Radiation Physics and Chemistry 78 (2009) 852–858

Contents lists available at ScienceDirect

Radiation Physics and Chemistry

journal homepage: www.elsevier.com/locate/radphyschem

Monte Carlo methods for design and analysis of radiation detectors

William L. Dunn *, J. Kenneth Shultis

Department of Mechanical and Nuclear Engineering, Kansas State University, 3002 Rathbone Hall, Manhattan, KS 66506, USA

ARTICLE INFO

Keywords:
Monte Carlo
Simulation
Radiation detectors
Inverse problems
Detector design

ABSTRACT

An overview of Monte Carlo as a practical method for designing and analyzing radiation detectors is provided. The emphasis is on detectors for radiation that is either directly or indirectly ionizing. This overview paper reviews some of the fundamental aspects of Monte Carlo, briefly addresses simulation of radiation transport by the Monte Carlo method, discusses the differences between direct and inverse detection problems, and illustrates how various Monte Carlo methods can be used in design and analysis of radiation detectors.

1. Introduction

The term "Monte Carlo (MC)", as a method of using sample means to estimate population means, was introduced by Metropolis and Ulam (1949), two pioneers in modern simulation practices. Others, such as Lord Kelvin and Enrico Fermi, previously had used methods that are now known as Monte Carlo to estimate the outcomes of certain experiments (real or hypothetical), but the term Monte Carlo was coined as scientists including John von Neumann, Stanislas Ulam, and Nicolas Metropolis formalized the use of the method with the ENIAC digital computer in nuclear weapon studies (Metropolis, 1985).

The present overview concentrates on essential features of MC that allow it to be applied to detector design and analysis, while other papers in this issue describe the capabilities of some of the general-purpose MC codes and address specific detector applications more fully. The treatment is directed to those who have little or no *a priori* knowledge of MC methods. It is risky to use general-purpose codes without clearly understanding the principles on which they are based, and this introduction seeks to elucidate those principles. We provide a brief overview of MC, review some fundamental concepts required to apply MC, consider its application to radiation transport in general, and indicate some of the differences between direct (analysis) and inverse (design) problems.

1.1. What is Monte Carlo?

An overview of Monte Carlo should include a brief discussion of various interpretations of the method. On one hand, the MC method can be thought of as essentially a robust form of numerical quadrature, i.e., a way to specify the abscissas and weights in a finite summation to estimate definite integrals that define statistical expectations. The basic theorems of Monte Carlo address this interpretation and we will develop the fundamentals from this point of view. In this interpretation, the problem need not involve a stochastic variable; ordinary definite integrals can be estimated by MC techniques. For an m-dimensional definite integral, errors in standard quadrature schemes using N prescribed abscissas vary, at best, as $N^{-1/m}$ (Mosegaard and Sambridge, 2002). However, errors in MC estimates vary, at worst, as $N^{-1/2}$. Hence, Monte Carlo is superior to other quadrature schemes for integrals of dimension greater than two! Further, there are a host of ways, called variance reduction techniques, to reduce the MC estimation errors even more.

The vast majority of problems addressed by Monte Carlo, however, are not posed as integrals and so many practitioners prefer to view MC as a way to simulate physical phenomena. In this view, Monte Carlo can be thought of as an "experiment" carried out on a digital computer or through a game, such as throwing darts or flipping a coin, that mimics some other physical process. In fact, perhaps the earliest implementation of Monte Carlo was to drop needles of uniform length L onto a plane surface that had lines at constant spacing D. De Buffon (1777) showed that the probability P_{cut} that a needle would cut (overlap or intersect) a line could be expressed, for $D \geq L$, as

$$P_{cut} = \frac{2L}{\pi D}. \tag{1}$$

Laplace (1786) realized that this procedure could be used to estimate the value of π by counting the number n_c of needles cutting a line out of N random needle drops. Since $P_{cut} \simeq n_c/N$, it follows that

$$\pi \simeq \frac{2L}{D}\frac{N}{n_c}. \tag{2}$$

* Corresponding author. Tel.: +1 785 532 5628; fax: +1 785 532 7057.
E-mail address: dunn@mne.ksu.edu (W.L. Dunn).

0969-806X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.radphyschem.2009.04.030

The terms *deterministic* Monte Carlo (for MC as quadrature) and *probabilistic* Monte Carlo (for MC as simulation of stochastic processes) have sometimes been used to distinguish between these different views; however, it is unnecessary to draw such distinctions. Either interpretation permits MC to be used for design and analysis of radiation detectors, because detector responses are derivable from quantities, such as particle flux or current densities, that can be interpreted as expected values, and because the transport and interaction of radiation can be simulated using digital computers.

