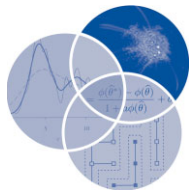
Advanced Review



Computational solution of stochastic differential equations

Timothy Sauer*

Stochastic differential equations (SDEs) provide accessible mathematical models that combine deterministic and probabilistic components of dynamic behavior. This article is an overview of numerical solution methods for SDEs. The solutions are stochastic processes that represent diffusive dynamics, a common modeling assumption in many application areas. We include a description of fundamental numerical methods and the concepts of strong and weak convergence and order for SDE solvers. In addition, we briefly discuss the extension of SDE solvers to coupled systems driven by correlated noise.

How to cite this article:
WIREs Comput Stat 2013, 5:362–371. doi: 10.1002/wics.1272

Keywords: stochastic differential equations; computational methods; diffusion problems

INTRODUCTION

Stochastic differential equations (SDEs) have become standard models for diffusive processes in the physical and biological sciences as well as economics and finance. Diffusion processes represent heat transfer and the movement and mixing of molecules in physics and chemistry, and transport of substances across membranes in cellular biology. In modern finance, the Black–Scholes formula for options pricing and other fundamental asset price models are based on SDEs where the diffusion coefficient represents price volatility.

Computational methods to solve SDEs are roughly analogous to solvers for ordinary differential equations, adapted for the probabilistic context. We include a self-contained brief introduction to stochastic calculus, and then survey the development of SDE solvers, beginning with the analogue of the Euler method. The important concepts of order and strong and weak convergence are explained. A final section shows how the methods can be generalized to multidimensional equations.

SOLUTIONS OF SDES

Under appropriate conditions, ordinary differential equations have a unique solution for each initial condition. SDEs, on the other hand, have solutions that are continuous-time stochastic processes. Methods for the computational solution of SDEs are based on techniques for ordinary differential equations, but adapted to account for stochastic dynamics.

Some fundamental concepts from stochastic calculus are needed to describe the numerical methods. A set of random variables X_t indexed by real numbers $t \geq 0$ is called a *continuous-time stochastic process*. Each instance, or *realization* of the stochastic process is a choice from the random variable X_t for each t, and is therefore a function of t.

Any (deterministic) function $f(t)$ can be trivially considered as a stochastic process, with variance $V(f(t)) = 0$. An archetypal example that is ubiquitous in models from physics, chemistry, and finance is the *Wiener process* W_t, a continuous-time stochastic process with the following three properties: (1) For each t, the random variable W_t is normally distributed with mean 0 and variance t. (2) For each $t_1 < t_2$, the normal random variable $W_{t_2} - W_{t_1}$ is independent of the random variable W_{t_1}, and in fact independent of

*Correspondence to: tsauer@gmu.edu

Department of Mathematics, George Mason University, Fairfax, VA, 22030, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

all $W_t, 0 \le t \le t_1$. (3) The Wiener process W_t can be represented by continuous paths.

The Wiener process is a formal version of random behavior first characterized by the botanist Robert Brown[1] in 1827, commonly called Brownian motion. Brownian motion is crucial in the modeling of stochastic processes since it represents the integral of idealized noise that is independent of frequency, called white noise. Often, the Wiener process is called upon to represent random, external influences on an otherwise deterministic system, or more generally, dynamics that for a variety of reasons cannot be deterministically modeled.

A typical *diffusion process* is modeled as a differential equation involving deterministic, or *drift* terms, and stochastic, or *diffusion* terms, the latter represented by a Wiener process, as in the equation

$$\mathrm{d}X = a\,(t, X)\,\mathrm{d}t + b\,(t, X)\,\mathrm{d}W_t. \tag{1}$$

SDEs are given in differential form, unlike the derivative form of ODEs. That is because many interesting stochastic processes, like Brownian motion, are continuous but not differentiable. Therefore, the meaning of the SDE (1) is, by definition, the integral equation

$$X\,(t) = X\,(0) + \int_0^t a\,(s, y)\,\mathrm{d}s + \int_0^t b\,(s, y)\,\mathrm{d}W_s,$$

where the meaning of the last integral, called an Ito integral, will be defined next.

Let $c = t_0 < t_1 < \cdots < t_{n-1} < t_n = d$ be a grid of points on the interval $[c,d]$. The Riemann integral is defined as a limit

$$\int_c^d f\,(x)\ \ \mathrm{d}x = \lim_{\Delta t \to 0} \sum_{i=1}^{n} f\left(t_{i'}\right) \Delta t_i,$$

where $\Delta t_i = t_i - t_{i-1}$ and $t_{i-1} \le t_{i'} \le t_i$. Similarly, the *Ito integral*[2] is the limit

$$\int_c^d f\,(t)\ \ \mathrm{d}W_t = \lim_{\Delta t \to 0} \sum_{i=1}^{n} f\,(t_{i-1})\,\Delta W_i$$

where $\Delta W_i = W_{t_i} - W_{t_{i-1}}$, a step of Brownian motion across the interval. The difference is that while the $t_{i'}$ in the Riemann integral may be chosen at any point in the interval (t_{i-1}, t_i), the corresponding point for the Ito integral is required to be the left endpoint of that interval.

Because f and W_t are random variables, so is the Ito integral $I = \int_c^d f\,(t)\ \ \mathrm{d}W_t$. The *differential* $\mathrm{d}I$ is a notational convenience; thus

$$I = \int_c^d f\,\mathrm{d}W_t$$

is expressed in differential form as $\mathrm{d}I = f\,\mathrm{d}W_t$. The differential $\mathrm{d}W_t$ of Brownian motion W_t is called *white noise*. A typical solution is a combination of drift and the diffusion of Brownian motion.

