

MONTE CARLO SIMULATION OF DIFFUSIONS

Peter W. Glynn

Department of Management Science and Engineering
 Stanford University
 Stanford, C.A. 94305, U.S.A.

ABSTRACT

This tutorial is intended to provide an overview of the key algorithms that are used to simulate sample paths of diffusion processes, as well as to offer an understanding of their fundamental approximation properties.

1 INTRODUCTION

Diffusion process and their generalization to jump diffusions are widely used as mathematical models within the physical sciences, engineering, finance, and economics. The modern perspective on diffusions is to view such stochastic processes as arising as \mathbb{R}^d -valued solutions $X = (X(t) : t \geq 0)$ of stochastic differential equations (SDE's). The simplest such SDE's evolve on a state space S that is an open subset of \mathbb{R}^d (e.g. $S \subseteq \mathbb{R}^d$), in which X never touches the boundary of S . In such cases, the process X is fully described by a stochastic equation of the form

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t) \quad (1.1)$$

subject to $X(0) = x \in S$, where $\mu : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ are deterministic functions specified by the modeler (known as the "drift" and "volatility" functions, respectively), and $B = (B(t) : t \geq 0)$ is a \mathbb{R}^m -valued standard Brownian motion (i.e. B is a continuous path Gaussian process with $EB(0) = 0$ and $EB(t)B(t)^T = t \cdot I$, where x^T denotes the transpose of x and I is the identity matrix). When X can reach the boundary of S with positive probability, the modeler must specify the behavior of X at the boundary, leading to an "SDE with boundaries" (or, equivalently, a "diffusion with boundaries").

The rigorous mathematical meaning of (1.1) is to view (1.1) as a short-hand for requiring that X satisfy the stochastic integral equation

$$X(t) - x = \int_0^t \mu(s, X(s)) ds + \int_0^t \sigma(s, X(s)) dB(s) \quad (1.2)$$

for $t \geq 0$. In this integral form, it is evident that one must make mathematical sense of the stochastic integral

$$\int_0^t \sigma(s, X(s)) dB(s) \quad (1.3)$$

in order to construct a mathematically rigorous theory of the SDE's (and, hence, diffusions). As is well known, there are several different mathematically satisfactory ways of defining such stochastic integrals which, unfortunately, do not coincide with one another (in the sense that the distribution of the random variable (rv) defined by (1.3) typically depends on the definition that is used). However, for reasons that will become clearer in the next section, the Itô stochastic integral is generally the one that is understood to underlie the stochastic integral appearing in (1.2).

2 THE CONNECTION TO PDE'S

There exists a large number of probabilities and expectations for diffusions that can be computed by solving partial differential equations (PDE's). To get a sense of the connection, suppose that $\mu(t, X(t))$ and $\sigma(t, X(t))$ depends only on $X(t)$ (and not explicitly on t) and that we wish to compute

$$u^*(x) = E_x \int_0^T \exp\left(-\int_0^t \alpha(X(s)) ds\right) g(X(t)) dt + E_x \exp\left(-\int_0^T \alpha(X(t)) dt\right) f(X(T)),$$

where $T = \inf\{t \geq 0 : X(t) \in A\}$ is the "first hitting time" of the "target set" $A \subseteq S$, and f, g , and α are given functions. Clearly, $u^*(x) = f(x)$ for $x \in A$. For $x \in A^c$, we note that

for $h > 0$,

$$u^*(x) = E_x \int_0^{h \wedge T} \exp\left(-\int_0^t \alpha(X(s)) ds\right) g(X(t)) dt + E_x \exp\left(-\int_0^{h \wedge T} \alpha(X(s)) ds\right) u^*(X(h \wedge T)) \quad (2.1)$$

where $h \wedge T \triangleq \min(h, T)$. Since we expect the solution X of (1.2) to have continuous paths, it should follow that for $T > h$

$$\int_0^{h \wedge T} \exp\left(-\int_0^t \alpha(X(s)) ds\right) g(X(t)) dt = g(X(0))h + o(h) \quad (2.2)$$

and

$$\exp\left(-\int_0^{T \wedge h} \alpha(X(s)) ds\right) = 1 - \alpha(X(0))h + o(h) \quad (2.3)$$