2. Review of fundamentals

2.1. Probability distributions

It is first necessary to establish some nomenclature and conventions. Consider an n-component vector $\mathbf{x}$ of random variables defined over a "volume" V. A joint probability density function (PDF) of $\mathbf{x}$ is a function that is defined and non-negative on V and is normalized such that

$$\int_V f(\mathbf{x})\,d\mathbf{x} = 1. \tag{3}$$

The PDF has the interpretation that $f(\mathbf{x})\,d\mathbf{x}$ is the probability that a random sample selected from f will have a value in $d\mathbf{x}$ about $\mathbf{x}$. Of great utility are PDFs of a single random variable $f(x)$. If $x \in [a,b]$, the *cumulative* distribution function (CDF) $F(x)$ is defined as

$$F(x) = \int_a^x f(x')\,dx'. \tag{4}$$

The CDF is a monotonic increasing function beginning at $F(a) = 0$ and ending at $F(b) = 1$, and $F(x)$ represents the probability a random ample has a value less than or equal to x.

If $\mathbf{x}$ is decomposed such that $\mathbf{x} = \{\mathbf{x}_j, \mathbf{x}_k\}$ where $j + k = n$ and $\mathbf{x}_j$ and $\mathbf{x}_k$ are j- and k-component vectors, respectively, of the random variables, then $f_j(\mathbf{x}_j) = \int_{V_k} f(\mathbf{x})\,d\mathbf{x}_k$ is called the *marginal* PDF of $\mathbf{x}_j$. The *conditional* PDF of $\mathbf{x}_k$ given $\mathbf{x}_j$ is $f(\mathbf{x}_k|\mathbf{x}_j) = f(\mathbf{x})/f_j(\mathbf{x}_j)$. Finally, if $f(\mathbf{x}) = f(x_1)f(x_2)\cdots f(x_n)$, then all n random variables are independent.

Similar definitions apply to discrete stochastic variables. If a stochastic variable x can assume n discrete values $\{x_1, x_2, \ldots, x_n\}$, where n can be finite or infinite, then the PDF of x is such that $f_i > 0$ for $i = 1, 2, \ldots, n$ and has the property $\sum_{i=1}^n f_i = 1$.

2.2. Means and variances

Two important measures of any random variable distributed on an interval $[a,b]$ are its mean or expected value (the *population mean*)

$$\langle x \rangle = \int_a^b x f(x)\,dx \tag{5}$$

and its variance (the *population variance*) defined as

$$\sigma^2(x) = \int_a^b [x - \langle x \rangle]^2 f(x)\,dx, \tag{6}$$

which can be written in terms of expected values as $\sigma^2(x) = \langle x^2 \rangle - \langle x \rangle^2$. The quantity $\sigma(x)$ is called the *population standard deviation*.

A function $z(\mathbf{x})$, where $\mathbf{x}$ is a vector of random variables with a joint PDF $f(\mathbf{x})$, is itself a random variable. Thus, the population mean of such a function is

$$\langle z \rangle = \int_V z(\mathbf{x}) f(\mathbf{x})\,d\mathbf{x}, \tag{7}$$

and the population variance is

$$\sigma^2(z) = \int_v [z(\mathbf{x}) - \langle z \rangle]^2 f(\mathbf{x})\,d\mathbf{x} = \langle z^2 \rangle - \langle z \rangle^2. \tag{8}$$

2.3. The basis of Monte Carlo

The MC method is based on the law of large numbers, first enunciated by Jacob Bernoulli in 1689 and published posthumously (Bernoulli, 1713), which states that the mean value of a sample of N trials selected from a distribution $f(x)$, i.e.,

$$\overline{x} \equiv \frac{1}{N}\sum_{i=1}^N x_i, \tag{9}$$

where x_i is the result of the ith trial, approaches the true or population mean $\langle x \rangle$ as N gets large. The right-hand side of Eq. (9) is sometimes called the *tally* or *score* of the MC game used to generate the N values of x_i.

There are weak and strong versions of the law of large numbers but the differences need not concern us here. The basic fact is that, very generally,

$$\lim_{N\to\infty} \overline{x} = \langle x \rangle, \tag{10}$$

as long as the x_i are drawn from the governing PDF. We call $\overline{x}$ the *sample mean* and define the *sample variance* to be

$$s^2 = \frac{1}{N-1}\sum_{i=1}^N [x_i - \overline{x}]^2. \tag{11}$$

The law of large numbers also can be used to show that the sample mean for a function of a vector of random variables, namely

$$\overline{z} \equiv \frac{1}{N}\sum_{i=1}^N z(\mathbf{x}_i) \tag{12}$$

also converges to the population mean given by Eq. (7) if the x_i are sampled from f.

The law of large numbers, however, does not address how large N must be in order to obtain good estimates. To estimate how rapidly $\overline{z}$ converges to $\langle z \rangle$ as N increases, we turn to the central limit theorem. One form of the central limit theorem states (Kendall and Stuart, 1977) that

$$\lim_{N\to\infty} \mathrm{Prob}\left\{ p \le \frac{\overline{z} - \langle z \rangle}{\sigma/\sqrt{N}} \le q \right\} = \frac{1}{\sqrt{2\pi}} \int_p^q e^{-u^2/2}\,du. \tag{13}$$

One can estimate the uncertainty in the MC estimate, using the calculated sample standard deviation s to approximate the usually unknown population standard deviation σ and re-write the central limit theorem in the form

$$\mathrm{Prob}\{|\overline{z} - \langle z \rangle| \le \lambda s/\sqrt{N}\} \simeq \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-u^2/2}\,du. \tag{14}$$

