JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 18, No. 1, January–February 1995

Runge-Kutta Algorithm for the Numerical Integration of Stochastic Differential Equations

N. Jeremy Kasdin*
Stanford University, Stanford, California 94305-4085

This paper presents a new Runge-Kutta (RK) algorithm for the numerical integration of stochastic differential equations. These equations occur frequently as a description of many mechanical, aerospace, and electrical systems. They also form the basis of modern control design using the linear quadratic regulator/Gaussian (LQR/LQG) method. It is convenient, and common practice, to numerically simulate these equations to generate sample random processes that approximate a solution of the system (often called Monte Carlo simulations). It is shown in the paper that the standard deterministic solution techniques are inaccurate and can result in sample sequences with covariances not representative of the correct solution of the original differential equation. A new set of coefficients is derived for a RK-type solution to these equations. Examples are presented to show the improvement in mean-square performance.

I. Introduction

CONSIDER for simulation the Ito stochastic differential equation:

$$\dot{\xi}(t) = F(\xi(t), t) + G(\xi(t), t)w(t) \tag{1}$$

where $w(t)$ is a vector of Gaussian, white-noise random variables with autocorrelation matrix $Q(t)\,\delta(\tau)$ [where $Q(t)$ is diagonal and equal to the spectral density of the white-noise random variables], $\xi(t)$ is the $n \times 1$ state vector, and $F(\xi(t), t)$ and $G(\xi(t), t)$ are the dynamics and input distribution matrix functions, respectively.

It is common practice in design and evaluation of control systems to numerically simulate these equations, that is, to produce a series of sample sequences $\{x_k\}$ with appropriate mean-square properties as a simulation of a typical member of the process ensemble, $\xi(t)$. This is particularly true when applying linear quadratic regulator/Gaussian (LQR/LQG) design techniques, and Eq. (1) is either inherently linear or linearized about a nominal operating point:

$$\dot{\xi}(t) = F(t)\xi(t) + G(t)w(t) \tag{2}$$

Although this linear case provides for an exact solution of Eq. (1),[1-7] and thus a simple set of difference equations that are easily computed, it is very often more convenient to numerically approximate a sample solution of Eq. (2), either because the large size of the system precludes easy evaluation of the exact solution or because the linear model is itself an approximation and must be compared against a more exact nonlinear representation.

In recent years, a large number of commercial computer packages have become available for simulating complex systems, some via block diagram algebra (e.g., Simulink, MatrixX, ACSL). After reduction of the block diagram to a set of describing differential state equations, these programs provide a number of standard numerical integration routines for approximating the solution (such as the variable-step, variable-order Runge-Kutta-Fehlberg algorithm in Ref. 8). These programs need to be used with great care, however, when treating stochastic problems as the standard integration methods employed can incur significant errors. When the inhomogeneous part of Eq. (1) is used as an input to the simulation, this white-noise process is approximated by a zero-order hold. There is also no clear method for selecting the covariance of this input noise sequence. These problems, with potential solutions, are expanded upon below.

Received Oct. 31, 1992; revision received Aug. 27, 1993; accepted for publication June 6, 1994.

*Project Manager and Chief Systems Engineer, Gravity Probe B Science Mission, W. W. Hansen Experimental Physics Laboratory.

In this paper a new Runge-Kutta (RK) algorithm is derived for the integration of stochastic differential equations that is shown to improve the accuracy of stochastic simulations over the traditional approaches. The next section reviews some numerical analysis material and describes the proper criterion to use for deriving the RK method. This is followed by the presentation of third- and fourth-order RK methods for linear time-invariant and time-varying differential equations. A brief discussion of the application to nonlinear equations is then presented. Lastly, a number of examples show the improvement in accuracy using the new coefficients.

II. Definitions

The key issue in developing a method for the discrete simulation of stochastic processes represented by Eq. (1) or (2) is finding an evaluation criterion to use in determining the accuracy of the simulation. Although it is very common to use frequency-domain measures such as the spectral density, a number of errors can be incurred.[9,10] Rather, a better approach for time-invariant systems is to require that the discrete process have an autocorrelation function in the steady state that exactly samples that of the continuous process.[9,10] This can be troublesome, though, for transient problems or time-varying and nonlinear equations where it is difficult (if not impossible) to define an autocorrelation function. Instead, the simulation criterion is limited to the requirement that the discrete, time-varying covariance sample the continuous covariance of the solution to Eq. (1). This is stated formally as Definition 1.

Definition 1. A zero-mean, discrete stochastic process $\boldsymbol{x}_k$ is said to simulate the continuous stochastic process given by Eq. (1), $\xi(t)$, if the discrete covariance matrix P_k is equal to samples of the continuous covariance matrix $P(t)$ within the error of the integration routine. That is,

$$P_k = P(t_k) + \mathcal{O}(h^{n+1}) \tag{3}$$

where the covariances are defined by the expected-value operations

$$P(t) = E\{\xi(t)\xi^T(t)\} \qquad P_k = E\left\{\boldsymbol{x}_k\boldsymbol{x}_k^T\right\}$$

In order for this definition to be valuable, it is necessary to have an equation for $P(t)$ given only Eq. (1). Such an equation is straightforward for the linear case of Eq. (2)[1,2]:

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) \tag{4}$$

As mentioned, it is possible to solve the linear case exactly, resulting in a sequence $\boldsymbol{x}_k$ satisfying

$$\boldsymbol{x}_{k+1} = \Phi_k\boldsymbol{x}_k + \boldsymbol{w}_k \tag{5}$$

where $\Phi(t)$ is the state transition matrix associated with Eq. (1), $\boldsymbol{w}_k$ is an independent and identically distributed (i.i.d.) vector random process with covariance matrix Q_d, and $\boldsymbol{x}_k$ has a known covariance matrix that satisfies Definition 1. This is a well-known result the details of which can be found in Refs. 1, 2, and 9. The resulting discrete covariance equation for $\boldsymbol{x}_k$ is

$$P_d(t_{k+1}) = \Phi(t_k)P_d(t_k)\Phi^T(t_k) + Q_d \tag{6}$$

It is important to note from Eqs. (5) and (6) that the exact solution provides a very precise alternative to numerical approximation techniques. By solving for the state transition matrix and the full-rank discrete covariance matrix Q_d (see Refs. 1, 2, 9, and 10), the solution can be simulated with very high accuracy by simply computing the recursive discrete equations (5) with a full vector of process noise. It is essential to use the full covariance solution as it accounts for the mixing of the noise among all the states during the sample interval. A particularly useful algorithm for computing Q_d is given by Van Loen.[11] Although for linear systems this approach benefits from high accuracy, it does require a matrix description of the full system equations and possibly a very large matrix exponential to find the solution. The solution is therefore not easily modified for changing parameters, and it cannot be applied to nonlinear systems of equations. The RK numerical approach thus has some advantages even for linear systems and is commonly used.