To solve SDEs analytically, we need to introduce the chain rule for stochastic differentials, called the *Ito formula*.[2] Let X be defined as in (1) and let $Y = f(t,X)$. Then

$$\begin{aligned}\mathrm{d}Y = &\frac{\partial f}{\partial t}\,(t, X)\,\mathrm{d}t + \frac{\partial f}{\partial x}\,(t, X)\,\mathrm{d}X \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\,(t, X)\,\mathrm{d}X\,\mathrm{d}X\end{aligned} \tag{2}$$

where the $\mathrm{d}X\ \mathrm{d}X$ term is interpreted from the identities

$$\begin{gathered}\mathrm{d}t\ \ \mathrm{d}t = \mathrm{d}t\ \ \mathrm{d}W_t = \mathrm{d}W_t\ \ \mathrm{d}t = 0 \\ \mathrm{d}W_t\ \ \mathrm{d}W_t = \mathrm{d}t.\end{gathered} \tag{3}$$

For example, to show that $Y = W_t^2$ is a solution to the SDE $\mathrm{d}Y = 1\,\mathrm{d}t + 2W_t\,\mathrm{d}W_t$, apply Ito's formula with $f(t,x) = x^2$ and $X = W_t$. Then

$$\begin{aligned}\mathrm{d}Y &= \frac{\partial f}{\partial t}\,(t, X)\,\mathrm{d}t + \frac{\partial f}{\partial x}(t, X)\,\mathrm{d}X \\ &\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X)\,\mathrm{d}X\,\mathrm{d}X \\ &= 0 + 2W_t\,\mathrm{d}W_t + \frac{1}{2}2\,\mathrm{d}W_t\,\mathrm{d}W_t \\ &= 1\,\mathrm{d}t + 2W_t\,\mathrm{d}W_t.\end{aligned}$$

The Ito formula is the stochastic analogue of the chain rule of differential calculus. Although it is expressed in differential form for ease of understanding, its meaning is precisely the equality of the Ito integral of both sides of the equation. It is proved under rather general hypotheses by referring the equation back to the definition of Ito integral. More complete details on Ito integrals and stochastic calculus can be found in a number of texts, including Refs 3–7.

For a second example, consider the stochastic differential equation

$$\begin{cases}\mathrm{d}X = \mu X\,\mathrm{d}t + \sigma X\,\mathrm{d}W_t \\ X\,(0) = X_0\end{cases} \tag{4}$$

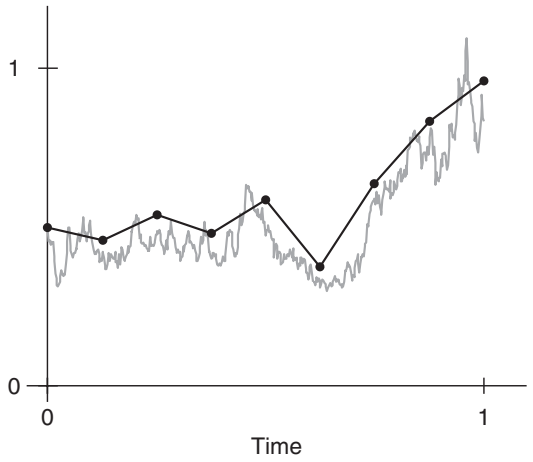


FIGURE 1 | Solution to the Black–Scholes stochastic differential equation (4). The exact solution (5) is plotted as a gray curve. The Euler–Maruyama approximation with time step $\Delta t = 1/8$ is plotted as a dark curve. The drift and diffusion parameters are set to $\mu = 0.2$ and $\sigma = 1$, respectively.

with constants μ and σ. The solution of this SDE is geometric Brownian motion

$$X(t) = X_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t}. \tag{5}$$

To check this, write $X = f(t,Y) = X_0 e^Y$, where $Y = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t$. By the Ito formula,

$$dX = X_0 e^Y \; dY + \frac{1}{2} e^Y \; dY \; dY$$

where $dY = \left(\mu - \frac{1}{2}\sigma^2\right) \; dt + \sigma \; dW_t$. Using the differential identities from the Ito formula,

$$dY \; dY = \sigma^2 \; dt,$$

and therefore

$$\begin{aligned} dX &= X_0 e^Y \left(r - \frac{1}{2}\sigma^2\right) dt + X_0 e^Y \sigma \, dW_t + \frac{1}{2}\sigma^2 e^Y dt \\ &= X_0 e^Y \mu \; dt + X_0 e^Y \sigma \; dW_t \\ &= \mu X \; dt + \sigma X \; dW_t \end{aligned}$$

as claimed.

Figure 1 shows a realization of geometric Brownian motion with constant drift coefficient μ and diffusion coefficient σ. Similar to the case of ordinary differential equations, relatively few SDEs have closed-form solutions. It is often necessary to use numerical approximation techniques, the central focus of this article.

