JOURNAL OF COMPUTATIONAL PHYSICS 113, 75–81 (1994)

Some Experiments on Numerical Simulations of Stochastic Differential Equations and a New Algorithm

W. P. PETERSEN

Interdisziplinäres Projektzentrum für Supercomputing, ETH, Zürich, Switzerland

Received February 21, 1992; revised August 31, 1993

In this paper we compare three second order methods for the numerical integration of Itô stochastic differential equations. One of these methods is new. We consider stability of implicit vs. explicit methods and compare the effects of step size, sample size, and type of increment used to approximate the Brownian motion.

1. INTRODUCTION

Recent reformulations of quantum statistical mechanics in terms of Langevin equations has spurred interest in numerical procedures for simulating the differential equations of continuous Markov processes [1, 2]. In this motivation (e.g., [3]), one is interested in computing an ensemble average[1]

$$\langle f \rangle = \frac{\int e^{-S(x)} f(x) \prod_{i=1}^{m} dx^i}{\int e^{-S(x)} \prod_{i=1}^{m} dx^i},$$

where $f(x)$ is some physical quantity, S is an action on m-dimensional phase space $\{x^i, i = 1, ..., m\}$, with m usually very large. The Langevin approach to approximate $\langle f \rangle$ sets up a convergent stochastic process $x(t)$ varying with simulated time t,

$$dx^i = -\frac{1}{2}\frac{\partial S}{\partial x^i}\,dt + d\omega^i(t) \tag{1}$$

and computes a long-time average (large T)

$$\langle f \rangle \approx \frac{1}{T}\int_0^T dt f(x(t)).$$

In the Langevin equation (1), ω is a m-dimensional Brownian motion. Whenever the m-dimensional phase space is non-Euclidean (e.g., confined to the surface of a torus or a sphere), the unit coefficient of $d\omega$ in (1) will appear as a diffusion coefficient matrix depending on the process x [2].

Although the above example illustrates a familiar motivation, Monte Carlo simulations of such processes have a wider scope in approximations of many types of multidimensional partial differential equations of elliptic and parabolic type using generalizations of the Feynman–Kac formula [4].

Thus, here we describe simulations of stochastic differential equations of a general form

$$dx = b(x)\,dt + \sigma(x)\,d\omega(t), \tag{2}$$

where process $\omega(t)$ is a vector-valued Brownian motion whose properties are entirely determined by the conditions ($r, s = 1, ..., m$, where m is the size of vectors x, ω):

$$\omega^r(0) = 0$$

$$\langle \omega^r(t) \rangle = 0$$

$$\langle \omega^r(t_1)\,\omega^s(t_2) \rangle = \min(t_1, t_2)\,\delta_{rs}.$$

Drift coefficient $b(x)$ is a vector and diffusion coefficient $\sigma(x)$ is a matrix.

The diffusion coefficient $\sigma(x)$ depending on x introduces difficulties. Since the Brownian motion (Wiener process) $\omega(t)$ is not of bounded variation, one must be careful defining the meaning of the diffusion. Here, we have chosen the Itô definition [5] wherein (2) is shorthand for

$$x(t) = x(0) + \int_0^t b(x(\tau))\,d\tau + \int_0^t \sigma(x(\tau))\,d\omega(\tau), \tag{3}$$

with the diffusion term interpreted as a belated stochastic integral

$$\int_0^t \sigma(x(\tau))\,d\omega(\tau) = \lim_{\Delta t \to 0} \sum_{i=0}^{n-1} \sigma(x(t_i))[\omega(t_{i+1}) - \omega(t_i)].$$

[1] Superscripts are used to indicate vector indices in order to distinguish them from discrete time-steps denoted by subscripts.

0021-9991/94 $6.00

Here, $\Delta t = \max_j (t_j - t_{j-1})$ on the ordered lattice $0 = t_0 < t_1 < \cdots < t_n = t$ of time points.

2. NUMERICAL APPROXIMATIONS

For simplicity we consider first the case that process x is a scalar complex variable. Generalizations to vector-valued complex processes are fairly straightforward and follow below. We wish to find a discrete approximation in the finite interval (**stepsize**) h. Our new algorithm will rely entirely upon the following decomposition.

$$\Delta x = \int_t^{t+h} b(x(\tau))\, d\tau + \int_t^{t+h} \sigma(x(\tau))\, d\omega(\tau)$$
$$= \Delta A + \Delta M.$$

The increment ΔA (drift) is of bounded variation $O(h)$, and ΔM is a martingale [5] of $O(h^{1/2})$ in probability. This means ΔM has the expectation values $\langle \Delta M \rangle = 0$ and $\langle (\Delta M)^2 \rangle = O(h)$.

