

Monte Carlo filters for identification of nonlinear structural dynamical systems

C S MANOHAR* and D ROY

Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, India

e-mail: {manohar, royd}@civil.iisc.ernet.in

Abstract. The problem of identification of parameters of nonlinear structures using dynamic state estimation techniques is considered. The process equations are derived based on principles of mechanics and are augmented by mathematical models that relate a set of noisy observations to state variables of the system. The set of structural parameters to be identified is declared as an additional set of state variables. Both the process equation and the measurement equations are taken to be nonlinear in the state variables and contaminated by additive and (or) multiplicative Gaussian white noise processes. The problem of determining the posterior probability density function of the state variables conditioned on all available information is considered. The utility of three recursive Monte Carlo simulation-based filters, namely, a probability density function-based Monte Carlo filter, a Bayesian bootstrap filter and a filter based on sequential importance sampling, to solve this problem is explored. The state equations are discretized using certain variations of stochastic Taylor expansions enabling the incorporation of a class of non-smooth functions within the process equations. Illustrative examples on identification of the nonlinear stiffness parameter of a Duffing oscillator and the friction parameter in a Coulomb oscillator are presented.

Keywords. Nonlinear structural system identification; particle filters; stochastic differential equations.

1. Introduction

The problem of linear structural system identification is widely studied; both time and frequency domain methods have been developed for this purpose (Ljung 1997; Lieven & Ewins 2001; Peeters & Roeck 2001; Pintelon & Schoukens 2001). Thus, for instance, the techniques for measurement of frequency and impulse response functions and the subsequent offline extraction of modal characteristics from these measured response functions are well

*For correspondence

This paper is dedicated to Prof R N Iyengar of the Indian Institute of Science on the occasion of his formal retirement.

established in the vibrations literature (Heylen *et al* 1997; Ewins 2000). Problems of online identification of structures have also received notable attention (Imai *et al* 1989; Ghanem & Shinozuka 1995; Shinozuka & Ghanem 1995; Ljung 1997). The theory of Kalman filtering provides one of the corner stones in the development of these methods (Kalman 1960; Kailath 1981; Brown & Hwang 1992; Chui & Chen 1998; Grewal & Andrews 2001). The problem of identifying parameters of nonlinear vibrating systems is of fundamental importance in understanding structural behaviour under extreme environmental effects such as strong motion earthquakes and high winds, and in studies on structural health monitoring and vibration control. These problems are of interest in the context of detection and characterization of structural damages, such as the occurrence of cracks, using vibration data. The subject is of particular relevance in the context of problems of earthquake engineering wherein structures are designed to possess controlled inelastic behaviour. In the context of nonlinear system identification problems, again, both frequency domain and time domain methods have been explored. The former class of studies includes development and analysis of higher order spectra (Worden & Manson 2001) and has been basically projected as extensions of classical modal analysis methods (Ewins 2000). On the other hand, the time domain methods have mainly focused on state estimation techniques. Here, linearization tools, combined with the traditional Kalman filtering techniques, have been developed for dealing with the nonlinear dynamic state estimation problems (Yun & Shinozuka 1980; Brown & Hwang 1997). It is of interest to note that the problem of state estimation of dynamical systems is widely encountered in many branches of sciences, including economics, object tracking, digital communication channels and signal processing. Recent research in these fields have been focused on dealing with nonlinear systems and non-Gaussian noise models (Gordon *et al* 1993; Tanizaki 1996; Doucet 1998; Liu & Chen 1998; Pitt & Shephard 1999; Doucet *et al* 2000; Iba 2001; Crisan & Doucet 2002; Ristic *et al* 2004). Of particular interest has been the development of the so-called particle filters which employ Monte Carlo simulation procedures for state estimation problems in dynamical systems. These methods are capable of handling general forms of nonlinearities and non-Gaussian additive/multiplicative noises. While the roots of these methods date back to the 1950s' (Doucet 1998), recent developments in this field have been spurred by the wide availability of inexpensive and fast computational power. To the best of our knowledge, the applicability of these methods to problems of system identification of relevance in structural engineering has not yet been explored in the existing literature. Accordingly, in the present work, we apply three simulation-based filtering strategies to the problem of system parameter identification in two typical nonlinear oscillators, namely, the Duffing oscillator and the Coulomb oscillator. These filters include: the density-based Monte Carlo filter as developed by Tanizaki (1996), the Bayesian bootstrap algorithm due to Gordon *et al* (1993) and the sequential importance sampling based method as developed by Doucet *et al* (2000). The paper provides a summary of the development of these filters before numerical illustrations are presented. It also attempts to bring in several concepts in stochastic calculus in order to reduce the stochastic differential equations for the process and/or observation variables to discrete stochastic maps, suitable for being applied within the filtering algorithms.

2. Problem formulation

Statement of the problem of structural system identification consists of two entities: first, the governing equation of motion, or the process equation, which is derived based on principles of mechanics and, second, an observation (or measurement) equation which relates

the observed quantity to the state variables. Both the process and measurement equations are taken to anticipate, with reasonable accuracy, the behaviour of the structure and the relationship between state variables and the measured quantities. Notwithstanding this prospect, both these equations include noise terms to represent effects of errors in formulation of the respective equations. Furthermore, these noise terms are modelled as Gaussian white noise processes and, consequently, the process equation is assumed to be governed by the Ito stochastic differential equation of the form,

$$dX(t) = \tilde{H}(X(t), \theta, t)dt + \tilde{G}(X(t), \theta, t)dB_I(t); \quad X(0) = X_0. \quad (1)$$

Here the sizes of the vectors $X(t)$, \tilde{H} , \tilde{G} and $dB_I(t)$ are respectively, $n_X \times 1$, $n_X \times 1$, $n_X \times m_X$ and $m_X \times 1$, θ is a $n_\theta \times 1$ vector of structural parameters to be identified, and $dB_I(t)$ is a vector of increments of standard Brownian motion processes. The vector of initial conditions X_0 is modelled as a vector of random variables, which is assumed to be independent of $dB_I(t)$. The measurement equation is similarly modelled as

$$dY(t) = \tilde{F}(X(t), \theta, t)dt + \tilde{Q}(X(t), \theta, t)dB_{II}(t). \quad (2)$$

Here the sizes of $Y(t)$, \tilde{F} , \tilde{Q} and $dB_{II}(t)$ are respectively, $n_Y \times 1$, $n_Y \times 1$, $n_Y \times m_Y$, and $m_Y \times 1$; $dB_{II}(t)$ is a vector of increments of standard Brownian motion processes that is independent of θ and X_0 . Further more, it is assumed that $dB_I(t)$ and $dB_{II}(t)$ are mutually independent. The noise terms in the process equation (1) represent effect of modelling errors, and errors made in applying the test signals. Similarly, the noise term in the measurement equation (2) represents effects of measurement noise and also modelling errors made in relating observed quantities with the state variables of the problem. The problem on hand consists of estimating the vector θ given the observations $dY(\tau)$, $0 < \tau \leq t$. We begin by treating θ as being a random process governed by

$$d\theta(t) = dB_{III}(t); \quad \theta(0) = \theta_0. \quad (3)$$

Here $dB_{III}(t)$ is a $n_\theta \times 1$ vector of increments of standard Brownian motion processes that are independent of $dB_I(t)$ and $dB_{II}(t)$; also, θ_0 is taken to be a vector of random variables that is independent of X_0 , $dB_I(t)$, $dB_{II}(t)$ and $dB_{III}(t)$. We define an augmented state vector $\chi = [X(t) \theta(t)]'$ and recast (1) through (3) in the form:

$$\begin{aligned} d\chi(t) &= H(\chi(t), t)dt + G(\chi(t), t)dB(t); \quad \chi(0) = \chi_0, \\ dY(t) &= F(\chi(t), t)dt + Q(\chi(t), t)d\tilde{B}(t). \end{aligned} \quad (4a,b)$$

We now consider the problem of estimation of the dynamic state $\chi(t)$ given the observations $dY(\tau)$, $0 < \tau \leq t$. It is clear that the solution to this problem includes the solution to the problem of estimating the parameter vector θ . It may be remarked here that the representation of the measurement equation in the above equations is formal in nature. In practice, the observations would be made at a set of discrete time instants, in which case, the measurement equation could be more realistically represented as

$$y(t_k) = F_k(\chi(t_k)) + Q_k(\chi(t_k))v_k; \quad k = 1, 2, \dots, n. \quad (5)$$

Here $\{v_k\}_{k=1}^n$ represents a sequence of independent random variables.

