SAND2019-11524R

Developing and evaluating Malliavin estimators for intrusive sensitivity analysis of Monte Carlo radiation transport*

Stephen Bond, Brian Franke, R. B. Lehoucq, Darby Smith
Sandia National Labs

11 September 2019

Abstract

We will develop Malliavin estimators for Monte Carlo radiation transport by formulating the governing jump stochastic differential equation and deriving the applicable estimators that produce sensitivities for our equations. Efficient and effective sensitivity can be used for design optimization and uncertainty quantification with broad utilization for radiation environments. The technology demonstration will lower development risk for other particle-based simulation methods.

Contents

*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.

1 Introduction

Traditional Monte Carlo development is based upon simulating the physical process, framed as the approximate solution to the Boltzmann equation, which provides the average of interest. Our approach is more fundamental and first formulates the stochastic jump differential equation (SJDE) that generates the Boltzmann equation. Second, given the SJDE, a Malliavin estimator can be derived and then implemented in a Monte Carlo code, since the latter is a representation of the SJDE. The estimator provides a sensitivity for the Monte Carlo simulation.

The significant practical impact is an efficient and on-the-fly capability to perform Monte Carlo sensitivity analysis, an elusive goal for some important parameters. The potential impact is significant—understanding the change in a quantity of interest with respect to another quantity is of paramount importance. Our deliverable paves the path towards a currently unavailable capability, setting the stage toward design, optimization and uncertainty quantification, staples of modeling and simulation.

The contribution of our FY19 exploratory LDRD *Developing and evaluating Malliavin estimators for intrusive sensitivity analysis of Monte Carlo radiation transport* is to express the

1. Boltzmann neutron transport equation in a form suitable for stating the underlying jump stochastic differential equation;

2. angular flux density, the solution of the transport equation, as the conditional expectation of the solution of the JSDE and the initial condition;

3. derive a relationship that if satisfied demonstrates that the differential operator sensitivity may be identified with a Malliavin derivative.

All three items are, to the best of our knowledge, not available in the general literature. Moreover, our contribution sets the stage for an algorithmic and mathematical basis for sensitivity analysis of Monte Carlo simulations of radiation transport. In particular, by exploiting techniques in modern probability, we provide a mathematical basis for our third contribution given in the above enumerated list.

Our report is organized as follows. Section 2 provides a brief summary of neutron transport and the Monte Carlo approximation of the underlying particle path. The reader is referred to [3] or, for a more comprehensive treatment, the textbook [1]. Section 3 reformulates the neutron transport equation enabling standard results to be applied, forming the jump stochastic differential equation (SDE) for the particle path. Section 4 compares and contrasts the two approaches. Our report concludes with section 5 where we derive the relationship that must hold so that the differential operator approach determines a Malliavin derivative for the desired sensitivity.

2 Neutron transport and Monte Carlo particle approximation

We adopt the standard assumption that a neutron is a point particle where intra-neutron interactions can be ignored. The position and velocity of the particle is given by the vectors

$$\boldsymbol{r} = \begin{pmatrix} x, & y, & z \end{pmatrix} \text{ and } \boldsymbol{v} = v\,\boldsymbol{\Omega} = v\begin{pmatrix} \sin\theta\cos\varphi, & \sin\theta\sin\varphi, & \cos\theta \end{pmatrix}, \tag{1}$$

respectively, where the unbolded v denotes the particle speed, an unsigned quantity. A population of neutrons is described by angular density $N(\boldsymbol{r}, \boldsymbol{\Omega}, E, t)$ so that

$$N(\boldsymbol{r}, \boldsymbol{\Omega}, E, t)\, dV\, d\boldsymbol{\Omega}\, dE$$

is defined to be the expected, or probable, number of neutrons at time t in a volume dV about $\boldsymbol{r}$ with directions $d\boldsymbol{\Omega}$ about $\boldsymbol{\Omega}$ and energies dE about E. However, a more useful quantity is the angular flux density

$$\Phi(\boldsymbol{r}, \boldsymbol{\Omega}, E, t) = vN(\boldsymbol{r}, \boldsymbol{\Omega}, E, t), \tag{2}$$

or the expected number of neutrons traveling with speed v in the phase space element $dV\, d\boldsymbol{\Omega}\, dE$. For the remainder of our report, we make two simplifying assumptions:

1. mono-energetic transport so that the angular flux density is given by $\Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t)$;
2. the source term S functions as an initial condition and does not contribute to the time evolving angular flux. More informally, S acts like an delta function in time. Hence $S\, dV\, d\boldsymbol{\Omega}\, dE$ represents the expected rate of neutrons appearing in the phase space element $dV\, d\boldsymbol{\Omega}\, dE$ due to external sources at the initial time of $t = 0$.