We will be content with a *weak solution*.[2] That is, $\forall f(x) \in \mathbf{C}^4$ we want [7]

$$\left\langle f\left(x_0 + \int_t^{t+h} dx(\tau)\right)\right\rangle$$
$$= \langle f(x_0) + f'(x_0)\,\Delta x + \tfrac{1}{2} f''(x_0)\,\Delta x^2 + \tfrac{1}{6} f'''(x_0)\,\Delta x^3$$
$$+ \tfrac{1}{24} f^{(4)}(x_0)\,\Delta x^4 \rangle + O(h^3). \tag{4}$$

Having chosen the Itô definition of the diffusion, we insist that the martingale ΔM satisfy

$$\langle \Delta M \rangle = 0 \tag{5}$$

to all orders in h.

3. THE SCHEME OF MIL'SHTEIN

We start with Mil'shtein's [7] simplified Taylor series. Repeated substitutions of the left-hand side of (3) for $t = h$ into the right-hand side of (3) yields the following series (where ξ is some model for $\Delta\omega$):

$$\Delta x = \underbrace{b_0 h + \tfrac{1}{2} b_0' \sigma_0 h\xi + \tfrac{1}{2}(bb' + \tfrac{1}{2} b''\sigma^2)_0\, h^2}_{\text{drift}}$$
$$+ \underbrace{\sigma_0 \xi + \tfrac{1}{2}\sigma\sigma_0'(\xi^2 - h) + \tfrac{1}{2}\sigma_0' b_0 h\xi + \tfrac{1}{4}\sigma_0^2 \sigma_0'' h\xi}_{\text{diffusion}}. \tag{6}$$

The tricks used to obtain this approximation are (from [7]) as follows (readers unfamiliar with the Itô calculus are referred to the book by Kloeden and Platen [6]):

- In the second term in the **diffusion**, the Itô rule gives

$$d(\omega^2) = 2\omega\, d\omega + dt;$$

hence,

$$\int_0^h \omega\, d\omega = \tfrac{1}{2}(\xi^2 - h). \tag{7}$$

- To $O(h^2)$ in terms of (4) we approximate

$$\int_0^h \tau\, d\omega(\tau) \approx \tfrac{1}{2} h\xi.$$

- Likewise (integrate by parts since τ is deterministic),

$$\int_0^h \omega(t)\, d\tau \approx \tfrac{1}{2} h\xi.$$

- And, because the integral is belated,

$$\int_0^h \omega(\tau)^2\, d\omega(\tau) \approx \tfrac{1}{2} h\xi \approx \tfrac{1}{2}\xi_0^2 \xi,$$

where ξ_0 is independent of and identically distributed to ξ.

- Finally, one ignores the $O(h^2)$ term in diffusion. Since this term has vanishing expectation (5), no *inner product* (within (4)) with any stochastic term of at least $O(h^{1/2})$ will contribute to $O(h^2)$. Furthermore, no inner product with any deterministic term will contribute to this order either.

In this paper we use a **model**, ξ, for increments of the Brownian motion,

$$\xi \approx \Delta\omega = \int_t^{t+h} d\omega(\tau),$$
$$= \sqrt{h}\, z,$$

where z is a zero mean, unit variance, symmetrically distributed random variable (e.g., gaussian).

4. A NEW ALGORITHM

Starting with Mil'shtein's Taylor series (6), we can derive a Runge–Kutta type algorithm. Note that an explicit trapezoidal rule

$$\frac{h}{2}\left(b(x_0) + b(x_0 + \sigma_0\xi_1 + b_0 h)\right)$$
$$= b_0 h + \frac{1}{2} b_0' \sigma_0 h\xi_1 + \frac{1}{2}\left(bb' + \frac{1}{2} b''\sigma^2\right)_0 h^2$$

obtains all of the drift in (6) correct to $O(h^2)$.

[2] A detailed discussion of weak vs strong solutions is in Kloeden and Platen [6].

Now look at the diffusion: with ξ_0 an estimation for an intermediate $\Delta\omega$ step, we start with

$$\sigma(x_0+\alpha\sigma_0\xi_0+\beta b_0 h)\,\xi_1 = \{\sigma_0+\alpha\sigma_0'\sigma_0\xi_0+\beta\sigma_0' b_0 h+\tfrac{1}{2}\alpha^2\sigma_0''(\sigma_0)^2\,\xi_0^2 + \alpha\beta\sigma_0''\sigma_0 b_0\xi_0 h+\tfrac{1}{3!}\alpha^3\sigma_0'''(\sigma_0)^3\,\xi_0^3\}\,\xi_1+O(h^2).$$

