International Mathematical Forum, 2, 2007, no. 6, 251 - 262

An Estimate of the Error for Strong Solutions of Stochastic Differential Equations

M. Namjoo and Ali R. Soheili

Department of Mathematics
University of Sistan and Baluchestan, Zahedan, Iran
namjoo@math.usb.ac.ir, soheili@math.usb.ac.ir

Abstract

Stochastic differential equations (SDEs) arise from modelling physical system by incorporating random elements in differential equation, such that Randomness can be included in the initial value for the problem or in function describing the physical system in order to model can be made more realistic. In this paper, we first give some techniques to obtain solution of SDEs. Then we use numerical simulations to estimate the error of an approximation by the absolute error criterion that is expectation of the absolute value of the difference between the Itô approximation and the exact solution SDE at a finite terminal time T. In continuation, we study the behavior of variation this estimate and confidence intervals of this error versus step size.

Keywords: Stochastic differential equations; Strong convergence; Itô method; Numerical simulation.

1 Introduction

Many physical systems are modeled by SDEs, where random effects being modeled by a Wiener process (see, for example, [7], [5], [4]) that is nowhere differentiable. The general form of SDEs, is given by

$$\begin{aligned} dY(t) &= g_0(Y(t))\,dt + \sum_{j=1}^{d} g_j(Y(t))\,dW_j(t)\ ,\ t \in [t_0, T]\ ,\ Y \in \mathbb{R}^m \qquad (1) \\ Y(t_0) &= Y_0, \end{aligned}$$

where $g_j(Y),\ j = 0, \ldots, d$, are m–vector–valued functions, and the $W_j(t),\ j = 1, \ldots, d$, are independent Wiener processes, and the solution $Y(t)$ is an m–vector process. Equation (1) can be written as a stochastic integral equation

$$Y(t) = Y_0 + \int_{t_0}^{t} g_0(Y(s))\,ds + \sum_{j=1}^{d} \int_{t_0}^{t} g_j(Y(s))\,dW_j(s)\ , \qquad (2)$$

where the d stochastic integrals in (2) by reason of infinite variation of the sample paths of a Wiener process cannot follows the usual rules of Riemann–Stieltjes calculus. They are defined as the limit (in the mean square sense), as $N \to \infty$, of the approximating sums

$$\sum_{i=1}^{N} g_j(Y(\xi_i))\ (W_j(t_i) - W_j(t_{i-1})),$$

where $\xi_i = \theta t_i + (1-\theta)t_{i-1}$, for a $\theta \in [0,1]$ and, for $\{t_0, \ldots, t_N\}$ be a partition of $[t_0, t]$, with $t_i = t_0 + \frac{i(t-t_0)}{N}$, $i = 0, \ldots, N$. The most common choices for the parameter θ are $\theta = 0$, and $\theta = \frac{1}{2}$, which gives Itô and Stratonovich integral, respectively. The Stratonovich interpretation satisfies the usual rules of calculus, while Itô integral forms a martingale, that provides some computational advantage (see [2]). While it is not always obvious whether a SDE, should be considered in Itô or Stratonovich form, which there is a relationship between them. Indeed, the solution of the Itô SDE (1) with $d = 1$, and the solution of the related Stratonovich SDE is given by

$$dY(t) = \bar{g}_0(Y(t))\ dt + g_1(Y(t))\ o dW(t)\ ,\ Y(t_0) = Y_0\ , t \in [t_0, T]\ ,\ Y \in \mathbb{R}^m$$

where

$$\bar{g}_0(Y) = g_0(Y) - \frac{1}{2}\,\frac{\partial g_1(Y)}{\partial Y}\ g_1(Y),$$

under different rules of calculus, have the same solution. There are different numerical methods to solve these kinds of differential equations (see, for example, [1], [3], [8]). A outline of this paper is as follows: In section 2 we briefly review some techniques to obtain analytical solution of SDEs. In section 3, we will concentrate on numerical simulations to the estimate of the absolute error numerical solution of SDE. In continuation the behavior this estimation and the confidence intervals of this error versus step size will be studied.