The central limit theorem implies that the probability the MC estimate differs from the true expected value by an amount less than $\lambda\sigma/\sqrt{N}$ is independent of $f(\mathbf{x})$, for large N, and, in fact, eventually behaves as a normal distribution with zero mean and unit standard deviation. Further, one can evaluate the integral in Eq. (14) numerically to generate the results given in Table 1. These results indicate that, for instance, the sample mean will differ from the population mean by less than $\sigma/\sqrt{n}$ with 68% confidence and by less than twice that with 95% confidence. In either case, one can estimate the error, for a given N, and then can reduce

Table 1
Confidence limits for various values of λ obtained from the central limit theorem.

λ	$\text{Prob}\left\{\lvert\overline{z}-\langle z\rangle\rvert\le\lambda\frac{\sigma}{\sqrt{N}}\right\}$	Confidence limit (%)
0.25	0.1974	20
0.5	0.3829	38
1	0.6827	68
2	0.9545	95
3	0.9973	99

the error by increasing N until an acceptable precision is obtained.[1]

2.4. Sampling

The MC method relies inherently on the concepts of sampling and scoring. Sampling is the process of selecting a particular value of a random variable from its governing PDF. There are several ways to sample from a continuous PDF. Computer-generated random numbers are almost always used in MC calculations. Such numbers are not truly random, since they are generated using deterministic algorithms. Thus, they are called pseudo-random numbers. A good pseudo-random number generator can produce extremely long sequences of different numbers that pass many statistical tests for randomness. Most sampling techniques are based on having a pseudo-random number generator that can produce sequences of pseudo-random numbers that are uniformly distributed on the unit interval.

The simplest way to generate random numbers from an arbitrary PDF of a single random variable is called straightforward sampling. This method begins by selecting a number, call it ρ_i, sampled randomly and uniformly between 0 and 1 using one of the standard pseudo-random number generators (see Gentle, 2000, or Fishman, 2003, for summaries of pseudo-random number generators). Then the desired value of x_i, distributed as $f(x)$, is obtained by solving

$$F(x_i) = \rho_i, \tag{15}$$

whose solution is, formally, $x_i = F^{-1}(\rho_i)$. This is an efficient method for generating random variates with a given distribution $f(x)$ but it is limited to those distributions whose inverse CDF is obtainable. This method, also known as the inverse CDF method, exploits the fact that CDFs are monotonic and thus it is possible to establish a one-to-one correspondence between two CDFs, i.e.,

$$F(x_i) = G(y_i), \tag{16}$$

where x is distributed according to f and y is distributed according to g. If G is the unit CDF on the unit interval then Eq. (16) reduces to Eq. (15).

Many other sampling algorithms exist that can be applied when it is difficult to invert the CDF. One is called rejection or acceptance–rejection sampling (e.g., see Fishman, 2003) and consists in selecting a sample and then checking an acceptance criterion. This is very powerful but is less efficient than straightforward sampling because multiple pseudo-random numbers must be generated for each accepted sample. Nevertheless, it has found use in radiation transport simulations, for instance, to sample the angle of scatter in Compton interactions.

[1] As N gets very large, computer roundoff in the MC scores eventually may limit the attainable precision; thus, the central limit theorem gives the *theoretical* bound to the precision of a MC estimate.

Sampling from a joint PDF usually proceeds by writing the joint PDF in terms of a marginal PDF and conditional PDFs, namely

$$f(\mathbf{x}) = f(x_n|\mathbf{x}_{n-1})f(x_{n-1}|\mathbf{x}_{n-2}),\ldots,f(x_3|\mathbf{x}_2)f(x_2|x_1)f_1(x_1). \tag{17}$$

Then one samples x_1 from the marginal PDF f_1, x_2 from the conditional PDF for x_2 given x_1, and so on until x_n is sampled from the conditional PDF for x_n given $\mathbf{x}_{n-1}$. Another scheme to sample joint PDFs is based on Markov chains and thus is called Markov chain Monte Carlo (MCMC). A Markov chain is a sequence of stochastic variables for which the next entry depends only on the current entry and is independent of all previous entries. The MCMC formalism uses rejection–acceptance criteria and "random-walk" concepts in which one dimension is selected at random and a step in that direction is then sampled. Hence, MCMC often is used for hypothesis testing of multi-dimensional models of complex processes (Gilks et al., 1996); however, it is not widely used in radiation transport simulation.

2.5. Variance reduction

It is possible to bias the sampling during a simulation history in order to reduce the sample or tally variance. Of course, a biased simulation would lead to an incorrect result, so the tally must be corrected to account for the bias. This is done by introducing a weight W_i to correct the bias given to the ith sample z_i. For example, if a bias is introduced that makes it twice as likely to obtain the value z_i, its weight is set to 1/2. Thus, an unbiased estimate is given by

$$\overline{z} = \frac{1}{N}\sum_{i=1}^{N} W_i z_i. \tag{18}$$

When no biasing is introduced into the simulation, $W_i = 1$ and Eq. (18) reduces to Eq. (12).