Thus, to $O(h^2)$, we examine the pair

$$\tfrac{1}{2}\{\sigma(x_0+\alpha\sigma_0\xi_0+\beta b_0 h)+\sigma(x_0-\alpha\sigma_0\xi_0+\beta b_0 h)\}\,\xi_1 = \sigma_0\xi_1+\beta\sigma_0' b_0 h\xi_1+\tfrac{1}{2}\alpha^2\sigma_0''(\sigma_0)^2\,\xi_0^2\xi_1$$

with choice $\alpha=2^{-1/2}$ and $\beta=\frac{1}{2}$. In evaluations of correlations in (4), with respect to *inner-products* with deterministic terms and the only $O(h^{1/2})$ term to contribute to $O(h^2)$, namely $\xi_1:\xi_0^2\approx h$. Finally, adding the Itô term $\frac{1}{2}\sigma_0\sigma_0'(\xi_1^2-h)$ obtains all of the diffusion correct from (6). Hence, we have the following second-order scheme:

$$\begin{aligned} x_h = x_0 &+ \frac{h}{2}\,(b(x_0)+b(x_0+\sigma_0\xi_1+b_0 h)) \\ &+ \frac{1}{2}\left\{\sigma\left(x_0+\sqrt{1/2}\,\sigma_0\xi_0+\frac{1}{2}b_0 h\right)\right. \\ &\left.+\,\sigma\left(x_0-\sqrt{1/2}\,\sigma_0\xi_0+\frac{1}{2}b_0 h\right)\right\}\xi_1 \\ &+ \frac{1}{2}\sigma_0\sigma_0'(\xi_1^2-h). \end{aligned} \tag{8}$$

Furthermore, replacing the $b(x_0+\sigma_0\xi_0+b_0 h)$ term by $b(x_h)$ yields an **implicit** form [8, 9] which is also weakly $O(h^2)$ accurate. This is straightforward and shown by repeated substitution and comparing with the Mil'shtein Taylor series. Maple [10] was used for this repeated substitution. Thus,

$$\begin{aligned} x_h = x_0 &+ \frac{h}{2}\,(b(x_0)+b(x_h)) \\ &+ \frac{1}{2}\left\{\sigma\left(x_0+\sqrt{1/2}\,\sigma_0\xi_0+\frac{1}{2}b_0 h\right)\right. \\ &\left.+\,\sigma\left(x_0-\sqrt{1/2}\,\sigma_0\xi_0+\frac{1}{2}b_0 h\right)\right\}\xi_1 \\ &+ \frac{1}{2}\sigma_0\sigma_0'(\xi_1^2-h) \end{aligned} \tag{9}$$

is also second order, but due to the appearance of x_h in the right-hand side, a Newton method or something similar must be used to find x_h at each step. For so-called stiff equations, those with widely differing time scales, this form is significantly more stable. Section 6 contains an example.

5. THE ALGORITHM OF TALAY

Another interesting algorithm, accurate to $O(h^2)$ in the weak sense, is the method of Talay [11]. It uses two random variables per time step and is a two-stage method:

$$\begin{aligned} x_1 &= 1/\sqrt{2}\,\sigma_0\xi_0+\tfrac{1}{2}(b-\tfrac{1}{2}\sigma\sigma')_0\,h+\tfrac{1}{4}\sigma_0\sigma_0' h\xi_0^2 \\ x_h &= x_0+\{\sqrt{2}\,\sigma_0\xi_0+\sqrt{2}\,\sigma(x_1)\,\xi_1-1/\sqrt{2}\,\sigma_0(\xi_0+\xi_1)\} \\ &\quad+(b-\tfrac{1}{2}\sigma\sigma')(x_1)\,h \\ &\quad+\{\tfrac{1}{2}\sigma_0\sigma_0'\xi_0^2 h+\tfrac{1}{2}\sigma\sigma'(x_1)\,\xi_1^2 h-\tfrac{1}{4}\sigma_0\sigma_0'(\xi_0+\xi_1)^2 h\}. \end{aligned} \tag{10}$$

6. MULTIVARIATE CASE

The generalization to the vector-valued case is straightforward and follows the expansion in a Taylor series according to the method of Mil'shtein [7, 8]. The only difficulty is to model the integrals (vector indices run $r, s = 1, \ldots, m$, where m is the size of the $\omega(t)$ vector),

$$\int_0^h \omega^r\,d\omega^s = hZ^{rs}.$$

A simple model (e.g., Talay [11], or [6] Section 5.12) satisfies the calculus to $O(h^2)$:

$$Z^{rs}=\begin{cases}\frac{1}{2}(z_1^r z_1^s-1) & \text{if } r=s,\\ \frac{1}{2}(z_1^r z_1^s-\tilde{z}^{rs}) & \text{for } r>s,\\ \frac{1}{2}(z_1^r z_1^s+\tilde{z}^{rs}) & \text{for } r<s.\end{cases}$$