2 Analytic Methods For Solving SDEs

In this section, we concentrate on Itô formula with its applications. This Theorem is due to Itô that can be used for solving some SDEs. In order to simplify some notational details, for the rest of this paper it will be assumed that $d = m = 1$.

Theorem 2.1 (Itô's formula) *Let $Y(t)$ be an exact solution of (1) where $g_0(Y(t)) \in L^1(0,T)$ and $g_1(Y(t)) \in L^2(0,T)$. Suppose $F : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ have continuous partial derivatives $\frac{\partial F}{\partial t}$, $\frac{\partial F}{\partial y}$ and $\frac{\partial^2 F}{\partial y^2}$. Then*

$$\begin{aligned} dF(t, Y(t)) &= \left(\frac{\partial F(t,Y(t))}{\partial t} + \frac{\partial F(t,Y(t))}{\partial y}\ g_0(Y(t)) + \frac{1}{2}\ \frac{\partial^2 F(t,Y(t))}{\partial y^2}\ g_1^2(Y(t))\right) dt \\ &+ \frac{\partial F(t,Y(t))}{\partial y}\ g_1(Y(t))\ dW(t). \end{aligned}$$

Proof: See [6].
By Theorem 2.1 we can solve a particular class of SDEs. Suppose a SDE of the form

$$\begin{cases} dY(t) = g_0(Y(t))\ dt + g_1(Y(t))\ dW(t), \quad t \in [0,T] \\ Y(0) = Y_0, \end{cases} \tag{3}$$

be given and let h is a continuous function. Let us first solve SDE

$$\begin{cases} dX(t) = h(t)\ f(X(t))\ dt + h(t)\ dW(t), \quad t \in [0,T] \\ X(0) = X_0, \end{cases} \tag{4}$$

where f will be selected later, and try to find a function F such that $Y(t) := F(X(t))$, solves SDE (3). Assuming for the moment F and f are known, by using Itô's formula we have

$$dY(t) = (h(t)\ f(X(t))\ F'(X(t)) + \frac{1}{2}h^2(t)\ F''(X(t)))\ dt + h(t)\ F'(X(t))\ dW(t).$$

Thus $Y(t)$ solves SDE (3) provided

$$\begin{cases} h(t)\ f(X(t))\ F'(X(t)) + \frac{1}{2}h^2(t)\ F''(X(t)) = g_0(F(X(t))), \\ h(t)\ F'(X(t)) = g_1(F(X(t))), \\ F(X_0) = Y_0. \end{cases} \tag{5}$$

Example 2.1 *Consider SDE given by*

$$\begin{cases} dY(t) = g(t)\ Y(t)\ dW(t), \quad t \in [0,T] \\ Y(0) = Y_0, \end{cases} \tag{6}$$

where g is a continuous function (not a random variable), for this, $g_0 \equiv 0$ and $g_1(Y(t)) = g(t)\ Y(t)$. Let $h := g$ then by (5)

$$F(X(t)) = Y_0\ e^{X(t)-X_0}, \quad f(X(t)) = -\frac{1}{2}g(t),$$

and therefore (4) *gives*

$$X(t) = X_0 - \frac{1}{2}\int_0^t g^2(s)\ ds + \int_0^t g(s)\ dW(s).$$

This shows that

$$Y(t) = F(X(t)) = Y_0\ e^{-\frac{1}{2}\int_0^t g^2(s)\ ds + \int_0^t g(s)\ dW(s)},$$

is a solution of (6).

The next Lemma can be applied to solve some SDEs.

