JOURNAL OF COMPUTATIONAL PHYSICS **145,** 406–431 (1998)
ARTICLE NO. CP986039

A Kinetic Theory for Nonanalog Monte Carlo Particle Transport Algorithms: Exponential Transform with Angular Biasing in Planar-Geometry Anisotropically Scattering Media

Taro Ueki* and Edward W. Larsen

Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, Michigan 48109
E-mail: t.ueki@iri.tudelft.nl and edlarsen@engin.umich.edu

Received January 7, 1998; revised May 28, 1998

We show that Monte Carlo simulations of neutral particle transport in planar-geometry anisotropically scattering media, using the exponential transform with angular biasing as a variance reduction device, are governed by a new "Boltzmann Monte Carlo" (BMC) equation, which includes particle weight as an extra independent variable. The weight moments of the solution of the BMC equation determine the moments of the score and the mean number of collisions per history in the nonanalog Monte Carlo simulations. Therefore, the solution of the BMC equation predicts the variance of the score and the figure of merit in the simulation. Also, by (i) using an angular biasing function that is closely related to the "asymptotic" solution of the linear Boltzmann equation and (ii) requiring isotropic weight changes at collisions, we derive a new angular biasing scheme. Using the BMC equation, we propose a universal "safe" upper limit of the transform parameter, valid for any type of exponential transform. In numerical calculations, we demonstrate that the behavior of the Monte Carlo simulations and the performance predicted by deterministically solving the BMC equation agree well, and that the new angular biasing scheme is always advantageous.

I. INTRODUCTION

For many years, Monte Carlo methods have been used to simulate the interaction of neutral particle radiation with matter [1]. In deep penetration "source-detector" problems,

*Present address: Interfaculty Reactor Institute, Delft University of Technology, Mekelweg 15, 2629 JB Delft, The Netherlands.

0021-9991/98 $25.00

nonanalog schemes are frequently used to render the Monte Carlo calculation more efficient. These schemes often employ *geometric splitting* (the user assigns geometric surfaces, across which particles split if they travel in a favorable direction or are Russian-rouletted if they travel in an unfavorable direction) or *weight windows* (the user assigns upper and lower weight windows in the system, and particle weights are maintained within these windows via splitting and Russian roulette) [2]. Both of these nonanalog schemes may require that a large number of biasing parameters be specified prior to the nonanalog simulation. The *exponential transform* is a simpler nonanalog scheme that relies on one biasing parameter. For general multidimensional, energy-dependent problems, the exponential transform has limited effectiveness. However, it has recently been shown that if, like the other biasing schemes, biasing (exponential transform) parameters are defined “locally,” within different subregions of phase space, then the resulting nonanalog scheme can be very efficient [3–5]. However, it is again the case that many biasing parameters must be specified prior to the simulation. In all these nonanalog schemes, obtaining the optimal biasing parameters is a difficult but important practical problem.

In this paper, we propose a new theory for the “classical” exponential transform with angular biasing applied to 1-D monoenergetic deep penetration transport problems with anisotropic scattering. This nonanalog scheme depends on a single biasing parameter. As discussed later, our theory can be extended to much more general transport problems. However, to introduce our ideas, we consider simpler problems in this paper. Our approach extends previous work treating transport problems with isotropic scattering [6, 7]. In our earlier and this present work, we develop a new *Boltzmann Monte Carlo* (BMC) equation that predicts the statistical behavior of Monte Carlo particles governed by the nonanalog Monte Carlo transport process occurring within the computer (but not, of course, in the real world). The new BMC equation yields the first through fourth score-moments of transmitted current estimates, and the mean number of collisions per Monte Carlo particle. Thus, the solution of the BMC equation enables one to predict the variance, the variance of the sample variance, and the Figure of Merit of the nonanalog scheme, for any biasing parameter.

In earlier work, theoretical approaches to predict the second or higher moments of the score for various values of the exponential transform parameter have employed the approximation of discretized flight directions [8–10]. This previous work explains some common features observed in the exponentially transformed simulation, such as the minimization of variance and the divergence and unreliability of variance estimates for large values of the transform parameter. The rigorous prediction of variance was first accomplished by Amster and Djomehri using transport-like integral equations for a score distribution in analog simulations [11]. Sarkar and Prasad [12] extended this approach to the exponential transform and predicted the variance for various values of the transform parameter. More recently, Sarkar and Rief [13] have combined this approach with a technique in sensitivity analysis [14] to predict a value of the transform parameter that minimizes the variance.

This prior work utilizes an adjoint theory, in which one focuses on the calculation of an integral response. The relationship between the adjoint-based work and the present BMC formulation is discussed in detail in [7] but will not be repeated here. In brief, the connection is that one can formulate an adjoint BMC equation and derive the previous equations for the score moment by simply taking weight moments of the adjoint BMC equation. Thus, the BMC equation yields a more basic theory, from which previous approaches can be derived. Also, the BMC equation provides additional information, such as the distribution of Monte Carlo particles and the Figure of Merit.

Variance reduction methods introduce particle weight, and reducing weight fluctuations is very important. In this paper, we propose a new angular biasing scheme in which the same weight change occurs in a collision for any flight direction. This "Generalized Dwivedi Transform" (GDT) generalizes earlier methods proposed by Dwivedi [15, 16] for isotropic scattering problems and Depinay [5] for anisotropic scattering. We use the BMC equation to predict the performance of the GDT scheme. Also, we compare these predictions to direct Monte Carlo simulations and show that the BMC predictions are correct.

In addition, we discuss a connection between the GDT scheme and the "asymptotic" solution of the Boltzmann equation for source-free media. This enables us to determine a "safe" upper limit of the transform parameter, above which the variance-of-the-sample-variance of the nonanalog solution will become infinite if the system is sufficiently large.

The remainder of this paper is organized as follows. In Section II, we develop the BMC equation for monoenergetic 1-D transport problems with anisotropic scattering, subject to the exponential transform with angular biasing. In Section III we develop the GDT scheme, and in Section IV we use the BMC equation to determine the maximum safe values of the GDT transform parameter. In Section V we develop an approximate GDT scheme, which may be used in problems with a high order of anisotropic scattering. In Section VI we discuss the optimization of the Figure of Merit. In Section VII we present numerical results, comparing the performance of various exponential transform schemes with theoretical predictions using the BMC equation. We discuss theoretical features of the GDT scheme in Section VIII and conclude with a general discussion in Section IX.

II. EXPONENTIAL TRANSFORM WITH ANGULAR BIASING

Let us consider the following mono-energetic planar-geometry transport problem with anisotropic scattering,

$$\mu\frac{\partial\psi}{\partial x}(x,\mu)+\sigma_t(x)\psi(x,\mu)$$
$$=\frac{\sigma_{s0}(x)}{2}\int_{-1}^{1}\left[1+\sum_{m=1}^{N}(2m+1)\theta_m(x)P_m(\mu)P_m(\mu')\right]\psi(x,\mu')\,d\mu',$$
$$0<x<X,\ -1\le\mu\le 1,\quad(1)$$

$$\psi(0,\mu)=\frac{\delta(\mu-\mu_{in})}{\mu_{in}},\qquad 0<\mu\le 1,\tag{2}$$

$$\psi(X,\mu)=0,\qquad -1\le\mu<0.\tag{3}$$

Here $P_n(\mu)$ is the nth Legendre polynomial and μ_{in} is fixed on (0, 1]. This problem is driven by a unit monodirectional current on the left boundary. We wish to compute the transmitted current:

$$J^+(X)=\int_0^1\mu\psi(X,\mu)\,d\mu.\tag{4}$$

We apply the following transformation to Eqs. (1)–(3),

$$\Psi(x,\mu)\equiv I(\mu)\,e^{\lambda\int_0^x\sigma_t(x')\,dx'}\psi(x,\mu),\tag{5}$$

where λ is a freely chosen parameter and $I(\mu)$ satisfies

$$I(\mu_{in}) = 1. \tag{6}$$

(The form of I is specified later.) Equation (1) becomes

$$\mu \frac{\partial \Psi}{\partial x}(x, \mu) + \sigma_t (1 - \lambda\mu)\Psi(x, \mu) = \int_{-1}^{1} \sigma_t (1 - \lambda\mu')\Psi(x, \mu')g(\mu', \mu)\, d\mu', \quad 0 < x < X, -1 \le \mu \le 1,$$

where

$$g(\mu', \mu) \equiv \frac{1}{1 - \lambda\mu'} \frac{I(\mu)}{I(\mu')} \frac{\sigma_{s0}}{2\sigma_t} \left[1 + \sum_{m=1}^{N} (2m + 1)\theta_m P_m(\mu) P_m(\mu') \right].$$

The variable x may also appear as an argument in $g(\mu', \mu)$; the omission of this dependence will cause no ambiguity. We rewrite $g(\mu', \mu)$ as the product

$$g(\mu', \mu) = f(\mu', \mu)\rho(\mu'),$$

where $f(\mu', \mu)$ and $\rho(\mu')$ are defined as

$$f(\mu', \mu) \equiv \frac{I(\mu)\left[1 + \sum_{m=1}^{N} (2m + 1)\theta_m P_m(\mu) P_m(\mu')\right]}{\int_{-1}^{1} I(\mu'')\left[1 + \sum_{m=1}^{N} (2m + 1)\theta_m P_m(\mu'') P_m(\mu')\right] d\mu''}, \tag{7}$$

and

$$\rho(\mu') \equiv \frac{1}{(1 - \lambda\mu')I(\mu')} \frac{\sigma_{s0}}{2\sigma_t} \int_{-1}^{1} I(\mu'') \left[1 + \sum_{m=1}^{N} (2m + 1)\theta_m P_m(\mu'') P_m(\mu') \right] d\mu''. \tag{8}$$

Then, $f(\mu', \mu)\, d\mu$ is the probability that in the system governed by the transformed equation, a particle scatters into $d\mu$ about μ assuming it scattered at (x, μ'), and $\rho(\mu')$ is the multiplication factor, or the mean number of particles exiting a collision. For simplicity, we have omitted the notational dependence of these functions on λ, $I(\mu)$, and x. Equations (1)–(3) now may be written

$$\mu \frac{\partial \Psi}{\partial x}(x, \mu) + \sigma_t (1 - \lambda\mu)\Psi(x, \mu) = \int_{-1}^{1} \sigma_t (1 - \lambda\mu')\Psi(x, \mu')\rho(\mu') f(\mu', \mu)\, d\mu', \quad 0 < x < X, -1 \le \mu \le 1, \tag{9}$$

$$\Psi(0, \mu) = \frac{\delta(\mu - \mu_{in})}{\mu_{in}}, \qquad 0 < \mu \le 1, \tag{10}$$

$$\Psi(X, \mu) = 0, \qquad -1 \le \mu < 0, \tag{11}$$

and the transmitted current in the original problem, Eq. (4), is expressed as

$$J^+(X) = e^{-\lambda \int_0^X \sigma_t(x')\, dx'} \int_0^1 \frac{\mu}{I(\mu)} \Psi(X, \mu)\, d\mu. \tag{12}$$

We simulate the system described by Eqs. (9)–(11) using survival biasing [2], in which upon collision, the statistical weight of a particle is multiplied by the multiplication factor ρ.

