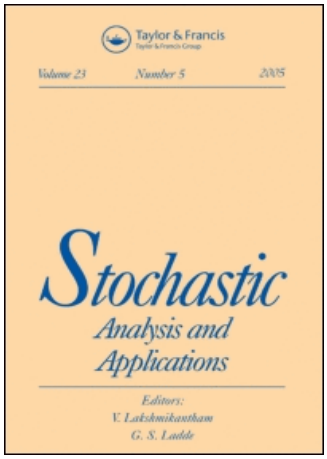

Stochastic Analysis and Applications
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597300

A second-order Monte Carlo method for the solution of the Ito stochastic differential equation

D. C. Haworth [a]; S. B. Pope [a]
[a] Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York

Online Publication Date: 01 January 1986
To cite this Article: Haworth, D. C. and Pope, S. B. (1986) 'A second-order Monte Carlo method for the solution of the Ito stochastic differential equation', Stochastic Analysis and Applications, 4:2, 151 - 186
To link to this article: DOI: 10.1080/07362998608809086
URL: http://dx.doi.org/10.1080/07362998608809086

STOCHASTIC ANALYSIS AND APPLICATIONS, 4(2), 151-186 (1986)

A SECOND-ORDER MONTE CARLO METHOD FOR THE SOLUTION OF THE ITO STOCHASTIC DIFFERENTIAL EQUATION

D.C. Haworth and S.B. Pope

Sibley School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York 14853

ABSTRACT

A difference approximation that is second-order accurate in the time step h is derived for the general Ito stochastic differential equation. The difference equation has the form of a second-order random walk in which the random terms are non-linear combinations of Gaussian random variables. For a wide class of problems, the transition pdf is joint-normal to second order in h; the technique then reduces to a Gaussian random walk, but its application is not limited to problems having a Gaussian solution. A large number of independent sample paths are generated in a Monte Carlo solution algorithm; any statistical function of the solution (e.g., moments or pdf's) can be estimated by ensemble averaging over these paths.

1. INTRODUCTION

The theory of stochastic differential equations and their use as models of physical phenomena originated with the development of a mathematical description of Brownian motion [1]. The theory blossomed with the introduction of the Ito and Stratonovich calculi [2,3] and these equations are now used in the modeling of a broad range of physical processes. The purpose of the present work is to develop a numerical algorithm applicable to an important class of stochastic ordinary differential equations. Specifically, we seek solutions to equations of the form

$$du_i = A_i(\underline{u},t)dt + B_{ij}(\underline{u},t)dW_j(t) \,, \tag{1}$$

for $t \geq t_0$, subject to the initial conditions

$$\underline{u}(t_0) = \underline{U}^0 \,. \tag{2}$$

 0736-2994/86/0402-0151$3.50/0

In Eq.(1), $\underline{u}(t)$ is a vector-valued random process, $A_i(\underline{u},t)$ and $B_{ij}(\underline{u},t)$ are functions of the random processes and of time, and $\underline{W}(t)$ is an isotropic Weiner process [4,5]: $d\underline{W}(t)$ has a joint-normal distribution with zero means and an isotropic covariance matrix

$$\langle dW_i(t)\rangle = 0 , \quad \langle dW_i(t)dW_j(t)\rangle = dt\delta_{ij} , \tag{3}$$

where δ_{ij} denotes the Kronecker delta. $\underline{U}^0$ of Eq.(2) may be a random vector. Here and in the remainder of the paper, Cartesian tensor notation is used where convenient.

Equation (1) is a general form of Ito's stochastic differential equation. Included in this form are coupled stochastic ordinary differential equations of arbitrary order where the randomness is expressed as a vector of independent Gaussian white noise processes, the derivative of the isotropic Weiner process. This type of equation has applications in a variety of fields including stochastic control theory [6], filtering theory [7], and fluid mechanics [8]. Of particular interest in applications is the Langevin equation, a special case of Eq.(1) in which the deterministic term is linear in the dependent variable and the coefficient of the random term is independent of the dependent variable.

We assume that the coefficient functions A_i and B_{ij} satisfy all necessary conditions for the existence of a unique diffusion process solution to Eq.(1) [4,9]. The coefficient functions must also satisfy the following differentiability requirements for the numerical method developed in this paper to be applicable: A_i and C_{ij} possess bounded continuous partial derivatives with respect to the components of $\underline{u}$ up to at least the second order and with respect to t up to at least the first order in the interval of interest, where C_{ij} is a symmetric positive semi-definite matrix,

$$C_{ij} \equiv B_{il}B_{jl} . \tag{4}$$

The numerical scheme can be simplified if, in addition to these constraints, C_{ij} satisfies the equation

$$C_{ij,l}C_{kl} + C_{jk,l}C_{il} + C_{ki,l}C_{jl} = 0 . \tag{5}$$

Necessary and sufficient conditions for the satisfaction of Eq.(5) are difficult to deduce and do not appear to have a simple interpretation in terms of the coefficients B_{ij}. However, we cite two important classes of equations for which the restriction is satisfied. First, it follows trivially if B_{ij} is a function of time only: $B_{ij}(\underline{u},t) = B_{ij}(t)$. In the second class of problems, the random processes u_i , $i = 1, \cdots, M$ have a natural ordering such that the coefficients $B_{ij}(\underline{u},t)$ satisfy $B_{ij}(\underline{u},t) = 0$ for $1 \le i,j \le m$, and $B_{ij}(\underline{u},t) = b^{(i)}(u_1, \cdots ,u_m,t)\delta_{ij}$ for $i, j > m$. That is, for $i > m$, the i^{th} component of the Weiner process (W_i) only affects the i^{th} component of the stochastic process (u_i); and, the coefficient $B_{ij} = b^{(i)}\delta_{ij}$ of the i^{th} process is independent of u_k, $k > m$. This includes generalized Langevin equations for vector-valued stochastic processes (see Haworth and Pope [10], for example). These are not the most general conditions for the satisfaction of Eq.(5). For example, any linear transformation of $\underline{u}$, $\tilde{u}_i = T_{ij}u_j$, in the second class of equations cited also suffices.

The differential notation of Eq.(1) is shorthand for the stochastic integral equation

$$u_i(t) = u_i(t_0) + \int_{t_0}^{t} A_i(\underline{u}(s),s)ds + \int_{t_0}^{t} B_{ij}(\underline{u}(s),s)dW_j(s) . \tag{6}$$

In Eq.(6), the meaning of the stochastic integral is ambiguous: we interpret the integral in the Ito sense [4,5,11]. In terms of the increment in $\underline{u}$ in an interval h,

$$\Delta\underline{u} \equiv \underline{u}(t+h) - \underline{u}(t) , \tag{7}$$

the solution of Eqs.(1) and (6) is then a diffusion process whose infinitesimal drift and diffusion parameters are

$$\lim_{h\to 0}\frac{1}{h}<\Delta u_i \mid \underline{u}(t)=\underline{U}> = A_i(\underline{U},t) , \tag{8}$$

$$\lim_{h\to 0}\frac{1}{h}<\Delta u_i \Delta u_j \mid \underline{u}(t)=\underline{U}> = C_{ij}(\underline{U},t) . \tag{9}$$

Results applicable to the alternative Stratonovich interpretation may be deduced by a transformation of the drift parameter [4,5,11].

Accurate and efficient solution methods for the Ito stochastic differential equation are required in practical applications. Our motivation for developing a new numerical algorithm for the solution of Eq.(1) is that a six-dimensional stochastic process whose evolution is governed by an equation of the Langevin type can be used to model the behavior of fluid particles in a turbulent flow [10,12]. Because of the large dimensionality and because the random terms are not necessarily small compared to the deterministic terms, most currently available solution methods are inadequate for our purposes. We were further compelled to develop a method that is second-order accurate in the computational time step h after tests using a first-order difference approximation were found to require an impractically small step size (see Section 4). For a difference approximation to a differential equation to have an error of $O(h^2)$ at a fixed time t, as sought here, the truncation error on a single time step must be of $O(h^3)$. (A quantity $q(h)$ is of $O(h^n)$ if $\lim_{h\to 0}\frac{q(h)}{h^{n-\epsilon}} = 0$ for all $\epsilon>0$ while $\lim_{h\to 0}\frac{q(h)}{h^{n+\epsilon}} = \infty$ for all $\epsilon>0$.) It is the one-time statistics of the solution that are of most interest in applications. Thus the minimum amount of information that we seek is the one-time pdf of the solution $f(\underline{V};t)$, the joint pdf of the event $\{\underline{u}(t) = \underline{V}\}$. The solution method developed here yields substantially more information than this: any one-time or multiple-time statistic can in principle be extracted from the numerical solution.

We begin in Section 2 by reviewing currently available numerical solution techniques for stochastic differential or integral equations. It is shown that none of the deterministic methods is ideally suited to the general Ito equation. A case is made for solving Eq.(1) via sample path solutions, that is, by a Monte Carlo method. Section 3 contains the principal theoretical result of this paper that is the basis for the second-order solution algorithm. The new method is compared with an alternative second order Monte Carlo algorithm to point out advantages of the approach adopted here. Three sample problems are then presented in Section 4. Two of these examples satisfy the constraint of Eq.(5) while the third does not. A comparison is made between first-order and second-order solutions; both statistical errors and time differencing errors are discussed. Finally, the results are summarized and conclusions are drawn in Section 5.

2. SURVEY OF SOLUTION TECHNIQUES

A wide variety of numerical schemes are available for the solution of stochastic differential or integral equations. Analytic solutions are also available, but as for deterministic equations, they are applicable only to a limited class of problems and are not considered here. In this section, the development of a new numerical procedure for the solution of the Ito stochastic differential equation is justified by demonstrating that currently available techniques are not well suited to this problem in its general form. Most techniques either: 1) yield insufficient information about the solution; 2) are strongly dependent on the specific form of the coefficient functions or on the dimensionality of the problem; 3) are not amenable to a priori error analysis; or 4) fail to take full advantage of the special characteristics of this equation, e.g., the Gaussian nature of the random terms. Many of the methods described here are applicable to a wider class of problems than stochastic differential equations of the Ito type; no general indictment of these methods is intended.

Numerical solution techniques for stochastic differential equations can be classified into two categories. We define "direct" methods as those that deal with the original stochastic equations. These methods use a discretization analogous to that used in obtaining numerical solutions to deterministic ordinary differential equations. "Indirect" methods transform the given set of random ordinary differential equations into either a single deterministic partial differential equation for the pdf of the solution or into a set of deterministic ordinary or partial differential equations for moments of the solution. These deterministic differential equations are then solved by standard techniques.

We consider indirect methods first. For (almost) any continuous time Markovian stochastic process including the current problem, it is possible to derive an equation governing the evolution of the probability density function [11]. The pdf equation can in principle be solved using finite-difference methods (see Smith [13], for example). However, the finite-difference solution of a partial differential equation in more than three independent variables is beyond the scope of practical computations. Hence, such an approach is limited to stochastic processes of low dimensionality.

Another class of indirect methods makes use of moment equations. A set of deterministic differential equations for the evolution of the moments or cumulants of a stochastic process can be derived either from its pdf evolution equation or directly from the given stochastic differential equations [14,15]. Except in simple cases, however, this set is not closed: higher-order moments appear in the evolution equations for the lower-order moments. Hierarchy methods overcome this difficulty by expressing the higher-order moments or cumulants as functions of lower-order moments or cumulants, truncating the infinite set of equations to a finite number that can then by solved by conventional techniques [16]. Such methods are not attractive here since only a limited amount of information about the solution is obtained, typically only the first few moments, and no a priori error estimates are in general available.

A review of several direct computational methods is given by Boyce [17]. In this paper, two techniques are discussed in detail; both are developed for the case of a scalar random process. The first treats the stochastic process as a collection of random variables, discretizing

both the state space and the index parameter (time), then marching from one time level to the next by determining transition probabilities for the time increment h (see also Kohler and Boyce [18]). This method fails to make full use of the special features of Eq.(1) and additional complications arise for a vector-valued random process. The second technique treats the process as a collection of sample functions by solving deterministic ordinary differential equations for fixed values of the random terms in the original equation, then assigning probabilities to each of these solutions based on the given distribution functions of the random terms (see also Barry and Boyce [19]). This method appears to be more promising but is not an efficient technique for solving the vector Ito equation.

A third direct method is the method of moments developed by Lax and Boyce [20,21]. This is an extension of Galerkin's method for linear deterministic equations to linear random ordinary differential or integral equations. The stochastic process is approximated by a linear combination of basis functions computed from the original equation. The solution consists of a finite number of moments, inadequate for our purposes, a priori error estimates do not appear to be available, and the technique is limited to linear equations.

The above methods are all deterministic, with a unique solution for a given set of initial conditions. Alternatively, sample-path solutions can be obtained by numerically integrating a discrete approximation to the stochastic differential equation. A Taylor series expansion is performed to obtain the corresponding difference equations, and pseudo-random number generators are used in a Monte Carlo solution algorithm to approximate the random terms. A first-order scheme consists of simply replacing dt of Eq.(1) by the finite time step h and replacing $d\underline{W}(t)$ by a vector of independent Gaussian random variables, each with zero mean and variance h: the stochastic differential equation is reduced to a Gaussian random walk. This approximation has been used extensively but often requires an extremely small time step in practice (see Section 4); we limit our discussion to higher-order methods. Second and higher order algorithms have been presented by several workers [22-28] for subsets of Eq.(1). Rao et al. [22] consider a scalar-valued Ito equation, expanding $u(t)$ in a Taylor series and retaining terms to $O(h^2)$. Mil'shtein [24] treats the same problem by expanding the transition pdf in a time step h rather than the stochastic process itself. This gives a simpler numerical algorithm that yields the same statistical information with the same accuracy as that of Rao et al. The approach adopted by Helfand [25] and expanded on by Greenside and Helfand [26] is similar to that of Mil'shtein. They treat a vector-valued Ito equation but restrict their analysis to the case where B_{ij} of Eq.(1) is isotropic and independent of $\underline{u}(t)$. This is also the case considered by Drummond et al. [27,28] who add the further restriction that B_{ij} be independent of t.

The present work extends the approach of Mil'shtein [24] to a general vector-valued Ito stochastic differential equation. The analysis is presented in a different way, however, with particular emphasis on the important subset of problems for which Eq.(5) is satisfied: the numerical method then reduces to a Gaussian random walk although its application is not restricted to problems having a Gaussian or joint-normal solution. The statistical error in approximating the stochastic process by a finite ensemble of sample paths is treated in detail here while the earlier works (except Drummond et al. [28]) dealt primarily with the time differencing error. Systematic verification of the analytic results for both statistical and discret-

ization error is given via sample problems (see Section 4). It is important to justify the use of the second-order algorithm by showing that the larger time step allowed compared to a first-order scheme compensates for the additional computational complexity. This is also done using sample problems.