It is more difficult to derive the covariance equation for the general nonlinear case. The probability density function $p(\xi, t)$ for $\xi(t)$ defined by Eq. (1) is a solution of the n-dimensional Fokker-Planck equation[4]

$$\frac{\partial p(\xi, t \mid \xi_0, t_0)}{\partial t} = -\sum_{j=1}^{n} \frac{\partial}{\partial \xi_f}[p_j(\xi, t)F] + \sum_{i,j=1}^{n} \frac{\partial^2}{\partial \xi_i \, \partial \xi_f}\left[(GQG^T)_{ij}p\right] \tag{7a}$$

or, written in matrix form,[12]

$$\frac{\partial p(\xi, t \mid \xi_0, t_0)}{\partial t} = -\frac{\partial p}{\partial \xi}F - p\left[\operatorname{tr}\left(\frac{\partial F}{\partial \xi}\right)\right] + \frac{1}{2}\operatorname{tr}\left(\frac{\partial^2}{\partial \xi^2}(GQG^T p)\right) \tag{7b}$$

Because Eq. (1) is Markovian, the density function at time t depends only upon the state at a single previous time, ξ_0 and t_0.

The Fokker-Planck equation can be used to derive an expression for the moment of any general function $H(\xi(t), t)$.[4] This is then used to find the covariance equation

$$\dot{P}(t) = E\{F(\xi(t), t)\xi^T(t)\} + E\{\xi(t)F^T(\xi(t), t)\} + E\{G(\xi(t), t)Q(t)G^T(\xi(t), t)\} \tag{8}$$

Unlike the linear case, it is not possible to solve these expectations without explicit knowledge of the functions F and G.

Given this criterion in Definition 1, the goal is to find a numerical integration routine for approximating a sample solution to Eq. (1) [or Eq. (2)]. There are two questions to answer in doing this: 1) how to evaluate the effectiveness of the method and, if possible, perform step-size control, and 2) how to pick a method and the mean and variance of the random input sequence to satisfy the criterion determined in 1. (See, e.g., Ref. 7.) Question 1 is partially answered by Definition 1 above. The remainder of this paper derives an answer to question 2.

The integration problem has often been simplified in two ways. Sometimes, a simple, first-order Euler integration is performed, because the properties of this integrator are easily examined and the driving noise covariance matrix is easily computed to be $Q_d = Q/h$. More frequently, a higher order method is used but with only a single call to the input noise generator per step. That is, although there may be numerous function evaluations to compute a single integration step [$F(\xi, t)$ and $G(\xi, t)$ above], the input noise is only computed once at the beginning of the step and given the Euler value Q/h. This is equivalent to a zero-order hold on the input noise and is particularly convenient when using commercial simulation software that allows external inputs to be applied to a block diagram simulation. This second approach is referred to as the "classical" method throughout this paper. It will be shown that, independent of the order of the integration method, accuracy is lost by using this approach for the input noise.

The problem of selecting a method is usually avoided in the same manner—the input noise is passed through a zero-order hold and standard integration packages are used. Numerical integration methods are normally grouped into two categories: single-step and multistep methods. The most common single-step method by far is the RK integrator. Most implementations of single-step methods, however, allow for variable-step-size selection. That is, at each step of the integration the step size h is modified based upon a local error estimate. Such techniques are inappropriate for stochastic systems, though, as it is impossible to determine a local error based upon a single member of the ensemble of random sequences. The correct error estimate is for the covariance of the solution.

The second category of methods, multistep methods, predicts the solution at step k based upon the value not only at the previous step, $k-1$, but some j earlier times. The most common of these are the predictor-corrector methods.[13–15] Multistep methods normally achieve larger step sizes by using smoothness assumptions of the solution to fit rational polynomials to previous steps (similar to Simpson's method for function integration). As a result, they are also inappropriate for stochastic integration, as the solution is not well described by a low-order rational (or analytic) function.

Thus, single-step, fixed-step-size methods are the best approach for simulating stochastic differential equations. Again, the most common integrators of this class are RK algorithms. Although there are many different RK-type routines of different orders available, the general deterministic algorithm is given as follows.

If $\boldsymbol{x}(t)$ is the solution of the differential equation

$$\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}, t) \tag{9}$$

then $\boldsymbol{x}(t)$ is approximated at time t_{k+1}, given its value at t_k, by the following set of equations:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha_1\boldsymbol{k}_1 + \alpha_2\boldsymbol{k}_2 + \cdots + \alpha_4\boldsymbol{k}_4 \\ \boldsymbol{k}_1 &= hf(t_k, \boldsymbol{x}_k) \\ \boldsymbol{k}_j &= hf\left(t_k + c_j h, \boldsymbol{x}_k + \sum_{i=1}^{j-1} a_{ji}\boldsymbol{k}_i\right) \end{aligned} \tag{10}$$

where h is the time step size.

These equations are an order-n RK integrator. The coefficients α_i, c_i, and a_{ji} are chosen in the deterministic case to ensure that $\boldsymbol{x}_k$ simulates the solution $\boldsymbol{x}(t)$ with error of order h^{n+1}. That is,

$$\boldsymbol{x}(t_k) = \boldsymbol{x}_k + \mathcal{O}(h^{n+1}) \tag{11}$$

The stochastic version of Eqs. (10), using Eq. (1), is given by

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha_1\boldsymbol{k}_1 + \alpha_2\boldsymbol{k}_2 + \cdots + \alpha_n\boldsymbol{k}_n \\ \boldsymbol{k}_1 &= hF(t_k, \boldsymbol{x}_k) + hG(t_k, \boldsymbol{x}_k)\boldsymbol{w}_1 \\ \boldsymbol{k}_j &= hF\left(t_k + c_j h, \boldsymbol{x}_k + \sum_{i=1}^{j-1} a_{ji}\boldsymbol{k}_i\right) \\ &\quad + hG\left(t_k + c_j h, \boldsymbol{x}_k + \sum_{i=1}^{j-1} a_{ji}\boldsymbol{k}_i\right)\boldsymbol{w}_j \end{aligned} \tag{12}$$

where $\boldsymbol{w}_j$ is an i.i.d. vector random process.

