MATHEMATICS OF COMPUTATION
VOLUME 49, NUMBER 180
OCTOBER 1987, PAGES 523–542

Numerical Solution of Stochastic Differential Equations with Constant Diffusion Coefficients

By Chien-Cheng Chang

Abstract. We present Runge-Kutta methods of high accuracy for stochastic differential equations with constant diffusion coefficients. We analyze L_2 convergence of these methods and present convergence proofs. For scalar equations a second-order method is derived, and for systems a method of order one-and-one-half is derived. We further consider a variance reduction technique based on Hermite expansions for evaluating expectations of functions of sample solutions. Numerical examples in two dimensions are presented.

1. Introduction. Recently, the numerical solution of stochastic differential equations has attracted the attention of researchers in many fields, both in probability theory and in its applications. Most of the methods that have been developed are of Taylor series type, in which one needs to evaluate derivatives, and thus has difficulties in applications to practical problems. Therefore, we are especially interested in methods of Runge-Kutta type, i.e., one-step methods which require no approximation to the derivatives of the functions involved. In this paper we consider the d-dimensional stochastic differential equation

$$d\mathbf{X} = \mathbf{f}(\mathbf{X})\,dt + \nu d\mathbf{W}_t, \qquad 0 < t < T, \tag{1-1}$$

where $\nu \geqslant 0$ is a constant, $\mathbf{f} = \mathbf{f}(\mathbf{x})$ is a sufficiently smooth function satisfying a Lipschitz condition in $\mathbf{x}$, and $\mathbf{W}_t$ ($t \geqslant 0$) is a d-dimensional Wiener process (Brownian motion). This equation can be interpreted either in the sense of Ito or in that of Stratonovich [1]. Our results extend readily to the case where $\mathbf{f}$ depends also on the time by introducing t as an additional dependent variable and imagining that the associated component of the Wiener process is zero.

Equation (1-1) occurs in the study of several physical phenomena, e.g., in the motion of a particle in the collision theory of chemical reactions [2], in blood clotting [11], in stellar dynamics [3], signal modeling in communication systems [14], and the stochastic behavior of fluid particles in turbulence theory [8].

We develop and analyze high-accuracy methods of constructing sample solutions of (1-1), and we further consider a variance reduction technique for evaluating expectations of functions of these sample solutions. Consider the partition of the interval $[0, T]$ given by

$$\Pi = (0 = t_0, \ldots, t_n, t_{n+1} = t_n + h, \ldots, t_N = T).$$

Received December 5, 1985; revised July 11, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65U05; Secondary 65C05.

0025-5718/87 $1.00 + $.25 per page

Let E denote expectation and $\|\cdot\|$ the two-norm in R^d space. We say that a random variable $\mathbf{Z}$ is of order p in the L_q sense if there is a constant C such that

$$(E\|\mathbf{Z}\|^q)^{1/q} \leqslant Ch^p$$

for all sufficiently small $h \geqslant 0$, and that a numerical scheme has order p in the L_q sense if $\mathbf{X}^{(n)} - \mathbf{X}(t_n)$ is of order p in the L_q sense, where $\mathbf{X}^{(n)}$ and $\mathbf{X}(t_n)$ are, respectively, the numerical and the exact solution of the stochastic differential equation at time t_n. In the present paper we adopt an L_2-norm analysis because it can best exhibit the nonanticipating property [1] of the solutions of stochastic differential equations. Our main results are a second-order scheme for scalar equations,

$$\text{(T)}\quad \begin{aligned} &\mathrm{P}^{(n)} = v\sqrt{\theta - \beta^2}, \\ &\mathrm{Q}^{(n)} = \mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)}) + v\sqrt{h}\,\beta, \\ &\mathrm{X}^{(n+1)} = \mathrm{X}^{(n)} + v\Delta W^{(n)} + \tfrac{1}{2}h\big[f(\mathrm{Q}^{(n)} + \sqrt{h}\,\mathrm{P}^{(n)}) + f(\mathrm{Q}^{(n)} - \sqrt{h}\,\mathrm{P}^{(n)})\big]; \end{aligned}$$

and a scheme of order one-and-one-half for systems,

$$\text{(S)}\quad \begin{aligned} &\mathbf{Q}^{(n)} = \mathbf{X}^{(n)} + \tfrac{1}{2}h\mathbf{f}(\mathbf{X}^{(n)}), \\ &\mathbf{Q}^{*(n)} = \mathbf{X}^{(n)} + \tfrac{1}{2}h\mathbf{f}(\mathbf{X}^{(n)}) + \tfrac{3}{2}v\sqrt{h}\,\boldsymbol{\beta}, \\ &\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + v\Delta\mathbf{W}^{(n)} + \tfrac{1}{3}h\big[\mathbf{f}(\mathbf{Q}^{(n)}) + 2\cdot\mathbf{f}(\mathbf{Q}^{*(n)})\big]. \end{aligned}$$

Here, $\Delta\mathbf{W}^{(n)} \equiv \mathbf{W}_{t_{n+1}} - \mathbf{W}_{t_n}$, and θ are $\boldsymbol{\beta}$ are defined in (3-4) and (5-3) as integrals of increments of the Wiener process over the nth subinterval. Scheme (T) does not give second-order accuracy when applied to systems [4]. The precise statements of the conditions under which these orders of convergence are proved are contained in Theorem 1 of Section 3 and Theorem 3 of Section 5, respectively.

Our analysis is based on Taylor expansion of the solution, followed by derivation of an approximation formula whose Taylor expansion coincides to some order with that of the solution. This is similar to the method used by Chorin [6] in the approximation of Wiener integrals.

In practice, a stochastic scheme with high accuracy would still be less competitive without a substantial reduction of statistical errors in evaluating expectations of functions of sample solutions. We discuss, in Section 6, a variance reduction technique, suggested by Chorin [5], based on Hermite polynomial expansions, and its application to stochastic differential equations. To effect the reduction, we make use of the nonanticipating property of the solutions and apply Chorin's technique to the successive differences of the functions concerned at each time step. In Section 7 we give two two-dimensional numerical examples.

The difficulty in solving the stochastic differential equation (1-1) accurately arises from the nondifferentiability of the Wiener process $\mathbf{W}_t$. To take a closer look at this difficulty, we define the variable

$$\mathbf{Y}(t) = \mathbf{X}(t) - v\mathbf{W}_t, \qquad 0 \leqslant t \leqslant T.$$

Equation (1-1) then reduces to an infinite set of ordinary differential equations,

$$\text{(1-2)}\qquad \frac{d\mathbf{Y}}{dt} = \mathbf{f}(\mathbf{Y} + v\mathbf{W}_t), \qquad 0 < t < T,$$

for almost every path of the Wiener process $\mathbf{W}_t$. The theory of ordinary differential equations assures the existence of the solutions $\mathbf{Y}(t)$ of these equations, which are only once differentiable as functions of t. However, since the error estimates of high-order accuracy methods involve high-order derivatives of $\mathbf{Y}(t)$, it is not clear how one is able to obtain methods with high-order accuracy for solving (1-2) or (1-1).

Historically, let us recall some numerical methods for solving the stochastic differential equation (1-1). The most commonly used methods are splitting schemes (see Chorin [7], [8], Franklin [12]). For these schemes, at each time step one approximates for each sample path of the Wiener process the differential equation

$$d\mathbf{X} = \mathbf{f}(\mathbf{X})\,dt \tag{1-3}$$

by a method for solving ordinary differential equations, and then one adds to the approximate solution an independent increment of the Wiener process $\nu\mathbf{W}_t$. The simplest example of a splitting scheme is Euler's method which is given by

$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + h\mathbf{f}(\mathbf{X}^{(n)}) + \nu\Delta\mathbf{W}^{(n)}. \tag{E}$$

Another example of a splitting scheme, based on the mid-point rule, is

$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + h\mathbf{f}\big(\mathbf{X}^{(n)} + \tfrac{1}{2}h\mathbf{f}(\mathbf{X}^{(n)})\big) + \nu\Delta\mathbf{W}^{(n)}. \tag{1-4}$$

These splitting schemes are only first-order accurate in the L_2 sense, no matter how accurately one solves the nonrandom part (1-3). This will be clear from the analysis for the scheme (1-4) in Section 2. To obtain more accurate numerical schemes, McShane [16] has extended the idea of Runge-Kutta methods for ordinary differential equations to stochastic differential equations. For (1-1), he proposed (see also Fahrmeir [10])

$$\mathbf{Q}^{(n)} = \mathbf{X}^{(n)} + h\mathbf{f}(\mathbf{X}^{(n)}) + \nu\Delta\mathbf{W}^{(n)},$$
$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + \nu\Delta\mathbf{W}^{(n)} + \tfrac{1}{2}h\big[\mathbf{f}(\mathbf{X}^{(n)}) + \mathbf{f}(\mathbf{Q}^{(n)})\big].$$

However, this scheme has the same order of accuracy as the splitting scheme mentioned above. In fact, based on a one-step error analysis, Rümelin [23] has shown that, if at each time step only the information $\Delta\mathbf{W}^{(n)}$ is used, then for a wide class of Runge-Kutta methods one can have at best a first-order accuracy. In this paper we do not discuss the Taylor series method, referring instead to Rao et al. [22], Mil'shtein [18], [19], Platen [20] and Platen and Wagner [21] for its development. In most of these references, the authors work on general Ito stochastic differential equations. In principle, by using an analysis paralleling that of Section 2, we can derive methods of Taylor series type with arbitrary order of accuracy for Eq. (1-1).

The paper is organized as follows. In Section 2 we analyze a splitting scheme based on the mid-point rule. In Section 3 we derive the second-order Runge-Kutta method (T); and in Section 4 we prove its convergence. In Section 5 we prove the convergence of the method (S). Section 6 is devoted to the study of variance reduction. Finally, we present numerical examples in Section 7.