To analyze the numerical process occurring in this simulation, we define a "Monte Carlo" (MC) particle as one with the following properties [6, 7]:

(1) Statistical weight "w" is an independent variable, just like x and μ.

(2) All MC particles are born with unit weight.

(3) The total cross section at (x, μ) is $\sigma_t(x)(1-\lambda\mu)$. (To ensure that this quantity is positive, we require $|\lambda| < 1$, and to ensure that particles are biased in the positive x-direction, we require $0 < \lambda < 1$.)

(4) At collisions a MC particle always scatters.

(5) Upon collision at (x, μ'), the weight "w" of a MC particle changes by the multiplicative factor $\rho(\mu')$.

(6) The distribution of flight directions after a collision at (x, μ') is $f(\mu', \mu)$.

For these MC particles, we define the angular flux:

$\phi(x, \mu, w)\, dx\, d\mu\, dw \equiv$ angular flux due to MC particles in $(dx, d\mu, dw)$ about (x, μ, w).

We require that the angular flux of MC particles, $\phi(x, \mu, w)$, is related to that of the system of Eqs. (9)–(11) by

$$\Psi(x, \mu) = \int_0^\infty w\phi(x, \mu, w)\, dw. \tag{13}$$

This makes our definition of MC particles consistent with the convention that the weighted particle density in simulations is equal to the particle density in the physical problem.

By the above properties (3)–(6), the collision process of MC particles is expressed by

$\sigma_t(1-\lambda\mu')\phi(x, \mu', w')\, dw'\, d\mu' =$ collision rate due to MC particles in $d\mu'$ about μ' and in dw' about w',

and

$\delta[w - \rho(\mu')w'] f(\mu', \mu)\, dw\, d\mu =$ the probability that when a MC particle with direction μ' and weight w' scatters, its new direction will lie in $d\mu$ about μ and its new weight will lie in dw about w.

Therefore, the Monte Carlo Boltzmann (BMC) equation is

$$\mu\frac{\partial\phi}{\partial x}(x, \mu, w) + \sigma_t(1-\lambda\mu)\phi(x, \mu, w)$$
$$= \int_{-1}^{1}\int_0^\infty \sigma_t(1-\lambda\mu')\phi(x, \mu', w')\delta[w - w'\rho(\mu')] f(\mu', \mu)\, dw'\, d\mu'.$$

The integral on the right side of this equation specifies the weight and direction changes of

Monte Carlo particles when they undergo collisions. Rewriting this integral using $\delta(ax) = \delta(x)/a$ and including the boundary condition, the BMC problem for Eqs. (9)–(11) becomes

$$\mu\frac{\partial\phi}{\partial x}(x,\mu,w) + \sigma_t(1-\lambda\mu)\phi(x,\mu,w)$$
$$= \int_{-1}^{1}\sigma_t(1-\lambda\mu')\phi\left(x,\mu',\frac{w}{\rho(\mu')}\right)\frac{f(\mu',\mu)}{\rho(\mu')}\,d\mu', \qquad 0<x<X, -1\le\mu\le 1, \quad (14)$$

$$\phi(0,\mu,w) = \frac{\delta(\mu-\mu_{in})}{\mu_{in}}\delta(w-1), \qquad 0<\mu\le 1, \quad (15)$$

$$\phi(X,\mu,w) = 0, \qquad -1\le\mu<0. \quad (16)$$

(The second δ-function in Eq. (15) is due to property (2) stated above.)

Now, let us define the nth weight moment of the MC particle flux:

$$\Phi^{(n)}(x,\mu) \equiv \int_0^\infty w^n\phi(x,\mu,w)\,dw, \qquad n = 0, 1, \ldots. \quad (17)$$

Operating on Eqs. (14)–(16) by $\int_0^\infty w^n(\cdot)\,dw$, we obtain the following problem for $\Phi^{(n)}(x,\mu)$:

$$\mu\frac{\partial\Phi^{(n)}}{\partial x}(x,\mu)+\sigma_t(1-\lambda\mu)\Phi^{(n)}(x,\mu) = \int_{-1}^{1}\sigma_t(1-\lambda\mu')\Phi^{(n)}(x,\mu')\rho(\mu')^n f(\mu',\mu)\,d\mu',$$
$$0<x<X, -1\le\mu\le 1, \quad (18)$$

$$\Phi^{(n)}(0,\mu) = \frac{\delta(\mu-\mu_{in})}{\mu_{in}}, \qquad 0<\mu\le 1, \quad (19)$$

$$\Phi^{(n)}(X,\mu) = 0, \qquad -1\le\mu<0. \quad (20)$$

Thus, the problems for the different weight moments are uncoupled.

For $n = 1$, we obtain Eqs. (9)–(11):

$$\Phi^{(1)}(x,\mu) = \Psi(x,\mu). \quad (21)$$

For $n = 0$, integrating Eq. (18) over μ, and using Eq. (17), we obtain

$$\frac{d}{dx}\int_{-1}^{1}\int_0^\infty \mu\phi(x,\mu,w)\,dw\,d\mu = 0.$$

Integrating this equation over $0 < x < X$ and using the boundary conditions (Eqs. (15) and (16)) to eliminate the incident MC fluxes, we obtain

$$1 = \int_0^1\int_0^\infty \mu\phi(X,\mu,w)\,dw\,d\mu + \int_{-1}^{0}\int_0^\infty |\mu|\phi(0,\mu,w)\,dw\,d\mu. \quad (22)$$

Therefore, if we define the probability density function $P(\mu,w)$ as

$P(\mu,w)\,dw\,d\mu \equiv$ the probability that a MC particle introduced at the left boundary will exit the system, with the direction of flight in $d\mu$ about μ, and with weight in dw about w, then

$$P(\mu, w) = \begin{cases} \mu\phi(X, \mu, w), & \text{for } 0 < \mu \le 1, 0 < w, \\ 0, & \text{for } \mu = 0, 0 < w, \\ |\mu|\phi(0, \mu, w), & \text{for } -1 \le \mu < 0, 0 < w. \end{cases} \tag{23}$$

Using Eqs. (17) and (21), we obtain from Eq. (12)

$$\begin{aligned} J^+(X) &= e^{-\lambda \int_0^X \sigma_t(x)\,dx} \int_0^1 \int_0^\infty \frac{\mu w}{I(\mu)} \phi(X, \mu, w)\, dw\, d\mu \\ &= \int_{-1}^1 \int_0^\infty S(\mu, w) P(\mu, w)\, dw\, d\mu, \end{aligned} \tag{24}$$

where $S(\mu, w)$, the score for the transmitted current estimate, is

$$S(\mu, w) \equiv \begin{cases} e^{-\lambda \int_0^X \sigma_t(x)\,dx} \frac{w}{I(\mu)}, & \text{for } 0 < \mu \le 1, 0 < w, \\ 0, & \text{for } -1 \le \mu \le 0, 0 < w. \end{cases} \tag{25}$$

The nth moment of the transmitted current estimate is

$$E[S^n(\mu, w)] \equiv \int_{-1}^1 \int_0^\infty S^n(\mu, w) P(\mu, w)\, dw\, d\mu, \tag{26}$$

and

$$J^+(X) = E[S(\mu, w)].$$

Using Eqs. (23), (25), and (17), we find that Eq. (26) becomes

$$E[S^n(\mu, w)] = \int_0^1 S^n(\mu, 1) \mu \Phi^{(n)}(X, \mu)\, d\mu. \tag{27}$$

Thus, using the solution of Eqs. (18)–(20), we can compute the nth moment of transmitted current estimates by Eq. (27).

We can simplify Eqs. (18)–(20) by defining

$$\xi^{(n)}(x, \mu) \equiv \frac{e^{-\lambda \int_0^x \sigma_t(x')\,dx'}}{I(\mu)} \Phi^{(n)}(x, \mu). \tag{28}$$

Then Eqs. (18)–(20) become

$$\begin{aligned} &\mu \frac{\partial \xi^{(n)}}{\partial x}(x, \mu) + \sigma_t \xi^{(n)}(x, \mu) \\ &\quad = \frac{\sigma_{s0}}{2} \int_{-1}^1 \left[1 + \sum_{m=1}^N (2m+1)\theta_m P_m(\mu') P_m(\mu) \right] \rho^{n-1}(\mu') \xi^{(n)}(x, \mu')\, d\mu', \\ &\qquad\qquad 0 < x < X, -1 \le \mu \le 1, \end{aligned} \tag{29}$$

$$\xi^{(n)}(0, \mu) = \frac{\delta(\mu - \mu_{in})}{\mu_{in}}, \qquad 0 < \mu \le 1, \tag{30}$$

$$\xi^{(n)}(X, \mu) = 0, \qquad -1 \le \mu < 0. \tag{31}$$

Also, by Eqs. (25), (27), and (28), the nth moment of the transmitted current estimate is

$$E[S^n(\mu, w)] = \int_0^1 S^{n-1}(\mu, 1)\mu\xi^{(n)}(X, \mu)\, d\mu. \tag{32}$$

We use Eqs. (29)–(32) to predict the score moments of the transmitted current, particularly the variance $E[S^2(\mu, w)] - E^2[S(\mu, w)]$.

III. ISOTROPIC MULTIPLICATIVE FACTOR

For problems with isotropic scattering, Dwivedi's importance transform [15], which consists of the exponential transform with angular biasing, is much more efficient than the plain exponential transform with no angular biasing, $[I(\mu) = 1]$ [7]. (Dwivedi's approximation to the importance function that is determined by the zero variance problem [16] becomes an exact asymptotic solution of the adjoint transport equation for a special value of the biasing parameter λ; see Eq. (43).) The angular part of Dwivedi's importance transform is

$$I(\mu) = \frac{1 - \lambda\mu_{in}}{1 - \lambda\mu}. \tag{33}$$

When Eq. (33) is applied to a problem with isotropic scattering ($N = 0$), the multiplicative factor $\rho(\mu)$ (Eq. (8)) becomes

$$\rho(\mu) = \rho_0 = \frac{\sigma_{s0}}{2\sigma_t} \int_{-1}^1 \frac{d\mu'}{1 - \lambda\mu'}. \tag{34}$$

Thus, the change of weight at a collision is independent of the direction of flight before the collision. Also, the distribution of flight directions after the collision $f(\mu', \mu)$ (Eq. (7)) is biased toward the positive direction:

$$f(\mu', \mu) = \frac{1/(1 - \lambda\mu)}{\int_{-1}^1 (d\mu'/(1 - \lambda\mu'))}.$$

These two features make the fluctuation of the score, $S(\mu, w)$ in Eq. (25), very small, yielding a significant reduction of variance. In this section, we propose a generalization of Dwivedi's transform to problems with anisotropic scattering. We call our proposed scheme the Generalized Dwivedi Transform (GDT).