The coefficients in Eqs. (10) are found by matching coefficients of a Taylor expansion of both Eqs. (9) and (10) to order h^n. It is well known that the resulting equations are underdetermined and that there are thus many order-n RK integrators. Often, additional

criteria are proposed for selecting solutions.[14] Because the solution becomes unwieldy for orders greater than 4 or 5, the most common techniques are of order 4. Coefficients for the most popular solutions of order $n = 1$ (Euler), $n = 2$ (trapezoidal), $n = 3$, and $n = 4$ can be found in many references, e.g., 13 and 14.

A number of difficulties arise when applying the usual RK algorithms to stochastic differential equations. Foremost is selecting the proper discrete covariance for the process noise, w_j, at each function evaluation in Eqs. (12). As mentioned, the "classical" solution is to assume a constant input over the step interval with covariance matrix $Q_d = Q/h$ (one pseudo-random-number generation per step). This approach is particularly well suited to block-diagram-type simulations. The development below, however, shows that this method is only a solution for the single-step, order-1 Euler integrator. With the more accurate and desirable higher order algorithms it can produce significant error.

Riggs and Phillips[7] investigated this problem and proposed a formula for the discrete covariance matrix of w_j as a function of integration method and order, $Q_d = \Gamma Q/h$, where $\Gamma = 1/\sum_{i=1}^{n} \alpha_i^2$. Unlike in the classical method, here a random number is generated at each function evaluation. Although this addresses the problem of choosing the variance, the authors correctly point out that the solution is still not $\mathcal{O}(h^{n+1})$ accurate. When using the standard deterministic RK solution formulas, the resulting sequence in Eqs. (12) is only accurate to at most order h^2 no matter what routine is being used.

The problem in both of these approaches lies in using the deterministic RK derivation of Eqs. (10) without considering the implications of a stochastic problem. When applied to a stochastic system, the standard approach of finding the coefficients by matching a Taylor series expansion of the exact continuous solution $\boldsymbol{x}(t)$ to a discrete series $\boldsymbol{x}_k$ is inappropriate. There is no single exact solution $\boldsymbol{x}(t)$ but, rather, an ensemble of sample sequences $\{x_i(t)\}$, all with different values but equivalent statistical properties. Rather, Definition 1 states that the covariance of the exact and approximate solutions should be computed. This is done in the next sections.

III. Linear Stochastic Integration

As in Ref. 7, this paper focuses primarily on the integration of time-varying linear stochastic differential equations described by Eq. (2). When F and G are independent of time, Eq. (2) is the more common time-invariant linear stochastic differential equation. The solution approach for the complete nonlinear case is more difficult and is discussed briefly in Sec. IV.

The stochastic RK integration formulas are found by matching a Taylor series expansion of the covariance of the discrete approximation in Eqs. (12) with that of the continuous covariance from Eq. (4) to corresponding order in h. The complete, time-varying covariance equation (4) can be solved at step $k+1$ in terms of $P(t_k)$ using a Taylor series expansion [for brevity, the solution is carried out only to third order in h and P is written for $P(t_k)$]:

$$
\begin{aligned}
P(t_{k+1}) &= P(t_k) + h(FP + PF^T) \\
&+ \frac{h^2}{2}(F^2P + PF^{2T} + 2FPF^T + \dot{F}P + P\dot{F}^T) \\
&+ \frac{h^3}{6}\begin{pmatrix} F^3P + PF^{3T} + 3FPF^{2T} + 3F^2PF^T \\ + F\dot{F}P + P\dot{F}^TF^T + 2\dot{F}FP + 2PF^T\dot{F}^T \\ + 3\dot{F}PF^T + 3FP\dot{F}^T + \ddot{F}P + P\ddot{F}^T \end{pmatrix} \\
&+ \cdots \\
&+ hGQG^T + \frac{h^2}{2}\begin{pmatrix} FGQG^T + GQG^TF^T \\ + \dot{G}QG^T + GQ\dot{G}^T \end{pmatrix} \\
&+ \frac{h^3}{6}\begin{pmatrix} F^2GQG^T + GQG^TF^{2T} + 2FGQG^TF^T \\ + F\dot{G}QG^T + FGQ\dot{G}^T + \dot{G}QG^TF^T \\ + GQ\dot{G}^TF^T + 2\dot{F}GQG^T + 2GQG^T\dot{F}^T \\ + \ddot{G}QG^T + GQ\ddot{G}^T + 2\dot{G}Q\dot{G}^T \end{pmatrix} \\
&+ \cdots
\end{aligned} \tag{13}
$$

The RK solution for $\boldsymbol{x}_k$ is given by the linear version of the stochastic equations (12):

$$
\begin{aligned}
\boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha_1\boldsymbol{k}_1 + \alpha_2\boldsymbol{k}_2 + \cdots + \alpha_n\boldsymbol{k}_n \\
\boldsymbol{k}_1 &= hF_k\boldsymbol{x}_k + hG_k\boldsymbol{w}_1 \\
\boldsymbol{k}_j &= hF(t_k + c_jh)\left[\boldsymbol{x}_k + \sum_{i=1}^{j-1} a_{ji}\boldsymbol{k}_i\right] + hG(t_k + c_jh)\boldsymbol{w}_j
\end{aligned} \tag{14}
$$

After expanding Eqs. (14), squaring, and taking the expected value, the coefficients α_i, c_i, and a_{ij} as well as the covariance of each $\boldsymbol{w}_j$ are determined by comparing a Taylor series expansion to that in Eq. (13). The following sections treat three special cases. First, the simple Euler integrator is derived as background for the classical approach. Then, the third- and fourth-order integrators are presented for the time-invariant case of Eq. (2). Finally, a fourth-order integrator is presented for the complete time-varying problem.

A. Time-Invariant, Euler Integration

The Euler integrator can be simply solved by writing the derivative in Eq. (2) as $(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k)/h$:

$$\boldsymbol{x}_{k+1} = (I + Fh)\boldsymbol{x}_k + hG\boldsymbol{w}_k \tag{15}$$

By computing the covariance of Eq. (15) and comparing to Eq. (13) to first order in h, the discrete covariance of $\boldsymbol{w}_k$, Q_d, is found equal to Q/h. This same result is often arrived at using a frequency-domain argument that limits the white noise to a band defined by the Nyquist frequency.[9] This value for the discrete covariance is used in what is called here the "classical" approach. That is, a standard third-, fourth-, or fifth-order RK solver using Eqs. (14) is used, rather than only the first-order equation (15) for which it was derived, with the covariance of the input noise given by Q/h and held constant over the sampling interval (each w_j is the same and equal to a single pseudo-random-number with covariance Q/h).[7] The examples in Sec. V will compare the results of this classical approach with the new equations presented below.

