SANDIA REPORT

SAND2022-12721
Printed September 2022

Sandia National Laboratories

Sensitivity Analyses for Monte Carlo Sampling-Based Particle Simulations

S. D. Bond, B. C. Franke, R. B. Lehoucq, J. D. Smith
and S. A. McKinley (Tulane University)

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185
Livermore, California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods

ABSTRACT

Computational design-based optimization is a well-used tool in science and engineering. Our report documents the successful use of a particle sensitivity analysis for design-based optimization within Monte Carlo sampling-based particle simulation—a currently unavailable capability. Such a capability enables the particle simulation communities to go beyond forward simulation and promises to reduce the burden on overworked analysts by getting more done with less computation.

Gradient-based methods crucially depend upon sensitivities, which are synonymous with the calculation of a derivative that measures the (instantaneous) change in a quantity with respect to an (instantaneous) change in another quantity. We exploit the link between the deterministic and particle models to pose an equivalent optimization problem containing a stochastic differential equation as a constraint. We then demonstrate that the sensitivities can be approximated for the gradient-based optimization problem by reusing the particle trajectories employed for the existing Monte Carlo approximation. Additional forward simulations—as in a finite difference approach—are unnecessary.

Our approach is general and so enables us to impact a broad spectrum of Monte Carlo particle simulation codes by a clean separation of the mathematical and the application-specific details. Our general approach develops capability largely unexplored as a Sandia analysis tool, and can be leveraged across several mission applications.

Acknowledgements

The authors acknowledge helpful conversations with Drew Kouri and Aaron Olson of Sandia National Laboratories, Tim Kelley of North Carolina State University, Petr Pechac of the University of Delaware, Ilham Variansyah and Ryan McClarren of the University of Notre Dame, and Todd Palmer of the Oregon State University.

The author Scott McKinley's affiliation is the Department of Mathematics, Tulane University, 6823 St Charles Ave, New Orleans, LA 70118.

Our report was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Adminstration under contract DE-NA0003525.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Computational design-based optimization is a well-used tool in science and engineering. Our report documents the successful use of a particle sensitivity analysis for design-based optimization within Monte Carlo sampling-based particle simulation—a currently unavailable capability. Such a capability enables the particle simulation communities to go beyond forward simulation and promises to reduce the burden on overworked analysts by getting more done with less computation. Gradient-based methods crucially depend upon sensitivities, which are synonymous with the calculation of a derivative that measures the (instantaneous) change in a quantity with respect to an (instantaneous) change in another quantity.

An important class of optimization problems contain partial integral differential equations as constraints so that gradient-based methods require sensitivities for the various functionals involving the solution of these equations. Such problems include models for plasmas, radiation transport, low-density fluids where the partial integral differential equation is the Boltzmann equation or the Fokker-Planck equation when molecular motion is of interest. These equations embody a deterministic model for the aggregate behavior of particles where the solution represents the density of particles. For an important class of partial integral differential equations, however, a stochastic model can be used to approximate the density of particles.

The stochastic, or equivalently, particle model represents a statistical approach. We exploit the link between the deterministic and particle models to pose an equivalent optimization problem containing a stochastic differential equation as a constraint. We then demonstrate that the sensitivities can be adequately approximated for the gradient-based optimization problem by reusing the particle trajectories available for the existing Monte Carlo approximation. Additional forward simulations—as in a finite difference approach—are unnecessary. This obviates the need to determine an optimal step-size or the complications that arise for multi-parameter optimization.

Our approach is general and so enables us to impact a broad spectrum of Monte Carlo particle simulation codes by a clean separation of the mathematical and the application-specific details. Our general approach develops capability largely unexplored as a Sandia analysis tool, and can be leveraged across several mission applications.

Our report is organized as follows. Section 2 introduces the optimization problem of interest at a high-level. Section 3 provides more details for a model problem originating in radiation transport including a demonstration calculation. The high-level presentation described §2 is in large part motivated by the somewhat more involved discussion of §3 and is substantially developed in a manuscript to be submitted for publication. Section 4 describes an optimization problem originating in equilibrium molecular dynamics and compares the reuse of trajectories to approximate a sensitivity with a finite difference approach. Section 5 describes a stochastic calculus approach

to Boltzmann linear transport. This is the first such treatment for linear transport with important consequences for sampling efficiency and will be submitted for publication.

2. A GRADIENT-BASED CONSTRAINED OPTIMIZATION PROBLEM

We first review the constrained optimization problem of interest in §2.1. We first consider the deterministic formulation, or equivalently a partial integral differential equation (PIDE) constrained optimization problem. We then introduce a stochastic formulation, what we introduce as a SDE constrained optimization problem amenable to Monte Carlo estimation. The class of PIDE includes important DOE mission problems for plasmas, radiation transport, low-density fluids where the PIDE is the Boltzmann equation or the Fokker-Planck equation when molecular motion is of interest.

A gradient-based optimization approach requires sensitivities. We explain how the sensitivities, and those for the adjoint formulation, can be approximated by reusing the particle trajectories employed for the Monte Carlo approximation in §2.2.

We confess that in order the convey the main ideas, our presentation is informal and focused upon the crucial relationships. Section3 provides more details for a model problem originating in radiation transport. The high-level presentation described in this section is in large part motivated by the more involved discussion of §3.

2.1. Constrained optimization problem

Our interest is in the values of θ that minimize the least-squares functional

$$\frac{1}{2}\sum_{i=1}^{m}\left(q_i - \int_0^{t_*}\int_\Omega q(x;\theta)\,u(t,x;\theta)\,dx\,dt\right)^2 \tag{2.1}$$

using gradient-based optimization methods. The residual is the difference between observed values q_i, typically obtained via measurement and a response functional modeling an observation. The density q selects the important portion of u needed for the observation q_i. Both the density q and u are parameterized by the vector θ.

The function u is also nonnegative and satisfies the PIDE

$$\begin{cases} \dfrac{\partial}{\partial t}u + L^* u = 0 & \text{over } (0,t_*)\times\Omega, \\ u(t,x;\theta) = 0 & (t,x)\in(0,t_*)\times\partial\Omega \\ u(0,x;\theta) = f(x;\theta) & x\in\Omega\,. \end{cases} \tag{2.2a}$$

where f is a density and L^* is the formal adjoint for the backward Kolmogorov operator L defined by

$$L\,u(t,x) = \frac{a(x)}{2}\Delta u(t,x) + b(x)\cdot\nabla u(t,x) + \sigma(x)\,u(t,x) + \int \big(u(t,x+\gamma(x,y)) - u(t,x)\big)\eta(y,x)\,dy\,. \quad (2.2b)$$

A related PIDE contains the density q as the initial condition.

$$\begin{cases} \dfrac{\partial}{\partial t}w - L\,w = 0 & \text{over } (0,t_*)\times\Omega, \\ w(t,x;\theta) = 0 & (t,x)\in(0,t_*)\times\partial\Omega \\ w(0,x;\theta) = q(x;\theta) & x\in\Omega\,. \end{cases} \quad (2.2c)$$

Both PIDEs (2.2a) and (2.2c) are deterministic models for diffusion, drift, absorption and jump-diffusion. We remark that for a steady-state formulation, $u(t,x) = u(x)$ is the equilibrium distribution so that the initial condition f can be identified as a source for the steady-state problem.

A particle model is given by the stochastic differential equation (SDE)

$$dX_t = b(t,X_t)\,dt + \sqrt{a(t,X_t)}\,dW_t + \gamma(t,X_t)\,dP_t \qquad 0 < t \leqslant \min(\tau_{\text{abs}},\tau) \quad (2.3)$$

where the random variable τ_{abs} is an exponentially distributed absorption time with rate σ, and τ is the exit-time when the particle hits the boundary $\partial\Omega$. The variables W_t and P_t are Wiener and Poisson processes, respectively.

The representation (i.e., Feynman-Kac) formulas

$$w(t,x;\theta) = \mathbb{E}^x\left[q(X_t;\theta)\,e^{\int_0^\tau \sigma(X_s)\,ds}\right] = \mathbb{E}[q(X_t;\theta)\,e^{\int_0^\tau \sigma(X_s)\,ds}\Big|X_0 = x] \quad (2.4a)$$

$$u(t,x;\theta) = \mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde\sigma(Y_s)\,ds}\right] = \mathbb{E}\Big[f(Y_t;\theta)\,e^{\int_0^\tau \tilde\sigma(Y_s)\,ds}\Big|Y_0 = x\Big] \quad (2.4b)$$

relate the deterministic (2.2) and particle (2.3) models where $\tilde\sigma$ is the absorption function associated with L^*, and the dual process Y_t is given by

$$Y_t = X_{t_*-t} \quad (2.4c)$$

for $0 \leqslant t \leqslant t_*$. In other words, the process Y_t is the process X_t run backwards in time. The relationships (2.4a)–(2.4c) imply that the density of particles u and w can be approximated with a Monte Carlo estimator using the SDE or via a numerical method for the PIDE.

By exploiting the representation formulas (2.4a)–(2.4c) with $X_0 \sim f$ and $Y_0 \sim q$, we can express the adjoint relationship

$$\int_0^{t_*}\int_\Omega q(x;\theta)\,u(t,x;\theta)\,dx\,dt = \int_0^{t_*}\int_\Omega f(x;\theta)\,w(t,x;\theta)\,dx\,dt \quad (2.5a)$$

as

$$\int_0^{t_*}\int_\Omega q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt \\ = \int_0^{t_*}\int_\Omega f(x;\theta)\,\mathbb{E}^x\left[q(X_t;\theta)\,e^{\int_0^\tau \sigma(X_s)\,ds}\right]dx\,dt\,. \quad (2.5b)$$

Both formulations of the response functional are useful. For instance, simulating the SDE for various initial conditions f can be avoided in favor of simulating X_t for a given selector function q; see Giles and Pierce [2000] for an introduction on the role of adjoints within design-based optimization for the deterministic model.

We now pose the SDE constrained optimization

$$\begin{cases} \min_\theta \dfrac{1}{2}\displaystyle\sum_{i=1}^m \left(q_i - \int_0^{t_*}\int_\Omega q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt\right)^2 \\ \text{subject to the particle trajectory } Y_t \text{ satisfying (2.4c) with } Y_0 \sim q. \end{cases} \quad (2.6)$$

The relationship (2.4a) then implies that the optimization problem (2.6) can be equivalently expressed as

$$\begin{cases} \min_\theta \dfrac{1}{2}\displaystyle\sum_{i=1}^m \left(q_i - \int_0^{t_*}\int_\Omega q(x;\theta)\,u(t,x;\theta)\,dx\,dt\right)^2 \\ \text{subject to } u \text{ satisfying the PIDE (2.2a).} \end{cases} \quad (2.7)$$

This latter formulation of the optimization problem is a conventional approach; see los Reyes [2015] for an introduction to PDE constrained optimization. A standard approach in gradient-based optimization algorithms is to use an adjoint method, which can now be easily accomplished via the equalities (2.5).

Justification in favor of the SDE constrained optimization problem (2.6) is that q and f are often localized so that the solution (2.4a) is needed only over a small portion of Ω. Hence a Monte Carlo approach avoids the discretization of a PIDE (2.2a) or (2.2c), a potentially daunting numerical computation. Two examples occur when the Boltzmann equation is considered over non-trivial geometries and for optimization problems originating in equilibrium molecular dynamics; see § 4 where a deterministic approach leads to an extremely large dimensional Fokker-Planck equation.

An interesting question is whether the relationship (2.4c) between X_t and Y_t enables us to simulate both simultaneously. We show in §5 that for the specific case of linear transport, the answer is yes. And so a further justification in support of the SDE constrained optimization problem (2.6) is that the solution of both (2.2a) and (2.2c) is avoided.

2.2. Monte Carlo sensitivities

Gradient-based optimization approaches for problem (2.6) depend upon gradients—sensitivities—with respect to θ in the least squares functional, i.e.,

$$\begin{aligned}\nabla_\theta \frac{1}{2}\sum_{i=1}^{m}\left(q_i - \int_0^{t_*}\int_\Omega q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt\right)^2 \\ = \sum_{i=1}^{m}\left(q_i - \int_0^{t_*}\int_\Omega q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt\right) \\ \times \nabla_\theta \int\int q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt\,.\end{aligned}$$

so that the crucial sensitivity is

$$\begin{aligned}\nabla_\theta \int_0^{t_*}\int_\Omega q(x;\theta)\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt = \int_0^{t_*}\int_\Omega \left(\nabla_\theta\, q(x;\theta)\right)\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt \\ + \int_0^{t_*}\int_\Omega q(x)\,\nabla_\theta\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]dx\,dt\,. \quad (2.8)\end{aligned}$$

We remark that the use of an adjoint gradient-based optimization algorithm is easily accomplished via the equalities (2.5). The challenge is to approximate the last sensitivity. We now explain three different approximations or what amounts to estimators for the last sensitivity. The third estimator enables us to reuse the trajectories used to estimate $\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right]$; this is an extremely important reuse since the application groups already using Monte Carlo sampling have well-established codes for estimating these expectations.