Lemma 2.1 *If*

$$\begin{cases} dY_1(t) = f_1(Y_1(t))\ dt + g_1(Y_1(t))\ dW(t), \\ dY_2(t) = f_2(Y_2(t))\ dt + g_2(Y_2(t))\ dW(t), \end{cases}$$

where $t \in [0, T]$ *and* $f_i(Y_i(t)) \in L^1(0, T),\ g_i(Y_i(t)) \in L^2(0, T),\ i = 1, 2,$ *then*

$$d(Y_1(t)\ Y_2(t)) = Y_2(t)\ dY_1(t) + Y_1(t)\ dY_2(t) + g_1(Y_1(t))\ g_2(Y_2(t))\ dt.$$

Proof: See [2].

Example 2.2 *consider SDE*

$$\begin{cases} dY(t) = d(t)\ Y(t)\ dt + g(t)\ Y(t)\ dW(t), \quad t \in [0, T] \\ Y(0) = Y_0, \end{cases} \tag{7}$$

where the coefficients d and g are specified functions of time t or constants such that they are Lebesgue measurable and bounded on interval $[0, T]$. Suppose that solution having the product form $Y(t) = Y_1(t)\ Y_2(t)$, where

$$\begin{cases} dY_1(t) = g(t)\ Y_1(t)\ dW(t), \quad t \in [0, T] \\ Y_1(0) = Y_0, \end{cases} \tag{8}$$

and

$$\begin{cases} dY_2(t) = c(t)\ dt + e(t)\ dW(t), \quad t \in [0, T] \\ Y_2(0) = 1, \end{cases} \tag{9}$$

such that the functions c and e will be selected later. By Lemma 2.1 and (8) we have

$$dY(t) = (Y_1(t)\ dY_2(t) + e(t)\ g(t)\ Y_1(t)\ dt) + g(t)\ Y(t)\ dW(t).$$

Now from (7), we can choose c and e such that

$$dY_2(t) + e(t)\ g(t)\ dt = d(t)\ Y_2(t)\ dt.$$

It follows from (9) that

$$e \equiv 0, \quad c(t) = d(t)\ Y_2(t),$$

and therefore the solution of (9) is

$$Y_2(t) = e^{\int_0^t d(s)\ ds}.$$

Since the solution of (8) according to the previous example given by

$$Y_1(t) = Y_0\ e^{-\frac{1}{2}\int_0^t g^2(s)\ ds + \int_0^t g(s)\ dW(s)},$$

we conclude that

$$Y(t) = Y_1(t)\ Y_2(t) = Y_0\ e^{\int_0^t \{d(s) - \frac{1}{2}g^2(s)\}\ ds + \int_0^t g(s)\ dW(s)}, \tag{10}$$

is a solution of SDE (7).

3 Path-wise Approximations For SDEs

In the previous section, we presented some analytical methods for solving SDEs, while other methods exist, SDEs are generally difficult to solve and so numerical methods are required such that these should be designed to perform with a certain order of accuracy. There are two ways for measuring the accuracy of numerical solution of a SDE: strong convergence and weak convergence. Only strong convergency will be considered in this paper. Strong convergence is required, when each trajectory of the numerical solution must be closed to the exact solution.

Definition 3.1 *Let y_N be the numerical approximation of $Y(t_N)$ after N steps with constant step size $h = \frac{t_N - t_0}{N}$, then y is said to convergence strongly to Y with strong global order p if there exist constants K_G, and $\delta > 0$ such that*

$$E(|y_N - Y(t_N)|) \leq K_G\ h^p,$$

for all $h < \delta$.

As with the deterministic case, an entire family of numerical schemes for solving SDEs can be derived from a stochastic Taylor expansion. There are several possibilities to obtain a stochastic Taylor expansion, most notably is the Itô–Taylor expansion that is based on the iterated application of the Itô's formula. All of the strong Taylor approximation schemes are derived by truncating the Itô–Taylor expansion at an appropriate term. The simplest of these methods is Euler–Maruyama method, which is derived by truncating the Itô–Taylor expansion after one deterministic and one stochastic term. As the order of the Euler–Maruyama method is low, the numerical results are inaccurate unless a small step size is used, and clearly more efficient methods are needed. More accurate methods (which require more derivative evaluations) can be obtained by using truncated forms of the stochastic Taylor series expansion but at the cost of derivative evaluations. The most famous of these methods is Milstein method, which is derived by truncating the Itô–Taylor expansion after one deterministic and two stochastic term. In the case $d = 1$, this takes the form

