LA-UR- 98- " 4755

Approved for public release;
distribution is unlimited.

RECEIVED

AUG 18 1999

Title: ADAPTIVELY LEARNING AN IMPORTANCE FUNCTION
USING TRANSPORT CONSTRAINED MONTE CARLO (U)

Author(s): Thomas E. Booth

Submitted to: Proceedings of the Third International Conference on Monte Carlo
and Quasi-Monte Carlo, Claremont, CA, June 22-26, 1998

Los Alamos

NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Adaptively Learning an Importance Function Using Transport Constrained Monte Carlo

Thomas E. Booth

Los Alamos National Laboratory, Code Integration Group XCI, Mail Stop F663,
Los Alamos, New Mexico, 87545 USA email teb@lanl.gov

Abstract. It is well known that a Monte Carlo estimate can be obtained with zero-variance if an exact importance function for the estimate is known. There are many ways that one might iteratively seek to obtain an ever more exact importance function. This paper describes a method that has obtained ever more exact importance functions that empirically produce an error that is dropping exponentially with computer time. The method described herein constrains the importance function to satisfy the (adjoint) Boltzmann transport equation. This constraint is provided by using the known form of the solution, usually referred to as the Case eigenfunction solution.

1 Introduction

The particle transport portion of the *Monte Carlo Methods for the 21st Century* project seeks to obtain a statistical error that decreases exponentially with computer time rather than the current $1/\sqrt{N}$ dictated by the central limit theorem of statistics. We break free of the central limit theorem by adaptively biasing the Monte Carlo calculation based on what has been learned from previous samples. Thus, the samples are no longer independent and identically distributed and the central limit theorem does not apply. Specifically, the convergence rate is no longer constrained to $1/\sqrt{N}$.

It is well known [1,2] that zero variance solutions can be obtained when the sampled probability for any event is proportional to its natural probability times the expected score subsequently produced given the event is sampled. In particle transport terminology, the expected score function is called the importance function. A learning process is introduced whereby an importance function that is estimated at the n -1st iteration is used to bias the n -th iteration's estimate of the importance function. This iterative process produces an estimated importance function that is rapidly converging to the true importance function.

2 Problem Definitions

Definition 1. $\sigma = 1$ is the macroscopic total cross section (i.e., collision probability per unit distance)

Definition 2. c = survival probability upon collision

Definition 3. θ = the particle's angle with respect to the x axis.

Definition 4. μ = the particle's direction cosine with respect to the x axis.

Definition 5. x = spatial position of the particle

Definition 6. $\Phi(x, \mu)$ = the angular flux

Definition 7. $\Phi^\dagger(x, \mu)$ = the importance, or expected score from a particle at (x, μ) .

Definition 8. T = the slab thickness in mean free paths

Definition 9. $\delta(x)$ = the Dirac delta function

Definition 10. δ_{mn} = the Kronecker delta function

This paper considers one-speed particle transport in slab geometry. That is, a particle may be anywhere in the spatial region described by $0 \leq x \leq T$, $-\infty < y < \infty$, and $-\infty < z < \infty$. Additionally, the particle has a direction cosine, $\mu = \cos(\theta)$, with respect to the x axis. The particle source is at $x = 0$ and is uniform in $\mu \geq 0$. That is,

$$Q(\mu) = 1 \text{ for } \mu \geq 0 \text{ and } 0 \text{ otherwise} \quad (1)$$

The desired estimate is the fraction of source particles that exit the slab at $x = T$. (Note that once the particles exit the slab, either at $x = 0$ or at $x = T$, they never re-enter the slab.)

3 Particle Transport Basics

The time-independent Boltzmann transport equation for the particle flux, $\Phi(x, \mu)$, for this problem is (see Eq. 2.5 of [3] and note that the source is zero except at $x=0$)

$$\mu \frac{\partial \Phi(x, \mu)}{\partial x} + \Phi(x, \mu) = \frac{c}{2} \int_{-1}^1 \Phi(x, \mu') d\mu' \quad (2)$$

The time-independent equation for the importance is

$$-\mu \frac{\partial \Phi^\dagger(x, \mu)}{\partial x} + \Phi^\dagger(x, \mu) = \frac{c}{2} \int_{-1}^1 \Phi^\dagger(x, \mu') d\mu' \quad (3)$$

Note that if $\Phi(x, \mu)$ satisfies Eq. 2 then

$$\Phi^\dagger(x, -\mu) = \Phi(x, \mu) \quad (4)$$

satisfies Eq. 3 and vice-versa.