B. Time-Invariant Equations

In Ref. 7, Riggs and Phillips allowed for independent random numbers $\boldsymbol{w}_j$ at each function evaluation of the standard RK algorithms. However, by putting unnecessary restrictions on the routine (such as identical input noise covariance at each evaluation and limiting the routine to the well-known sets of coefficients), they did not achieve full $\mathcal{O}(h^{n+1})$ accuracy as per Definition 1. Here, the full set of coefficients from Eq. (14) has been assumed and the RK derivation process has been performed for third and fourth order on the time-invariant system to develop a new set of equations for the coefficients as well as the input noise covariance. Because the differential equation has been assumed time invariant, the RK equations are independent of the c_j. Also, the process noise covariance at each function evaluation, has been given the form

$$Q_{d_i} = E\{\boldsymbol{w}_i\boldsymbol{w}_i^T\} = q_i(Q/h) \tag{16}$$

Thus, it is necessary to solve for the coefficients α_i, a_{ji}, and q_i.

For a third-order algorithm there are nine unknowns, $\alpha_1, \alpha_2, \alpha_3$, $a_{21}, a_{31}, a_{32}, q_1, q_2$, and q_3. The solution process produces eight equations that must be simultaneously solved, leaving one free parameter (see the Appendix for the equations). This is in contrast to the deterministic, time-invariant third-order algorithm with six equations and three unknowns. For this work, the extra degree of freedom was eliminated by arbitrarily setting a_{31} to zero. See Table 1 for the resulting third-order coefficients. These coefficients were found using a numerical root finder[8] on the defining equations in the Appendix.

For a fourth-order solution, there are 14 unknowns and 13 equations (again, see the Appendix for the coefficient equations). Table 1 lists one possible set of coefficients with the extra degree of freedom eliminated by requiring that $a_{41} = 0$. Note that all solutions are restricted so that $q_i > 0$. This time-invariant solution is referred to as method A.

Table 1 Runge-Kutta coefficients for third- and fourth-order time-invariant solutions (method A)

Coefficients	Third order	Fourth order
α_1	0.53289582961739	0.47012396888046
α_2	0.25574324768195	0.36597075368373
α_3	0.21136092270067	0.08906615686702
α_4	—	0.07483912056879
a_{21}	1.52880952525675	2.71644396264860
a_{31}	0	−6.95653259006152
a_{32}	0.51578733443615	0.78313689457981
a_{41}	—	0
a_{42}	—	0.48257353309214
a_{43}	—	0.26171080165848
q_1	1.87653936176981	2.12709852335625
q_2	3.91017166264989	2.73245878238737
q_3	4.73124353935667	11.22760917474960
q_4	—	13.36199560336697

Table 2 Fourth-order, time-varying RK coefficients (method B)

Coefficients	Fourth order
α_1	0.25001352164789
α_2	0.67428574806272
α_3	−0.00831795169360
α_4	0.08401868181222
a_{21}	0.66667754298442
a_{31}	0.63493935027993
a_{32}	0.00342761715422
a_{41}	−2.32428921184321
a_{42}	2.69723745129487
a_{43}	0.29093673271592
q_1	3.99956364361748
q_2	1.64524970733585
q_3	1.59330355118722
q_4	0.26330006501868

C. Time-Varying Equations

The set of coefficients in Table 1 are only correct for the special case of time-invariant differential equations and should be used for those problems. However, it is often necessary to integrate sets of time-varying differential equations where the coefficients in Table 1 are no longer correct and it is necessary to find the proper values for the c_i. It can be shown that the second-order RK integrator is solvable with six equations and six unknowns, $\alpha_1, \alpha_2, \alpha_{21}, q_1, q_2$, and c_2. This solution, with error $\mathcal{O}(h^3)$, is given by

$$\alpha_1 = \alpha_2 = \tfrac{1}{2}, \qquad a_{21} = 1, \qquad q_1 = q_2 = 2, \qquad c_2 = 1$$

This integrator is commonly called Euler-Cauchy or trapezoidal integration. It is also identical to the second-order method presented in Ref. 7 and matches the deterministic second-order algorithm. Because of its low order (and thus low accuracy unless very small time steps are used), second-order algorithms are rarely used in practice.

No solution exists, however, for the third-order problem. There are 17 conditions that must be satisfied with only 11 parameters. The alternative is to use the fourth-order RK algorithm with coefficients matching only the h^3 terms. In this case there are now 17 unknowns for the 17 equations, but the error is $\mathcal{O}(h^4)$. This is an important peculiarity of the time-varying stochastic system. This fourth-order set of coefficients has been solved (again using a numerical root finder[8]; see the Appendix for the defining equations) and is listed in Table 2. The time-varying solution is referred to as method B. The c are found from the standard conditions:

$$c_2 = a_{21}$$

$$c_3 = a_{31} + a_{32}$$

$$c_4 = a_{41} + a_{42} + a_{43}$$

It is important to note again that this integrator has error $\mathcal{O}(h^4)$. Future work should be directed at solving the fifth-order RK problem to find a fourth-order accurate method for time-varying equations.

IV. Non-linear Stochastic Differential Equations

The coefficients in Table 2 and RK algorithms have been rigorously derived only for linear stochastic differential equations. It would be very useful to likewise prove their utility for the more general nonlinear Ito equation with covariance matrix given by Eq. (8). This is a formidable task and has yet to be accomplished. The difficulty is in performing the expectations in Eq. (8) after a Taylor series expansion. In the linear case, the probability density of the state variables remains Gaussian for all time. Thus, the covariance completely describes the process and the propagation equations are easy to solve. In contrast, for a nonlinear system the probability density changes with time despite the driving noise being normal. Thus, it is in general impossible to find the expectations in Eq. (8) without simultaneously solving the Fokker-Planck equation for the probability density function. In fact, simply solving for the covariance is not enough. A precise solution of order n would compute the propagation equations for the first $n + 1$ moments [using Eq. (11)] of the random state variables.

Nevertheless, it is proposed that the same set of coefficients derived for the time-varying linear case and listed in Table 2 are appropriate for the nonlinear case as well. This proposition is arrived at after some minor algebraic manipulations of Eq. (8) and comparison to the linear case and by requiring that a linearization of Eq. (8) for small state values and times be solved by the set of equations presented above. Example 3 provides some evidence for this assertion.

V. Examples

The new RK coefficients proposed in this paper were tested on three examples: a linear equation; a linear time-varying equation; and a nonlinear dynamic system. The results of these simulations are presented below.