$$y_{n+1} = y_n + h\ g_0(y_n) + \Delta W_n\ g_1(y_n) + \frac{1}{2}((\Delta W_n)^2 - h)\ g_1'(y_n)\ g_1(y_n),$$

with initial value y_0, such that $\Delta W_n = W(\tau_{n+1}) - W(\tau_n)$, for equidistant discretization times $\tau_n = nh$ with $h = \frac{T}{N}$ for some integer N to be large enough such that $h \in (0, 1)$. To avoid this the computational cost, a great deal of attention has been paid to developing derivative–free schemes. One approach is to replace the derivative by differences (see [2]), which for the Milstein method with $d = 1$ leads to the Itô method

$$\begin{aligned} Y_1 &= y_n + \sqrt{h}\ g_1(y_n), \\ y_{n+1} &= y_n + h\ g_0(y_n) + \Delta W_n\ g_1(y_n) + \frac{\sqrt{h}}{2}\ ((\frac{\Delta W_n}{\sqrt{h}})^2 - 1)\ (g_1(Y_1) - g_1(y_n)) \end{aligned} \quad (11)$$

Since $\Delta W_n \sim N(0,h)$, hence $\frac{\Delta W_n}{\sqrt{h}} \sim N(0,1)$, so $\Delta W_n = \sqrt{h}\ R_n$, where $R_n \sim N(0,1)$ and $W(0) = 0$. In order to calculate the error of the Itô approximation we can use the absolute error criterion,

$$\epsilon = E(|y_T - Y(T)|),$$

where y_T and $Y(T)$ denote the Itô approximation and the exact solution at finite terminal time T, for the same sample path of the Wiener process, respectively. Now, in order to investigate ϵ consider SDE

$$\begin{cases} dY(t) = \mu Y(t)\ dt + \sigma Y(t)\ dW(t), & t \in [0,T], \\ Y(0) = Y_0, \end{cases} \tag{12}$$

where μ and σ, are the positive integer numbers. From (10) the exact solution SDE (12) is

$$Y(t) = Y_0\ \exp((\mu - \frac{1}{2}\sigma^2)t + \sigma\ W(t)). \tag{13}$$

Now we apply the scheme (11) for the SDE (12) and then using (13) for determining the corresponding values of the exact solution for the same sample path of the Wiener, that is

$$Y(\tau_n) = Y_0\ \exp((\mu - \frac{1}{2}\sigma^2)\tau_n + \sigma \sum_{i=1}^{n} \Delta W_{i-1}), \quad n = 0, \ldots, N. \tag{14}$$

Figure 1 shows the exact solution and the Itô approximation on the interval $[0,1]$ with stepsize $h = 2^{-7}$ and $Y_0 = y_0 = 1$, $\mu = 1.5$ and $\sigma = 1$, for the same sample path of the Wiener process.

We will try to estimate ϵ statistically. For this purpose, we apply N_1 trajectories for simulation sample paths of process $Y(t)$ and their Itô approximations corresponding on the same sample paths of the Wiener process. Then ϵ be estimated by:

$$\hat{\epsilon} = \frac{1}{N_1} \sum_{k=1}^{N_1} |y_{T,k} - Y(T,k)|,$$

where $y_{T,k}$ and $Y(T,k)$ denote the Itô approximation and the exact solution at terminal time T on the k–th simulated trajectory, respectively. In order to calculate the $\hat{\epsilon}$, we simulate $N_1 = 100$ trajectories of the process $Y(t)$ that satisfying in (12) with $Y_0 = 1,\ \mu = 1.5,\ \sigma = 1$, and then calculate their Itô approximations with initial value $y_0 = 1$, and stepsizes $h = 2^{-k}, k = 4, 5, \ldots, 13$ corresponding to the same paths of the Wiener process on the time interval $[0,T]$ for $T = 1$. The numerical results are shown in Table 1. Here, the

h	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}	2^{-13}
$\hat{\epsilon}$	0.4057	0.1916	0.1196	0.0450	0.0227	0.0191	0.0061	0.0027	0.0017	0.0007