In the GDT scheme, we require $\rho(\mu)$ in Eq. (8) to be isotropic:

$$\rho(\mu) = \rho_0. \tag{35}$$

(Here, we omit showing the dependence of ρ_0 on λ and other parameters in $I(\mu)$ for notational simplicity.) We obtain the following eigenvalue problem for ρ_0 and $I(\mu)$,

$$\rho_0(1 - \lambda\mu)I(\mu) = \frac{c}{2} \int_{-1}^1 I(\mu') \left[1 + \sum_{m=1}^{N} (2m + 1)\theta_m P_m(\mu') P_m(\mu)\right] d\mu', \tag{36}$$

where $c = \sigma_{s0}/\sigma_t$. When $\lambda = 0$, $I(\mu) = 1$ and $\rho_0 = c$, as expected. In the following analysis, we assume $\lambda \neq 0$ unless specifically mentioned.

III.I. *Linearly Anisotropic Scattering*

When $N=1$, Eq. (36) becomes

$$\rho_0(1-\lambda\mu)I(\mu)=\frac{c}{2}\int_{-1}^{1}I(\mu')(1+3\theta_1\mu'\mu)\,d\mu'. \tag{37}$$

Operating on Eq. (37) by $\int_{-1}^{1}(\cdot)\,d\mu$, we obtain

$$\int_{-1}^{1}\mu' I(\mu')\,d\mu'=\frac{\rho_0-c}{\rho_0\lambda}\int_{-1}^{1}I(\mu')\,d\mu'. \tag{38}$$

Combining Eq. (37) with (38), $I(\mu)$ is expressed as

$$I(\mu)=\frac{c}{2\rho_0}\left(\frac{1+3\theta_1((\rho_0-c)/\rho_0\lambda)\mu}{1-\lambda\mu}\right)\int_{-1}^{1}I(\mu')\,d\mu'. \tag{39}$$

Operating on Eq. (39) by $\int_{-1}^{1}(\cdot)\,d\mu$, we obtain

$$\rho_0=\frac{c}{2}\int_{-1}^{1}\frac{1+(3\theta_1/\rho_0\lambda)(\rho_0-c)\mu}{1-\lambda\mu}\,d\mu.$$

This yields the following quadratic equation for ρ_0:

$$\rho_0^2-\left(\frac{c}{2}\int_{-1}^{1}\frac{1+(3\theta_1/\lambda)\mu}{1-\lambda\mu}\,d\mu\right)\rho_0+\left(\frac{c^2}{2}\int_{-1}^{1}\frac{(3\theta_1/\lambda)\mu}{1-\lambda\mu}\,d\mu\right)=0. \tag{40}$$

When $\theta_1=0$ (isotropic scattering), Eq. (40) reduces to Eq. (34). When $\theta_1>0$, Eq. (40) has two positive solutions, one of which is larger than c because the left side of Eq. (40) is positive at $\rho_0=0$ and negative at $\rho_0=c$. We choose the ρ_0 which is larger than c because due to Eq. (38), this choice is consistent with angular biasing toward the positive direction.

For each value of λ, we calculate ρ_0 in this manner and obtain the angular part of the GDT by Eqs. (39) and (6),

$$I(\mu)=\frac{1-\lambda\mu_{in}}{1+\alpha_1\mu_{in}}\frac{1+\alpha_1\mu}{1-\lambda\mu}, \tag{41}$$

where

$$\alpha_1=3\theta_1\frac{\rho_0-c}{\rho_0\lambda}. \tag{42}$$

When $\theta_1=0$, $\alpha_1=0$, and Eq. (41) reduces to Eq. (33). We define λ^* as the value of λ that yields $\rho_0=1$:

$$1=\frac{c}{2}\int_{-1}^{1}\frac{1+(3\theta_1/\lambda^*)(1-c)\mu}{1-\lambda^*\mu}\,d\mu. \tag{43}$$

When $\theta_1=0$, Eq. (43) reduces to Eq. (34) with $\rho(\mu)=\rho_0=1$, the corresponding equation for $N=0$. For $\lambda=\lambda^*$, there is no fluctuation in a particle's weight: this weight is initially unity and remains unity after arbitrarily many collisions.

III.II. *Quadratically Anisotropic Scattering*

When $N = 2$, Eq. (36) becomes

$$\rho_0(1 - \lambda\mu)I(\mu) = \frac{c}{2}\int_{-1}^{1} I(\mu')[1 + 3\theta_1\mu'\mu + 5\theta_2 P_2(\mu')P_2(\mu)]\,d\mu'. \tag{44}$$

Operating on Eq. (44) by $\int_{-1}^{1}(\cdot)\,d\mu$, we obtain Eq. (38). Operating on Eq. (44) by $\int_{-1}^{1}\mu(\cdot)\,d\mu$, we obtain

$$(\rho_0 - c\theta_1)\int_{-1}^{1}\mu I(\mu)\,d\mu = \rho_0\lambda\int_{-1}^{1}\mu^2 I(\mu)\,d\mu. \tag{45}$$

Using the recursion relation of Legendre polynomial $P_1(\mu) = \mu^2 = \frac{2}{3}P_2(\mu) + \frac{1}{3}P_0(\mu)$ and Eq. (38), Eq. (45) becomes

$$\int_{-1}^{1} P_2(\mu)I(\mu)\,d\mu = \left[\frac{3}{2\rho_0^2\lambda^2}(\rho_0 - c\theta_1)(\rho_0 - c) - \frac{1}{2}\right]\int_{-1}^{1} I(\mu)\,d\mu. \tag{46}$$

If we define

$$A(\lambda, \rho_0) \equiv \frac{\rho_0 - c}{\rho_0\lambda} \tag{47}$$

and

$$B(\lambda, \rho_0) \equiv \frac{3}{2\rho_0^2\lambda^2}(\rho_0 - c\theta_1)(\rho_0 - c) - \frac{1}{2}, \tag{48}$$

then combining Eq. (44) with Eqs. (38) and (46)–(48), we obtain

$$I(\mu) = \left(\frac{c}{2\rho_0}\right)\frac{1 + 3\theta_1 A(\lambda, \rho_0)\mu + 5\theta_2 B(\lambda, \rho_0)P_2(\mu)}{1 - \lambda\mu}\int_{-1}^{1} I(\mu)\,d\mu. \tag{49}$$

Finally, integrating Eq. (49) over μ, we obtain

$$\rho_0 = \frac{c}{2}\left[L_0(\lambda) + 3\theta_1 A(\lambda, \rho_0)L_1(\lambda) + 5\theta_2 B(\lambda, \rho_0)L_2(\lambda)\right],$$

where

$$L_n(\lambda) \equiv \int_{-1}^{1}\frac{P_n(\mu)\,d\mu}{1 - \lambda\mu}.$$

This yields the following cubic equation for ρ_0:

$$\rho_0^3 - \frac{c}{2}\left(L_0(\lambda) + \frac{3\theta_1 L_1(\lambda)}{\lambda} + \frac{15\theta_2 L_2(\lambda)}{2\lambda^2} - \frac{5\theta_2 L_2(\lambda)}{2}\right)\rho_0^2 + \frac{c^2}{2}\left(\frac{3\theta_1 L_1(\lambda)}{\lambda} + \frac{15\theta_2(\theta_1 + 1)L_2(\lambda)}{2\lambda^2}\right)\rho_0 - \left(\frac{15c^3\theta_1\theta_2 L_2(\lambda)}{4\lambda^2}\right) = 0. \tag{50}$$

As in the previous subsection, we choose the value of ρ_0 that is greater than c.

By Eqs. (49) and (6), the angular part of the GDT becomes

$$I(\mu) = \frac{1-\lambda\mu_{in}}{1+\alpha_1\mu_{in}+\alpha_2 P_2(\mu_{in})}\,\frac{1+\alpha_1\mu+\alpha_2 P_2(\mu)}{1-\lambda\mu}, \tag{51}$$

where

$$\alpha_1 \equiv 3\theta_1 A(\lambda, \rho_0), \tag{52}$$

$$\alpha_2 \equiv 5\theta_2 B(\lambda, \rho_0), \tag{53}$$

and ρ_0 is calculated by Eq. (50). When $\theta_2 = 0$, $\alpha_2 = 0$, Eq. (50) reduces to Eq. (40), and Eq. (51) reduces to Eq. (41). As before, we define λ^* to be such that when $\lambda = \lambda^*$, $\rho_0 = 1$,

$$1 = \frac{c}{2}[L_0(\lambda^*) + 3\theta_1 A(\lambda^*, 1)L_1(\lambda^*) + 5\theta_2 B(\lambda^*, 1)L_2(\lambda^*)]. \tag{54}$$

For $\theta_2 = 0$, Eq. (54) reduces to Eq. (43), the corresponding equation for $N = 1$. When $\lambda = \lambda^*$, each particle's weight is always unity, as before.

IV. MAXIMUM "SAFE" VALUES OF THE TRANSFORM PARAMETERS

The parameters introduced in the previous sections should be chosen to lie in the range for which $E[S^n(\mu, w)]$ is finite for $n = 2, 3$, and 4; this guarantees that both the variance and the variance of the variance estimate are finite [7]. By Eq. (32), these conditions are satisfied if and only if $\xi^{(n)}(X, \mu)$ is finite for $n = 2, 3$, and 4. Thus, Eqs. (29)–(31) must have bounded solutions for $n = 2, 3$, and 4. When the fourth moment is finite, the first through third moments are finite, because for $0 < \alpha < \beta$, $E[|S|^\alpha]^{\frac{1}{\alpha}} \le E[|S|^\beta]^{\frac{1}{\beta}}$ (Lyapounov's inequality [7]). Therefore, it suffices to consider a case such that $\xi^{(4)}(X, \mu)$ is "borderline" finite.

Let us consider Eqs. (29)–(31) for $X = \infty$ (semi-infinite medium). If these equations represent "non-multiplying" media, then their solutions will decrease to zero as $X \to \infty$, and hence will remain bounded. The borderline finite case is when for $n = 4$, these equations have a solution that approaches a nonzero constant as $X \to \infty$. When this happens, this solution is finite for all x, but if the multiplicative factor $\rho(\mu')$ increases by even the slightest amount, it may become infinite as $X \to \infty$.