The evolution of the angular flux density is then assumed to satisfy the six-dimensional Boltzmann neutron transport equation

$$\frac{1}{v}\frac{\partial \Phi}{\partial t} + \boldsymbol{\Omega} \cdot \nabla\Phi + \big(\sigma_s + \sigma_a\big)\Phi = A\Phi + S \tag{3a}$$

where

$$C\,\Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t) = \int \Phi(\boldsymbol{r}, \boldsymbol{\Omega}', t)\sigma_s(\boldsymbol{r})\, p(\boldsymbol{\Omega}' \to \boldsymbol{\Omega})\, d\boldsymbol{\Omega}' \tag{3b}$$

is the expected rate of neutrons entering $(\boldsymbol{r}, \boldsymbol{\Omega})$ at time t due to scattering. The parameters σ_s, σ_a are the scattering, absorption interaction cross sections, p is the angular scattering joint probability density, and ∇ is the gradient operator with respect to the coordinate directions x, y, z comprising the particle position $\boldsymbol{r}$. We also remark that the integral over angle cannot be replaced with a differential operator because a collision results in discontinuous change in direction, e.g., see [1, p.21].

Since our interest is in Monte Carlo particle approximation, we re-express (3a) as an integral equation of the second kind that enables us to make explicit the underlying stochastic process. The method of characteristics, see e.g., [3, §3.4] or [1, §1.2], is applied to (3a) to obtain

$$\Phi = T\,C\,\Phi + T\,S\,, \tag{4a}$$

where the density rate $C\,\Phi$ was given by the integral equation (3b) and

$$T\,\Upsilon(\boldsymbol{r},\boldsymbol{\Omega},t) = \int_0^\infty \exp\Big(-\int_0^{\ell'}\big(\sigma_a(\boldsymbol{r}-\ell\boldsymbol{\Omega})+\sigma_s(\boldsymbol{r}-\ell\boldsymbol{\Omega})\big)d\ell\Big)\Upsilon(\boldsymbol{r}-\ell'\boldsymbol{\Omega},\boldsymbol{\Omega},t)\,d\ell' \tag{4b}$$

transforms an angular flux density rate Y to another traveling with speed v. This implies that the angular flux density is given by the formal solution of (4a) via the Neumann series

$$\Phi = (I - T\,C)^{-1}T\,S = \sum_{n=0}^{\infty}\Phi_n\,, \tag{4c}$$

where $\Phi_0 = T\,S\,, \Phi_1 = T\,C\,\Phi_0\,, \cdots, \Phi_n = T\,C\,\Phi_{n-1}$. By (2), the quantity

$$\frac{1}{v}\Phi_n(\boldsymbol{r},\boldsymbol{\Omega},t)\,dV d\boldsymbol{\Omega} \tag{4d}$$

is the expected number of neutrons that undergo n collisions in the phase space element $dV d\boldsymbol{\Omega}$ at time t so that the Neumann series sums over all the possible number of collisions. The Monte Carlo approximation is now given by sampling an exponentially distributed random variable with parameter $\sigma_a + \sigma_s$ for the distance traveled to a collision or absorption, events that we assume are independent. Absorption is deemed to have occurred if a uniformly sampled random variable on the interval $(0,1)$ is less than $\sigma_a/(\sigma_a+\sigma_s)$ otherwise the interaction is deemed to have been a collision at which point scattering occurs.[1]

3 A Jump SDE for neutron transport

We first rewrite (3a), using the equality (1) to obtain

$$\begin{cases} \dfrac{\partial\Phi}{\partial t} + \boldsymbol{v}\cdot\nabla\Phi + v\sigma_a\Phi = v\,(C-\sigma_s I)\Phi\,, \\ \qquad\qquad \Phi(\boldsymbol{r},\boldsymbol{\Omega},0) = v\,S(\boldsymbol{r},\boldsymbol{\Omega}) \end{cases} \tag{5}$$

where we continue to make the simplifying assumption that the source term functions as an initial condition and does not contribute to the time evolving angular flux. If we assume p is a joint probability density, then p satisfies

$$\int p(\boldsymbol{\Omega}'\to\boldsymbol{\Omega})\,d\boldsymbol{\Omega}' = 1$$

