MONTE CARLO METHODS : AN INTRODUCTION

K.P.N. Murthy
School of Physics, University of Hyderabad
Hyderabad - 500134, India

and

Reactor Research Centre, Kalpakkam 603 102
Tamil Nadu, India

1. Introduction

Monte Carlo can be defined as a numerical method that uses random numbers to solve a problem. The methods of Monte Carlo have become quite popular in the recent times, thanks to the availability of high speed computers, and are being applied to a growing variety of problems. Deterministic as well as stochastic problems can be handled. For solving a deterministic problem, we construct first a stochastic model. Then we carry out a Monte Carlo simulation of this artificial stochastic model on a computer. Multi-dimensional integration, matrix inversion, solution of Dirichlets problem etc. are a few examples of the deterministic problems that have been successfully solved using Monte Carlo.

For stochastic problems, however, we can resort to direct simulation. A notable example of this is the study of neutron transport in nuclear reactors. Indeed, the identification and systematic exploitation of Monte Carlo as a powerful tool for solving practical problems started with nuclear applications and the first investigators were von Neumann, Ulam, Fermi, Harris, Herman Kahn, Metropolis and others. A brief account of the interesting history of Monte Carlo and reference to some of the early works can be found in the monograph of Hammersley and Handscomb [1].

If the stochastic problem on hand is complex (or if we are not interested in the microscopic details of the solution process), we can construct a simpler stochastic model of the problem and then simulate this model on a computer. Monte Carlo simulation of Brownian motion modelled by the Langevin equation, study of kinetic Ising models, percolation problems etc. come under this category.

In this lecture I shall try to present briefly the basic features of Monte Carlo and illustrate its application to the study of nonlinear Brownian motion.

2. Elements of Monte Carlo

The important features of Monte Carlo can be easily understood by considering the following simple problem.

Let x be a random variable and f(x) its probability density function. We set to determine the average of a function h(x) given by

$$\mu = \int_{-\infty}^{\infty} h(x) f(x) dx \qquad (1)$$

Monte Carlo method of evaluating μ consists of selecting a set of N independent random values x_i, i = 1, N from the density f(x) and carrying out the following summation

$$\bar{h}_N = \frac{1}{N} \sum_{i=1}^{N} h(x_i) \qquad (2)$$

$\bar{h}_N$ given above is a Monte Carlo estimate of the desired answer μ. Sampling of x from the density f(x) is accomplished most conveniently as follows. First generate a set of random numbers between 0 and 1 from a uniform density. (We shall refer to these as random numbers in all subsequent discussions, and denote them by ξ). Having generated a random number ξ, we shall perform some appropriate transformation to obtain x for use in (2). Such techniques are called random sampling techniques. For example in the inversion technique, we set $X = F^{-1}(\xi)$ where F is the probability distribution (or cumulative density) function of X. There are numerous random sampling techniques like rejection technique, equiprobability table method, composition and decomposition methods etc. for different kinds of density functions. For a review of these techniques see [2]. There are many algorithms to generate random numbers on a computer. These numbers should be strictly called pseudo random numbers since there is nothing random about their source; for, they are produced by deterministic procedures. We are justified in using these numbers in our problem since our concern lies not in the source of these numbers but on whether they are 'correctly' distributed. A sequence of numbers that pass through the required set of statistical tests for randomness should suffice for our purpose. Examples of (pseudo) random number generators are the early mid square method of Metropolis and von Neumann (see [3]), the multiplicative congruential method of Lehmer [4] etc. We shall not go into the details of these techniques. Those interested can see [1-5] and the references therein. See also Chaitin [6] for an interesting discussion on randomness and mathematical proof.