COMPUTATIONAL METHODS FOR SDES

The simplest effective computational method for the approximation of ordinary differential equations is the Euler method; see Ref 8. The Euler–Maruyama method[9] is the analogue of the Euler method for ordinary differential equations. To develop an approximate solution on the interval $[c,d]$, assign a grid of points

$$c = t_0 < t_1 < t_2 < \cdots < t_n = d.$$

Approximate X values

$$x_0, \quad x_1, \quad x_2, \quad \ldots, \quad x_n$$

will be determined at the respective t points. Given the SDE initial value problem

$$\begin{cases} dX(t) = a(t,X)dt + b(t,X)dW_t \\ X(c) = X_c \end{cases} \tag{6}$$

the approximate solution is computed as follows:

Euler–Maruyama Method

$$\begin{aligned} x_0 &= X_0 \\ x_{i+1} &= x_i + a(t_i, x_i)\,\Delta t_{i+1} + b(t_i, x_i)\,\Delta W_{i+1} \end{aligned} \tag{7}$$

where

$$\Delta t_{i+1} = t_{i+1} - t_i$$

$$\Delta W_{i+1} = W(t_{i+1}) - W(t_i). \tag{8}$$

The Brownian motion is modeled by the increments ΔW_i, which are determined from a normal random number generator. Define $N(0,1)$ to be the standard random variable that is normally distributed with mean 0 and standard deviation 1. The random increment ΔW_i is computed as

$$\Delta W_i = z_i \sqrt{\Delta t_i} \tag{9}$$

where z_i is chosen from $N(0,1)$.

This is a clear difference from the Euler method in the deterministic ordinary differential equation case. Each set of $\{w_0, \ldots, w_n\}$ produced by the Euler–Maruyama method is an approximate realization of the solution stochastic process $X(t)$ which depends on the random numbers z_i that were chosen. Since W_t is a stochastic process,

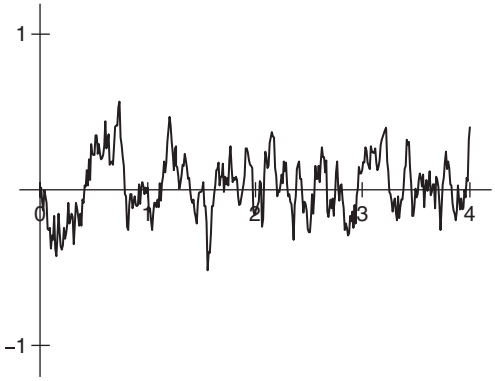


FIGURE 2 | Solution to Langevin equation (11). The path is the solution approximation for parameters $\mu = 10$, $\sigma = 1$, computed by the Euler–Maruyama method with stepsize $\Delta t_i = 0.01$ for all i.

each realization will be different and so will the approximations.

As a first example, the Euler–Maruyama method is applied to the SDE (4). The Euler–Maruyama equations (7) have the form

$$x_0 = X_0 \qquad (10)$$
$$x_{i+1} = x_i + \mu x_i \Delta t_i + \sigma x_i \Delta W_i.$$

We will use the drift coefficient $\mu = 0.2$ and diffusion coefficient $\sigma = 1$. An exact realization, generated from the solution (5), along with the corresponding Euler–Maruyama approximation, are shown in Figure 1. By corresponding, we mean that the approximation used the same Brownian motion realization as the true solution. Note the close agreement between the solution and the approximating points.

As another example, consider the *Langevin* equation[10]

$$\mathrm{d}X(t) = -\mu X(t)\ \mathrm{d}t + \sigma\ \mathrm{d}W_t \qquad (11)$$

where μ and σ are positive constants. In this case, it is not possible to analytically derive the solution in terms of elementary processes. The solution of the Langevin equation is a stochastic process called the *Ornstein-Uhlenbeck* process.[11] Figure 2 shows one realization of the approximate solution. It was generated from an Euler–Maruyama approximation, using the steps

$$x_0 = X_0 \qquad (12)$$
$$x_{i+1} = x_i - \mu x_i \Delta t_i + \sigma \Delta W_i$$

for $i = 1, \ldots, n$. This SDE is used to model systems that tend to revert to a particular state, in this case the state $X = 0$, in the presence of a noisy background.

STRONG CONVERGENCE OF SDE SOLVERS

The definition of convergence for SDE approximation methods is similar to convergence for ordinary differential equation solvers, aside from the differences caused by the fact that a solution to an SDE is a stochastic process, and each computed trajectory is only one realization of that process. Each approximate solution path $w(t)$, gives a random value at T, so that $w(T)$ is a random variable as well. The difference between the values at time T, $\mathrm{e}(T) = X(T) - x(T)$, is therefore a random variable.

A discrete-time approximation is said to *converge strongly* to the solution $X(t)$ at time T if

$$\lim_{\Delta t \to 0} E\{|X(T) - x_{\Delta t}(T)|\} = 0$$

where $x_{\Delta t}$ is the approximate solution computed with constant stepsize Δt, and E denotes expected value. For strongly convergent approximations, we further quantify the rate of convergence by the concept of order. An SDE solver *converges strongly with order* m if the expected value of the error is of mth order in the stepsize, i.e., if for any time T,

$$E\{|X(T) - x_{\Delta t}(T)|\} = O\left((\Delta t)^m\right)$$

for sufficiently small stepsize Δt. This definition generalizes the standard convergence criterion for ordinary differential equations, reducing to the usual definition when the stochastic part of the equation goes to zero.

Although the Euler method for ordinary differential equations has order 1, the strong order for the Euler–Maruyama method for SDEs is 1/2. This fact was proved [8] under appropriate conditions on the functions a and b in standard form (6). In particular, the drift term $a(t,x)$ must be globally Lipschitz continuous or grow at most linearly in the variable x.