As a first step towards obtaining the estimates of the dynamic states, we discretize (4) in time to obtain the system

$$\begin{aligned}x_{k+1} &= h_k(x_k) + g_k(x_k)w_k; \quad k = 0, 1, \dots, n, \\y_k &= f_k(x_k) + q_k(x_k)v_k; \quad k = 1, 2, \dots, n.\end{aligned}\tag{6a,b}$$

Here the sizes of the vectors x_k , h_k , g_k , w_k , y_k , f_k , q_k , v_k are respectively, $n_x \times 1$, $n_x \times 1$, $n_x \times n_w$, $n_w \times 1$, $n_y \times 1$, $n_y \times n_v$, $1 \times n_v$ and $n_v \times 1$; further, $n_x = n_X + n_\theta$, $n_w = m_X + m_\theta$, $n_y = n_Y$, $n_v = m_Y$. Depending upon the discretization scheme used, it is possible that the second term on the right hand side of the above equations could be of the form $g_k(x_k, w_k)$ and $q_k(x_k, v_k)$ respectively. The steps involved in deriving the discretized equations (6) from (4) will be taken up later in the paper when we consider two specific illustrative examples (§§ 4 and 5). It may be asserted here that the process x_k possesses strong Markov properties, and, it may be assumed that the probability density functions $p(x_k|x_{k-1})$ and $p(y_k|x_k)$ are deducible from (6). We introduce the notation $x_{0:k} = \{x_i\}_{i=0}^k$ and $y_{1:k} = \{y_i\}_{i=1}^k$. The problem on hand consists of determining the conditional probability density function (PDF) $p(x_{0:k}|y_{1:k})$ which represents the PDF of the state vector conditioned on observations available up to time k . Specifically, we are interested in the marginal PDF $p(x_k|y_{1:k})$, known as the filtering density, and the expectations

$$\begin{aligned}a_{k|k} &= E[x_k|y_{1:k}] = \int x_k p(x_k|y_{1:k}) dx_k, \\ \Sigma_{k|k} &= E[(x_k - a_{k|k})(x_k - a_{k|k})^t] = \int (x_k - a_{k|k})(x_k - a_{k|k})^t p(x_k|y_{1:k}) dx_k.\end{aligned}\tag{7a,b}$$

The integrations in the above equations are multidimensional in nature (in the present case it is $n_x \times 1$) and the limits of integrations here and, in similar contexts in the rest of this paper, are not explicitly specified for the sake of simplifying the notations. Analytical solution to this problem is generally not possible with exceptions, for instance, arising when the process and measurement equations are linear and noises are Gaussian and additive and, in such a case, the traditional Kalman filter provides the exact solution to the problem on hand. For the general case, a *formal* solution can be obtained by considering (Gordon *et al* 1993)

$$\begin{aligned}p(x_k|y_{1:k-1}) &= \int p(x_k, x_{k-1}|y_{1:k-1}) dx_{k-1} \\ &= \int p(x_k|x_{k-1}, y_{1:k-1}) p(x_{k-1}|y_{1:k-1}) dx_{k-1}.\end{aligned}\tag{8}$$

From (6a) it follows that $p(x_k|x_{k-1}, y_{1:k}) = p(x_k|x_{k-1})$, and, therefore, the above equation can be simplified to get

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1}) p(x_{k-1}|y_{1:k-1}) dx_{k-1}.\tag{9}$$

This equation represents the *prediction* equation. Upon the receipt of the measurement y_k , we can derive the *update* equation as follows:

$$\begin{aligned}p(x_k|y_{1:k}) &= \frac{p(x_k, y_{1:k})}{p(y_{1:k})} = \frac{p(x_k, y_k, y_{1:k-1})}{p(y_k, y_{1:k-1})} \\ &= \frac{p(y_k|x_k, y_{1:k-1}) p(x_k|y_{1:k-1}) p(y_{1:k-1})}{p(y_k|y_{1:k-1}) p(y_{1:k-1})}.\end{aligned}\tag{10}$$

From (6b) it follows that $p(y_k|x_k, y_{1:k-1}) = p(y_k|x_k)$, and, by interpreting the denominator as a marginal PDF, one gets the updation equation as

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{\int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k}. \quad (11)$$

A recursive equation for the evolution of the multi-dimensional posteriori probability density function can be obtained by noting (Doucet *et al* 2000)

$$p(x_{0:k+1}|y_{1:k+1}) = \frac{p(x_{0:k+1}, y_{1:k+1})}{p(y_{1:k+1})} = \frac{p(y_{1:k+1}|x_{0:k+1})p(x_{0:k+1})}{p(y_{1:k+1})}. \quad (12)$$

Since $\{x_i\}_{i=1}^k$ constitutes a Markov sequence, it follows that

$$p(x_{0:k+1}) = p(x_{k+1}, x_{0:k}) = p(x_{k+1}|x_{0:k})p(x_{0:k}) = p(x_{k+1}|x_k)p(x_{0:k}). \quad (13)$$

Furthermore, by writing $p(y_{1:k+1}) = p(y_{k+1}, y_{1:k}) = p(y_{k+1}|y_{1:k})p(y_{1:k})$, (12) may be recast as

$$\begin{aligned} p(x_{0:k+1}|y_{1:k+1}) &= \frac{p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})} p(y_{1:k+1}|x_{0:k+1}) \frac{p(x_{0:k})}{p(y_{1:k})} \\ &= \frac{p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})} \frac{p(x_{0:k})}{p(y_{1:k})} p(y_{k+1}|x_{k+1}, x_{0:k}, y_{1:k}) p(y_{1:k}|x_{k+1}, x_{0:k}). \end{aligned} \quad (14)$$

From (6b) it follows that $p(y_{k+1}|x_{k+1}, x_{0:k}, y_{1:k}) = p(y_{k+1}|x_{k+1})$; thus we get,

$$\begin{aligned} p(x_{0:k+1}|y_{1:k+1}) &= \frac{p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})} \frac{p(x_{0:k})}{p(y_{1:k})} p(y_{k+1}|x_{k+1}) p(y_{1:k}|x_{k+1}, x_{0:k}) \\ &= \frac{p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})} \frac{p(x_{0:k})}{p(y_{1:k})} p(y_{k+1}|x_{k+1}) \frac{p(y_{1:k}, x_{0:k}|x_{k+1})}{p(x_{0:k})} \\ &= \frac{p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})} p(y_{k+1}|x_{k+1}) p(x_{0:k}|y_{1:k}, x_{k+1}) \frac{p(y_{1:k}|x_{k+1})}{p(y_{1:k})}. \end{aligned} \quad (15)$$

By noting that $p(x_{0:k}|y_{1:k}, x_{k+1}) = p(x_{0:k}|y_{1:k})$ and $p(y_{1:k}|x_{k+1}) = p(y_{1:k})$, we obtain the recursive formula,

$$p(x_{0:k+1}|y_{1:k+1}) = p(x_{0:k}|y_{1:k}) \frac{p(y_{k+1}|x_{k+1})p(x_{k+1}|x_k)}{p(y_{k+1}|y_{1:k})}. \quad (16)$$

The above equation for the multi-dimensional posteriori probability density function, or, (9) and (11) together, for the filtering density function, offer formal solutions to the estimation problem on hand. The actual solutions of these equations, for instance, by using numerical quadrature, are, however, infeasible because of the unwieldy dimensions of the integrals involved. Consequently, alternatives based on Monte Carlo simulation strategies have been proposed in the existing literature to tackle this problem.

3. Filtering via Monte Carlo simulations

In this section, we provide the details of formulation of three alternative filtering strategies, namely, the density based Monte Carlo filter (Tanizaki 1996), Bayesian bootstrap filter (Gordon *et al* 1993) and the sequential Monte Carlo sampling filters (also called the particle filters) (Doucet *et al* 2000).