The method of sample-path solutions, i.e., a Monte Carlo method, appears to be most appropriate for our purposes. By obtaining N sample solutions, all statistics can be obtained with a statistical error of order $N^{-1/2}$, and the order of the discretization error is determined by the level at which the series expansions are truncated in the derivation of the difference equations. Problems of large dimensionality are well suited to this numerical technique [29]: computational requirements increase only linearly with the number of independent variables, the best that can be achieved.

3. DERIVATION OF NUMERICAL METHOD

In this section, we derive the second-order Monte Carlo algorithm for obtaining sample-path solutions to the Ito stochastic ordinary differential equation. This numerical method is based on a Taylor series expansion of the transition joint pdf corresponding to Eq.(1). The advantage of this approach over expanding the stochastic process itself is illustrated by comparing the algorithm derived here with that of Rao et al. [22]. A method for extracting the one-time moments of the solution from the sample paths is described and estimates of the resulting statistical error are given.

We begin by considering the evolution of the joint pdf. The transition pdf $f^*(\underline{V};s,t \mid \underline{U})$ is the joint pdf of the event $\{\underline{u}(t) = \underline{V}\}$ conditional on the event $\{\underline{u}(s) = \underline{U}\}$, $(t \geq s)$. Because Eq.(1) describes a Markov process, all finite-dimensional distributions of the solution can be expressed in terms of f^* [30]. Thus while we concentrate on obtaining the one-time moments of the solution in this paper, knowledge of the transition pdf to $O(h^n)$ is sufficient information to deduce all finite-dimensional moments and distributions of the solution with an error of $O(h^n)$. For any diffusion process, the transition pdf evolves according to the corresponding Fokker-Planck equation or forward Kolmogorov equation [4,5,11]. The Fokker-Planck equation corresponding to Eq.(1) is

$$\frac{\partial f^*}{\partial t} + \frac{\partial [f^* A_i(\underline{V},t)]}{\partial V_i} - \frac{1}{2}\frac{\partial^2 [f^* C_{ij}(\underline{V},t)]}{\partial V_i \partial V_j} = 0 \,. \tag{10}$$

Equation (10) also governs the evolution of $f(\underline{V};t)$ for $t \geq s$ if the state at time s is fixed.

We consider first the case where the constraint of Eq.(5) is satisfied. The basis for the numerical method is then a result derived in the Appendix: subject to the deterministic initial conditions $\underline{u}(t) = \underline{U}$, and provided the coefficient functions satisfy Eq.(5), Eq.(10) causes $f(\underline{V};t+h)$ to evolve as a joint-normal distribution to second-order in h:

$$f(\underline{V};t+h) = f^{JN}(\underline{V};t) + O(h^3) \,, \tag{11}$$

where $f^{JN}(\underline{V};t)$ denotes a joint-normal distribution. This result implies that second-order sample paths can be formed by choosing $\Delta\underline{u}$, the increment in $\underline{u}$ in each interval h (Eq. 7), from a joint-normal distribution whose means and covariances are correct to $O(h^2)$. Second-order accurate approximations for the means and covariances of $\Delta\underline{u}$ can be derived in several ways but they have already been found as part of the proof that appears in the Appendix (Eqs. A34 - A35):

$$<\Delta u_i> = A_i h + \frac{1}{2}(\dot{A}_i + A_{i,l}A_l + \frac{1}{2}A_{i,lm}C_{lm})h^2 + O(h^3) , \tag{12}$$

$$<\Delta u'_i\Delta u'_j> = C_{ij}h + \frac{1}{2}(\dot{C}_{ij} + C_{ij,l}A_l + A_{i,l}C_{jl} + A_{j,l}C_{il} + \frac{1}{2}C_{ij,lm}C_{lm})h^2 + O(h^3) . \tag{13}$$

In these equations, all terms on the right-hand sides are evaluated at the current state $\underline{u}(t)$ at the current time t and primes denote fluctuations about the mean (Eq. A29). The $O(h^3)$ term in Eq.(13) can always be chosen such that $<\Delta u'_i\Delta u'_j>$ is a valid covariance matrix (see Appendix).

Second-order accurate sample paths can be constructed by selecting the increment $\Delta\underline{u}$ from a joint-normal distribution whose means and covariances are given by Eqs.(12) and (13), and marching from time level t to time level $t+h$ by adding this increment to $\underline{u}(t)$. It is more convenient computationally to express the sample paths as a second-order random walk; this also facilitates comparison with Rao et al. Assume that it is possible to devise a random walk of the form

$$u_i(t+h) = u_i(t) + F_i(\underline{u}(t),t,h) + f_{ij}(\underline{u}(t),t,h)\xi_j , \tag{14}$$

such that the transition pdf of Eq.(14) corresponds to that of Eq.(1) to $O(h^2)$. Here F_i and f_{ij} are deterministic functions of the state at time t and $\underline{\xi}$ is a vector of independent standardized Gaussian random variables, i.e., $\underline{\xi}$ is a joint-normal random vector with means and covariances

$$<\xi_i> = 0 , \qquad <\xi_i\xi_j> = \delta_{ij} . \tag{15}$$

This is a plausible assumption. The resulting pdf of $\Delta\underline{u}$ will be joint-normal since Eq.(14) expresses the increment as a linear combination of joint-normal random variables. Also, the two functions F_i and f_{ij} provide enough degrees of freedom that the means and covariances can be made to evolve by Eqs.(12) and (13) to $O(h^2)$. The means and covariances of $\Delta\underline{u}$ corresponding to Eq.(14) are

$$<\Delta u_i> = F_i , \qquad <\Delta u'_i\Delta u'_j> = f_{ik}f_{jk} . \tag{16}$$

To satisfy Eqs.(12) and (13) with an error of $O(h^3)$, we find that

$$F_i(\underline{u}(t),t,h) = A_i h + \frac{1}{2}(\dot{A}_i + A_{i,l}A_l + \frac{1}{2}A_{i,lm}C_{lm})h^2 , \tag{17}$$

$$f_{ij}(\underline{u}(t),t,h) = B_{ij}h^{1/2} + \frac{1}{2}[\dot{B}_{ij} + B_{ij,l}A_l + A_{i,l}B_{lj} + \frac{1}{2}(B_{ij,lm} + B^{-1}_{jk}B_{kp,l}B_{ip,m})C_{lm}]h^{3/2} . \tag{18}$$

The existence of B^{-1}_{ij} is not an additional restriction on the numerical method. However, sample paths cannot be constructed using Eq.(14) unless $B^{-1}_{jk}B_{kp,l}B_{ip,m}C_{lm}$ is finite whenever B_{ij} is

singular. Equation (18) effectively selects the $O(h^3)$ truncation term in Eq.(13) in a way that guarantees realizability, i.e., so that $<\Delta u'_i \Delta u'_j>$ is symmetric and positive semi-definite.

We next consider the case where Eq.(5) is not satisfied. The transition pdf is not joint-normal to second order in h and a Gaussian random walk as expressed by Eq.(14) is no longer appropriate. A second-order random walk can still be constructed by extending the idea proposed by Mil'shtein [24] for a scalar random process. Equation (14) is replaced by a random walk of the form

$$u_i(t+h) = u_i(t) + G_i(\underline{u}(t),t,h) + g_{ij}(\underline{u}(t),t,h)\xi_j + \gamma_{ijk}(\underline{u}(t),t,h)\xi_j\xi_k \,, \tag{19}$$

where G_i, g_{ij}, and γ_{ijk} are deterministic functions of the state at time t and $\underline{\xi}$ is again a vector of independent standardized Gaussian random variables. As before, the goal is to find functions G_i, g_{ij}, and γ_{ijk} such that the distribution of the random variable $\Delta\underline{u}$ corresponding to Eq.(19) agrees to $O(h^2)$ with the transition pdf of Eq.(1). Equation (A36) shows that to satisfy this, it is enough that the first four moments of $\Delta\underline{u}$ from Eq.(19) agree with those of Eq.(1) to $O(h^2)$ and that all higher moments of $\Delta\underline{u}$ be of $O(h^3)$ or higher. The required functions are

$$G_i(\underline{u}(t),t,h) = (A_i - \frac{1}{2}B_{ip,l}B_{lp})h + \frac{1}{2}(\dot{A}_i + A_{i,l}A_l + \frac{1}{2}A_{i,lm}C_{lm})h^2 \,, \tag{20}$$

$$\begin{aligned} g_{ij}(\underline{u}(t),t,h) = B_{ij}h^{1/2} + \frac{1}{2}\{&-\frac{1}{4}B^{-1}_{js}B_{sp,m}B_{mq}B_{iq,n}B_{np} + \dot{B}_{ij} \\ &+ B_{ij,l}A_l + A_{i,l}B_{lj} + \frac{1}{2}(B_{ij,lm} + \frac{1}{2}B^{-1}_{jq}B_{qr,l}B_{ir,m})C_{lm}\}h^{3/2} \,, \end{aligned} \tag{21}$$

$$\gamma_{ijk}(\underline{u}(t),t,h) = \frac{1}{2}B_{ik,p}B_{pj}h \,. \tag{22}$$

To summarize, it has been shown that the second-order random walk

$$\begin{aligned} u_i(t+h) = u_i(t) &+ (A_i - \frac{1}{2}B_{ip,l}B_{lp})h + \frac{1}{2}(\dot{A}_i + A_{i,l}A_l + \frac{1}{2}A_{i,lm}C_{lm})h^2 \\ &+ \{B_{ij}h^{1/2} + \frac{1}{2}[-\frac{1}{4}B^{-1}_{js}B_{sp,m}B_{mq}B_{iq,n}B_{np} + \dot{B}_{ij} \\ &+ B_{ij,l}A_l + A_{i,l}B_{lj} + \frac{1}{2}(B_{ij,lm} + \frac{1}{2}B^{-1}_{jq}B_{qr,l}B_{ir,m})C_{lm}]h^{3/2}\}\xi_j \\ &+ \frac{1}{2}B_{ik,p}B_{pj}h\xi_j\xi_k \,, \end{aligned} \tag{23}$$

yields the proper evolution of the joint pdf of the increment $\Delta\underline{u}$ to second-order in h. When the constraint of Eq.(5) is satisfied, this simplifies to the Gaussian random walk (Eqs. 14-18)

$$\begin{aligned} u_i(t+h) = u_i(t) &+ A_i h + \frac{1}{2}(\dot{A}_i + A_{i,l}A_l + \frac{1}{2}A_{i,lm}C_{lm})h^2 \\ &+ \{B_{ij}h^{1/2} + \frac{1}{2}[\dot{B}_{ij} + B_{ij,l}A_l + A_{i,l}B_{lj} + \frac{1}{2}(B_{ij,lm} + B^{-1}_{jk}B_{kp,l}B_{ip,m})C_{lm}]h^{3/2}\}\xi_j \,. \end{aligned} \tag{24}$$

In the numerical implementation, Eq.(1) is integrated in time by a marching algorithm where the state at time $t+h$ is obtained from the state at time t using Eq.(23) or (24). These

difference equations are similar to those resulting from a discrete approximation to a deterministic ordinary differential equation except for the presence of the random vector $\underline{\xi}$. These random terms present no difficulty, however, as computationally efficient methods for generating a sequence of pseudo-random numbers having a normal distribution are well known (see Knuth [31], for example). Samples of the Gaussian random variables are generated at each time step. A large number N of independent sample paths are constructed, each representing an independent realization of the stochastic process or an independent "particle" path in the M-dimensional $\underline{V}$ space. Ensemble averaging over these particles is then used to obtain estimates of the statistics of the solution. The initial condition $\underline{U}^0$ of Eq.(2), if random, is also computed using an appropriate pseudo-random number generator.

In the second-order method of Rao et al. for a scalar random variable, derivatives of the coefficients of the random term up to the third order with respect to u and the u,t mixed derivatives appear. Also, three correlated Gaussian random variables need to be generated. By contrast, Eq.(23) (in the scalar case) contains derivatives up to just the second order with respect to u and only a single Gaussian random variable is necessary. In Rao et al., it is the sample paths themselves that are expanded in a Taylor series while here (and in Mil'shtein [24]) it is the transition pdf that is expanded. Thus the sample paths of Rao et al. converge in probability as h^2 to sample solutions of the given stochastic differential equation while here the sample paths converge in distribution, a weaker form of stochastic convergence [30]. This does not imply any loss of information with the new method, but a gain in efficiency: by expanding the transition pdf instead of the sample paths, we obtain the same statistical information about the solution with the same accuracy as Rao et al. using a simpler numerical algorithm.

We now consider the evaluation of one-point statistics from the sample solutions. In addition to the discretization error of $O(h^3)$ for each time step, there is a statistical error associated with approximating the stochastic process by an ensemble of sample paths. Consider an arbitrary function of the random variables $\underline{u}(t)$, $g(\underline{u},t)$. The mean of $g(\underline{u},t)$ can be approximated as an ensemble average over the N independent sample paths or particles:

$$<g(\underline{u},t)> \approx <g(\underline{u},t)>_N \equiv \frac{1}{N}\sum_{i=1}^{N} g(\underline{u}^{(i)},t) , \tag{25}$$

where a superscript "(i)" refers to the i^{th} particle. A consequence of the central limit theorem [32] is that, as N tends to infinity, the distribution of $<g(\underline{u},t)>_N$ approaches a normal distribution with mean and variance

$$E_g \equiv <<g(\underline{u},t)>_N> = <g(\underline{u},t)> , \tag{26}$$

$$\sigma_g^2 \equiv <(<g(\underline{u},t)>_N)'^2> = \frac{1}{N}<g'(\underline{u},t)^2> . \tag{27}$$

σ_g is a measure of the statistical error incurred in approximating mean values by ensemble averages: Eq.(27) shows that this error decreases as $N^{-1/2}$. Setting $g(\underline{u},t) = u_\alpha(t)$, we obtain the

standard statistical error for the mean of u_α:

$$\sigma_{<u_\alpha>} = \frac{1}{\sqrt{N}} <u_\alpha'^2>^{1/2} . \tag{28}$$

Similarly, we obtain the standard error for the second central moments:

$$\sigma_{<u'_\alpha u'_\beta>} = \frac{1}{\sqrt{N}} (<u'^2_\alpha u'^2_\beta> - <u'_\alpha u'_\beta>^2)^{1/2} , \tag{29}$$

(no sum on α or on β). A normalized standard statistical error for the covariances is defined by dividing Eq.(29) by $<u'_\alpha u'_\beta>$. This yields, for $\alpha = \beta$,

$$\varepsilon_{<u'_\alpha u'_\alpha>} = \frac{1}{\sqrt{N}} (K_{u_\alpha} - 1)^{1/2} , \tag{30}$$

where K is the flatness factor or kurtosis:

$$K_{u_\alpha} \equiv \frac{<u'^4_\alpha>}{<u'^2_\alpha>^2} . \tag{31}$$

For a joint-normal distribution, K is identically equal to 3.