One way to produce higher-order solvers for ordinary differential equations is to make use of higher-order terms in the Taylor series of the solution. We use a similar approach to build a strong order 1 method for SDEs. The analogue of the Taylor series for SDEs is the Ito-Taylor expansion. See Ref 12 for details. The Milstein method,[13–15,16] introduced in the mid-1970s, includes one more term from this expansion. Consider the SDE

$$\begin{cases} \mathrm{d}X(t) = a(X,t)\mathrm{d}t + b(X,t)\mathrm{d}W_t \\ X(0) = X_0. \end{cases} \qquad (13)$$

TABLE 1 | Average Error Versus Step Size for the Euler–Maruyama and Milstein Approximations of Eq. (15)

Δt	Euler–Maruyama	Milstein
2^{-1}	0.169369	0.063864
2^{-2}	0.136665	0.035890
2^{-3}	0.086185	0.017960
2^{-4}	0.060615	0.008360
2^{-5}	0.048823	0.004158
2^{-6}	0.035690	0.002058
2^{-7}	0.024277	0.000981
2^{-8}	0.016399	0.000471
2^{-9}	0.011897	0.000242
2^{-10}	0.007913	0.000122

Milstein Method

$$x_0 = X_0$$

$$x_{i+1} = x_i + a\left(x_i, t_i\right)\Delta t_i + b(x_i, t_i)\Delta W_i + \frac{1}{2}b\left(x_i, t_i\right)\frac{\partial b}{\partial x}(x_i, t_i)(\Delta W_i^2 - \Delta t_i). \quad (14)$$

The Milstein method has order 1, meaning that it will converge to the correct stochastic solution process more quickly than Euler–Maruyama as the stepsize Δt_i goes to 0. The Milstein method is identical to the Euler–Maruyama method if there is no X term in the diffusion part $b(X,t)$ of the equation.

To compare the Euler–Maruyama and Milstein methods, we apply them to the Black–Scholes SDE

$$\mathrm{d}X = \mu X\ \mathrm{d}t + \sigma X\ \mathrm{d}W_t. \quad (15)$$

We discussed the Euler–Maruyama approximation above. The Milstein method becomes

$$x_0 = X_0 \quad (16)$$

$$x_{i+1} = x_i + \mu x_i \Delta t_i + \sigma x_i \Delta W_i + \frac{1}{2}\sigma\left(\Delta W_i^2 - \Delta t_i\right).$$

Applying the Euler–Maruyama and the Milstein methods with decreasing stepsizes Δt results in successively improved approximations as shown in Table 1.

The two columns represent the average of the error $\{x(T) - X(T)\}$ at $T = 8$ over 100 realizations. The orders 1/2 for Euler–Maruyama and 1 for Milstein are clearly visible in the table. Cutting the stepsize by a factor of 4 is required to reduce the error by a factor of 2 with the Euler–Maruyama method. For the Milstein method, cutting the stepsize by a factor of 2 achieves the same result.

The Milstein method is an Ito-Taylor method, meaning that it is derived from a truncation of the stochastic Ito-Taylor expansion of the solution. This is in many cases a disadvantage, since the partial derivative appears in the approximation method, and must be provided explicitly by the user. This is analogous to higher-order Taylor methods for solving ordinary differential equations, which are seldom used in practice for that reason. To counter this problem, Runge–Kutta methods were developed for ODEs, which trade these extra partial derivatives in the Taylor expansion for extra function evaluations from the readily-available differential equation.

In the SDE context, the same trade can be made with the Milstein method, resulting in a strong order 1 method that requires evaluation of $b(X)$ at two places on each step. A heuristic derivation can be carried out by making the replacement

$$b_x\left(x_i\right) \approx \frac{b\left(x_i + b\left(x_i\right)\sqrt{\Delta t_i}\right) - b\left(x_i\right)}{b\left(x_i\right)\sqrt{\Delta t_i}}$$

in the Milstein formula (14), which leads to the Runge–Kutta method.

Strong Order 1.0 Runge–Kutta Method

$$x_0 = X_0$$

$$x_{i+1} = x_i + a\left(x_i\right)\Delta t_i + b(x_i)\Delta W_i + \frac{1}{2}[b(x_i + b\left(x_i\right)\sqrt{\Delta t_i}) - b(x_i)](\Delta W_i^2 - \Delta t_i)/\sqrt{\Delta t_i}.$$

A proper derivation can be found in Rumelin.[17] The orders of the methods introduced here for SDEs,[18] 1/2 for Euler–Maruyama and 1 for Milstein and the Runge–Kutta counterpart, would be considered low by ODE standards. Higher-order methods can be developed for SDEs, but become much more complicated as the order grows. As an example, consider the strong order 1.5 scheme for the SDE (13) proposed by Platen and Wagner.[19]

Strong Order 1.5 Taylor Method

$$x_0 = X_0$$

$$x_{i+1} = x_i + a\Delta t_i + b\Delta W_i + \frac{1}{2}bb_x\left(\Delta W_i^2 - \Delta t_i\right)$$

$$+ a_y\sigma\,\Delta Z_i + \frac{1}{2}\left(aa_x + \frac{1}{2}b^2a_{xx}\right)\Delta t_i^2 + \left(ab_x + \frac{1}{2}b^2b_{xx}\right)(\Delta W_i\Delta t_i - \Delta Z_i) + \frac{1}{2}b\left(bb_{xx} + b_x^2\right)(\frac{1}{3}\Delta W_i^2 - \Delta t_i)\Delta W_i \qquad (17)$$

where partial derivatives are denoted by subscripts, and where the additional random variable ΔZ_i is normally distributed with mean 0, variance $E\left(\Delta Z_i^2\right) = \frac{1}{3}\Delta t_i^3$ and correlated with ΔW_i with covariance $E\left(\Delta Z_i\Delta W_i\right) = \frac{1}{2}\Delta t_i^2$. Note that ΔZ_i can be generated as

$$\Delta Z_i = \frac{1}{2}\Delta t_i\left(\Delta W_i + \Delta V_i/\sqrt{3}\right)$$

where ΔV_i is chosen independently from $\sqrt{\Delta t_i}N(0,1)$.