The simplest form of variance reduction is called importance sampling. If one wishes to estimate

$$\langle z\rangle = \int_a^b z(x)f(x)\,dx \tag{19}$$

by Monte Carlo but cannot easily sample from f, one can multiply the integrand by $f^*(x)/f^*(x)$, which is unity, and re-arrange to obtain

$$\langle z\rangle = \int_a^b z(x)\frac{f(x)}{f^*(x)}f^*(x)\,dx. \tag{20}$$

Here $f^*(x)$ is a PDF defined on the interval $[a,b]$ but one from which it is easy to sample. The MC estimate of $\langle z\rangle$ can then be written

$$\overline{z} = \frac{1}{N}\sum_{i=1}^{N} W_i z(x_i), \tag{21}$$

where in this case the weight factor is given by $W_i = f(x_i)/f^*(x_i)$, and the x_i are sampled from f^* instead of f. It can be shown that although the sample mean tends to the population mean, the sample variance does not tend toward the population variance. If the importance function f^* is well chosen, the sample variance can be considerably smaller than that for unbiased sampling.

Gentle (2000) and many others discussed a variety of variance reduction schemes such as analytic reduction, antithetic variates, stratified sampling, common variates, constrained sampling, conditional Monte Carlo, and the random quadrature method. Russian roulette and splitting are easy to implement and can be very effective in reducing variance in radiation transport problems (Shultis and Faw, 2000). Indeed the "art" of using Monte Carlo is in finding effective variance reduction methods for the problem at hand.

3. Monte Carlo simulation of direct problems

3.1. Radiation transport simulation

Radiation detectors are used to capture information about a radiation field. Characterizing a radiation field can be of extreme importance, for instance, in understanding the output of high energy physics experiments to study the basic particles and forces of nature, in detecting radiation beams to create images of the human body, in monitoring the power level of power reactors, and in assessing the dose to astronauts due to cosmic radiation. If all of these situations, and others, could be treated adequately by analytical methods, there might be no need for Monte Carlo.

The diffusion or transport of radiation particles through a host medium is described by the linearized transport equation first enunciated in the 1870s by Ludwig Boltzmann. This equation is quite general and, therefore, quite difficult to solve. Analytical solutions of the governing transport equation are known only for highly simplified problems such as those involving a purely absorbing medium or monoenergetic particles in an infinite homogeneous medium. In most practical problems, the distribution in space, energy and direction (phase space $\mathbf{P}$) of the radiation cannot be obtained analytically. Sometimes numerical approximations to the transport equation are used to estimate the radiation field or detector response, but such deterministic methods are available only for relatively simple geometries. Monte Carlo provides a very practical way to estimate detector responses in arbitrary geometries involving complex radiation fields. Use of MC to predict detector responses, for example, is a form of "direct" Monte Carlo. Today the MC method has become, because of the maturation of digital computer technology, a widely used method for radiation transport simulation.

The steady-state integral transport equation (Duderstadt and Martin, 1979) can be expressed in the general form

$$\phi(\mathbf{P}) = \int \phi(\mathbf{P}')K(\mathbf{P}' \to \mathbf{P})\, d\mathbf{P}' + S(\mathbf{P}), \tag{22}$$

where $\mathbf{P}$ and $\mathbf{P}'$ represent some subsets of six-dimensional phase space, ϕ is flux density, K is a transport kernel, and S accounts for external sources. Suppose the response $R(E)$ of a detector of volume V_d to a radiation particle of energy E is sought. Formally, this is obtained by multiplying Eq. (22) by $\Re(\mathbf{r}_d, E, \mathbf{\Omega})$, the detector response produced by a unit flux density of energy E and direction $\mathbf{\Omega}$ at position $\mathbf{r}_d$ in the detector, and then integrating this result over all directions and the detector volume. The result is

$$R(E) = \int_{V_d} d\mathbf{r}_d \int_{4\pi} d\Omega \Big\{ \Re(\mathbf{r}_d, E, \mathbf{\Omega}) \int_{4\pi} d\Omega' \int_0^\infty dE' \int_V d\mathbf{r}' \times [\phi(\mathbf{r}', E', \mathbf{\Omega}')K(\mathbf{r}' \to \mathbf{r}_d, E' \to E, \mathbf{\Omega}' \to \mathbf{\Omega})] + S(\mathbf{r}_d, E, \mathbf{\Omega}) \Big\}. \tag{23}$$

Now Eq. (23) is in the form of an 11-fold expectation of the form of Eq. (7). The joint PDF is embedded in the kernel K. Hence, the integral form of the transport equation makes it clear that Monte Carlo is a natural method to estimate solutions to radiation transport problems.

It is, of course, not necessary to write explicit equations such as Eqs. (22) and (23) in order to solve radiation transport problems, but this approach does provide a tie to the earlier MC mathematical formulation. In fact, most radiation transport simulation proceeds on physical grounds, in which a radiation particle is tracked through its history from birth to death and the results of many histories are averaged to estimate the quantity of interest. Also, the quantity of interest often is not the flux density itself but rather a quantity that depends on flux density such as the response of a detector (as above), the energy deposited in a certain region, or the number of fissions produced in one region of a reactor from neutrons born in another.