Returning to (2) it is desirable to quantify the confidence we repose on $\bar{h}_N$ as an approximation to μ and this is afforded by the central limit theorem. It states that as the sample size N tends to ∞ the probability density of the random variable $\bar{h}_N$ tends to be Gaussian with mean μ and variance σ^2/N where σ^2 is the variance of the problem. Thus $\pm \sigma/\sqrt{N}$ gives one-sigma confidence interval (also called statistical error). Since σ^2 is unknown we use its Monte Carlo estimate S_N^2 (the sample variance) to obtain the statistical error. The expression for S_N^2 is

$$S_N^2 = \frac{1}{N-1} \sum_{i=1}^{N} h^2(x_i) - \frac{N}{N-1} (\bar{h}_N)^2 \qquad (3)$$

3. Variance Reduction Techniques

It is clear that the statistical error in Monte Carlo estimates can be decreased by increasing N, the sample size or decreasing σ. The latter implies that we distort the problem on hand so that the altered problem has a smaller variance but the same

mean. Such variance reduction techniques have become imperative to Monte Carlo simulation and considerable work is being carried out in this area, see for e.g. [7-9]. Let me illustrate the technique of variance reduction by considering importance sampling. This essentially consists of using an importance density function $g(\alpha, x)$ instead of the actual density $f(x)$ for evaluating μ (see (1)). α here is an adjustable parameter, significance of which will become clear soon. Now, for purpose of preserving mean, we consider the function $H(x)$ in place of $h(x)$. Clearly if $H(x)$ is chosen as $h(x)f(x)/g(\alpha, x)$ the mean of $H(x)$ in the distorted space of $g(\alpha, x)$ is the same as μ but the variance is

$$\sigma_1^2(\alpha) = \int_{-\infty}^{\infty} \left\{ \frac{h(x)f(x)}{g(\alpha, x)} \right\}^2 g(\alpha, x)\, dx \qquad (4)$$

A judicious choice of $g(\alpha, x)$ would lead to considerable variance reduction. Indeed one can do better by optimizing the importance function during the process of simulation itself, by splitting it into various stages and improving the parameter from stage to stage. This can be accomplished as follows. We rewrite (4) as

$$\sigma_1^2(\alpha) = \int_{-\infty}^{\infty} \frac{h(x)f(x)}{g(\hat{\alpha}, x)} \frac{h(x)f(x)}{g(\alpha, x)} g(\hat{\alpha}, x)\, dx \qquad (5)$$

where $\hat{\alpha}$ is an initial guess of the parameter α. The procedure consists of first generating a small sample of x from $g(\hat{\alpha}, x)$. From this sample, using (5) evaluate the sample variance at different values of α. Determine the value of α at which σ_1^2 is minimum. Use this value in the next stage and proceed as before. After a few stages, α converges to its optimum value and this can be used in the final stage where a comparatively large sample is chosen to estimate the mean. This powerful self learning technique due to Spanier has been successfully used in many problems [7-9,11].

Besides importance sampling there are many variance reduction techniques like systematic sampling, stratified sampling, correlated sampling, control variates, antithetic variates, regression methods etc. Ref.[10] provides a review of the numerous variance reduction techniques.

4. Monte Carlo Simulation of Non-Linear Brownian Motion

In this part we shall consider an application of Monte Carlo to study nonlinear Brownian motion, modelled by the Langevin equation,

$$\frac{dx}{dt} = c(x) + \eta(t) \qquad (6)$$

where $x(t)$ is the driven stochastic process, $c(x)$ is the nonlinear drift term and $\eta(t)$ is the driving Gaussian white noise satisfying the following properties

$$\langle \eta(t) \rangle = 0 \; ; \quad \langle \eta(t)\, \eta(t') \rangle = 2 \epsilon \delta(t-t') \qquad (7)$$

ϵ in the above is the diffusion constant. To simulate the stochastic process $x(t)$ we first discretise the time variable. Correspondingly the random force $\eta(t)$ is

also replaced by a set $\{\eta_i\}$ at discrete times $\{t_i\}$. The property (7) now becomes

$$\langle \eta_i \rangle = 0 \; ; \; \langle \eta_i \, \eta_j \rangle = \sigma^2 \, \delta_{ij} \tag{8}$$