Whether higher-order methods are needed in a given application depends on how the resulting approximate solutions are to be used. In the ordinary differential equation case, the usual assumption is that the initial condition and the equation are known with accuracy. Then it makes sense to calculate the solution as closely as possible to the same accuracy, and higher-order methods are called for. In the context of SDEs, in particular if the initial conditions are chosen from a probability distribution, the advantages of higher-order solvers are often less compelling, and if they come with added computational expense, may not be warranted.

WEAK CONVERGENCE OF SDE SOLVERS

Strong convergence allows approximations to be computed accurately on an individual realization basis. For some usages, such detailed pathwise information is required. In other applications such as Monte Carlo estimates,[20–23] the goal is to learn the probability distribution of the solution $X(T)$; single realizations may not be of interest.

Weak solvers are devised to fill this need. They are often simpler than corresponding strong methods, since their goal is to replicate the probability distribution only. In analogy with strong convergence, we offer the following definition.

A discrete-time approximation $x_{\Delta t}$ with stepsize Δt is said to *converge weakly* to the solution $X(T)$ if

$$\lim_{\Delta t\to 0} E\left\{f\left(x_{\Delta t}(T)\right)\right\} = E\left\{f\left(X(T)\right)\right\}$$

for all polynomials $f(x)$. According to this definition, all moments converge as $\Delta t \to 0$. If the stochastic part of the equation is 0 and the initial value is deterministic, the definition agrees with the strong convergence definition, and the usual ordinary differential equation definition.

Weakly convergent methods can also be assigned an order of convergence. We say that a solver *converges weakly with order m* if the error in the moments is of *m*th order in the stepsize, or

$$\left|E\left\{f\left(X(T)\right)\right\} - E\left\{f\left(x_{\Delta t}(T)\right)\right\}\right| = O((\Delta t)^m)$$

for sufficiently small stepsize Δt.

In general, the rates of weak and strong convergence do not agree. Unlike the case in ordinary differential equations, where the Euler method has order 1, the Euler–Maruyama method for SDEs is guaranteed to converge strongly with order $m = 1/2$, and converge weakly with order 1.

Higher-order weak methods can be much simpler than corresponding strong methods, and are available in several different forms. The most direct approach is to exploit the Ito-Taylor expansion referred above. An example SDE solver that converges weakly with order 2 is as follows.

Weak Order 2 Taylor Method

$$x_0 = X_0$$
$$x_{i+1} = x_i + a\Delta t_i + b\Delta W_i + \frac{1}{2}bb_x\left(\Delta W_i^2 - \Delta t_i\right) + a_xb\Delta Z_i + \frac{1}{2}\left(aa_x + \frac{1}{2}a_{xx}b^2\right)\Delta t^2 + \left(ab_x + \frac{1}{2}b_{xx}b^2\right)(\Delta W_i\Delta t_i - \Delta Z_i) \qquad (18)$$

where ΔW_i is chosen from $\sqrt{\Delta t_i}N(0,1)$ and ΔZ_i is distributed as in the Strong Order 1.5 method.

A second approach is to mimic the idea of Runge–Kutta solvers for ordinary differential equations. These solvers replace the explicit higher derivatives in the Ito-Taylor solvers with extra function evaluations at interior points of the current solution interval. Platen[24,25] proposed the weak order 2 solver of Runge–Kutta type.

Weak Order 2 Runge–Kutta Method

$$x_0 = X_0$$
$$x_{i+1} = x_i + \frac{1}{2}[a(u) + a(x_i)]\Delta t_i$$

$$+\frac{1}{4}[b(u_+)+b(u_-)+2b(x_i)]\Delta W_i$$
$$+\frac{1}{4}[b(u_+)-b(u_-)](\Delta W_i^2-\Delta t)/\sqrt{\Delta t_i} \qquad (19)$$

where

$$u = x_i + a\Delta t_i + b\Delta W_i$$
$$u_+ = x_i + a\Delta t_i + b\sqrt{\Delta t_i}$$
$$u_- = x_i + a\Delta t_i - b\sqrt{\Delta t_i}.$$

Figure 3 compares the Euler–Maruyama method, which converges with order 1 in the weak sense, to the Weak Order 2 Runge–Kutta method. Note the difference between strong and weak convergence. In Figure 3 the mean error of the estimate of the expected value $E[X(T)]$ is plotted, since we are comparing weak convergence of the methods. The weak orders are clearly revealed by the log–log plot.

Several other higher-order weak solvers can be found in Kloeden and Platen.[12,26] Weak Taylor methods of any order can be constructed, as well as Runge–Kutta analogues that reduce or eliminate the derivative calculations. In addition, standard Richardson extrapolation techniques [49] can be used to bootstrap weak method approximations of a given order to the next order.