Estimates of the one-time pdf $f(\underline{V};t)$ can also be extracted from the sample solutions. Each particle or realization can be considered as a delta function discrete representation of the pdf in the M dimensional $\underline{V}$ state space. The ensemble average pdf $f_N(\underline{V};t)$ is then the normalized number density of particles in this state space (see Pope [33], for example). This is not discussed further here since in what follows, we concentrate on obtaining the one-time moments of the solution.

A new second-order Monte Carlo algorithm for the solution to the Ito stochastic differential equation has been presented. The basis for this numerical method is an expansion of the joint pdf of the increment of the particle state in a time step h. By obtaining N sample-path solutions, moments and distributions can be obtained with a discretization error of $O(h^3)$ on each time step and a statistical error of order $N^{-1/2}$.

4. EXAMPLES

Three sample problems are solved numerically using the Monte Carlo method derived in Section 3. These examples retain the essential features of the general Ito stochastic differential equation while remaining simple enough that their solution can be studied in some detail. The first problem deals with a Langevin equation and the second treats the case of a non-linear deterministic term in Eq.(1). In these two examples, Eq.(5) is satisfied and the simplified Gaussian random walk of Eq.(24) is used. The final problem treats a two-dimensional stochastic process for which Eq.(5) is not satisfied; the general random walk of Eq.(23) is then used. An analytic solution is available for the first problem, and the second-order method is compared both with this analytic solution and with a first-order method. Only the one-time means and covariance matrix of the solution are discussed in detail; higher-order moments are examined only to study departures from joint-normality.

4.1. Homogeneous Langevin Equation

Consider the coupled set of stochastic differential equations

$$dx(t) = v(t)dt\ , \qquad dv(t) = -\alpha(t)v(t)dt + \beta(t)^{1/2}dW(t)\ . \tag{32}$$

$x(t)$ and $v(t)$ may represent, for example, the position and velocity of a particle subject to a deterministic drag force that is linear in the particle velocity and a random collisional force operating at a time scale much smaller than that of the macroscopic particle motion [1]. Since the coefficient functions $\alpha(t)$ and $\beta(t)$ are independent of position, the equation for $v(t)$ is referred to as a homogeneous Langevin equation. The increments Δx and Δv (Eq. 7) are joint-normal with means and covariances (Eqs. 12 and 13)

$$\langle\Delta x\rangle = vh - \frac{1}{2}\alpha vh^2 + O(h^3)\ , \qquad \langle\Delta v\rangle = -\alpha vh + \frac{1}{2}(\alpha^2 - \dot{\alpha})vh^2 + O(h^3)\ , \tag{33}$$

$$\langle\Delta x'^2\rangle = \frac{1}{3}\beta h^3 + O(h^4)\ , \qquad \langle\Delta v'^2\rangle = \beta h + (\frac{1}{2}\dot{\beta} - 2\alpha\beta)h^2 + O(h^3)\ , \tag{34}$$

$$\langle\Delta x'\Delta u'\rangle = \frac{1}{2}\beta h^2 + O(h^3)\ . \tag{35}$$

Here the Taylor series expansions in the Appendix have been extended to third order in h to obtain the leading order term in $\langle\Delta x'^2\rangle$. The resulting correlation coefficient r

$$r \equiv \frac{\langle\Delta x'\Delta v'\rangle}{[\langle\Delta x'^2\rangle\langle\Delta v'^2\rangle]^{1/2}}\ , \tag{36}$$

is then $r = \sqrt{3}/2 + O(h)$. It may be seen from Eqs.(34)-(35) that the variance of Δx must be $O(h^3)$ to satisfy the realizability condition $-1 \le r \le +1$. In the numerical implementation, however, $\langle\Delta x'^2\rangle$ may be anything of $O(h^3)$ without affecting the formal accuracy of the scheme. The second-order random walk of Eq.(24) for this case reduces to

$$x(t+h) = x + vh - \frac{1}{2}\alpha vh^2 + \frac{1}{2}(\beta h^3)^{1/2}\xi\ , \tag{37}$$

$$v(t+h) = v - \alpha vh + \frac{1}{2}(\alpha^2 - \dot{\alpha})vh^2 + [(\beta h)^{1/2} + \frac{\dot{\beta} - 2\alpha\beta}{4\beta^{1/2}}h^{3/2}]\xi\ , \tag{38}$$

where ξ is a standardized normal random variable (zero mean, unit variance). Equation (37) yields $\langle\Delta x'^2\rangle = 1/4\beta h^3 + O(h^4)$, so that the correlation coefficient of Eq.(36) is unity rather than $\sqrt{3}/2$ as $h \to 0$. This is a consequence of expressing the results as a random walk using a vector of independent Gaussian random variables, Eqs.(14)-(18) and (24). A first-order method is obtained by retaining only the terms to $O(h)$ in Eqs.(37) and (38):

$$x(t+h) = x + vh\ , \tag{39}$$

$$v(t+h) = v - \alpha vh + (\beta h)^{1/2}\xi\ . \tag{40}$$

To separate the effects of discretization error (non-zero time step h) and statistical error (finite sample size N), we derive deterministic difference equations for the first two moments

corresponding to the above sample path solutions. For the second-order method of Eqs.(37)-(38) we obtain

$$\langle x(t+h)\rangle = \langle x\rangle + \langle v\rangle h - \frac{1}{2}\alpha\langle v\rangle h^2 , \tag{41}$$

$$\langle v(t+h)\rangle = \langle v\rangle - \alpha\langle v\rangle h + \frac{1}{2}(\alpha^2-\dot{\alpha})\langle v\rangle h^2 , \tag{42}$$

$$\begin{aligned}\langle x'(t+h)^2\rangle = {} & \langle x'^2\rangle + 2\langle x'v'\rangle h + \langle v'^2\rangle h^2 - \alpha\langle x'v'\rangle h^2 \\ & + \frac{1}{4}\beta h^3 - \alpha\langle v'^2\rangle h^3 + \frac{1}{4}\alpha^2\langle v'^2\rangle h^4 ,\end{aligned} \tag{43}$$

$$\begin{aligned}\langle v'(t+h)^2\rangle = {} & \langle v'^2\rangle + \beta h - 2\alpha\langle v'^2\rangle h + \alpha^2\langle v'^2\rangle h^2 + (\alpha^2-\dot{\alpha})\langle v'^2\rangle h^2 \\ & + (\frac{1}{2}\dot{\beta}-\alpha\beta)h^2 - \alpha(\alpha^2-\dot{\alpha})\langle v'^2\rangle h^3 + \frac{1}{16}(\dot{\beta}\beta^{-1/2} - 2\alpha\beta^{1/2})^2 h^3 \\ & + \frac{1}{4}(\alpha^2-\dot{\alpha})^2\langle v'^2\rangle h^4 ,\end{aligned} \tag{44}$$

$$\begin{aligned}\langle x'(t+h)v'(t+h)\rangle = {} & \langle x'v'\rangle - \alpha\langle x'v'\rangle h + \langle v'^2\rangle h - \frac{3}{2}\alpha\langle v'^2\rangle h^2 \\ & + \frac{1}{2}(\alpha^2-\dot{\alpha})\langle x'v'\rangle h^2 + \frac{1}{2}\beta h^2 + \frac{1}{2}(\alpha^2-\dot{\alpha})\langle v'^2\rangle h^3 + \frac{1}{2}\alpha^2\langle v'^2\rangle h^3 \\ & + \frac{1}{8}(\dot{\beta} - 2\alpha\beta)h^3 - \frac{1}{4}\alpha(\alpha^2-\dot{\alpha})\langle v'^2\rangle h^4 ,\end{aligned} \tag{45}$$

where all terms on the right-hand sides are evaluated at time level t. Terms of all order in h have been retained in these difference equations: this does not imply that these are fourth-order accurate approximations for the variances and covariance. These higher-order terms are retained so that we have a set of difference equations that correspond exactly to the Monte Carlo solution with the statistical error removed. The set of difference equations corresponding to the first-order method of Eqs.(39) and (40) is computed similarly:

$$\langle x(t+h)\rangle = \langle x\rangle + \langle v\rangle h , \tag{46}$$

$$\langle v(t+h)\rangle = \langle v\rangle - \alpha\langle v\rangle h , \tag{47}$$

$$\langle x'(t+h)^2\rangle = \langle x'^2\rangle + 2\langle x'v'\rangle h + \langle v'^2\rangle h^2 , \tag{48}$$

$$\langle v'(t+h)^2\rangle = \langle v'^2\rangle + \beta h - 2\alpha\langle v'^2\rangle h + \alpha^2\langle v'^2\rangle h^2 , \tag{49}$$

$$\langle x'(t+h)v'(t+h)\rangle = \langle x'v'\rangle - \alpha\langle x'v'\rangle h + \langle v'^2\rangle h - \alpha\langle v'^2\rangle h^2 . \tag{50}$$

The joint pdf of $x(t)$ and $v(t)$ relaxes to joint-normal from any initial distribution in this homogeneous problem. If we choose a joint-normal or delta function initial condition, the pdf remains joint-normal for all time and the evolution of the first and second moments provides a complete one-time statistical description of the evolution of the system. A closed set of ordi-

nary differential equations governs the evolution of these one-time moments; these follow directly from the difference equations Eqs.(41)-(45) or (46)-(50). For reasonable choices of $\alpha(t)$ and $\beta(t)$, explicit solutions can be obtained for the time evolution of the means and covariances. We choose

$$\alpha(t) = \frac{\alpha_0}{\alpha_0 t+1} , \qquad \beta(t) = \beta_0(\alpha_0 t+1)^m . \tag{51}$$

The analytic solutions are then

$$\langle x(t)\rangle = \langle x(0)\rangle + \frac{\langle v(0)\rangle}{\alpha_0}\ln(\alpha_0 t+1) , \tag{52}$$

$$\langle v(t)\rangle = \frac{\langle v(0)\rangle}{(\alpha_0 t+1)} , \tag{53}$$

$$\begin{aligned}\langle x'(t)^2\rangle = \langle x'(0)^2\rangle &+ \frac{2\beta_0}{\alpha_0^3(m+3)^3}[(\alpha_0 t+1)^{m+3} - 1] \\ &+ \frac{2}{\alpha_0}[\langle x'(0)v'(0)\rangle - \frac{\beta_0}{\alpha_0^2(m+3)^2}]\ln(\alpha_0 t+1) \\ &+ \frac{1}{\alpha_0^2}[\langle v'^2(0)\rangle - \frac{\beta_0}{\alpha_0(m+3)}][\ln(\alpha_0 t+1)]^2 ,\end{aligned} \tag{54}$$

$$\langle v'(t)^2\rangle = \frac{\langle v'(0)^2\rangle}{(\alpha_0 t+1)^2} + \frac{\beta_0}{\alpha_0(m+3)}[(\alpha_0 t+1)^{m+1} - (\alpha_0 t+1)^{-2}], \tag{55}$$

$$\begin{aligned}\langle x'(t)v'(t)\rangle = \frac{\langle x'(0)v'(0)\rangle}{\alpha_0 t+1} &+ \frac{\beta_0}{\alpha_0^2(m+3)^2}[(\alpha_0 t+1)^{m+2} - (\alpha_0 t+1)^{-1}] \\ &+ \frac{1}{\alpha_0}[\langle v'(0)^2\rangle - \frac{\beta_0}{\alpha_0(m+3)}]\frac{\ln(\alpha_0+1)}{\alpha_0 t+1} .\end{aligned} \tag{56}$$

These solutions are valid for all $\alpha_0 \geq 0$, $\beta_0 \geq 0$, $m \neq -2$, and $m \neq -2½$. In general, the moment equations do not constitute a closed set. This would be the case if, for example, α and β of Eq.(32) were functions of x as well as of t.

The problem that we select for study is Eq.(32) where $\alpha(t)$ and $\beta(t)$ are given by Eq.(51), with $\beta_0 = 1$, $\alpha_0 = 1$, and $m = 3$, subject to the delta function initial conditions $x(0) = 0$, $v(0) = 1$. This provides a severe test for the Monte Carlo algorithm: the means change only gradually while the variances and covariance grow rapidly, hence the statistical error in the means grows rapidly (Eq. 28). The pdf is joint-normal for all time so that the kurtoses of $x(t)$ and of $v(t)$ are equal to 3 (Eq. 31) while the skewness factors

$$S_x \equiv \frac{\langle x'^3\rangle}{\langle x'^2\rangle^{3/2}} , \qquad S_v \equiv \frac{\langle v'^3\rangle}{\langle v'^2\rangle^{3/2}} , \tag{57}$$

are identically zero. Analytic solutions for the means and covariances (Eqs. 52-56) are shown in Figures 1 and 2.

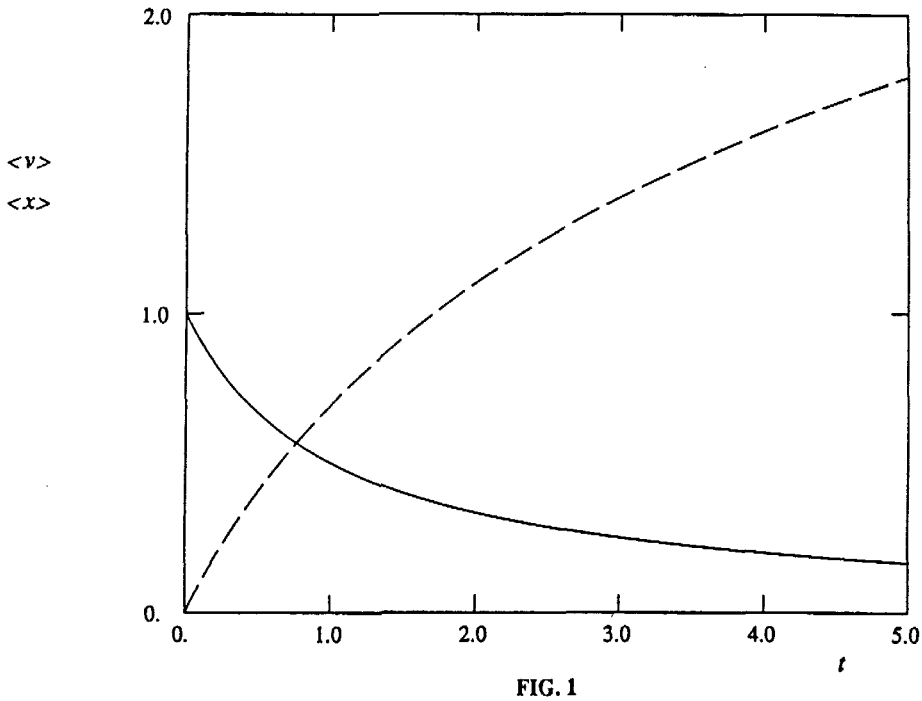

FIG. 1
Homogeneous Langevin, Analytic Evolution of Means

Two sets of computations are reported. In the first set, the difference equations Eqs.(41)-(45) and (46)-(50) are compared to the analytic solutions of Eqs.(52)-(56) to study the discretization error. In the second set, the random walk of Eqs.(37)-(38) is used with a small enough time step that discretization error is negligible to study the statistical error.