We may now consider three classes of estimators (see e.g., [Asmussen and Glynn, 2007, chap.VII]) suggested by the equalities

$$\begin{aligned}\nabla_\theta\,\mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\right] &= \mathbb{E}^x\left[\tilde{f}'(Y_t;\theta)\,\frac{\mathrm{d}Y_t}{\mathrm{d}\theta}\right] + \mathbb{E}^x\left[\frac{\partial}{\partial\theta}\tilde{f}(Y_t;\theta)\,Y_t\right] \\ &= \mathbb{E}^x\left[f(Y_t;\theta)\,e^{\int_0^\tau \tilde{\sigma}(Y_s)\,ds}\,Z_t\right]\end{aligned} \quad (2.9)$$

where $\tilde{f}(y) = f(y;\theta)\,e^{\int_0^\tau \tilde{\sigma}(y)\,ds}$. They are

1. a finite-difference approximation to the first expectation;
2. a derivative of f and the particle trajectories for the second expectation (also referred to as infinitesimal perturbation analysis or a pathwise derivative approach);
3. what amounts to a formal integration by parts for the third expectation (also referred to as likelihood ratio method or a Malliavin estimator or what we introduce as Stein equation estimators).

The third expectation depends upon the same measure as $\mathbb{E}^x\left[f(Y_t;\theta)\, e^{\int_0^\tau \tilde{\sigma}(Y_s)\, ds}\right]$. Therefore the Monte Carlo scheme used to approximate the latter expectation can be reused. The practical impact of this mathematical statement is significant—the nontrivial understanding involved in Monte-Carlo approaches for approximating the expectation $\mathbb{E}^x\left[f(Y_t;\theta)\, e^{\int_0^\tau \tilde{\sigma}(Y_s)\, ds}\right]$ can be applied to approximate the sensitivity and no additional forward simulations are necessary. Moreover, this obviates the need to determine an optimal step-size or the complications that arise when more than one parameter needs to be varied. The papers Fournié et al. [1999, 2001] also establish that the third class of estimators are of minimal variance for Brownian motion. We caution the reader that in practice, the decision on which class of estimators to use depends upon the sensitivity of interest among several issues; see Sheppard et al. [2012]. Section 4.4 compares the finite-difference and likelihood ratio methods on a model problem.

The interested reader is referred to the book [Glasserman, 2004, Chap.7] for applications to finance and [Asmussen and Glynn, 2007, Chap.VII] for a general treatment. These tools also include stochastic calculus based Malliavin estimators and what can also be introduced as Stein [1973] equation estimators. The former estimators originated in the mathematical probability literature to obtain regularity estimates of the probability measure in terms of the stochastic process [Malliavin, 1976]; these estimators have a well-understood application within mathematical finance; see e.g., [Chen and Glasserman, 2007]. The latter estimators originate in the mathematical statistics literature to obtain bounds on the distance between two probability distributions. See also the application-focused papers [Rief, 1994],[Plyasunov and Arkin, 2007],[Warren and Allen, 2012] that are interested in the reuse of sample trajectories.

3. A MONTE CARLO GRADIENT-BASED INVERSE METHOD FOR A LINEAR TRANSPORT PROBLEM

Consider a set of particles that each emerge from a random location in some source and travel at a random constant velocity until they are absorbed by the surrounding medium or hit the boundary of some domain Ω. We can model particle trajectories by the random ordinary differential equation (ODE)

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = \mu, \quad X(0) = \xi \sim h, \quad \mu \sim \nu \quad t < \tau_{\text{abs}} \tag{3.1}$$

where $h : \Omega \to \mathbb{R}_+$ and $\nu : \mathbb{R} \to \mathbb{R}_+$ are probability densities for the initial location and velocity. The random variable τ_{abs} is the absorption time, which we assume to be governed by a non-homogeneous (location-dependent) absorption rate function $\sigma : \Omega \to \mathbb{R}_+$. Throughout this work, we will assume that we can observe particles at a few locations, but the entire trajectory is not available for observation. The goal of the analysis is to find methods for inferring the location and scale of the source distribution given these restrictions on particle observation. In particular, we will consider the problem where we can only observations at the boundary of the domain, but we can nevertheless infer information about the source.

The standard approach used for inference problems models the location-dependent steady-state concentration of particles by the boundary value problem

$$\begin{cases} \partial_x(\mu\phi(x,\mu)) + \sigma(x)\phi(x,\mu) = s(x,\mu), \quad (x,\mu) \in (0,\ell)\times\mathbb{R}\setminus\{0\}, \\ \quad \phi(0,\mu) = 0,\ \mu > 0; \qquad \phi(\ell,\mu) = 0,\ \mu < 0 \end{cases} \tag{3.2}$$

where $s(x,\mu)$ is a term that summarizes the particle-source properties. The relationship between s and the particle densities h and ν will be developed later. Note that we do not prescribe boundary values for $\mu = 0$ because the solution takes on the special form $\phi(x,0) = s(x,0)/\sigma(x)$. A BVP like this has appeared, for example, as a model for one-speed radiation transport in a critical slab where the interaction is absorption with the background; see, e.g., Bell and Glasstone [1970]. The function ϕ is viewed as the distribution of the angular flux. The distribution of μ arises from the projection of the particle velocity on to the longitudinal direction of interest.

To model our observation process, we introduce a detector function $\Lambda : \Omega \to \mathbb{R}_+$ and a value $a > 0$. Define

$$\Phi(\Lambda) \coloneqq \int_\Omega \int_{\mathbb{R}\setminus\{0\}} \mu^a \phi(x,\mu)\Lambda(x)\mathrm{d}\mu\mathrm{d}x. \tag{3.3}$$

If $a = 0$ and $\Lambda(x) = \mathbb{1}_{[x_L,x_R]}(x)/(x_R - x_L)$, then $\Phi(x)$ is the average concentration of particles over the interval $[x_L, x_R]$. If $a = 1$ then $\Phi(x)$ is the current over $[x_L, x_R]$.

For our inverse problem of interest, suppose that there is some parameter θ that affects the source of the particles ($s(x) = s(x;\theta)$) and also possibly the absorption rate ($\sigma(x) = \sigma(x;\theta)$). There is

an unknown "true" source parameter value θ_* and given a set a detector functions $\{\Lambda_m\}_{m=1}^M$, we assume that the set of measurables we can observe for our system are

$$C_m = \Phi(\Lambda_m\,;\theta_*) \quad m \in \{1,\dots,M\}. \tag{3.4a}$$

Typically the detectors will be located on or near points in the boundary of the domain. In our one-dimensional setting, there will be two detectors set near $x = 0$ and $x = \ell$. One particular problem we consider is to take an appropriately defined sequence of detectors $\{\Lambda_m^{(i)}\}$ with decreasing width and constant mass so that $\lim_{i\to\infty}\Lambda_m^{(i)}(x) = \delta_{x_m}(x)$, where the right-hand side is a Dirac δ-function centered at the point x_m. Then the measurable C_m is a μ-weighted average of ϕ evaluated at a point x_m.

Whatever the set of measurables, the problem of estimating θ_* from $\{C_m\}_{m=1}^M$ can be articulated as a PDE-constrained optimization problem:

$$\begin{aligned}&\min_\theta \frac{1}{2}\sum_{m=1}^{m}\big(C_m - \Phi(\Lambda_m\,;\theta)\big)^2\\ &\text{subject to } \phi \text{ satisfying the BVP (3.2).}\end{aligned} \tag{3.4b}$$

Such a problem can be addressed through gradient-based inverse methods; see, e.g., los Reyes [2015], but when the detectors are δ-functions, this approach does not exploit that the BVP solution ϕ is only needed at a few points.

The contribution of our paper is to pose and develop a Monte Carlo approach to solve (3.4b) by instead exploiting an equivalent formulation using the random ODE (3.1). We will show that the source function $s(x,\mu\,;\theta)$ is related to the properties of the random ODE through the relationship $s(x,\mu\,;\theta) = c_0 h(x\,;\theta)\nu(\mu)$ for some $c_0 > 0$. Moreover, defining τ to be the time that a particle either exits the domain or is absorbed by the medium, we will show that

$$\begin{aligned}&\min_\theta \frac{1}{2}\sum_{i=1}^{n}\Big(C_m - c_0 E_\theta\Big(\int_0^\tau \mu^a\,\Lambda_m\big(X(t)\big)\mathrm{d}t\Big)\Big)^2\\ &\text{subject to } X(t) \text{ satisfying the random ODE (3.1).}\end{aligned} \tag{3.5}$$

is equivalent to the formulation (3.4b). In the above, E_θ means that we take mathematical expectation using the value θ in the particle source distribution and the absorption rate function.

It is already standard practice to use collections of simulated particle trajectories $\{X_n\}_{n=1}^N$ to produce Monte Carlo approximations for quantities like C_m. In our case,

$$\begin{aligned}C_m &= c_0 E_{\theta_*}\Big(\int_0^\tau \mu^a\,\Lambda_m\big(X(t)\big)\mathrm{d}t\Big)\\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} c_0\int_0^{\tau_n}\mu^a\,\Lambda_m\big(X_n(t)\big)\mathrm{d}t.\end{aligned} \tag{3.6}$$

The latter equality follows from the strong law of large numbers and the limit is an almost sure limit.

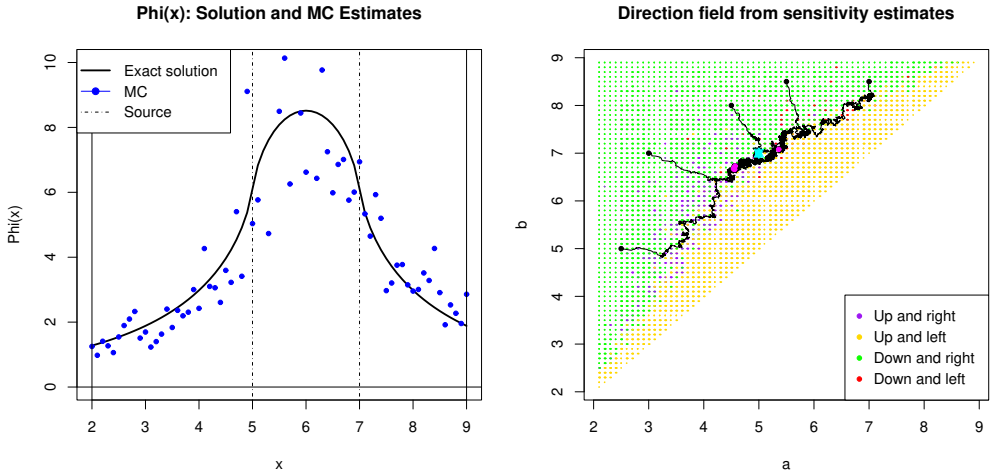


Figure 3-1. An example of Monte Carlo samples used to solve an inverse problem in which information is known only at the boundary, but the size and location of the source can be inferred. Here, the source of particles is uniformly distributed over the interval $[a,b] = [5,7]$. We assume constant velocity and absorption rate. Detectors are placed at $x = 2$ and $x = 9$. Left. Exact solution over $x \in [2,9]$ (black curve) and Monte Carlo estimates at various values of x (blue circles) each using $N = 100$ independent samples. Right. A color-coded direction field constructed using Monte Carlo sensitivity estimates ($N = 100$) at each point. Each black trajectory is a random walk (start point, black; end point, magenta) guided by the local gradients in the direction field. Despite the small sample sizes, the walkers quickly find a neighborhood of the true values for the source endpoints, a and b.

Given this relationship between particle simulations and the measurables of interest, we demonstrate how to exploit established methods of stochastic differentiation to re-use the particle simulations to simultaneously produce estimates for

$$\nabla_\theta E_\theta\Big(\int_0^\tau \mu^a \Lambda_m\big(X(t)\big)\mathrm{d}t\Big). \tag{3.7}$$

for any proposed parameter value θ. This coupled calculation of the measurable and its gradient for any given value of θ immediately invites the use of stochastic gradient descent methods to search for the true value θ_*.

To the best of our knowledge, a particle approach to solve (3.5) is unavailable in the literature. Our approach is sufficiently general, ultimately relying upon the probabilistic basis for an important class of partial differential integral equations that the reader will see the potential to other deterministic systems equivalently formulated as stochastic system.

The benefits of a particle over a deterministic approach is that former exploits that ϕ may only be needed at a small number of points. Moreover, if the two-dimensional PDE for the BVP (3.2) is of three or more dimensions and σ varies substantially, then the numerical solution of the BVP is computationally intensive because of the refined mesh needed. Although Monte Carlo schemes are beholden to sampling error, the restriction of computation to small neighborhoods of particle locations significantly decreases number of σ-evaluations.

In Figure 3-1 we share the result of applying the method to a simple problem. We assume that the source of the particles in uniformly distributed over a given interval $[a,b]$. We treat $\theta = (a,b)$

as the parameter vector of interest. We assume that the overall rate of particle production scales with the source interval size, $c(a,b) = c_0(b-a)$, and that the velocity distribution is ν is uniformly distributed over the interval $[-1,1]$. We assume a constant absorption rate σ. At left in the figure is the exact solution for the location-dependent concentration of particles $\Phi(\delta_x) = \int_{-1}^{1} \phi(x,\mu)\mathrm{d}\mu$ (black curve) when $[a,b] = [5,7]$, $c_0 = 4$, and $\sigma = 0.2$. The blue circles indicate Monte Carlo estimates at several values of x. To compute the MC estimate at a given value of x, we give the nth simulated particle trajectory a score of $\chi_n = 1$ if the particle passes the point x before being absorbed and a score of $\chi_n = 0$ if not. In our work, we are able to show that if the detector is a Dirac-δ function located at a position x, the MC estimate for the concentration of particles at x is

$$\Phi(\delta_x) = \frac{c_0}{N} \sum_{n=1}^{N} \frac{1}{\mu} \chi_n$$

In Figure 3-1 we used only $N = 100$ samples to demonstrate that despite noisy estimation of the original function, we can still succeed in the solving the inverse problem.