Thus, we define the *maximum safe value* of λ to be one for which Eq. (29) with $n = 4$ has a solution independent of x. This implies

$$\xi^{(4)}(\mu) = \frac{\sigma_{s0}}{2\sigma_t}\int_{-1}^{1}\left[1+\sum_{m=1}^{N}(2m+1)\theta_m P_m(\mu')P_m(\mu)\right]\rho(\mu')^3\xi^{(4)}(\mu')\,d\mu', \quad -1 \le \mu \le 1. \tag{55}$$

Because $\rho(\mu')$ in Eq. (8) is defined by $I(\mu')$ and λ as well as cross sections, the existence of a solution of Eq. (55) depends on the value of the transform parameter λ. When $\rho(\mu')$ does not depend on μ' ($\rho(\mu') = \rho_0 = $ const w.r.t. μ'), we obtain, by integrating Eq. (55) over μ,

$$1 = \frac{\sigma_{s0}}{\sigma_t}\rho_0^3.$$

In GDT, $I(\mu)$ is defined by Eqs. (33), (41), or (51) for $N = 0, 1$, or 2, respectively, and $\rho(\mu')$ does not depend on μ'. In this case, we can define the maximum safe value, λ_{max}, as the value which gives

$$\rho_0(\lambda_{max}) = \sqrt[3]{\frac{\sigma_t}{\sigma_{s0}}}. \tag{56}$$

Here $\rho_0(\lambda_{max})$ denotes the root of the linear equation (34), quadratic equation (40), or the cubic equation (50) when $\lambda = \lambda_{max}$.

V. APPROXIMATE METHOD

An inefficient aspect of the exponential transform for problems with anisotropic scattering is that sampling flight directions leads to employing the rejection method [1, 2]. To cope with this, for $I(\mu)$ defined by Eqs. (41) or (51), we sample a new flight direction μ from $f(\mu', \mu)$ in the following way:

(1) Sample μ from $h(\mu)$ by direct inversion, where

$$h(\mu) = \frac{1}{1-\lambda\mu} \Big/ \int_{-1}^{1} \frac{d\mu'}{1-\lambda\mu'}.$$

(2) Sample u uniformly on $(0, G(\mu'))$, where

$$G(\mu') = \max_{-1\le v\le 1} [1 + 3\theta_1 v\mu' + 5\theta_2 P_2(v) P_2(\mu')][1 + \alpha_1 v + \alpha_2 P_2(v)].$$

(3) Accept μ if $u < [1 + 3\theta_1\mu\mu' + 5\theta_2 P_2(\mu) P_2(\mu')][1 + \alpha_1\mu + \alpha_2 P_2(\mu)]$. Otherwise, go back to step (1) and repeat.

This logic is valid because

$$\begin{aligned} P(\mu \mid \text{accept}) &= \frac{P(\mu, \text{accept})}{P(\text{accept})} \\ &= \frac{h(\mu) \frac{[1+3\theta_1\mu\mu' + 5\theta_2 P_2(\mu) P_2(\mu')][1+\alpha_1\mu + \alpha_2 P_2(\mu)]}{G(\mu')}}{\int h(\mu'') \frac{[1+3\theta_1\mu''\mu' + 5\theta_2 P_2(\mu'') P_2(\mu')][1+\alpha_1\mu'' + \alpha_2 P_2(\mu'')]}{G(\mu')}\, d\mu''} \\ &= f(\mu', \mu), \end{aligned}$$

where $P(\mu \mid \text{accept})\, d\mu$ is the conditional probability of the direction cosine taking values in $d\mu$ about μ assuming that it is accepted, and $P(\mu, \text{accept})\, d\mu$ is the probability of the direction cosine taking values in $d\mu$ about μ and its being accepted. Since

$$[1 + 3\theta_1 v\mu' + 5\theta_2 P_2(v) P_2(\mu')][1 + \alpha_1 v + \alpha_2 P_2(v)]$$

is less steep than

$$[1 + 3\theta_1 v\mu' + 5\theta_2 P_2(v) P_2(\mu')][1 + \alpha_1 v + \alpha_2 P_2(v)]/(1 - \lambda v)$$

for most of the values of μ' on $[-1, 1]$, the above rejection method is more efficient than plain rejection sampling. Thus, if $I(\mu)$ (Eqs. (41) and (51)) has a reasonably good approximation, $I_{ap}(\mu)$, for which

$$I_{ap}(\mu) \Big/ \int_{-1}^{1} I_{ap}(\mu)\, d\mu$$

can be sampled by direct inversion, then sampling flight directions becomes efficient without sacrificing overall efficiency because

$$1 + 3\mu_1 v\mu' + 5\mu_2 P_2(v) P_2(\mu')$$

is less steep than

$$[1 + 3\mu_1 v\mu' + 5\mu_2 P_2(v) P_2(\mu')][1 + \alpha_1 v + \alpha_2 P_2(v)].$$

This logic is also valid because in (1), $h(\mu)$ is replaced by

$$I_{ap}(\mu) \Big/ \int_{-1}^{1} I_{ap}(\mu)\, d\mu,$$

in (2), $G(\mu')$ is replaced by

$$G(\mu') = \max_{-1\le v\le 1} [1 + 3\theta_1 v\mu' + 5\theta_2 P_2(v) P_2(\mu')],$$

and in (3),

$$u < [1 + 3\theta_1 \mu\mu' + 5\theta_2 P_2(\mu) P_2(\mu')][1 + \alpha_1 \mu + \alpha_2 P_2(\mu)]$$

is replaced by

$$u < [1 + 3\theta_1 \mu\mu' + 5\theta_2 P_2(\mu) P_2(\mu')].$$

Therefore, $P(\mu \mid accept)$ becomes $f(\mu', \mu)$ with $I(\mu) = I_{ap}(\mu)$.

Also, the "maximum entropy" form of the approximation to $I(\mu)$ defined by Eqs. (41) and (51) is exponential:

$$I(\mu) \approx I_{ap}(\mu) = e^{\beta(\mu-\mu_{in})}. \tag{57}$$

This guarantees direct inversion sampling. Because $|\lambda\mu| < 1$ and $|\lambda\mu_{in}| < 1$, Eq. (51) becomes

$$I(\mu) = \left[\left(1 - \frac{\alpha_2}{2}\right) + \left(\alpha_1 + \lambda - \frac{\alpha_2\lambda}{2}\right)\mu + \cdots\right] \Big/ \left[\left(1 - \frac{\alpha_2}{2}\right) + \left(\alpha_1 + \lambda - \frac{\alpha_2\lambda}{2}\right)\mu_{in} + \cdots\right].$$

Then, β may be approximated by

$$\beta = \frac{\alpha_1 + \lambda - \alpha_2\lambda/2}{1 - \alpha_2/2}. \tag{58}$$

This approximation is accurate for small values of λ because then α_1 and α_2 are small.

The exponential approximation to $I(\mu)$ may be useful in problems for which the exact algebraic expressions for $I(\mu)$, as are obtained in Section III, are too complex to warrant exact treatment. This can occur, for example, if the scattering process in the original transport equation has a high-order Legendre polynomial expansion.

VI. OPTIMIZATIONS

Now we show how the BMC equation can be used to optimize the choice of the biasing parameter λ. First, the mean number of collisions in the history of one MC particle, $CL(\lambda)$, is expressed as

$$CL(\lambda) = \int_0^X \int_{-1}^1 \int_0^\infty \sigma_t(1 - \lambda\mu)\phi(x, \mu, w)\, dw\, d\mu\, dx, \tag{59}$$

where $\phi(x, \mu, w)$ is the solution of Eqs. (14)–(16). (We assume that $I(\mu) = 1$ or $I(\mu)$ is defined by either of Eqs. (33), (41), (51), or (57).) Equation (59) is valid because Eqs. (14)–(16) imply that the integral in Eq. (59) is the collision rate for Monte Carlo particles [7]. Equation (59) may be rewritten as

$$CL(\lambda) = \int_0^X \sigma_t(x)\, e^{\lambda \int_0^x \sigma_t(x')\, dx'} \left[\int_{-1}^1 (1 - \lambda\mu) I(\mu) \xi^{(0)}(x, \mu)\, d\mu\right] dx, \tag{60}$$

using Eqs. (17) and (28).

When using variance reduction methods, the ultimate interest is to maximize the figure of merit (FOM) [18]:

$$\text{FOM} = \frac{1}{(\text{cpu time})(\text{variance})}.$$

If differences in the efficiency of rejection sampling for various values of λ can be ignored, the cpu time is roughly proportional to the mean number of flights per history, $CL(\lambda) + 1$. Therefore, we may define a quality factor, $Q(\lambda)$, as the FOM normalized to its value at $\lambda = 0$:

$$Q(\lambda) \equiv \frac{[CL(0) + 1][E(S^2) - E^2(S)]_{\lambda=0}}{[CL(\lambda) + 1][E(S^2) - E^2(S)]_\lambda}. \tag{61}$$

Here, $E[S^n]$ is computed using Eqs. (29)–(32); its dependence on λ through $I(\mu)$ is shown as subscripts. Equation (61) can be used as a guide to optimize the exponential transform with angular biasing; one searches to determine the value of the transform parameter λ that maximizes $Q(\lambda)$.

VII. NUMERICAL RESULTS

In this section we present numerical results of the Monte Carlo estimation and theoretical prediction of variance, the theoretical prediction of the fourth score moment, and the optimization of the transform parameter. In Monte Carlo simulations of the problem of Eqs. (1)–(4), we employed four nonanalog schemes:

(1) The plain exponential transform (exponential transform with no angular biasing), for which $I(\mu) = 1$ in Eq. (5).

(2) Dwivedi's transform, for which $I(\mu)$ is defined by Eq. (33).

(3) The Generalized Dwivedi Transform (GDT), for which $I(\mu)$ is defined by Eq. (41) for linearly anisotropic scattering and Eq. (51) for quadratically anisotropic scattering.

(4) The approximate method to the GDT, for which $I(\mu)$ is defined by Eqs. (57) and (58).

For the resulting transformed system, we simulated Eqs. (9)–(12) following the properties (1)–(6) stated in Section II. This is equivalent to simulating Eqs. (14)–(16) for Monte Carlo particles. In the theoretical prediction of the score moments and the mean number of flights per history by solving Eqs. (29)–(31) and computing Eqs. (32) and (60), we wrote a special-purpose discrete ordinate (S_N) code [19]. We chose $\theta_1 = \sigma_{s1}/\sigma_{s0} = 0.3$ for linearly anisotropic scattering problems, and $\theta_1 = 0.5$ and $\theta_2 = \sigma_{s2}/\sigma_{s0} = 0.1$ for quadratically anisotropic scattering problems, for various values of $c = \sigma_{s0}/\sigma_t$. The slab thickness is always taken to be 15 mean free paths.