Example 1. The first example is a simple, first-order, time-invariant stochastic differential equation:

$$\dot{x}(t) = -ax(t) + w(t) \tag{17}$$

where $Ew(t)w(t+\tau) = Q\delta(\tau)$.

The known solution for the variance is

$$P(t) = \frac{Q}{2a}(1 - e^{-2at}) \tag{18}$$

Equation (17) was numerically integrated using three methods: the new fourth-order RK integrator with the coefficients from Table 1 (method A), the classical method with a constant input over the sample interval with variance Q/h, and the equations of Ref. 7, with $Q_d = 3.6Q/h$. The results for a step size of 0.8 s are shown in Fig. 1, with $a = 2$ s and $Q = 4$ Hz^{-1} (the magnitude units are arbitrary). The resulting steady-state covariance is thus unity. Note that the previous solution methods do not reproduce the correct variance for this large step size relative to the system time constant. The variances in Fig. 1 were found by simulating Eq. (17) eight hundred times and averaging the results.

Figure 2 shows the accuracy of the three methods applied to Example 1 for a range of step sizes. It should be pointed out that all three methods became unstable for $h > 1.4$ seconds. (For linear differential equations, RK integration methods can be viewed as a continuous-to-discrete mapping of the s plane to the z plane. In this context, second-order RK integration is equivalent to a bilinear transformation and is thus guaranteed stable for all step sizes. However, higher order RK routines map the left half of the s plane to regions of the z plane outside of the unit circle and thus can result in unstable integration for larger step sizes.[16]) The results in Fig. 2 show that the coefficients presented here maintain an accuracy of better than 4% for a step size as large as 1 s. For the step sizes tested, the classical method was never better than 2% and rapidly increased to as much as 100%. For a 1-s step size the new method was a factor of 10 better than the classical. The Riggs method fared better for smaller step sizes, but after $h = 0.5$ s the error increases dramatically. Thus, the coefficients presented here, applied to this example, allow for a fourth-order RK integration with more accuracy and at least twice as large a step size and, consequently, half the computer time. It should be pointed out, however,

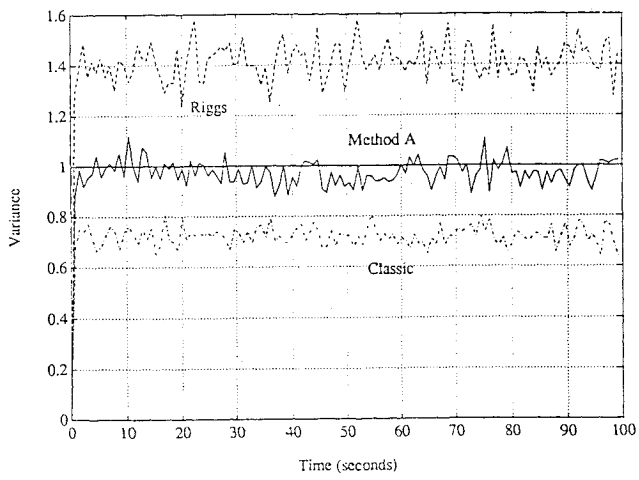

Fig. 1 Example 1, fourth-order stochastic Runge-Kutta solution, h = 8 s.

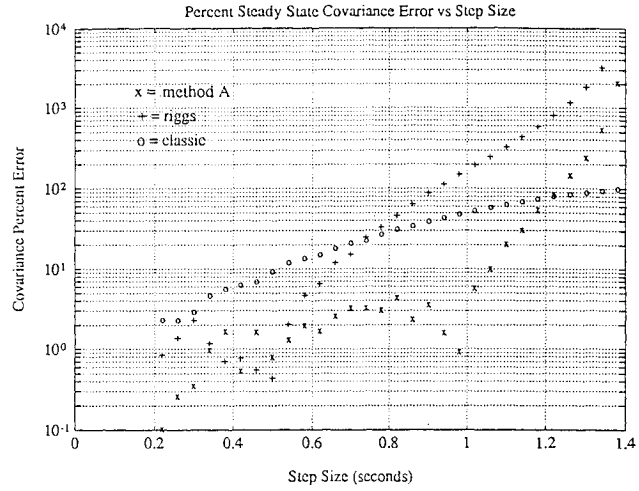

Fig. 2 Comparison of integration method error for example 1 vs step size.

that the classic method and the Riggs method are incorrect in that they apply fourth-order equations but are, in fact, only second-order accurate.

Example 2. This example is the second-order time-varying differential equation

$$\ddot{x}(t) = -\sin(\omega t)x(t) - b\dot{x}(t) + w(t) \tag{19}$$

This equation can be written in matrix form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\sin(\omega t) & -b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \tag{20}$$

While there is no closed-form solution for the covariance of this system, we can write the differential equations describing the covariance terms and integrate them with a standard, deterministic equation solver:

$$\begin{aligned} \dot{P}_1 &= 2P_{12} \\ \dot{P}_{12} &= P_2 - \sin(\omega t)P_1 - bP_{12} \\ \dot{P}_2 &= -2\sin(\omega t)P_{12} - 2bP_2 + Q \end{aligned} \tag{21}$$

Figure 3 shows the results for this example, again using the three different integrators. The covariance equations were solved using a fourth- and fifth-order variable-step/variable-order RK solver.[8] Note that here also the Riggs method falls far short of the other two. Although the classical method worked better here than in the time-invariant case, it is still not as accurate as the new method.

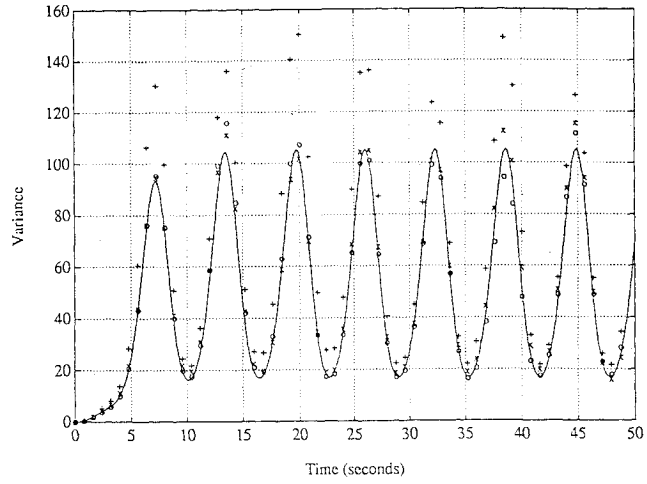

Fig. 3 Example 2, time-varying fourth-order Runge-Kutta solution, h = 8 s, x = method B, + = Riggs, o = classical.

