 2, based on an expansion
scalar flux as a sum of nine Legendre polynomials in \( x \) alone, was
after 40 sequential correlated sampling stages each consisting of
number \( W = 100,000 \) of random walks per stage as used to obta
1. It is quite evident that the resulting solution is much better beh-
when a double Legendre expansion was used, and this solution also
less work. Using the transport equation itself as a constraint on th
Fig. 1. Global flux based on double Legendre expansion, 9x9 coefficients

Fig. 2. Global flux based on single Legendre expansion, 9 coefficients
the solution has obviously improved our ability to capture the global flux with a minimum of effort. More evidence of the inherent precision of this method can be seen in Figure 3, in which the global flux is plotted over the entire $-1 \leq \mu \leq 1$ range, rather than just the half-range $0 \leq \mu$ of Figures 1 and 2. The nature of the singularity along the $\mu = 0$ line is clearly evident, and easily seen, in this graph.

Although Figures 1-3 were obtained using sequential correlated sampling, similar results were obtained using adaptive importance sampling. A study [13] in which the two methods were compared for a family of one-dimensional slab transport problems revealed consistently superior convergence rates for importance sampling when compared with correlated sampling. This is consistent with the results found when comparing the two algorithms. A family of discrete ordinates transport problems, as reported in [5] and earlier work.

6 Determining Global Scalar Fluxes

Rarely is the full angular flux $\phi(r, \Omega, E)$ required in nuclear design; other settings. More frequently, the scalar flux $\varphi(r, E)$ is sought.
For example, reaction rates are expressible as weighted integrals of the scalar flux:

\[ I = \int \int \Sigma(r, E) \phi(r, \Omega, E) dr dE d\Omega = \int \int \Sigma(r, E) \phi(r, E) dr dE. \]

In this case, further simplifications can be made.

We reconsider the integro-differential equation (22) again, setting \( Q = 0 \) (no internal source) as before. We obtain

\[
\mu \frac{\partial \phi}{\partial z} + \sigma \phi = \frac{\sigma}{\mu} \int_{-\mu}^{\mu} \phi(x, \mu') d\mu', \quad -1 \leq \mu \leq 1; \quad 0 < x < T \\
\phi(0, \mu) = Q_0(\mu), \quad 0 < \mu \leq 1, \\
\phi(T, \mu) = Q_T(\mu) = 0, \quad -1 \leq \mu < 0, \tag{25}
\]

Rewriting (25) (\( \mu \neq 0 \)), and integrating from \( x = 0 \) to \( x = x \) for \( \mu > 0 \) and from \( x = x \) to \( x = T \) for \( \mu < 0 \) gives

\[
\phi(x, \mu) = \begin{cases} 
\frac{\sigma}{\mu} \int_0^x e^{-\frac{\sigma}{\mu}(x-y)} dy \int_{-\mu}^{\mu} \phi(y, \mu') d\mu', & \mu > 0, \\
\frac{\sigma}{\mu} \int_T^x e^{-\frac{\sigma}{\mu}(x-y)} dy \int_{-\mu}^{\mu} \phi(y, \mu') d\mu', & \mu < 0.
\end{cases} \tag{26}
\]

Integrating \( \phi(x, \mu) \) from \( \mu = -1 \) to \( \mu = 1 \) then produces

\[
\varphi(x) = \int_0^T K(x, y) \varphi(y) dy + S(x), \tag{27}
\]

where

\[
\varphi(x) = \int_{-1}^{1} \phi(x, \mu) d\mu, \\
S(x) = \int_0^1 e^{-\frac{\sigma}{\mu}x} Q_0(\mu) d\mu \\
K(x, y) = \frac{\sigma}{2} \int_0^1 e^{-\frac{\sigma}{\mu}x - \frac{\sigma}{\mu}y} \frac{e^{-\frac{\sigma}{\mu}y}}{\mu} d\mu.
\]

To solve (27) using Monte Carlo methods, we can make use of the symmetry of the kernel \( K(x, y) = K(y, x) \), and the fact that for any \( y \),

\[
\int_{-\infty}^{+\infty} K(x, y) dx = \frac{\sigma}{2} \int_{-\infty}^{+\infty} dx \int_0^1 e^{-\frac{\sigma}{\mu}x - \frac{\sigma}{\mu}y} \frac{e^{-\frac{\sigma}{\mu}y}}{\mu} d\mu \\
= \sigma a.
\]

Now we can apply the methods described above to estimate a Legendre expansion of the scalar flux by adaptive Monte Carlo methods. After solving (27) we can again obtain the angular flux \( \phi(x, \mu) \) using (26).
We used this method for finding the scalar flux to solve a series of transport problems in which the amount of scattering was varied. The results are shown in Figure 4. The scalar flux at $x = 0.9$ for three different scattering rates is graphed in Figure 4. Once again, sequential correlated sampling was used to obtain these results, based on 30 sequential stages each consisting of $W = 200$ random walks. For each problem, a unit semisotropic source at $x = 0$ introduces particles into a 1 mean free path thick slab. The Neumann series that produces the solution of the integral equation is used, and the amount of scattering is set to $0.1$ and $0.9$ for the three problems. We see that the rates of convergence of the Neumann series improve as the amount of scattering decreases (amount of absorption increases). This is not surprising since fewer random walks are needed when the amount of absorption is high.

Clear evidence of geometric convergence is obtained in all cases.

7 Conclusions, Future Work

This paper has demonstrated that learning algorithms can be devised either on the use of correlated sampling or importance sampling, or both, for obtaining very accurate global solutions of model transport problems. The algorithms can learn successfully enough to achieve geometric convergence to a theoretically zero variance solution. Making use of the
equation itself as a constraint on the global solution, two dimensional \((x,\mu)\) calculations in planar geometry have been successfully performed.

If the scalar flux is all that is required, the complexity of the problem can be further reduced by applying the same algorithmic ideas to the integral equation for the scalar flux itself.

The use of non-polynomial basis functions is clearly indicated by the qualitative and quantitative issues encountered when very smooth basis functions are used to represent the flux. Indeed, the use of locally defined basis functions for the flux representation suggests itself, making the use of wavelet bases potentially very attractive.

An error analysis for these adaptive methods is needed to gain further knowledge of how to set parameters in running adaptive algorithms for systematic improvement of the results. Initial results along this line can be found in [8].

Finally, all of the results reported in this paper have been based on the use of pseudorandom sequences exclusively to generate the needed random walks. It seems clear that convergence characteristics will be improved by the selective substitution of low discrepancy sequences (or hybrid sequences) for pseudorandom ones.

Much more work is needed, and is anticipated, to further our understanding of these promising beginnings.

References

113


13. Hayakawa, C., Spanier, J.: Comparison of Monte Carlo Algorithms showing Geometric Convergence for Model Transport Problems. this vol