The form of the solution to Eqs. 2 and 3 is given by Eq. 23 of [4] as

$$\begin{aligned}\Phi^\dagger(x, -\mu) = \Phi(x, \mu) = & a_{+L}\psi_0^+(\mu)e^{-x/\nu_0} + a_{-R}\psi_0^-(\mu)e^{-(T-x)/\nu_0} \\ & + \int_{-1}^0 A_R(\nu)\psi_\nu(\mu)e^{(T-x)/\nu}d\nu + \int_0^1 A_L(\nu)\psi_\nu(\mu)e^{-x/\nu}d\nu\end{aligned}\quad (5)$$

where a_{+L} , a_{-R} , $A_L(\nu)$, and $A_R(\nu)$ are coefficients arising from a separation of variables solution of Eq. 2, $\psi_0^+(\mu)$, $\psi_0^-(\mu)$, and $\psi_\nu(\mu)$ are the orthogonal eigenfunctions, and N_0^+ , N_0^- , and N_ν are (Eqs. 11-13 [4]).

$$N_0^\pm = \int_{-1}^1 \mu\psi_0^\pm(\mu)\psi_0^\pm(\mu)d\mu = \pm \frac{c}{2}\nu_0^3 \left[\frac{c}{\nu_0^2 - 1} - \frac{1}{\nu_0^2} \right] \quad (6)$$

$$\int_{-1}^1 \mu\psi_{\nu'}(\mu)\psi_\nu(\mu)d\mu = N_\nu\delta(\nu - \nu') \quad (7)$$

with

$$N_\nu = \nu \left[\lambda^2(\nu) + \frac{\pi^2 c^2}{4} \nu^2 \right] \quad (8)$$

From [3] pages 70 and 72 note that ν_0 is the positive root of

$$1 = c\nu_0 \tanh^{-1} \frac{1}{\nu_0} \equiv \frac{c\nu_0}{2} \ln \frac{\nu_0 + 1}{\nu_0 - 1}, \quad (9)$$

and

$$\nu_0 > 1 \quad \text{for } 0 < c < 1 \quad (10)$$

With the importance written in the form of Eq. 5, it is now possible to get the importance $\Phi^\dagger(x, -\mu)$ for all x by estimating integrals that involve only the boundary importance.

The $A_L(\nu)$ and $A_R(\nu)$ are functions of the flux/importance at the left and right boundaries given by Eqs. 28 and 48 of [4]. The left and right boundary importances (for particles going into the slab) can be expanded in a double Legendre (P_m) expansion with $(b_0, b_1, b_2 \dots)$ and $(c_0, c_1, c_2, \dots)$ given by

$$b_m = (2m+1) \int_{-1}^0 P_m(-1-2\mu)\Phi^\dagger(0, -\mu)d\mu \quad (11)$$

$$c_m = (2m+1) \int_0^1 P_m(-1+2\mu)\Phi^\dagger(T, -\mu)d\mu, \quad (12)$$

(Note that the $A_L(\nu)$ and $A_R(\nu)$ components for particles exiting the slab can be determined directly (without Monte Carlo) from the boundary conditions; see sections 4.4 and 4.6. From Eq. 14 of [4] and Eq. 4, with subscript L indicating an expansion at the left boundary,

$$a_{+L} = \frac{1}{N_0^+} \int_{-1}^1 \Phi^\dagger(0, -\mu)\mu\psi_0^+(\mu)d\mu \quad (13)$$

Note that $\Phi(0, \mu)$ in Eq. 15 of [4] is the flux at the expansion point $x = 0$. If an expansion is done at the right boundary, then $\Phi(T, \mu)$ takes the place of $\Phi(0, \mu)$. (This can be viewed as a coordinate shift, $x' = x - T$, in which the new $x' = 0$ is at the right boundary.) Using Eq. 15 of [4] and Eq. 4, with subscript R indicating an expansion at the right boundary,

$$a_{-R} = \frac{1}{N_0^-} \int_{-1}^1 \Phi^\dagger(T, -\mu) \mu \psi_0^-(\mu) d\mu \quad (14)$$

The importance can thus be written as

$$\Phi^\dagger(x, -\mu) = \Phi(x, \mu) = F(a_{+L}, a_{-R}, b_0, b_1, b_2, \dots, c_0, c_1, c_2, \dots) \quad (15)$$

with the arguments of F all integrals that can be estimated by Monte Carlo.