Turning to estimating the sensitivity, it might seem to be a concern that χ is a step function. However, using the Malliavin derivative approach which relies on differentiating the probability measure underlying the expectation rather than the function itself, we can circumvent this issue and estimate $\nabla_\theta E_\theta(\chi)$ using the same function simulations $\{\chi_n\}_{n=1}^{N}$. As a proof of concept for a stochastic gradient descent method, we computed MC estimates for the gradient of the target function for a wide range of proposed (a,b) combinations and then simulated random walks that are biased by the local gradients. In Figure 3-1 we color-coded the local gradient estimations and showed five random walk traces which begin in different places (black dots) but quickly arrive near the true value (cyan star). This plot is for expository purposes: we do not propose this as an eventual method of inverse problem estimation. However, the plot makes clear the effectiveness of the MC method despite a limited number of samples, and we believe that there is tremendous potential for future theoretical development on the stochastic gradient descent aspect of this problem.

4. PARAMETER SENSITIVITY ANALYSIS FOR EQUILIBRIUM MOLECULAR DYNAMICS USING BROWNIAN DYNAMICS SIMULATIONS

We consider the task of designing an interatomic molecular dynamics potential to match macroscopic physical properties. Macroscopic physical properties of a molecular dynamics system can be expressed as expectations of functionals of microscopic trajectories. For example, at equilibrium, the equipartition theorem states that the temperature of the system can be computed from the expectation

$$T = \frac{1}{3k_B}\langle x_i \cdot \frac{\partial}{\partial x_i} U(\vec{x})\rangle,$$

where T is temperature, k_B is Boltzmann's constant, $x_i \in \mathbb{R}^3$ is the position of the ith particle, and U is the potential. A second example is the self-diffusion coefficient that can be computed from the integral of the velocity auto-correlation function,

$$D = \frac{1}{3}\lim_{t_f \to \infty} \int_0^{t_f} \langle v_i(0) \cdot v_i(\tau)\rangle d\tau,$$

where $v_i(t) = \dot{x}_i(t)$ is the velocity of the ith particle at time t.

To express the design task as an optimization problem, suppose we have a system of N particles evolving according to the (Brownian dynamics) stochastic differential equation

$$\gamma dX_t = F(X_t;\theta)\, dt + \sqrt{2\gamma T}\, dW_t \tag{4.1}$$

where $X_t \in \mathbb{R}^{3N}$ are the positions of the N particles, T is the temperature in energy units, γ is a damping parameter, and W is a vector of Wiener processes. The deterministic force among the N particles is given by

$$F(\vec{x};\theta) = -\nabla_{\vec{x}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} U_{ij}\left(\|x_i - x_j\|, \theta\right).$$

with a potential energy function $U_{ij}\left(\|x_i - x_j\|, \theta\right)$ parameterized by N_p parameters, $\theta \in \mathbb{R}^{N_p}$.

Now, suppose that we have N_c macroscopic quantities given by expectations of functionals over the particle trajectories, e.g., $\langle A_k\left(\vec{x}(t), \dot{\vec{x}}(t); \theta\right)\rangle$ and the $\bar{A}_k$ are known target values.

An instance of the SDE constrained optimization problem (2.6) is

$$\begin{cases} \min_{\theta} \dfrac{1}{2} \displaystyle\sum_{k=1}^{N_c} \left\| \bar{A}_k - \langle A_k\left(\vec{x}(t), \dot{\vec{x}}(t); \theta\right)\rangle \right\|^2 \\ \text{subject to } \vec{x} \text{ satisfying the SDE (4.1)}, \end{cases} \tag{4.2}$$

provides an estimate of the vector θ of parameters. A gradient-based optimization method requires derivatives of the expectations of the functionals with respect to the parameters θ. These derivatives are known as sensitivities.

An instance of the PIDE (2.2a) is given by the Fokker-Planck equation

$$\gamma \frac{\partial \rho}{\partial t} = \nabla \cdot [\nabla U(\vec{x})\, \rho] + k_B T \nabla \cdot \nabla \rho \tag{4.3}$$

associated with the Brownian dynamics equations (4.1). The solution of this equation is $\rho(\vec{x}, t)$, the probability density of finding the system in configuration $\vec{x} \in \mathbb{R}^{3N}$ at time t. To find the stationary (equilibrium) distribution, we set the time derivative of ρ to zero which results in

$$0 = \nabla \cdot [\nabla U(\vec{x}) + \rho k_B T \nabla \rho] = \nabla \cdot [\nabla U(\vec{x}) + k_B T \nabla \ln \rho] .$$

Using appropriate boundary and normalization conditions for ρ, we recover the equilibrium solution, or canonical ensemble

$$\rho(\vec{x}) = \frac{1}{C} \exp\left[-\frac{1}{k_B T} U(\vec{x})\right], \tag{4.4}$$

where C is the partition function. Due to the high-dimensionality of configuration space ($3N$ for N particles) it is computionally intractable to solve the Fokker-Planck equation (4.3) for large systems. Similarly, calculation of equilibrium averages with respect to the canonical ensemble (4.4) using traditional quadrature methods is prohibitively computationally expensive. For example, the cost of a traditional Cartesian-product quadrature rule with N_q points in each coordinate direction scales as N_q^{3N}, which is exponential in N. It is for this reason that random sampling methods (e.g. Metropolis Monte Carlo or Brownian Dynamics) are used to approximate molecular dynamics averages.

In the remainder of this section we focus on computing the sensitivities of a radial distribution function computed from Brownian dynamics simulations. The associated optimization problem is known as the Boltzmann inversion problem.

4.1. The Lennard-Jones potential

Consider the Lennard-Jones potential,

$$U_{ij}(r) = 4\epsilon \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \theta \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right),$$

where $r_{ij} = \|x_i - x_j\|$ is the distance between particles i and j, and θ is a parameter that controls the relative strength of the attractive force (typically set to one). The length and energy scale parameters, σ_{ij} and ϵ, are one in Lennard-Jones units. The Lennard-Jones potential is used to model simple liquids and gasses like Argon. In Figure 4-1 we plot the Lennard-Jones potential in Lennard-Jones units with $\theta = 1$.

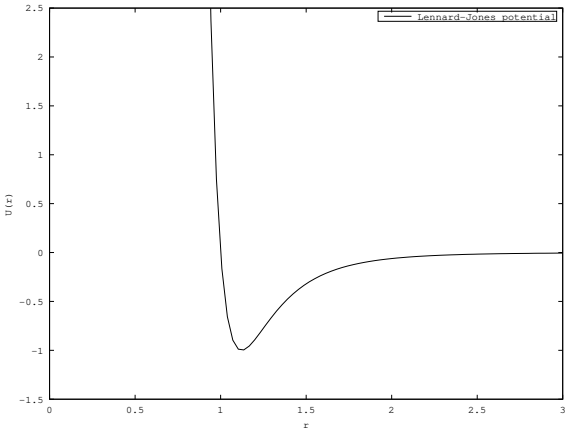


Figure 4-1. The Lennard-Jones potential with $\theta = 1$

4.2. Radial density function

For this report, we are interested calculating the sensitivity of the radial density function with respect to changes in the interaction potential. The radial density function is the density of particles at distance r from a given particle. Note that the radial density function divided by the bulk density is known as the radial distribution function. Here, we consider the radial density function,

$$\mathrm{RDF}(r) = 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{k=1}^{N_b}\frac{1}{V_k}\psi_k\left(\|x_j - x_i\|\right),$$

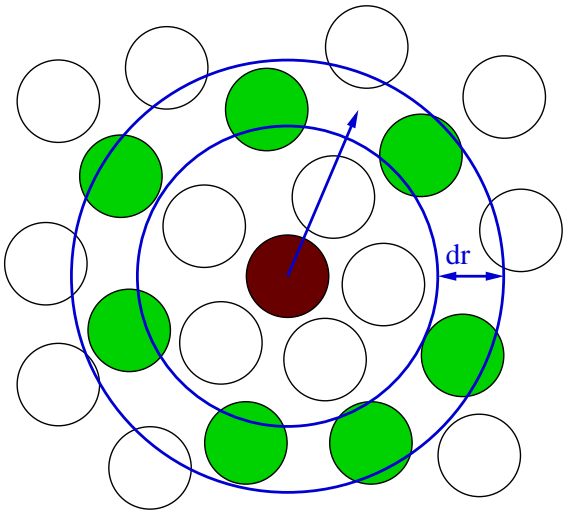


Figure 4-2. An illustration showing how the radial density function is calculated. The space around a red particle is decomposed into N_b spherical shells of thickness dr. The number density of green particles with centers within each shell is computed. The radial density function is the average of this quantity over all N center (red) particles. As the shell radius increases, the radial density function converges to the bulk density of the system.

where the histogram bin functions are

$$\psi_k(r) = \begin{cases} 1, & -\Delta r/2 \le r - r_k < \Delta r/2, \\ 0, & \text{otherwise}, \end{cases}$$

and the volume of each spherical shell is

$$V_k = \frac{4}{3}\pi(r_k + \Delta r/2)^3 - \frac{4}{3}\pi(r_k - \Delta r/2)^3.$$

The bins are centered at $\{r_1, r_2, \cdots, r_{N_b}\}$ and have uniform width of Δr.

For efficiency, the radial density function can be computed at the same time as the force and potential. Because the bins are non-overlapping, each particle-pair distance is assigned to at most one bin.

4.3. Equilibrium sensitivities

4.3.1. Correlated finite differences

A simple method for estimating sensitivities is to run two simulations with two difference parameter values and use finite differences to approximate the derivative,

$$\frac{d}{d\theta}\langle A\rangle \approx \frac{1}{\Delta\theta}\left(\langle A\rangle_{\theta+\Delta\theta} - \langle A\rangle_{\theta}\right).$$

The optimal finite difference step, $\Delta\theta$, depends on quantity of interest, number of samples, variance, ... The statistical accuracy of this approach can be improved by using the same random number sequence for both simulations. Although, the amount of improvement is highly system dependent.

4.3.2. Likelihood ratio / Malliavin

At equilibrium, our particles are distributed according to the Boltzmann distribution,

$$\langle A\rangle = \frac{1}{C}\int\cdots\int A(\vec{x}) \exp\left[-\frac{1}{T}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} U_{ij}\left(\|x_i - x_j\|\right)\right] d\vec{x}.$$

Differentiating with respect to the parameter θ we find

$$\frac{d}{d\theta}\langle A\rangle = \langle A\,\omega\rangle + \langle A\rangle\frac{d\ln C}{d\theta},$$

where

$$\omega(\vec{x}) := -\frac{1}{T}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{dU}{d\theta}\left(\|x_i - x_j\|\right),$$

is independent of A. In the special case when A is a constant equal to one, we have

$$0 = \frac{d}{d\theta}\langle 1 \rangle = \langle \omega \rangle + \frac{d \ln C}{d\theta},$$

or

$$\frac{d \ln C}{d\theta} = -\langle \omega \rangle.$$

Putting it all together, we get

$$\frac{d}{d\theta}\langle A \rangle = \langle A\, \omega \rangle - \langle A \rangle \langle \omega \rangle,$$

where the ω are the likelihood ratio weights. To further reduce the variance, we can shift A by a constant, $\bar{A}$ so that the mean of $A - \bar{A}$ is nearly zero, which results in

$$\frac{d}{d\theta}\langle A \rangle = \langle \left(A - \bar{A}\right)\, (\omega - \langle \omega \rangle) \rangle.$$

Shifting the quantity of interest to get mean zero is called "centering".

4.4. Numerical experiments

To compare correlated finite differencing with the likelihood ratio estimator, we calculated the sensitivity of the radial density function to changes in the Lennard-Jones attractive force multiplier, θ. The standard minimum image convention is used for computing distances between particles, and the potential is smoothly truncated to zero at half the period box length. We simulated a system of 256 particles in a periodic box at reduced temperature 1.5 and with bulk reduced density, 0.95. This corresponds to a supercritical fluid, and we expect the structure of the radial density to depend on the strength of the attractive force.

Each simulation is started from an equilibrated initial condition, and evolved using Brownian dynamics with a time-step of 0.005. The Brownian damping parameter, γ, was set to 10. The moderate length simulations were run for 2.0×10^5 time steps. The long simulations were run for 2.0×10^7 time steps.

We conclude the following based upon our numerical experiments:

1. Both correlated finite differences and the Malliavin estimator can be computed as an average of a time-series over the simulation (or pair of simulations for finite differences).

2. The accuracy achieved by the correlated finite difference method when the optimal step-size is used is nearly that of the Malliavin estimator (aka likelihood ratio).

3. The variance of correlated finite differences is roughly twice the size for finite differences compared to the Malliavin estimator. Our observations support the minimal variance for the Malliavin estimator (aka likelihood ratio) established by Fournié et al. [1999, 2001] .

4. Determining the optimal finite difference step is not trivial, requiring estimates for the statistical error and truncation error.