Numerical results for linearly anisotropic scattering with $c = 0.9$ are presented in Figs. 1 and 2. In Fig. 1 we observe that the variance is smaller using Dwivedi's transform than using the plain exponential transform, and that it is further reduced using the GDT. Figure 2 shows that the variance estimates become unreliable for large values of the transform parameter because the fourth moment diverges (the variance of the variance becomes infinite). In Fig. 1, we also show λ^*, defined by Eq. (43), and λ_{max}, defined by Eq. (56). We observe that in the GDT, the variance is minimized near $\lambda = \lambda^*$, and that $0 < \lambda < \lambda_{max}$ is indeed a

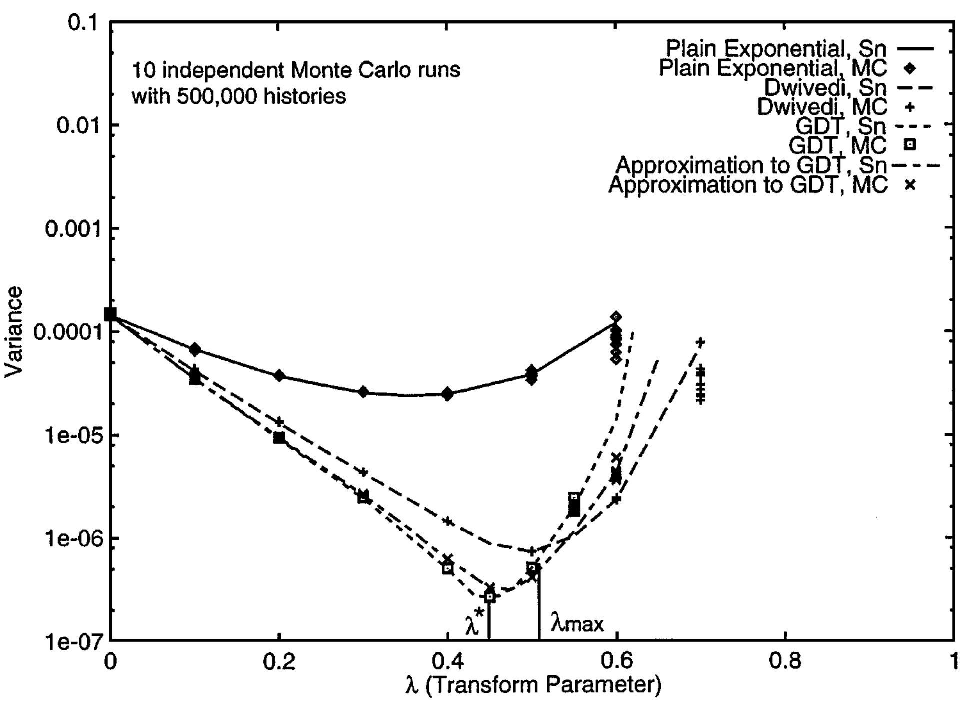

FIG. 1. Variance by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

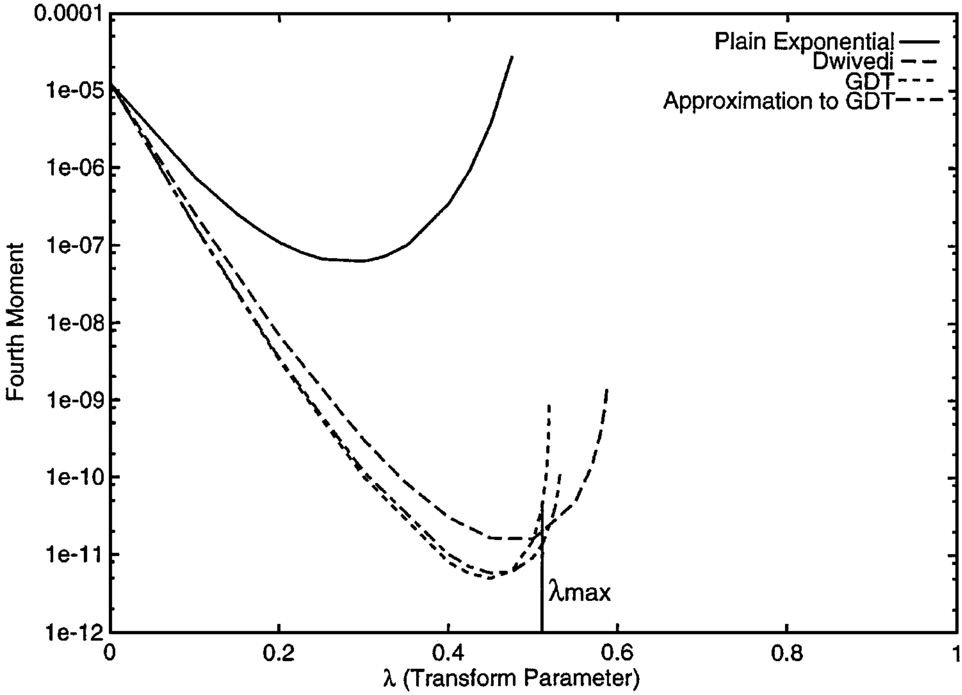

FIG. 2. Theoretical (S_N) predictions of fourth score moment for $c = 0.9$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

safe range because for such λ the fourth moment, the dominant term ensuring the finiteness of the variance of variance estimates [7], remains finite. The same numerical results are shown in Figs. 3 and 4 for quadratically anisotropic scattering with $c = 0.9$, with λ^* defined by Eq. (54). We observe a similar performance as in Figs. 1 and 2, and a smaller variance for the GDT than for Dwivedi's transform. We also observe in Figs. 1 and 3 that for $\lambda \leq \lambda^*$, the exponential approximation to GDT is efficient in the linearly anisotropic case and very efficient in the quadratically anisotropic case.

Results for linearly and quadratically anisotropic scattering with $c = 0.7$ are shown in Figs. 5–8. We observe similar performances as with $c = 0.9$, except for the exponential approximation to the GDT. This is not an efficient approximation for linearly anisotropic scattering with $\lambda \geq 0.6$, but it is reasonably efficient for quadratically anisotropic scattering with $\lambda \leq \lambda^*$.

We observe that the theoretical prediction of variance using the BMC equation is accurate and that the approximate method is reasonably close in performance to the GDT for values of the transform parameter, λ, up to about 0.6. Surprisingly, for "large" values of λ, greater than λ^*, the exponential approximation to the GDT has smaller variances than the GDT method itself. We also calculated the variance using the S_N code for $c = 0.98$ and 0.5 with linearly and quadratically anisotropic scattering. We observed that the merit of the GDT over Dwivedi's method is smaller for these scattering ratios compared to the results for $c = 0.9$ and 0.7. These results are not shown here because they are very similar to Figs. 1, 3, 5, and 7. We may conclude that in the GDT, variance is always minimized at near $\lambda = \lambda^*$, that this minimal value of the variance is smaller than the minimum value of the variance in other methods, and that for any value of λ on $[0, \lambda^*]$, the variance using the GDT is always smaller than the variance using Dwivedi's method.

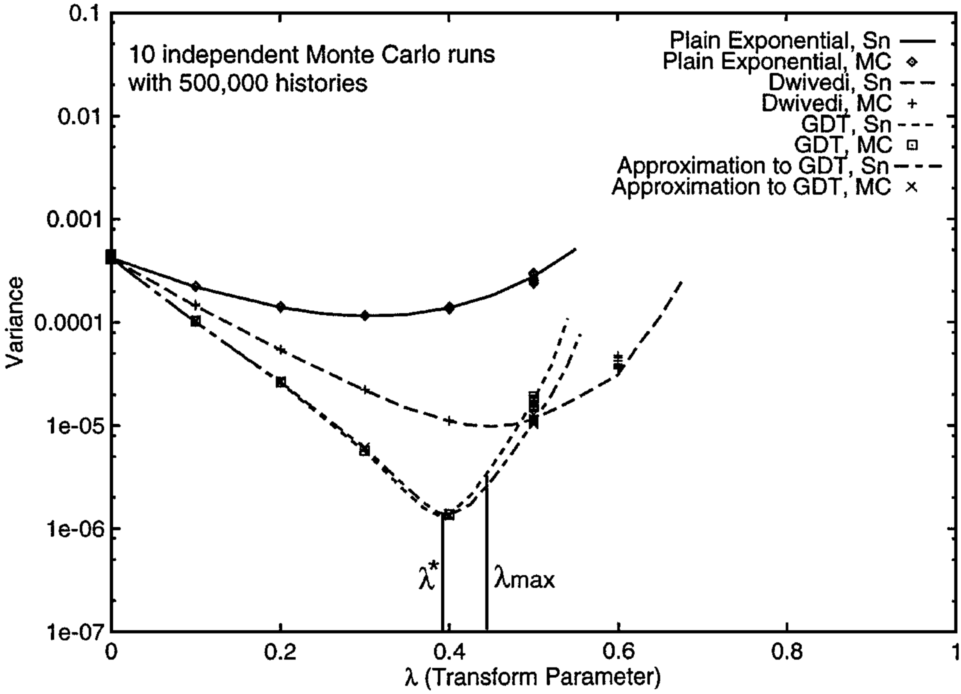

FIG. 3. Variance by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

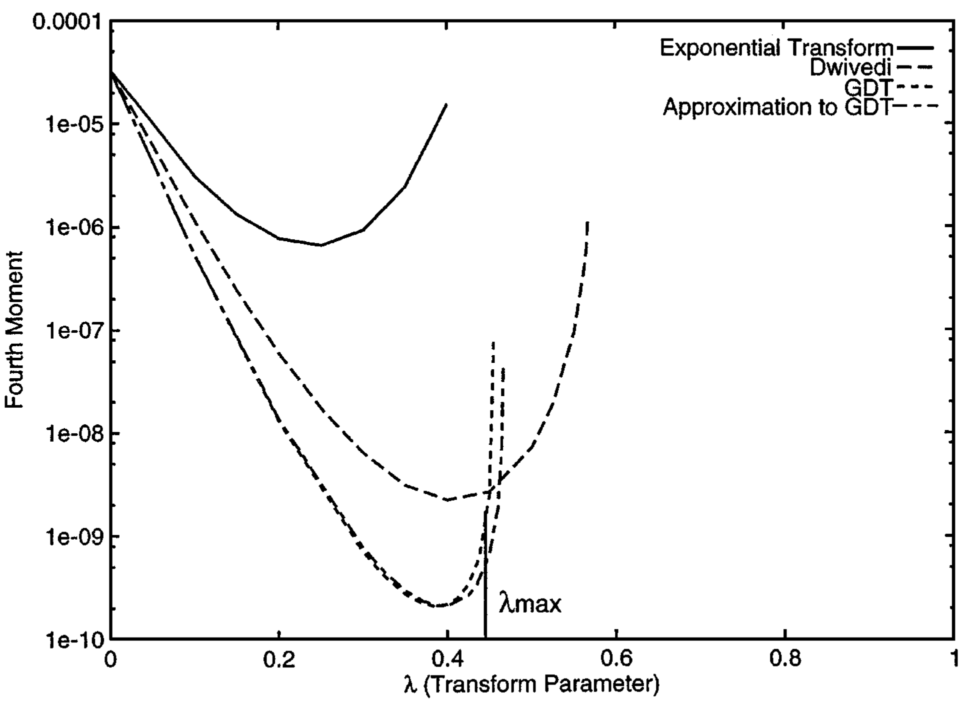

FIG. 4. Theoretical (S_N) predictions of fourth score moment for $c = 0.9$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