4 Zero Variance Estimation of a_+ , a_- , b_m , and c_m

A crucial element of an iterative zero variance biasing procedure is that if the exact importance function is *input* for the biasing on the i -th iteration, then the i -th iteration must *output* the exact importance function as input to the $i+1$ st iteration. If this does not happen, then there can be no convergence to the exact importance function necessary for zero variance biasing. Note that the importance function is determined by a_+ , a_- , b_m , and c_m , so that these coefficients must be known exactly (i.e. with zero variance) for the importance function to be known exactly. This section will show how to estimate all the coefficients simultaneously using the same set of random walks together with rejection techniques.

4.1 Zero Variance Estimation of a_{+L}

Note that $\Phi^\dagger(0, -\mu) = 0$ for $\mu > 0$ because any particle that is moving backward at $x = 0$ leaves the geometry forever with no score. Hence, Eq. 13 becomes

$$a_{+L} = \frac{1}{N_0^+} \int_{-1}^0 \Phi^\dagger(0, -\mu) \mu \psi_0^+(\mu) d\mu \quad (16)$$

This integral can be estimated by sampling μ_i (N times) from

$$p(\mu) = 1 \quad \text{for } -1 \leq \mu \leq 0 \quad (17)$$

and tallying $\Phi^\dagger(0, -\mu_i) \mu_i \psi_0^+(\mu_i)$ so that

$$a_{+L} \approx \frac{1}{N_0^+} \frac{1}{N} \sum_{i=1}^N \Phi^\dagger(0, -\mu_i) \mu_i \psi_0^+(\mu_i) \quad (18)$$

The only trouble with Eq. 18 is that $\Phi^\dagger(0, -\mu_i)$ is unknown. This can be handled by starting a random walk at μ_i and making a one particle estimate

of $\Phi^\dagger(0, -\mu_i)$. This one particle estimate, $\Phi_1^\dagger(0, -\mu_i)$ is then inserted in Eq. 18, in place of $\Phi^\dagger(0, -\mu_i)$ yielding:

$$a_{+L} \approx \frac{1}{N_0^+} \frac{1}{N} \sum_{i=1}^N \Phi_1^\dagger(0, -\mu_i) \mu_i \psi_0^+(\mu_i) \quad (19)$$

To completely define the estimate, note from Eq. 7 of [4] that

$$\psi_0^+(\mu) = \frac{c}{2} \frac{\nu_0}{\nu_0 - \mu} > 0 \quad (20)$$

Now consider using a biased pdf, $b_+(\mu)$, instead of $p(\mu)$ in Eq. 17. The biased pdf will make use of the estimated importance function. Let $\Phi_e^\dagger(0, -\mu)$ be an estimate of $\Phi^\dagger(0, -\mu)$. Note that

$$\Phi_e^\dagger(0, -\mu) (-\mu \psi_0^+(\mu)) \geq 0 \quad \text{for } -\mu \geq 0 \quad (21)$$

as long as

$$\Phi_e^\dagger(0, -\mu) \geq 0 \quad (22)$$

Eq. 22 can be enforced, if necessary, by increasing the constant b_0 until the estimated importance is nonnegative everywhere. Instead of $p(\mu)$ sample from

$$b_+(\mu) = \frac{\Phi_e^\dagger(0, -\mu) |\mu| \psi_0^+(\mu)}{-\int_{-1}^0 \Phi_e^\dagger(0, -\mu) \mu \psi_0^+(\mu) d\mu} \quad (23)$$

and weight the sample by

$$w_+(\mu) = \frac{p(\mu)}{b_+(\mu)} = \frac{\int_{-1}^0 \Phi_e^\dagger(0, -\mu) \mu \psi_0^+(\mu) d\mu}{\Phi_e^\dagger(0, -\mu) \mu \psi_0^+(\mu)} \quad (24)$$