Weak solvers can be used in diffusion models where information is carried primarily by means or variances, as in many biological and molecular dynamics applications. They are also an appropriate choice for practical financial modeling uses, in particular when the goal is to investigate the probability distribution of an asset price or interest rate, or when stochastic sampling is used to price a complicated derivative. In such cases it is typical to be primarily interested in one of the statistical moments of a stochastically defined quantity, and weak methods may be simpler and still sufficient for the sampling purpose.

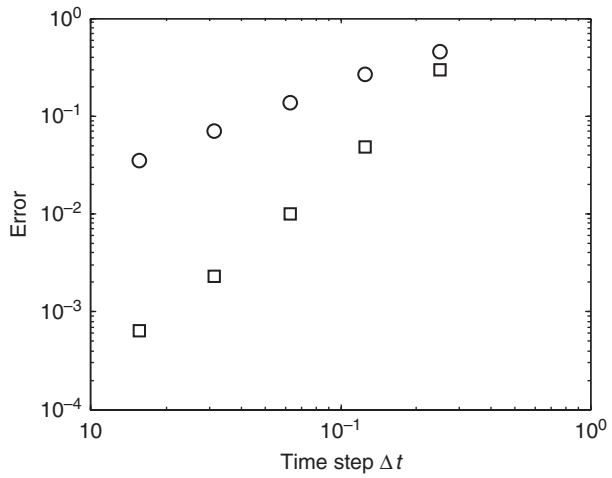


FIGURE 3 | The mean error of the estimation of $E(X(T))$ for SDE (15). The plot compares the Euler–Maruyama method (circles) which has weak order 1, and the weak order 2 Runge–Kutta type method (squares) given in (19). Parameters used were $X(0) = 10$, $T = 1$, $\mu = -3$, $\sigma = 0.2$.

Monte Carlo (MC) simulation[20,27,23] is used in many fields to compute quantities of interest in stochastic modeling scenarios. In simplest form, the quantity of interest is expressed as the expected value of a random variable, which is approximated as the average of the random variable over many random realizations. In many applications, the quantity of interest is a function of the solution value $X(T)$ of an SDE at time T, where $X(0)$ is known exactly, or chosen from a known distribution. For example, the values of exotic financial derivatives[28,29] are often computed in this way, where $X(t)$ represents the price of an underlying asset.

A recent key development in the research on MC approximation is the emergence of multilevel Monte Carlo (MLMC),[30,31] which is a way of accelerating convergence to the desired expected value. The method works as a type of variance reduction, where both a nearby quantity with lower variance and the difference between the two quantities are separately calculated, with reduced total effort relative to the original problem. A number of researchers[32–34] have reported useful applications of MLMC to improve MC approximation of quantities modeled by SDEs. The development of MLMC has stimulated the search for more efficient and more generally applicable strong and weak SDE solvers.[35]

Multidimensional SDEs

So far we have described solution methods for scalar SDEs. Extending the methods to computational approximation of coupled SDEs is straightforward, as long as the noise contributions are uncorrelated. We begin with such an example, and then discuss changes needed to handle correlated noise inputs.

A common asset model in contemporary finance is the Heston model[36]

$$dX_t = rX_t\ dt + \sqrt{V_t}X_t\ dW_t^1$$
$$dV_t = \kappa(\theta - V_t)\ dt + c\sqrt{V_t}\ dW_t^2. \qquad (20)$$

where X_t and V_t represent asset price and volatility, respectively. The Heston model is a generalization

of the Black–Scholes asset model[37] that allows volatility to change stochastically, yet yields an analytic form for the pricing of contingent assets like options. If we assume that W_t^1 and W_t^2 are independent Wiener processes, the Euler–Maruyama discrete version would be

$$\begin{aligned} x_0 &= X_0 \\ v_0 &= V_0 \\ x_{i+1} &= x_i + r x_i \ \Delta t_i + \sqrt{v_i} x_i \ \Delta W_i^1 \\ v_{i+1} &= v_i + \kappa(\theta - v_i) \ \Delta t_i + c\sqrt{v_i} \ \Delta W_i^2. \end{aligned} \tag{21}$$

with independent Brownian increments $\Delta W_i^1 = z_i^1 \sqrt{\Delta t_i}, \Delta W_i^2 = z_i^2 \sqrt{\Delta t_i}$, where z_i^1 and z_i^2 are independent draws from the standard normal distribution.

However, coupled SDEs often involve noise that is driven by a correlated multidimensional Wiener process. The solution of multidimensional SDEs is complicated in general; here we give only a glimpse of the issues involved.

For multifactor Wiener processes $\left(W_t^1, \ldots, W_t^k\right)$, the generalization of Ito's formula requires that (3) is replaced with

$$\begin{aligned} \mathrm{d}t \ \mathrm{d}t &= \mathrm{d}t \ \mathrm{d}W_t^i = \mathrm{d}W_t^i \ \mathrm{d}t = 0 \\ \mathrm{d}W_t^i \ \mathrm{d}W_t^j &= \rho_{ij} \ \mathrm{d}t \end{aligned} \tag{22}$$

where ρ_{ij} represents the statistical correlation between W_t^i and W_t^j. As usual, correlation ρ of two random variables X_1 and X_2 is defined as

$$\rho(X_1, X_2) = \frac{\mathrm{cov}(X_1, X_2)}{\sqrt{V(X_1)}\sqrt{V(X_2)}}.$$

Note that $\rho(X_1,X_1) = 1$, and X_1 and X_2 are uncorrelated if $\rho(X_1,X_2) = 0$.