2. Analysis of a Splitting Scheme Based on the Mid-Point Rule. For $\nu = 0$, the scheme (1-4) is the Runge-Kutta method based on the mid-point rule for the equation (1-1) with $\nu = 0$; it has second-order accuracy (see Gear [13]). However, in

this section, we show that if $\nu \neq 0$, the scheme (1-4) is not second-order in the L_p sense for any $p \geqslant 2$ for the stochastic differential equation (1-1). Without loss of generality, we may assume that $\nu = 1$ and consider the stochastic differential equation

$$dX = f(X)\,dt + dW_t, \qquad 0 < t < T; \tag{2-1}$$

and the splitting scheme

$$\mathrm{X}^{(n+1)} = \mathrm{X}^{(n)} + \Delta W^{(n)} + hf\big(\mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)})\big). \tag{2-2}$$

In analogy with the analysis for ordinary differential equations, we analyze the local truncation error D_n of the above scheme, which is defined by the equation

$$X(t_{n+1}) = X(t_n) + \Delta W^{(n)} + hf\big(X(t_n) + \tfrac{1}{2}hf(X(t_n))\big) - D_n. \tag{2-3}$$

For simplicity, we denote by ΔW_t the increment $W_t - W_{t_n}$, and we further define, for each subinterval $[t_n, t_{n+1}]$, the random variable

$$Y(t) = X(t) - \Delta W_t, \qquad t_n \leqslant t \leqslant t_{n+1} = t_n + h. \tag{2-4}$$

We note that $Y(t)$ is defined only for $t_n \leqslant t \leqslant t_n + h$ and that

$$Y(t_n) = X(t_n). \tag{2-5}$$

Substituting the definition in (2-4) into (2-1) and (2-3), we obtain

$$\frac{dY}{dt} = f(Y + \Delta W_t), \qquad t_n < t < t_{n+1} = t_n + h, \tag{2-6}$$

and in view of (2-5),

$$-D_n = Y(t_{n+1}) - Y(t_n) - hf\big(Y(t_n) + \tfrac{1}{2}hf(Y(t_n))\big). \tag{2-7}$$

For convenience of analysis we rewrite $-D_n$ in integral form. Integrating Eq. (2-6) from t_n to $t_n + h$ and substituting into (2-7), we obtain

$$-D_n = \int_{t_n}^{t_n+h} \big[f\big(Y(s) + \Delta W_s\big) - f\big(Y(t_n) + \tfrac{1}{2}hf(Y(t_n))\big)\big]\,ds. \tag{2-8}$$

With D_n in this form, further analysis can be made because of the differentiability of the function f. In the subsequent discussion, we will analyze D_n in the L_2 sense, for the reason stated in the introduction. For this, we make the following assumption:

$$\frac{\partial^\mu}{\partial x^\mu} f(x) \text{ are bounded}, \qquad 0 \leqslant \mu \leqslant 5. \tag{2-9}$$

As indicated in [23], this boundedness assumption is not a serious restriction, since actual computation requires weaker conditions on f. In the following, $O(h^p)$ will be employed to denote a stochastic quantity of order p in the L_2 sense. In this connection, we can easily see that (i) $(\Delta W_s)^p$ is $O(h^{0.5p})$, and (ii) $(Y(s) - Y(t_n))^p$ is $O(h^p)$; indeed,

$$Y(s) - Y(t_n) = (s - t_n) f_x(Y(t_n)) + O(h^{1.5}).$$

This suggests that we write the first integrand in (2-8) in the form

$$f\big(Y(s) + \Delta W_s\big) = f\big(Y(t_n) + \big[Y(s) - Y(t_n) + \Delta W_s\big]\big).$$

Expanding the right-hand side in a Taylor series about $Y(t_n)$, and taking into consideration (i) and (ii), we obtain

$$\begin{aligned} f(Y(s)+\Delta W_s) = {} & f(Y(t_n)) + f_x(Y(t_n))(Y(s)-Y(t_n)+\Delta W_s) \\ & + \tfrac{1}{2}f_{xx}(Y(t_n))(2(Y(s)-Y(t_n))\Delta W_s + \Delta W_s^2) \\ & + \tfrac{1}{6}f_{xxx}(Y(t_n))\Delta W_s^3 + O(h^2). \end{aligned} \tag{2-10}$$

We further expand the second integrand in (2-8) in a Taylor series about $Y(t_n)$ and obtain

$$f(Y(t_n) + \tfrac{1}{2}hf(X(t_n))) = f(Y(t_n)) + \tfrac{1}{2}hf_x(Y(t_n))f(Y(t_n)) + O(h^2). \tag{2-11}$$

Substituting (2-10) and (2-11) into D_n of (2-8), we can, after some cancellation, write D_n in increasing powers of ΔW_s:

$$-D_n = f_x(Y(t_n))\int_{t_n}^{t_n+h} \Delta W_s\, ds + \tfrac{1}{2}f_{xx}(Y(t_n))\int_{t_n}^{t_n+h} \Delta W_s^2\, ds - R_n,$$

or, in view of (2-5),

$$-D_n = f_x(X(t_n))\int_{t_n}^{t_n+h} \Delta W_s\, ds + \tfrac{1}{2}f_{xx}(X(t_n))\int_{t_n}^{t_n+h} \Delta W_s^2\, ds - R_n, \tag{2-12}$$

where we keep in $-D_n$ only the two terms of the expansion (2-11) of leading order in ΔW_s, and group all the other terms in the remainder

$$\begin{aligned} -R_n = {} & f_x(Y(t_n))\int_{t_n}^{t_n+h} [(Y(s)-Y(t_n)-\tfrac{1}{2}hf(Y(t_n)))]\, ds \\ & + f_{xx}(Y(t_n))\int_{t_n}^{t_n+h} (Y(s)-Y(t_n))\Delta W_s\, ds \\ & + \tfrac{1}{6}f_{xxx}(Y(t_n))\int_{t_n}^{t_n+h} \Delta W_s^3\, ds + O(h^3). \end{aligned} \tag{2-13}$$

We now want to show that the remainder $-R_n$ is $O(h^{2.5})$. Recalling (i) and (ii), we can see that the second and third terms in the right-hand side of (2-13) are $O(h^{2.5})$. That the first term is $O(h^{2.5})$ follows from the following calculations:

$$\begin{aligned} & \int_{t_n}^{t_n+h} [Y(s)-Y(t_n)-\tfrac{1}{2}hf(Y(t_n))]\, ds \\ & \quad = \int_{t_n}^{t_n+h}\int_{t_n}^{s} [f(Y(r)+\Delta W_r) - f(Y(t_n))]\, dr\, ds \\ & \quad = \int_{t_n}^{t_n+h}\int_{t_n}^{s} [f_x(Y(t_n)\Delta W_r) + O(h)]\, dr\, ds \\ & \quad = f_x(Y(t_n))\int_{t_n}^{t_n+h}\int_{t_n}^{s} \Delta W_r\, dr\, ds + O(h^3). \end{aligned}$$

With this expression, (2-5) and (ii), $-R_n$ can be written in terms of increments of W_t:

$$\begin{aligned} -R_n = {} & f_x^2(X(t_n))\int_{t_n}^{t_n+h}\int_{t_n}^{s} \Delta W_r\, dr\, ds \\ & + f(X(t_n))f_{xx}(X(t_n))\int_{t_n}^{t_n+h} (s-t_n)\Delta W_s\, ds \\ & + \tfrac{1}{6}f_{xxx}(X(t_n))\int_{t_n}^{t_n+h} \Delta W_s^3\, ds + O(h^3). \end{aligned} \tag{2-14}$$

For the sake of brevity, we introduce the random variables γ, δ and τ by

$$h^{2.5}\gamma \equiv \int_{t_n}^{t_n+h} (s - t_n)\Delta W_s\, ds, \quad h^{2.5}\delta \equiv \int_{t_n}^{t_n+h} \int_{t_n}^{s} \Delta W_r\, dr\, ds,$$
$$h^{2.5}\tau \equiv \int_{t_n}^{t_n+h} \Delta W_s^3\, ds, \tag{2-15}$$

suppressing time-dependence for convenience. From these definitions it is clear that the variables γ,δ, τ are all of order zero in the L_2 sense. Then the expression (2-14) can be abbreviated as

$$-R_n = h^{2.5}\{ f_x^2\delta + ff_{xx}\gamma + \tfrac{1}{6}f_{xxx}\tau \} + O(h^3), \tag{2-16}$$

where all the functions are evaluated at $X(t_n)$. Thus, together with (2-12), we see that D_n^2 is $O(h^3)$, and therefore (2-2) is at best of order 1 in the L_2 sense, and hence in the L_p sense for $p > 2$, by Liapunov's inequality [9]. Indeed, Scheme (2-2) is a first-order method, as can be shown [4]. The expressions of $-D_n$ in (2-12) and $-R_n$ in in (2-14) are illuminating in the sense that, for each term therein, the integral part (of increments of W_t) is independent of the preceding product of functions evaluated at $X(t_n)$. This is a consequence of the nonanticipating property of the solutions of stochastic differential equations [1].

Remark. In the above discussion we often encountered expressions of half-integral order, due to the appearance of ΔW_s to an odd power. Recalling the nonanticipating property of the solutions of stochastic differential equations, we conclude that the expectations of the leading terms in these expressions vanish. We will use this fact repeatedly in the later development.