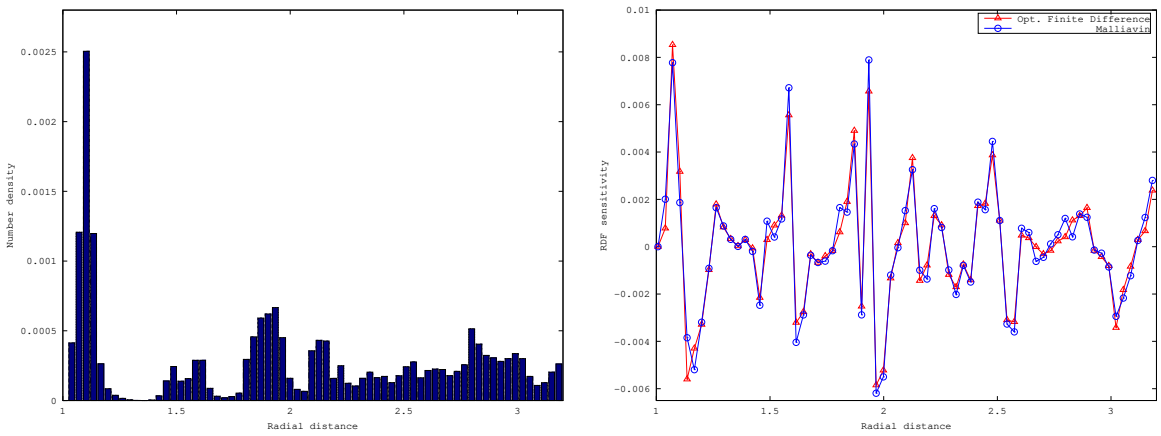


Figure 4-3. On the left, the computed radial density function over a long simulation. On the right, the computed sensitivity in the radial density function using (a) correlated finite differences and (b) Malliavin weights. The finite difference results use a finite difference step of 10^{-4}.

5. If the finite difference step is known, the runtime cost of correlated finite differences is roughly twice that of the Malliavin estimator. This is because finite differencing requires two simulations compared to one.

6. The additional runtime cost of computing the Malliavin sensitivity weight is negligible if it is computed at the same time as the force. This is similar to how the potential energy is computed at the same time as the force.

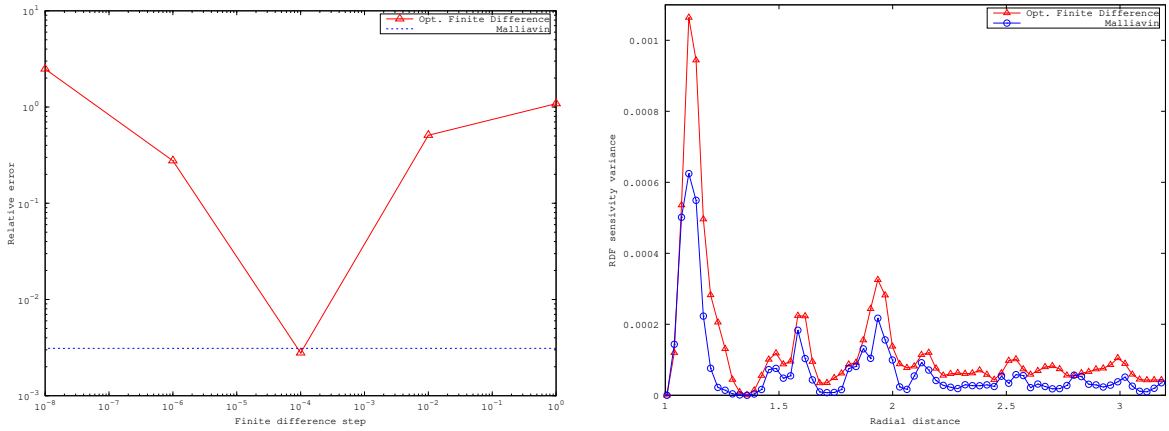


Figure 4-4. On the left, the relative error in the compute sensitivities for a moderate length simulation. The correlated finite difference error is minimized for a finite difference step of 10^{-4}. On the right, the statistical variance in the computed sensitivities for a long simulation.

5. A STOCHASTIC CALCULUS APPROACH TO BOLTZMANN TRANSPORT

This section uses stochastic calculus to describe Boltzmann transport. As discussed in Section 2, representing the Boltzmann transport solution as the expectation of a stochastic differential equation allows for the use of various sensitivity estimators. Here, we develop such a representation of transport by representing the particle simulated in traditional Monte Carlo Boltzmann transport. By developing this stochastic differential equation, we determine that the traditional particle naturally induces the adjoint problem rather than the forward.

This identification has two critical consequences. First, particles can be simulated a single time to obtain both the forward transport solution and the adjoint solution. Second, by employing the adjoint relation the traditional source iteration Monte Carlo method can be recovered, allowing for a stochastic differential equation representation of traditional simulation. Consequently, in the future such sensitivity estimators discussed in this report might be applied to Boltzmann transport. In addition to these extraordinary implications, the stochastic calculus representation developed here also implies that source terms and adjoint source terms may be interchanged without re-simulating particle trajectories.

The remainder of this section is organized as follows. We provide an introduction to Boltzmann transport and a background to the problem area in Section 5.1. In Section 5.2 we review the derivation of the traditional Monte Carlo simulation of Boltzmann transport through source iteration. Then, we develop a stochastic differential equation description of the physical particle process and demonstrate how it induces the adjoint Boltzmann transport equation in Section 5.3. Next, we recover the source iteration method, explicitly relating source iteration sampled particles and particles used in the stochastic calculus solution for the adjoint equation in Section 5.4. Finally we provide a simple numerical demonstration in Section 5.5, discuss the adjoint particle and simulation in Section 5.6, and conclude with a discussion on the importance of this stochastic form for further analysis in Section 5.7.

5.1. An introduction to Boltzmann transport

The Boltzmann transport equation was originally formulated by Ludwig Boltzmann to describe the kinetics of particular gasses. However, the equation has been applied in numerous areas since, including neutron, photon, and electron transport. Rigorous simulation of the Boltzmann equation for particle and radiation transport is critical for power distribution problems, radiation detector design, x-ray simulation, spectral response and analysis, and several others Dunn and Shultis [2009], Vaz [2009], Zhang and Li [2022].

Canonically, a population of neutrons is described through the neutron angular density $N(\mathbf{r}, \overrightarrow{\Omega}, E, t)$, which is the density of neutrons at the location $\mathbf{r}$ traveling in the direction $\overrightarrow{\Omega}$ with energy E at time t Dupree and Fraley [2002]. A related quantity of interest in radiation transport is the angular flux Φ given as $\Phi = vN$, where v is the particle speed. In this notation, a particle's velocity would be written as $\mathbf{v} = v\overrightarrow{\Omega}$.

The neutron angular flux follows the Boltzmann transport equation and may be written as:

$$\begin{aligned} \frac{1}{v}\frac{\partial}{\partial t}\Phi\left(\mathbf{r}, \overrightarrow{\Omega}, E, t\right) + \overrightarrow{\Omega}\cdot\nabla\Phi\left(\mathbf{r}, \overrightarrow{\Omega}, E, t\right) + \Sigma_t\left(\mathbf{r}, \overrightarrow{\Omega}, E\right)\Phi\left(\mathbf{r}, \overrightarrow{\Omega}, E, t\right) - S\left(\mathbf{r}, \overrightarrow{\Omega}, E, t\right) \\ = \int\int \Phi\left(\mathbf{r}, \overrightarrow{\Omega}', E', t\right)\Sigma_s\left(\mathbf{r}, \overrightarrow{\Omega}', E'\right) p\left(\overrightarrow{\Omega}, E \middle| \overrightarrow{\Omega}', E'\right) \mathrm{d}E'\mathrm{d}\overrightarrow{\Omega}'. \end{aligned} \tag{5.1}$$

In this notation, S is a source term and Σ_t gives the macroscopic total interaction cross section, or the probability per unit length of path that a particle will have some sort of interaction. In the case of non-fissionable material, there are only two interactions, scattering and absorption; hence we write $\Sigma_t = \Sigma_a + \Sigma_s$, where Σ_a and Σ_s are the absorption and scattering cross sections, respectively. The probability distribution $p\left(\Omega, E \middle| \overrightarrow{\Omega}', E'\right)$ gives the probability of changing from $(\overrightarrow{\Omega}', E')$ to $(\overrightarrow{\Omega}, E)$ after a scattering event has occurred. We note that in conventional notation, $\Sigma_s\left(\mathbf{r}, \overrightarrow{\Omega}', E'\right) p\left(\overrightarrow{\Omega}, E \middle| \overrightarrow{\Omega}' E'\right)$ is written as $\Sigma_s\left(\mathbf{r}; \overrightarrow{\Omega}', E' \rightarrow \overrightarrow{\Omega}, E\right)$ and called the scattering kernel Dupree and Fraley [2002]. This explicit writing of the probability distribution is necessary for the mathematical treatment presented in this work, and we merely remark that our notation naturally gives rise to the convention

$$\Sigma_s\left(\mathbf{r}, \overrightarrow{\Omega}', E'\right) = \int\int \Sigma_s\left(\mathbf{r}; \overrightarrow{\Omega}', E' \rightarrow \alpha, \beta\right)\mathrm{d}\alpha\mathrm{d}\beta.$$

While deterministic numerical solvers for certain forms of the Boltzmann transport equation do exist Abbassi et al. [2011], Soba et al. [2021], Monte Carlo approaches for numerical approximation remain the dominant approach to solving neutron transport problems Dunn and Shultis [2009], Péraud et al. [2014], Vaz [2009]. There are many flavors and variations of Monte Carlo approaches for neutron transport, including those using multilevel schemes Louvin et al. [2017] or employing additional sampling or counting sub-steps like the predictor-corrector method Leppänen [2009].

Though there are several ways to alter the Monte Carlo approach, the basic method is one of sampling particle paths and "scoring" or "tallying" them in detector regions Dupree and Fraley [2002], Lux and Koblinger [2018]. Briefly, this involves initializing a family of particles with a position $\mathbf{r}_0$ and direction $\overrightarrow{\Omega}_0$ based on the source term. Then, a distance or time to a collision point is drawn as an exponential random variable with a rate given by the total cross-section Σ_t. Depending on whether a time or distance is drawn, this rate may be modified by the appropriate variables. The particle moves forward to the location or time of the event. The event is then determined to be a scattering event with probability $\Sigma_s/(\Sigma_a + \Sigma_s)$ and an absorption event otherwise. In the event of a scattering event, a new direction and energy is selected using the distribution p. If the particle is absorbed, it is either removed from the simulation pool or it reduces its contribution by a precalculated weight involving Σ_a. This process continues until all particles escape from the domain of

interest $\mathcal{D}$ or are absorbed. The neutron flux density is calculated in a detector region by counting and keeping track of a tally of whenever a particle exists within the region.

When calculated in this way, sampled particle paths can be saved and used to calculate new tallies based on different detectors. However, as the method of initialization is tied to the source, the particle paths cannot be reused when the source is changed. Nonetheless, if the detector area is sufficiently small, many paths simulated from the source will be wasted and never encounter the detector Kersting et al. [2020]. As such, the adjoint flux density is sometimes a preferred quantity of interest in these cases. The adjoint simulation is similar to the Boltzmann transport simulation (which we will refer to as forward transport). Simulation is accomplished by initializing particles from the detector and simulating them backwards to the source Kersting et al. [2020], Vitali et al. [2018]. By starting at a detector region with a suitably defined source, more paths are likely to be useful to computation and tallying by hitting the source. Adjoint paths can be reused when sources are swapped out, but not detectors. Sometimes, both forward and adjoint information is desired. Unfortunately since simulations can have differing scattering rates – sometimes handled through altered rates and sometimes handled through precalculated tally weights – and likely have differing starting locations, forward transport sampled particle paths cannot be reused with these methodologies to calculate adjoint information, and vice-versa.

Despite this apparent incompatibility with current methods, traditionally sampled particle paths can be reused with a different averaging scheme to produce adjoint information. In this work, we develop a stochastic calculus based treatment for transport by starting with a stochastic differential equation (SDE) description of the forward particle process. Using this particle description, we create a probabilistic expression for a scored, or tallied, quantity of interest and then show that this quantity of interest is not the Boltzmann flux density, but the adjoint flux density. That is, **the description of the forward physical process naturally gives rise to the adjoint equation**.

5.2. A Review of Source Iteration

Before developing a stochastic calculus approach to Boltzmann transport, we take a moment to revisit source iteration for the uninitiated. It is at this point we will make several simplifications to both reduce notation and to simplify explanation. Namely, we will consider a one-dimensional steady-state version of (5.1). The steady-state, or time-integrated problem is called the particle fluence Dupree and Fraley [2002]. We will further reduce the number of multi-line equations and mathematical clutter by reducing to one-dimension, making $\mathbf{r}$ just x and $\overrightarrow{\Omega}$ just Ω, and by also assuming that our fluence problem has no energy dependence and that Σ_t, Σ_a, and Σ_s depend only on Ω and not on x.

We stress that despite these simplifications, all the results shown, both in this section and in Sec. 5.3, can be extended to the full time-dependent and steady-state Boltzmann equations with only a little more bookkeeping. We will further note that our choice to have the cross-sections depend only on Ω rather than only on x may feel a little unorthodox as it is more likely for the cross-sections to depend on position rather than direction, but choosing x over Ω in this simplified setting would obscure some of the nuance in the mathematics and in the differences between transport

modes. But again, all results showcased here are easily extended to higher dimensions and to full dependence in $\mathbf{r}$, $\overrightarrow{\Omega}$, and E.

With our one-dimensional and variable dependence assumptions, the steady-state Boltzmann transport equation for particle fluence is given by

$$\frac{\partial}{\partial x} v\Omega\Phi(x,\Omega) + v\Sigma_t(\Omega)\Phi(x,\Omega) - f(x,\Omega) = v\int \Phi(x,\Omega')\Sigma_s(\Omega')\, p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega', \tag{5.2}$$

where $f(x,\Omega) = vS(x,\Omega)$ is our new source term. We will now use source iteration to develop a Monte Carlo numerical approach to solving this equation.