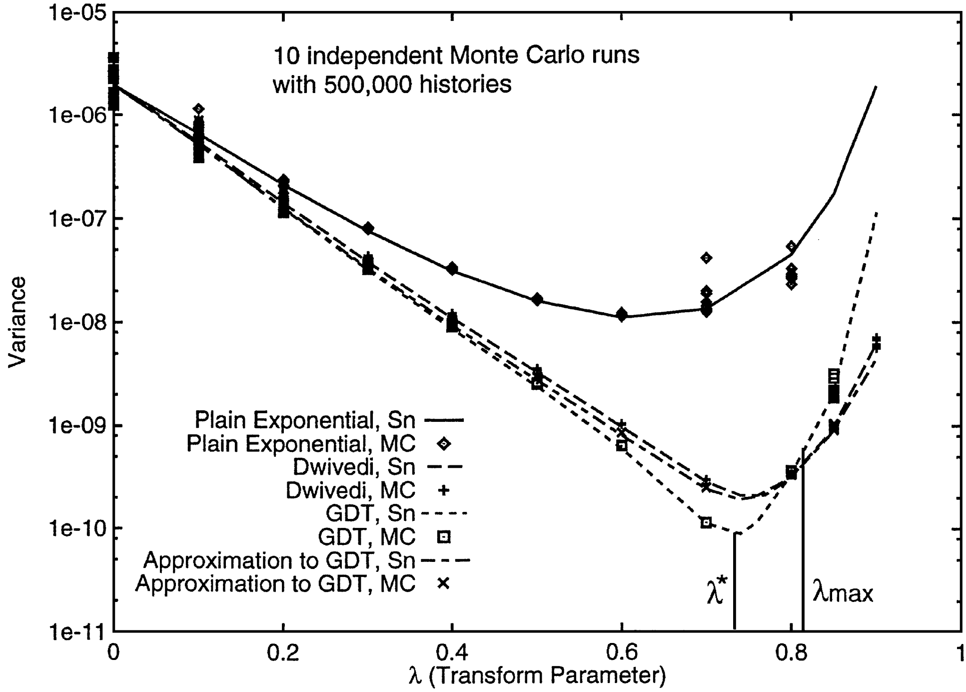

FIG. 5. Variance by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.7$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

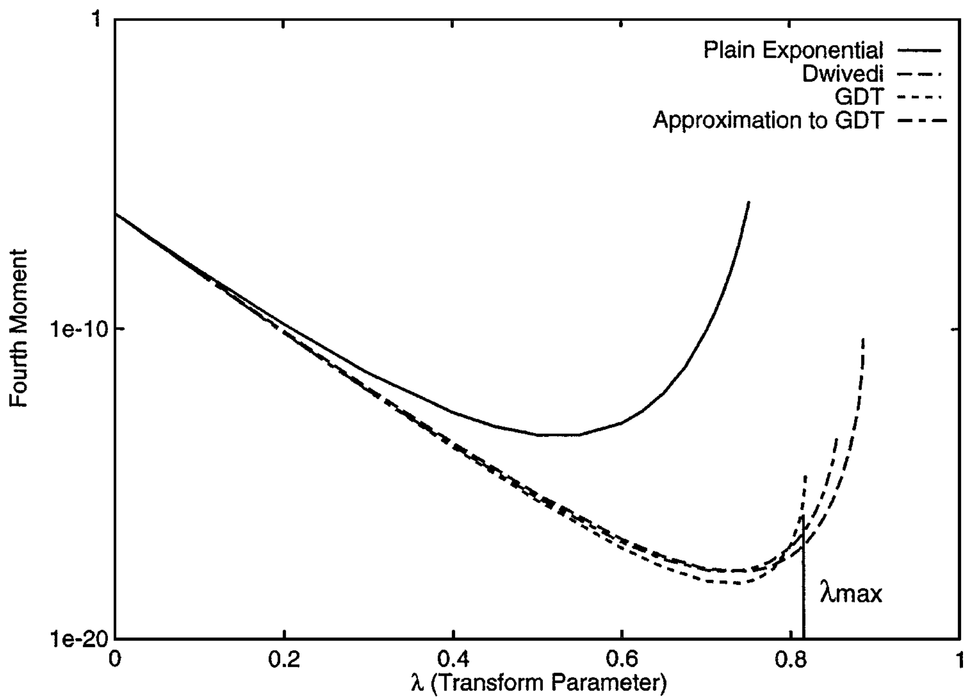

FIG. 6. Theoretical (S_N) predictions of fourth score moment for $c = 0.7$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

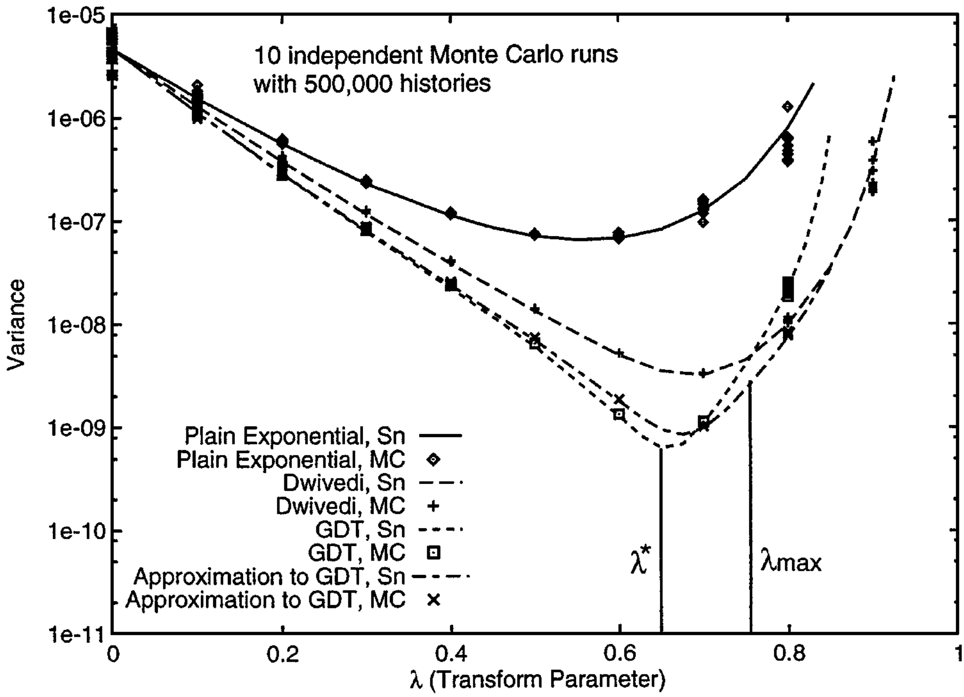

FIG. 7. Variance by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.7$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

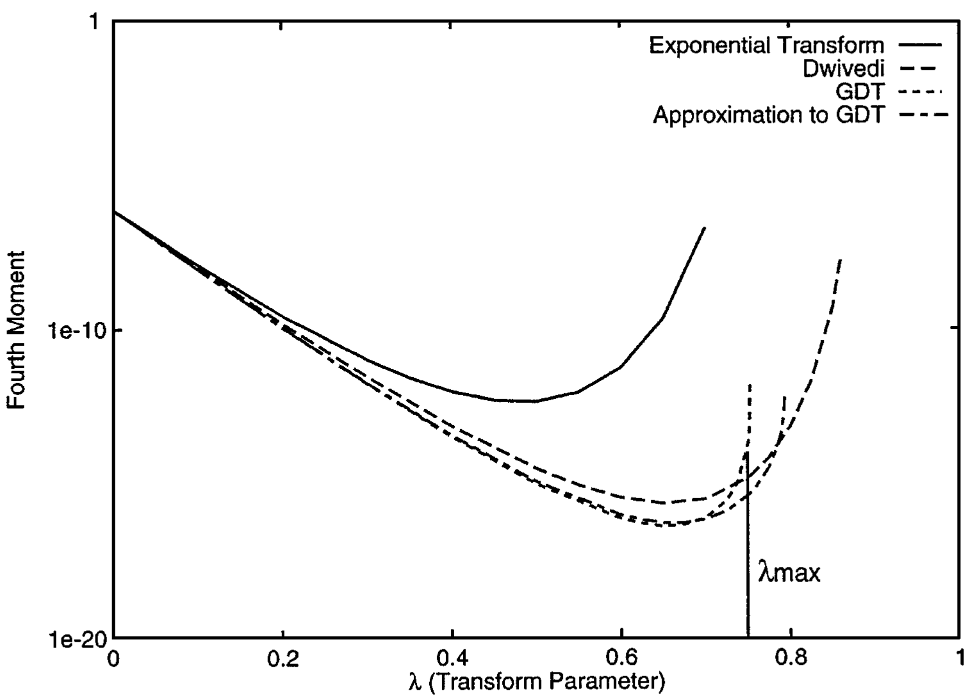

FIG. 8. Theoretical (S_N) predictions of fourth score moment for $c = 0.7$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

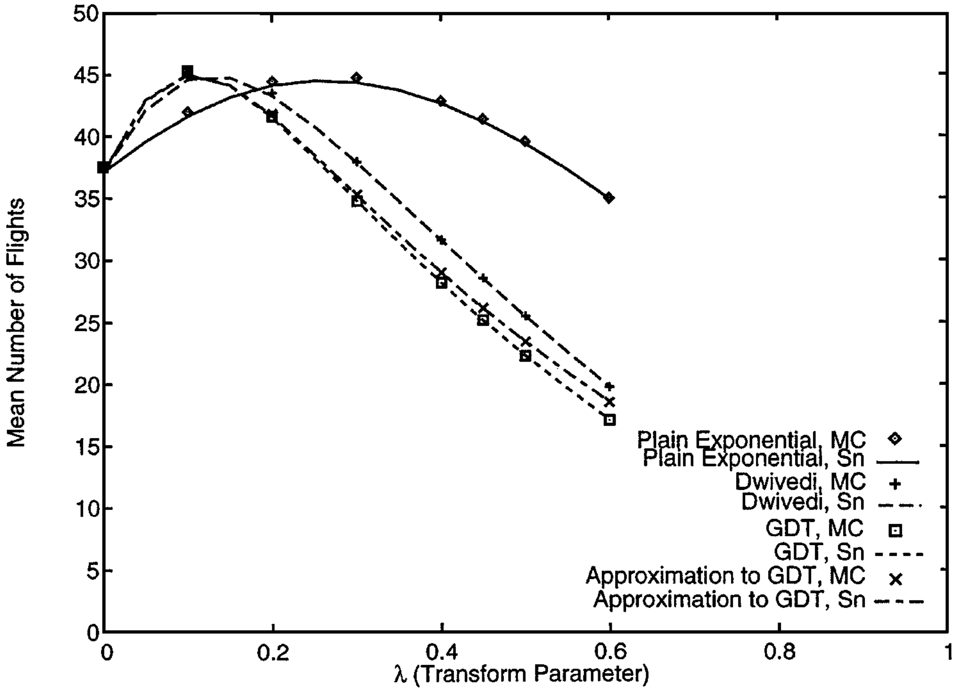

FIG. 9. Mean number of flights per history by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

Finally, we present Figure-of-Merit optimization results. In Fig. 9, the mean number of flights, which would be proportional to the cpu time if the difference between the efficiency of rejection sampling for various values of the transform parameter was negligible, is shown for $c = 0.9$ with linearly anisotropic scattering. We observe that for $\lambda \geq 0.3$, the mean number of flights in Dwivedi's transform is much smaller than that in the plain exponential transform, and that in the GDT this number is further reduced by at least 10%. (For small values of λ, the mean number of collisions per history is an increasing function of λ because the biasing scheme increasingly prevents Monte Carlo particles from leaking out the left side of the system, and hence having very short histories.) These results are expected, because of the various effects of angular biasing. We also observe that the approximate method efficiently reduces the mean number of flights. The quality factor (Eq. (61)) for this problem is shown in Fig. 10.