The integral in Eqs. 23 and 24 can be obtained from the estimated importance expansion. Let

$$U = \int_{-1}^0 \Phi_e^\dagger(0, -\mu) \mu \psi_0^+(\mu) d\mu \quad (25)$$

Using Eqs. 30 and 32 of [4] and Eqs. 4 and 20 herein, Eq. 25 becomes

$$U = \sum_{m=0}^{\infty} b_m \int_{-1}^0 P_m(-1-2\mu) \mu \frac{c}{2} \frac{\nu_0}{\nu_0 - \mu} d\mu \quad (26)$$

Noting that

$$\frac{\mu}{\nu_0 - \mu} = -1 + \frac{\nu}{\nu_0 - \mu} \quad (27)$$

Eq. 26 becomes

$$U = \sum_{m=0}^{\infty} b_m \frac{c\nu_0}{2} \left[-\int_{-1}^0 P_m(-1-2\mu) du + \nu_0 \int_{-1}^0 \frac{P_m(-1-2\mu)}{\nu_0 - \mu} d\mu \right] \quad (28)$$

Recalling Eq. 36 of [4]

$$J_m(-1-2\nu) = \int_{-1}^0 \frac{P_m(-1-2\mu)}{\nu-\mu} d\mu \quad (29)$$

and changing variables in the first integral, Eq. 28 becomes

$$U = \sum_{m=0}^{\infty} b_m \frac{c\nu_0}{2} \left[-\frac{1}{2} \int_{-1}^1 P_m(y) dy + \nu_0 J_m(-1-2\nu_0) \right] \quad (30)$$

Noting that the J_m are related to the Legendre functions of the second kind, Q_m , (for $\nu > 0$) as [5]

$$\begin{aligned} J_m(-1-2\nu) &= \int_{-1}^0 P_m(-1-2\mu) \frac{d\mu}{\nu-\mu} = \int_{-1}^1 P_m(y) \frac{dy}{y+2\nu+1} \\ &= (-1)^m \int_{-1}^1 P_m(x) \frac{dx}{2\nu+1-x} = (-1)^m 2Q_m(2\nu+1) \end{aligned} \quad (31)$$

and using the orthogonality of the P_m , Eq. 30 becomes

$$U = \sum_{m=0}^{\infty} b_m \frac{c\nu_0}{2} [-\delta_{m0} + \nu_0 (-1)^m 2Q_m(2\nu_0+1)] \quad (32)$$

With U determined by Eq. 32, then $b_+(\mu)$ is determined by Eq. 23, and can thus be sampled.

A rejection scheme is used to sample μ from $b_+(\mu)$. First sample μ from

$$\frac{\Phi_e^\dagger(0, -\mu)}{\int_{-1}^0 \Phi_e^\dagger(0, -\mu) du} \quad (33)$$

For the rejection part, note two things from Eq. 20

$$-\mu\psi_0^+(\mu) \geq 0 \quad \text{for } -1 \leq \mu \leq 0 \quad (34)$$

and

$$\max_{-1 \leq \mu \leq 0} (\mu\psi_0^+(\mu)) = \psi_0^+(-1) \quad (35)$$

thus μ is accepted with probability

$$\frac{-\mu\psi_0^+(\mu)}{\psi_0^+(-1)} = -\mu \frac{\nu_0+1}{\nu_0-\mu} \quad (36)$$

The estimate of a_{+L} is then

$$a_{+L} \approx \frac{1}{N_0^+} \frac{1}{N} \sum_{i=1}^N w_+(\mu_i) \Phi_1^\dagger(0, -\mu_i) \mu_i \psi_0^+(\mu_i) \quad (37)$$

4.2 Zero Variance Estimation of a_{-R}

The estimation of a_{-R} is very similar to that for a_{+L} . At the boundary, $\Phi^\dagger(T, -\mu)$ is known for $-\mu > 0$ and Eq. 14 becomes

$$a_{-R} = \frac{1}{N_0^-} \int_{-1}^0 \Phi^\dagger(T, -\mu) \mu \psi_0^-(\mu) d\mu + \frac{1}{N_0^-} \int_0^1 \Phi^\dagger(T, -\mu) \mu \psi_0^-(\mu) d\mu \quad (38)$$