3. A Second-Order Runge-Kutta Method. We remarked in the introduction that a splitting scheme can at best have first-order accuracy in the L_2 sense. However, in this section we will show that by interlacing the function f and the Wiener process, it is possible to get Runge-Kutta methods of higher accuracy. As a preliminary step, the analysis of the splitting scheme (2-2) with the local truncation error D_n in (2-12) suggests that we consider the following Taylor series method:

$$\mathrm{Q}^{(n)} = \mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)}),$$
$$\mathrm{X}^{(n+1)} = \mathrm{X}^{(n)} + \Delta W^{(n)} + hf(\mathrm{Q}^{(n)}) + f_x(\mathrm{X}^{(n)})\int_{t_n}^{t_n+h} \Delta W_s\, ds + \tfrac{1}{2}f_{xx}(\mathrm{X}^{(n)})\int_{t_n}^{t_n+h} \Delta W_s^2\, ds. \tag{3-1}$$

The local truncation error of the scheme is given by R_n in (2-16), i.e., the exact solution $X = X(t)$ of Eq. (2-1) satisfies

$$X(t_{n+1}) = X(t_n) + \Delta W^{(n)} + hf(Q(t_n)) + f_x(X(t_n))\int_{t_n}^{t_n+h} \Delta W_s\, ds + \tfrac{1}{2}f_{xx}(X(t_n))\int_{t_n}^{t_n+h} \Delta W_s^2\, ds - R_n, \tag{3-2}$$

where we define

$$Q(t_n) = X(t_n) + \tfrac{1}{2}hf(X(t_n)). \tag{3-3}$$

Since R_n is of order 2.5 in the L_2, and thus in the L_1 sense, we would expect that Scheme (3-1) has order 1.5 in the L_1 sense due to the accumulation of the local truncation errors. However, an L_2 analysis shows that the scheme considered is in

fact a second-order method. Nevertheless, our analysis will not be made directly on the scheme (3-1). This scheme is an intermediate step which leads to a more satisfactory method of Runge-Kutta type. Before we go further, let us define, for the sake of brevity, random variables β and θ by

$$h^{1.5}\beta \equiv \int_{t_n}^{t_n+h} \Delta W_s\, ds, \qquad h^2\theta \equiv \int_{t_n}^{t_n+h} \Delta W_s^2\, ds, \tag{3-4}$$

where, again, the time-dependence is suppressed. Clearly, β and θ are all of order zero in the L_2 sense, and (3-1) can be rewritten as

$$\begin{aligned} \mathrm{Q}^{(n)} &= \mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)}), \\ \mathrm{X}^{(n+1)} &= \mathrm{X}^{(n)} + \Delta W^{(n)} + hf(\mathrm{Q}^{(n)}) + h^{1.5}\beta f_x(\mathrm{X}^{(n)}) + \tfrac{1}{2}h^2\theta f_{xx}(\mathrm{X}^{(n)}), \end{aligned} \tag{3-5}$$

which has a more convenient form to develop into a Runge-Kutta method. First, we add a term involving β to $\mathrm{Q}^{(n)}$ so that the first derivative term in $\mathrm{X}^{(n+1)}$ will appear implicitly. Observe that

$$hf\big(\mathrm{Q}^{(n)} + \sqrt{h}\,\beta\big) = hf(\mathrm{Q}^{(n)}) + h^{1.5}\beta f_x(\mathrm{X}^{(n)}) + \tfrac{1}{2}h^2\beta^2 f_{xx}(\mathrm{X}^{(n)}) + O(h^{2.5}),$$

which leads us to consider the following scheme:

$$\begin{aligned} \mathrm{Q}'^{(n)} &= \mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)}) + \sqrt{h}\,\beta, \\ \mathrm{X}^{(n+1)} &= \mathrm{X}^{(n)} + \Delta W^{(n)} + hf(\mathrm{Q}'^{(n)}) + \tfrac{1}{2}h^2(\theta - \beta^2) f_{xx}(\mathrm{X}^{(n)}). \end{aligned} \tag{3-6}$$

Its local truncation error V_n is defined by

$$X(t_{n+1}) = X(t_n) + \Delta W^{(n)} + hf\big(Q'(t_n)\big) + \tfrac{1}{2}h^2(\theta - \beta^2) f_{xx}\big(X(t_n)\big) - V_n, \tag{3-7}$$

where

$$Q'(t_n) = Q(t_n) + \sqrt{h}\,\beta = X(t_n) + \tfrac{1}{2}hf\big(X(t_n)\big) + \sqrt{h}\,\beta. \tag{3-8}$$

Here we have been careful to make the local truncation error V_n of Scheme (3-6) have the same order (in the L_2 sense) as R_n of Scheme (3-1) (or (3-5)). This can be seen by analyzing V_n further. As a starting point for seeing that V_n and R_n are of the same order, we carry out the Taylor expansion:

$$\begin{aligned} hf\big(Q'(t_n)\big) &= hf\big(Q(t_n) + \sqrt{h}\,\beta\big) \\ &= hf\big(Q(t_n)\big) + h^{1.5}\beta f_x\big(Q(t_n)\big) + \tfrac{1}{2}h^2\beta^2 f_{xx}\big(Q(t_n)\big) \\ &\quad + \tfrac{1}{6}h^{2.5}\beta^3 f_{xxx}\big(Q(t_n)\big) + O(h^3). \end{aligned} \tag{3-9}$$

Recall the definition of $Q(t_n)$ in (3-3). Each term on the right-hand side of the above equation is then expanded in a Taylor series about $X(t_n)$, and this gives

$$\begin{aligned} hf\big(Q'(t_n)\big) &= hf\big(Q(t_n)\big) + h^{1.5}\beta f_x\big(X(t_n)\big) \\ &\quad + \tfrac{1}{2}h^{2.5}\beta f\big(X(t_n)\big) f_{xx}\big(X(t_n)\big) + \tfrac{1}{2}h^2\beta^2 f_{xx}\big(X(t_n)\big) \\ &\quad + \tfrac{1}{6}h^{2.5}\beta^3 f_{xxx}\big(X(t_n)\big) + O(h^3). \end{aligned} \tag{3-10}$$

Substituting this result into (3-7), we obtain, after some cancellations,

$$
\begin{aligned}
X(t_{n+1}) = X(t_n) &+ \Delta W^{(n)} + hf(Q(t_n)) \\
&+ h^{1.5}\beta f_x(X(t_n)) + \tfrac{1}{2}h^2\theta f_{xx}(X(t_n)) \\
&+ \tfrac{1}{2}h^{2.5}\beta f(X(t_n))f_{xx}(X(t_n)) \\
&+ \tfrac{1}{6}h^{2.5}\beta^3 f_{xxx}(X(t_n)) - V_n + O(h^3).
\end{aligned} \tag{3-11}
$$

Recalling the definitions of β and θ in (3-4) and comparing this expression with (3-2), we can relate V_n and R_n by the equation

$$
-R_n = -V_n + h^{2.5}\left(\tfrac{1}{2}\beta f f_{xx} + \tfrac{1}{6}\beta^3 f_{xxx}\right) + O(h^3), \tag{3-12}
$$

where all the functions are evaluated at $X(t_n)$. Since R_n is of order 2.5 in the L_2 sense, so is V_n. However, it is still not clear how one is able to derive a Runge-Kutta method from the scheme (3-6), because it contains a second derivative of f with a coefficient containing $\theta - \beta^2$. Fortunately, the Cauchy-Schwarz inequality yields the interesting relationship

$$
\left\{\int_{t_n}^{t_n+h} \Delta W_s\, ds\right\}^2 \leqslant h\int_{t_n}^{t_n+h} \Delta W_s^2\, ds, \tag{3-13}
$$

which implies that $\theta - \beta^2$ is nonnegative. Following this observation, (3-6) suggests the following Runge-Kutta method:

$$
\begin{aligned}
&\mathrm{P}^{(n)} = \sqrt{\theta - \beta^2}, \\
&\mathrm{Q}'^{(n)} = \mathrm{X}^{(n)} + \tfrac{1}{2}hf(\mathrm{X}^{(n)}) + \sqrt{h}\,\beta, \\
&\mathrm{X}^{(n+1)} = \mathrm{X}^{(n)} + \Delta W^{(n)} + \tfrac{1}{2}h\left[f(\mathrm{Q}'^{(n)} + \sqrt{h}\,\mathrm{P}^{(n)}) + f(\mathrm{Q}'^{(n)} - \sqrt{h}\,\mathrm{P}^{(n)})\right],
\end{aligned} \tag{3-14}
$$

with β and θ defined in (3-4). This scheme is obtained by a symmetry consideration so that we need to evaluate only one intermediate value, i.e., $\mathrm{Q}'^{(n)}$ at each time step. We now state the main result of this section.

THEOREM 1. *Let f be a smooth function satisfying the condition stated in* (2-9). *Then the above scheme is second-order in the L_2 sense, provided that the initial condition is imposed with second-order accuracy in the L_2 sense.*

Note that, in Scheme (3-14), if we replace β by $\nu\beta$, θ by $\nu^2\theta$ and $\mathrm{P}^{(n)}$ by $\nu\mathrm{P}^{(n)}$, then we obtain the corresponding scheme (T) for solving Eq. (1-1). As ν tends to zero, this scheme reduces to the ordinary mid-point Runge-Kutta method, as we would expect. Before we prove Theorem 1, we will analyze the local truncation error T_n of Scheme (3-14), which is defined by the equation

$$
\begin{aligned}
X(t_{n+1}) = X(t_n) &+ \Delta W^{(n)} \\
&+ \tfrac{1}{2}h\left[f\left(Q'(t_n) + \sqrt{h}\,\mathrm{P}^{(n)}\right) + f\left(Q'(t_n) - \sqrt{h}\,\mathrm{P}^{(n)}\right)\right] - T_n.
\end{aligned} \tag{3-15}
$$

We start by considering the Taylor expansions

$$
\begin{aligned}
&f\left(Q'(t_n) \pm \sqrt{h}\,\mathrm{P}^{(n)}\right) \\
&\quad = f\left(Q'(t_n)\right) \pm \sqrt{h}\,\mathrm{P}^{(n)} f_x\left(Q'(t_n)\right) \\
&\qquad + \tfrac{1}{2}h(\mathrm{P}^{(n)})^2 f_{xx}\left(Q'(t_n)\right) \pm \tfrac{1}{6}h^{3/2}(\mathrm{P}^{(n)})^3 f_{xxx}\left(Q'(t_n)\right) + O(h^2).
\end{aligned} \tag{3-16}
$$