3.1 Density based filters

Here we focus attention on determining $a_{k|k}$ and $\Sigma_{k|k}$, (7a,b), by using the equations

$$\begin{aligned} p(x_k|y_{1:k}) &= \left[\int p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k-1} \right] / \left[\int p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k} \right], \\ a_{k|k} &= \left[\int x_k p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k} \right] / \left[\int p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k} \right], \\ \Sigma_{k|k} &= \frac{\left[\int (x_k - a_{k|k})(x_k - a_{k|k})^t p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k} \right]}{\left[\int p(y_{1:k}|x_{0:k}) p(x_{0:k}) dx_{0:k} \right]}. \end{aligned} \quad (17a,b,c)$$

The basic idea of this method consists of the evaluation of the last two integrals in the above equation by using Monte Carlo simulations (Tanizaki 1996). For this purpose, samples of $x_{0:k}$ denoted by $\{x_{i,0:k}\}_{i=1}^N$ are drawn from $p(x_{0:k})$. This, in turn, is achieved by drawing $\{x_{i,0}\}_{i=1}^N$ from the initial PDF $p(x_0)$ and propagating the N samples through the process equation (6a) to obtain $\{x_{i,0:k}\}_{i=1}^N$. Here, it is assumed that $p(y_{1:k}|x_{0:k})$ is determinable based on (6b) and the known PDF of $\{v_k\}_{k=1}^N$. Thus we get

$$a_{k|k} = \left[\frac{1}{N} \sum_{i=1}^N x_{i,k} p(y_{1:k}|x_{i,0:k}) \right] / \left[\frac{1}{N} \sum_{i=1}^N p(y_{1:k}|x_{i,0:k}) \right]. \quad (18)$$

By noting that $p(y_{1:k}|x_{0:k}) = \prod_{s=1}^k p(y_s|x_s)$ the above equation can be written as

$$a_{k|k} = \left[\sum_{i=1}^N x_{i,k} \prod_{s=1}^k p(y_s|x_{i,s}) \right] / \left[\sum_{i=1}^N \prod_{s=1}^k p(y_s|x_{i,s}) \right]. \quad (19)$$

This process can be cast in a recursive format as follows:

$$\begin{aligned} a_{k|k} &= \sum_{i=1}^N x_{i,k} \omega_{i,k-1} \\ \omega_{i,k} &= \left[p(y_k|x_{i,k}) \omega_{i,k-1} \right] / \left[\sum_{j=1}^N p(y_k|x_{j,k}) \omega_{j,k-1} \right]; \\ \omega_{i,0} &= (1/N) \forall i = 1, 2, \dots, N. \end{aligned} \quad (20)$$

Similarly, the expression for the covariance is obtained as

$$\Sigma_{k|k} = \sum_{i=1}^N \{x_{i,k} - a_{k|k}\} \{x_{i,k} - a_{k|k}\}^t \omega_{i,k-1}. \quad (21)$$

This filter is applicable for any form of the nonlinearity in the process and measurement equations and also for general non-Gaussian models for the noise.

3.2 Bayesian bootstrap filter

The formulation here is as developed by Gordon *et al* (1993) and is based on an earlier result by Smith & Gelfand (1992). This result can be stated as follows (Gordon *et al* 1993): suppose that samples $\{x_i^*(i) : i = 1, 2, \dots, N\}$ are available from a continuous density function

$G(x)$ and that samples are required from the pdf proportional to $L(x)G(x)$, where $L(x)$ is a known function. The theorem states that a sample drawn from the discrete distribution over $\{x_i^*(i) : i = 1, 2, \dots, N\}$ with probability mass function $L(x_k^*(i))/\sum_{j=1}^N L(x_k^*(j))$ on $x_k^*(i)$, tends in distribution to the required density as $N \rightarrow \infty$.

The algorithm for implementing the filter is as follows:

- (1) Set $k = 0$. Generate $\{x_{i,0}^*\}_{i=1}^N$ from the initial pdf $p(x_0)$ and $\{w_{i,0}\}_{i=1}^N$ from $p(w)$.
- (2) Obtain $\{x_{i,k+1}^*\}_{i=1}^N$ by using the process equation (6a).
- (3) Consider the k th measurement y_k . Define $q_i = p(y_k|x_{i,k}^*)/\{\sum_{j=1}^N p(y_k|x_{j,k}^*)\}$.
- (4) Define the probability mass function $P[x_k(j) = x_{i,k}^*] = q_i$; generate N samples $\{x_{i,k}\}_{i=1}^N$ from this discrete distribution.
- (5) Evaluate $a_{k|k} = (1/N) \sum_{i=1}^N x_{i,k}$ and $\Sigma_{k|k} = [1/(N-1)] \sum_{i=1}^N (x_{i,k} - a_{k|k})(x_{i,k} - a_{k|k})^t$.
- (6) Set $k = k + 1$ and go to step 2 if $k < n$; otherwise, stop.

The central contention in the formulation of this algorithm is that the samples $\{x_{i,k}\}_{i=1}^N$ drawn in step 4 above are approximately distributed as per the required PDF $p(x_k|y_{0:k})$. The justification for this is provided by considering the updation equation (11) in conjunction with the result due to Smith and Gelfand (stated above) with $G(x)$ identified with $p(x_k|y_{1:k-1})$ and $L(x)$ with $p(x_k|y_k)$. This method has the advantage of being applicable to nonlinear processes and measurement equations, and, non-Gaussian noise models. It is also important note that the samples in this algorithm naturally get concentrated near regions of high probability densities.

3.3 Filtering using sequential importance sampling

As has already been noted, the basic aim of state estimation problem is to evaluate $p(x_{0:k}|y_{1:k})$, its associated marginal PDFs such as $p(x_k|y_{1:k})$ and expectations of the form

$$I[\psi_k] = E_p[\psi(x_{0:k})|y_{1:k}] = \int \psi(x_{0:k}) p(x_{0:k}|y_{1:k}) dx_{0:k}. \quad (22)$$

Here $E_p[\cdot]$ denotes the expectation operator defined with respect to the PDF $p(x_{0:k}|y_{1:k})$. In this section we follow the procedure given by Doucet *et al* (2000) to formulate a filter based on importance sampling to evaluate the integral in the above equation. The basic idea here consists of choosing an importance sampling PDF, denoted by $\pi(x_{0:k}|y_{1:k})$, such that, if $p(x_{0:k}|y_{1:k}) > 0 \Rightarrow \pi(x_{0:k}|y_{1:k}) > 0$. In this case the integral in (22) can be written as

$$\begin{aligned} I[\psi_k] &= \int \psi(x_{0:k}) [p(x_{0:k}|y_{1:k})/\pi(x_{0:k}|y_{1:k})] \pi(x_{0:k}|y_{1:k}) dx_{0:k} \\ &= \int \psi(x_{0:k}) w^*(x_{0:k}) \pi(x_{0:k}|y_{1:k}) dx_{0:k} = E_\pi[\psi(x_{0:k}) w^*(x_{0:k})|y_{1:k}]. \end{aligned} \quad (23)$$

Here $E_\pi[\cdot]$ denotes expectation with respect to the PDF $\pi(x_{0:k}|y_{1:k})$. If $\{x_{i,0:n}\}_{i=1}^N$ are samples drawn from $\pi(x_{0:k}|y_{1:k})$, and

$$\hat{I}_N[\psi_{0:k}] = \frac{1}{N} \sum_{i=1}^N \psi(x_{i,0:k}) w^*(x_{i,0:k}),$$

with

$$\begin{aligned} w^*(x_{i,0:k}) &= \frac{p(x_{i,0:k}|y_{0:k})}{\pi(x_{i,0:k}|y_{1:k})} = \frac{p(y_{0:k}|x_{i,0:k}) p(x_{i,0:k})}{\pi(x_{i,0:k}|y_{1:k}) p(y_{1:k})} \\ &= \frac{p(y_{0:k}|x_{i,0:k}) p(x_{i,0:k})}{\pi(x_{i,0:k}|y_{1:k}) \int p(y_{0:k}|x_{0:k}) p(x_{0:k}) dx_{0:k}}, \end{aligned} \quad (24)$$

it can be shown that $\lim_{N \rightarrow \infty} \hat{I}_N[\psi_k] \rightarrow I[\psi_k]$ almost surely (Doucet 1998). It may be emphasized that the evaluation of the normalization constant $p(y_{1:k}) = \int p(y_{0:k}|x_{0:k})p(x_{0:k})dx_{0:k}$ in (24) presents a major difficulty in implementing the above procedure since it is seldom possible to express this constant in a closed form. However, if we evaluate this integral too *via* importance sampling, we get,

$$\begin{aligned} \hat{I}_N[\psi_{0:k}] &= \left[\frac{1}{N} \sum_{i=1}^N \psi(x_{i,0:k}) \frac{p(y_{0:k}|x_{i,0:k})p(x_{i,0:k})}{\pi(x_{i,0:k}|y_{1:k})} \right] / \left[\frac{1}{N} \sum_{j=1}^N \frac{p(y_{0:k}|x_{j,0:k})p(x_{j,0:k})}{\pi(x_{j,0:k}|y_{1:k})} \right] \\ &= \frac{\sum_{i=1}^N \psi(x_{i,0:k})w(x_{i,0:k})}{\sum_{j=1}^N w(x_{j,0:k})} = \sum_{i=1}^N \psi(x_{i,0:k})\tilde{w}(x_{i,0:k}), \end{aligned} \quad (25)$$

with the new weights $\tilde{w}(x_{i,0:k})$ given by

$$\tilde{w}(x_{i,0:k}) = [w(x_{i,0:k})] / \left[\sum_{j=1}^N w(x_{j,0:k}) \right],$$

with

$$w(x_{i,0:k}) = [p(y_{1:k}|x_{i,0:k})p(x_{i,0:k})] / [\pi(x_{i,0:k}|y_{1:k})]. \quad (26)$$