The punchline of the source iteration method is to take (5.2) and to write it in the form of $\Phi = K\Phi + f'$ for some operator K and an attenuated source function f'. Once in this form, we can write

$$\Phi = \sum_{n=0}^{\infty} \Phi_n, \tag{5.3}$$

where $\Phi_0 = f'$ and $\Phi_n = K\Phi_{n-1}$, provided that $\sum_{n=0}^{\infty} K^n$ converges in the operator norm. We will soon see that (5.3) can be interpreted in a Monte Carlo scheme. However, we must first identify our operator K.

Consider a point x_0 and the line passing through this point in direction Ω. Parameterizing this line with the variable s, we can write

$$\frac{\mathrm{d}}{\mathrm{d}s}\Phi(x_0+s\Omega,\Omega) + \Sigma_t(\Omega)\Phi(x_0+s\Omega,\Omega) = f(x_0+s\Omega,\Omega) + \int \Phi(x_0+s\Omega,\Omega')\Sigma_s(\Omega')\, p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega'.$$

We will identify the whole right-hand side of this equation as $q(x_0+s\Omega,\Omega)$. Assuming that as $s \to -\infty$ we will have Φ tend to zero – that is, as we get further and further away in the negative direction from our given point, the fluence approaches zero – we can solve for Φ at the point x_0:

$$\Phi(x_0,\Omega) = \int_{-\infty}^{0} e^{\Sigma_t(\Omega)s} q(x_0+s\Omega,\Omega)\,\mathrm{d}s.$$

By changing variables where $x_0 = x$ and changing the sign in the s variable, we obtain

$$\Phi(x,\Omega) = \int_0^{\infty} e^{-\Sigma_t(\Omega)s} \int \Phi(x-s\Omega,\Omega')\Sigma_s(\Omega')\, p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega'\,\mathrm{d}s + \int_0^{\infty} e^{-\Sigma_t(\Omega)s} f(x-s\Omega,\Omega)\,\mathrm{d}s,$$

where the first term on the right-hand side represents our operator $K\Phi$ and the second term our attenuated source, f'. This is a derivation that applies to our one-dimensional special case; the full derivation may be found in Chapter 3 of Dupree and Fraley [2002].

As mentioned previously, (5.3) has an interpretation as a Monte Carlo scheme with our found operator K. First, consider the initial term:

$$\Phi_0 = f' = \int_0^{\infty} e^{-\Sigma_t(\Omega)s} f(x-s\Omega,\Omega)\,\mathrm{d}s.$$

Reading this equation, we see that Φ_0 describes all possible particles from the source traveling in the direction Ω, arriving at the point x in s "steps", attenuated by the probability that they have **not** been scattered or absorbed in those s "steps." That is, Φ_0 describes the fluence of particles arriving at point x traveling in direction Ω that have undergone no scattering or absorption events. Similarly, consider the next term:

$$\Phi_1 = Kf' = \int_0^\infty e^{-\Sigma_t(\Omega)s} \left(\int \left(\int_0^\infty e^{-\Sigma_t(\Omega)s'} f\left(x - s\Omega - s'\Omega', \Omega'\right) \mathrm{d}s' \right) \Sigma_s\left(\Omega'\right) p\left(\Omega \,|\, \Omega'\right) \mathrm{d}\Omega' \right) \mathrm{d}s.$$

Again, reading this equation we see that Φ_1 describes all possible particles from the source that begin in some direction Ω', have a single collision after s' steps, changing their direction to Ω and arriving at position x after another s steps, where each portion of the journey is attenuated by the probability of having no other absorption or scattering events in that time. In other words, Φ_1 describes the fluence of particles at the point x traveling in the direction Ω that have undergone exactly one scattering event. Indeed, this process continues with Φ_n representing the fluence of particles that have undergone exactly n scattering events.

The sum in (5.3) then provides the total fluence of particles arriving at position x in direction Ω having undergone any number of scattering events. To recap our discussion in the introduction, informally simulation occurs by selecting a particle from the source function f. An event point is determined, often by selecting an exponential random distance using Σ_t. Then, the type of event is determined with scattering changing the direction of the particle using the distribution p and absorption either removing the particle from simulation or to attenuate its contribution by the appropriate Σ_a dependent weight. Equation 5.3 is written for the fluence density. To calculate this in simulation, the appropriately normalized cumulative particle tally of particles reaching (x, Ω) is calculated. Calculation is often not for a single point, like implied, but for the integral over a region of points making up the detector; this too is handled with cumulative tally count.

Now that we have obtained the traditional Monte Carlo approach and described the particle process, we will formally describe the particle as an SDE.

5.3. The Physical Particle Stochastic Process and the Adjoint Equation

We will now produce a stochastic calculus description of the particle process used in traditional Boltzmann transport simulation. Once more, we will limit our discussion to the special one-dimensional case previously described for clarity; but note that the stochastic calculus work depicted here can be extended to higher dimensions with only a little more bookkeeping. For an introduction to SDEs, see Wiersema [2008]; for an in-depth treatment of stochastic calculus and integration, see Protter [1992]; for generalized stochastic calculus with jump diffusions, see Hanson [2007].

To that end, let us construct a process $X(t) = (X_1(t), X_2(t))^\top$ that behaves similar to the particles described by the Boltzmann equation. We will interpret $X_1(t)$ as the position and $X_2(t)$ as the direction. Hence, if v is the maximum speed of the particles, then $vX_2(t)$ is the velocity of the particle.

To mimic the physical particles used in the simulation of the Boltzmann equation, we will allow the process to change its direction $X_2(t)$ upon scattering events. Since the inter-arrival times/distances of scattering events are exponential, the scattering events themselves should arrive according to a Poisson arrival process. Let $P(t, \omega, X_2(t))$ be the Poisson arrival process for scattering events with rate $\Sigma_s(X_2(t))$. When a scattering event occurs, a new direction ω after scattering is selected given the pre-scattering direction $X_2(t)$ from a distribution with density function $p(\cdot \mid X_2(t))$. This arrival process is a compound Poisson process – when an event occurs it returns a random value of ω. Following an event, the process X_2 adds a reward function $h(X_2(t), \omega) = \omega - X_2(t)$ to its value. Hence, when a Poisson event occurs, the direction of the process is replaced by the random value ω, selected from a distribution dependent on the pre-scattering direction.

Much like in traditional simulation, once the particle has exited the domain of interest $\mathcal{D}$ we no longer want to track the particle. We accomplish this by defining a stopping time for our stochastic process, $\tau_{x,\Omega}$, which is necessarily dependent on the starting location and direction of the process (x, Ω). This stopping time is the first time the stochastic process exits the domain of interest. In stochastic differential equation (SDE) form, the dynamics of $X(t)$ can be written as:

$$\begin{aligned} \mathrm{d}X_1(t) &= vX_2(t)\mathrm{d}t, \\ \mathrm{d}X_2(t) &= (\omega - X_2(t))\,\mathrm{d}P(t, \omega, X_2(t)), \\ \mathbb{E}\left[\mathrm{d}P(t, \omega, X_2(t)) \mid X_2(t) = \Omega\right] &= \Sigma_s(\Omega)\,\mathrm{d}t \\ \omega &\sim p(\cdot \mid X_2(t)) \\ \tau_{x,\Omega} &= \inf\{t > 0 \mid X_1(0) = x, X_2(0) = \Omega, (X_1(t), X_2(t)) \notin \mathcal{D}\}. \end{aligned} \tag{5.4}$$

Intuitively, the equation states that the change in the position of the process $X_1(t)$ increases by $vX_2(t)$ per change in time, and that the change in direction of the process $X_2(t)$ increases by $\omega - X_2(t)$ per jump of the Poisson process. In this sense, $\mathrm{d}P$ is a counting process that increments randomly, returning a new direction ω. Notably, we have not described absorption in this description. This is because absorption "kills" the particle and is not quite as straightforward to describe in an SDE framework. Instead we opt to handle absorption through the weighted contribution approach. Note that while the probability of a scattering event occurring at exactly $\tau_{x,\Omega}$ is zero, in any practical discretized simulation this may occur. If this does occur in simulation it is imperative to assume that the particle hits the boundary first. This ensures that X_2 is continuous at the boundary point.

Using this stochastic representation of the particle, we will now demonstrate that this particle gives rise to the adjoint transport equation, not the Boltzmann equation. This result is essentially the Feynman-Kac formula for jump processes applied to the adjoint transport equation; however we will build up the equation by means of calculation of a quantity of interest.

To that end, suppose that we wish to calculate a quantity of interest from many simulated paths of (5.4). In a Monte-Carlo fashion, we would like to start a mass of particles at some initial location x with initial direction Ω, let them evolve according to (5.4), contribute to our quantity of interest for the amount of time they remain in the support of some "detector" function g, and attenuate that contribution by the probability that the particles might have been absorbed, according to a Poisson arrival process with rate Σ_a, by the time they enter the support of g. Our final quantity of interest is the average contribution, or expected value, across all simulated particles. Writing in probabilistic

form and conditioning on the starting coordinate for particles (x, Ω), we express our quantity of interest Ψ as:

$$\Psi(x,\Omega) = \mathbb{E}\left[\int_0^{\tau_{x,\Omega}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right) \mathrm{d}u\right) \mathrm{d}t \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right]. \quad (5.5)$$

If g really were an indicator of some detector or sensor, then the preceding quantity could be interpreted as the response of a sensor to a point source located in space-direction space at (x, Ω). Indeed, the construction of the quantity feels a lot like traditional radiation transport: we simulate particles by advecting them and scattering them according to an appropriate process; these particles are counted if they reach some region defined by g; and we either absorb them according to a Poisson process or attenuate their contribution appropriately. However, if we view (5.5) as a function of both x and Ω, then this interpretation starts to make less sense as we would be calculating the contribution based on point sources at every location in position-direction space.

Instead, if we interpret g as an adjoint source, or an importance function, then our quantity of interest Ψ makes more sense as a function of (x, Ω). If particles starting at (x, Ω) often make it to or loiter in the support of g before absorption, then that would mean that the location (x, Ω) is rather important in generating a response in the support of g. That is to say, $\Psi(x, \Omega)$ solves an adjoint transport problem with an implied zero boundary condition. We shall now prove this is the case in a setting with a general boundary value condition.

Theorem 5.3.1. *Let $X(t)$ be given by* (5.4)*. Suppose that there exists a classical solution $\Psi : \mathcal{D} \to \mathbb{R}$, $\mathcal{D}$ compact, that solves the boundary value problem*

$$\begin{aligned} -v\Omega\frac{\partial}{\partial x}\Psi(x,\Omega) + v\Sigma_t(\Omega)\Psi(x,\Omega) - v\int \Psi(x,\Omega')\Sigma_s(\Omega)p(\Omega'\,|\,\Omega)\,\mathrm{d}\Omega' = g(x,\Omega), \\ \Psi(x,\Omega) = B(x,\Omega), \qquad (x,\Omega) \in \partial\mathcal{D}. \end{aligned} \quad (5.6)$$

Then, if Σ_a, Σ_s, and g are continuous almost everywhere and bounded, and if $\mathbb{E}\left[\tau_{x,\Omega}\right] < \infty$ for all $(x, \Omega) \in \mathcal{D}$, then the solution Ψ must be able to be written as

$$\begin{aligned} \Psi(x,\Omega) =& \mathbb{E}\left[B\left(X_1\left(\tau_{x,\Omega}\right), X_2\left(\tau_{x,\Omega}\right)\right) \exp\left(-\int_0^{\tau_{x,\Omega}} v\Sigma_a\left(X_2(t)\right)\mathrm{d}t\right) \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right] \\ &+ \mathbb{E}\left[\int_0^{\tau_{x,\Omega}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right]. \end{aligned} \quad (5.7)$$

Proof. For ease, we rewrite the boundary value problem as

$$\begin{aligned} -v\Omega\frac{\partial}{\partial x}\Psi(x,\Omega) + v\Sigma_a(\Omega)\Psi(x,\Omega) - v\Sigma_s(\Omega)\int \left(\Psi(x,\Omega') - \Psi(x,\Omega)\right)p(\Omega'\,|\,\Omega)\,\mathrm{d}\Omega' = g(x,\Omega), \\ \Psi(x,\Omega) = B(x,\Omega), \qquad (x,\Omega) \in \partial\mathcal{D}. \end{aligned} \quad (5.8)$$

In order to prove the result, we will need to appeal more strongly to the stochastic calculus. This will require the use of the Poisson random measure $\mathcal{P}$ associated with our compound Poisson process P. The measure can be related to the counting process $\mathrm{d}P$ as:

$$\int_{\theta\in[-1,1]} \mathcal{P}\left((t+\mathrm{d}t], (\theta+\mathrm{d}\theta], X_2(t), t\right) = \mathrm{d}P\left(t, \omega, X_2(t)\right). \tag{5.9}$$

Here, the assumed space that the random variable ω (and indeed the variable Ω) lives in is $[-1,1]$ and the integration is performed in this space; the value θ can be any realization of the random variable ω. The Poisson random measure can be written in terms of a mean-zero Poisson random measure, $\widehat{\mathcal{P}}$, and its mean:

$$\mathcal{P}\left((t+\mathrm{d}t], (\theta+\mathrm{d}\theta], X_2(t), t\right) = \widehat{\mathcal{P}}\left((t+\mathrm{d}t], (\theta+\mathrm{d}\theta], X_2(t), t\right) + \Sigma_s\left(X_2(t)\right) p\left(\theta \,|\, X_2(t)\right) \mathrm{d}\theta\mathrm{d}t. \tag{5.10}$$