The first integral can be integrated directly without Monte Carlo because the importance function is known from the boundary condition; only the second integral needs to be estimated. Let

$$\alpha = \frac{1}{N_0^-} \frac{1}{N_0^-} \int_0^1 \Phi^\dagger(T, -\mu) \mu \psi_0^-(\mu) d\mu \quad (39)$$

This integral can be estimated by sampling μ_i from

$$p(\mu) = 1 \quad \text{for } 0 \leq \mu \leq 1 \quad (40)$$

and tallying $\Phi^\dagger(T, -\mu_i) \mu_i \psi_0^-(\mu_i)$ so that

$$\alpha \approx \frac{1}{N_0^-} \frac{1}{N} \sum_{i=1}^N \Phi^\dagger(T, -\mu_i) \mu_i \psi_0^-(\mu_i) \quad (41)$$

The only trouble with Eq. 41 is that $\Phi^\dagger(T, -\mu_i)$ is unknown. This can be handled by starting a random walk at μ_i and making a one particle estimate of $\Phi^\dagger(T, -\mu_i)$. This one particle estimate, $\Phi_1^\dagger(T, -\mu_i)$ is then inserted in Eq. 41, in place of $\Phi^\dagger(T, -\mu_i)$ yielding:

$$\alpha \approx \frac{1}{N_0^-} \frac{1}{N} \sum_{i=1}^N \Phi_1^\dagger(T, -\mu_i) \mu_i \psi_0^-(\mu_i) \quad (42)$$

To completely define the estimate, note from Eq. 7 of [4] that

$$\psi_0^-(\mu) = \frac{c}{2} \frac{\nu_0}{\nu_0 + \mu} > 0 \quad (43)$$

Note that

$$\Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu) \geq 0 \quad (44)$$

as long as

$$\Phi_e^\dagger(0, -\mu) \geq 0 \quad (45)$$

Equation 45 can be enforced, if necessary, by increasing the constant c_0 until the estimated importance is nonnegative everywhere. Instead of sampling from $p(\mu)$, sample from

$$b_-(\mu) = \frac{\Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu)}{\int_0^1 \Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu) d\mu} \quad (46)$$

and weight the sample by

$$w_-(\mu) = \frac{p(\mu)}{b_-(\mu)} = \frac{\int_0^1 \Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu) d\mu}{\Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu)} \quad (47)$$

The integral in Eqs. 46 and 47 and can be obtained from the estimated importance expansion. Let

$$V = \int_0^1 \Phi_e^\dagger(0, -\mu) \mu \psi_0^-(\mu) d\mu \quad (48)$$

Using Eqs. 50 and 52 of [4] and Eqs. 4 and 43 herein, Eq. 48 becomes

$$V = \sum_{m=0}^{\infty} c_m \int_0^1 P_m(-1+2\mu) \mu \frac{c}{2} \frac{\nu_0}{\nu_0 + \mu} d\mu \quad (49)$$

Noting that

$$\frac{\mu}{\nu_0 + \mu} = 1 - \frac{-\nu_0}{-\nu_0 - \mu} \quad (50)$$

Eq. becomes

$$V = \sum_{m=0}^{\infty} c_m \frac{c\nu_0}{2} \left[\int_0^1 P_m(-1+2\mu) d\mu + \nu_0 \int_0^1 \frac{P_m(-1+2\mu)}{-\nu_0 - \mu} d\mu \right] \quad (51)$$

Recalling Eq. 56 of [4]

$$L_m(-1+2\nu) = \int_0^1 \frac{P_m(-1+2\mu)}{\nu - \mu} d\mu \quad (52)$$

and changing variables in the first integral, Eq. becomes

$$V = \sum_{m=0}^{\infty} c_m \frac{c\nu_0}{2} \left[\frac{1}{2} \int_{-1}^1 P_m(y) dy + \nu_0 L_m(-1-2\nu_0) \right] \quad (53)$$

From Eqs. 29 and 52 note that $L_m(-1+2\nu) = -J_m(-1+2\nu)$, then using Eqs. 31 and 52, and the orthogonality of the P_m .

$$V = \sum_{m=0}^{\infty} c_m \frac{c\nu_0}{2} [\delta_{m0} + \nu_0 (-1)^{m+1} 2Q_m(2\nu_0+1)] \quad (54)$$