3.2. Detector spectral analysis

Exemplary of the power of MC methods in application of radiation detectors is the use to which MC has been put in developing detector response functions. The spectral response of a detector can be expressed in the following general form:

$$R(E) = S_o \int_0^\infty \chi(E_o)K(E_o \to E)\, dE_o, \tag{24}$$

where S_o is the number of source particles incident on the detector during the counting time, $\chi(E_o)$ is the PDF defining the energy distribution of source particles incident on the detector, $K(E_o \to E)$ is the detector response function or the response at energy E due to a unit source at energy E_o, and $R(E)$ is the response of the detector per unit energy at energy E. For data obtained with multichannel analyzers, Eq. (24) is usually written in a discrete form, namely

$$c_i = S_o \sum_{j=1}^{J} \chi_j K_{j\to i}, \quad i = 1, 2, \ldots, N, \tag{25}$$

where c_i is the number of counts in channel i during the counting time, N is the number of channels, χ_j is the discrete PDF describing the number of source particles incident on the detector within the jth energy interval, J is the number of energy intervals considered, and $K_{j\to i}$ is the response matrix describing the probability that a particle within energy interval j produces a count in energy channel i. This formulation recognizes that, for instance, a photon of energy 1 MeV incident on a detector can produce a count in essentially any channel below the one corresponding to 1 MeV and, because of the finite resolution of the detector, to several above that channel.

Spectral analysis used to consist in estimating the total number of counts under a full-energy peak and relating this to the source intensity $S_o\chi_j$. However, this neglects the information available in rest of the spectrum. The use of Monte Carlo to estimate detector response functions has revolutionized spectral analysis in the recent years (see e.g., Gardner and Sood, 2004). The reason Monte Carlo is so important in this respect is that it is difficult to obtain real sources that emit monoenergetic photons at all energies of interest. However, once a MC code has been developed to calculate detector response functions accurately, it is trivial to specify any source energy.

4. Inverse and optimization models

Inverse problems abound in the design of detectors and analysis of detector measurements. A design inverse problem might be to use Monte Carlo to design a detector that will have desired or optimum response characteristics. For instance, Monte Carlo is widely used in the high energy physics community to design complex detectors that surround interaction points in large accelerators and colliders. Another type of detector inverse problem to which Monte Carlo can be applied uses MC techniques to determine some parameter or parameters of the radiation's host medium (such as density and composition) from one or more measured detector responses.

A general class of such inverse problems can be expressed in the form

$$R(y_k) = \int_V z(\mathbf{x}, y_k) f(\mathbf{x}|\boldsymbol{\alpha})\, d\mathbf{x}, \quad k = 1, 2, \ldots, K, \tag{26}$$

where R is a detector's response as a function of some independent variable y (e.g., energy or position), f has the properties of a probability density function over the n-dimensional volume V, and $\boldsymbol{\alpha}$ is a vector of J unknown parameters. Then Eq. (26) is an

inverse problem for $\boldsymbol{\alpha}$ given values of R at K values of y. In principle, the α_j can be parameters in a known functional form (see Section 4.1) or the amplitudes of a histogram approximation of an unknown function (see Section 4.2).

The traditional way to use MC to solve inverse problems is to iterate. Specify a design $\boldsymbol{\alpha}$ and then calculate the response R for the K assumed values of y. If the calculated response does not match the desired response, change the design parameters and estimate a new set of responses. Iterate until the MC estimate matches the desired response with sufficient precision. However, very many iterations might be required in order to estimate the solution if the number of design parameters is large. If each iteration requires a lengthy simulation, the computational expense may become inordinately high.

It is possible, in principle, to use MC to solve inverse problems with only one (or at least only a few) MC simulation steps (Dunn, 1981). Conceptually, this method proceeds as follows. Construct a single forward simulation by first sampling from an assumed PDF $f^*(\mathbf{x})$ also defined on V, and then weight the score of each history in the same manner as used in importance sampling to obtain

$$\overline{R}_k = \frac{1}{N}\sum_{i=1}^{N} z(\mathbf{x}_i, y_k)\frac{f(\mathbf{x}_i|\boldsymbol{\alpha})}{f^*(\mathbf{x}_i)}, \quad k = 1, 2, \ldots, K. \tag{27}$$

Here, the unknown parameters are embedded in the history scores and Eq. (27) forms a system of K equations in J unknowns, which can be solved by standard techniques if $J \le K$. Note that only one forward simulation is required, in principle. We refer to this form of solving inverse problems by MC as symbolic Monte Carlo (SMC). Symbolic Monte Carlo has found application, e.g., in radiative transfer (Dunn, 1983), in medical imaging (Floyd et al., 1985, 1986), in energy-dispersive X-ray fluorescence analysis (Yacout and Dunn, 1987), and in photon beam modifier design (Dunn and Womersley, 1994).