Here, the mutually independent z_1^r variables are the zero mean, variance = 1, random variables which appear in (11). The $\tilde{z}^{rs}$ random variables are independent of the set $\{z_1^r, r=1, \ldots, m\}$ and likewise are mean zero and unit variance. The construction of the Z matrix requires $m(m-1)/2$ of these $\tilde{z}$ variables. For example, in a two-dimensional case, the approximation for Z in $\int_0^h \omega^r\,d\omega^s = hZ^{rs}$ looks like

$$\begin{bmatrix} Z^{11}, & Z^{12}\\ Z^{21}, & Z^{22}\end{bmatrix}=\begin{bmatrix}\frac{1}{2}(z_1^1 z_1^1-1), & \frac{1}{2}(z_1^1 z_1^2+\tilde{z})\\ \frac{1}{2}(z_1^2 z_1^1-\tilde{z}), & \frac{1}{2}(z_1^2 z_1^2-1)\end{bmatrix},$$

where z_1^1, z_1^2, and $\tilde{z}=\tilde{z}^{12}$ are the needed random variables. The z_1 variables are the same as those appearing in the drift and outside of the diffusion in (11). All the z_1 and $\tilde{z}$ variables are mutually independent and, furthermore, mutually independent of the vector z_0 which appears in the inner portion of (11). Hence, (11) requires $(m^2+3M)/2$ univariant, zero-mean variables. Section 8 discusses the forms of these variables in more detail.

Altogether, the algorithm for the multivariate case ($k=1, \ldots, m$) is

$$x_h^k = x_0^k + \frac{h}{2}\left(b^k(x_0) + b^k(x_0 + \sqrt{h}\,\sigma_0 z_1 + b_0 h)\right) + \frac{\sqrt{h}}{2}\sum_{r=1}^{m}\left\{\sigma^{kr}\left(x_0 + \sqrt{h/2}\,\sigma_0 z_0 + \frac{h}{2}b_0\right) + \sigma^{kr}\left(x_0 - \sqrt{h/2}\,\sigma_0 z_0 + \frac{h}{2}b_0\right)\right\} z_1^r + h\sum_{l,r,s=1}^{m}(\partial^l \sigma_0^{kr})\,\sigma_0^{ls} Z^{rs}. \tag{11}$$

Indices in (11) for inner function arguments are suppressed. In the scalar case, an alternate form uses $z_0 = 1$; both forms are weakly $O(h^2)$ accurate. The random variable matrix Z^{rs} is constructed from z_1^r, z_1^s, and $\tilde{z}^{rs}$ as above. Also, the notation $\partial^l(\arg)$ means $\partial(\arg)/\partial x^l$. Algorithm (11) has been tested for a $d = 2 + 1$ lattice fermion model by Klauder and Lee [12].

Implicit Forms

An implicit form for the multivariate case analogous to (9) is the same as (11) with the $b^k(x_0 + \sqrt{h}\,\sigma_0 z_1 + b_0 h)$ term in the drift replaced by $b^k(x_h)$. Solving this for the vector x_h usually requires an implicit numerical method (see, e.g., [13]), except when b is linear.

A simple example illustrating the increase in stability by using implicit methods is given by a two-dimensional Ornstein–Uhlenbeck equation. This problem is a straightforward generalization of the stiffness example for ordinary differential equations found in Bulirsch and Stoer [14, p. 462],

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \frac{dt}{2}\begin{bmatrix} \lambda_1 + \lambda_2, & \lambda_1 - \lambda_2 \\ \lambda_1 - \lambda_2, & \lambda_1 + \lambda_2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}. \tag{12}$$

Written in matrix form it is

$$d\mathbf{u} = dt\, A\mathbf{u} + d\mathbf{w},$$

where $\mathbf{u} = (x, y)^{\mathrm{T}}$ and $\mathbf{w} = (w_1(t), w_2(t))^{\mathrm{T}}$. The solution to (12) is formally

$$\mathbf{u}(t) = e^{At}\mathbf{u}(0) + \int_0^t e^{A(t-\tau)}\, d\mathbf{w}(\tau). \tag{13}$$

This process becomes stationary for large t when the eigenvalues of A have negative real parts. As $t \to \infty$, the initial value part of (13), $e^{At}\mathbf{u}_0 \to 0$. In that case, the asymptotic correlation matrix is

$$\langle u^i(\infty)\, u^j(\infty)\rangle \equiv \lim_{t\to\infty}\langle u^i(t)\, u^j(t)\rangle = \lim_{t\to\infty}\sum_k\sum_l \int_0^t\int_0^t [e^{A(t-\tau)}]^{ik}\,[e^{A(t-\eta)}]^{jl} \times \langle dw^k(\tau)\, dw^l(\eta)\rangle. \tag{14}$$