Table 1: Absolute errors $\hat{\epsilon}$ for different stepsizes h.

random numbers generate by $\texttt{randn}(N_1, \#step)$ in Matlab. The command $\texttt{randn}(N_1, \#step)$ creates a $N_1 \times \#step$ matrix of independent $N(0,1)$ samples. In order to make a repeatable simulation, Matlab allows to generate the same random numbers, when the initial states are the same. Here we set the initial state, arbitrarily, to be 50 with the command $\texttt{randn}('state', 50)$. Hence different simulations can be performed by resetting the initial state. Now if we use three different initial states, arbitrarily, to be 68, 94, 123 and consider $\hat{\epsilon}_1$, $\hat{\epsilon}_2$, $\hat{\epsilon}_3$ respectively, are the estimation value of the corresponding absolute errors with these three different initial states then we obtain the numerical results in Table 2. However, these estimates are random variables and take

h	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}	2^{-13}
$\hat{\epsilon}_1$	0.4921	0.2779	0.1787	0.0617	0.0339	0.0175	0.0059	0.0032	0.0013	0.0006
$\hat{\epsilon}_2$	0.4522	0.1917	0.1199	0.0510	0.0162	0.0104	0.0050	0.0020	0.0014	0.0005
$\hat{\epsilon}_3$	0.4676	0.1976	0.1287	0.0579	0.0350	0.0148	0.0059	0.0033	0.0012	0.0004

Table 2: Absolute errors $\hat{\epsilon}_1$, $\hat{\epsilon}_2$, $\hat{\epsilon}_3$ for different stepsizes h.

different values while initial states are chosen different. For large N_1 by the Central Limit Theorem, we conclude that the error $\hat{\epsilon}$ becomes a Gaussian random variable asymptotically and converges in distribution to an non–random expectation ϵ, as $N_1 \to \infty$. Since in practice, it is impossible to generate an infinite number of trajectories, hence, we can estimate the variance σ_ϵ^2 of $\hat{\epsilon}$ and then obtain a confidence interval for the absolute error ϵ. In order to do this, M batches be considered with N_1 simulation paths for each batch and the variance of $\hat{\epsilon}$ be estimated as follows:
Suppose $y_{T,k,j}$ and $Y(T,k,j)$ denote the Itô approximation and the exact solution at time T for k–th simulated trajectory in the j–th batch, respectively. Then the average errors

$$\hat{\epsilon}_j = \frac{1}{N_1} \sum_{k=1}^{N_1} |y_{T,k,j} - Y(T,k,j)|,$$

of the M batches $j = 1, \ldots, M$, are independent and approximately Gaussian distribution for N_1 large enough. Now we can use the Student t–distribution to construct confidence intervals for a sum of independent Gaussian distribution with unknown variance. The mean of the batch averages is estimated by:

$$\bar{\hat{\epsilon}} = \frac{1}{M}\sum_{j=1}^{M}\hat{\epsilon}_j = \frac{1}{M\ N_1}\sum_{j=1}^{M}\sum_{k=1}^{N_1}|y_{T,k,j} - Y(T,k,j)|$$

and then use the formula

$$\hat{\sigma}_\epsilon^2 = \frac{1}{M-1}\sum_{j=1}^{M}(\hat{\epsilon}_j - \bar{\hat{\epsilon}})^2,$$