In sampling the direction of flights in Monte Carlo simulations, we employed direct inversion for the plain exponential transform and no transform ($\lambda = 0$), and rejection sampling for the other transforms with non-zero λ. We observe that the quality factors estimated in Monte Carlo simulations are slightly lower than those predicted by the S_N code except for the plain exponential transform. This is due to the extra computational cost of rejection sampling. The maximum quality factors of the GDT and its approximation are almost equal to each other, and both exceed that of Dwivedi's transform by more than 50%. The quality factor of the plain exponential transform is much lower than that of the other methods.

The results for quadratically anisotropic scattering are shown in Figs. 11 and 12. By comparing Figs. 9–12, we observe that the merit of the GDT is larger in quadratically anisotropic scattering than in linearly anisotropic scattering problems. In Monte Carlo

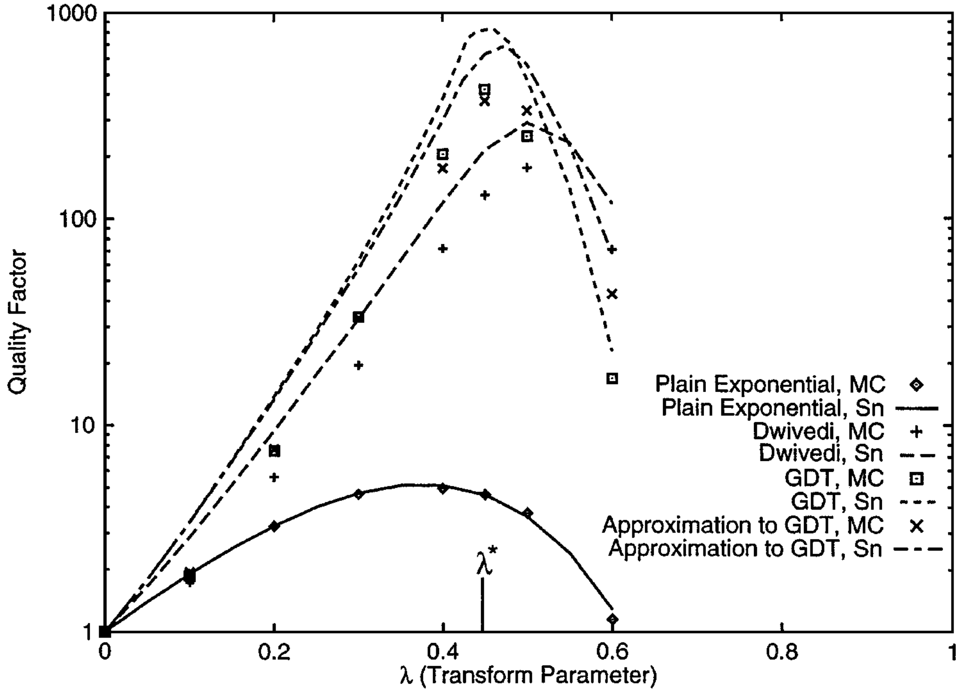

FIG. 10. Quality factor by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.3$, and $\theta_2 = 0.0$.

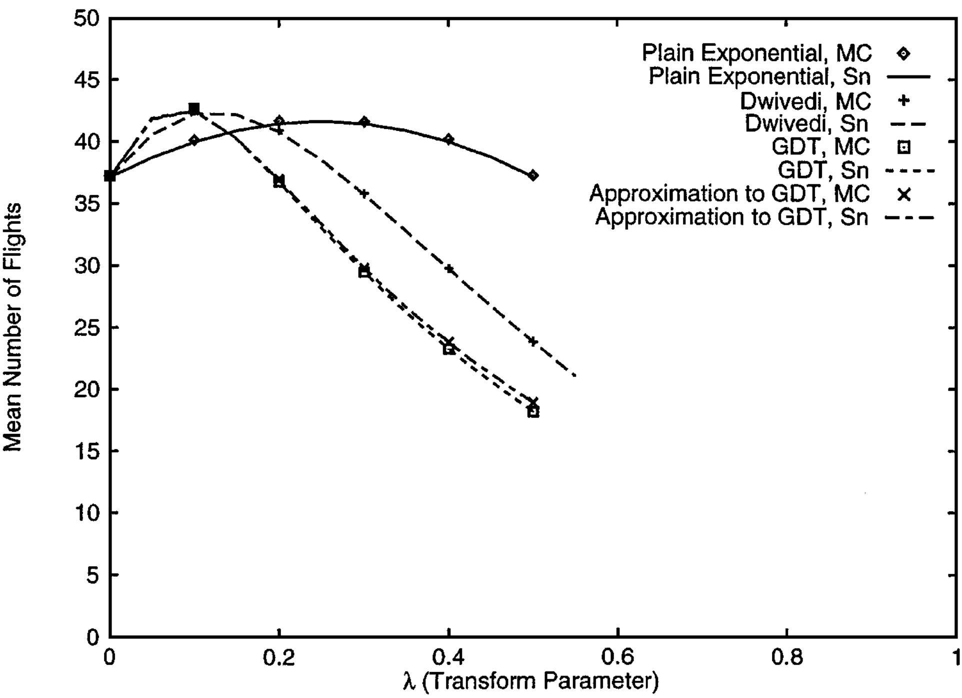

FIG. 11. Mean number of flights per history by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

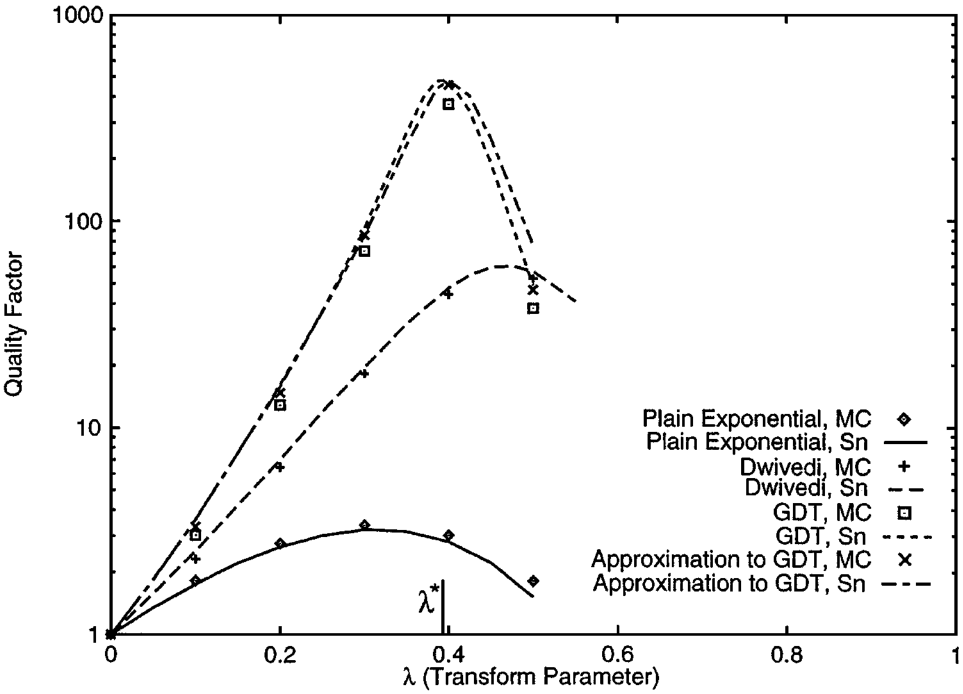

FIG. 12. Quality factor by theoretical (S_N) predictions and Monte Carlo (MC) simulations for $c = 0.9$, $\theta_1 = 0.5$, and $\theta_2 = 0.1$.

simulations for quadratically anisotropic scattering, we used rejection sampling to determine the direction of flights for all transform schemes. We could have used direct inversion for the plain exponential transform, but did not because the root selection rule of the cubic equation arising in the direct inversion of Eq. (7) with $I(\mu) = 1$ depends on the values of incident angles, μ', as well as θ_1 and θ_2, and this additional algebra makes its merit over rejection sampling small. Also, rejection sampling is a natural choice for more highly anisotropic scattering. The difference between the quality factors estimated by Monte Carlo and calculated by S_N is slight. One notable feature is that in Fig. 12, the maximum quality factor of the approximate method in Monte Carlo simulations is larger than that of the GDT. This is due to the efficient rejection sampling as stated in Section V. We believe that for general transport problems with higher-order anisotropic scattering, the approximate GDT method may yield the largest Quality Factors (hence, the largest Figures of Merit).

VIII. DISCUSSION

When $\lambda = \lambda^*$, Eq. (9) with $X = \infty$ and $I(\mu)$ defined by Eqs. (41) and (51) becomes

$$\mu \frac{\partial \Psi}{\partial x}(x, \mu) + \sigma_t (1 - \lambda^* \mu) \Psi(x, \mu) = \int_{-1}^{1} \sigma_t (1 - \lambda^* \mu') \Psi(x, \mu') f(\mu', \mu)\, d\mu', \quad 0 < x < \infty. \tag{62}$$

As $x \to \infty$, this equation has the non-zero position-independent solution,

$$\Psi(x, \mu) = \Psi(\mu) \equiv K \frac{I(\mu)}{1 - \lambda^* \mu} \int_{-1}^{1} I(\mu'') \left[1 + \sum_{m=1}^{N} (2m+1) \theta_m P_m(\mu'') P_m(\mu) \right] d\mu'',$$

TABLE I

λ^* for Various Types of Scattering and Scattering Ratios

Type of scattering	Scattering ratio (c)			
	0.98	0.9	0.7	0.5
Isotropic	0.24	0.53	0.83	0.96
Linearly anisotropic ($\theta_1 = 0.3$)	0.20	0.45	0.74	0.90
Quadratically anisotropic ($\theta_1 = 0.5, \theta_2 = 0.1$)	0.17	0.39	0.67	0.84

by Eq. (7), where K is a constant. Using Eq. (36) with $\lambda = \lambda^*$ and $\rho_0 = 1$, $\Psi(\mu)$ becomes

$$\Psi(\mu) = K'[I(\mu)]^2,$$

where $K' = 2K/c$. Therefore, by Eq. (5), when $X = \infty$, the original equation, Eq. (1), has the asymptotic solution far away from the boundary $x = 0$,

$$\psi(x, \mu) = K' e^{-\lambda^* \sigma_t x} I(\mu), \tag{63}$$

where $I(\mu)$ is defined by Eqs. (33), (41), and (51) for $N = 0, 1, 2$, respectively.