This can be easily evaluated using $\langle dw^k(\tau)\, dw^l(\eta)\rangle = \delta_{kl}\,\delta(\tau - \eta)\, d\tau\, d\eta$, giving an integral of a matrix exponential. Our matrix A is symmetric, commutes with its transpose, thus by (14) and the ergodic theorem

$$\langle u^i(\infty)\, u^j(\infty)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T u^i(t)\, u^j(t)\, dt = -\frac{1}{2}[A^{-1}]^{ij}. \tag{15}$$

The second-order Runge–Kutta method (**explicit trapezoidal rule**) (11) applied to (12) is

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \tfrac{1}{2}\{A\mathbf{u}_k + A(\mathbf{u}_k + hA\mathbf{u}_k + \sqrt{h}\,\mathbf{z}_k)\} + \sqrt{h}\,\mathbf{z}_k = (1 + hA + \tfrac{1}{2}h^2A^2)\,\mathbf{u}_k + (1 + hA/2)\sqrt{h}\,\mathbf{z}_k, \tag{16}$$

while an **implicit trapezoidal rule** (implicit form of (11)) for (12) is

$$\mathbf{u}_{k+1} = (1 - hA/2)^{-1}\,[(1 + hA/2)\,\mathbf{u}_k + \sqrt{h}\,\mathbf{z}_k)]. \tag{17}$$

The stability condition is that the drift part of the one-step fundamental matrix is contracting (see [14] or [9]). Which for the explicit rule (16) can be shown for $|\lambda_2| > |\lambda_1|$ to be

$$\|1 + hA + \tfrac{1}{2}h^2A^2\| = |1 + h\lambda_2 + (h\lambda_2)^2/2| < 1.$$

This condition is $0 > \lambda_2 > -2/h$. On the other hand, the stability condition for the implicit trapezoidal rule (17) is

$$\|(1 - hA/2)^{-1}(1 + hA/2)\| = |(1 + h\lambda_2)/(1 - h\lambda_2)| < 1$$

which is always satisfied for $\lambda_2 < 0$. We shall see in the results, Section 8, that indeed, the implicit form is absolutely stable (for big $|\lambda_2|$), while the explicit form fails precisely where expected.

7. TEST PROBLEMS

The first (**A**) illustrates the stability argument of the previous section. Namely, we show that for (**A**), when $\lambda_2 < -2/h$, the usual argument for ordinary differential equations holds in the stochastic case. Implicit algorithms show better linear stability. Algorithms Mil'shtein (6), Talay (10), and (8) were tested on the following additional problems. Problem (**B**) illustrates the simplest linear diffusion coefficient case. Both α and β were allowed to be complex. Problems (**C**) and (**D**) are examples of semi-bounded and bounded processes, respectively. We chose to examine *accuracy* in these Monte-Carlo simulations by looking at root-mean-square errors (over a range of k) in characteristic functions,[3] $\phi(k) = \langle e^{ikx(t)}\rangle$ compared to sample averages

[3] These were also computed numerically. The infinite range integrals were done using QUADPACK [15], finite range distributions by discrete Fourier transform.

$(1/N)\sum_{j=1}^{N} e^{ikx^{[j]}(t)}$. In addition, root-mean-square comparisons of the distributions of the processes' real and imaginary parts were also made.

The characteristic function as a vehicle for measuring accuracy had the following advantages: (i) $|\phi(k)| \leqslant 1$ is bounded, which avoids round-off errors which plague moment comparisons, (ii) all of the properties of the distribution are tested, and (iii) rms estimates of the characteristic function in complex linear cases examined the distributions of the processes' real and imaginary parts. Advantage (i) is important because moment comparisons are dangerous. For example, in the linear problem below, when $\beta = 0$ and $\alpha = 1$, as t increases this martingale develops increasingly larger populations of very small-valued paths, but increasingly fewer and growing (in size) large-valued ones. When averages are taken, contributions from small-valued paths are truncated by round-off error when added to the large-valued ones. All higher order moments look too low [16] because contributions from the small ones are underestimated. Advantage (iii) means that separate histograms of real parts and imaginary parts do not show any relationship between these parts.

The Test Problems

A. $\mathbf{u}(t) = e^{At}\mathbf{u}(0) + \int_0^t e^{A(t-\tau)}\, d\mathbf{w}(\tau)$;

SDE, $d\mathbf{u} = dt\, A\mathbf{u} + d\mathbf{w}$,

B. $x = e^{(\beta - (1/2)\alpha^2)t + \alpha\omega}$;

SDE, $dx = \beta x\, dt + \alpha x\, d\omega$,

C. $x = \cosh(\omega)$;

SDE, $dx = \frac{1}{2}x\, dt + \sqrt{x^2 - 1}\, d\omega$,

D. $x = 1/(a + \omega^2)$;

SDE, $dx = -(4ax^3 - 3x^2)\, dt - 2x\sqrt{x - ax^2}\, dw$.