Recall the definition of $Q'(t_n)$. Further manipulation yields

$$\tfrac{1}{2}h\big[f\big(Q'(t_n)+\sqrt{h}\,\mathrm{P}^{(n)}\big)+f\big(Q'(t_n)-\sqrt{h}\,\mathrm{P}^{(n)}\big)\big]$$
$$=hf\big(Q'(t_n)\big)+\tfrac{1}{2}h^2(\mathrm{P}^{(n)})^2f_{xx}\big(X(t_n)\big)+\tfrac{1}{2}h^{2.5}\beta(\mathrm{P}^{(n)})^2f_{xxx}\big(X(t_n)\big)+O(h^3).$$

Substituting the above result into (3-15), with $\mathrm{P}^{(n)}$ defined in (3-14), we obtain

$$X(t_{n+1})=X(t_n)+\Delta W^{(n)}+hf\big(Q'(t_n)\big)$$
$$+\tfrac{1}{2}h^2(\theta-\beta^2)f_{xx}\big(X(t_n)\big)+\tfrac{1}{2}h^{2.5}\beta(\mathrm{P}^{(n)})^2f_{xxx}\big(X(t_n)\big)-T_n+O(h^3).$$

By comparing this expression with (3-7) we can relate T_n and V_n by the equation

$$-V_n=-T_n+\tfrac{1}{2}h^{2.5}(\theta-\beta^2)f_{xxx}+O(h^3). \tag{3-17}$$

As before, f_{xxx} is evaluated at $X(t_n)$. Now we are ready to write down explicitly the local truncation error T_n of Scheme (3-14), since we have the relationship (3-12) between R_n and V_n and the relationship (3-17) between V_n and T_n. The result is

$$-T_n=h^{2.5}\big\{\tfrac{1}{2}(2\gamma-\beta)ff_{xx}+\delta f_x^2+\tfrac{1}{6}(\tau-3\beta\theta+2\beta^3)f_{xxx}\big\}+h^3T_n', \tag{3.18}$$

where γ, δ and τ are the random variables defined in (2-15), where T_n' is of order zero and all the functions are evaluated at $X(t_n)$.

4. Convergence of the Second-Order Runge-Kutta Method (Scalar). In this section we prove Theorem 1. Let us write down the numerical scheme considered,

$$\mathrm{X}^{(n+1)}=\mathrm{X}^{(n)}+\Delta W^{(n)}+\tfrac{1}{2}h\big[f(\mathrm{Q}'^{(n)}+\sqrt{h}\,\mathrm{P}^{(n)})+f(\mathrm{Q}'^{(n)}-\sqrt{h}\,\mathrm{P}^{(n)})\big], \tag{4-1}$$

and the exact equation with the local truncation error,

$$X(t_{n+1})=X(t_n)+\Delta W^{(n)}$$
$$+\tfrac{1}{2}h\big[f\big(Q'(t_n)+\sqrt{h}\,\mathrm{P}^{(n)}\big)+f\big(Q'(t_n)-\sqrt{h}\,\mathrm{P}^{(n)}\big)\big]-T_n. \tag{4-2}$$

Let e_n denote $\mathrm{X}^{(n)}-X(t_n)$. As in the theory of numerical solution of ordinary differential equations, we subtract Eq. (4-2) from Eq. (4-1). This gives

$$e_{n+1}=e_n+\tfrac{1}{2}hv_n+T_n, \tag{4-3}$$

where we define

$$v_n\equiv v_{n,+}+v_{n,-}$$

and

$$v_{n,\pm}\equiv f\big(Q'(t_n)\pm\sqrt{h}\,\mathrm{P}^{(n)}\big)-f(\mathrm{Q}'^{(n)}\pm\sqrt{h}\,\mathrm{P}^{(n)}).$$

To make an L_2-norm analysis, we square both sides of (4-3) to get

$$e_{n+1}^2=e_n^2+he_nv_n+\tfrac{1}{4}h^2v_n^2+2e_nT_n+hv_nT_n+T_n^2. \tag{4-4}$$

We now estimate the expectations of each term on the right-hand side of the above equation. By assumption, f satisfies a Lipschitz condition,

$$|f(x)-f(y)|\leqslant L|x-y|,\qquad x,y\in R,$$

where $L>0$ is a constant. Consider $v_{n,+}$ and $v_{n,-}$. Recalling the definition (3-8) of $Q'(t_n)$ and applying the Lipschitz condition on f, we obtain

$$|v_{n,\pm}|\leqslant L\big|Q(t_n)-\mathrm{Q}^{(n)}\big|\leqslant L\big(1+\tfrac{1}{2}hL\big)|e_n|.$$

Therefore, the second term on the right-hand side of (4-4) can be estimated as

$$(4\text{-}5)\quad |E(he_nv_n)| \leqslant hE|e_nv_n| \leqslant hE\big(|e_n|\,|v_{n,+} + v_{n,-}|\big) \leqslant 2hL\big(1 + \tfrac{1}{2}hL\big)E\big(e_n^2\big).$$

The estimation of the third term is similar, and we have

$$(4\text{-}6)\qquad E\big(\tfrac{1}{4}h^2v_n^2\big) \leqslant \tfrac{1}{4}h^2E\big(v_n^2\big) \leqslant h^2L^2\big(1 + \tfrac{1}{2}hL\big)^2E\big(e_n^2\big).$$

Next comes the fourth term, where we need to take into account T_n given in (3-18), thus T_n':

$$(4\text{-}7)\qquad |E(2e_nT_n)| = 2h^3\big|E\big(e_nT_n'\big)\big| \leqslant \varepsilon_1 hLE\big(e_n^2\big) + \varepsilon_1^{-1}L^{-1}h^5E\big(T_n'^2\big),$$

where we employ the inequality $2ab \leqslant a^2 + b^2$ with $a = (\varepsilon_1 hL)^{1/2}e_n$ and $b = (\varepsilon_1 hL)^{-1/2}h^2T_n'$, and ε_1 is an appropriate positive number. A similar trick can also be applied to the fifth term and yields

$$(4\text{-}8)\qquad |E(hv_nT_n)| \leqslant \tfrac{1}{2}\varepsilon_2 hL\big(1 + \tfrac{1}{2}hL\big)^2E\big[e_n^2\big] + O(h^6),$$

where, again, ε_2 is an appropriate positive number. In this estimation we use the obvious fact that T_n^2 is $O(h^5)$; actually one can show [4]

$$(4\text{-}9)\qquad E\big(T_n^2\big) \leqslant E\big(G_n^2\big)h^5 + O(h^6),$$

where

$$G_n^2 \equiv \frac{1}{40}f^2f_{xx}^2 + \frac{3}{20}f_x^4 + \frac{11}{30340}f_{xxx}^2.$$

Finally we reach the stage of estimating the whole equation (4-4). By taking expectations on both sides of (4-4) and collecting the results from (4-5)–(4-9), we obtain

$$(4\text{-}10)\quad E\big(e_{n+1}^2\big) \leqslant B(h)E\big(e_n^2\big) + \big[E\big(G_n^2\big) + \tfrac{1}{2}\varepsilon_1^{-1}L^{-1}E\big(T_n'^2\big)\big]h^5 + O(h^6),$$

where

$$B(h) = 1 + (2 + \varepsilon_1 + \varepsilon_2)hL + (2 + \varepsilon_2/2)h^2L^2 + \big(1 + \tfrac{1}{8}\varepsilon_2\big)h^3L^3 + \tfrac{1}{4}h^4L^4.$$

To have a common bound for all time steps, for fixed $h_0 < 1$, we define $A \equiv \max_{n,h \leqslant h_0} E(G_n^2)$ and $B \equiv \max_{n,h \leqslant h_0} E(T_n'^2)$, and set $M = A + \tfrac{1}{2}\varepsilon_1^{-1}L^{-1}B$. Then the inequality (4-10) becomes

$$E\big(e_{n+1}^2\big) \leqslant e^{(2+\varepsilon)hL}E\big(e_n^2\big) + Mh^5 + O(h^6),$$

where we choose $\varepsilon \equiv \varepsilon_1 + \varepsilon_2$ so that $B(h) \leqslant e^{(2+\varepsilon)hL}$. This is a recursive relation that we often encounter in the theory of numerical solution of ordinary differential equations. An elementary calculation shows

$$(4\text{-}11)\qquad E\big(e_n^2\big) \leqslant \frac{e^{(2+\varepsilon)t_nL} - 1}{(2 + \varepsilon)L}Mh^4 + e^{(2+\varepsilon)t_nL}E\big(e_0^2\big) + O(h^5).$$

The right-hand side of this inequality is of order 4, provided that the initial condition is properly imposed. Suppose that $E(e_0^2) \leqslant C_0h^4$, where C_0 is a constant. Substituting this into the above equation and taking square roots, we complete the proof with $[E(e_n^2)]^{1/2} \leqslant Ch^2$, where

$$C = \sup_{h \leqslant h_0}\left\{\frac{M}{(2 + \varepsilon)L}\big(e^{(2+\varepsilon)TL} - 1\big) + C_0e^{(2+\varepsilon TL)} + O(h)\right\}^{1/2}. \quad \square$$

Remark. There are two reasons for introducing the two positive numbers ε_1 and ε_2: to keep track of the "interaction" between T_n and e_n (see (4-7)) or v_n (see (4-8)), and to balance the error contributions from the initial error and local truncation errors (see (4-11)) in the hope that the constant C can be minimized with suitable choice of ε, though we have not done so.

5. Runge-Kutta Methods of Order One-and-One-Half (Scalar & System). There are two main difficulties with Scheme (3-14): the first is that we do not have an efficient way to sample systematically the Gaussian variables β, $\Delta W^{(n)}$ and the non-Gaussian random variable θ (see [15]), and the second is that it will not be a second-order method when applied to a system.