A rejection scheme is used to sample μ from $b_-(\mu)$. First sample μ from

$$\frac{\Phi_e^\dagger(0, -\mu)}{\int_0^1 \Phi_e^\dagger(0, -\mu) d\mu}, \quad (55)$$

For the rejection part, note two things

$$\mu \psi_0^-(\mu) \geq 0 \quad \text{for } 0 \leq \mu \leq 1 \quad (56)$$

and

$$\max_{0 \leq \mu \leq 1} (\mu \psi_0^-(\mu)) = \psi_0^-(1) \quad (57)$$

thus μ is accepted with probability

$$\frac{-\mu \psi_0^-(\mu)}{\psi_0^-(1)} = \mu \frac{\nu_0 + 1}{\nu_0 + \mu} \quad (58)$$

The estimate of α is then

$$\alpha \approx \frac{1}{N_0^-} \frac{1}{N} \sum_{i=1}^N w_-(\mu_i) \Phi_1^\dagger(0, -\mu_i) \mu_i \psi_0^-(\mu_i) \quad (59)$$

Hence, from Eqs. 38, 39, and 59

$$a_{-R} \approx \frac{1}{N_0^-} \int_{-1}^0 \Phi^\dagger(T, -\mu) \mu \psi_0^-(\mu) d\mu + \frac{1}{N_0^-} \frac{1}{N} \sum_{i=1}^N w_-(\mu_i) \Phi_1^\dagger(0, -\mu_i) \mu_i \psi_0^-(\mu_i) \quad (60)$$

4.3 Estimating the b_m

This section shows how all the b_m can be estimated simultaneously with the same set of starting random walks at the left boundary used to estimate a_{+L} and a_{-R} . The basic technique is to use a rejection technique to importance weight the sampling of the particles' starting μ_i 's.

Because the estimates of the b_m do not involve $\psi_0^\pm(\mu)$ or $\psi_\nu(\mu)$, it is convenient to simplify the notation. At the left boundary

$$y = -1 - 2\mu \quad (61)$$

Rewriting all the importance functions as functions of y . The true importance is,

$$\eta(y) \doteq \Phi^\dagger(0, (y+1)/2) = \Phi^\dagger(0, -\mu) \quad (62)$$

The estimated importance from the previous iteration is,

$$\eta_e(y) \doteq \Phi_e^\dagger(0, (y+1)/2) = \Phi_e^\dagger(0, -\mu) \quad (63)$$

The one particle estimate of the importance of a particle started at y is

$$\eta_1(y) \doteq \Phi_1^\dagger(0, (y+1)/2) = \Phi_1^\dagger(0, -\mu) \quad (64)$$

A zero variance estimation of

$$b_m = \frac{2m+1}{2} \int_{-1}^1 P_m(y) \eta(y) dy \quad (65)$$

for $m > 0$ is problematical because the integrand in Eq. 65 is not always positive. To get around this problem define

$$b'_m \doteq (2m+1) \int_{-1}^1 \frac{1+P_m(y)}{2} \eta(y) dy \quad (66)$$

so that by Eqs. 65 and 66

$$b'_m = b_m + (2m+1)b_0 \quad (67)$$

Note that because $|P_m(y)| \leq 1$ and $\int_{-1}^1 P_m(y) dy = 0$ for $m > 0$

$$g_m(y) = \frac{1+P_m(y)}{2} \quad \text{for } -1 \leq y \leq 1 \quad \text{and } m > 0 \quad (68)$$

is a pdf for $m > 0$. Thus b'_m could be estimated by sampling $y_i = -1 - 2\mu_i$ from $g_m(y)$ and then

$$b'_m \approx (2m+1) \frac{1}{N} \sum_{i=1}^N \eta(y_i) \quad (69)$$

A zero variance estimate of b'_m requires sampling from an importance weighted pdf; that is,

$$d_m(y) = C_m \frac{1+P_m(y)}{2} \eta_e(y) \quad \text{for } -1 \leq y \leq 1 \quad \text{and } m > 0 \quad (70)$$

with C_m being the normalization constant

$$C_m^{-1} = \int_{-1}^1 \frac{1+P_m(y)}{2} \eta_e(y) dy \quad \text{for } -1 \leq y \leq 1 \quad \text{and } m > 0 \quad (71)$$