We consider the discretization error first. Figures 3-5 show the normalized error in the means and covariances at time $t = 5.0$ as a function of the time step h for Eqs.(41)-(45) and Eqs.(46)-(50). The normalized discretization error ε in Figures 3-5 is

$$\varepsilon \equiv \frac{|\textit{analytic solution} - \textit{numerical solution}|}{\textit{analytic solution}} . \tag{58}$$

As expected, the slope of the lines corresponding to the $O(h)$ method is equal to unity on the log-log plots, and the slope of those corresponding to the $O(h^2)$ method is equal to two. These calculations were performed in 64-bit precision to minimize roundoff error. We conclude from Figures 3-5 that to keep the normalized discretization errors to less than, say, 1%, the $O(h^2)$ method permits a time step h roughly 10 times that permitted by the $O(h)$ method; for 0.1% error, the second-order method allows a step 100 times larger than the first-order method.

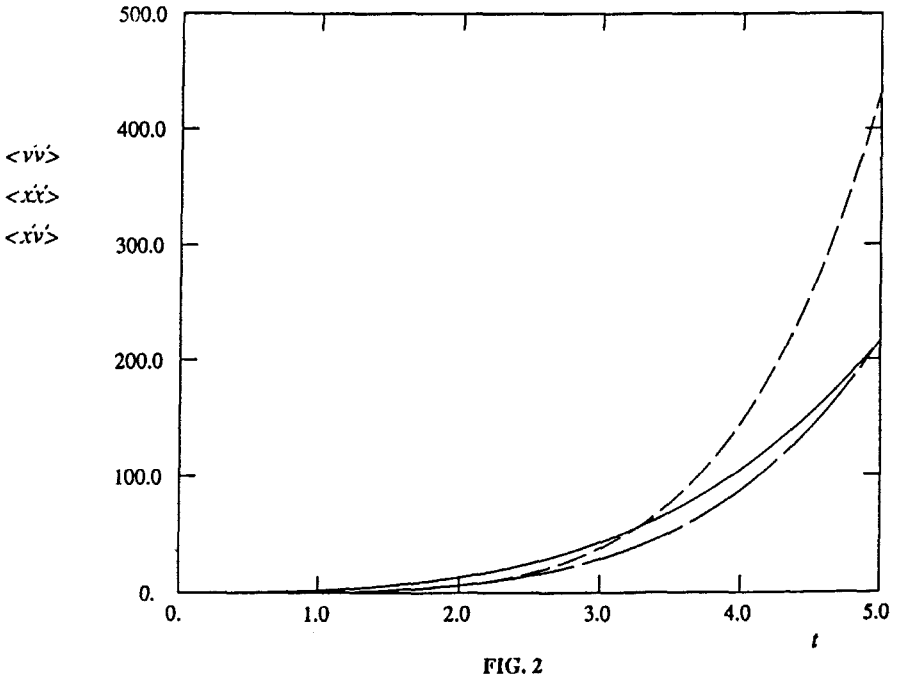

FIG. 2
Homogeneous Langevin, Analytic Evolution of Variances and Covariance

velocity variance
position variance
position-velocity covariance

We now consider the statistical error. Second-order Monte Carlo runs were made with $h = 0.01$ to keep the discretization error to less than 0.01%. In all figures showing Monte Carlo results, only a few time points are plotted for clarity. Figures 6 and 7 show the normalized statistical errors for the mean and variance of $v(t)$ as a function of time for $N = 1{,}000$, $N = 10{,}000$, and $N = 100{,}000$. Results for $\langle x \rangle$, $\langle x'^2 \rangle$, and $\langle x'v' \rangle$ are similar. Here the normalized statistical error is

$$\varepsilon \equiv \frac{\textit{analytic solution} - \textit{numerical solution}}{\sigma_\varepsilon} , \tag{59}$$

where σ_ε is defined in Eqs.(28) and (29). Also shown are error bars corresponding to the normalized standard statistical error. The computed statistical error is in good agreement with the predictions.

For a given time step h and to integrate to a fixed time level, the second-order Monte Carlo algorithm requires about 10% more CPU time than the first-order method for this

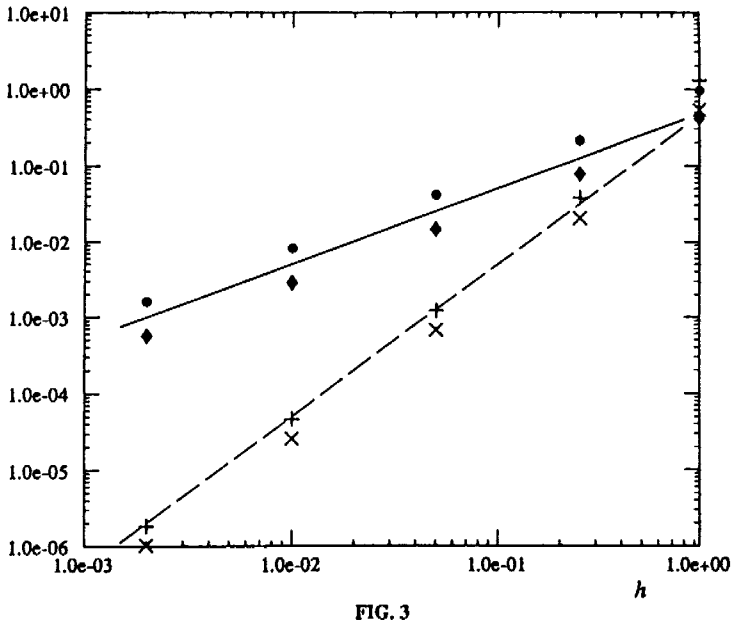

FIG. 3
Homogeneous Langevin, Discretization Error in the Means @ t=5

- mean velocity, first-order method
- reference line with slope = 1
- mean velocity, second-order method
- mean position, first-order method
- mean position, second-order method
- reference line with slope = 2

example; this difference is greater for more complex problems. If a 1% error is acceptable, then the second-order method allows a time step roughly ten times as large as the first-order method and approximately 10,000 particles are needed, a reasonable number for practical computations. The second-order method is clearly justified in this case. However, to further reduce the statistical error requires a significant increase in N since the statistical error is of order $N^{-\frac{1}{2}}$ (Eqs. 27-30). At first glance, it appears reasonable to argue that there is no point in choosing h so small or N so large that one type of error is much smaller than the other. There is an important distinction between the two types of error, however. Discretization error is systematic while statistical error is random and, for the example considered here, unbiased. Methods such as spline fitting may be used to effectively decrease the statistical error [34,35]. It may make sense then to keep the discretization error at a level smaller than the standard statistical error, and in that case the higher-order method becomes more attractive.

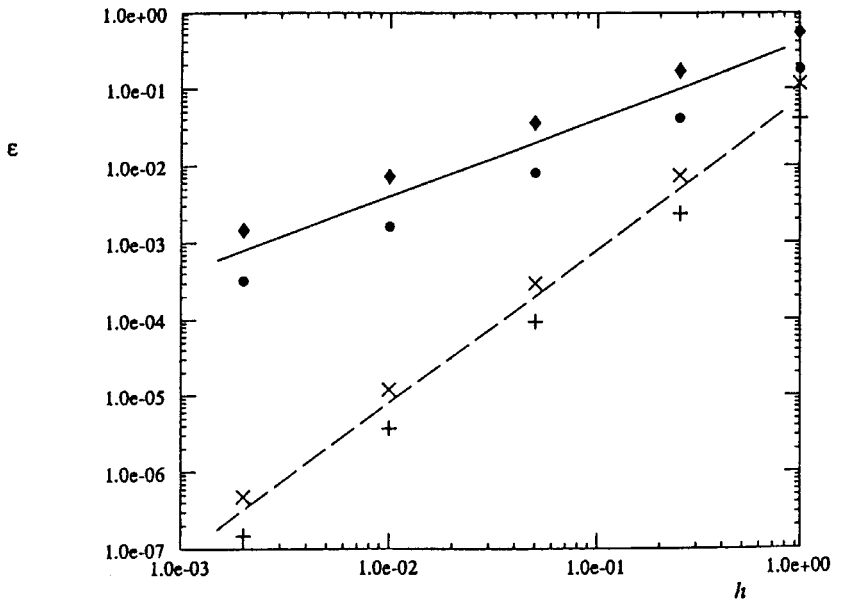

FIG. 4
Homogeneous Langevin, Discretization Error in the Variances @ t=5

velocity variance, first-order method
reference line with slope = 1
velocity variance, second-order method
position variance, first-order method
position variance, second-order method
reference line with slope = 2

4.2. Non-linear Deterministic Term

We next consider a case where the deterministic term is non-linear in the dependent variable. The pdf is no longer joint-normal and the moment equations no longer constitute a closed set. Having discussed discretization and statistical errors in detail in the previous example, we content ourselves here with demonstrating computationally that the second-order method converges as $N \to \infty$ and as $h \to 0$. It is expected that as the departure from joint-normal behavior becomes more severe, the second-order method becomes more attractive. We select a one-dimensional problem where there is a significant departure from a Gaussian pdf:

$$dv(t) = \alpha(v,t)dt + \beta(t)dW(t) , \tag{60}$$

with

$$\alpha(v,t) = \alpha_0[2 + \sin(\omega_\alpha t)][1 + erf(\frac{v}{v_0})] , \qquad \beta(t) = \beta_0 \cos(\omega_\beta t) , \tag{61}$$

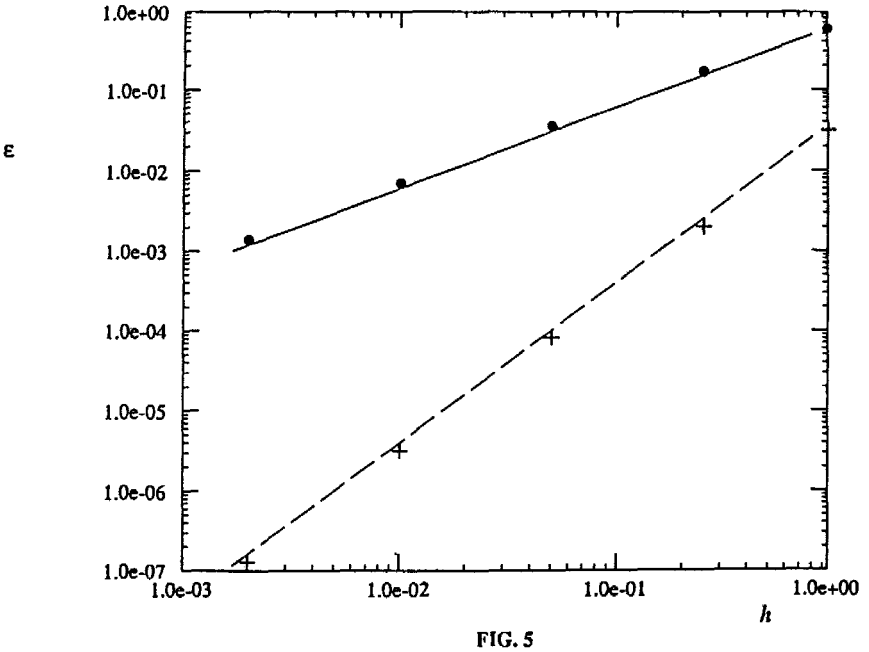

FIG. 5
Homogeneous Langevin, Discretization Error in the Covariance @ t=5

- position-velocity covariance, first-order method
- reference line with slope = 1
- position-velocity covariance, second-order method
- reference line with slope = 2

and $\alpha_0 = 1$, $\omega_\alpha = \pi$, $v_0 = 1$, $\beta_0 = 2.5$, $\omega_\beta = \pi$. The initial condition is $v(0) = 1$. Equation (5) is satisfied; the transition pdf is Gaussian to $O(h^2)$ but the pdf $f(V;t)$ is not Gaussian for all time. The second-order random walk for this problem is (Eq. 24)

$$v(t+h) = v + \alpha h + \frac{1}{2}(\dot{\alpha} + \alpha_{,v}\alpha + \frac{1}{2}\alpha_{,vv}\beta^2)h^2 + [\beta + \frac{1}{2}(\dot{\beta} + \alpha_{,v}\beta)h]h^{1/2}\xi \,, \tag{62}$$

where a subscript v denotes a partial derivative with respect to v, and ξ is again a standardized normal random variable. Three sets of Monte Carlo calculations were performed with $N = 1{,}000$, $N = 10{,}000$, and $N = 100{,}000$. The time step in each case is $h = 0.01$. Results for the one-time moments of $v(t)$ up to the fourth are plotted in Figures 8-10. The line in each figure is a least squares cubic splines fit through the numerical solution for $N = 100{,}000$. Clearly, the solution for these one-time moments converges as N increases. Note that there is a significant departure from Gaussianity evident in the evolution of the skewness and flatness factors (Figure 10). To check that the solution algorithm is $O(h^2)$, we plot in Figure 11 the

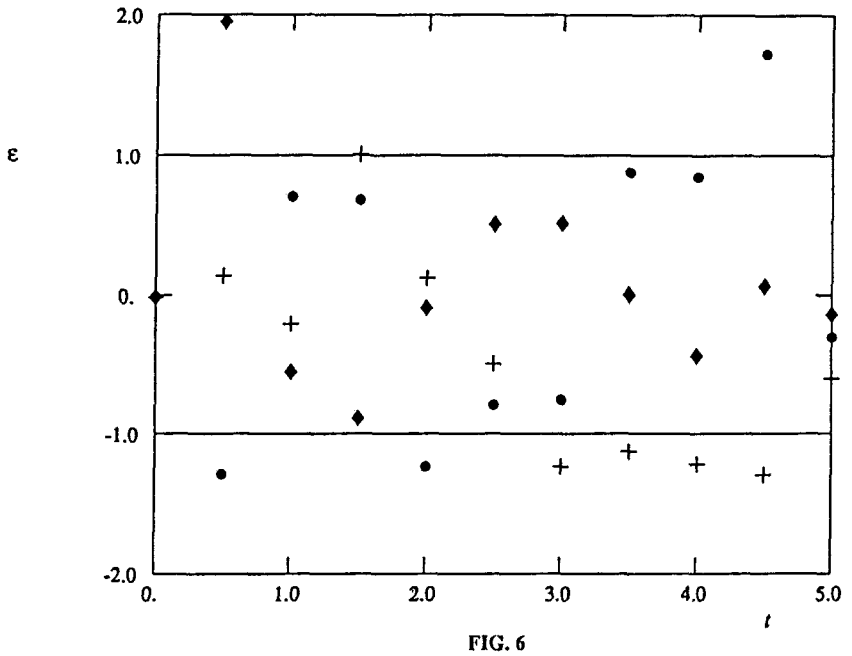

FIG. 6
Homogeneous Langevin, Statistical Error in Mean Velocity

• Monte Carlo calculation, N = 1,000
\+ Monte Carlo calculation, N = 10,000
♦ Monte Carlo calculation, N = 100,000
——— normalized standard statistical error

normalized error in the mean and variance (Eq. 58) at $t = 5.0$ vs. the time step h for $N = 1{,}000{,}000$. Here the "analytic solution" is taken as the Monte Carlo solution for $h = 0.02$, $N = 1{,}000{,}000$ so that the normalized statistical error is about 0.1%. Figure 11 verifies that the error in the variance decreases as the square of the time step. However, the error in the mean is barely above the threshold of statistical noise so that we cannot draw meaningful conclusions about the discretization error for $\langle v \rangle$. Both error curves depart from linearity at large values of h as higher-order terms become significant.