To sample only Gaussian random variables, one must be content with schemes with less accuracy. In this section we provide such schemes, of order 1.5 in the L_2 sense. The main advantage with the schemes is that they will maintain the order of accuracy when extended to a system of stochastic differential equations. We consider the scheme (S) and prove the following theorem.

THEOREM 2. *Suppose that* $\mathbf{f}$ *has bounded partial derivatives up to fifth order. The scheme* (S) *is of order one-and-one-half in the* L_2 *sense, provided that the initial condition is imposed with accuracy of order one-and-one-half.*

Proof. First, we consider the scalar case. There is no substantial difference between this proof and that of Theorem 1. We need only to check whether the technique used in the latter can be applied to this case. The key point is to examine the local truncation error of Scheme (S) for the scalar case, which is defined by the equation

$$X(t_{n+1}) = X(t_n) + \Delta W^{(n)} + \tfrac{1}{3}h\left[f(Q(t_n)) + 2\cdot f(Q^*(t_n))\right] - T_n^*, \tag{5-1}$$

where we define

$$Q^*(t_n) = X(t_n) + \tfrac{1}{2}hf(X(t_n)) + \tfrac{3}{2}\sqrt{h}\,\beta.$$

To make an error analysis, we carry out the following Taylor expansion:

$$\begin{aligned} hf(Q^*(t_n)) &= hf\left(Q(t_n) + \tfrac{3}{2}\sqrt{h}\,\beta\right) \\ &= hf(Q(t_n)) + \tfrac{3}{2}h^{1.5}\beta f_x(X(t_n)) + \tfrac{9}{8}h^2\beta^2 f_{xx}(X(t_n)) + O(h^{2.5}) \\ &= hf(Q(t_n)) + \tfrac{3}{2}f_x(X(t_n))\int_{t_n}^{t_n+h}\Delta W_s\,ds + \tfrac{9}{8}h^2\beta^2 f_{xx}(X(t_n)) + O(h^{2.5}). \end{aligned}$$

Replacing $f(Q^*(t_n))$ in (5-1) by the above expression, we obtain

$$\begin{aligned} X(t_{n+1}) = X(t_n) &+ \Delta W^{(n)} + hf(Q(t_n)) \\ &+ f_x(X(t_n))\int_{t_n}^{t_n+h}\Delta W_s\,ds + \tfrac{3}{4}h^2\beta^2 f_{xx}(X(t_n)) - T_n^* + O(h^{2.5}). \end{aligned}$$

Comparing the above expression with (3-2) and recalling the definition of θ and that R_n is of order 2.5, we arrive at

$$T_n^* = \tfrac{1}{2}h^2 f_{xx}(X(t_n))\left(\theta - \tfrac{3}{2}\beta^2\right) + O(h^{2.5}).$$

One major fact about T_n^* is that its expectation is of order 3. The reason is that (i) the expectations of those terms of 2.5 are zero, and (ii) $E[(\beta^2)] = \frac{1}{3}$ and $E(\theta) = \frac{1}{2}$, which make the leading terms in T_n^* cancel each other after taking expectations.

With this fact in mind, the rest of the proof proceeds exactly as in the proof of Theorem 1. Note that $(T_n^*)^2$ is of order 4. The result is

$$E(e_n^2) \leqslant \frac{1}{120} \cdot \frac{e^{(2+\varepsilon)TL} - 1}{(2+\varepsilon)L} D^2 h^3 + e^{(2+\varepsilon)TL} E(e_0^2) + O(h^4),$$

where $D^2 \equiv \max_{n,h \leqslant h_0} E[f_{xx}^2(X(t_n))]$, and $\varepsilon > 0$ is defined in a similar way as in Theorem 1. □

System Case. Via an elaboration, one can show that for the system case the corresponding local truncation error for the scheme (S), in component form, is

$$T_n^{*i} = \tfrac{1}{2} h^2 f_{,jk}^i(\mathbf{X}(t_n))\left(\theta^{jk} - \tfrac{3}{2}\beta^j\beta^k\right) + O(h^{2.5}), \tag{5-2}$$

where a subscript with a comma denotes differentiation, the summation convention is assumed, and the random variables $\boldsymbol{\beta} \equiv \{\beta^j\}$ and θ^{jk} are defined as follows,

$$h^{1.5}\beta^j \equiv \int_{t_n}^{t_n+h} \Delta W_s^j \, ds, \qquad h^2\theta^{jk} \equiv \int_{t_n}^{t_n+h} \Delta W_s^j \Delta W_s^k \, ds, \tag{5-3}$$

in analogy with the definitions of β and θ in (3-4). Since the components of a Wiener process are independent, $E(T_n^{*i})$ is of order 3, due to the nonanticipating property. This is the key to the proof of Theorem 2 for both scalar equations and systems. Some work shows that the error is bounded by (see [4])

$$\begin{aligned} E(\|\mathbf{e}_n\|^2) \leqslant{}& \frac{1}{240} \cdot \frac{e^{(2\sqrt{d}+\varepsilon)TL} - 1}{(2\sqrt{d}+\varepsilon)L} d^2(d+1) D^2 h^3 \\ &+ e^{(2\sqrt{d}+\varepsilon)TL} E(\|\mathbf{e}_0\|^2) + O(h^4), \end{aligned} \tag{5-4}$$

where $D^2 \equiv \max_{n,h \leqslant h_0} E[\sum_{i,j,k} (f_{,jk}^i(\mathbf{X}(t_n)))^2]$. This completes the proof of Theorem 2. □

From the expression (5-4) we see that, if the initial error is sufficiently small, then for $h \approx 0.01$, Scheme (S) is practically of order 2 in the L_2 sense. We conclude this section by indicating that the general idea in designing a scheme of order one-and-one-half, like Scheme (S), is to consider the family

$$\begin{aligned} \mathbf{Q}^{(n)} &= \mathbf{X}^{(n)} + \tfrac{1}{2} h \mathbf{f}(\mathbf{X}^{(n)}) + k\sqrt{h}\,\boldsymbol{\beta}, \\ \mathbf{Q}^{*(n)} &= \mathbf{X}^{(n)} + \tfrac{1}{2} h \mathbf{f}(\mathbf{X}^{(n)}) + l\sqrt{h}\,\boldsymbol{\beta}, \\ \mathbf{X}^{(n+1)} &= \mathbf{X}^{(n)} + \Delta\mathbf{W}^{(n)} + h\left[a\mathbf{f}(\mathbf{Q}^{(n)}) + b\mathbf{f}(\mathbf{Q}^{*(n)})\right], \end{aligned} \tag{5-5}$$

where a, b, k, l are parameters to be determined. In a similar way as we did in the proof of Theorem 2, we find that the exact solution of the stochastic differential equation (1-1) satisfies

$$\begin{aligned} \mathbf{X}(t_{n+1}) = \mathbf{X}(t_n) &+ \Delta\mathbf{W}^{(n)} + h(a+b)\mathbf{f}(\mathbf{Q}(t_n)) \\ &+ h(a \cdot k + b \cdot l)\beta^j \mathbf{f}_{,j}(\mathbf{X}(t_n)) \\ &+ \tfrac{1}{2} h(a \cdot k^2 + b \cdot l^2)\beta^j\beta^k \mathbf{f}_{,jk}(\mathbf{X}(t_n)) - \mathbf{T}^{*(n)} + O(h^{2.5}). \end{aligned}$$

Now we make a choice so that a Taylor expansion of $\mathbf{X}(t_{n+1})$ matches the right-hand side up to order 1.5 and find

$$a + b = 1, \qquad a \cdot k + b \cdot l = 1. \tag{5-6}$$

Then the local truncation errors $\mathbf{T}_n^*$ of the schemes in (5-5), in component form, are

$$T_n^{*i} = \tfrac{1}{2}h^2 f_{,jk}^i(\mathbf{X}(t_n))\big(\theta^{jk} - (a \cdot k^2 + b \cdot l^2)\beta^j\beta^k\big) + O(h^{2.5}).$$

However, as follows from the proof of Theorem 2 (or 1), we may wish to minimize the contribution of the local truncation error T_n^{*i}. One way to achieve this is to choose the parameters so that the expectations of the leading terms of T_n^* are zero (e.g., in (5-2)). This leads to

$$a \cdot k^2 + b \cdot l^2 = \tfrac{3}{2}. \tag{5-7}$$

The case corresponding to Scheme (S) is $a = \frac{1}{3}$, $b = \frac{2}{3}$, $k = 0$, $l = \frac{3}{2}$, which is clearly a solution of Eqs. (5-6) and (5-7), but this choice is not essential.

6. Variance Reduction Using Hermite Polynomials. In this section we consider a variance-reduction technique for evaluating expectations of functions of solutions of stochastic differential equations. Intrinsically, the numerical evaluation of expectations involves a sampling process, i.e., Monte-Carlo computation. Being a finite process, Monte-Carlo computation creates statistical errors due to imperfect sampling. The errors depend heavily on how one choose the estimators.

Our goal is thus to construct estimators with a small variance. We start by considering Chorin's variance reduction technique for evaluating functionals of Gaussian random variables. This technique exploits specific properties of the Hermite polynomials. Then we show how to implement Chorin's technique for functions of solutions of stochastic differential equations.