as $h \downarrow 0$, where $o(r(h))$ defines a function for which $o(r(h))/r(h) \rightarrow 0$ as $h \downarrow 0$. On the other hand, if A is closed and $x \in A^c$, the likelihood of hitting A from x in the interval $[0, h]$ is typically exponentially small, so that

$$P_x(T \leq h) \leq \exp(-c/h) \quad (2.4)$$

for some $c > 0$. Relations (2.2) and (2.4) suggest that

$$E_x \int_0^{h \wedge T} \exp\left(-\int_0^t \alpha(X(s)) ds\right) g(X(t)) dt = g(x)h + o(h), \quad (2.5)$$

as $h \downarrow 0$.

To handle the second term on the right-hand side of (2.1), assume that u^* is smooth. Write $X(t) = (X_1(t), \dots, X_d(t))^T$, so that $X_i(t)$ is the i 'th component of X . Similarly, $B_i(t)$ is the i 'th component of $B(t)$. The smoothness of u^* guarantees that when $T > h$,

$$\begin{aligned} & u^*(X(T \wedge h)) - u^*(X(0)) \\ &= u^*(X(h)) - u^*(X(0)) \\ &= \sum_{i=1}^d \frac{\partial u^*}{\partial x_i}(X(0))(X_i(h) - X_i(0)) \\ &+ \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 u^*}{\partial x_i \partial x_j}(X(0))(X_i(h) - X_i(0))(X_j(h) - X_j(0)) \\ &+ \frac{1}{6} \sum_{i,j,k=1}^d \frac{\partial^3 u^*}{\partial x_i \partial x_j \partial x_k}(X(0))(X_i(h) - X_i(0))(X_j(h) - X_j(0)) \\ &\cdot (X_k(h) - X_k(0)) + \dots \end{aligned} \quad (2.6)$$

The continuity of X suggests that

$$\int_0^h \mu(X(s)) ds = \mu(X(0))h + o(h). \quad (2.7)$$

Furthermore, we expect the stochastic integral in (1.3) to obey

$$\int_0^h \sigma(X(s)) dB(s) \approx \sigma(X(0))(B(h) - B(0)) \quad (2.8)$$

for h small, where \approx denotes ‘‘approximately equal to’’. We also note that $B(h) - B(0)$ has the same distribution as $h^{1/2}N(0, I)$, so the third term (and all the higher order terms) on the right-hand side of (2.6) should be of order $h^{3/2}$. In view of (2.6), (2.7), and (2.8), we arrive at

$$\begin{aligned} & u^*(X(T \wedge h)) - u^*(X(0)) \\ &\approx \sum_{i=1}^d \frac{\partial u^*}{\partial x_i}(X(0)) \left(\mu_i(X(0))h + \sum_{k=1}^m \sigma_{ik}(X(0))(B_k(h) - B_k(0)) \right) \\ &+ \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 u^*}{\partial x_i \partial x_j}(X(0)) \sum_{k,l=1}^m \sigma_{ik}(X(0)) \sigma_{jl}(X(0))(B_k(h) - B_k(0)) \\ &\cdot \sigma_{jl}(X(0))(B_l(h) - B_l(0)) + o(h) \end{aligned}$$

and consequently,

$$\begin{aligned} & \exp\left(-\int_0^{T \wedge h} \alpha(X(s)) ds\right) u^*(X(T \wedge h)) \\ &\approx u^*(X(0))(1 - \alpha(X(0))h) \\ &+ \sum_{i=1}^d \frac{\partial u^*}{\partial x_i}(X(0)) \left(\mu_i(X(0))h + \sum_{k=1}^m \sigma_{ik}(X(0))(B_k(h) - B_k(0)) \right) \\ &+ \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 u^*}{\partial x_i \partial x_j}(X(0)) \sum_{k,l=1}^m \sigma_{ik}(X(0)) \sigma_{jl}(X(0))(B_k(h) - B_k(0)) \\ &\cdot \sigma_{jl}(X(0))(B_l(h) - B_l(0)) + o(h) \end{aligned}$$

as $h \downarrow 0$. So, this suggests that

$$\begin{aligned} & E_x u^*(X(T \wedge h)) \exp\left(-\int_0^{T \wedge h} \alpha(X(s)) ds\right) \\ &= u^*(x)(1 - \alpha(x)h) + \sum_{i=1}^d \frac{\partial u^*}{\partial x_i}(x) \mu_i(x)h \\ &+ \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 u^*}{\partial x_i \partial x_j}(x) \sum_{k=1}^m \sigma_{ik}(x) \sigma_{jk}(x)h + o(h). \end{aligned} \quad (2.9)$$