If we select the importance sampling density function of the form

$$\pi(x_{0:k}|y_{1:k}) = \pi(x_{0:k-1}|y_{1:k})\pi(x_k|x_{0:k-1}, y_{1:k}), \quad (27)$$

it would be possible to compute the importance weights in a recursive manner. To see this, we begin by noting that $p(y_{1:k}|x_{0:k}) = \prod_{j=1}^k p(y_j|x_j)$ and $p(x_{0:k}) = p(x_0) \prod_{j=1}^n p(x_j|x_{j-1})$. Using these facts along with (27) in the second equation of (26) we get

$$\begin{aligned} w(x_{0:k}) &= \frac{\prod_{j=1}^k p(y_j|x_j)p(x_0) \prod_{j=1}^k p(x_j|x_{j-1})}{\pi(x_{0:k-1}|y_{1:k})\pi(x_k|x_{0:k-1}, y_{1:k})} \\ &= \frac{\prod_{j=1}^{k-1} p(y_j|x_j)p(x_0) \prod_{j=1}^{k-1} p(x_j|x_{j-1})}{\pi(x_{0:k-1}|y_{1:k})} \frac{p(y_k|x_k)p(x_k|x_{k-1})}{\pi(x_k|x_{0:k-1}, y_{1:k})} \\ &= w(x_{0:k-1}) \frac{p(y_k|x_k)p(x_k|x_{k-1})}{\pi(x_k|x_{0:k-1}, y_{1:k})}. \end{aligned}$$

One of the difficulties in implementing this algorithm in practice is that after a few time steps, most weights become small in comparison to a few which become nearly equal to unity. This implies that most of the computational effort gets wasted on trajectories which do not eventually contribute to the final estimate. This problem is widely discussed in the existing literature and one way to remedy this problem is to introduce the notion of an effective sample size (Doucet 1998) given by

$$N_{\text{eff}} = 1 / \sum_{i=1}^N \{\tilde{w}_{i,0:k}\}^2. \quad (28)$$

It may be noted that if all weights are equal, $N_{\text{eff}} = N$ and if all weights excepting one are zero, $N_{\text{eff}} = 1$. Thus, when the effective sample at any time step falls below a threshold N_{thres} , a step involving resampling is implemented with a view to mitigate the effect of depletion of samples.

Furthermore, the choice of the importance sampling density function plays a crucial role in the success of the algorithm. It has been shown (Doucet 1998) that $\pi(x_k | x_{i,0:k-1}, y_{1:k}) = p(x_k | x_{i,k-1}, y_k)$ is the importance function, which minimizes the variance of the non-normalized weights, $w_{i,k}^*$, conditioned upon $x_{i,0:k-1}$ and $y_{1:k}$. This can be shown by noting that

$$\begin{aligned} E_{\pi}[w(x_{0:k}) | x_{i,0:k}, y_{1:k}] &= \int \frac{p(y_k | x_k) p(x_k | x_{i,k-1})}{\pi(x_k | x_{i,0:k-1}, y_{1:k})} \pi(x_k | x_{i,0:k-1}, y_{1:k}) dx_k \\ &= \int p(y_k | x_k) p(x_k | x_{i,k-1}) dx_k \\ &= p(y_k | x_{i,k-1}). \end{aligned} \quad (29)$$

Similarly, the conditional variance of the weights can be obtained as

$$\text{Var}_{\pi}[w(x_{0:k}) | x_{i,0:k}, y_{1:k}] = E_{\pi} \left[\left\{ \frac{p(y_k | x_k) p(x_k | x_{i,k-1})}{\pi(x_k | x_{i,0:k-1}, y_{1:k})} - p(y_k | x_{i,k-1}) \right\}^2 \right]. \quad (30)$$

It follows that this variance will be zero when

$$\pi(x_k | x_{i,0:k-1}, y_{1:k}) = \frac{p(y_k | x_k) p(x_k | x_{i,k-1})}{p(y_k | x_{i,k-1})}. \quad (31)$$

By noting that $p(y_k | x_k) = p(y_k | x_k, x_{i,0:k-1})$, we can write

$$\begin{aligned} \pi(x_k | x_{i,0:k-1}, y_{1:k}) &= \frac{p(y_k | x_k, x_{i,0:k-1}) p(x_k | x_{i,k-1})}{p(y_k | x_{i,k-1})} \\ &= \frac{p(x_k, y_k | x_{i,0:k-1}) p(x_k | x_{i,k-1})}{p(x_k | x_{i,k-1}) p(y_k | x_{i,k-1})} \\ &= p(x_k | y_k, x_{i,k-1}). \end{aligned} \quad (32)$$

In general, it is not feasible to deduce the ideal importance sampling density function $p(x_k | y_k, x_{i,k-1})$. However, when the process and measurement equations assume the form,

$$\begin{aligned} x_k &= f(x_{k-1}) + v_k; & v_k &\equiv N[0_{n_v \times 1} \quad \Sigma_v], \\ y_k &= Cx_k + w_k; & w_k &\equiv N[0_{n_w \times 1} \quad \Sigma_w], \end{aligned} \quad (33)$$

it can be shown that the ideal importance density function is a normal density function with mean m_k and covariance Σ given by,

$$\begin{aligned} m_k &= \Sigma(\Sigma_v^{-1} f(x_{k-1}) + C^t \Sigma_w^{-1} y_k), \\ \Sigma^{-1} &= \Sigma_v^{-1} f(x_{k-1}) + C^t \Sigma_w^{-1} C. \end{aligned} \quad (34)$$

It may be noted that $x \equiv N(a, B)$ denotes that x is a vector of normal random variables with mean vector a and covariance matrix B . Equations of the type (33) are very likely to be

encountered in problems of structural system identification and, therefore, the above result is of notable significance. When dealing with more general forms of process and measurement equations, as in (6a,b), suitable strategies need to be evolved in choosing the importance function. In the present study, however, we limit our attention to application of this method to equations of the form (33).

Thus, in order to implement the filtering with sequential importance sampling the following steps are adopted:

- (1) Set $k = 0$; draw samples $\{x_{i,0}\}_{i=1}^N$ from $p(x_0)$.
- (2) Set $k = k + 1$;
- (3) Sample $\{\hat{x}_{i,k}\}_{i=1}^N$ from $\pi(x_k | x_{i,0:k-1}, y_{1:k})$ and set $\hat{x}_{i,0:k} = \{x_{i,0:k-1}, \hat{x}_{i,k}\}$ for $i \in [1, N]$.
- (4) For $i \in [1, N]$ compute the importance weights

$$w_{i,k}^* = w_{i,k-1}^* [p(y_k | \hat{x}_{i,k}) p(\hat{x}_{i,k} | \hat{x}_{i,k-1})] [\pi(\hat{x}_{i,k} | \hat{x}_{i,0:k-1}, y_{1:k})]$$

- (5) For $i \in [1, N]$, normalize the weights

$$\tilde{w}_{i,k} = w_{i,k}^* / \sum_{j=1}^N w_{j,k}^*$$

- (6) Evaluate the expectations of interest, for instance,

$$a_{k|k} = \sum_{i=1}^N x_{i,k} \tilde{w}_{i,k}; \quad \Sigma_{k|k} = \sum_{i=1}^N (x_{i,k} - a_{k|k})(x_{i,k} - a_{k|k})^t \tilde{w}_{i,k}$$

- (7) Evaluate the effective sample N_{eff} using (28).
- (8) If $N_{\text{eff}} > N_{\text{thres}}$, set $x_{i,0:k} = \hat{x}_{i,0:k}$; go to step 2 if $k < n$, otherwise end.
- (9) If $N_{\text{eff}} \leq N_{\text{thres}}$, implement the following:
 - For $i \in [1, N]$, sample an index $j(i)$ distributed according to the discrete distribution with N elements satisfying $P[j(i) = l] = \tilde{w}_{l,k}$ for $l = 1, 2, \dots, N$.
 - For $i \in [1, N]$, set $x_{i,0:k} = \tilde{x}_{j(i),0:k}$ and $w_{i,k} = 1/N$.
 - Go to step 2 if $k < n$; otherwise, end.