Consider the function

$$w\left(X_1(t), X_2(t)\right) = \Psi\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\mathrm{d}u\right),$$

and define

$$T_{x,\Omega,n} = \min\left\{n, \tau_{x,\Omega}\right\}.$$

Then, for $t \le T_{x,\Omega,n}$, Itō's rule (a chain rule for stochastic calculus) yields

$$\begin{aligned}
&\mathrm{d}w\left(X_1(t), X_2(t)\right) \\
&\quad= \exp\left(-\int_0^t v\Sigma_a\left(X(u)\right)\mathrm{d}u\right)\left[\left(vX_2(t)\frac{\partial}{\partial x}\Psi\left(X_1(t), X_2(t)\right) - \Sigma_a\left(X_2(t)\right)\Psi\left(X_1(t), X_2(t)\right)\right)\mathrm{d}t\right. \\
&\quad\left.+ \int_{\theta\in[-1,1]} \left(\Psi\left(X_1(t), X_2(t)+\omega-X_2(t)\right) - \Psi\left(X_1(t), X_2(t)\right)\right)\mathcal{P}\left((t+\mathrm{d}t], (\theta+\mathrm{d}\theta], X_2(t), t\right)\right]. \\
&\quad= \exp\left(-\int_0^t v\Sigma_a\left(X(u)\right)\mathrm{d}u\right)\left[-g\left(X_1(t), X_2(t)\right)\mathrm{d}t\right. \\
&\quad\left.+ \int_{\theta\in[-1,1]} \left(\Psi\left(X_1(t), \omega\right) - \Psi\left(X_1(t), X_2(t)\right)\right)\widehat{\mathcal{P}}\left((t+\mathrm{d}t], (\theta+\mathrm{d}\theta], X_2(t), t\right)\right].
\end{aligned} \tag{5.11}$$

The move from the first equality to the second comes from expanding $\mathcal{P}$ and then by employing (5.8). Now, we integrate both sides from 0 to $T_{x,\Omega,n}$ and take an expectation, conditioning on the starting values of X_1 and X_2. Since the expected value of the integral involving a mean-zero Poisson random measure is zero, this leaves us with

$$\begin{aligned}
&\mathbb{E}\left[w\left(X_1\left(T_{x,\Omega,n}\right), X_2\left(T_{x,\Omega,n}\right)\right) \,\middle|\, X_1(0)=x, X_2(0)=\Omega\right] - \mathbb{E}\left[w\left(X_1(0), X_2(0)\right) \,|\, X_1(0)=x, X_2(0)=\Omega\right] \\
&\quad= \mathbb{E}\left[-\int_0^{T_{x,\Omega,n}} g\left(X_1(t), X_2(t)\right)\exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0)=x, X_2(0)=\Omega\right].
\end{aligned} \tag{5.12}$$

The second expectation simplifies to $w(x,\Omega) = \Psi(x,\Omega)$. Solving for Ψ, this leaves

$$\begin{aligned}\Psi(x,\Omega) = \mathbb{E}\left[w\left(X_1\left(T_{x,\Omega,n}\right), X_2\left(T_{x,\Omega,n}\right)\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ + \mathbb{E}\left[\int_0^{T_{x,\Omega,n}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right) \mathrm{d}u\right) \mathrm{d}t \middle| X_1(0) = x, X_2(0) = \Omega\right].\end{aligned} \tag{5.13}$$

We would like to take a limit as $n \to \infty$ of this result. Since $\mathbb{E}\left[\tau_{x,\Omega}\right] < \infty$, as $n \to \infty$ we will have $T_{x,\Omega,n} \to \tau_{x,\Omega}$. Since X_1 is always continuous, $X_1\left(T_{x,\Omega,n}\right) \to X_1\left(\tau_{x,\Omega}\right)$. Since we assume X_2 hits the boundary before a scattering event occurs, X_2 is continuous at the boundary point and $X_2\left(T_{x,\Omega,n}\right) \to X_2\left(\tau_{x,\Omega}\right)$. In order to take a limit of this equation and push it inside both expectations, by dominated convergence it is sufficient to show that the interiors of each expectation are bounded for all n.

We begin with the first term. Note that for a given $t < T_{x,\Omega,n}$ that $X_2(t)$ must have finitely many discontinuities. If it didn't, then there were infinite arrivals of a Poisson process with a bounded rate Σ_s in a finite time period, a contradiction. Next, since Σ_a is continuous almost everywhere and since $X_2(u)$ has finitely many discontinuities, $\Sigma_a\left(X_2(t)\right)$ must also be continuous almost everywhere and hence integrable. Therefore $\exp\left(-\int_0^{T_{x,\Omega,n}} \Sigma_a\left(X_2(t)\right) \mathrm{d}t\right)$ is continuous. Furthermore, this term is bounded since Σ_a is bounded and since $T_{x,\Omega,n}$ is finite for all n. Since Ψ is a classical solution to (5.6), it is also continuous. Since $\mathcal{D}$ is compact, Ψ is also bounded. Hence w is continuous and $w\left(X_1\left(T_{x,\Omega,n}\right), X_2\left(T_{x,\Omega,n}\right)\right)$ is bounded for all n. Ergo:

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{E}\left[w\left(X_1\left(T_{x,\Omega,n}\right), X_2\left(T_{x,\Omega,n}\right)\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ &\quad = \mathbb{E}\left[\lim_{n\to\infty} w\left(X_1\left(T_{x,\Omega,n}\right), X_2\left(T_{x,\Omega,n}\right)\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ &\quad = \mathbb{E}\left[w\left(\lim_{n\to\infty} X_1\left(T_{x,\Omega,n}\right), \lim_{n\to\infty} X_2\left(T_{x,\Omega,n}\right)\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ &\quad = \mathbb{E}\left[w\left(X_1\left(\tau_{x,\Omega}\right), X_2\left(\tau_{x,\Omega}\right)\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ &\quad = \mathbb{E}\left[\Psi\left(X_1\left(\tau_{x,\Omega}\right), X_2\left(\tau_{x,\Omega}\right)\right) \exp\left(-\int_0^{\tau_{x,\Omega}} v\Sigma_a\left(X_2(t)\right) \mathrm{d}t\right) \middle| X_1(0) = x, X_2(0) = \Omega\right] \\ &\quad = \mathbb{E}\left[B\left(X_1\left(\tau_{x,\Omega}\right), X_2\left(\tau_{x,\Omega}\right)\right) \exp\left(-\int_0^{\tau_{x,\Omega}} v\Sigma_a\left(X_2(t)\right) \mathrm{d}t\right) \middle| X_1(0) = x, X_2(0) = \Omega\right].\end{aligned}$$

Next we consider the integral term. For $t < T_{x,\Omega,n}$, similar arguments to the preceding show that $\exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right) \mathrm{d}u\right)$ is both continuous and bounded. Since g is continuous almost everywhere, their product is also continuous almost everywhere and therefore integrable. Since g is also bounded and since $T_{x,\Omega,n}$ is finite for all n, the integral in the expectation in (5.13) is both

continuous in n and bounded for all n. Hence:

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[\int_0^{T_{x,\Omega,n}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\,\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right] \\
&\quad = \mathbb{E}\left[\lim_{n\to\infty}\int_0^{T_{x,\Omega,n}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\,\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right] \\
&\quad = \mathbb{E}\left[\int_0^{\tau_{x,\Omega}} g\left(X_1(t), X_2(t)\right) \exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\,\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0) = x, X_2(0) = \Omega\right].
\end{aligned}$$

Therefore, taking the limit in n of (5.13) produces (5.7), proving the result. □

We will note that this result does not prove that a solution Ψ exists and only includes the assumptions necessary to write the solution in probabilistic form. Existence of solutions for such problems is well-studied. See Pao [1973], for example. Often conditions for existence of solution involve balancing conditions with the boundary condition. One sufficient condition is that if the cross sections are bounded and the source term is square-integrable, then the equation has a solution with any boundary condition Pao [1973]. This is not too far off from the assumptions in our result. We will remark that our assumptions were geared toward a dominated convergence argument and that it may be possible to weaken our assumptions so that a similar limit in n could be taken with a uniformly integrable argument.

In a sense, Theorem 5.3.1 is telling us that when the adjoint transport equation has a solution, the underlying stochastic process is actually the physical particle and **not** the one that we would derive via adjoint source iteration. This is, of course, from the stochastic calculus viewpoint. Both methods are correct and we will recover the Boltzmann transport solution in the sequel using the stochastic calculus representation above and the properties of adjoints.

Combined with the traditional method of simulation, what this representation tells us is more profound. **If particles are sampled for the Boltzmann transport equation by means of the physical particle and the source iteration representation, then those same particles can be reused to obtain adjoint information at their points of origin using** (5.7)**.** Regarding their points of origin, the traditional method of simulation for Boltzmann transport does require initializing particles so that their density approximates some source term. This means the number of particles initialized at the same location can vary from location to location. When reusing these paths in (5.7), those locations that initialize few samples will have poor estimates. In Section 5.5, we demonstrate that sampled particles can be used or reused to obtain either the adjoint or forward transport solutions.

We briefly note again that the result above can be extended to the full adjoint transport equation in all dimensions with full dependence on $\mathbf{r}$, Ω, and E with significantly more bookkeeping. See Hanson [2007], Chapter 7 for the details necessary for extension.

5.4. Connecting Boltzmann Transport and the Adjoint Sample Solution

Connecting the developed stochastic calculus representation of the physical particle and the adjoint solution with the Boltzmann transport solution is merely a manner of employing the definition of the adjoint.

First, we will rewrite (5.6) in terms of the operator $\mathcal{L}^*$:

$$\mathcal{L}^*\Psi(x,\Omega) = g(x,\Omega), \tag{5.14}$$

where

$$\mathcal{L}^*\Psi(x,\Omega) = -v\Omega\frac{\partial}{\partial x}\Psi(x,\Omega) + v\Sigma_t(\Omega)\Psi(x,\Omega) - v\int \Psi(x,\Omega')\Sigma_s(\Omega)\,p(\Omega'\,|\,\Omega)\,\mathrm{d}\Omega'. \tag{5.15}$$

The adjoint of this operator is

$$\mathcal{L}\Phi(x,\Omega) = v\Omega\frac{\partial}{\partial x}\Phi(x,\Omega) + v\Sigma_t(\Omega)\Phi(x,\Omega) - v\int \Phi(x,\Omega')\Sigma_s(\Omega')\,p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega', \tag{5.16}$$

and the Boltzmann transport equation (5.2) may be written as

$$\mathcal{L}\Phi(x,\Omega) = f(x,\Omega), \tag{5.17}$$

where f is the source term. The adjoint and Boltzmann fluence can be related through the use of the adjoint property:

$$\int g\Phi = \langle g,\Phi\rangle = \langle \mathcal{L}^*\Psi,\Phi\rangle = \langle \Psi,\mathcal{L}\Phi\rangle = \langle \Psi, f\rangle = \int \Psi f. \tag{5.18}$$

Provided $f \geq 0$, we can use the above relation to reconnect the traditional method of Boltzmann transport with our stochastic calculus representation of the adjoint fluence. Suppose that g is an indicator function marking the domain of some sensor. Then the quantity $\langle g,\Phi\rangle$ gives the sensor response, or total particle fluence over the domain of the sensor. Dividing f by its integral, $\|f\|$, may be interpreted as a probability density on the pairs (x,Ω) in the domain of f. This allows us to interpret $\langle \Psi, f/\|f\|\rangle$ as an expected value. Hence,

$$\begin{aligned}
\langle g,\Phi\rangle &= \langle \Psi, f\rangle, \\
&= \|f\|\cdot\mathbb{E}_{f/\|f\|}\left[\Psi\left(\widehat{x},\widehat{\Omega}\right)\right], \\
&= \|f\|\cdot\mathbb{E}_{f/\|f\|}\left[\mathbb{E}\left[\int_0^{\tau_{\widehat{x},\widehat{\Omega}}} g\left(X_1(t),X_2(t)\right)\exp\left(-\int_0^t v\Sigma_a\left(X_2(u)\right)\,\mathrm{d}u\right)\mathrm{d}t \,\middle|\, X_1(0)=\widehat{x}, X_2(0)=\widehat{\Omega}\right]\right].
\end{aligned} \tag{5.19}$$

Here, particles with initial coordinate $\left(\widehat{x},\widehat{\Omega}\right)$ are initialized so that their density approximates $f/\|f\|$. These particles are simulated according to the physical process $X(t)$ and scored when they enter the domain of the sensor, given by the indicator function g. This corresponds exactly with how traditional Boltzmann fluence simulation is carried out today.

Equation 5.19 allows us to plainly see that calculation of the adjoint and Boltzmann quantities are inherently linked through a stochastic calculus representation. In the next section, we demonstrate this with a numerical simulation. While the preceding focuses solely on the physical particle process and how it gives rise to the adjoint equation, a similar notion can be mined from an adjoint particle process. Indeed, the adjoint particle process can be shown to give rise to the Boltzmann equation. For completeness, we demonstrate this in Section 5.6.