Plugging (2.9) and (2.6) into (2.1), subtracting $u^*(x)$ from both sides, dividing by h , and sending $h \downarrow 0$, we conclude

that u^* should satisfy

$$-g(x) = \sum_{i=1}^d \mu_i(x) \frac{\partial u^*}{\partial x_i}(x) + \frac{1}{2} b_{ij}(x) \frac{\partial^2 u^*}{\partial x_i \partial x_j}(x) - \alpha(x) u^*(x) \tag{2.10}$$

for $x \in A^c$ (where $b(x) \triangleq \sigma(x)\sigma(x)^T$), subject to $u^*(x) = f(x)$ for $x \in A$.

Of course, the above derivation is heuristic and non-rigorous. To make this connection between the PDE (2.10) and SDE (1.1) rigorous, it turns out that we must adopt a specific definition for the stochastic integral (1.3), namely the Itô definition. The Itô definition of (1.3) defines the stochastic integral as the limit (in mean-square) of

$$\sum_{i=0}^{n-1} \sigma \left(X \left(\frac{it}{n} \right) \right) \left(B \left(\frac{(i+1)t}{n} \right) - B \left(\frac{it}{n} \right) \right)$$

as $n \rightarrow \infty$. Use of other definitions will typically lead to a different PDE for the above expectation.

The connection between PDE's and diffusions is of fundamental mathematical importance, and offers the opportunity to compute certain probabilities and expectations by numerically solving the appropriate PDE. Since numerical computation of the solution to a PDE involves a time-step integration over the state space of X , it is clear that such integration schemes become inefficient when the dimension d is of moderate size or larger. In such settings, (as often occur in finance), Monte Carlo schemes become computationally attractive.

3 MONTE CARLO SIMULATION OF DIFFUSIONS

The process X evolves in continuous time, and evolves randomly within any time interval of positive length. As a consequence, we can not simulate X exactly as a continuous time process (unlike Markov jump process, where one can take advantage of the piecewise-constant sample paths by simulating the state only at the jump epochs). Hence, the typical goal in the diffusion setting is to simulate the process X only at some discrete sequence of times, usually a lattice of the form $0, h, 2h, \dots$

Unfortunately, the state of the diffusion X can be exactly generated at $0, h, 2h, \dots$ only for a very special subclass of diffusions (e.g. Brownian motion, the Ornstein-Uhlenbeck process, etc.) Rather, the typical approach used to simulate X is to generate a discrete-time approximation $X_h = (X_h(kh) : k \geq 0)$, just as numerical solution of deterministic differential equations usually evolves computing a discrete-time numerical approximation to the continuous time solution. The simplest such approximation is the Euler

approximation defined by the recursion

$$\begin{aligned} X_h((k+1)h) - X_h(kh) \\ = \mu(X_h(kh))h + \sigma(X_h(kh))(B((k+1)h) - B(kh)). \end{aligned}$$

We can then define X_h between the lattice points $0, h, 2h, \dots$ via a linear interpolation.

Given a functional φ defined on paths of X , computing $\alpha = E\varphi(X(s) : 0 \leq s \leq t)$ involves generating X_h up to time t , computing $\varphi(X_h(s) : 0 \leq s \leq t)$, simulating n independent and identically distributed (iid) copies of $\varphi(X_h(s) : 0 \leq s \leq t)$, and averaging the n copies to produce the final estimator

$$\alpha_n(h) = \frac{1}{n} \sum_{i=1}^n \varphi(X_h^i(s) : 0 \leq s \leq t).$$

Because we are simulating an approximation to X rather than X itself, $\alpha_n(h)$ includes a systematic error (i.e. bias), in addition to the sampling error that is typical of Monte Carlo schemes. The bias $E\alpha_n(h) - \alpha$ typically is of the order h^p for some $p > 0$; the parameter p is called the order of the method/estimator. Given that the computational complexity of generating $\alpha_n(h)$ is of order n/h , it comes as no surprise that p has a dramatic effect on the optimal rate of convergence (when optimized over combinations of n and h subject to the constraint $n/h = c$); see [Duffie and Glynn \(1995\)](#).