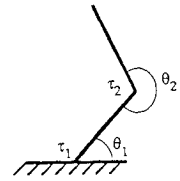

Fig. 4 Two-link robot example.

Example 3. Lastly, the RK integrator of method B was applied to a nonlinear set of stochastic differential equations. Despite the difficulty in proving the appropriateness of these coefficients as a solution, the example does show they reproduce the correct covariance and that the time-varying solution of method B is needed.

The problem being simulated is a two-link robot arm with torques applied at the joints and arms of unit length (see Fig. 4). A simple full state feedback (PD) controller is used with unity gains on all of the four states. A white-noise random torque with unit spectral density is applied at the middle joint (τ_2). The equations of motion for the joint angles and angular rates are given by

$$\begin{aligned} \dot{\theta}_1 &= \omega_1 \\ \dot{\theta}_2 &= \omega_2 \\ \dot{\omega}_1 &= \frac{-\sin\theta_2}{1+\sin^2\theta_2}\left[(\omega_1+\omega_2)^2 - \cos\theta_2\omega_1^2\right] \\ &\quad + \frac{1}{1+\sin^2\theta_2}\left(\frac{\tau_1}{2} - \frac{\tau_2}{2} + \cos\theta_2\frac{\tau_2}{2}\right) \\ \dot{\omega}_2 &= \frac{\sin\theta_2(1-\cos\theta_2)}{1+\sin^2\theta_2}\left[(\omega_1+\omega_2)^2 - \cos\theta_2\omega_1^2\right] \\ &\quad + \sin\theta_2\omega_1^2 + \frac{\tau_2}{2} \\ &\quad - \frac{1-\cos\theta_2}{1+\sin^2\theta_2}\left[\frac{\tau_1}{2} - \frac{\tau_2}{2} + \cos\theta_2\frac{\tau_2}{2}\right] \end{aligned} \tag{22}$$

The applied torques τ_1 and τ_2 are given by

$$\begin{aligned} \tau_1 &= -\omega_1 - (\theta_1 - \theta_{\text{ref1}}) \\ \tau_2 &= w - \omega_2 - (\theta_2 - \theta_{\text{ref2}}) \end{aligned} \tag{23}$$

It can be shown, using Lyapunov theory, that this closed-loop system is globally asymptotically stable.

Before performing the random simulation, the step response of the system was numerically integrated. This verified stability, checked the various integration programs, as they should work equally well on deterministic systems, and provided information

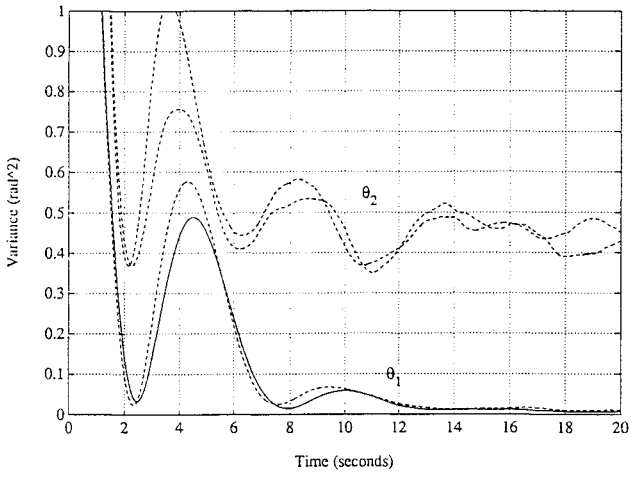

Fig. 5 Deterministic step response of robot example.

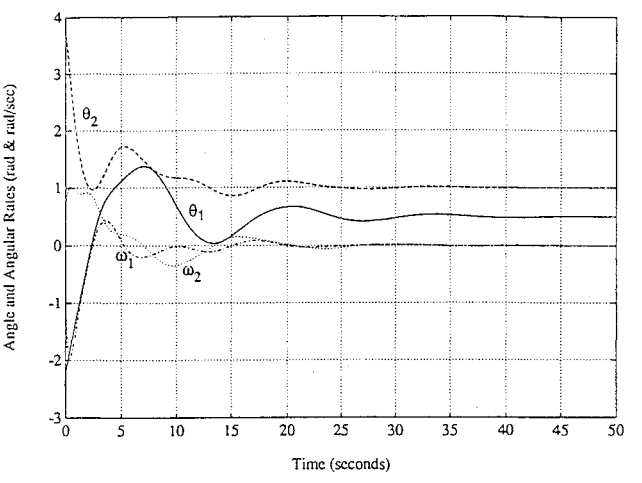

Fig. 6 Variance of robot example for nonlinear (solid) and linearized (dashed) methods.

about the system time constants. This step response is shown in Fig. 5 for $\theta_{\text{ref1}} = 0.5$ and $\theta_{\text{ref2}} = 1.0$ and nonzero initial conditions. This integration was performed using a variable-step, fourth- and fifth-order RK integrator with relative accuracy of 10^{-6},[8] the standard fixed-step, fourth-order RK with a step size of 0.1 s, and the method B Runge-Kutta integration routine proposed in this paper for a fixed step size of 0.1 s. There was no discernible difference in the results.

Figure 5 shows that the time constants of this closed-loop system are on the order of 20–25 s. Thus, unlike Examples 1 and 2, the step size used in the integration (0.1 s) is very short compared to the system time constant. For this nonlinear example, all of the considered RK routines became unstable for $h > 0.4$ s. It is thus not possible to compare the various methods for larger step sizes, but it is possible to verify the efficacy of the proposed RK coefficients on nonlinear systems. This is done by simulating Eqs. (22) for the parameters described and comparing to a simulation using a local linearization technique.[17]

The local linearization technique involves linearizing Eqs. (8) at a given time t_0 about the state values at that time, $\boldsymbol{x}(t_0)$. This linear system is then simulated over a single time step using the RK integration techniques described earlier (method A) or solved exactly by computing the discrete state transition and input covariance matrices to find the new state values at $t_0 + h$. The process is then repeated, with a new linearization, for the sequence of desired times. For sufficiently small times this produces an accurate simulation of the stochastic system (within the accuracy of the linearization). (Note that this approach can be used in general to simulate nonlinear differential systems driven by noise. However, the linearization process can become quite involved and, particularly for very large systems, very unwidely. Thus, it is neither practical nor convenient for most applications.)

Figure 6 shows the variance results for a simulation of the full nonlinear equations using the coefficients in Table 2 for time-varying systems and of the local linearization technique. Note that the linearized system was solved via the RK integrator and an exact discrete solution with no discernible difference. The strong agreement adds validity to the assertion that the RK integrator proposed here is also correct for nonlinear systems.