4.3. General Two-Dimensional Problem

As a final example, consider a two-dimensional stochastic process not satisfying Eq.(5):

$$dx(t) = A_1(x,v)dt + B_1(x)dW_1(t) , \qquad dv(t) = A_2(x,v,t)dt + B_2(v,t)dW_2(t) , \qquad (63)$$

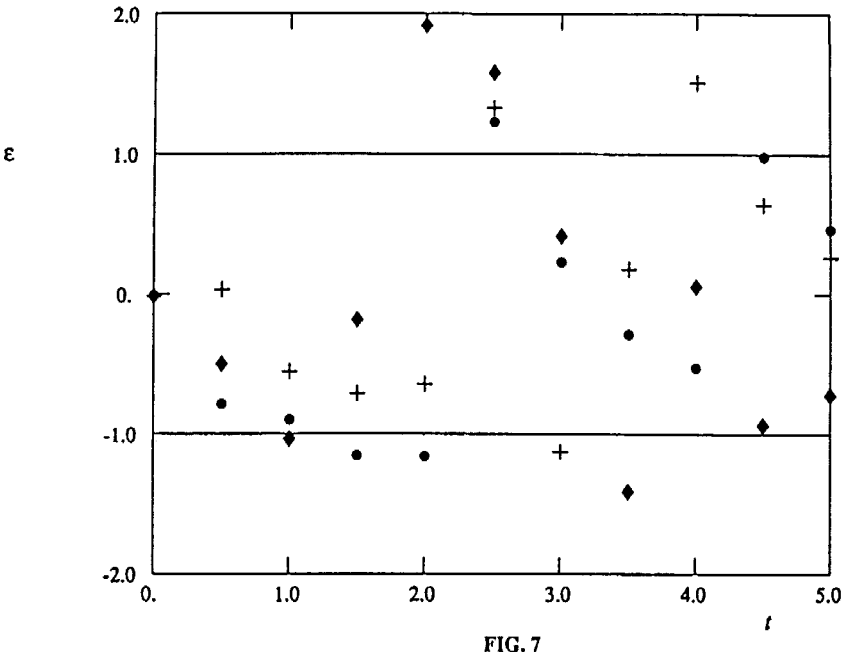

FIG. 7
Homogeneous Langevin, Statistical Error in Velocity Variance

- • Monte Carlo calculation, N = 1,000
- + Monte Carlo calculation, N = 10,000
- ♦ Monte Carlo calculation, N = 100,000
- ——— normalized standard statistical error

where the coefficient functions are

$$A_1(x,v) = \alpha_1[1 + erf(\frac{x}{x_0})][2 + \cos(v)] , \qquad B_1(x) = \beta_1[2 + \sin(x)] , \tag{64}$$

$$A_2(x,v,t) = \alpha_2 v[2 + \sin(\omega t)]\cos(x) , \qquad B_2(v,t) = \beta_2[2 + \cos(\omega t)][2 + \sin(v)] , \tag{65}$$

with $\alpha_1 = 1$, $\alpha_2 = 1$, $\beta_1 = 2$, $\beta_2 = 1/4$, $x_0 = 1$, and $\omega = \pi$. The second-order random walk now has the form (Eq. 23)

$$\begin{aligned} x(t+h) = x(t) &+ (A_1 - \frac{1}{2}B_{1x}B_1)h + \frac{1}{2}(A_{1x}A_1 + A_{1v}A_2 + \frac{1}{2}A_{1xx}B_1^2 + \frac{1}{2}A_{1vv}B_2^2)h^2 \\ &+ \{B_1 h^{1/2} + [\frac{1}{2}A_1 B_{1x} + \frac{1}{2}A_{1x}B_1 + \frac{1}{4}B_{xx}B_1^2]h^{3/2}\}\xi_1 \\ &+ \frac{1}{2}A_{1v}B_2 h^{3/2}\xi_2 + \frac{1}{2}B_{1x}B_1 h\xi_1^2 , \end{aligned} \tag{66}$$

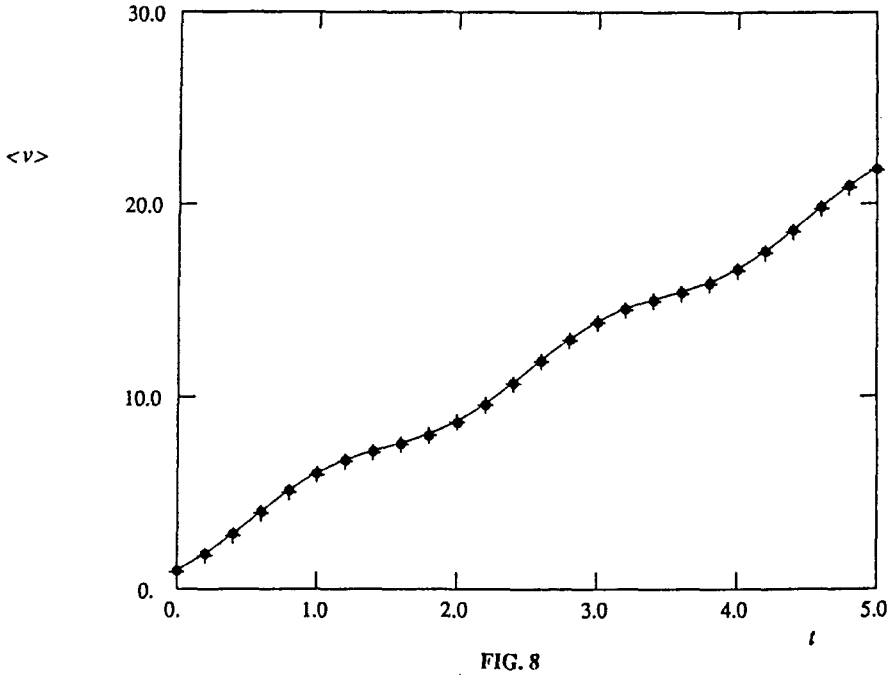

FIG. 8
Non-linear Example, Evolution of the Mean

• Monte Carlo calculation, N = 1,000
\+ Monte Carlo calculation, N = 10,000
——— spline fit through Monte Carlo calculation, N = 100,000

$$v(t+h) = v(t) + (A_2 - \frac{1}{2}B_{2v}B_2)h + \frac{1}{2}(\dot{A}_2 + A_{2x}A_1 + A_{2v}A_2 + \frac{1}{2}A_{2xx}B_1^2)h^2$$
$$+ \{B_2h^{1/2} + [\frac{1}{2}\dot{B}_2 + \frac{1}{2}A_2B_{2v} + \frac{1}{2}A_{2v}B_2 + \frac{1}{4}B_{2vv}B_2^2]h^{3/2}\}\xi_2 \qquad (67)$$
$$+ \frac{1}{2}A_{2x}B_1h^{3/2}\xi_1 + \frac{1}{2}B_{2v}B_2h\xi_2^2 .$$

As before, subscripts x and v denote partial derivatives with respect to x and v respectively, and ξ_1 and ξ_2 are independent standardized Gaussian random variables. The delta function initial conditions are $x(0) = 1$, $v(0) = 1$. Monte Carlo calculations were performed using the same values of h and N as in the previous example; convergence of the statistical error is evident in Figures 12-14. Only the first four moments of $v(t)$ are shown but similar results are obtained for the moments of $x(t)$ and for the joint statistics of $x(t)$ and $v(t)$.

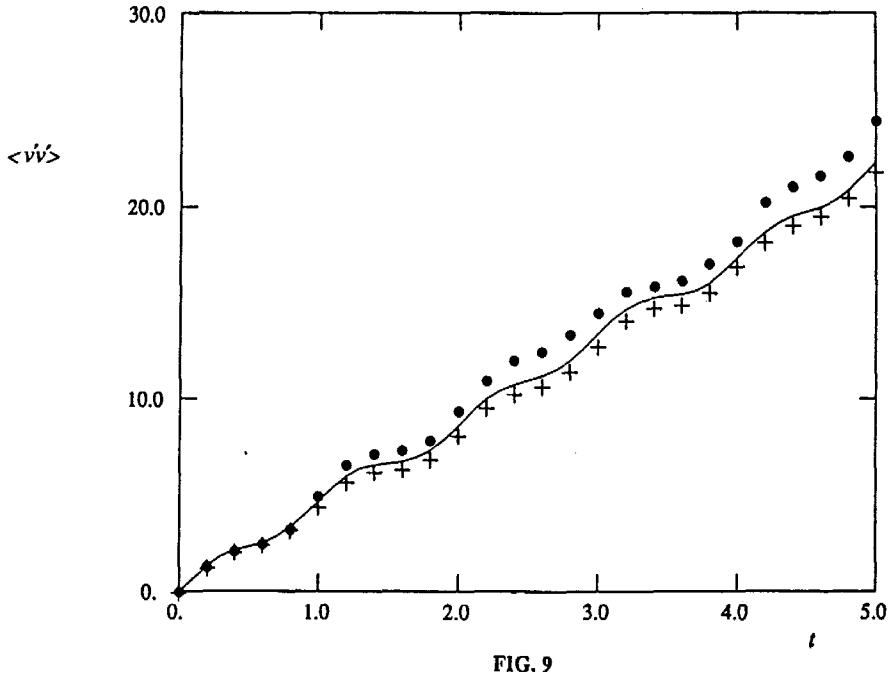

FIG. 9
Non-linear Example, Evolution of the Variance

• Monte Carlo calculation, N = 1,000
\+ Monte Carlo calculation, N = 10,000
——— spline fit through Monte Carlo calculation, N = 100,000

Taking as the "analytic solution" a Monte Carlo calculation for $h = 0.005$, $N = 1{,}000{,}000$, the second-order nature of the algorithm is checked in Figure 15 by plotting the error ε (Eq. 58) at $t = 1.0$ versus the step size h for $N = 1{,}000{,}000$. The slopes of the lines corresponding to $\langle v \rangle$, $\langle v'^2 \rangle$, and $\langle \alpha \rangle$ are equal to 2, at least for time steps large enough that the error is above the level of statistical noise ($\approx 0.1\%$) and small enough that higher-order terms are not important. Results for $\langle x'^2 \rangle$ and especially $\langle x'v' \rangle$ are, unfortunately, not as convincing. There are three sources of error in these calculations. The normalized statistical error in $\langle x'v' \rangle$ is estimated to be 0.4% (σ of Eq. 29 divided by $\langle x'v' \rangle$). This can be reduced by increasing the number of particles N. Errors due to higher-order truncated terms in h can be reduced by decreasing the time step, but it is necessary to simultaneously increase N so that discretization error can be distinguished from statistical error. Finally, more severe demands are placed on the random number generators here than in the previous examples; up to fourth moments of the Gaussian random variables affect the second moments of $\Delta \underline{u}$ for the random walk of Eq.(23),

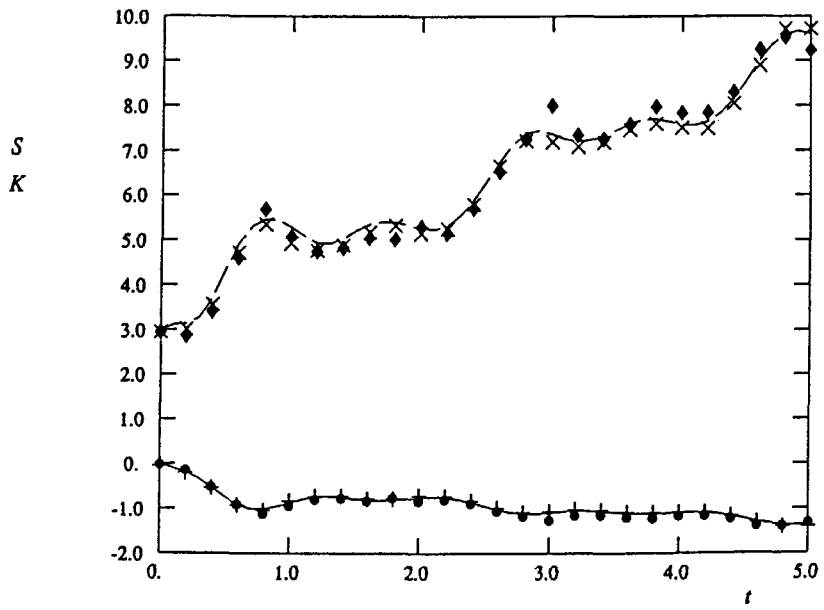

FIG. 10
Non-linear Example, Evolution of the Skewness and Kurtosis

- • skewness, Monte Carlo calculation, N = 1,000
- \+ skewness, Monte Carlo calculation, N = 10,000
- ——— skewness, spline fit through Monte Carlo calculation, N = 100,000
- • kurtosis, Monte Carlo calculation, N = 1,000
- × kurtosis, Monte Carlo calculation, N = 10,000
- — — — - kurtosis, spline fit through Monte Carlo calculation, N = 100,000

while only up to the second moments of the Gaussian random variables affect the second moments of $\Delta\underline{u}$ for the simplified random walk of Eq.(24). Because these numerical tests consume large amounts of computer time, we did not attempt to improve on the results of Figure 15.