To construct discretized correlated Wiener processes for use in SDE solvers, one begins with a desired correlation matrix

$$\mathbf{R} = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1k} \\ \vdots & & \vdots \\ \rho_{k1} & \cdots & \rho_{kk} \end{bmatrix}$$

for the Wiener processes $W^1, \ldots, W^k$. The matrix $\mathbf{R}$ is symmetric positive semi-definite (all eigenvalues are nonnegative) with units on the main diagonal. A straightforward way to create noise processes with a specified correlation is by taking a matrix square root of $\mathbf{R}$. One can use the Cholesky factorization $\mathbf{R} = CC^\tau$ for this purpose, or alternatively one can use the singular value decomposition (SVD) (see Ref 8 for description of both).

To create a noise process with correlation $\mathbf{R}$, begin with k independent, uncorrelated Wiener processes $Z_1, \ldots, Z_k$, satisfying $\mathrm{d}Z_i \mathrm{d}Z_i = \mathrm{d}t, \mathrm{d}Z_i \mathrm{d}Z_j = 0$ for $i \neq j$. Define the column vector $\mathbf{dW} = C\mathbf{dZ}$, and check that the covariance matrix, and therefore the correlation matrix, of $\mathbf{dW}$ is

$$\begin{aligned} \mathbf{dW}\mathbf{dW}^\tau &= C\mathbf{dZ}\left(C\mathbf{dZ}\right)^\tau \\ &= C\mathbf{dZ}\mathbf{dZ}^\tau C^\tau \\ &= CC^\tau \mathrm{d}t = \mathbf{R} \ \mathrm{d}t. \end{aligned}$$

For example, we could consider two molecular processes X_1 and X_2 that are driven by negatively correlated noise, say with correlation matrix

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

where $\rho = -0.8$. The matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$$

is the Cholesky square root of $\mathbf{R}$, so the noise can be generated as

$$\begin{aligned} \mathrm{d}W^1 &= \mathrm{d}Z^1 \\ \mathrm{d}W^2 &= \rho \ \mathrm{d}Z^1 + \sqrt{1-\rho^2} \ \mathrm{d}Z^2. \end{aligned} \tag{23}$$

CONCLUSION

Numerical methods for the solution of SDEs are essential for the analysis of random phenomena. Strong solvers are necessary when exploring characteristics of systems that depend on trajectory-level properties. Several approaches exist for strong solvers, in particular Taylor and Runge–Kutta methods, although these methods tend to increase greatly in complexity for orders greater than one.

In many applications, major emphasis is placed on the probability distribution of solutions, and in particular mean and variance of the distribution. In such cases, weak solvers may be sufficient, and have the advantage of comparatively less computational overhead.

In addition to the choice of SDE solver, methods of variance reduction exist that may increase

computational efficiency. MLMC methods offer a promising means of variance reduction and increases in efficiency in simulation of SDE trajectories. Moreover, the replacement of pseudorandom numbers with quasirandom numbers[8,22] from low-discrepancy sequences[38] is applicable as long as statistical independence along trajectories is maintained.

A more in-depth explanation of the mathematics behind the algorithms, including the Ito-Taylor expansions, can be found in Kloeden and Platen.[12] Another readable survey of SDE solvers which discusses implementation issues is due to Higham.[39] Approximation of solutions of SDEs depends on the existence of powerful normal random number generators, typically produced by a uniform random generator together with the Box-Muller method.[40] Computational techniques for the generation of random numbers is itself an interesting story; see Refs 21, 41–44 for general information on random number generation. We have emphasized evenly-spaced time stepping in our discussion, but there is a growing literature on adaptive time-step methods.[45–47] Progress on error and stability issues for SDE solvers[48–54] is an important and ongoing research area.

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation grant DMS-1216568.

REFERENCES

1. Brown R. A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philos Mag* 1828, 4:161–173.
2. Ito K. Stochastic integral. *Proc Imp Acad Tokyo* 1944, 20:518–524.
3. Gikhman I, Skorokhod A. *Stochastic Differential Equations*. Berlin: Springer; 1972.
4. Klebaner F. *Introduction to Stochastic Calculus with Applications*. London: Imperial College Press; 1998.
5. Oksendal B. *Stochastic Differential Equations: An Introduction With Applications*. 5th ed. Berlin, Heidelberg and New York: Springer; 1998.
6. Steele JM. *Stochastic Calculus and Financial Applications*. New York: Springer; 2001.
7. Mao X. *Stochastic Differential Equations and Applications*. 2nd ed. Chichester, UK: Horwood; 2007.
8. Sauer T. *Numerical Analysis*. 2nd ed. Boston: Pearson; 2012.
9. Maruyama G. Continuous Markov processes and stochastic equations. *Rend Circ Math Palermo* 1955, 4:48–90.
10. Langevin P. On the theory of Brownian motion. *C R Acad Sci (Paris)* 1908, 146:530–533.
11. Uhlenbeck GE, Ornstein L. On the theory of Brownian motion. *Phys Rev* 1930, 36:823–841.
12. Kloeden P, Platen E. *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer; 1992.
13. Milstein G. Approximate integration of stochastic differential equations. *Theory Probab Appl* 1975, 19:557–600.
14. Milstein G. *Numerical Integration of Stochastic Differential Equations*. Dordrecht: Kluwer; 1995.
15. Milstein G, Tretyakov M. *Stochastic Numerics for Mathematical Physics*. Berlin, Heidelberg and New York: Springer; 2004.
16. Milstein G, Tretyakov M. Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients. *SIAM J Num Anal* 2005, 43:1139–1154.
17. Rumelin W. Numerical treatment of stochastic differential equations. *SIAM J Numer Anal* 1982, 19:604–613.
18. Tocino A, Ardanuy R. Runge-Kutta methods for numerical solution of stochastic differential equations. *J Comput Appl Math* 2002, 138:219–241.
19. Platen E, Wagner W. On a Taylor formula for a class of Ito processes. *Probab Math Stat* 1982, 3:37–51.
20. Fishman GS. *Monte Carlo: Concepts, Algorithms, and Applications*. Berlin, Heidelberg and New York: Springer; 1996.
21. Gentle JE. *Random Number Generation and Monte Carlo Methods*. 2nd ed. Berlin, Heidelberg and New York: Springer; 2003.
22. Niederreiter H. *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia: SIAM Publications; 1992.
23. Rubinstein RY. *Simulation and the Monte Carlo Method*. New York: Wiley; 1981.
24. Platen E. Derivative-free numerical methods for stochastic differential equations. *Lect Notes Control Inform Sci* 1987, 96:187–193.