Chorin's Estimator. Consider a random function $g(\xi) = g(\xi^1, \ldots, \xi^d)$, where $\xi = (\xi^1, \ldots, \xi^d)$ is an R^d-valued Gaussian random variable with distribution $N(0, I_d)$. The expectation of $g(\xi)$ is

$$E[g(\xi)] = E\big[g(\xi^1, \ldots, \xi^d)\big] = (2\pi)^{-d/2}\int g(\mathbf{u})e^{-\|u\|^2/2}\,d\mathbf{u},$$

where $\mathbf{u} = (u^1, \ldots, u^d)$, $d\mathbf{u} = du^1 \cdots du^d$, and we recall that $\|\mathbf{u}\|$ is the 2-norm of $\mathbf{u}$ in the R^d space. The Gaussian random variable ξ can be readily sampled (see Section 7). The usual Monte-Carlo estimator for $E[g(\xi)]$ is given by

$$N^{-1}\sum_{j=1}^{N} g(\xi_j) = N^{-1}\sum_{j=1}^{N} g\big(\xi_j^1, \ldots, \xi_j^d\big),$$

where $\{\xi_j^k\}$ are drawn from the Gaussian distribution with mean 0 and variance 1. The standard deviation of this estimator, which yields the order of magnitude of the error, is

$$N^{-1/2}\big(E[g^2(\xi)] - [Eg(\xi)]^2\big)^{1/2}, \tag{6-1}$$

which is proportional to $N^{-1/2}$, thus may not be acceptable for reasonable size N. Hence, an estimator of $E[g(\xi)]$ with smaller standard deviation is needed to achieve more accuracy in Monte-Carlo computation.

Chorin [5] proposed a method to obtain an estimator for obtaining $E[g(\xi)]$ with a substantial reduction in standard deviation. The idea is to use a finite Hermite series for the goal function g to design an estimator of control variate type for $E[g(\xi)]$.

The set of Hermite polynomials

$$H_n(z) = \frac{(-1)^n}{\sqrt{n!}} e^{z^2/2} \frac{d^n}{dz^n} e^{-z^2/2}, \qquad n = 1, 2, \ldots,$$

form a family of orthonormal functions in the space $L_2(R)$ of square integrable functions defined on R with respect to the weight $e^{-z^2/2}/\sqrt{2\pi}$. That is,

$$(2\pi)^{-1/2} \int H_n(z) H_m(z) e^{-z^2/2}\, dz = \delta_{nm}.$$

In general, let $\mathbf{m} = (m^1, \ldots, m^d)$ with m^j nonnegative integers, and set $|\mathbf{m}| = m^1 + \cdots + m^d$. We define the product polynomials

$$H_{\mathbf{m}} = H_{(m^1, \ldots, m^d)}(\mathbf{u}) = H_{m^1}(u^1) \cdots H_{m^d}(u^d).$$

Then the family of functions

$$H_{\mathbf{m}}(\mathbf{u}) \cdot e^{-\|\mathbf{u}\|^2/2}, \qquad 0 \leqslant |\mathbf{m}| < \infty, \tag{6-2}$$

forms a complete orthonormal set in the space $L_2(R^d)$ of all square integrable functions defined on R^d with respect to the weight $(2\pi)^{-d/2} e^{-\|\mathbf{u}\|^2/2}$. For a more detailed analysis of the family of Hermite polynomials $H_{\mathbf{m}}$, see Chorin [8], Maltz and Hitzl [17]. Assuming that the function $g(\mathbf{u}) e^{-\|\mathbf{u}\|^2/2}$ is square integrable, we can expand it in terms of the orthonormal functions in (6-2):

$$g(\mathbf{u}) e^{-\|\mathbf{u}\|^2/2} = \sum_{\mathbf{m}} a_{\mathbf{m}} H_{\mathbf{m}}(\mathbf{u}) e^{-\|\mathbf{u}\|^2/2},$$

i.e.,

$$g(\mathbf{u}) = \sum_{\mathbf{m}} a_{\mathbf{m}} H_{\mathbf{m}}(\mathbf{u}),$$

where we have

$$a_{\mathbf{m}} = E\left[H_{\mathbf{m}}(\xi) g(\xi)\right] = (2\pi)^{-d/2} \int H_{\mathbf{m}}(\mathbf{u}) g(\mathbf{u}) e^{-\|\mathbf{u}\|^2/2}\, d\mathbf{u} \tag{6-3}$$

because of the orthonormality

$$E\left[H_{\mathbf{n}}(\xi) H_{\mathbf{m}}(\xi)\right] = E\left[H_{n^1, \ldots, n^d}(\xi) H_{m^1, \ldots, m^d}(\xi)\right] = \delta_{\mathbf{n},\mathbf{m}} = \delta_{n^1 m^1, \ldots, n^d m^d}.$$

We also notice that (i) $a_0 = E[g(\xi)]$ and (ii) $E[H_{\mathbf{m}}(\xi)] = 0$ if $m \neq 0$. Therefore,

$$E\left[g(\xi)\right] = b_0 + E\left\{ g(\xi) - \sum_{|\mathbf{m}| \leqslant |\mathbf{p}|} b_{\mathbf{m}} H_{\mathbf{m}}(\xi) \right\} \tag{6-4}$$

for any set of numbers $\{b_{\mathbf{m}}\}$ and $\mathbf{p}$. In actual computation, we will take $\{b_{\mathbf{m}}\}$ to be $\{a_{\mathbf{m}}\}$. The success of Chorin's variance reduction lies in the fact that this identity does not imply that the Monte-Carlo estimators on both sides will have the same standard deviations. Chorin's idea is to make a first sampling to determine the coefficients $b_{\mathbf{m}}$ in (6-4) according to the formula (6-3), then a second sampling to simulate the Gaussian variables that appear in the argument of g and the polynomials $H_{\mathbf{m}} = H_{m^1, \ldots, m^d}$ on the right-hand side of (6-4). Specifically, we have

$$a^*_{\mathbf{m}} = \frac{1}{N} \sum_{j=1}^{N} \left(H_{\mathbf{m}}(\xi_j) g(\xi_j) \right) \tag{6-5}$$

and

$$a_0^* + \frac{1}{N}\sum_j \left\{ g(\hat{\xi}_j) - \sum_{|\mathbf{m}|\leqslant|\mathbf{p}|} a_{\mathbf{m}}^* H_{\mathbf{m}}(\hat{\xi}^j) \right\}, \tag{6-6}$$

where $\xi_j = \{\xi_j^i\}$ and $\hat{\xi}_j = \{\hat{\xi}_j^i\}$ are two sets of independent samples drawn from the Gaussian distribution with mean 0 and variance 1. The formulas (6-5) and (6-6) are called *Chorin's estimator* for $E[g(\xi)]$. In order to see the standard deviation of Chorin's estimator, let us define the remainder

$$r_{\mathbf{p}}(\mathbf{u}) = g(\mathbf{u}) - \sum_{|\mathbf{m}|\leqslant|\mathbf{p}|} a_{\mathbf{m}} H_{\mathbf{m}}(\mathbf{u}). \tag{6-7}$$

Maltz and Hitzl [17] have shown that Chorin's estimator has the following standard deviation:

$$N^{-1/2}\left(E\|r_{\mathbf{p}}\|^2 + N^{-1}\sum_{|\mathbf{m}|\leqslant|\mathbf{p}|} \sigma_{\mathbf{m}}^2 \right)^{1/2}, \tag{6-8}$$

where $\sigma_{\mathbf{m}}^2$ is the variance of $a_{\mathbf{m}}^*$ in (6-5) with $N = 1$, i.e., the single sample variance for the Monte-Carlo estimate of $a_{\mathbf{m}}$. Note that $E\|r_{\mathbf{p}}\|^2$ and $\sigma_{\mathbf{m}}^2$ are complementary in the sense that as $|\mathbf{p}|$ increases the former goes up while the latter goes down. For each N, an optimum $\mathbf{p}$ could be determined. See Maltz and Hitzl [17] for a detailed discussion.

Partial Variance Reduction. Let ϕ be a sufficiently smooth function. We implement Chorin's variance reduction technique to evaluate the expectation $E[\phi(\mathbf{X}^{(n)})]$ where $\mathbf{X}^{(n)}$ is a numerical solution of the d-dimensional version of (2-1). Replace $\Delta\mathbf{W}^{(n)}$ by $\sqrt{h}\,\boldsymbol{\alpha}^{(n)}$; then the methods that we consider in this paper have the form

$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + \sqrt{h}\,\boldsymbol{\alpha}^{(n)} + h\boldsymbol{\Psi}^{(n)}. \tag{6-9}$$

We note that $\boldsymbol{\beta}^{(n)}$ is included in $\boldsymbol{\Psi}^{(n)}$ except in Euler's method, so that $\mathbf{X}^{(n)}$, thus $\phi(\mathbf{X}^{(n)})$, is a function of the n pairs of R^d-valued Gaussian random variables $\boldsymbol{\alpha}^{(k)}$, $\boldsymbol{\beta}^{(k)}$, $0 \leqslant k \leqslant n$, since we implement the scheme (6-9) $n+1$ times. That is, $\phi(\mathbf{X}^{(n)})$ is a function of $2\cdot(n+1)\cdot d$ (scalar) Gaussian random variables. Therefore, it would be too expensive, even if the variance technique considered in the previous section were applied only once to all these Gaussian variables to evaluate the expectation $[\phi(\mathbf{X}^{(n)})]$. Instead, we wish to do only partial variance reduction, i.e., to determine an expression for $E[\phi(\mathbf{X}^{(n+1)})]$ with certain distinguished Gaussian variables, and apply Chorin's variance reduction technique to them only. First, we observe that

$$\begin{aligned} \phi(\mathbf{X}^{(n+1)}) &= \phi(\mathbf{X}^{(n)} + \sqrt{h}\,\boldsymbol{\alpha}^{(n)} + h\boldsymbol{\Psi}^{(n)}) \\ &= \phi\big(\mathbf{X}^{(0)} + \sqrt{h}\,[\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(n)}] + h[\boldsymbol{\Psi}^{(0)} + \cdots + \boldsymbol{\Psi}^{(n)}]\big), \end{aligned} \tag{6-10}$$

from which we see that the accumulating random variable $\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(n)}$ plays a major role in determining $\phi(\mathbf{X}^{(n+1)})$, while the individual $\boldsymbol{\alpha}^{(k)}$, $0 \leqslant k \leqslant n$, play only a minor role. Hence, the obvious strategy is to apply Chorin's estimator to evaluate $E[\phi(\mathbf{X}^{(n)})]$ at each time step with respect to $\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(n)}$ only. The main drawbacks with this strategy are (i) that variance reduction is only done with respect to $(\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(n)})$ and (ii) that there is no connection between two successive evaluations $E[\phi(\mathbf{X}^{(n)})]$ and $E[\phi(\mathbf{X}^{(n+1)})]$. To improve this variance reduction

technique and 'link' $\{E[\phi(\mathbf{X}^{(n)})]\}$, we write

$$\phi(\mathbf{X}^{(n+1)}) = \left[\phi(\mathbf{X}^{(n+1)}) - \phi(\mathbf{X}^{(n)})\right] + \cdots + \left[\phi(\mathbf{X}^{(k+1)}) - \phi(\mathbf{X}^{(k)})\right] + \cdots + \phi(\mathbf{X}^{(0)}).$$