8. RESULTS

Not all of the experimental results can be illustrated here simply because of space. The qualitative conclusions, however, are supported by the figures shown.

Stability

Figure 1 shows long-time ergodic averages $\langle u^1(\infty)\, u^1(\infty)\rangle$ and $\langle u^1(\infty)\, u^2(\infty)\rangle$ given by the right-hand side of (14). These were computed by discrete sums of process (16) and process (17), respectively, while the lines were computed from (15). Figure 1 shows that the explicit rule (16) is not stable when the time constants λ_1 and λ_2 have widely differing scales. Conversely, the implicit rule (17) is absolutely stable for all reasonable ratios $|\lambda_2|/|\lambda_1|$. It was gratifying to see that Fig. 1 also shows that (16) fails when

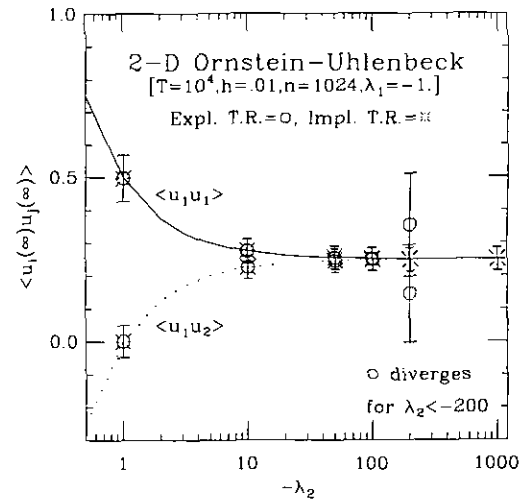

FIG. 1. Stability of explicit trapezoidal rule (16) (○) vs implicit trapezoidal rule (17) (*).

$|\lambda_2| > 2/h$, and this inequality is sharp. This result was expected from discretized ordinary differential equations [14] and Section 6 above. Each simulation (16) and (17) used a sample of $n = 1024$, and from (15) $T = 10000$ (10^6 timesteps), with stepsize $h = 0.01$. Error bars indicate the variance among the $n = 1024$ sample of long time averages.

Sample Size

It should not come as any shock, but Fig. 2 clearly shows that the best results come from large sample sizes. In this figure, the $1/\sqrt{N}$ behavior of the error is clearly illustrated. The example is from the problem **B** with the parameters shown. (The notation is RK = Runge–Kutta (8); GNM = Mil'shtein (6); and DT = Talay (10).) For the other test examples the results are similar; namely, there are no essen-

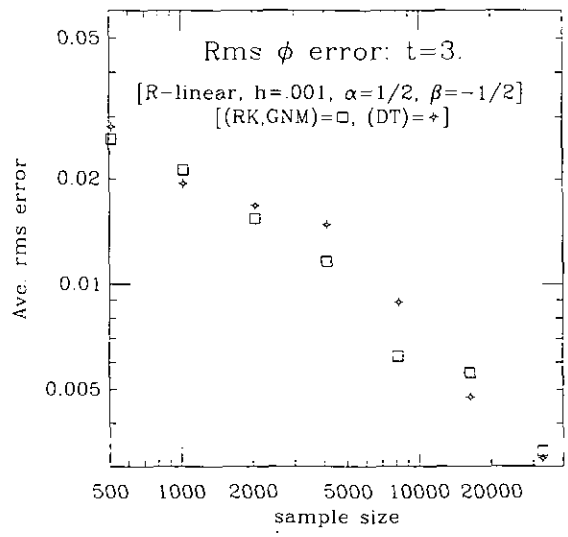

FIG. 2. Effect of sample size for real linear problem.

tial accuracy differences between the algorithms when the problem is well posed and stable.

Vector/parallel computers make taking larger samples more attractive. That is, simultaneous paths were computed in inner loops according to

$$x^{[i]} = x_0^{[i]} + \Delta A^{[i]} + \Delta M^{[i]}, \qquad i = 1, \ldots, N,$$

where, random variables z in (8) were computed by pairs using the Box–Muller [17] method. Both lagged Fibonacci series [18] and standard linear congruential generator methods were tested. There were no discernible differences in the results—except on the Cray Y-MP (sn1522 at ETH, Zürich) and the NEC SX-3 in Manno the lagged Fibonacci series method is faster.

Effects of $\Delta\omega$ Increment

Results were fairly insensitive to the form of the $\Delta\omega$ approximation. Several forms were tried: gaussians (as above), simple uniforms ($\Delta\omega \approx 2\sqrt{3h}\,u$, where $u \in (-\frac{1}{2}, +\frac{1}{2})$), and *hats* ($\Delta\omega \approx \sqrt{6h}\,(u_1 + u_2)$, where $u_1, u_2 \in (-\frac{1}{2}, +\frac{1}{2})$).