to estimate the variance $\hat{\sigma}_\epsilon^2$ of the batch averages. For the Student t–distribution with $M-1$ degrees of freedom an $100(1-\alpha)\%$ confidence interval for ϵ has the form $(\bar{\hat{\epsilon}} - \Delta\hat{\epsilon}, \bar{\hat{\epsilon}} + \Delta\hat{\epsilon})$ with

$$\Delta\hat{\epsilon} = t_{1-\alpha,M-1}\sqrt{\frac{\hat{\sigma}_\epsilon^2}{M}}\ , \tag{15}$$

where $t_{1-\alpha,M-1}$ is determined from the Student t–distribution with $M-1$ degrees of freedom. For simulate $M = 10$ batches and each with $N_1 = 100$ trajectories of the process $Y(t)$ that satisfying in (12) with $Y_0 = 1$, $\mu = 1.5$, $\sigma = 0.1$ and their Itô approximations with stepsize $h = 2^{-4}$ corresponding to the same sample paths of the Wiener process on interval $[0, 1]$, we can obtain 90% confidence interval for the absolute error ϵ. If we repeat this procedure for $M = 20, 30, 40, 60$ and 100 batches, such that in each case using the batches already simulated then the results are shown in Figure 2.

Moreover we can obtain the relationship between the absolute error of Itô approximation and the step size. Here, we simulate $M = 20$ batches, each with $N_1 = 100$ trajectories of the process $Y(t)$ that satisfying in SDE (12) with $Y_0 = 1$, $\mu = 1.5$, $\sigma = 0.1$ and their Itô approximations with stepsize $h = 2^{-3}$ corresponding to the same sample paths of the Wiener process on the interval $[0, 1]$, and then evaluate the 90% confidence interval for the absolute error ϵ. If we repeat this for stepsizes $h = 2^{-4}, 2^{-5}, 2^{-6}$ and plot the confidence intervals of ϵ versus stepsize axes, and $log_2(\epsilon)$ versus $log_2(h)$ axes, then the results are shown in Figures 3 and 4.

4 Conclusion

The results in Tables 1 and 2 show that in all cases, the estimate of the absolute error decrease with decreasing stepsize h. Moreover results plotted in Figure 2, indicate that the length of the confidence interval for the absolute error decrease as the number of batches increase. More precisely by (15), we roughly need to increase the number of batches fourfold in order to halve the length of the confidence interval. Finally Figure 3 shows that the stepsize h has a definite effect on the length of the confidence interval. However we could include in Figure 3 the graph of a function $\tilde{\epsilon}(h) = Kh$ for an appropriate constant K which would suggest that the absolute error is proportional to the stepsize. This can see more clearly in Figure 4 the becomes a straight line with slope approximately 1. In fact the exponent 1 of h is corresponding to order of strong convergence Itô method.

References

[1] K. Burrage, P. M. Burrage, *High strong explicit Runge–Kutta methods for stochastic ordinary differential equations,* Appl. Numer. Math. **22** (1996), 81–101.

[2] K. Burrage, P. M. Burrage, *Numerical methods for stochastic differential equations with applications, Advanced computational modelling center,* Univ. Queensland, Australia, 2002.

[3] K. Burrage, T. H. Tian, *Predictor–corrector methods of Runge–Kutta type for stochastic differential equations,* SIAM Numer. Anal. **40** (2002), 1516–1537.

[4] T. C. Gard, *Introduction to stochastic differential equations,* Marcel Dekker, New York, 1988.

[5] P. E. Kloeden, E. Platen, *Numerical solution of stochastic differential equations,* Springer, Berlin, 1995.

[6] B. Oksendal, *Stochastic differential equations, an introduction with applications,* Springer, 1998.

[7] K. Sobczyk, *Stochastic differential equations, with applications to Physics and Engineering,* Kluwer Academic Publisher, Dordrecht, 1991.

[8] T. H. Tian, K. Burrage, *Two–stage stochastic Runge–Kutta methods for stochastic differential equations,* BIT. **42** (2002), 625–643.

Received: April 21, 2006