We consider finite difference approximation to (6) given by

$$x_j = x_{j-1} + \Delta t . \, c(x_{j-1}) + \Delta t \eta_{j-1} \tag{9}$$

Monte Carlo simulation of x(t) is carried out as follows. A sample track $\{x_o, x_1, x_2 \; x_3 \ldots\ldots\}$ is generated by first sampling x_o from the given initial distribution of the random process. Then we add successively a deterministic increment, $c \, \Delta t$ and a random increment $\eta \, \Delta t$ where η is sampled from a Gaussian of mean 0 and variance σ^2. The evolution of x(t) is considered upto the desired time. The whole process is repeated and a sample of N tracks is obtained. The required time dependent statistics of the driven stochastic process like mean, variance, correlations etc. are calculated by explicit averaging over the finite ensemble of N tracks.

In the above, random force η is to be sampled from a Gaussian of variance σ^2. The relation between σ^2 and the diffusion constant ϵ is obtained as follows. It is clear that

$$\int_o^t \eta(t) dt = \lim_{\Delta t \to o} \sum_i^{Lt} \eta_i \Delta t \tag{10}$$

Then (7) yields

$$\left\langle \int_o^t \eta(t') \, \eta(t'') \, dt'' \right\rangle = 2\epsilon \tag{11}$$

from which it follows

$$\sum_i \langle \eta_i \eta_j \rangle \, \Delta t = 2\epsilon \tag{12}$$

Thus we get

$$\sigma^2 = 2\epsilon / \Delta t \tag{13}$$

For generating the random forces $\eta_j, j = 1,2 \ldots\ldots$ we use central limit theorem. This consists of selecting M random numbers and taking their sum $S = \sum \xi_i$. The desired Gaussian variate (mean zero and variance σ^2) is given by

$$\eta = \sigma \left\{ \sqrt{12/M} \; - S - \sqrt{3.M} \right\} \tag{14}$$

For large M, η has the required Gaussian density. In practice M = 12 is sufficient [5].

For computing deterministic evolution of the process over the time interval Δt, one can use sophisticated algorithms instead of the finite difference scheme shown in (9). This would help improve the simulation considerably. Recently we studied the problem of diffusion in bistable potential [12,13] where the nonlinear drift term of (6) has the form $c(x) = \gamma x - g x^3$. Here γ and g are positive real constants. We have used Runge-Kutta Hill method for the deterministic evolution. In Fig.1 we

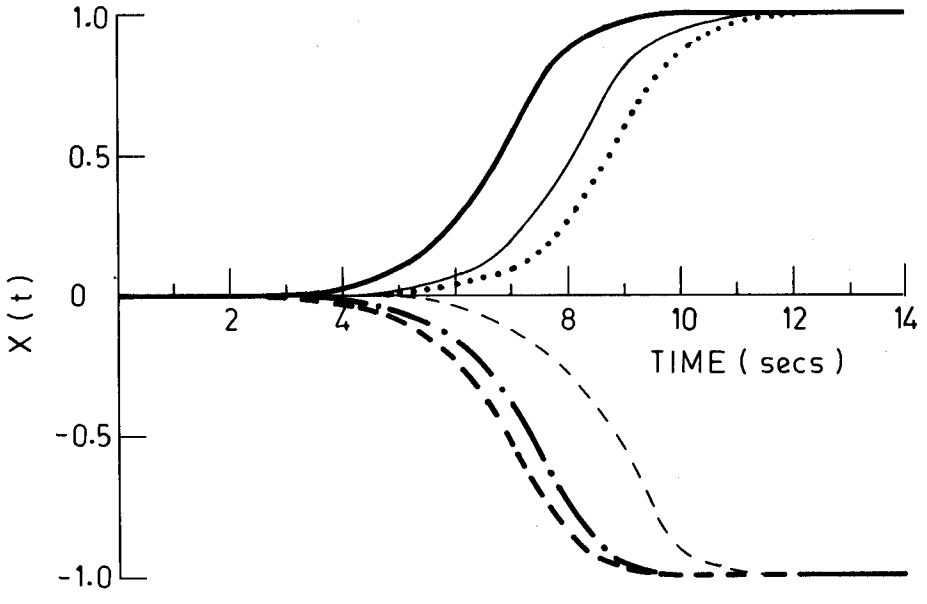

Fig.1 Sample tracks of Nonlinear Brownian Motion from an initial unstable steady state