4.1. An SMC linear PDF example

Assume that a particular response can be expressed as

$$R(y) = C\int_a^b e^{-\lambda x/y} f(x)\,dx, \tag{28}$$

where a, b, λ, and C are specified constants. Here $f(x) = \alpha_1 + \alpha_2 x$ is a PDF but α_1 and α_2 are unknown. There are, of course, infinitely many linear PDFs on any finite interval $[a, b]$, but they all obey the normalization condition, which requires that

$$\alpha_1 = \frac{1}{b-a} - \alpha_2\frac{b+a}{2}. \tag{29}$$

Thus, we seek the value of α_2 that best reproduces a given response. The SMC tally for this problem can be written as

$$\overline{R}_k = C\frac{b-a}{N}\sum_{i=1}^{N} e^{-\lambda x_i/y_k}\left[\frac{1}{b-a} + \alpha_2\left(x_i - \frac{b+a}{2}\right)\right] \tag{30}$$

for any value y_k, where the x_i are sampled from the uniform PDF $f^*(x) = 1/(b-a)$. To evaluate $\overline{R}_k$ two summations must be accumulated during the MC simulation, namely

$$S_1 = \frac{1}{N}\sum_{i=1}^{N} e^{-\lambda x_i/y_k} \quad \text{and} \quad S_2 = \frac{1}{N}\sum_{i=1}^{N} x_i\, e^{-\lambda x_i/y_k}. \tag{31}$$

With values of S_1 and S_2, Eq. (30) can be solved to yield

$$\alpha_2 = \frac{R_k - CS_1}{C(b-a)S_2 - C(b^2-a^2)S_1/2}, \tag{32}$$

where the known or measured R_k replaces the symbol $\overline{R}_k$.

We considered the case $C = 10{,}000$, $a = 1$, $b = 10$, $\lambda = 0.2$, $K = 1$, and $y = 1$, for which we find $R_1 = 2920.3$ (which can be calculated from Eq. (28) for the case in which $\alpha_1 = 0.00111$ and $\alpha_2 = 0.02$). The results in Table 2 show the SMC solution of Eq. (32) for α_2 and the corresponding value of α_1 obtained from Eq. (29). The SMC solution appears to converge toward the correct answer. It is interesting to note that a relatively small number of histories (in this case $N = 100$) is often sufficient to recover the unknown parameter(s).

4.2. Estimating the slope of a flattening filter

Filters are widely used in photon beam therapy to produce desired dose distributions in phantoms and patients. Although very refined beam conditioning techniques are now used, we illustrate SMC for the simple case of a flattening filter placed in a downward collimated conical beam so as to produce in a phantom isodose profiles that are essentially parallel ("flat") at some angle to the horizontal. This treatment is slightly different from the solution given by Dunn and Womersley (1994), in order to illustrate a multi-dimensional inversion problem.

Consider the irradiation of a homogeneous slab phantom by a collimated photon source. The source beam is modified by a wedge filter, as shown schematically in Fig. 1. The dose D as a function of position $\mathbf{r}$ within the phantom can be expressed in the general form

$$D(\mathbf{r}) = \int_{\Delta\boldsymbol{\Omega}} K_D[\mathbf{r}, t(\boldsymbol{\Omega})] I_0(\boldsymbol{\Omega})\, d\boldsymbol{\Omega}, \tag{33}$$

where $\boldsymbol{\Omega}$ is the solid angle variable that specifies photon direction, $\Delta\boldsymbol{\Omega}$ is the solid angle interval that defines the beam collimation, $I_0(\boldsymbol{\Omega})$ is the source intensity in direction $\boldsymbol{\Omega}$, $t(\boldsymbol{\Omega})$ is the filter thickness along the ray emanating from the source in direction $\boldsymbol{\Omega}$, and $K[\mathbf{r}, t(\boldsymbol{\Omega},)]$ is a kernel that accounts for the transport of photons through the filter and the phantom to energy deposition

Table 2
Symbolic Monte Carlo results for α_2, Eq. (32), and α_1, Eq. (29), for various numbers of histories.

N	α_2	α_1
10^2	0.02021	−0.00007
10^4	0.02005	0.00081
10^6	0.02001	0.00117

The exact values are $\alpha_1 = 0.02000$ and 0.00111.

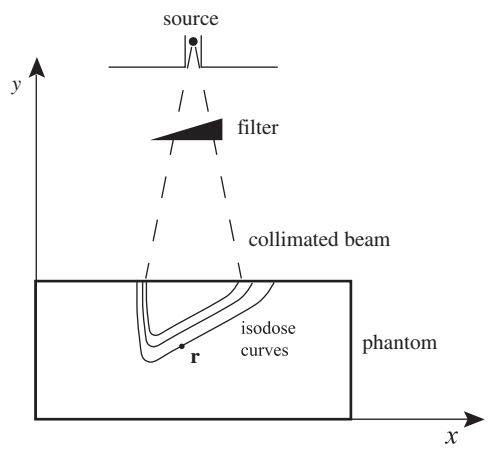

Fig. 1. A conically collimated photon beam passes through a wedge filter to produce in a phantom isodose profiles that are inclined at some specified angle with respect to the horizontal.

within a small volume about position **r** in the phantom. This is an inverse problem for $t(\mathbf{\Omega})$, given measurements of D at various positions **r**.