In implementing this method it is assumed that it is possible to analytically deduce $p(x_k | x_{k-1})$ and $p(y_k | x_k)$ using (33).

4. Discretization of the system or observation equations

In the context of dynamical systems of interest in structural dynamics, the system equations are often derived using a continuum approach based on either a variational principle or a Newtonian force balance. Such equations are almost always in the form of a set of partial differential equations (PDEs) that are spatially discretized within a finite element framework to finally result in a system of ordinary differential equations (ODEs). However, in the context of a filtering problem, these ODEs are taken to be polluted with a system of noise processes, modelled through a set of white noise processes, as in (1). Such ODEs are referred to as stochastic differential equations (SDEs) and we are then required to appropriately discretize the SDEs so as to arrive at discrete maps as in (6) for further processing using nonlinear filtering

algorithms. A white noise process mathematically exists only as a measure and not in a path-wise sense and as such it is formally defined as the ‘derivative’ of a Wiener process, which is itself not differentiable (appendix A). Accordingly, it is often preferable to write a system of SDEs in an incremental form involving only increments of Brownian noise processes (see, for instance, (36) below). Just as a deterministic ODE is discretized into a map (such as the Newmark or the Runge–Kutta map) using variations of a Taylor expansion, discretization of the system and observation SDEs similarly requires the application of the stochastic Taylor expansion (appendix B contains a reasonably self-contained account of how such expansions can be derived for functions of solutions of a given system of SDEs).

As applications of the filters described in the previous section, we presently consider two single-degree of freedom nonlinear oscillators, namely, the Duffing oscillator and one with Coulomb friction damping. The equations of motion for these oscillators respectively, are of the form

$$\begin{aligned} \ddot{x} + 2\eta\omega\dot{x} + \omega^2x + \alpha x^3 &= f(t); & x(0) &= x_0, \dot{x}(0) = \dot{x}_0, \\ \ddot{x} + \mu \operatorname{sgn}(\dot{x}) + \omega^2x &= f(t); & x(0) &= x_0, \dot{x}(0) = \dot{x}_0. \end{aligned} \quad (35a,b)$$

In particular, we consider the problem of identifying the parameters α and μ based on noisy observations made either on the displacement response or on the reaction transferred to the support. We begin by considering the Duffing oscillator where the augmented governing process equations may be written in the following incremental form,

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= (-2\eta\omega x_2 - \omega^2 x_1 + x_3 x_1^3 + P \cos \lambda t) dt + \sigma dB_1(t), \\ dx_3 &= dB_2(t), \\ x_1(0) &= x_{10}; x_2(0) = x_{20}; x_3(0) = 0. \end{aligned} \quad (36)$$

Now, following the procedure as outlined in appendix B, we may write the following (explicit forms of) stochastic Taylor expansion for the displacement and velocity components over the interval $(t_k, t_{k+1}]$ with the step-size $h = t_{k+1} - t_k$ being uniform:

$$\begin{aligned} x_{1(k+1)} &= x_{1k} + x_{2k}h + a_{2k} \frac{h^2}{2} - P\lambda \sin \lambda t_k \frac{h^3}{6} - \omega^2 x_{2k} \frac{h^3}{6} - x_{3k} x_{1k}^2 x_{2k} \frac{h^3}{2} \\ &\quad + \sigma I_{10} + 2\eta\omega\sigma I_{100} + \rho_{1k}, \\ x_{2(k+1)} &= x_{2k} + \frac{P}{\lambda} (\sin \lambda t_{k+1} - \sin \lambda t_k) - 2\eta\omega \left(x_{2k}h + a_{2k} \frac{h^2}{2} \right) \\ &\quad - \omega^2 \left(x_{1k}h + x_{2k} \frac{h^2}{2} \right) - x_{3k} \left(x_{1k}^3 h + 3x_{1k}^2 x_{2k} \frac{h^2}{2} \right) \\ &\quad + \sigma I_1 - 2\eta\omega\sigma I_{10} + \rho_{2k}, \\ x_{3(k+1)} &= x_{3k} + \sigma_\alpha I_2, \\ a_{2k} &= -2\eta\omega x_{2k} - \omega^2 x_{1k} - x_{3k} x_{1k}^3 + P \cos \lambda t_k, \end{aligned} \quad (37)$$

where $I_1 = \int_{t_k}^{t_{k+1}} dB_1(t)$, $I_2 = \int_{t_k}^{t_{k+1}} dB_2(t)$ etc. are the multiple stochastic integrals (MSIs) as explained in appendix B and ρ_{1k} , ρ_{2k} are the remainder terms associated with the displacement and velocity maps respectively. The displacement and velocity maps are then given by

the first two of equations (37) without considering (i.e., by truncating) the remainder terms. Note that the remainder terms merely decide the local and global error orders associated with the maps. Before these stochastic maps can be numerically implemented, one also needs to model the MSIs, which are all Ito integrals and hence zero-mean Gaussian random variables. It is then only needed to find the covariance matrix for the MSIs. We may use Ito's formula, restricted to the interval $(t_k, t_{k+1}]$, to determine the elements of this covariance matrix. To explain this further, consider the following system of scalar SDEs:

$$du = dB_1(t), \quad dv = udt \quad \text{and} \quad dw = vdt \quad (38)$$

subject to trivial (zero) initial conditions. It immediately follows that $w_{k+1} = I_{100}$. Now, use of Ito's formula leads to the following scalar SDE for $w^2(t)$:

$$dw^2 = 2vw dt, \quad (39)$$

so that $E[w_{k+1}^2] = E[I_{100}^2] = \int_{t_k}^{t_{k+1}} E(vw) dt$. Similarly, we may write down the following SDE for vw :

$$dvw = (uw + v^2)dt, \text{ or, } E[vw]_{k+1} = E[I_{10}I_{100}] = \int_{t_k}^{t_{k+1}} E(uw + v^2)dt. \quad (40)$$

We also need to write down the following SDEs for uw , v^2 and uv :

$$d(uw) = uvdt + wdB_1, \text{ or, } E[uw]_{k+1} = E[I_1I_{100}] = \int_{t_k}^{t_{k+1}} E(uv)dt, \quad (41)$$

$$d(uv) = u^2dt + vdB_1, \text{ or, } E[uv]_{k+1} = E[I_1I_{10}] = \int_{t_k}^{t_{k+1}} E(u^2)dt = h^2/2, \quad (42)$$

$$d(v^2) = 2uvdt, \text{ or, } E[v^2]_{k+1} = E[I_{10}^2] = 2 \int_{t_k}^{t_{k+1}} E(uv)dt = h^3/3. \quad (43)$$

Now, working backwards, we can readily obtain expressions for rest of the elements of the covariance matrix. Finally, we may thus claim that I_1 , I_{10} , I_{100} and I_2 are normal random variables of the form

$$\begin{Bmatrix} I_1 \\ I_{10} \\ I_{100} \\ I_2 \end{Bmatrix} \equiv N \left[\mathbf{0}_{4 \times 1}, \begin{bmatrix} h & \frac{h^2}{2} & \frac{h^3}{6} & 0 \\ \frac{h^2}{2} & \frac{h^3}{3} & \frac{h^4}{8} & 0 \\ \frac{h^3}{6} & \frac{h^4}{8} & \frac{h^5}{20} & 0 \\ 0 & 0 & 0 & h \end{bmatrix} \right]. \quad (44)$$

Finally, a word about the local and global error orders corresponding to the stochastic maps (37) would be in order. Following Milstein (1995), one may show that the local and global orders of accuracy of the displacement map are respectively $O(h^3)$ and $O(h^{2.5})$ respectively. Similarly, for the velocity map, these orders are respectively $O(h^2)$ and $O(h^{1.5})$.