[1] We are exploiting the fact that the minimum of two independent exponentially distributed random variable is also exponentially distributed with rate $\sigma_a+\sigma_s$ and that the probability that the minimum is absorption or scattering is $\sigma_a/(\sigma_a+\sigma_s)$ and $\sigma_s/(\sigma_a+\sigma_s)$, respectively.

and so

$$\begin{aligned} v\,(C-\sigma_s I)\,\Phi(\boldsymbol{r},\boldsymbol{\Omega},t) &= \int \big(\Phi(\boldsymbol{r},\boldsymbol{\Omega}',t)-\Phi(\boldsymbol{r},\boldsymbol{\Omega},t)\big)\,v\sigma_s(\boldsymbol{r})p(\boldsymbol{\Omega}'\to\boldsymbol{\Omega})\,d\boldsymbol{\Omega}' \\ &= \int \big(\Phi(\boldsymbol{r},\boldsymbol{\Omega}+\boldsymbol{\omega},t)-\Phi(\boldsymbol{r},\boldsymbol{\Omega},t)\big)\,v\sigma_s(\boldsymbol{r})p(\boldsymbol{\omega}\to 0)\,d\boldsymbol{\omega}\,, \end{aligned} \tag{6}$$

where the second equality follows from a change of integration variable via the substitution $\boldsymbol{\omega}+\boldsymbol{\Omega}=\boldsymbol{\Omega}'$. We remark that the quantity

$$v\,\sigma_s(\boldsymbol{r})p(\boldsymbol{\Omega}'\to\boldsymbol{\Omega})$$

represents a rate of angular scattering and is the critical quantity needed to define the jump SDE that describes particle transport. If we continue to assume that absorption is independent from scattering and has no dependence upon the state of the neutron, the Feynman–Kac theorem, see e.g. [4, Chap.7], implies that

$$\Phi(\boldsymbol{r},\boldsymbol{\Omega},t)=\mathbb{E}[v\,e^{-v\sigma_a(\boldsymbol{r})t}S\big(\boldsymbol{R}(t),\boldsymbol{D}(t)\big)\,|\,\boldsymbol{R}(0)=\boldsymbol{r}\,,\boldsymbol{D}(0)=\boldsymbol{\Omega}] \tag{7a}$$

where the neutron path is given by the compound Poisson processes

$$\boldsymbol{R}_t=\boldsymbol{r}+v\sum_{k=1}^{N_t}\boldsymbol{D}_k \text{ and } \boldsymbol{D}_t=\boldsymbol{\Omega}+\sum_{k=1}^{N_t}\boldsymbol{\omega}_k \text{ for } 0\leqslant t\leqslant\tau_a \tag{7b}$$

where N_t is a Poisson process with parameter $v\sigma_s(\boldsymbol{r})t$ and $\boldsymbol{\omega}_k$ is a trigonometric random variable, and τ_a is the time to absorption. In words, the neutron path is piecewise linear with random changes in direction until absorption. Moreover, the angular flux density is the expected value of the random variable $e^{-v\sigma_a(\boldsymbol{r})t}\,S\big(\boldsymbol{R}(t),\boldsymbol{D}(t)\big)$ conditioned on the event that the neutron is located at $\boldsymbol{r}$ with direction $\boldsymbol{\Omega}$ at time $t=0$. The relations (7) demonstrate that we have achieved the first two items listed in the introduction §1.

The evolution of $\boldsymbol{R}_t$ and $\boldsymbol{D}_t$ are equivalently given by the jump SDEs

$$\begin{cases} d\boldsymbol{R}_t=v\boldsymbol{D}_t\,dt \text{ with } \boldsymbol{R}_0=\boldsymbol{r} \\ d\boldsymbol{D}_t=\boldsymbol{\omega}_t\,dP_t \text{ with } \boldsymbol{D}_0=\boldsymbol{\Omega}\,, \end{cases} \tag{8a}$$

where the above is notation for

$$\begin{cases} \boldsymbol{R}_t=\boldsymbol{r}+v\displaystyle\int_0^t\boldsymbol{D}_s\,ds \\ \boldsymbol{D}_t=\boldsymbol{\Omega}+\displaystyle\int_0^t\boldsymbol{\omega}_s\,dP_s\,, \end{cases} \tag{8b}$$

where dP_s is a Poisson differential. In practical computation, the process $\boldsymbol{D}_t$ is not simulated because it requires the distribution for the angle change $\boldsymbol{\omega}_s$ and the distribution for the direction change is given by $p(\boldsymbol{\Omega}'\to\boldsymbol{\Omega})$.