5.5. Reuse of Trajectories Example

The utility of Theorem 5.3.1 is multi-fold. At its heart, it allows for the simulation of adjoint fluence through the simulation of standard particles. This is quite impactful in that we can now simulate multiple adjoint fluence responses from various adjoint sources by utilizing a single set of trajectories. This is because the standard particle simulation is initialized in the region where the adjoint fluence is desired and is scored in the adjoint source. In traditional adjoint simulation, adjoint particles are initialized in a manner dependent on the adjoint source. Hence, changing the adjoint source will require new trajectory simulation in traditional settings.

Another benefit of this result discussed in the previous section is that particle trajectories generated for a traditional Boltzmann fluence simulation can be reused for an adjoint fluence simulation with no alteration. This implies that if both forward and adjoint information are desired, one need only simulate a single set of trajectories. In this section, we demonstrate the utility of the single trajectory set for simulation of both fluence modes.

The Boltzmann fluence problem we simulate is given by (5.2) with the value assignments given in Table 5-1 and the following boundary condition:

$$\begin{aligned} \Phi(-1,\Omega) &= 0, \qquad \text{if } \Omega \geq 0 \\ \Phi(1,\Omega) &= 0, \qquad \text{if } \Omega \leq 0. \end{aligned} \tag{5.20}$$

Table 5-1. Table of variable values for numerical example. Note, χ is the indicator function for the given rectangles and $\mathcal{A}$ represents the area of the specified rectangle.

Variable	Value	Units
$\Sigma_a(\Omega)$	5.00	cm^{-1}
$\Sigma_s(\Omega)$	2.50	cm^{-1}
$f(x,\Omega)$	$\mathcal{A}(R_1)\chi_{R_1}(x,\Omega)$	$\text{cm}^{-3}\text{steradian}^{-1}$
R_1	$[0.29,0.69]\times[-1,1]$	Rectangle in (x,Ω) space
$g(x,\Omega)$	$\mathcal{A}(R_2)\chi_{R_2}(x,\Omega)$	$\text{cm}^{-3}\text{steradian}^{-1}$
R_2	$[-0.22,-0.06]\times[-1,1]$	Rectangle in (x,Ω) space
$p(\Omega\|\Omega')$	Uniform, no dependence on Ω'.	Probability density

5.5.1. Traditional Boltzmann Fluence Simulation

To simulate Boltzmann transport classically, we initialize a population of $M = 500,000$ particles so that their density approximates the source term f. Since our select source term is flat, we allocate an equal number of particles across a mesh with spatial discretization $\Delta s = 0.01$ and direction discretization $\Delta a = 0.01$. Rounding so that all locations receive an equal number of particles in the source, we initialized 61 particles per location, giving an effective total of $500,200$ particles in this simulation.

Particles are simulated in discrete time with step size $\Delta t = 0.01$. Once initialized, a time to scattering event is selected as an exponential random variable with parameter $(v\Sigma_s)^{-1}$. The particle increments its location by $v\Omega\Delta t$ until just before the event time step or until the particle exits the domain of interest. The event time step is determined to be the first time step of simulation that is greater than the randomly determined scattering event time. For each completed time step, a tally is recorded if the particle is within a mesh point of the sensor rectangle R_2. This tally is weighted by the probability that the particle would not have been absorbed by the current time step; for example, if the particle were in a mesh point for the sensor location at the j^{th} time step, the tally for that mesh location would be increased by $1 \cdot e^{-v\Sigma_a j\Delta t}$. Once particles reach the event time step, the position is updated as usual and then a new direction Ω is selected by uniform random draw. Once more we update the tally if the particle ends the event time step in a mesh point for the sensor.

This process is then repeated with a new event time drawn. Particles are simulated until they exit the domain of interest. For our implementation, we also incorporated a maximum number of time steps of simulation, 1500, so that we did not need to compute with dynamic arrays.[1] Once all tallies are complete, the tallies are normalized. For this simulation, this requires multiplying by the integral of the source term (equal to 1), multiplying by the time step size Δt (this is the fluence problem, so each tally is per time step), and dividing by the total number of walkers $(500,200)$ times Δa and Δs. This final division makes it a density of all walkers in the current mesh size.

We carried out this simulation as described using MATLAB. The results of simulation are given in Figure 5-1a.

5.5.2. Adjoint Fluence from Forward Particle Trajectories

Employing Eq. 5.5, we can use the trajectories generated in the traditional Boltzmann fluence simulation (Fig. 5-1a) to produce an adjoint simulation.

From Eq. 5.5, we can simulate the adjoint fluence solution at a location (x, Ω) by initializing a population of particles at the same location (x, Ω). Since we are reusing particles from the previous simulation, we only have particles initialized within those mesh points comprising the domain of the source function f. Therefore, we estimate the adjoint fluence over the mesh points within the source domain using those particles initialized within those mesh points.

In a discretized fashion, we examine trajectories beginning within a mesh point. Tracing those trajectories, we keep a score, adding the constant value $\mathcal{A}(R_2)\, e^{-v\Sigma_a j\Delta t}$ to our score whenever a

[1] Only 4.88% of simulated particles did not exit the domain after this many time steps.

trajectory is in the rectangle R_2 on the j^{th} time step. Once all trajectories for all particles starting within the mesh point have been traced and the score calculated, we divide that score by the total number of walkers starting within the mesh point (in the case of this simulation, this is always 61). This division takes the expected value over all particles initialized and gives us the adjoint fluence value for that mesh point.

Using our MATLAB generated trajectories, we traced through and completed this calculation, giving an adjoint fluence simulation. This is plotted in Figure 5-1b.

5.5.3. Traditional Adjoint Simulation

For comparison purposes, we completed a traditional adjoint simulation. Additional details on the adjoint particle can be found in Section 5.6. For the purposes of this discussion, it is sufficient to say that for this particular problem, the adjoint particle moves in the reverse of the direction Ω and scatters and absorbs as normal.

Much in the same manner as our traditional Boltzmann fluence simulation, we initialize particles so that their density approximates the adjoint source function g. Again, in this case, our function g is flat, so we will allocate an even number of walkers per mesh location. For this simulation we allocated 147 particles per mesh location for a total number of $499,800$ particles simulated.

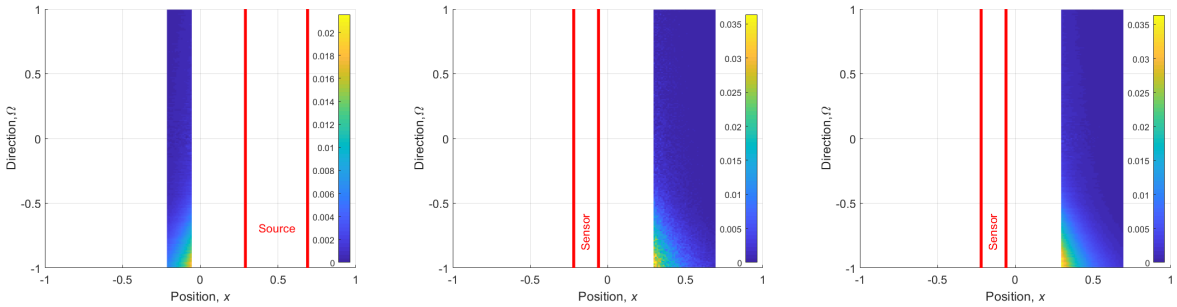
Simulation occurs in a similar manner as well. A time to scattering event is drawn as an exponential random variable with parameter $(v\Sigma_s)^{-1}$. Particles update their position by the increment $-v\Omega\Delta t$ until they leave the domain or reach the scattering event. Note the negative sign for the update increment; this is how the adjoint particle differs in this particular case. The scattering event assigns a new direction from a uniform draw. For each time step j a particle exists within the rectangle R_1, a tally is updated by $1 \cdot e^{-v\Sigma_a j\Delta t}$. The tallies are normalized by multiplying by the integral of the adjoint source (equal to 1 in our case), multiplying by Δt, and dividing by the product of the total number of particles simulated, Δa, and Δs.[2]

The results of this adjoint simulation are given in Figure 5-1c. Immediately we see striking visual agreement, though the reuse case seems noisier. This is to be expected – the reuse of trajectories calculated the value for each position in the domain of f based on the number of samples initialized on that location. This was a flat 61, a rather small number of samples for a Monte Carlo method. Nonetheless, this small number of samples still produces a nice approximation. The norm of difference between the two numerical solutions is 0.043.

5.6. The Adjoint Particle and Boltzmann Fluence

This section will produce a foil to Section 5.3 in that it will develop the Boltzmann fluence equation from a stochastic description of the adjoint particle. We will continue one-dimensional assumptions and assumptions on variable dependence established for the forward problem.

[2] Again, a max time step of 1500 was implemented; only 5.93% of simulated particles did not exit the domain by this time.



(a) Traditional Boltzmann **(b) Adjoint, Reused Trajectories** **(c) Traditional Adjoint**

Figure 5-1. Numerical particle fluence simulations. (a) Traditional Boltzmann fluence simulation generated using forward particles. Particles travel from the source (marked in red) to the sensor (area of solution). (b) Adjoint fluence simulation generated through the reuse of trajectories generated in (a). Particles travel from the source (area of solution) and are scored in the sensor (marked in red). (c) Traditional adjoint fluence simulation generated using adjoint particles. Particles move from the sensor (marked in red) to the source (area of solution). The reuse of forward particles in (b) generates a surprisingly good simulation, with the norm of difference between (b) and (c) equal to 0.043**.**

Heuristically speaking, the adjoint particle moves "backward in time." It moves with a speed v in the opposite direction compared to the physical particle. In addition to movement, scattering and absorption also need to occur in a backward manner. In traditional simulation, an adjoint scattering kernel is calculated through the use of the forward kernel and absorption differences are handled through appropriate weight factors (see Kersting et al. [2020], for example).

To describe our adjoint particle process, we will define the new scattering and absorption cross-sections using our notation. To develop the adjoint scattering cross-section, we must consider that we are no longer scattering from an initial direction Ω' to a new direction Ω, but rather we are starting at direction Ω and reverse scattering to some previous direction Ω'. Hence, we should construct our cross-section by weighting scattering from all previous Ω' values into a given value of Ω:

$$S_\Sigma(\Omega) = \int p(\Omega \,|\, \Omega') \, \Sigma_s(\Omega') \, \mathrm{d}\Omega'. \tag{5.21}$$

Next, since $p(\Omega \,|\, \Omega')$ gives the probability of taking direction Ω given that a scattering event occurred with initial direction Ω', we need to find a new distribution: $q(\Omega' \,|\, \Omega)$, the probability of having come from direction Ω' given a scattering event occurred and direction Ω was taken. We define this distribution q so that the probability of direction is balanced with the rates of scattering from particular directions:

$$q(\Omega' \,|\, \Omega) \, S_\Sigma(\Omega) = p(\Omega \,|\, \Omega') \, \Sigma_s(\Omega'). \tag{5.22}$$

The distribution q is defined in this way precisely because if no particles ever scatter from a direction Ω', then the reverse distribution should never choose such a direction. Note, while it is not always the case that $S_\Sigma = \Sigma_s$, the total cross-section for adjoint transport will remain the same. That is, in the reverse process or the forward process, the total interaction for scattering or absorption should remain constant. Therefore, the adjoint absorption cross-section will be equal to

$A_\Sigma = \Sigma_t - S_\Sigma$. Do note, however, that A_Σ need not always be positive, leading to weights in simulation larger than 1. Additionally, while our notation requires the writing of probability distributions, this notation is consistent with standard notation using kernels over cross-sections.

With these adjoint values in hand, we are ready to describe the adjoint stochastic process. Let $Y_1(t)$ represent the adjoint particle position and $Y_2(t)$ represent the adjoint particle direction. The particle travels at a velocity of $-vY_2(t)$, experiences scattering events according to a Poisson arrival process with rate $S_\Sigma(Y_2(t))$. When a scattering event occurs, a new direction ζ is selected given the pre-scattering direction $Y_2(t)$ from a distribution with density function $q(\cdot|Y_2(t))$. Once more, we will view this as a compound Poisson process $Q(t,\omega,Y_2(t))$. Following the event, the process Y_2 adds the reward $\zeta - Y_2(t)$ to its current value. The domain of interest $\mathcal{D}$ is assumed to be the same as for the forward problem, and the process terminates when it leaves this domain. Defining a new stopping time $\sigma_{x,\Omega}$ we can write the adjoint process as an SDE:

$$\begin{aligned} \mathrm{d}Y_1(t) &= -vY_2(t)\mathrm{d}t \\ \mathrm{d}Y_2(t) &= (\zeta - Y_2(t))\,\mathrm{d}Q(t,\zeta,Y_2(t)) \\ \mathbb{E}\left[\mathrm{d}Q(t,\zeta,Y_2(t))\,|\,Y_2(0)=\Omega\right] &= S_\Sigma(\Omega)\,\mathrm{d}t, \\ \zeta &\sim q(\cdot\,|\,Y_2(t)), \\ \sigma_{x,\Omega} &= \inf\{t>0\,|\,Y_1(0)=x, Y_2(0)=\Omega, (Y_1(t),Y_2(t))\notin\mathcal{D}\}. \end{aligned} \tag{5.23}$$

This SDE can be used to motivate the Boltzmann equation much in the same way the forward particle SDE was used to motivate the adjoint equation in Section 5.3. We will now state and prove a corollary demonstrating this connection.