We may now consider the Coulomb oscillator and, as before, begin by recasting the equations in the following incremental form:

$$\begin{aligned}
 dx_1 &= x_2 dt, \\
 dx_2 &= (-x_3 \operatorname{sgn}(x_2) - \omega^2 x_1 + P \cos \lambda t) dt + \sigma dW_1(t), \\
 dx_3 &= dW_2(t), \\
 x_1(0) &= x_{10}; x_2(0) = x_{20}; x_3(0) = 0.
 \end{aligned} \tag{45}$$

A notable difference in this oscillator as compared to the Duffing oscillator is that the present system has a signum function that is not even continuous everywhere. Thus a direct application of stochastic Taylor expansion to obtain a map with a similar order of accuracy (as in the previous case) is ruled out. In order to avoid a direct usage of Ito's formula, we may apply the stochastic Heun scheme (Gard 1988) to obtain the velocity update first and then attempt to derive a higher order map for the displacement update using a combined stochastic Taylor expansion and the available information on the velocity update. Thus, the velocity map may be written as:

$$\begin{aligned}
 x_{2(k+1)} &= x_{2k} - x_{3k} \operatorname{sgn} \left\{ \frac{1}{2} (x_{2k} + x_{2(k+1)}^e) \right\} h \\
 &\quad - \omega^2 \left(x_{1k} h + x_{2k} \frac{h^2}{2} \right) + \frac{P}{\lambda} (\sin \lambda t_{k+1} - \sin \lambda t_k) + \sigma I_1,
 \end{aligned} \tag{46}$$

where,

$$\begin{aligned}
 a_{2k} &= -x_{3k} \operatorname{sgn}(x_{2k}) - \omega^2 x_{1k} + P \cos \lambda t_k, \\
 x_{2(k+1)}^e &= x_{2k} + a_{2k} h + \sigma I_1.
 \end{aligned} \tag{47}$$

At this stage, the displacement map may be derived using an application of the stochastic Taylor expansion with the distinction that any differentiation with respect to the signum function is to be replaced by a discrete gradient. The displacement map thus takes the following form:

$$\begin{aligned}
 x_{1(k+1)} &= x_{1k} + x_{2k} h + a_{21} \frac{h^2}{2} - \omega^2 x_{2k} \frac{h^3}{6} - x_{3k} \frac{\operatorname{sgn}(x_{2(k+1)}) - \operatorname{sgn}(x_{2k})}{x_{2(k+1)} - x_{2k}} \frac{h^3}{6} \\
 &\quad + P \cos \lambda t_k \frac{h^3}{6} + \sigma I_{10} + \sigma \frac{a_{2(k+1)}^e - a_{2k}}{x_{2(k+1)} - x_{2k}} I_{100}, \\
 x_{3(k+1)} &= x_{3k} + \sigma_\mu I_2, \\
 a_{2(k+1)}^e &= -x_{3k} \operatorname{sgn}(x_{2(k+1)}) - \omega^2 x_{1(k+1)}^e + P \cos \lambda t_{k+1}, \\
 x_{1(k+1)}^e &= x_{1k} + \frac{1}{2} (x_{2k} + x_{2(k+1)}) h.
 \end{aligned} \tag{48}$$

Note that I_1 , I_{10} , I_{100} and I_2 are the same normal random variables with properties as given in (44). A detailed analysis of the orders of accuracy of the displacement and velocity maps is beyond the scope of the present study and we shall take it up elsewhere.

5. Numerical results and discussion

We consider the problem of determining the nonlinearity parameters α and μ in the above models based on the density based Monte Carlo filter (filter 1), bootstrap filter (filter 2) and sequential importance sampling filter (filter 3). In implementing these filters, the measurements either on the displacement response or on the reaction transferred to the support are assumed to be made. In the present study, these measurement time histories are synthetically generated.

5.1 Duffing's oscillator

The measurement equation is taken to be of one of the following forms

$$y_k = 2\eta\omega x_{2k} + \omega^2 x_{1k} + x_3 x_{1k}^3 + \sigma_m I_m; \quad I_m \equiv N(0, 1),$$

$$y_k = x_k + \sigma_m I_m; \quad I_m \equiv N(0, 1).$$

Thus, in this example the process equation is nonlinear and the measurement equation could be linear/nonlinear depending upon the choice of the measurement equation made and the noise processes being additive. In the numerical work we take $\eta = 0.04$, $\omega = 4\pi$ rad/s, $h = (2\pi/80\lambda)$ s, $\lambda = 0.8\omega$, $\sigma = P/10$, $P = 10$ N, and $n = 400$. The initial displacement and velocity are taken to be zero mean, mutually independent Gaussian random variables, with standard deviations $\sigma_{10} = 0.02$ m and $\sigma_{20} = \lambda\sigma_{10}$ respectively. Synthetic data on the reaction transferred to the supports is generated with an assumed value of $\tilde{\alpha} = 1.5791 \times 10^5$ N/m³. This time history is used as the measured response for implementing filters 1 and 2. The measurement noise here is taken to have zero mean and standard deviation $\sigma_m = 2.2992$ N. In implementing filter 3, the importance sampling density function as given in (34) is used. This, as has been already noted, is valid only when the measurement equation is linear in the state variables. Thus, for implementing filter 3 it is assumed that the measurement is made on displacement of the oscillator, (49b) with measurement noise having zero mean

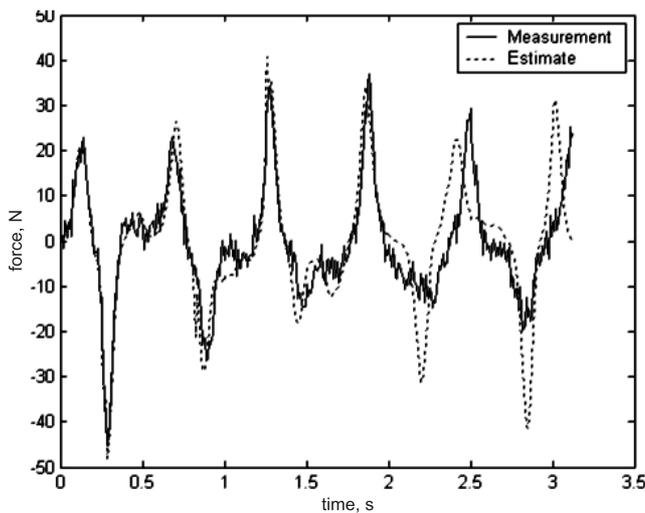


Figure 1. Study on the Duffing oscillator; filter 1; the measured and estimated reaction transferred to the support.

and $\sigma_m = 0.0037$ m. In the filter computations, it is assumed that $\sigma_\alpha = 0.05\tilde{\alpha}$. In the implementations of the filters it becomes necessary to assume the PDF of x_3 at $k = 0$. The results in this paper are presented for the case, when x_3 at $k = 0$ is taken to be independent of x_1 and x_2 , at $k = 0$ and it is assumed that x_3 is distributed uniformly in the range $[0.8\tilde{\alpha}, 1.6\tilde{\alpha}]$. Figures 1–7 show the results obtained using filters 1–3. Figures 1 and 3 show the estimate of the force transferred to the support obtained using the filters 1 and 2 respectively; similar results on estimate of displacement obtained using filter 3 are shown in figure 5. Also shown in these figures is the time history of the measured force (figures 1,3)/displacement (figure 5). Similarly, results on the estimate of α , and the associated standard deviation, obtained using the three filters, are shown in figures 2, 4 and 6. It was observed that these estimates depended