Let β_m be the estimate of b_m from the previous iteration so that

$$\eta_e(y) = \sum_{m=0}^{\infty} \beta_m P_m(y), \quad (72)$$

(Note that β_m will be taken as zero unless the previous iteration produced a statistically significant estimate of b_m .) Using the orthogonality of the Legendre polynomials and Eqs. 71 and 72

$$C_m^{-1} = \beta_0 + \frac{\beta_m}{2m+1} \quad \text{for } m > 0 \quad (73)$$

Because $d_m(y)$ is sampled rather than $g_m(y)$ there is a weight factor

$$w_m(y) = \frac{g_m(y)}{d_m(y)} = \frac{1}{C_m \eta_e(y)} \quad \text{for } m > 0 \quad (74)$$

Thus, y is sampled from $d_m(y)$ and b'_m is estimated as

$$b'_m \approx \frac{1}{N} \sum_{i=1}^N w_m(y_i) \eta_1(y_i) \quad (75)$$

For $m = 0$

$$b_0 = \int_{-1}^1 \frac{1}{2} \eta(y) dy \quad (76)$$

An estimate of b_0 is obtained by sampling from the importance weighted pdf

$$d_0(y) = C_0 \frac{1}{2} \eta_e(y) \quad (77)$$

instead of $1/2$ as the pdf. Using the orthogonality of the Legendre polynomials,

$$C_0^{-1} = \beta_0 \quad (78)$$

and the importance weighting is

$$w_0(y) = \frac{1/2}{d_0(y)} = \frac{\beta_0}{\eta_e(y)} \quad (79)$$

Thus, y is sampled from $d_0(y)$ and b_0 is estimated as

$$b_0 \approx \frac{1}{N} \sum_{i=1}^N w_0(y_i) \eta_1(y_i) \quad (80)$$

Note that $d_m(y)$ can be sampled by first sampling from

$$h(y) \doteq \frac{\eta_e(y)}{\int_{-1}^1 \eta_e(y) dy}, \quad (81)$$

and then accepting with probability

$$a_m \doteq \frac{1 + P_m(y)}{2} \quad \text{for } m \geq 0 \quad (82)$$

4.4 $A_L(\nu)$ at the Left Boundary

Note that particles exiting the left boundary make no tally and thus contribute nothing to $A_L(\nu)$. Thus, the $A_L(\nu)$ is given by Eq. 40 of [4]

$$A_L(\nu) = N_\nu^{-1} \sum_{m=0}^{\infty} b_m \frac{c\nu}{2} [-\delta_{m0} + \nu J_m(-1 - 2\nu)] \quad (83)$$

4.5 Estimating the c_m

Estimating the c_m is exactly the same as estimating the b_m except that the particles are started on the right boundary moving inward with $-\mu > 0$. For the c_m , $y = -1 + 2\mu$. Rewriting all the importance functions as functions of y . The true importance is,

$$\kappa(y) \doteq \Phi^\dagger(0, -(y+1)/2) = \Phi^\dagger(0, -\mu) \quad (84)$$

The estimated importance from the previous iteration is,

$$\kappa_e(y) \doteq \Phi_e^\dagger(0, -(y+1)/2) = \Phi_e^\dagger(0, -\mu) \quad (85)$$

The one particle estimate of the importance of a particle started at y is

$$\kappa_1(y) \doteq \Phi_1^\dagger(0, -(y+1)/2) = \Phi_1^\dagger(0, -\mu) \quad (86)$$

The analysis of the preceding section follows if the η 's are replaced with the corresponding κ 's above.

4.6 The Direct Component of $A_R(\nu)$ at the Right Boundary

The c_m , and the $A(\nu)$ component associated with the c_m , are due to the particles directed inward at the right boundary. There is also a component of $A(\nu)$ due to particles directed outward at the right boundary. This component can be obtained by direct integration. Let $A_0(\nu)$ denote this direct component. That is, using Eq. 17 of [4] and Eq. 4 herein

$$A_0(\nu) = N_\nu^{-1} \int_{-1}^0 \Phi^\dagger(T, -\mu) \mu \psi_\nu(\mu) d\mu \quad (87)$$

so that $A_R(\nu)$ is using Eq. 40 of [4] and Eq. 87

$$A_R(\nu) = A_0(\nu) + N_\nu^{-1} \sum_{m=0}^{\infty} c_m \frac{c\nu}{2} [-\delta_{m0} + \nu L_m(-1 + 2\nu)] \quad (88)$$

5 Sample Results

A slab of thickness $T = 1$ was chosen for illustration because both the discrete eigenfunction components, $\psi_0^\pm(\mu)$, and the continuum eigenfunction component, $\psi_\nu(\mu)$, are significant.