The only exception to this is the three-state form often seen in the literature:

$$\Delta\omega = \begin{cases} +\sqrt{3h} & \text{with probability } \frac{1}{6}, \\ -\sqrt{3h} & \text{with probability } \frac{1}{6}, \\ 0 & \text{with probability } \frac{2}{3}. \end{cases}$$

In Fig. 3, we see a plot of in $\mathrm{Re}(\phi(k))$ for 64 points in the range shown. The three-state model for $\Delta\omega$ shows a lot of scatter for the high wave numbers. The discontinuous distribution three-state model for $\Delta\omega$ introduces unwanted high-frequency modes to which $\phi(k)$ is sensitive. Since this result was at first surprising, it is presented as a note of caution. Studies using uniform distributions and *hats* were indistinguishable from gaussians.

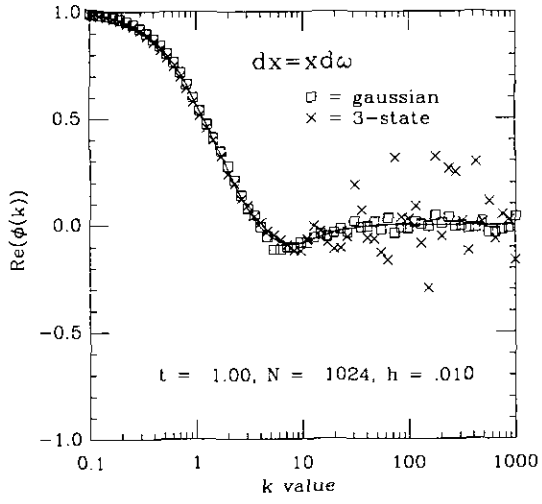

FIG. 3. Discrete $\Delta\omega$ comparison.

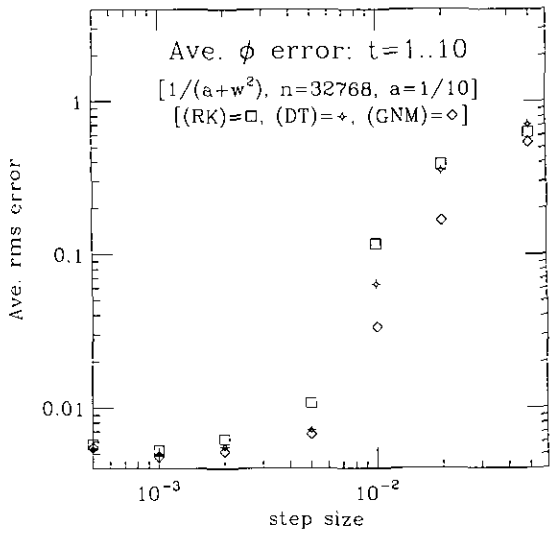

FIG. 4. Rms error vs h for $1/(a+\omega^2)$ problem.

Effects Due to Step Size

The drift increment (ΔA) is $O(h)$ while the diffusion (ΔM) is $O(h^{1/2})$. Thus, if the drift coefficient $b(x)$ and the diffusion coefficient $\sigma(x)$ are relatively the same size, ΔM is $O(1/\sqrt{h})$ larger than ΔA! This is significant because *unless the drift coefficient is large, the step-size will not be very important.* For this reason, the importance of order ($O(h^2)$ in the algorithms here) takes a very different importance than in ordinary differential equations.

This explains why there are many simulations of Langevin equations which use the forward Euler method fairly successfully. However, when $b(x)$ is large, the step size and order become important. In Figs. 4 and 5 we illustrate

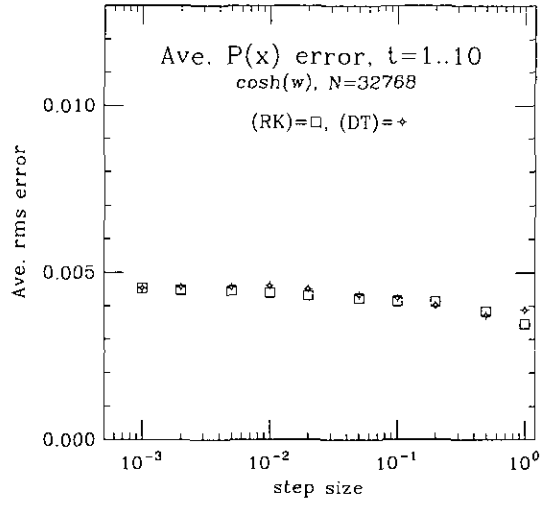

FIG. 5. Rms error vs h for $\cosh(w)$ problem.