Three sample problems have been solved using the new second-order Monte Carlo algorithm. These examples illustrate the convergence of the numerical solution as $h \to 0$ and as $N \to \infty$ and verify (in most cases) that the method is second-order in the time step even in cases where there are departures from a joint-normal pdf and the restriction of Eq.(5) does not hold. Although fixed initial conditions were used for convenience in these examples, the method is equally applicable to random initial conditions.

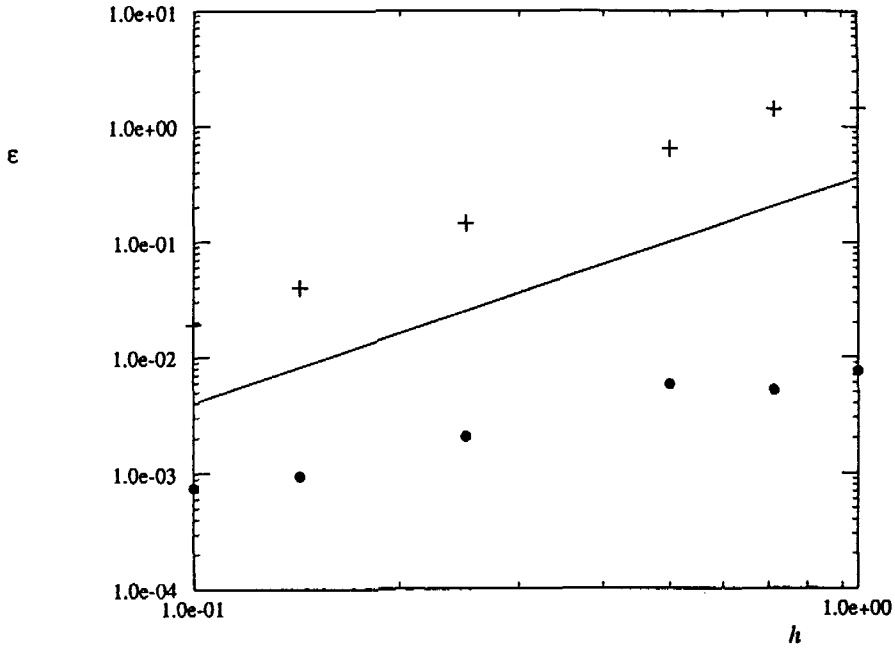

FIG. 11
Non-linear Example, Discretization Error @ t=5

• mean of v
——— reference line with slope = 2
\+ variance of v

5. CONCLUSIONS

A numerical solution algorithm applicable to the general Ito stochastic differential equation has been developed. The method is based on a Taylor series expansion of the transition joint pdf in a time increment h corresponding to Eq.(1); for a wide class of coefficient functions, this pdf is joint-normal to second order in h. By numerically integrating N sample paths, all one-time and multiple-time statistics of the solution can be extracted with a statistical error of order $N^{-\frac{1}{2}}$ and a discretization error of $O(h^3)$ on each time step. Results were presented for three sample problems, verifying the analytic error estimates and demonstrating convergence even in cases where there is a significant departure from joint-normality.

The present work is a generalization to vector-valued random processes of the method developed by Mil'shtein [24]. In working with the transition pdf instead of the sample paths themselves as was done by Rao et al. [22], we are able to extract the same statistical information from the solution via a computationally more efficient algorithm.

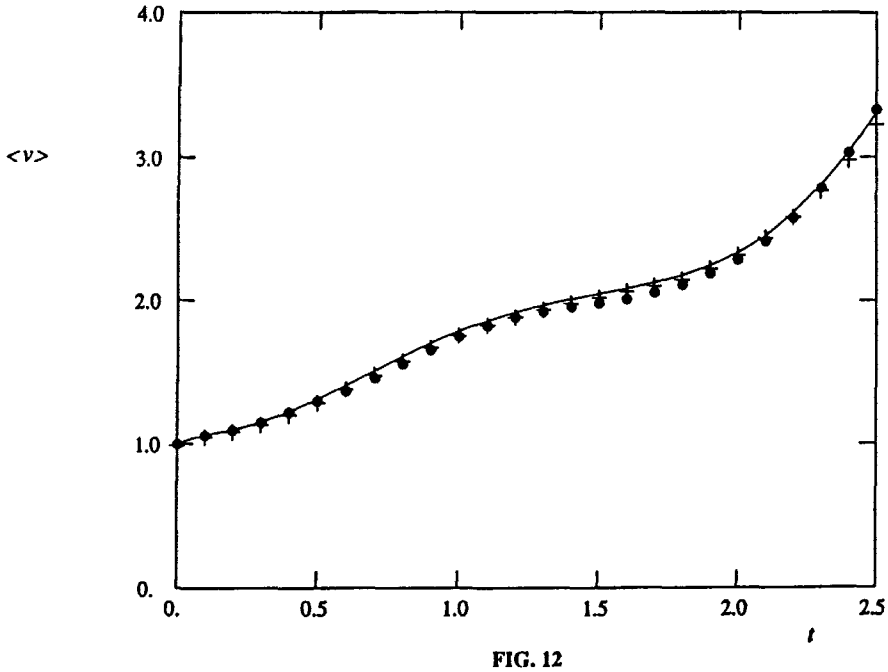

FIG. 12
Two-Dimensional Example, Evolution of Mean of v

• Monte Carlo calculation, N = 1,000
\+ Monte Carlo calculation, N = 10,000
——— spline fit through Monte Carlo calculation, N = 100,000

Compared to other currently available numerical solution techniques, the Monte Carlo method takes advantage of the Gaussian nature of the random terms in Eq.(1) and is readily applicable to problems of high dimensionality. The strength of the Monte Carlo method is that computational requirements increase only linearly with the number of independent variables; the penalty is the slow convergence of the statistical error as $N^{-1/2}$. Higher order methods in h can be derived by carrying additional terms in the Taylor series expansions [24-26]. The computational complexity of the second-order method compared to a first-order method appears to be warranted. That is, the increase in the time step allowed by the higher-order scheme compensates for the extra terms that must be calculated at each time step. In statistically stationary problems where statistics can be generated by time averaging over particle trajectories rather than by ensemble averaging over a large number of realizations, higher-order schemes may be

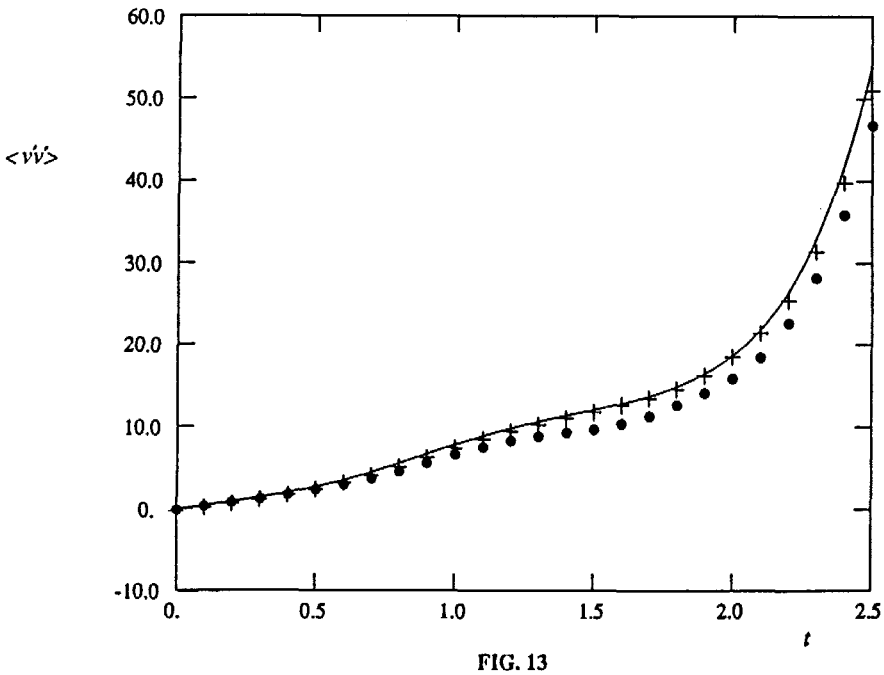

FIG. 13
Two-Dimensional Example, Evolution of Variance of v

• Monte Carlo calculation, N = 1,000
\+ Monte Carlo calculation, N = 10,000
——— spline fit through Monte Carlo calculation, N = 100,000

desirable. Otherwise, there is little incentive to extending to $O(h^3)$ or higher as the number of particles required to keep the statistical error correspondingly low would be prohibitive.

At each time step of the random walk, a vector of independent standardized Gaussian random variables is needed. A random number generator is also required for the initial condition, if random. Convergence of the numerical calculations to the analytic solution as $h \to 0$ and as $N \to \infty$ demands that the error of these pseudo-random number generators in approximating the desired distributions be negligible. No attempt to quantify this source of error has been made here. In practice, it is usually kept small compared to the discretization error and statistical error using standard computational algorithms (see Knuth [31], for example). As pointed out in Section 4.3 however, the general random walk of Eq.(23) places severe demands on the random number generators. The calculation of the random numbers can take a significant fraction of the total execution time in simple problems, about 50% for the example of Section 4.1, but

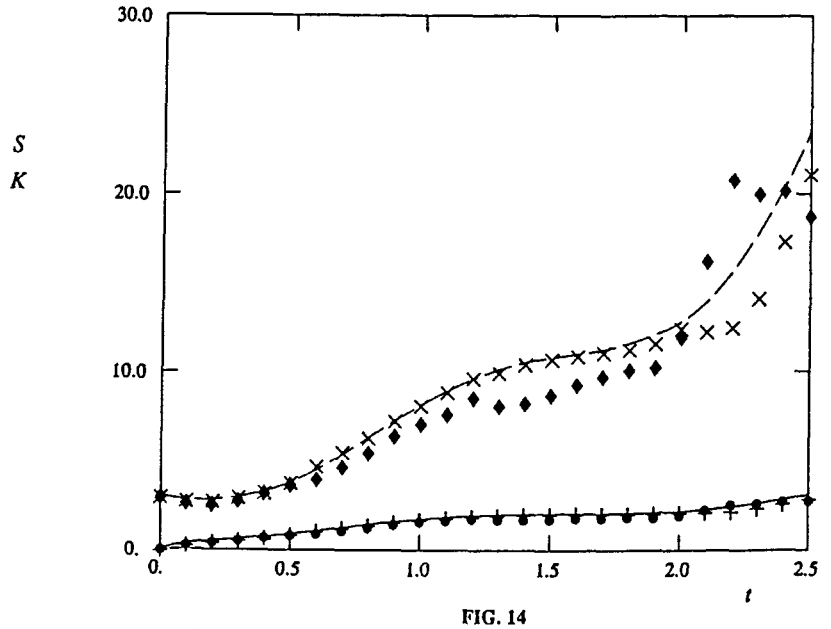

FIG. 14
Two-Dimensional Example, Evolution of Skewness and Kurtosis of v

- • skewness, Monte Carlo calculation, N = 1,000
- \+ skewness, Monte Carlo calculation, N = 10,000
- ——— skewness, spline fit through Monte Carlo calculation, N = 100,000
- ♦ kurtosis, Monte Carlo calculation, N = 1,000
- × kurtosis, Monte Carlo calculation, N = 10,000
- — — — - kurtosis, spline fit through Monte Carlo calculation, N = 100,000

this fraction decreases as the complexity of the coefficient functions increases. For the problem of Section 4.2, only 1-2% of the computational time is spent in generating the Gaussian random variables.

As a final comment, it may not always be convenient to use the difference equations with the derivatives of the coefficient functions explicitly in evidence, as in Eqs.(23) and (24). This is especially important in cases where the coefficients A_i and B_{ij} are functions of the statistics of the solution and hence are not known a priori as explicit functions of time. It is possible to overcome this difficulty by rewriting the difference equations as a predictor-corrector or Runge-Kutta scheme. Runge-Kutta methods are presented by Mil'shtein [24], by Greenside and Helfand [25,26], and by Drummond et al. [28] for the subsets of problems considered in those papers. The derivation of Runge-Kutta methods is straightforward when B_{ij} is independent of $\underline{u}$. For example, the difference equation for the sample problem of Section 4.2 (Eq. 62)

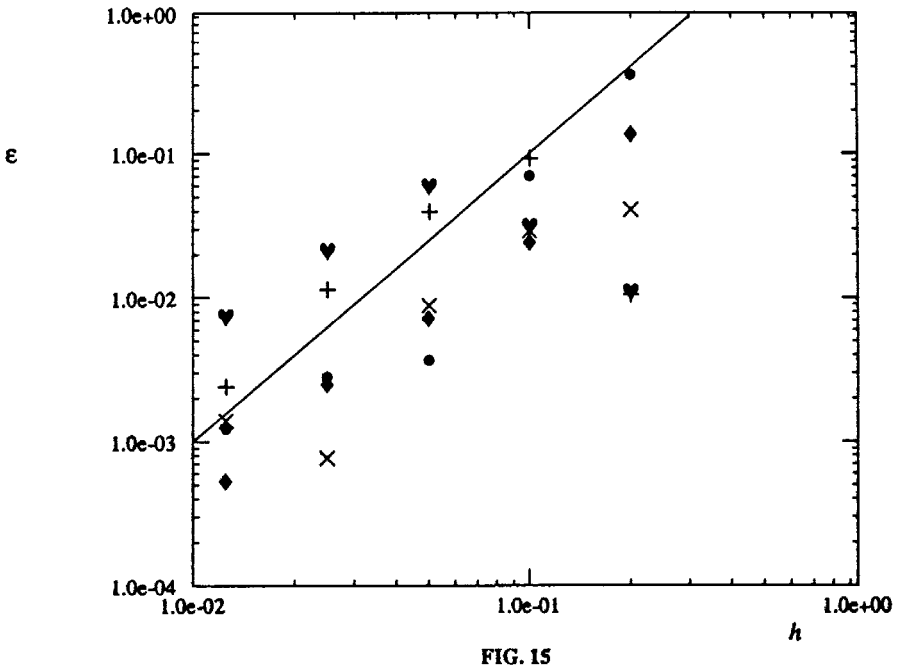

FIG. 15
Two-Dimensional Example, Discretization Error @ t=1

- mean of v
- reference line with slope = 2
- variance of v
- mean of x
- variance of x
- covariance of x and v

can be written in the alternative form

$$v^{(1)} = v^{(0)} + \alpha^{(0)}h + \beta^{(0)}h^{1/2}\xi\,, \tag{68}$$

$$v(t+h) = v^{(0)} + \frac{1}{2}(\alpha^{(0)} + \alpha^{(1)})h + \frac{1}{2}(\beta^{(0)} + \beta^{(1)})h^{1/2}\xi\,, \tag{69}$$

where quantities with a superscript "(0)" are evaluated at time level t and at $v = v(t)$, while those with a superscript "(1)" are evaluated at time level $t + h$ and at $v = v^{(1)}$. Such modifications are permissible since all that is required to retain the formal accuracy of the method is that the increment in $\underline{u}$ in the time step h has a transition pdf that corresponds to that of Eq.(A36) to second-order in h, or equivalently, that the moments of $\Delta\underline{u}$ are equal to the

moments of Eq.(A36) to second order in h. Extension to the general case is more difficult; no formulation eliminating all derivatives in Eqs.(23) or (24) is offered here. The structure of a Runge-Kutta method that eliminates all derivatives from the random walk when B_{ij} is a function of $\underline{u}$ as well as of t is a topic for further study.