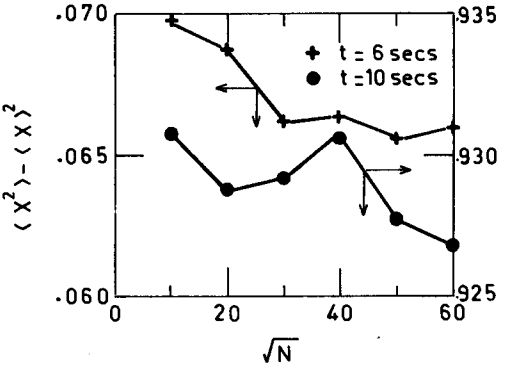

Fig.2 Variation of ensemble average of fluctuations with sample size at two typical times

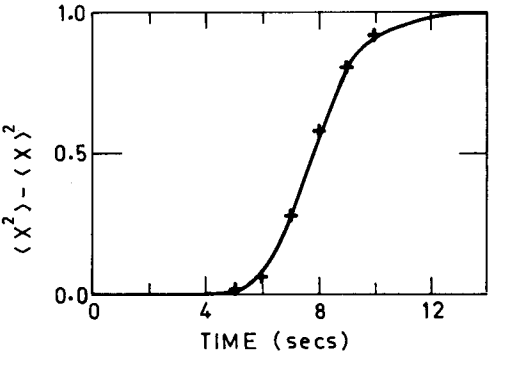

Fig.3 Fluctuations as a function of time. Ensemble average over 3600 tracks.

show a few of the tracks that were generated with the parameters $\epsilon = 0.5 \times 10^{-6}$, $\gamma = g = 1$ and the deterministic initial condition $x_o = 0$. Figure 2 depicts Monte Carlo estimates of the variance at a two typical times as a function of the sample size N. This illustrates the nature of convergence of the sample estimates. We show the fluctuation averaged over an ensemble of 3600 tracks, as a function of time in Fig.3.

5. Conclusions

I have tried to present the elementary notions of the method of Monte Carlo. This presentation undoubtedly suffers from my personal bias. Unlike other numerical techniques, Monte Carlo has to be developed each time taking into account the special requirements of the problem. There is indeed a great scope for ingenuity both in modelling of the problem and in developing special methods of simulation.

References

1. Hammersley J.M. and Handscomb D.C., Monte Carlo Methods, Chapman and Hall (1979).
2. McGrath E.J., and Irving D.C., Random Number Generation for Selected Probability Distributions, Vol.II of Techniques of Efficient Monte Carlo Simulation, Report ORNL-RSIC-38, ORNL (1975)
3. Forsythe G.E., NBS Applied Mathematics Series 12, 13 (1951)
4. Lehmer D.H., Ann. Comp. Lab. Harvard Univ. 26, 141 (1951)
5. James F., Monte Carlo Theory and Practice, Report DD/80/6, CERN Data Handling Division (1980)
6. Chaitin G.J., Scientific American 232, 47 (1975)
7. Spanier J., SIAM J. Appl. Maths., 18, 172 (1970)
8. Murthy K.P.N., Ann. Nucl. Ener. 7, 389 (1980)
9. Ragheb M.M.H et al, Atomkernenergie, Kertechnik 37, 188 (1981)
10. McGrath E.J. and Irving D.C., Variance Reduction Vol.III of Techniques of Efficient Monte Carlo Simulation, Report ORNL-RSIC-38, ORNL (1975)
11. Murthy K.P.N., Atomkernenergie, Kettechnik 34, 125 (1979)
12. Suzuki M., Adv. Chem. Phys. 46, 195 (1981) and the references therein
13. Valsakumar M.C., Murthy K.P.N. and Ananthakrishna G., J. Stat. Phys. 30, 637 (1983)