Isodose distributions were measured in a water tank exposed to a ^{60}Co source that was modified by each of two wedge filters made of lead. One was a so-called 30° wedge (the descriptor referring to the slope of the isodose curves at depth, not to the wedge angle) and the other a 60° wedge. The fields were of square cross section (10×10 cm) and the doses were measured within the principal plane of the collimated beam (the $z = constant$ plane in Fig. 1 that contains the source). The doses at $K = 60$ points were recorded for each of the wedges, yielding dose vectors $\mathbf{D} = \{D_k, k = 1, 2, \ldots, K\}$.

The SMC solution assumes that the wedge filter can be represented by a series of vertical steps perpendicular to the principal plane, each with a thickness t_j, $j = 1, \ldots, J$, as depicted schematically in Fig. 2 for $J = 8$. It is further assumed that multiple scatters in the filter can be ignored and, thus, the measured doses are due either to radiation transmitted through the wedge or to radiation that scatters once within the wedge. There are ways to relax this approximation, but consideration of this simplified case is sufficient for present purposes. The additional assumption is made that the dose distribution is insensitive to the precise depth of each scatter in the filter, which is a reasonable approximation for optically thin filters and large filter-to-phantom distances.

The SMC model selects a photon from the source, determines within which filter interval j it enters the filter, then samples to determine if the photon interacts in the filter. Since the thickness t_j is unknown, some approximate PDF must be specified, such as

$$f_j^* = \begin{cases} f_0, & \text{if no interaction occurs,} \\ 1 - f_0, & \text{if an interaction occurs,} \end{cases} \tag{34}$$

where $0 < f_0 < 1$. The true PDF is, of course,

$$f_j = \begin{cases} e^{-\mu t_j}, \text{if no interaction occurs,} \\ 1 - e^{-\mu t_j}, \text{if an interaction occurs,} \end{cases} \tag{35}$$

where μ is the total interaction coefficient of the filter material and t_j is the unknown thickness of filter segment j. To remove the error or bias introduced by using the wrong PDF requires multiplying the history weight by the factor f/f^*. Now, for those histories for which no interaction occurs in the filter, the photon is transported to the phantom surface and, within the phantom, the dose tally is incremented if the photon interacts within a small volume about one of the detector points. For those histories in which an interaction occurs in the filter, a scatter is forced by multiplying the history weight by the scattering probability σ/μ (where σ is the scattering interaction coefficient), a scatter angle is sampled, and the simulation proceeds with the assumption that the photon escapes the filter.

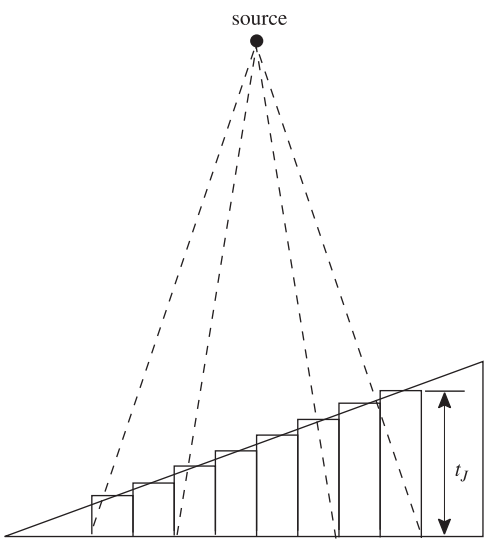

Fig. 2. The wedge filter is approximated by a series of steps, each with a thickness $t_j, j = 1, \ldots, J$.

The estimator for the dose at the kth detector position in the phantom can then be written as

$$\overline{D}_k = C\left\{\sum_{j=1}^{J} \frac{\alpha_j}{f_0 N_{0j}} \sum_{i=1}^{N_{0j}} A_{ki}^j + \sum_{j=1}^{J} \frac{1-\alpha_j}{(1-f_0)N_{1j}} \sum_{i=1}^{N_{1j}} B_{ki}^j\right\}, \tag{36}$$

where C is a conversion constant that scales the MC scores to the measured dose values, N_{0j} the number of histories that intersect filter interval j but do not interact in the filter, A_{ki} is the contribution to the dose at detector position k due to gamma ray i that was transmitted through region j of the filter, N_{1j} is the number of histories that intersect region j of the filter and scatter once in the filter, B_{ki} is the contribution to the dose at detector position k due to gamma ray i that scattered in filter region j, and $\alpha_j = e^{-\mu t_j}$. Eq. (36) forms a system of $K = 60$ algebraic for the J unknown values of α_j.