The tutorial will largely be focused on a discussion of the interaction between the choice of the discretization scheme (leading to a specific approximation X_h), the functional φ , and the order p . In particular, when φ takes the form $\varphi(X(s) : 0 \leq s \leq t) = k(X(t))$ (for k smooth), a discretization scheme is said to have *weak error* of order p if

$$Ek(X_h(t)) = Ek(X(t)) + O(h^p)$$

as $h \downarrow 0$ (where $O(r(h))$ is a function for which $O(r(h))/r(h)$ remains bounded as $h \downarrow 0$). On the other hand, if φ is Lipschitz in the sense that

$$|\varphi(x(s) : 0 \leq s \leq t) - \varphi(y(s) : 0 \leq s \leq t)| \leq c \sup_{0 \leq s \leq t} |x(s) - y(s)|$$

(for some constant $c < \infty$), then the discretization X_h is said to have *strong error* of order p if

$$E\varphi(X_h(s) : 0 \leq s \leq t) = E\varphi(X(s) : 0 \leq s \leq t) + O(h^p)$$

as $h \downarrow 0$. Finally, the scheme X_h is said to have *microscopic error* of order p if for any bounded function φ ,

$$E\varphi(X_h(kh) : 0 \leq kh \leq t) = E\varphi(X(kh) : 0 \leq kh \leq t) + O(h^p)$$

as $h \downarrow 0$. In other words, the class of φ 's for which a discretization scheme offers a given order p becomes larger

(and more demanding) as one moves from a consideration of weak error to strong error to microscopic error. Our goal in this tutorial is to discuss the interaction between the choice of discretization, the order of convergence, and the choice of functional φ ; the classical reference for discussion of weak and strong error is [Kloeden and Platen \(1992\)](#). See [Perez \(2004\)](#) for an account of microscopic error.

As one might expect, the presence of boundaries for the diffusion complicates the discussion. Specifically, a scheme having error of order p when implemented on a diffusion without boundaries may have an error of order $q < p$ when implemented on a diffusion with boundaries. The tutorial will therefore also provide some discussion of this fundamental issue; see [Lepingle \(1995\)](#) for details in the reflected case and [Gobet \(2000\)](#) for details in the absorbing setting.

REFERENCES

- Duffie, D., and P. W. Glynn. 1995. Efficient Monte Carlo simulation of security prices. *Annals of Applied Probability* 5:897–905.
- Gobet, E. 2000. Weak convergence of killed diffusion using Euler schemes. *Stochastic Processes and their Applications* 87:167–197.
- Kloeden, P. E., and E. Platen. 1992. *Numerical Solutions of Stochastic Differential Equations*. New York: Springer-Verlag.
- Lepingle, D. 1995. Euler scheme for reflected stochastic differential equations. *Mathematics and Computers in Simulation* 38:119–126.
- Perez, J. A. 2004. *Convergence of numerical schemes in the total variation sense*. Ph.D. thesis, Courant Mathematics and Computing Laboratory, New York University.

AUTHOR BIOGRAPHY

PETER W. GLYNN received his Ph.D. in Operations Research from Stanford University in 1982. He then joined the faculty of the University of Wisconsin at Madison, where he held a joint appointment between the Industrial Engineering Department and Mathematics Research Center, and courtesy appointments in Computer Science and Mathematics. In 1987, he returned to Stanford, where he is now the Thomas Ford Professor of Engineering in the Department of Management Science and Engineering. He also has a courtesy appointment in the Department of Electrical Engineering and serves as Director of the Stanford Institute of Computational and Mathematical Engineering. He is a member of INFORMS and a fellow of the Institute of Mathematical Statistics and his research interests in computational probability, simulation, queueing theory, statistical inference for stochastic processes, and stochastic modeling.