25. Platen E. An introduction to numerical methods for stochastic differential equations. *Acta Numer* 1999, 8:197–246.
26. Kloeden P, Platen E, Schurz H. *Numerical Solution of SDE Through Computer Experiments*. Berlin: Springer; 1994.
27. Glasserman P. *Monte Carlo Methods in Financial Engineering*. New York: Springer; 2004.
28. Sauer T. Numerical solution of stochastic differential equations in finance. In: Duan J-C, Hardle W, Gentle J, eds. *Handbook of Computational Finance*. Springer, Berlin-Heidelberg; 2012, 529–550.
29. Longstaff F, Schwartz E. Valuing American options by simulation: a simple least-squaares approach. *Rev Financ Stud*. 2001, 14:113–148.
30. Giles M. Multi-level Monte Carlo path simulation. *Oper Res* 2008, 56:607–617.
31. Heinrich S. Multilevel Monte Carlo methods. In: *Lecture Notes on Large Scale Scientific Computing*. Berlin: Springer; 2001, 58–67.
32. Higham DJ, Mao X, Roj M, Song Q, Yin G. Mean exit times and the multilevel Monte Carlo method. *J Uncertain Quant* 2013, 1:2–18.
33. Giles M, Waterhouse B. Multi-level quasi-Monte Carlo path simulation. *Radon Series Comp Appl Math* 2009, 8:165–181.
34. Giles M, Higham D, Mao X. Analysing mujlei-level Monte Carlo for options with non-global Lipschitz payoff. *Finance Stoch* 2009, 13:403–413.
35. Hutzenthaler M, Jentzen A, Kloeden P. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann Appl Probab* 2012, 22:1611–1641.
36. Heston S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev Financ Stud* 1993, 6:327–343.
37. Black F, Scholes M. The pricing of options and corporate liabilities. *J Polit Econ* 1973, 81:637–654.
38. Halton JH. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer Math* 1960, 2:84–90.
39. Higham DJ. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev* 2001, 43:525–546.
40. Box GEP, Muller M. A note on the generation of random normal deviates. *Ann Math Stat* 1958, 29:610–611.
41. Park S, Miller K. Random number generators: good ones are hard to find. *Commun ACM* 1988, 31:1192–1201.
42. Hellekalek P. Good random number generators are (not so) easy to find. *Math Comput Simul* 1998, 46:485–505.
43. Marsaglia G, Zaman A. A new class of random number generators. *Ann Appl Probab* 1991, 1:462–480.
44. Marsaglia G, Tsang WW. The ziggurat method for generating random variables. *J Stat Softw* 2000, 5:1–7.
45. Romisch W, Winkler R. Stepsize control for mean-square numerical methods for stochastic differential equations with small noise. *SIAM J Sci Comput* 2006, 28:604–625.
46. Lamba H, Mattingly JC, Stuart A. An adaptive Euler-Maruyama scheme for SDEs: convergence and stability. *IMA J Numer Anal* 2007, 27:479–506.
47. Hoffman N, Müller-Gronbach T, Ritter K. Step size control for the uniform approximation of systems of stochastic differential equations with additive noise. *Ann Appl Probab* 2000, 10:616–633.
48. Higham DJ, Kloeden P. Numerical methods for nonlinear stochastic differential equations with jumps. *Numerische Math* 2005, 101:101–119.
49. Higham DJ, Mao X, Stuart A. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J Numer Anal* 2002, 40:1041–1063.
50. Burrage K, Burrage PM, Mitsui T. Numerical solutions of stochastic differential equations—implementation and stability issues. *J Comput Appl Math* 2000, 125:171–182.
51. Burrage K, Burrage PM, Tian T. Numerical methods for strong solutions of stochastic differential equations: an overview. *Proc R Soc Lond A* 2004, 460:373–402.
52. Saito Y, Mitsui T. Stability analysis of numerical schemes for stochastic differential equations. *SIAM J Numer Anal* 1996, 33:2254–2267.
53. Talay D, Tubaro L. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch Anal Appl* 1990, 8:483–509.
54. Jentzen A, Kloeden P, Neuenkirch A. Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients. *Numerische Math* 2008, 112:41–64.