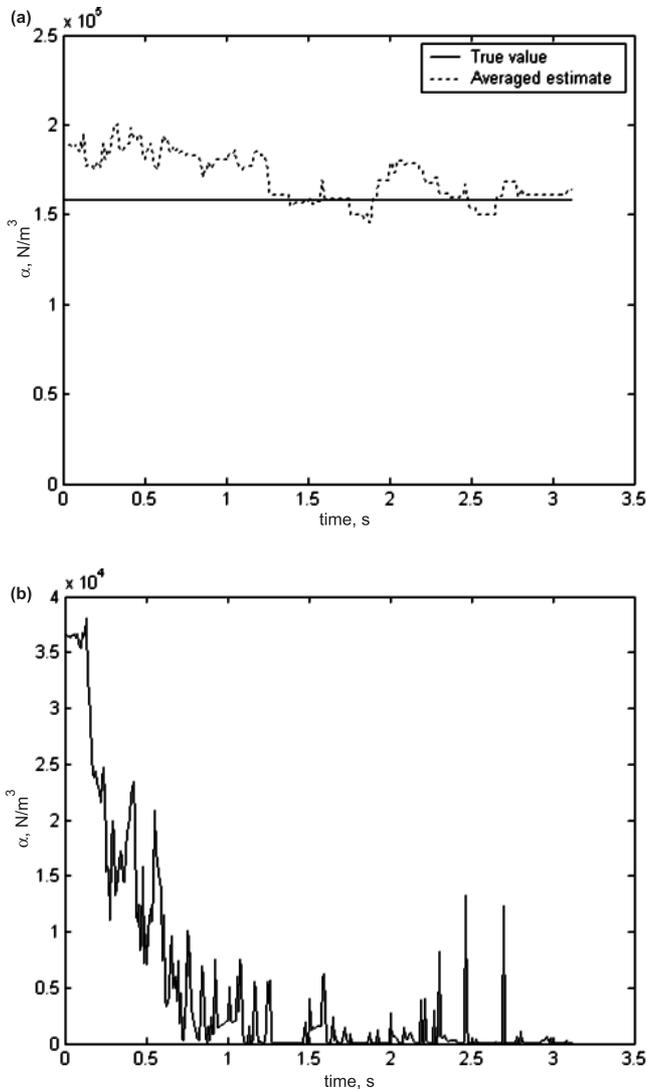


Figure 2. Study on the Duffing oscillator; filter 1; $\alpha_0 \equiv U(0.8\tilde{\alpha}, 1.6\tilde{\alpha})$; $\tilde{\alpha} = 1.5791 \times 10^5$ N/m³; mean of $\{x_{i,3}\}$ across 10 realizations = 1.7035×10^5 N/m³, estimates on mean of α (a) and standard deviation of α (b).

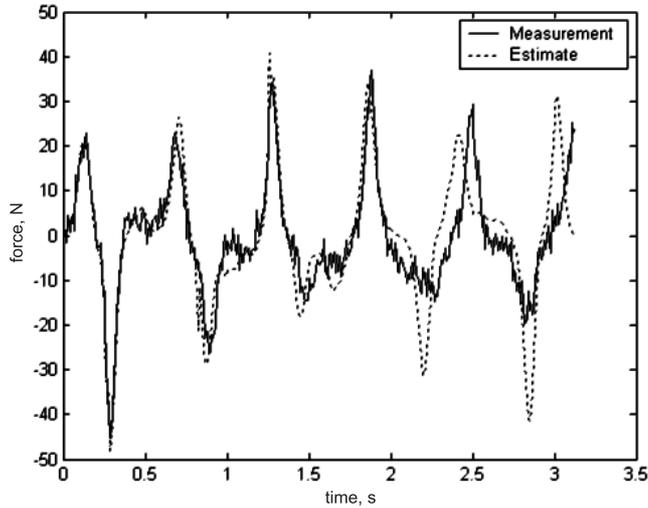


Figure 3. Study on the Duffing oscillator; filter 2; the measured and estimated reaction transferred to the support.

upon (a) choice made on the PDF of the variable at $t = 0$, (b) the sample size used, (c) the threshold sample size used for invoking the resampling (in filter 3). These factors have more pronounced effects upon the estimate of the nonlinearity parameter than on the estimate of the force/displacement. In implementing filter 1, the sampling fluctuations associated with estimates for α , for a given sample size, is considered. The results shown in figures 2a and 2b correspond to an ensemble average across 10 estimates that are obtained by running the filter 10 times each time with a sample size of 500. In implementing filter 3, the threshold of sample size to decide upon the need for resampling was taken to be 1/3rd the sample

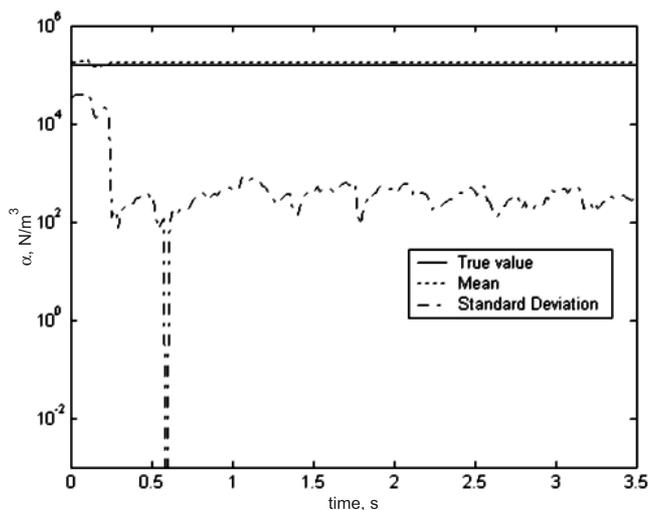


Figure 4. Study on the Duffing oscillator; filter 2; estimates on α ; $\alpha_0 \equiv U(0.8\tilde{\alpha}, 1.6\tilde{\alpha})$; $\tilde{\alpha} = 1.5791 \times 10^5 \text{ N/m}^3$; $\text{mean}\{x_{i,3}\} = 1.7666 \times 10^5 \text{ N/m}^3$.

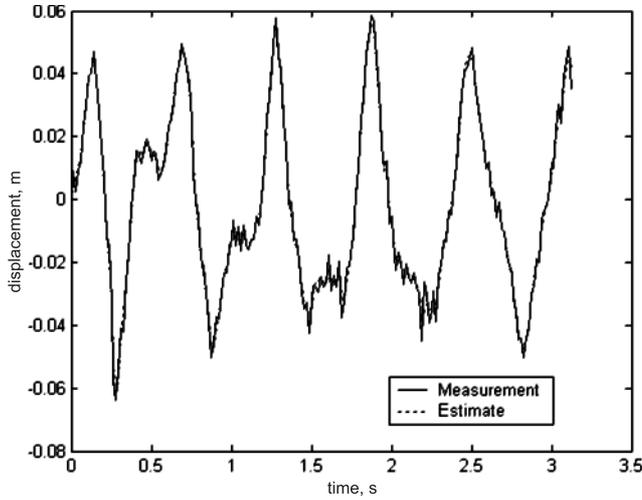


Figure 5. Study on the Duffing oscillator; filter 3; the measured and estimated displacement.

size used. In numerical implementation, it is observed that at 64 out of 200 time instants, resampling was needed. The time-averaged mean of the mean estimate of α using the three methods is obtained in the range of about 1.65×10^5 to 1.7×10^5 N/m³. Given the presence of measurement and process noise, this estimate can be considered acceptable approximation to the true value of $\tilde{\alpha} = 1.5791 \times 10^5$ N/m³.

5.2 Coulomb's oscillator

In the numerical work we take $\omega = 4\pi$ rad/s, $h = (2\pi/200\lambda)$ s, $\lambda = 0.8\omega$, $\sigma = P/10$, $P = 10$ N, and $n = 1000$. The initial displacement and velocity are taken to be zero mean, mutually

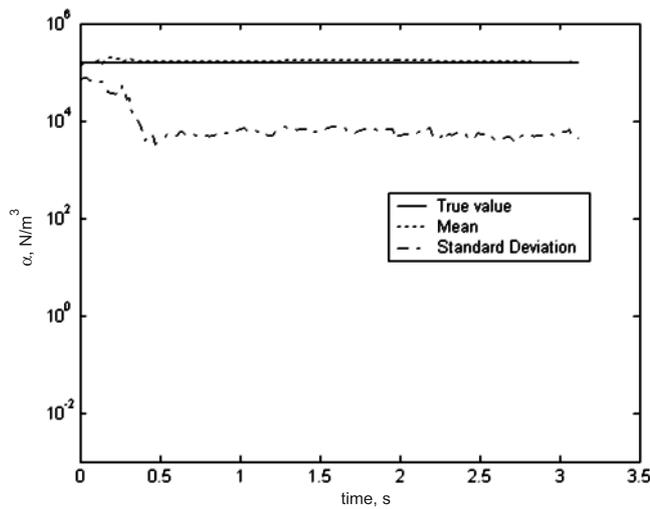


Figure 6. Study on the Duffing oscillator; filter 1; estimates on α ; $\alpha_0 \equiv U(0.8\tilde{\alpha}, 1.6\tilde{\alpha})$; $\tilde{\alpha} = 1.5791 \times 10^5$ N/m³; $\text{mean}\{x_{i,3}\} = 1.6997 \times 10^5$ N/m³.