For each piece $\phi(\mathbf{X}^{(k+1)}) - \phi(\mathbf{X}^{(k)})$, we carry out the Taylor expansion of $\phi(\mathbf{X}^{(k+1)})$ about $\mathbf{X}^{(k)}$:

$$\phi(\mathbf{X}^{(k+1)}) - \phi(\mathbf{X}^{(k)}) = \phi_{,J}(\mathbf{X}^{(k)})[\sqrt{h}\,\boldsymbol{\alpha}^{(k)} + h\Psi^{(k)}]^J + \tfrac{1}{2}h\phi_{,Jl}(\mathbf{X}^{(k)})\alpha^{(k)J}\alpha^{(k)l} + O(h^{1.5}),$$

where $\boldsymbol{\alpha}^{(k)} = \{\alpha^{(k)J}\}$ is the random variable sampled at the kth time step. Moving the first term on the right-hand side to the left and denoting the resultant expression by $\Phi^{(k)}$, we have

$$\begin{aligned}\Phi^{(k)} &= \phi(\mathbf{X}^{(k+1)}) - \phi(\mathbf{X}^{(k)}) - \sqrt{h}\,\phi_{,J}(\mathbf{X}^{(k)})\alpha^{(k)J} \\ &= h\phi_{,J}(\mathbf{X}^{(k)})\Psi^{(k)j} + \tfrac{1}{2}h\phi_{,Jl}(\mathbf{X}^{(k)})\alpha^{(k)J}\alpha^{(k)l} + O(h^{1.5}).\end{aligned} \tag{6-11}$$

Note the independence between $\boldsymbol{\alpha}^{(k)}$ and $\mathbf{X}^{(k)}$. Taking expectations on both sides of the first equality in (6-11) and summing the results over k from 0 to n, we have

$$E\left[\phi(\mathbf{X}^{(n+1)})\right] = E\left[\boldsymbol{\Phi}^{(n)}\right] + \cdots + E\left[\boldsymbol{\Phi}^{(0)}\right] + E\left[\phi(\mathbf{X}^{(0)})\right], \tag{6-12}$$

which is equivalent to

$$E\left[\phi(\mathbf{X}^{(n+1)})\right] = E\left[\phi(\mathbf{X}^{(n)})\right] + E\left[\boldsymbol{\Phi}^{(n)}\right]. \tag{6-13}$$

Thus we obtain a recursive relation between $E[\phi(\mathbf{X}^{(n)})]$ and $E[\phi(\mathbf{X}^{(n+1)})]$. From the second equality of (6-11) we see that for each fixed k, $\boldsymbol{\alpha}^{(k)}$ plays a leading role in determining $\boldsymbol{\Phi}^{(k)}$. By a similar argument, following (6-10), one can easily see that $\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(k-1)}$ plays a major role in determining $\phi_{,J}(\mathbf{X}^{(k)})$ and $\phi_{,Jk}(\mathbf{X}^{(k)})$. Hence we have the following

Strategy (S): We evaluate the expectation $E[\phi(\mathbf{X}^{(n)})]$ by applying Chorin's estimator (6-5, 6) to evaluate $E[\boldsymbol{\Phi}^{(n)}]$ in (6-13) with respect to $\boldsymbol{\alpha}^{(n)}$ and $(\boldsymbol{\alpha}^{(0)} + \cdots + \boldsymbol{\alpha}^{(n)})/\sqrt{n}$ (normalized $N(0, I_d)$), where $\boldsymbol{\Phi}^{(n)}$ is computed according to the first equality in (6-11), and adding the result to the estimate of $E[\phi(\mathbf{X}^{(n)})]$ obtained from the previous time step.

To see how the standard deviation, at each time step, of the estimate in Strategy (S) will accumulate, and whether this accumulation will destroy the accuracy of the variance reduction, we need the following lemma.

LEMMA. *Let $z_1, z_2, \ldots, z_n$ be n random variables; then their variances and the variance of their sum satisfy the inequality $\sigma^2_{z_1+\cdots+z_n} \leqslant [\sigma_{z_1} + \cdots + \sigma_{z_n}]^2$. Hence, by the Cauchy-Schwarz inequality, we have $\sigma^2_{z_1+\cdots+z_n} \leqslant n[\sigma^2_{z_1} + \cdots + \sigma^2_{z_n}]$.*

From the second equality in (6-11) we may write $\Phi^{(k)} = h\mathrm{G}^{(k)}$ for each fixed k, where $\mathrm{G}^{(k)}$ is of order zero. Then from (6-8) we see that the standard deviation $\mathrm{SD}^{(k)}$ of Chorin's estimator for each $E[\Phi^{(k)}]$ is of order

$$h \cdot \mathrm{SD}^{(k)} = h \cdot \frac{1}{\sqrt{N}}\left[E\|r_{\mathbf{p}}\|^2 + \frac{1}{N}\sum_{|\mathbf{m}|\leqslant|\mathbf{p}|}\sigma^2_{\mathbf{m}}\right]^{1/2} \tag{6-14}$$

for some $\mathbf{p}$, where $r_{\mathbf{p}}$ is defined similar to (6-7), with $g = G^{(k)}$, and where we suppress the dependence of $r_{\mathbf{m}}$ on k. Let the maximum of (6-14) over k be $SD^{(k_0)}$ for some k_0; then by the lemma we have the bound $n \cdot hSD^{(k_0)} = t_n \cdot SD^{(k_0)}$ for the estimate in Strategy (S). Hence we have

THEOREM 3. *The standard deviation of the estimator in Strategy* (S) *for evaluating* $E[\phi(\mathbf{X}^{(n)})]$ *with N samplings is of the form* (6-8), *which is proportional to* t_n *at the nth step, i.e., the piecewise application of Chorin's variance reduction technique to each summand in* (6-12), *produces a standard deviation as in* (6-8).

7. Numerical Implementation. In order to compare their accuracy, we implement the schemes (E) and (S) given in the introduction. We present two two-dimensional examples: one is a linear equation, and the other is nonlinear. Recall that we assumed boundedness on $\mathbf{f}$ and its first few partial derivatives in Theorems 1 and 2. However, it is clear that these theorems still hold if $\mathbf{f}$ satisfies a Lipschitz condition and all the expectations appearing in the proofs are uniformly bounded in the L_2 sense. Our first test equation satisfies this less restricted condition, while it is also interesting to observe the numerical results for the second test equation.

As in (6-9), for each time step, we set $\Delta\mathbf{W} = \sqrt{h}\,\boldsymbol{\alpha}$. To simulate the Gaussian random variables $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in Euler's method and Scheme (S), we write

$$\boldsymbol{\alpha} = \boldsymbol{\xi}, \qquad \boldsymbol{\beta} = \tfrac{1}{2}\boldsymbol{\xi} + \frac{\sqrt{3}}{6}\boldsymbol{\eta},$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are two independent R^d-valued Gaussian variables with distribution $N(0, I_d)$. These expressions give the correct correlation between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Then $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are sampled according to the Box-Muller formula

$$\xi^l = \cos(2\pi u^l)\left[-2\log(v^l)\right]^{1/2}, \qquad \eta^l = \sin(2\pi u^l)\left[-2\log(v^l)\right]^{1/2},$$

where $\mathbf{u}$ and $\mathbf{v}$ are two independent R^d-valued uniform distributions over $[0,1]^d$. For different time steps, we sample independent pairs $\{\mathbf{u}, \mathbf{v}\}$.

Our first computational example is the 2×2 system of linear equations

$$dX^1 = -X^2\,dt + dW_t^1, \qquad dX^2 = -X^1\,dt + dW_t^2$$

with zero initial data $X^1(0) = X^2(0) = 0$. Adding these equations together, we find upon integration

$$X^1(t) + X^2(t) = \int_0^t e^{-(t-s)}\,d\left(W_s^1 + W_s^2\right)$$

which, for each fixed t, is a Gaussian random variable with mean 0 and variance $1 - \exp(-2t)$. In this example, we evaluate numerically

$$E\left[\cos\left(X^1(t) + X^2(t)\right)\right] = \exp\left(\tfrac{1}{2}(1 - e^{-2t})\right).$$

The second computational example is the 2×2 system of nonlinear equations

$$dX^1 = \tfrac{1}{2}e^{-(X^1+X^2)}\,dt + dW_t^1, \qquad dX^2 = \tfrac{1}{2}e^{-(X^1+X^2)}\,dt + dW_t^2$$

with the zero initial data $X^1(0) = X^2(0) = 0$. By a calculation we obtain

$$e^{(X^1(t)+X^2(t))} = e^{(W_t^1+W_t^2)}\left(1 + \int_0^t e^{-(W_s^1+W_s^2)}\,ds\right).$$

In this example, we evaluate numerically

$$E\left[e^{(X^1(t)+X^2(t))}\right] = 2e^t - 1.$$

For each scheme we compute the expectations in two ways: (i) the usual Monte-Carlo estimator, and (ii) Chorin's estimator in Strategy (S) of Section 6. The errors depend on the time stepsize Δt and the number of simulations, N, for which we use 10000.