4 Comparing both approaches

One consequence of the discussion following (4c) is that the probabilistic interpretation associated with the Neumann series allows us to identify a random walk undertaken by a neutron. The walk is constant in velocity in the random direction given by $\boldsymbol{\Omega}$ until a collision occurs at a distance specified by the exponential distribution. In contrast, although the walk defined by the compound Poisson process (7b) is also constant in velocity, the change in direction occurs at an exponentially distributed time.

Without loss of generality, we now re-express (3a) and (5) in terms of the angular density $N = N(\boldsymbol{r}, \boldsymbol{\Omega}, E, t)$ under the assumption that the source term functions as an initial condition and does not contribute to the time evolving angular flux. This will enable us to compare the random process, or walk, associated with each Monte Carlo sample. First, we note that (3a) can be re-expressed as

$$\begin{cases} \dfrac{\partial N}{\partial t} + \boldsymbol{v} \cdot \nabla N + (v\sigma_a + v\sigma_s)N = \tilde{C}N \,, \\ N(\boldsymbol{r}, \boldsymbol{\Omega}, 0) = S(\boldsymbol{r}, \boldsymbol{\Omega}) \,, \end{cases} \tag{9a}$$

where

$$\tilde{C}\, N(\boldsymbol{r}, \boldsymbol{\Omega}, t) = C\, \Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t) = \int N(\boldsymbol{r}, \boldsymbol{\Omega}', t) v\sigma_s(\boldsymbol{r}) p(\boldsymbol{\Omega}' \to \boldsymbol{\Omega})\, d\boldsymbol{\Omega}' \,. \tag{9b}$$

Second, we note that (5) is equivalent to

$$\begin{cases} \dfrac{\partial N}{\partial t} + \boldsymbol{v} \cdot \nabla N + v\sigma_a N = (\tilde{C} - v\sigma_s I)N \,, \\ N(\boldsymbol{r}, \boldsymbol{\Omega}, 0) = S(\boldsymbol{r}, \boldsymbol{\Omega}) \,, \end{cases} \tag{9c}$$

or equivalently, a simple rearrangement of (9a). From (7a), we conclude that the angular density

$$N(\boldsymbol{r}, \boldsymbol{\Omega}, t) = \mathbb{E}[e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big) \,|\, \boldsymbol{R}(0) = \boldsymbol{r}\,, \boldsymbol{D}(0) = \boldsymbol{\Omega}] \tag{10a}$$

is induced by the solution of the jump SDE, i.e., the compound Poisson process (7b), and then by (4c)

$$N = \sum_{n=0}^{\infty} \frac{1}{v} \Phi_n \,. \tag{10b}$$

The above analysis can be summarized in the following result.

Lemma 1. *Let S be the initial condition for the compound Poisson process (7b) with killing (or absorption) rate $-v\sigma_a$. Then the solution of the Boltzmann equation (9a) is*

$$N(\boldsymbol{r}, \boldsymbol{\Omega}, t) = \mathbb{E}[e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big) \,|\, \boldsymbol{R}(0) = \boldsymbol{r}\,, \boldsymbol{D}(0) = \boldsymbol{\Omega}] = \frac{1}{v} \sum_{n=0}^{\infty} \Phi_n(\boldsymbol{r}, \boldsymbol{\Omega}, t) \,, \tag{11}$$

where $\Phi_0 = T\,S\,, \Phi_1 = T\,C\,\Phi_0\,, \cdots, \Phi_n = T\,C\,\Phi_{n-1}$ and the operators C and T be given by (3b) and (4b).

An important, practical, conclusion is that although simulations of the JSDE can be identified with Monte Carlo sampling of the integral equation (4a), there are important distinctions to be made. Chief among these are that the

- Monte Carlo method for the integral equation (4a) samples for the distance between collisions, both due to scattering and absorption;
- compound Poisson process (7b) samples for the number and location of collisions due to scattering through time t and until absorption.

However, if the method of characteristics underlying the integral equation (4a) was instead applied to the transport problem (9a), then the resulting Monte Carlo method samples for the time between collisions. The resulting Monte Carlo history then represents a sampling of the compound Poisson process (7b).