ACKNOWLEDGMENTS

This work was funded in part by the General Electric Corporation Research and Development Division under the grant "Turbulent Flow Modeling". Computations supporting this research were performed on the Cornell Production Supercomputer Facility, which is supported in part by the National Science Foundation and the IBM Corporation.

REFERENCES

1. N. Wax, *Selected Papers on Noise and Stochastic Processes*, Dover Publications, New York, 1954.

2. K. Ito, *On Stochastic Differential Equations*, American Mathematical Society, New York, 1951 *(Memoirs, American Mathematical Society, No. 4)*.

3. R.L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control* 4 , 1966, 362-371.

4. L. Arnold, *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, New York, 1974.

5. Z. Schuss, *Theory and Applications of Stochastic Differential Equations*, John Wiley & Sons, New York, 1980.

6. H.J. Kushner, *Introduction to Stochasic Control*, Holt, New York, 1971.

7. A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

8. P.A. Durbin, "Stochastic differential equations and turbulent dispersion," NASA RP-1103, 1983.

9. B.G. Zhang and W.J. Padgett, "The existence and uniqueness of solutions to stochastic differential- difference equations," *Stoch. Anal. App.* 2 , 1984, 335-345.

10. D.C. Haworth and S.B. Pope, "A generalized Langevin model for turbulent flows," *Phy. Fl.*, 1985 (in print).

11. S. Karlin and H.M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York, 1981.

12. S.B. Pope, "A Lagrangian two-time probability density function equation for inhomogeneous turbulent flows," *Phys. Fl.* 26 , 1983, 3448-3450.

13. G.D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford University (Clarendon) Press, Oxford, 1978.

14. J. McKenna and J.A. Morrison, "Moments and correlation functions of solutions of a stochastic differential equation," *J. Math. Phys.* 11 , 1970, 2348-2360.

15. J. McKenna and J.A. Morrison, "Moments of solutions of a class of stochastic differential equations," *J. Math. Phys.* 12 , 1971, 2126-2136.

16. J.M. Richardson, "The application of truncated hierarchy techniques in the solution of a stochastic linear differential equation," in *Stochastic Processes in Mathematical Physics and Engineering, Proc. of Symp. in Applied Math., vol. 16,* ed. R. Bellman, American Math. Soc., Providence, 1964, 290-302.

17. W.E. Boyce, "Approximate solution of random ordinary differential equations," *Adv. App. Prob.* 10 , 1978, 172-184.

18. W.E. Kohler and W.E. Boyce, "A numerical analysis of some first order stochastic initial value problems," *SIAM J. App. Math.* 27 , 1974, 167-179.

19. M.R. Barry and W.E. Boyce, "Numerical solution of a class of random boundary value problems," *J. Math. Anal. App.* 67 , 1979, 96-119.

20. M.D. Lax and W.E. Boyce, "The method of moments for linear random initial value problems," *J. Math. Anal. App.* 53 , 1976, 111-132.

21. M.D. Lax, "Method of moments approximate solutions of random linear integral equations," *J. Math. Anal. App.* 12 , 1977, 46-55.

22. N.J. Rao, J.D. Borwankar, and D. Ramkrishna, "Numerical solution of Ito integral equations," *SIAM J. Control* 12 , 1974, 124-139.

23. G.N. Mil'shtein, "Approximate integration of stochastic differential equations," *Theory Prob. App.* 19 , 1974, 557-562.

24. G.N. Mil'shtein, "A method of second-order accuracy integration of stochastic differential equations," *Theory Prob. App.* 23 , 1978, 396-401.

25. E. Helfand, "Numerical integration of stochastic differential equations," *Bell Sys. Tech. J.* 58 , 1975, 2289-2299.

26. H.S. Greenside and E. Helfand, "Numerical integration of stochastic differential equations - II," *Bell Sys. Tech. J.* 60 , 1981, 1927-1940.

27. I.T. Drummond, S. Duane, and R.R. Horgan, "The stochastic method for numerical simulations: higher order corrections," *Nuc. Phys.* B220[FS8] , 1983, 119-136.

28. I.T. Drummond, S. Duane, and R.R. Horgan, "Scalar diffusion in simulated helical turbulence with molecular diffusivity," *J. Fluid Mech.* 138 , 1984, 75-91.

29. G. Dahlquist and A. Bjorck, *Numerical Methods,* Prentice-Hall, New York, 1974.

30. S. Karlin and H.M. Taylor, *A First Course in Stochastic Processes,* 2nd ed., Academic Press, New York, 1975.

31. D.E. Knuth, *The Art of Computer Programming, Vol. 2,* Addison-Wesley, Reading, 1981.

32. B.V. Gnedenko, *The Theory of Probability,* MIR Publishers, Moscow, 1978.

33. S.B. Pope, "Pdf methods in turbulent reactive flows," *Prog. Energy Comb. Sci.* 11 , 1985, 119-192.

34. C. de Boor, *A Practical Guide to Splines*, Springer-Verlag, New York, 1978.

35. S.J. Bean and C.P. Tsokos, "Developments in nonparametric density estimation," *Int. Stat. Review* 48 , 1980, 267-287.

APPENDIX

The evolution equation for the one-time pdf $f(\underline{V};t)$ corresponding to the Ito stochastic ordinary differential equation of Eq.(1) is given by Eq.(10):

$$\frac{\partial f}{\partial t} + \frac{\partial [f A_i(\underline{V},t)]}{\partial V_i} - \frac{1}{2}\frac{\partial^2 [f C_{ij}(\underline{V},t)]}{\partial V_i \partial V_j} = 0 , \tag{A1}$$

where C_{ij} is defined in Eq.(4). It is shown that subject to the deterministic initial conditions

$$\underline{u}(t_0) = \underline{U}^0 , \tag{A2}$$

and for a wide class of coefficient functions $A_i(\underline{V},t)$ and $C_{ij}(\underline{V},t)$ (restrictions on the coefficient functions are introduced in the course of this proof), Eq.(A1) causes $f(\underline{V};t)$ for $t \geq t_0$ to evolve as a joint-normal distribution to second order in $h \equiv t-t_0$:

$$f(\underline{V};t) = f^{JN}(\underline{V};t) + O(h^3) . \tag{A3}$$

In Eq.(A3), $f^{JN}(\cdot)$ denotes a joint-normal distribution. There are four steps in this proof:

1) Equation (A1) is Fourier transformed to obtain the evolution equation for the characteristic function $g(\underline{W},t)$;

2) the evolution equation for the logarithm of the characteristic function $q(\underline{W},t) \equiv \ln[g(\underline{W},t)]$ is derived;

3) $q(\underline{W},t)$ is expanded in a Taylor series about the initial state of Eq.(A2); and

4) it is shown that $q(\underline{W},t)$ corresponds to a joint-normal distribution to $O(h^2)$.

For arbitrary coefficient functions, the pdf is not necessarily joint-normal to second order in h. The extension to this general case is also discussed.

1) Fourier Transform

First, we Fourier transform Eq.(A1) in $\underline{V}$:

$$\mathbf{F}(f) \equiv \int e^{i\underline{V}\cdot\underline{W}} f(\underline{V};t) d\underline{V} = g(\underline{W},t) , \tag{A4}$$

$$\mathbf{F}^{-1}(g) \equiv \left(\frac{1}{2\pi}\right)^M \int e^{-i\underline{V}\cdot\underline{W}} g(\underline{W},t) d\underline{W} = f(\underline{V};t) . \tag{A5}$$

In these equations, $\underline{W}$ denotes the transform variable and $i \equiv \sqrt{-1}$. $d\underline{V} = dV_1 dV_2 ... dV_M$ and similarly for $d\underline{W}$, where M is the dimension of vectors $\underline{V}$ and $\underline{W}$. The integrations are over the entire M dimensional space. By the definition of the Fourier transform, Eqs.(A4) and (A5):

$$\mathbf{F}\left(\frac{\partial^n f}{\partial V_{j_1} \partial V_{j_2} \cdots \partial V_{j_n}}\right) = (-i)^n W_{j_1} W_{j_2} \cdots W_{j_n} g \,, \tag{A6}$$

$$\mathbf{F}(f V_{j_1} V_{j_2} \cdots V_{j_n}) = (-i)^n \frac{\partial^n g}{\partial W_{j_1} \partial W_{j_2} \cdots \partial W_{j_n}} \,. \tag{A7}$$

We multiply Eq.(A1) by $e^{i\underline{V}\cdot\underline{W}}$ and integrate over all $\underline{V}$ using property (A6) to obtain

$$\frac{\partial g}{\partial t} - iW_j \mathbf{F}(fA_j) + \frac{1}{2} W_j W_k \mathbf{F}(fC_{jk}) = 0 \,. \tag{A8}$$

Next the coefficient functions are expanded in a Taylor series about $\underline{U}^0$:

$$A_j(\underline{V},t) = A_j^0(t) + A_{j,l}^0(t)(V_l - U_l^0) + \frac{1}{2} A_{j,lm}^0(t)(V_l - U_l^0)(V_m - U_m^0) + \cdots \,, \tag{A9}$$

$$C_{jk}(\underline{V},t) = C_{jk}^0(t) + C_{jk,l}^0(t)(V_l - U_l^0) + \frac{1}{2} C_{jk,lm}^0(t)(V_l - U_l^0)(V_m - U_m^0) + \cdots \,. \tag{A10}$$

A superscript "0" means that the function is evaluated at $\underline{U}^0$. The truncation after the second derivatives is justified below. Using Eqs.(A9), (A10), and property (A7), Eq.(A8) becomes

$$\begin{aligned} &\frac{\partial g}{\partial t} - iW_j\left[\left(A_j^0 - A_{j,l}^0 U_l^0 + \frac{1}{2} A_{j,lm}^0 U_l^0 U_m^0\right) g - i\left(A_{j,l}^0 - A_{j,lm}^0 U_m^0\right)\frac{\partial g}{\partial W_l}\right. \\ &\quad \left. - \frac{1}{2} A_{j,lm}^0 \frac{\partial^2 g}{\partial W_l \partial W_m} + \cdots\right\} \\ &\quad + \frac{1}{2} W_j W_k\left[\left(C_{jk}^0 - C_{jk,l}^0 U_l^0 + \frac{1}{2} C_{jk,lm}^0 U_l^0 U_m^0\right) g - i\left(C_{jk,l}^0 - C_{jk,lm}^0 U_m^0\right)\frac{\partial g}{\partial W_l}\right. \\ &\quad \left. - \frac{1}{2} C_{jk,lm}^0 \frac{\partial^2 g}{\partial W_l \partial W_m} + \cdots\right\} + \cdots = 0 \,, \end{aligned} \tag{A11}$$

where the time argument (t) of the coefficient functions and their derivatives has been omitted for convenience. $g(\underline{W},t)$ is the characteristic function: it is the Fourier transform of the pdf $f(\underline{V};t)$ with respect to $\underline{V}$ [32].