TABLES 7.1–7.3

Numerical results for $E[\cos(X^1(t) + X^2(t))]$, *errors/standard deviations are in exponential form.*

TABLE 7.1

Example 1: t = 0.2 N = 10,000 Exact Value = 0.8480				
Δ t	Euler's Method		Scheme (S)	
0.2000	-2.85-2/2.33-3	-2.92-2/2.16-4	2.77-3/1.96-3	2.31-3/1.67-4
0.1000	-1.50-2/2.14-3	-1.36-2/5.15-4	-7.32-4/1.97-3	5.26-4/3.86-4
0.0500	-8.46-3/2.09-3	-6.66-3/2.21-4	-1.70-3/2.01-3	3.28-5/1.89-4
0.0250	-3.87-3/2.05-3	-3.07-3/1.73-4	-5.74-4/2.01-3	1.83-4/1.80-4
0.0125	6.15-4/2.00-3	-1.56-3/2.25-4	2.22-3/1.98-3	4.46-5/2.64-4

TABLE 7.2

Example 1: t = 0.4 N = 10,000 Exact Value = 0.7593				
Δ t	Euler's method		Scheme (S)	
0.2000	-4.14-2/3.39-3	-3.88-2/1.55-3	6.83-4/1.18-2	2.88-3/2.24-3
0.1000	-2.10-2/3.22-3	-1.87-2/7.56-4	-1.85-3/2.96-3	3.05-4/1.17-3
0.0500	-9.16-3/3.14-3	-8.52-3/5.00-4	-6.40-5/3.04-3	4.80-4/5.48-4
0.0250	-7.54-4/3.03-3	-4.52-3/7.08-4	3.67-3/2.98-3	-6.60-5/7.07-4
0.0125	5.67-4/3.01-3	-2.22-3/7.62-4	2.71-3/2.98-3	-2.87-5/7.58-4

TABLE 7.3

Example 1: t = 0.8 N = 10,000 Exact value = 0.6710				
Δ t	Euler's Method		Scheme (S)	
0.2000	-4.40-2/4.28-3	-4.20-2/2.29-3	1.96-2/3.88-3	1.80-2/1.79-3
0.1000	-1.79-2/4.13-3	-1.88-2/1.07-3	2.29-3/3.94-4	1.25-3/1.21-3
0.0500	-4.57-3/3.91-3	-1.08-2/1.92-3	5.26-3/3.86-4	-9.18-4/1.89-3
0.0250	-1.00-3/3.91-3	-5.24-3/2.44-3	3.69-3/3.87-3	-3.57-4/2.11-3
0.0125	1.57-3/3.87-3	-2.36-3/2.20-3	3.94-3/3.85-3	5.74-5/2.18-3

TABLES 7.4–7.6

Numerical results for $E[\exp(X^1(t) + X^2(t))]$, *errors/standard deviations are in exponential form.*

TABLE 7.4

Example 2: t = 0.2 N = 10,000 T = 1.4428				
Δ t	Euler's method		Scheme (S)	
0.2000	1.68-1/1.30-2	1.58-1/3.24-4	-3.13-3/1.00-2	-1.08-2/9.49-4
0.1000	7.72-3/1.13-2	6.96-2/1.19-3	3.14-3/1.01-2	-3.63-3/8.35-4
0.0500	3.97-2/1.07-2	3.29-2/9.63-4	5.37-3/1.02-2	-1.18-3/2.20-4
0.0250	2.12-2/1.04-2	1.62-2/1.17-3	4.77-3/1.01-2	-2.38-4/1.12-3
0.0125	1.73-2/1.03-2	7.91-3/1.22-3	9.17-3/1.02-2	3.14-4/1.20-3

TABLE 7.5

Example 2: t = 0.4 N = 10,000 T = 1.9836				
Δ t	Euler's method		Scheme (S)	
0.2000	3.52-1/2.76-2	3.34-1/7.51-3	-7.94-3/2.18-2	-2.23-2/5.84-3
0.1000	1.67-1/2.44-2	1.45-1/3.90-3	1.17-2/1.96-2	-7.97-3/3.32-3
0.0500	7.80-2/2.26-2	6.99-2/4.71-3	6.79-3/2.16-2	-1.41-3/4.44-3
0.0250	5.12-2/2.25-2	3.39-2/4.85-3	1.66-2/2.20-2	-1.06-3/4.72-3
0.0125	6.97-3/2.21-2	1.61-2/4.79-3	9.46-3/2.19-2	-8.97-4/4.73-3

TABLE 7.6

Example 2: t = 0.8 N = 10,000 T = 3.4511				
Δ t	Euler's method		Scheme (S)	
0.2000	8.16-1/6.34-2	7.45-1/2.65-2	4.90-3/6.70-2	-5.24-2/2.12-2
0.1000	3.46-2/6.76-2	3.37-1/2.66-2	-1.04-3/6.33-2	1.17-2/2.41-2
0.0500	1.98-1/6.92-2	1.62-1/2.65-2	3.61-2/6.61-2	-4.15-3/2.47-2
0.0250	5.12-2/6.88-2	7.38-2/2.53-2	-2.54-2/6.74-2	-4.86-3/2.47-2
0.0125	-6.27-3/6.11-2	3.28-2/2.53-2	-4.37-2/6.44-2	2.01-3/2.50-2

The situation is shown in Tables 7.1–7.6. In each table, we list the results at time 0.2, 0.4, and 0.8. For each scheme, in the first subcolumn we list the errors plus the standard deviations of the computed values obtained by using the usual Monte-Carlo estimator, and in the second column errors are listed for Chorin's estimator. We see that Chorin's estimator shows Euler's method to be precisely a first-order method. For Scheme (S), Chorin's estimator shows that they are roughly second-order methods. This is not surprising, because we are evaluating expectations of functions, which should not exhibit the 'half' part of the order of the accuracy of the scheme, due to the nonanticipating property of the solutions.

Acknowledgment. The author wishes to thank his thesis adviser, Alexandre Chorin, and Professor Ole Hald for many helpful comments and discussions. This work was supported in part by the Applied Mathematics Sciences subprogram of the Office of Energy Research, U. S. Department of Energy under contract DE-AC03-76SF00098.

Lawrence Berkeley Laboratory
Berkeley, California 94720

Department of Mathematics
University of California
Berkeley, California 94720

1. L. ARNOLD, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, 1974.
2. S. W. BENSON, *The Foundations of Chemical Kinetics*, McGraw-Hill, New York, 1980.
3. S.CHANDRASEKHAR, "Stochastic problems in physics and astronomy," *Noise and Stochastic Processes* (N. Wax, ed.), Dover, New York, 1954.
4. C. C. CHANG, *Numerical Solution of Stochastic Differential Equations*, Ph.D. Dissertation, University of California, Berkeley, 1985.
5. A. J. CHORIN, "Hermite expansions in Monte-Carlo computation," *J. Comput. Phys.*, v. 8, 1971, pp. 472–482.
6. A. J. CHORIN, "Accurate evaluation of Wiener integrals," *Math. Comp.*, v. 27, 1973, pp. 1–15.
7. A. J. CHORIN, "Numerical study of slightly viscous flow," *J. Fluid Mech.*, v. 57, 1973, pp. 785–786.
8. A. J. CHORIN, *Lectures on Turbulence Theory*, Publish or Perish, Boston, 1975.
9. K. L. CHUNG, *A Course in Probability Theory*, 2nd ed., Academic Press, New York, 1974.
10. L. FAHRMEIR, "Approximation von stochastischen Differentialgleichungen auf Digital- und Hybridrechnern," *Computing*, v. 16, 1976, pp. 359–371.
11. A. L. FOGELSON, "A mathematical model and numerical method for studying platelet adhesion and aggregation during blood clotting," *J. Comput. Phys.*, v. 50, 1984, pp. 111–134.
12. J. N. FRANKLIN, "Difference methods for stochastic ordinary differential equations," *Math. Comp.*, v. 19, 1965, pp. 552–561.
13. C. W. GEAR, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, N. J., 1971.
14. A. H. JAZWINSKI, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
15. P. LEVY, *Wiener Random Function and Other Laplacian Random Functions*, Proc. 2nd Berkeley Sympos. Math. Stat. & Prob., vol. 1, 1966, pp. 171–187.
16. E. J. MCSHANE, *Stochastic Differential Equations and Models of Random Processes*, Proc. 6th Berkeley Sympos. Math. Stat & Prob., vol. 1, 1966, pp. 263–294.
17. F. H. MALTZ & D. L. HITZL, "Variance reduction in Monte-Carlo computations using multi-dimensional Hermite polynomials," *J. Comput. Phys.*, v. 32, 1979, pp. 345–376.
18. G. N. MIL'SHTEIN, "Approximate integration of stochastic differential equations," *Theory Probab. Appl.*, v. 19, 1974, pp. 557–562.
19. G. N. MIL'SHTEIN, "A method of second-order accuracy integration of stochastic differential equations," *Theory Probab. Appl.*, v. 23, 1978, pp. 396–401.
20. E. PLATEN, "Weak convergence of approximation of Ito integral equations," *Z. Angew. Math. Mech.*, v. 60, 1980, pp. 609–614.
21. E. PLATEN & W. WAGNER, "On a Taylor formula for a class of Ito process," *Probab. Math. Statist.*, v. 3, 1982, pp. 37–51.
22. N. J. RAO, J. D. BORWANHAR & D. RAMAKRISHNA, "Numerical solution of Ito integral equations," *SIAM J. Control*, v. 12, 1974, pp. 124–139.
23. W. RÜMELIN, "Numerical treatment of stochastic differential equations," *SIAM J. Numer. Anal.*, v. 19, 1982, pp. 604–613.