The sample problem is defined by:

1. a uniform source of $Q(\mu) = 1$ for $\mu \geq 0$ at $x = 0$
2. a scattering probability of $c = 0.5$
3. a thickness of $T = 0$
4. a tally of particles penetrating the right boundary moving forward

The results are shown in Fig. 1. At each iteration, 2000 particles were started at each boundary. Estimates of a_+ , a_- , b_m , and c_m were used to infer an importance function $\Phi^1(x, -\mu)$ that was then used to bias the random walks of the particles on the next iteration as per Eq. 15. The sample standard deviation is the standard deviation of the scores of the individual source particles.

There are twenty graphs in Fig. 1, representing twenty independent Monte Carlo runs starting with different random number seeds. There are ten runs with the Legendre expansion truncated after $m = 4$ and ten runs with the Legendre expansion truncated after $m = 8$. Note four things

1. There is some statistical fluctuation in the convergence between different runs with the same Legendre order.
2. The slopes of the curves are independent of m so that exponential convergence rate does not deteriorate with increasing m .
3. The final sample standard deviations for the $m = 8$ cases are better than for the $m = 4$ cases.
4. The time per iteration is approximately constant, so that Fig. 1 represents an exponential convergence with computer time as well as with iteration.

6 Concluding Remarks

When this paper was given at the conference, this slab problem was one of several discussed in connection with a conundrum about why the approach on this problem produced exponential convergence and other approaches (on other transport problems) produced only power law convergence with computer time. After considerable interaction with Jerome Spanier, Rong Kong, Carole Hayakawa, and Yongzeng Lai of the Claremont Graduate University (CGU), the conundrum was resolved. (The other approaches that were discussed used a cosine basis expansion that has some singular properties at the endpoints.)

This work, by design, was complementary to the work at the Claremont Graduate University under Jerome Spanier. At the time the work started, it was unknown whether exponential convergence could be achieved for the slab problem. The most natural basis functions for the slab problem were the eigenfunctions used herein. If exponential convergence could not be obtained using this basis set, it was unlikely that any other basis set could achieve exponential convergence. For generality, Spanier's group initially chose to try a double Legendre expansion in both x and μ that would be more widely applicable than the eigenfunction basis herein.

The CGU approach initially ran into difficulties associated with trying to represent functions like $e^{-x/\mu}$ (that naturally occur in the slab problem) with a double Legendre expansion in x and μ . Spanier's group then rejected the double Legendre expansion approach and produced exponential convergence by a different, but still general, approach [6]. The approach herein, while

achieving exponential convergence first, is somewhat overshadowed by CGU's exponential convergence with a more general technique.

Although overshadowed by CGU's work for the particular application, there are still some interesting and worthwhile features of the technique herein. First, convergence can be obtained with very few basis functions. Second, this is a very tight coupling of the Monte Carlo method with an analytical transport method. Third, analytical methods in other fields (e.g. partial differential equations) often involve expansion coefficients that are integrals that can be treated by the zero variance biasing and simultaneous coefficient estimation techniques described herein.

References

1. Jerome Spanier and Ely M. Gelbard, "*Monte Carlo Principles and Neutron Transport Problems*," Addison-Wesley Publishing Company (1969).
2. Zero-Variance Solutions for Linear Monte Carlo, Thomas E. Booth, Nuclear Science and Engineering: 102, 332-340 (1989)
3. "Nuclear Reactor Theory," George I. Bell and Samuel Glasstone, Van Nostrand Reinhold Company, Copyright 1970 by Litton Educational Publishing
4. Monte Carlo Estimates of Transport Solutions to the Isotropic Slab Thomas E. Booth, Nuclear Science and Engineering: vol 130 Nov. 1998
5. "Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables", edited by Milton Abramowitz and Irene Stegun, National Bureau of Standards Applied Mathematics Series 55, June 1964
6. get reference from Spanier

T=1 c=.5 2000 particles/iteration