this point. In Fig. 4 the rms error in the characteristic function for the process $x = 1/(a+\omega^2)$ is plotted. This function was computed using discrete Fourier transform [19] for 128 bins of a uniform histogram. In Fig. 5, a similar rms error in the distribution function for $x = \cosh(\omega)$ (problem **B**) is plotted. The characteristic function for $x = \cosh(\omega)$ is highly oscillating and is hard to scale. The comparison of Fig. 4 and Fig. 5 is quite striking. Namely, Fig. 4 shows quadratic behavior in the error with respect to step-size, while in Fig. 5 the error shows flat behavior! Tests on the simple martingale $dx = x\,d\omega$ show the same behavior as in $x = \cosh(\omega)$: accuracy of the results are very insensitive to step-size; only sample size matters. On the other hand, for the $x = 1/(a+\omega^2)$ problem, the drift coefficient behaves initially ($t \sim 0$) like $-4/a^2$, which for small values of a is very large; hence we see a strong sensitivity to step-size.

9. CONCLUSIONS

The experiments presented here support the following conclusions:

1. The three algorithms tested here, the Taylor expansion of Mil'shtein (6), the two-stage algorithm of Talay (10), and the Runge–Kutta of Section 4 (8), show no essential accuracy differences. Due to the extra derivatives (function calls) in (6), this algorithm is slower. In the scalar process case (8) is fastest since it uses only one random variable per step. For the multivariate case the number of random variables and number of function calls per step in (8) and (10) is the same. However, example (**A**) illustrates that just as in discrete ordinary differential equations simulations, implicit stability can be important. That (9) and (11) lend themselves to implicit (and semi-implicit) variants is very attractive.

2. Sample size determines accuracy typical of Monte-Carlo—namely, $1/\sqrt{N}$ behavior.

3. The form of the increment model ($\xi \approx \Delta\omega$) for the Brownian motion (Fig. 3) did not seem very important—continuously distributed uniforms, gaussians, and *hats* worked equally well. However, a discontinuously distributed increment like the three-state model ($\pm\sqrt{3h}$, 0) introduced high-frequency problems that showed up right away in characteristic functions. This model also costs more to compute than a simple uniform model, $\Delta\omega \approx \sqrt{3h}\,(2u-1)$, where $u \in (0, 1)$ is from a standard Fortran library random number generator.

4. For the linear test problem $dx = \beta x\,dt + \alpha x\,d\omega$, a range of values for α and β were examined for which $\phi(k)$ is defined: (real α, β) *or* (α purely imaginary). From which we conclude complex processes seem to pose no obvious difficulties.

5. Finally, for strongly diffusion driven processes ($b < \sigma$) the simulations are fairly step-size insensitive.

The author thanks Sean Lee for valuable discussions and testing of the multivariate algorithm. The author's e-mail address is wpp@ips.ethz.ch.

REFERENCES

1. G. Parisi and Y.-S. Wu, *Sci. Sin.* **24**, 483 (1981).
2. H. Gausterer and J. R. Klauder, *Phys. Rev. D* **33**, 3678 (1986).
3. J. R. Klauder and W. P. Petersen, *J. Statist. Phys.* **39**, 53 (1985).
4. M. Freidlin, *Functional Integrals and Partial Differential Equations*, Ann. of Math. Stud., Vol. 109, (Princeton Univ. Press, Princeton, NJ, 1985).
5. K. Itô, *Zenkoku Shijo Sugaku Danwakai* **244**(1077), 1352 (1942).
6. P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992).
7. G. N. Mil'shtein, *Theory Probab. Appl.* **23**, 396 (1976).
8. G. N. Mil'shtein, *Theory Probab. Appl.* **30**(4), 750 (1985).
9. D. B. Hernandez and R. Spigler, "*BIT*" **33**, 654 (1993).
10. B. W. Char, K. O. Geddes, *et al.*, *Maple 5 Language Reference Manual* (Springer-Verlag, New York/Berlin, 1991).
11. D. Talay, *Math. Modeling Numer. Anal.* **20**(1), 141 (1986).
12. J. R. Klauder and S. Lee, *Phys. Rev. D* **45**(6), 2101 (1992).
13. J. Ortega and W. Rheinbolt, *Numerical Solutions of Non-linear Systems of Equations* (Academic Press, New York, 1970).
14. R. Bulirsch and J. Stoer, *Introduction to Numerical Analysis* (Springer-Verlag, New York, 1980).
15. R. Piessens, E. D. Doncker-Kapenga, *et al.*, *QUADPACK—A Subroutine Package for Automatic Integration* (Springer-Verlag, Berlin, 1983).
16. J. R. Klauder and W. P. Petersen, *SIAM J. Numer. Anal.* **22**(6) (1985).
17. M. E. Muller, *J. Assoc. Comput. Mach.* **6**, 376 (1959).
18. W. P. Petersen, *J. Supercomput.* **1**, 327 (1988).
19. *CFFT2, Complex Binary Radix FFT*, Cray Research, Inc., Technical Note 2240203, June 1978.