This example did point out an important discovery. When simulating nonlinear equations, the coefficients derived above for linear, time-varying systems must be used. This can be easily understood by returning to the local linearization technique. At each time, the linearization results in different F and G matrices depending upon the time and state variables. Thus, a nonlinear system can be interpreted to behave like a time-varying linear one for small enough time steps. To verify, this example was simulated using the time-invariant coefficients from Table 1 with erroneous results (the variance was in error by factors of 3 or more). For this reason, it would be very useful to derive higher order time-varying coefficients with accuracy better than third order.

VI. Conclusions

This paper has presented the derivation of specific RK algorithms of second, third, and fourth order for use on stochastic differential equations. Although these algorithms suffer slightly from being fixed step routines, thus requiring some *a priori* knowledge of the system time constants, they allow larger time steps and improved mean-square accuracy of the simulated random process over previous methods. These standard methods, particularly those in readily available commercial software packages, were shown to be inaccurate when applied to stochastic differential equations. Two examples, a linear time-invariant equation and a time-varying differential equation, were presented to verify the accuracy of the new coefficients. It was claimed by analogy to the linear case that these coefficients are most appropriate for nonlinear stochastic differential equations as well. A third example, a nonlinear dynamic system, was presented to show the accuracy of the algorithm on nonlinear stochastic differential equations.

Acknowledgments

This work was prepared under NASA contract NAS8-36125. I also express my gratitude to Brad Parkinson and Arthur E. Bryson for their helpful suggestions.

Appendix

Listed below are the equations used for determining the coefficients above. The coefficients α_i are given as a_{11}, a_{12}, etc. The q_i are given as $q(i)$.

Time Invariant, Third Order

$$a_{11} + a_{12} + a_{13} - 1$$

$$a_{12}a_{21} + a_{13}(a_{31} + a_{32}) - \tfrac{1}{2}$$

$$a_{13}a_{32}a_{21} - \tfrac{1}{6}$$

$$a_{11}a_{12}a_{21} + a_{11}a_{13}(a_{31} + a_{32}) + a_{12}a_{12}a_{21} + a_{12}a_{13}(a_{31} + a_{32} + a_{21}) + a_{13}a_{13}(a_{31} + a_{32}) - \tfrac{1}{2}$$

$$a_{11}a_{11}q(1) + a_{12}a_{12}q(2) + a_{13}a_{13}q(3) - 1$$

$$a_{11}a_{12}a_{21}q(1) + a_{11}a_{13}a_{31}q(1) + a_{12}a_{13}a_{32}q(2) - \tfrac{1}{2}$$

$$a_{11}a_{13}a_{32}a_{21}q(1) - \tfrac{1}{6}$$

$$(a_{12}a_{21} + a_{13}a_{31})(a_{12}a_{21} + a_{13}a_{31})q(1) + a_{13}a_{13}a_{32}a_{32}q(2) - \tfrac{1}{3}$$

Time Invariant, Fourth Order

$$a_{11} + a_{12} + a_{13} + a_{14} - 1$$

$$a_{12}a_{21} + a_{13}a_{31} + a_{13}a_{32} + a_{14}(a_{41} + a_{42} + a_{43}) - \tfrac{1}{2}$$

$$a_{13}a_{21}a_{32} + a_{14}(a_{21}a_{42} + a_{31}a_{43} + a_{32}a_{43}) - \tfrac{1}{6}$$

$$a_{11}a_{12}a_{21} + a_{11}a_{13}(a_{31} + a_{32}) + a_{12}a_{12}a_{21} + a_{12}a_{13}(a_{31} + a_{32} + a_{21}) + a_{13}a_{13}(a_{31} + a_{32}) + a_{11}a_{14}(a_{41} + a_{42} + a_{43}) + a_{12}a_{14}(a_{41} + a_{42} + a_{43} + a_{21}) + a_{13}a_{14}(a_{41} + a_{42} + a_{43} + a_{31} + a_{32}) + a_{14}a_{14}(a_{41} + a_{42} + a_{43}) - \tfrac{1}{2}$$

$$a_{14}a_{43}a_{32}a_{21} - \tfrac{1}{24}$$

$$a_{11}a_{13}a_{32}a_{21} + a_{12}a_{13}a_{32}a_{21} + a_{13}a_{13}a_{32}a_{21} + a_{14}(a_{42}a_{21} + a_{43}a_{31} + a_{43}a_{32}) + a_{14}a_{13}a_{32}a_{21} - \tfrac{1}{6}$$

$$a_{12}a_{12}a_{21}a_{21} + 2a_{12}a_{13}a_{21}(a_{31} + a_{32}) + a_{13}a_{13}(a_{31} + a_{32}) \times (a_{31} + a_{32}) + 2(a_{12}a_{14}a_{21} + a_{13}a_{14}a_{31} + a_{13}a_{14}a_{32})(a_{41} + a_{42} + a_{43}) + a_{14}a_{14}(a_{41} + a_{42} + a_{43})(a_{41} + a_{42} + a_{43}) - \tfrac{1}{4}$$

$$a_{11}a_{11}q(1) + a_{12}a_{12}q(2) + a_{13}a_{13}q(3) + a_{14}a_{14}q(4) - 1$$

$$a_{11}a_{12}a_{21}q(1) + a_{11}a_{13}a_{31}q(1) + a_{12}a_{13}a_{32}q(2) + a_{11}a_{14}a_{41}q(1) + a_{12}a_{14}a_{42}q(2) + a_{13}a_{14}a_{43}q(3) - \tfrac{1}{2}$$

$$a_{11}a_{13}a_{32}a_{21}q(1) + a_{11}a_{14}[a_{42}a_{21}q(1) + a_{43}a_{31}q(1)] + a_{12}a_{14}a_{43}a_{32}q(2) - \tfrac{1}{6}$$

$$(a_{12}a_{21} + a_{13}a_{31})(a_{12}a_{21} + a_{13}a_{31})q(1) + a_{13}a_{13}a_{32}a_{32}q(2) + 2a_{12}a_{14}a_{41}a_{21}q(1) + 2a_{13}a_{14}[a_{41}a_{31}q(1) + a_{42}a_{32}q(2)] + a_{14}a_{14}[a_{41}a_{41}q(1) + a_{42}a_{42}q(2) + a_{43}a_{43}q(3)] - \tfrac{1}{3}$$

$$a_{11}a_{14}a_{43}a_{32}a_{21}q(1) - \tfrac{1}{24}$$

$$a_{12}a_{13}a_{32}a_{21}a_{21}q(1) + a_{13}a_{13}a_{32}a_{31}a_{21}q(1) + a_{12}a_{14}[a_{42}a_{21}a_{21}q(1) + a_{43}a_{31}a_{21}q(1)] + a_{13}a_{14}[a_{42}a_{31}a_{21}q(1) + a_{43}a_{31}a_{31}q(1) + a_{41}a_{32}a_{21}q(1) + a_{43}a_{32}a_{32}q(2)] + a_{14}a_{14}[a_{42}a_{41}a_{21}q(1) + a_{43}a_{41}a_{31}q(1) + a_{43}a_{42}a_{32}q(2)] - \tfrac{1}{8}$$