For isotropic scattering, Eq. (43) reduces to

$$1 = \frac{c}{2} \int_{-1}^{1} \frac{d\mu'}{1 - \lambda^* \mu'},$$

and by Eq. (33), Eq. (63) reduces to

$$\psi(x, \mu) = \text{const} \times \frac{e^{-\lambda^* \sigma_t x}}{1 - \lambda^* \mu}.$$

This is the “asymptotic” solution of the linear Boltzmann equation (62) [20].

Thus, the GDT method with $\lambda = \lambda^*$ is equivalent to using the asymptotic solution with the exponential spatial factor replaced by its inverse. We consider $\lambda > \lambda^*$ to be an “over-transformed” region, although $\lambda^* < \lambda < \lambda_{max}$ is a statistically “safe” region for the GDT.

In Table I, λ^* is calculated for various types of scattering and various values of the scattering ratio. We observe that in general, λ^* becomes small for highly scattering media, and highly anisotropic scattering media with small angular deflections. In Table II, λ_{max} is shown for the GDT method. The same phenomena are observed. Therefore, the range of the transform parameter on which the optimization is done becomes restrictive for highly scattering media, or for media with highly forward-peaked scattering.

TABLE II

λ_{max} in the GDT Method for Various Types of Scattering and Scattering Ratios

Type of scattering	Scattering ratio (c)			
	0.98	0.9	0.7	0.5
Isotropic	0.28	0.59	0.89	0.98
Linearly anisotropic ($\theta_1 = 0.3$)	0.24	0.51	0.82	0.95
Quadratically anisotropic ($\theta_1 = 0.5, \theta_2 = 0.1$)	0.20	0.45	0.75	0.91

To prevent overbiasing, we propose that λ^* be used as an upper limit for any type of exponential transform. For the GDT method, we also propose that λ_{max} be viewed as a statistically "safe" upper limit. Thus, the range $0 < \lambda < \lambda^*$ would be considered "very safe," while the range $\lambda^* < \lambda < \lambda_{max}$ would be considered only "safe." In the "very safe" region, the variance will usually decrease as λ increases. In the "safe" region, the variance will usually increase as λ increases, but it should remain finite. However, at the right edge of the "safe" region, the variance of the variance is very nearly (or truly) infinite. For this and larger values of λ, the Monte Carlo estimates of the variance can no longer be trusted. As λ increases beyond this value, the variance will increase and at some point will itself become infinite.

In view of this discussion, and of our numerical observations that the value of λ that maximizes the Figure of Merit is very close to λ^*, one is tempted to ask the following question: Does it make practical sense to use any value of λ other than λ^*? For the idealized problems considered in this paper, the answer is probably not. For such problems, the value $\lambda = \lambda^*$ makes theoretical sense, it is safe, and it comes very close to yielding the maximum Figure of Merit. However, the same question has a more ambiguous answer when considering generalizations of the GDT scheme and the BMC equation to multidimensional, energy-dependent transport problems. Here, it is extremely unlikely that a nonanalog scheme based on a single biasing parameter will yield adequate computational efficiency.

Extending the GDT method to multidimensional geometries can be done using similar algorithms proposed by Turner and Larsen [3, 4] and Depinay [5]. In these nonanalog schemes, the physical system is divided into disjoint subregions, and different biasing parameters are determined for each subregion by means of a deterministic adjoint calculation that is performed prior to the Monte Carlo simulation. In each subregion, one must determine a value of λ *and* a direction vector in which the nonanalog scheme will bias particles. Depinay chooses $\lambda = \lambda^*$, so that particles undergo no weight changes within a subregion, and he defines the the direction vector using the adjoint calculation. In the Local Importance Function Transform (LIFT) of Turner and Larsen, λ and the direction vector are both obtained from the adjoint calculation. In both approaches, particles undergo weight changes as they flow from one region to another. In Depinay's method, these weight changes occur only at subregion boundaries; in the LIFT method they can also occur within a subregion. However, the LIFT method is an approximation of a zero-variance method, and it is not known whether the LIFT method or Depinay's method yields a larger Figure of Merit.

Similar statements apply to extensions of the GDT scheme to energy-dependent problems; this can be done using the LIFT algorithm [3, 4] with energy-dependent biasing parameters. Thus, for problems that are more realistic than the ones treated in this paper, it may indeed be logical to choose $\lambda \neq \lambda^*$.

IX. CONCLUSIONS

We have developed a new "Boltzmann Monte Carlo" (BMC) equation, which contains weight as an extra independent variable, and which describes all of the "physics" in nonanalog Monte Carlo particle transport simulations employing the exponential transform with an arbitrary form of angular biasing in anisotropically scattering media. By taking the weight-moments of the BMC equation, we have formulated reduced problems for nth moments of the score in particle transmission problems. The spatial integration of the zeroth weight moment of the BMC solution yields the mean number of particle flights per history. Therefore, using the weight moments of the solution of the BMC equation, one can deterministically predict the variance and Figure of Merit for the original transmission problem. The predictions by our deterministic S_N code agree with the results of Monte Carlo simulations. In principle, one can use the BMC equation to optimize various nonanalog Monte Carlo schemes using the exponential transform with various forms of angular biasing.

We have also developed a new variance reduction scheme by requiring that the weight changes of a particle upon collisions be independent of the direction of flight. We call this new scheme the Generalized Dwivedi Transform (GDT), because it is a natural extension of Dwivedi's transform, originally developed for isotropic scattering. Numerical results obtained from the BMC equation and direct Monte Carlo simulations show that the GDT method is advantageous over existing exponential transforms. We have also developed an approximation to the GDT method and have shown that it is efficient for highly scattering media. This approximation may be useful for problems with high-order anisotropic scattering.

Our derivation of the GDT naturally yields the asymptotic solution of the linear Boltzmann equation for a source-free, semi-infinite, and anisotropically scattering medium. This derivation provides an upper limit for the exponential transform parameter, above which the variance of the variance is infinite.

In an earlier paper [7] we showed that an adjoint BMC equation can be formulated and that this equation is useful in evaluating responses and variances of estimated responses. We could have used such an adjoint-based theory in this paper to calculate the variance in the transmission probability, but we chose not to do this for the sake of simplicity. However, for other types of responses that do not make use of a final event estimator, it may be necessary to employ an adjoint theory in order to calculate the associated variance.

Finally, we note that specific results in this paper have been derived only for monoenergetic planar geometry transport problems with linearly and quadratically anisotropic scattering. The extension of the GDT and the BMC equation to problems with higher-order anisotropic scattering is straightforward. Although the extension of the GDT scheme to energy-dependent, multidimensional problems is less straightforward, this can be done in a fully practical way by using a *local* scheme, in which the underlying phase space is divided into subregions and within each subregion, a biasing scheme like the GDT is employed, with its own "local" biasing parameters. If there are many subregions, then many biasing parameters must be pre-determined. However, this can be done automatically, by the computer itself, using a relatively crude deterministic calculation. This is the philosophy underlying the LIFT [3, 4] biasing scheme and a very similar biasing scheme proposed by Depinay [5].

In this way, the GDT scheme may be extended to fully practical problems. There is little doubt that the corresponding BMC equation can be developed to theoretically analyze this generalized scheme. Thus, the concepts introduced here can be viewed as one of the building blocks of this more general and practical theory, which we hope to pursue in future work.

REFERENCES

1. M. H. Kalos and P. A. Whitlick, *Monte Carlo Methods. Vol. I. Basics* (Wiley, New York, 1986).
2. I. Lux and L. Koblinger, *Monte Carlo Particle Transport Methods: Neutron and Photon Calculations* (CRC Press, Boca Raton, FL, 1991).

3. S. A. Turner and E. W. Larsen, Automatic variance reduction for 3-D Monte Carlo simulations by the local importance function transform. Part I. Analysis, *Nucl. Sci. Eng.* **127**, 22 (1997).
4. S. A. Turner and E. W. Larsen, Automatic variance reduction for 3-D Monte Carlo simulations by the local importance function transform. Part II. Numerical results, *Nucl. Sci. Eng.* **127**, 36 (1997).
5. J.-M. Depinay, An automatic Monte-Carlo variance reduction scheme for the monoenergetic transport equation, in *Proceedings of the Joint International Conference on Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs, New York, October 5–9, 1997*, Vol. 1, p. 691.
6. T. Ueki and E. W. Larsen, Deterministic theory of Monte Carlo variance, *Trans. Am. Nucl. Soc*. **75**, 133 (1996).
7. T. Ueki and E. W. Larsen, A kinetic theory for nonanalog Monte Carlo algorithms: Exponential transform with angular biasing, *Nucl. Sci. Eng.*, in press; A kinetic theory for nonanalog Monte Carlo algorithms: Exponential transform with angular biasing, in *Proceedings of the Joint International Conference on Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs, New York, October 5–9, 1997*, Vol. 1, p. 701.
8. J. Spanier, An analytic approach to variance reduction, *SIAM J. Appl. Math*. **18**, 172 (1970).
9. T. E. Booth, A caution on reliability using 'optimal' variance reduction parameters, *Trans. Am. Nucl. Soc*. **66**, 278 (1992).
10. T. E. Booth, *Analytic Score Distributions and Moments for a Spatially Continuous Tridirectional Monte Carlo Transport Problem*, LA-12570, Los Alamos National Laboratory, July 1993.
11. H. J. Amster and M. J. Djomehri, Prediction of statistical error in Monte Carlo transport calculations, *Nucl. Sci. Eng*. **60**, 131 (1976).
12. P. K. Sarkar and M. A. Prasad, Prediction of statistical error and optimization of biased Monte Carlo transport calculations, *Nucl. Sci. Eng*. **70**, 243 (1979).
13. P. K. Sarkar and H. Rief, Optimization of nonanalog Monte Carlo games using differential operator sampling, *Nucl. Sci. Eng*. **124**, 291 (1996).
14. H. Rief, A synopsis of Monte Carlo perturbation algorithms, *J. Comput. Phys*. **111**, 33 (1994).
15. S. R. Dwivedi, A new importance biasing scheme for deep penetration Monte Carlo, *Ann. Nucl. Energy* **9**, 359 (1982).
16. S. R. Dwivedi, Zero variance biasing schemes for Monte Carlo calculations of neutrons and radiation transport problem, *Nucl. Sci. Eng*. **80**, 172 (1982).
17. P. Billingsley, *Probability and Measure* (Wiley, New York, 1995), 3rd ed., pp. 81, 277.
18. P. W. Glynn and W. Whitt, The asymptotic efficiency of simulation estimators, *Oper. Res.* **40**, 505 (1992).
19. E. E. Lewis and W. F. Miller, *Computational Methods of Neutron Transport* (Wiley, New York, 1984; reprinted, American Nuclear Society, Inc., La Grange Park, IL, 1993).
20. K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison–Wesley, Reading, MA, 1967), Chap. 4.