5 Differential Operator Sensitivity and Malliavin Derivative

The paper [5] by Rief describes an approach for computing a sensitivity of the equilibrium angular flux density upon another quantity. More precisely,

$$\frac{\partial \Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t; \rho)}{\partial \rho}$$

where the notation $\Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t; \rho)$ emphasizes the dependence of the angular flux upon the scalar quantity ρ and unlike Rief, we consider the non-equilibrium case. We then have that

$$\frac{\partial \Phi(\boldsymbol{r}, \boldsymbol{\Omega}, t; \rho)}{\partial \rho} = \sum_{n=0}^{\infty} \frac{\partial}{\partial \rho} \Phi_n(\boldsymbol{r}, \boldsymbol{\Omega}, t; \rho)\,. \tag{12a}$$

Rief's assertion that the requisite (non-equilibrium) sensitivity can be computed during the Monte Carlo simulation used to approximate the angular flux occurs upon the realization that the derivative with respect to ρ can be propagated through to the random walk employed for the estimate of Φ_n.

We can now conclude via Lemma 1 that the sensitivity of the angular flux is given by

$$v \frac{d}{d\rho} \mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}}[e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big)] \tag{12b}$$

where for brevity the notation $\mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}}$ denotes the conditional expectation expressed in the Lemma. Formally, we can invoke Ito's rule followed by a stochastic integration by parts to obtain

$$v \frac{d}{d\rho} \mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}}\big[\, e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big)\big] = v\, \mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}}\big[\nabla_{\boldsymbol{r},\boldsymbol{\Omega}}\big(e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big)\big) \cdot \boldsymbol{Y}_t\big] \tag{12c}$$

$$= v\, \mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}}\big[e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big)\, H_t\big]\,, \tag{12d}$$

where $\boldsymbol{Y}_t = \big(d\boldsymbol{R}(t)/d\rho,\, d\boldsymbol{D}(t)/d\rho\big)$ is the first variation process and H_t is an adjoint process. The latter equality explains that the sensitivity can be estimated during the compound Poisson process (7b). The adjoint process H_t is the Malliavin derivative of the first variation process $\boldsymbol{Y}_t = \big(d\boldsymbol{R}(t)/d\rho,\, d\boldsymbol{D}(t)/d\rho\big)$ and informally represents the divergence of $\boldsymbol{Y}_t$. See the papers [2, 6] for interesting introductions within the context of finance and statistical mechanics including many citations to the literature.

We have formally established the important relationship

$$\sum_{n=0}^{\infty} \frac{\partial}{\partial \rho} \Phi_n(\boldsymbol{r}, \boldsymbol{\Omega}, t; \rho) = v\, \mathbb{E}_{\boldsymbol{r},\boldsymbol{\Omega}} \big[e^{-v\sigma_a(\boldsymbol{r})t} S\big(\boldsymbol{R}(t), \boldsymbol{D}(t)\big)\, H_t \big]\,, \tag{13}$$

a consequence of Lemma (1). The identification between the differential operator sensitivity as a Malliavin derivative ultimately depends upon the distinction between the Monte Carlo method applied to the integral equation (4a) with the compound Poisson process (7b) described following Lemma 1 so that an adjoint process H_t can be determined. An appropriate adjoint process is determined if either the method of characteristics underlying the integral equation (4a) was instead applied to the transport problem (9a) or the compound Poisson process (7b) was suitably scaled to sample for the number and location of collisions due to scattering through a specified distance until absorption. We remark that we have achieved the third, and final, item listed in the introduction §1.

References

[1] G. I. Bell and S. Glasstone. *Nuclear Reactor Theory*. Van Nostrand Rehinhold, 1970.

[2] N. Chen and P. Glasserman. Malliavin greeks without Malliavin calculus. *Stochastic Processes and their Applications*, 117(11):1689–1723, 2007. Recent Developments in Mathematical Finance.

[3] S. Dupree and S. Fraley. *A Monte Carlo Primer: A Practical Approach to Radiation Transport*. Number v. 1. Springer US, 2002.

[4] F. B. Hanson. *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation*. Advances in Design and Control. SIAM, 2007.

[5] H. Rief. A synopsis of Monte Carlo perturbation algorithms. *Journal of Computational Physics*, 111(1):33–48, 1994.

[6] P. B. Warren and R. J. Allen. Malliavin weight sampling: A practical guide. *Entropy*, 16(1):221–232, 2014.