Eq. (36) was solved for the unknown α_j by fitting the model to the measured doses, i.e., $\overline{D}_k$ were set to the measured D_k. With the α_j determined, the J thicknesses were then found from $t_j = -\ln\alpha_j/\mu$. In the analysis $K = 60$ detector measurements and $J = 30$ filter thicknesses were used. Since the case considered is for a regular filter with a constant slope, a straight line was fit to the filter thicknesses. For the "30° wedge", whose actual wedge angle was 7.2°, an angle of 7.9° was estimated and for the "60° wedge", whose actual wedge angle was 16.1°, an angle of 17.5° was estimated. The fact that better agreement was not achieved is due to the many assumptions made (as evidenced by the fact that worse results were obtained for the thicker wedge). Nevertheless, this problem serves to illustrate that SMC can be used to solve, with only a single MC simulation, an inverse problem in which many parameters are unknown.

5. Summary

We have reviewed a formalism in which the concepts of MC can be used to solve both direct and inverse radiation detection problems. Monte Carlo is especially well suited to radiation transport simulation because it is highly efficient for multi-dimensional integration and radiation transports in a phase space of six spatial, energy, and direction dimensions. Further, the laws governing radiation transport in general are not only probabilistic in nature but also well understood and so can be sampled and scored in the MC sense. Realistically, MC is one of the most powerful tools that the radiation specialist has available to perform quantitative analysis of detectors or detector measurements.

Powerful general-purpose MC codes such as EGSnrc, FLUKA, GEANT, MCNP, MCSHAPE, PENELOPE, and many more exist for radiation transport simulation. Some of these codes are described in related papers in this issue. It is also possible to construct special-purpose codes that address specific problems. One particular advantage of constructing a special-purpose code is that the developer must have a thorough understanding of the details of the physics involved. Hence, MC can be a very useful vehicle for the education of future radiation specialists.

Users of MC realize that it is both a science and an art. The science requires that the concepts be well understood and

properly implemented. The art allows the user to use intuition in the model and take creative shortcuts in solving a given problem. The reach of Monte Carlo makes the method a powerful tool and the flexibility of MC makes it a joy to use.

References

Bernoulli, J., 1713. Ars Conjectandi, Impensis Thurnisiorum, Fratrun, Basilea.

de Buffon, C., 1777. Essai d'Arithmetique Morale. Supplement à l'Histoire Naturelle, vol. 4.

Duderstadt, J.J., Martin, W.R., 1979. Transport Theory. Wiley, New York, NY.

Dunn, W.L., 1981. Inverse Monte Carlo analysis. J. Comput. Phys. 41 (1), 154–166.

Dunn, W.L., 1983. Inverse Monte Carlo solutions for radiative transfer in inhomogeneous media. J. Quant. Spectrosc. Radiat. Transfer 29 (1), 19–26.

Dunn, W.L., Womersley, J., 1994. Solutions of inverse and optimization problems in high energy and nuclear physics using inverse Monte Carlo. In: Dragovitsch, P., Linn, S.L., Burbank, M. (Eds.), MC93 International Conference Monte Carlo Simulation in High Energy and Nuclear Physics. World Scientific, Singapore, pp. 78–87.

Fishman, G.S., 2003. Monte Carlo Concepts, Algorithms, and Applications, Corrected Printing. Springer, New York, NY.

Floyd, C.E., Jaszczak, R.J., Coleman, R.E., 1985. Inverse Monte Carlo: a unified reconstruction algorithm for SPECT. IEEE Trans. Nucl. Sci. NS-32, 779.

Floyd, C.E., Jaszczak, R.J., Greer, K.L., Coleman, R.E., 1986. Inverse Monte Carlo as a unified reconstruction algorithm for ECT. J. Nucl. Med. 27 (10), 1577–1585.

Gentle, J.E., 2000. Random Number Generation and Monte Carlo Methods. Springer, New York, NY.

Gardner, R.P., Sood, A., 2004. A Monte Carlo simulation approach for generating NaI detector response functions (DRFs) that accounts for non-linearity and variable flat continua. Nucl. Instr. and Meth. Phys. Res. B 213, 87–99 (NY).

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall, New York, NY.

Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics, vol. 1. Macmillan, New York, NY.

Laplace, P.S., 1786. Theorie analytique des probabilities. Livre 2, 356–366. In: Oeuvres Completes de Laplace, de L'Academie des Sciences, vol. 7 (part 2).

Metropolis, N., 1985. Monte Carlo: in the beginning and some great expectations. In: Proceedings of Joint Los Alamos National Laboratory-Commissariat à l'Energie Atomique Meeting, Cadarache Castle, Provence, France, Lecture Notes in Physics, vol. 240. Springer, New York, NY, pp. 62–70.

Metropolis, N., Ulam, S., 1949. The Monte Carlo method. J. Am. Stat. Assoc. 44 (227), 335–341.

Mosegaard, K., Sambridge, M., 2002. Monte Carlo analysis of inverse problems. Inverse Probl. 18, R29–R54.

Shultis, J.K., Faw, R.E., 2000. Radiation Shielding. American Nuclear Society, LaGrange Park, IL.

Yacout, A.M., Dunn, W.L., 1987. Application of the inverse Monte Carlo method to energy-dispersive X-ray fluorescence. Adv. X-Ray Anal. 30, 113–120.