Time Varying, Fourth Order

$$a_{11} + a_{12} + a_{13} + a_{14} - 1$$

$$a_{12}a_{21} + a_{13}a_{31} + a_{13}a_{32} + a_{14}(a_{41} + a_{42} + a_{43}) - \tfrac{1}{2}$$

$$a_{13}a_{21}a_{32} + a_{14}(a_{21}a_{42} + a_{31}a_{43} + a_{32}a_{43}) - \tfrac{1}{6}$$

$$a_{11}a_{12}a_{21} + a_{11}a_{13}(a_{31} + a_{32}) + a_{12}a_{12}a_{21}a_{12}a_{13}(a_{31} + a_{32} + a_{21}) + a_{13}a_{13}(a_{31} + a_{32}) + a_{11}a_{14}(a_{41} + a_{42} + a_{43}) + a_{12}a_{14}(a_{41} + a_{42} + a_{43} + a_{21}) + a_{13}a_{14}(a_{41} + a_{42} + a_{43} + a_{31} + a_{32}) + a_{14}a_{14}(a_{41} + a_{42} + a_{43}) - \tfrac{1}{2}$$

$$a_{11}a_{11}q(1) + a_{12}a_{12}q(2) + a_{13}a_{13}q(3) + a_{14}a_{14}q(4) - 1$$

$$a_{11}a_{12}a_{21}q(1) + a_{11}a_{13}a_{31}q(1) + a_{12}a_{13}a_{32}q(2) + a_{11}a_{14}a_{41}q(1) + a_{12}a_{14}a_{42}q(2) + a_{13}a_{14}a_{43}q(3) - \tfrac{1}{2}$$

$$a_{11}a_{13}a_{32}a_{21}q(1) + a_{11}a_{14}[a_{42}a_{21}q(1) + a_{43}a_{31}q(1)] + a_{12}a_{14}a_{43}a_{32}q(2) - \tfrac{1}{6}$$

$$q(1)(a_{12}a_{21} + a_{13}a_{31} + a_{14}a_{41})^2 + q(2)(a_{13}a_{32} + a_{14}a_{42})^2 + a_{14}a_{14}a_{43}a_{43}q(3) - \tfrac{1}{3}$$

$$c_2 = a_{21}$$

$$c_3 = a_{31} + a_{32}$$

$$c_4 = a_{41} + a_{42} + a_{43}$$

$$a_{12}a_{21}c_2 + a_{13}c_3(a_{31} + a_{32}) + a_{14}c_4(a_{41} + a_{42} + a_{43}) - \tfrac{1}{3}$$

$$a_{13}a_{32}c_2 + a_{14}a_{42}c_2 + a_{14}a_{43}c_3 - \tfrac{1}{6}$$

$$a_{11}a_{12}a_{21}c_2q(1) + a_{11}a_{13}a_{31}c_3q(1) + a_{12}a_{13}a_{32}c_3q(2) + a_{11}a_{14}a_{41}c_4q(1) + a_{14}a_{12}a_{42}c_4q(2) + a_{14}a_{13}a_{43}c_4q(3) - \tfrac{1}{3}$$

$$a_{12}a_{12}c_2q(2) + a_{13}a_{13}c_3q(3) + a_{14}a_{14}c_4q(4) - \tfrac{1}{2}$$

$$a_{12}a_{12}c_2c_2q(2) + a_{13}a_{13}c_3c_3q(3) + a_{14}a_{14}c_4c_4q(4) - \tfrac{1}{3}$$

$$a_{12}a_{13}a_{32}c_2q(2) + a_{14}a_{12}a_{42}c_4q(2) + a_{14}a_{13}a_{43}c_4q(3) - \tfrac{1}{6}$$

References

[1]Bryson, A. E., and Ho, Y. C., *Applied Optimal Control*, Hemisphere, Washington, DC, 1975.

[2]Stengel, R. F., *Stochastic Optimal Control, Theory and Applications*, Wiley, New York, 1986.

[3]Wong, E., and Hajek, B., *Stochastic Processes in Engineering Systems*, Springer-Verlag, New York, 1985.

[4]Soong, T. T., *Random Differential Equations in Science and Engineering*, Academic, New York, 1973.

[5]Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 2nd ed., McGraw-Hill, New York, 1984.

[6]Hoel, P. G., Port, S. C., and Stone, C. J., *Introduction to Stochastic Processes*, Waveland, IL, 1972.

[7]Riggs, T. L., Jr., and Phillips, C. L., "Modeling Continuous Noise Sources in Digital Simulations," *Simulation*, Vol. 48, No. 1, 1987.

[8]Matlab™, Mathworks, South Natick, MA, 1992.

[9]Kasdin, N. J., "Discrete Simulation of Colored Noise and Stochastic Processes and $1/f^{\alpha}$ Power Law Noise Generation," *Proceedings of the IEEE* (in preparation).

[10]Kasdin, N. J., and Walter, T., "Discrete Simulation of Power Law Noise," *Proceedings of the IEEE Frequency Control Symposium* (Hershey, PA), 1992, p. 294.

[11]Van Loen, C. F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions of Automatic Control*, Vol. AC-23, No. 3, 1978.

[12]Schmidt, G. C., "Designing Nonlinear Filters Based on Daum's Theory," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 16, No. 2, 1993.

[13]Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vertterling, W. T., *Numerical Recipes*, Cambridge Univ. Press, Cambridge, England, UK, 1986.

[14]Jain, M. K., *Numerical Methods for Scientific and Engineering Computation*. Wiley, New York, 1985.

[15]Conte, S. D., and de Boor, C., *Elementary Numerical Analysis, an Algorithmic Approach*, McGraw-Hill, New York, 1980.

[16]Franklin, G. F., Powell, J. D., and Workman, M. L., *Digital Control of Dynamic Systems*, 2nd ed., Addison-Wesley, Reading, MA, 1990.

[17]Bryson, A. E., private communications.