2) Logarithm of Characteristic Function

The function $q(\underline{W},t)$ denotes the natural logarithm of the characteristic function:

$$q(\underline{W},t) \equiv \ln[g(\underline{W},t)] \,. \tag{A12}$$

To obtain an evolution equation for this function, we divide Eq.(A11) by $g(\underline{W},t)$ and make use of the identities

$$\frac{dg}{g} = d[\ln(g)] = dq \,, \tag{A13}$$

$$\frac{1}{g}\frac{\partial^2 g}{\partial W_j \partial W_k} = \frac{\partial^2 q}{\partial W_j \partial W_k} + \frac{\partial q}{\partial W_j}\frac{\partial q}{\partial W_k} \,. \tag{A14}$$

The resulting evolution equation for $q(\underline{W},t)$ is

$$\begin{aligned}
&\frac{\partial q}{\partial t} - iW_j\{(A_j^0 - A_{j,l}^0 U_l^0 + \frac{1}{2}A_{j,lm}^0 U_l^0 U_m^0) - i(A_{j,l}^0 - A_{j,lm}^0 U_m^0)\frac{\partial q}{\partial W_l} \\
&\quad - \frac{1}{2}A_{j,lm}^0(\frac{\partial^2 q}{\partial W_l \partial W_m} + \frac{\partial q}{\partial W_l}\frac{\partial q}{\partial W_m}) + \cdots\} \\
&\quad + \frac{1}{2}W_j W_k\{(C_{jk}^0 - C_{jk,l}^0 U_l^0 + \frac{1}{2}C_{jk,lm}^0 U_l^0 U_m^0) - i(C_{jk,l}^0 - C_{jk,lm}^0 U_m^0)\frac{\partial q}{\partial W_l} \\
&\quad - \frac{1}{2}C_{jk,lm}^0(\frac{\partial^2 q}{\partial W_l \partial W_m} + \frac{\partial q}{\partial W_l}\frac{\partial q}{\partial W_m}) + \cdots\} + \cdots = 0 .
\end{aligned} \tag{A15}$$

3) Taylor Series Expansion

Equation (A15) is used as the basis for a Taylor series expansion of $q(\underline{W},t)$ about the deterministic initial state of Eq.(A2). Retaining terms to $O(h^2)$:

$$q(\underline{W},t_0+h) = q^0(\underline{W}) + \dot{q}^0(\underline{W})h + \frac{1}{2}\ddot{q}^0(\underline{W})h^2 + O(h^3) , \tag{A16}$$

where "$\cdot$" denotes a time derivative and superscript "0" means that the function is evaluated at time t_0. First, we evaluate $\ddot{q}(\underline{W},t)$ from Eq.(A15):

$$\begin{aligned}
&\frac{\partial^2 q}{\partial t^2} - iW_j\{(\dot{A}_j^0 - \dot{A}_{j,l}^0 U_l^0 + \frac{1}{2}\dot{A}_{j,lm}^0 U_l^0 U_m^0) - i(\dot{A}_{j,l}^0 - \dot{A}_{j,lm}^0 U_m^0)\frac{\partial q}{\partial W_l} \\
&\quad - i(A_{j,l}^0 - A_{j,lm}^0 U_m^0)\frac{\partial \dot{q}}{\partial W_l} - \frac{1}{2}\dot{A}_{j,lm}^0(\frac{\partial^2 q}{\partial W_l \partial W_m} + \frac{\partial q}{\partial W_l}\frac{\partial q}{\partial W_m}) \\
&\quad - \frac{1}{2}A_{j,lm}^0(\frac{\partial^2 \dot{q}}{\partial W_l \partial W_m} + 2\frac{\partial \dot{q}}{\partial W_l}\frac{\partial q}{\partial W_m}) + \cdots\} \\
&\quad + \frac{1}{2}W_j W_k\{(\dot{C}_{jk}^0 - \dot{C}_{jk,l}^0 U_l^0 + \frac{1}{2}\dot{C}_{jk,lm}^0 U_l^0 U_m^0) - i(\dot{C}_{jk,l}^0 - \dot{C}_{jk,lm}^0 U_m^0)\frac{\partial q}{\partial W_l} \\
&\quad - i(C_{jk,l}^0 - C_{jk,lm}^0 U_m^0)\frac{\partial \dot{q}}{\partial W_l} - \frac{1}{2}\dot{C}_{jk,lm}^0(\frac{\partial^2 q}{\partial W_l \partial W_m} + \frac{\partial q}{\partial W_l}\frac{\partial q}{\partial W_m}) \\
&\quad - \frac{1}{2}C_{jk,lm}^0(\frac{\partial^2 \dot{q}}{\partial W_l \partial W_m} + 2\frac{\partial \dot{q}}{\partial W_l}\frac{\partial q}{\partial W_m}) + \cdots\} + \cdots = 0 .
\end{aligned} \tag{A17}$$

The deterministic initial conditions of Eq.(A2) correspond to a delta function initial pdf:

$$f(\underline{V};t_0) = \delta(\underline{V}-\underline{U}^0) . \tag{A18}$$

The corresponding initial $q(\underline{W},t)$ is found from Eqs.(A4) and (A12):

$$q^0(\underline{W}) = iW_j U_j^0 . \tag{A19}$$

It follows that

$$\frac{\partial q^0}{\partial W_l} = iU_l^0 , \tag{A20}$$

$$\frac{\partial^n q^0}{\partial W_{j_1} \cdots \partial W_{j_n}} = 0 \text{ for } n \geq 2 . \tag{A21}$$

Equation (A15) then yields

$$q^0(\underline{W}) = iW_j A_j^0 - \frac{1}{2} W_j W_k C_{jk}^0 , \tag{A22}$$

so that

$$\frac{\partial q^0}{\partial W_l} = iA_l^0 - C_{jl}^0 W_j , \tag{A23}$$

$$\frac{\partial^2 q^0}{\partial W_l \partial W_m} = -C_{ml}^0 , \tag{A24}$$

$$\frac{\partial^n q^0}{\partial W_{j_1} \partial W_{j_1} \cdots \partial W_{j_n}} = 0 \quad \text{for } n \geq 3 . \tag{A25}$$

The superscript "0" in the coefficient functions now denotes that the function is evaluated at time t_0 as well as at $\underline{U}^0$. From Eq.(A17)

$$\begin{aligned} \ddot{q}^0(\underline{W}) = {} & iW_j\{\dot{A}_j^0 + A_{j,l}^0 A_l^0 + \frac{1}{2} A_{j,lm}^0 C_{ml}^0\} \\ & - \frac{1}{2} W_j W_k \{\dot{C}_{jk}^0 + C_{jk,l}^0 A_l^0 + 2A_{j,l}^0 C_{kl}^0 + \frac{1}{2} C_{ml}^0 C_{jk,lm}^0\} \\ & - iW_j W_k W_p \{\frac{1}{2} C_{jk,l}^0 C_{pl}^0\} . \end{aligned} \tag{A26}$$

The second-order approximation to $q(\underline{W},t)$ is constructed from Eq.(A16) with Eqs.(A19), (A22), and (A26):

$$\begin{aligned} q(\underline{W},t_0+h) = {} & iW_j\{U_j^0 + A_j^0 h + \frac{1}{2}(\dot{A}_j^0 + A_{j,l}^0 A_l^0 + \frac{1}{2} A_{j,lm}^0 C_{lm}^0)h^2 + O(h^3)\} \\ & - \frac{1}{2} W_j W_k \{C_{jk}^0 h + \frac{1}{2}(\dot{C}_{jk}^0 + C_{jk,l}^0 A_l^0 + A_{j,l}^0 C_{kl}^0 + A_{k,l}^0 C_{jl}^0 + \frac{1}{2} C_{jk,lm}^0 C_{lm}^0)h^2 + O(h^3)\} \\ & - iW_j W_k W_p \{\frac{1}{4} C_{jk,l}^0 C_{pl}^0 h^2 + O(h^3)\} + O(h^3) . \end{aligned} \tag{A27}$$

In the derivation of this equation, it is implicitly assumed that the coefficient functions possess bounded continuous partial derivatives up to at least the second order with respect to the components of $\underline{\mu}$ and to at least the first order with respect to t.

4) Joint-Normal Characteristic Function

The logarithm of the characteristic function corresponding to a joint-normal random vector $\underline{\Phi}$ is [32]

$$q(\underline{\Psi}) = i\Psi_\alpha<\Phi_\alpha> - \frac{1}{2}\Psi_\alpha\Psi_\beta<\Phi_\alpha'\Phi_\beta'> , \tag{A28}$$

where a prime denotes a fluctuation about the mean

$$\Phi_\alpha' \equiv \Phi_\alpha - <\Phi_\alpha> , \tag{A29}$$

and $<\Phi'_\alpha\Phi'_\beta>$ is a symmetric positive semi-definite matrix (to satisfy realizability). It may be seen on comparison of Eqs.(A27) and (A28) that to first order in h, $q(\underline{W},t_0+h)$ is joint-normal with means $U_j^0 + A_j^0 h$ and covariances $C_{jk}^0 h$. Since C_{jk} is symmetric and positive semi-definite (Eq. 4), $C_{jk}^0 h$ corresponds to a valid covariance matrix.

To $O(h^2)$, Eq.(A27) fits the general form of Eq.(A28) as long as $C_{jk}(\underline{u},t)$ satisfies the restriction

$$C_{jk,l}C_{pl} + C_{kp,l}C_{jl} + C_{pj,l}C_{kl} = 0 . \tag{A30}$$

The superscript "0" has been dropped since this constraint must hold for arbitrary t_0. It is the symmetry of $W_jW_kW_p$ with respect to all permutations of its indices that leads to this condition. For an M-dimensional third-rank tensor T_{jkp} that is symmetric in its first two indices to satisfy the condition $T_{jkp} + T_{kpj} + T_{pjk} = 0$, the M diagonal components must be equal to zero and $(M^3 - M)/3$ of the off-diagonal components can be chosen arbitrarily, the remainder being fixed by symmetry and by this condition. Two specific examples of coefficients B_{ij} that satisfy Eq.(A30) are given in Section 1.

It is not obvious that the coefficient of W_jW_k in Eq.(A27) corresponds to a valid covariance matrix when terms of second or higher order in h are retained. The coefficient is clearly symmetric to $O(h^2)$, and as $h \to 0$ at least, is positive semi-definite as well. However, for non-zero h, the truncated coefficient is not necessarily positive semi-definite. We argue that since Eq.(A27) is derived from a valid pdf evolution equation (Eq. A1) subject to valid initial conditions (Eq. A2), then for some choice of the $O(h^3)$ term, this coefficient can be made positive semi-definite. One such choice results from continuing the Taylor series expansion of Eq.(A16) to higher order in h but this imposes stricter differentiability requirements on the coefficient functions than are necessary; any $O(h^3)$ term that ensures realizability is satisfactory. This is sufficient to conclude that Eq.(A27) corresponds to a joint-normal distribution to second-order in h. In Section 4.1, an example is given where the covariance matrix is made positive semi-definite by carrying a term of $O(h^3)$ that is different from the third-order term in the Taylor series expansion.

It has been proven that subject to the deterministic initial conditions of Eq.(A2) and for a wide class of coefficient functions A_i and C_{ij}, Eq.(A1) causes the pdf $f(\underline{V};t_0+h)$ to evolve as a joint-normal distribution to second-order in h. The mean vector and covariance matrix defining the joint-normal distribution are found by matching Eq.(A27) with Eq.(A28):

$$<u_j> = U_j^0 + A_j^0 h + \frac{1}{2}(\dot{A}_j^0 + A_{j,l}^0 A_l^0 + \frac{1}{2}A_{j,lm}^0 C_{lm}^0)h^2 + O(h^3) , \tag{A31}$$

$$<u'_j u'_k> = C_{jk}^0 h + \frac{1}{2}(\dot{C}_{jk}^0 + C_{jk,l}^0 A_l^0 + A_{j,l}^0 C_{kl}^0 + A_{k,l}^0 C_{jl}^0 + \frac{1}{2}C_{jk,lm}^0 C_{lm}^0)h^2 + O(h^3) . \tag{A32}$$

It can be shown that retaining higher-order terms in the Taylor series expansions of the coefficient functions (Eqs. A9 and A10) does not change this conclusion: q^0, $\dot{q}^0$, and $\ddot{q}^0$ of Eqs.(A19), (A22), and (A26) are unaffected by the higher-order terms.

It is convenient to express the results in terms of the increments

$$\Delta \underline{u} \equiv \underline{u}(t_0+h) - \underline{u}(t_0) \; . \tag{A33}$$

Since the initial conditions $\underline{u}(t_0)$ are deterministic (Eq. A2), we conclude immediately that the pdf of the increment $\Delta \underline{u}$ is joint-normal to second order in h with means and covariances

$$<\Delta u_j> = A_j^0 h + \frac{1}{2}(\dot{A}_j^0 + A_{j,i}^0 A_i^0 + \frac{1}{2}A_{j,lm}^0 C_{lm}^0)h^2 + O(h^3) \; , \tag{A34}$$

$$<\Delta u'_j \Delta u'_k> = C_{jk}^0 h + \frac{1}{2}(\dot{C}_{jk}^0 + C_{jk,i}^0 A_i^0 + A_{j,i}^0 C_{ki}^0 + A_{k,i}^0 C_{ji}^0 + \frac{1}{2}C_{jk,lm}^0 C_{lm}^0)h^2 + O(h^3) \; . \tag{A35}$$

Extension to the case where the constraint of Eq.(A30) is not satisfied is not immediate. Equation (A27) then corresponds to no known distribution. In principle, the corresponding pdf can be found by taking the inverse Fourier transform but this is likely to be of little use: the truncated pdf will almost certainly not be realizable in the sense of Bochner's theorem [32]. (In the case where the cubic term in $\underline{W}$ of Eq.(A27) is zero, Bochner's theorem reduces to the condition that the coefficient of the quadratic term in $\underline{W}$ be symmetric and positive semi-definite, as discussed above). We can no longer exploit the features of the joint-normal pdf in devising a numerical scheme, i.e., that any linear combination of joint-normal random variables is joint-normal, and that the joint-normal pdf is parameterized by its first two moments (see Section 3). One approach is to find a realizable random variable whose characteristic function agrees with that of Eq.(A27) to $O(h^2)$. Motivated by the work of Mil'shtein [24], we construct a non-linear combination of Gaussian random variables that yields a realizable random variable with the desired distribution in Section 3.

To expedite this, the characteristic function of the increment $\Delta \underline{u}$ is expanded in powers of $i\underline{W}$. This brings the moments of $\Delta \underline{u}$ explicitly into evidence, since the coefficient of the term $i^n W_{j_1} W_{j_2} \cdots W_{j_n}$ in the power series is the n^{th} moment of $\Delta \underline{u}$ divided by $n!$, $<\Delta u_{j_1} \Delta u_{j_2} \cdots \Delta u_{j_n}>/n!$ [24]. The characteristic function of the increment, $g(\underline{W},h)$, is equivalent to $g(\underline{W},t_0+h)$ (Eq. A4) if $\underline{U}^0$ of Eq.(A2) is equal to zero. Expanding the characteristic function $g(\underline{W},h)$ in this way yields

$$\begin{aligned} g(\underline{W},h) = 1 &+ iW_j\{A_j^0 h + \frac{1}{2}(\dot{A}_j^0 + A_{j,p}^0 A_p^0 + \frac{1}{2}A_{j,pq}^0 C_{pq}^0)h^2\} \\ &- \frac{1}{2}W_j W_k\{C_{jk}^0 h + (A_j^0 A_k^0 + \frac{1}{2}\dot{C}_{jk}^0 + \frac{1}{2}C_{jk,p}^0 A_p^0 + \frac{1}{2}A_{j,p}^0 C_{kp}^0 + \frac{1}{2}A_{k,p}^0 C_{jp}^0 + \frac{1}{4}C_{jk,pq}^0 C_{pq}^0)h^2\} \\ &- \frac{1}{6}iW_j W_k W_l\{[A_j^0 C_{kl}^0 + A_k^0 C_{jl}^0 + A_l^0 C_{jk}^0 + \frac{1}{2}(C_{jk,p}^0 C_{lp}^0 + C_{jl,p}^0 C_{kp}^0 + C_{kl,p}^0 C_{jp}^0)]h^2\} \\ &+ \frac{1}{24}W_j W_k W_l W_m\{(C_{jk}^0 C_{lm}^0 + C_{jl}^0 C_{km}^0 + C_{jm}^0 C_{kl}^0)h^2\} + O(h^3) \; . \end{aligned} \tag{A36}$$

This follows from Eq.(A27) since $q(\underline{W},t) \equiv \ln[g(\underline{W},t)]$ (Eq. A12).