Theorem 5.6.1 (Corollary)**.** *Let $Y(t)$ be given by* (5.23)*. Suppose that there exists a classical solution $\Phi\colon \mathcal{D}\to\mathbb{R}$, $\mathcal{D}$ compact, that solves the boundary value problem*

$$\begin{aligned} \frac{\partial}{\partial x}v\Omega\Phi(x,\Omega) + v\Sigma_t(\Omega)\Phi(x,\Omega) - \int \Phi(x,\Omega')\Sigma_s(\Omega')\,p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega' &= f(x,\Omega), \\ \Phi(x,\Omega) &= H(x,\Omega), \qquad (x,\Omega)\in\partial\mathcal{D}. \end{aligned} \tag{5.24}$$

Then, if Σ_a, Σ_s, and f are continuous almost everywhere and bounded and if $\mathbb{E}\left[\sigma_{x,\Omega}\right]<\infty$ for all $(x,\Omega)\in\mathcal{D}$, then the solution Φ must be able to be written as

$$\begin{aligned} \Phi(x,\Omega) = \mathbb{E}\left[H\left(Y_1\left(\sigma_{x,\Omega}\right),Y_2\left(\sigma_{x,\Omega}\right)\right)\exp\left(-v\int_0^{\sigma_{x,\Omega}}(\Sigma_t - S_\Sigma)(Y_2(t))\,\mathrm{d}t\right)\middle|\,Y_1(0)=x,Y_2(0)=\Omega\right] \\ +\mathbb{E}\left[\int_0^{\sigma_{x,\Omega}} f(Y_1(t),Y_2(t))\exp\left(-v\int_0^t(\Sigma_t - S_\Sigma)(Y_2(u))\,\mathrm{d}u\right)\mathrm{d}t\middle|\,Y_1(0)=x,Y_2(0)=\Omega\right]. \end{aligned} \tag{5.25}$$

Proof. The proof rests on rearranging (5.24) to fit the form of (5.6) in Theorem 5.3.1. Up to constant multiples, the only item out of place is the integration in the probability distribution. We manipulate the integral by using (5.22):

$$\int \Phi(x,\Omega')\Sigma_s(\Omega')\,p(\Omega\,|\,\Omega')\,\mathrm{d}\Omega' = \int \Phi(x,\Omega')\,S_\Sigma(\Omega)\,q(\Omega'\,|\,\Omega)\,\mathrm{d}\Omega'.$$

Since Σ_s is bounded and continuous almost everywhere, so is S_Σ by (5.21). Hence Theorem 5.3.1 applies. □

As with Theorem 5.3.1, this result can also be extended into the full dimensional problem with dependence on all variables.

To conclude this asection, we will revisit our numerical example. In Section 5.5, we simulated a traditional adjoint fluence problem. Using the exact same trajectories generated for Figure 5-1c we can employ (5.25) to obtain the Boltzmann fluence.

Equation (5.25) simulates Boltzmann fluence at a location (x, Ω) by initializing a population of adjoint particles at the same location. Reusing the adjoint trajectories from Fig. 5-1c, we only have particles initialized within mesh points corresponding to the domain of the sensor (adjoint source) function g.

We examine all trajectories initializing from a mesh point. Those trajectories are traced and a score is kept. The score is incremented on the j^{th} time step by the constant value $\mathcal{A}(R_1)\, e^{-\Sigma_a j \Delta t}$ if the particle is in the rectangle R_1.[3] Once all trajectories are examined, the score is divided by the total number of walkers that started from the examined mesh point (in the case of this simulation, 147). From our MATLAB generated adjoint particle trajectories we produced the Boltzmann fluence plotted in Figure 5-2.

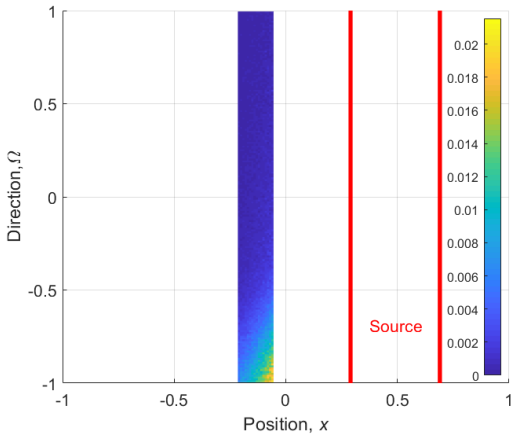


Figure 5-2. Boltzmann fluence simulation generated through the reuse of adjoint particle trajectories used in Fig. 5-1c. Particles travel from the sensor (area of solution) and are scored in the source (marked in red). The reuse of particles generates a good approximation, with the norm of the difference between this solution and the one in Fig. 5-1c equal to 0.013**.**

When comparing Fig. 5-2 to Fig. 5-1c, we again see nice agreement. This simulation benefited from a greater number of particles simulated per position. The norm of the difference of the two adjoint simulations is 0.013.

[3] For this particular problem, $\Sigma_s = S_\Sigma$, so $\Sigma_t - S_\Sigma = \Sigma_a$.

5.7. Discussion

Enabling the reuse of forward trajectories for adjoint calculation is a boon. In traditional simulation, forward trajectories can be reused to calculate fluence when the sensor region is changed. Using the framework developed here, these trajectories can be further reused to calculate adjoint information using any number of adjoint source terms. This is yet another benefit – in traditional adjoint simulation, new trajectories must be computed whenever the adjoint source is changed.

Similarly, Section 5.6 implies that adjoint trajectories may be reused to calculate Boltzmann fluence quantities, and that these trajectories may be reused for any desired source term. This is in contrast to traditional Boltzmann simulation which needs to calculate new trajectories when the source is altered.

Obviously, there is benefit in saving trajectories so that multiple calculations for adjoint and forward transport can be computed from a single set of simulated particles. We did save our particle trajectories for demonstrating reuse in Figures 5-1 and 5-2. However, appropriately organized code could simulate a single set of particles and calculate tallies for both forward and adjoint information simultaneously, removing the need to reprocess stored trajectories.

Beyond the perks of reuse, the stochastic calculus representation is useful on its own. As already mentioned, having a choice of simulation mode (forward or adjoint) gives one the ability to swap out either source terms or adjoint source terms without having to recompute trajectories. Additionally, as Section 5.6 highlights, simulating the adjoint particle often requires the calculation of several additional weights and quantities. This calculation may not always be straightforward to compute. Having this explicit relation between forward particles and the adjoint equation allows for one to use traditional particle trajectories for adjoint information rather than utilizing a potentially more complicated adjoint process.

It is also worth noting that the physics of the problem may make it so that many trajectories are wasted in one simulation mode or the other. For instance, the adjoint source term may not be ideal for hitting the adjoint scoring area or sensor, leading to many simulated adjoint particles never contributing to the quantity of interest. Using the forward particle and the stochastic calculus interpretation allows one to simulate trajectories starting in the adjoint source and ensures that all simulated particles provide meaningful information. The reverse of this situation is also true with forward source terms and sensors.

We must, however, remark that caution must be taken when explicitly reusing forward samples generated from a traditional source-initialized simulation for adjoint information. The examples in this work were for uniform, flat source terms. This means during initialization, we set an equal number of particles across starting locations. In practice, source terms may not be so nicely defined. If a source term is not uniform, then some regions may have fewer particles initialized than others. Since the stochastic calculus method is an expectation based on where particles are initialized, such locations may have a poorer estimate than others. While this may be a concern, we do remark that there are ways of artificially increasing the number of particles that initiate from a particular voxel. One method might be to take the tail of particles trajectories that enter a given voxel at some point in their trajectory.

At large, this stochastic calculus representation provides a rich expansion of Boltzmann transport. From the work presented here, we have two explicit stochastic calculus ways to approach Boltzmann transport and an analogous two ways to approach adjoint transport. Having these SDEs provides more than simulation choice. Namely, having such representations opens up new analysis tools for transport. As detailed elsewhere in this report, having explicit forms for the SDEs and expectation representations for the solutions allows for the use of Malliavin estimators. Hence, further stochastic analysis of these forms can lead one to novel sensitivity methods for transport. This is an active area of research and a main target for this group.

REFERENCES

M Abbassi, A Zolfaghari, A Minuchehr, and M Yousefi. An adaptive finite element approach for neutron transport equation. *Nuclear engineering and design*, 241(6):2143–2154, 2011.

S. Asmussen and P.W. Glynn. *Stochastic Simulation: Algorithms and Analysis*. Stochastic Modelling and Applied Probability. Springer New York, 2007. ISBN 9780387690339. doi: 978-0-387-69033-9.

Goerge Bell and Samuel Glasstone. *Nuclear Reactor Theory*. Van Nostrand Reinhold, 1970.

Nan Chen and Paul Glasserman. Malliavin greeks without Malliavin calculus. *Stochastic Processes and their Applications*, 117(11):1689–1723, 2007. ISSN 0304-4149. doi: 10.1016/j.spa.2007.03.012. Recent Developments in Mathematical Finance.

William L Dunn and J Kenneth Shultis. Monte carlo methods for design and analysis of radiation detectors. *Radiation Physics and Chemistry*, 78(10):852–858, 2009.

Stephen A Dupree and Stanley K Fraley. *A Monte Carlo primer: A Practical approach to radiation transport*. Springer Science & Business Media, 2002.

Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. Applications of malliavin calculus to monte carlo methods in finance. *Finance and Stochastics*, 3: 391–412, 1999.

Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. Applications of malliavin calculus to monte-carlo methods in finance. ii. *Finance and Stochastics*, 5: 201–236, 2001.

Michael B Giles and Niles A. Pierce. An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion*, 65:393–415, 2000.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Applications of mathematics : stochastic modelling and applied probability. Springer, 2004. ISBN 9780387004518. doi: 10.1007/978-0-387-21617-1.

Floyd B Hanson. *Applied stochastic processes and control for jump-diffusions: modeling, analysis and computation*. SIAM, 2007.

Luke J Kersting, Alex Robinson, Eli Moll, Philip Britt, Lewis Gross, and Douglass Henderson. Single scattering adjoint monte carlo electron transport in frensie. *Nuclear Science and Engineering*, 194(5):350–372, 2020.

Jaakko Leppänen. Two practical methods for unionized energy grid construction in continuous-energy monte carlo neutron transport calculation. *Annals of Nuclear Energy*, 36(7):878–885, 2009.

Juan Carlos De los Reyes. *Numerical PDE-Constrained Optimization*. Springer International Publishing, 2015.

Henri Louvin, Eric Dumonteil, Tony Lelièvre, Mathias Rousset, and Cheikh M Diop. Adaptive multilevel splitting for monte carlo particle transport. In *EPJ Web of Conferences*, volume 153, page 06006. EDP Sciences, 2017.

Iván Lux and László Koblinger. *Monte Carlo particle transport methods: neutron and photon calculations*. CRC press, 2018.

Paul Malliavin. Stochastic calculus of variations and hypoelliptic operators. In *Proc. Int. Symp. on Stochastic Differential Equations (Kyoto, 1976)*, pages 195–263, New York, 1976. Wiley.

CV Pao. On nonlinear neutron transport equations. *Journal of Mathematical Analysis and Applications*, 42(3):578–593, 1973.

Jean-Philippe M Péraud, Colin D Landon, and Nicolas G Hadjiconstantinou. Monte carlo methods for solving the boltzmann transport equation. *Annual Review of Heat Transfer*, 17, 2014.

Sergey Plyasunov and Adam P. Arkin. Efficient stochastic sensitivity analysis of discrete event systems. *Journal of Computational Physics*, 221(2):724–738, 2007. ISSN 0021-9991. doi: 10.1016/j.jcp.2006.06.047.

Philip Protter. *Stochastic Integration and Differential Equations, A New Approach*. Springer-Verlag, 1992.

Herbert Rief. A synopsis of Monte Carlo perturbation algorithms. *Journal of Computational Physics*, 111(1):33–48, 1994. ISSN 0021-9991. doi: 10.1006/jcph.1994.1041.

Patrick W. Sheppard, Muruhan Rathinam, and Mustafa Khammash. A pathwise derivative approach to the computation of parameter sensitivities in discrete stochastic chemical systems. *The Journal of Chemical Physics*, 136(3):034115, 2012. doi: 10.1063/1.3677230.

Alejandro Soba, Mauricio E Cazado, Guillaume Houzeaux, Albert Gutierrez-Milla, Mervi J Mantsinen, and Xavier Saez. Validations of the radiation transport module neutro: A deterministic solver for the neutron transport equation. *Fusion Engineering and Design*, 169:112497, 2021.

Charles M Stein. Estimation of the mean of a multivariate normal distribution. *Stanford Univ. Technical Report No. 48*, 1973.

P Vaz. Neutron transport simulation (selected topics). *Radiation Physics and Chemistry*, 78(10): 829–842, 2009.

Vito Vitali, Sandra Dulla, Piero Ravetto, and Andrea Zoia. Comparison of monte carlo methods for adjoint neutron transport. *The European Physical Journal Plus*, 133(8):1–19, 2018.

Patrick B. Warren and Rosalind J. Allen. Malliavin weight sampling for computing sensitivity coefficients in Brownian dynamics simulations. *Physical Review Letters*, 109(25):250601, 2012. ISSN 0031-9007. doi: 10.1103/physrevlett.109.250601.

Ubbo F Wiersema. *Brownian motion calculus*. John Wiley & Sons, 2008.

Tengfei Zhang and Zhipeng Li. Variational nodal methods for neutron transport: 40 years in review. *Nuclear Engineering and Technology*, 2022.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Technical Library	1911	sanddocs@sandia.gov

Hardcopy—Internal

Number of Copies	Name	Org.	Mailstop
1	L. Martin, LDRD Office	1910	0359

Hardcopy—External

Number of Copies	Name(s)	Company Name and Company Mailing Address

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.