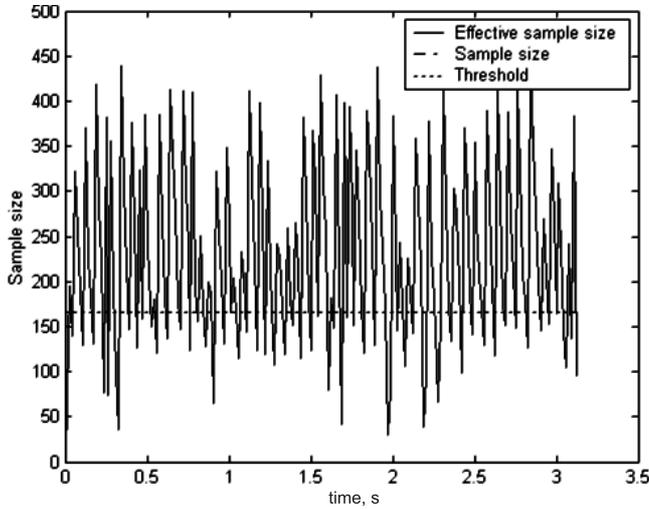


Figure 7. Study on the Duffing oscillator; filter 3; details of time instants at which resampling was necessary; $N = 500$; $N_{\text{thres}} = N/3$; number of resampling steps = 64 out of possible 200 slots.

independent Gaussian random variables with standard deviations $\sigma_{10} = 0.02$ m and $\sigma_{20} = \lambda\sigma_{10}$ respectively. The measurement equation is taken to be of the form,

$$y_k = -x_{3k}\text{sgn}(x_{2k}) + \omega^2 x_{1k} + \sigma_m I_m; \quad I_m \equiv N(0, 1). \quad (49)$$

The measurement noise is taken to have zero mean and standard deviation $\sigma_m = 3.6790$ N. Synthetic data on the reaction transferred to the supports is generated with an assumed value of $\tilde{\mu} = 0.05$. In the filter computations it is assumed that $\sigma_\alpha = 0.05\tilde{\mu}$. It is assumed that x_3 at $k = 0$ is taken to be independent of x_1 and x_2 at $k = 0$ and also that it is distributed uniformly in the range $[0.02\tilde{\mu}, 0.08\tilde{\mu}]$. As has been already noted, the application of filter 3 requires the determination of $p(x_{k+1}|x_k)$ and $p(y_k|x_k)$. In the present case, a scrutiny of (48) would reveal that in the equation for $x_{1(k+1)}$ there appear terms in which the noise term σI_1 appears as an argument of a signum function. Consequently, the problem of determining $p(x_{k+1}|x_k)$ in this case becomes complicated. In the present study we do not address this issue, and, therefore, results here are presented only for filters 1 and 2. These results are shown in figure 8–15. Figures 8 and 13 show the estimate of the mean displacement along with the measured signal. Figures 9 and 10 (filter 1) and figure 14 (filter 2) show the results on mean and standard deviation of estimate of μ . These results are obtained for one simulation run with 1000 samples. As is obvious, these estimates are subject to sampling fluctuations. Figures 11 and 12 (filter 1) and figure 15 (filter 2) show the ensemble averages, across 10 simulation runs, each with 1000 samples, of the estimates on mean and standard deviation of μ . These averaged results, especially on standard deviation of estimate of μ , are seen to be more well behaved than individual results from single runs of the simulations. Using the two filters, the results on the estimate of μ are found to be in the range of 0.055 to 0.068: this, again, can be considered as being reasonable approximations to the true value of $\mu = 0.05$.

6. Conclusions

Nonlinear filtering techniques have received considerable attention in the fields of economics, image processing, object tracking, robotics, automatic control, and applied statistics. The

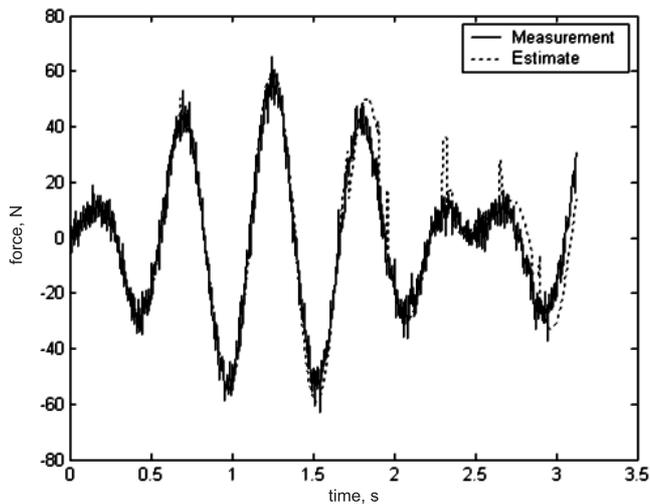


Figure 8. Study on Coulomb oscillator; filter 1; measured and estimated reaction transferred to the support.

existing literature on methods of nonlinear filtering is therefore quite vast and these developments have a significant bearing on many structural engineering problems—system identification, structural health monitoring using vibration signatures, earthquake strong motion data processing, and development of active control strategies for earthquake/wind loaded structures, to name a few. Such applications, however, do not seem to have received much attention in the existing structural dynamics literature and the present paper is an effort in demonstrating the utility of these filtering tools in nonlinear system identifications of relevance in structural mechanics. Moreover, in the context of structural dynamics, the process equations are often derived from a continuum approach and thus these equations are often in

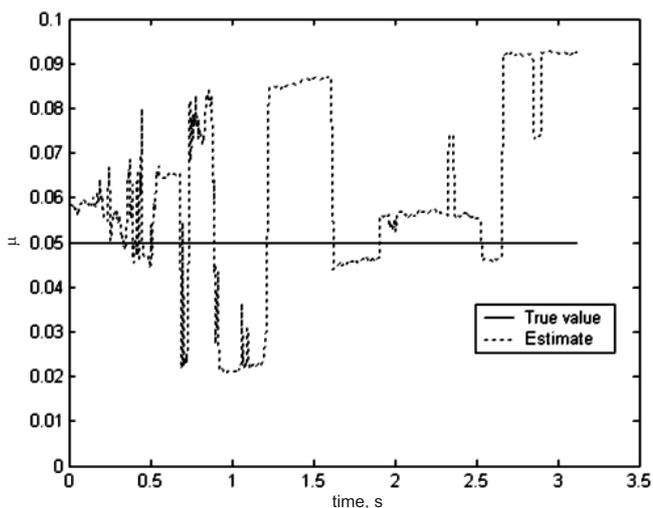


Figure 9. Studies on the Coulomb oscillator; filter 1; estimation of the nonlinearity parameter; results on the mean estimate; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$; $\text{mean}(a_{3k|k}) = 0.0604$.

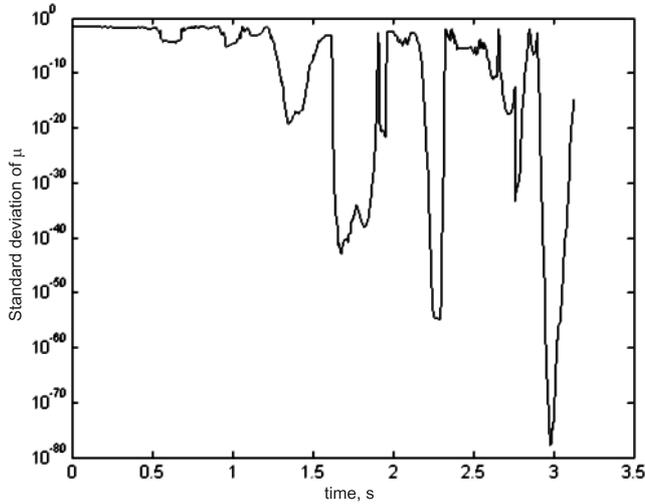


Figure 10. Studies on the Coulomb oscillator; filter 1; estimation of the nonlinearity parameter; results on the standard deviation of the estimate; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$; mean (α_3) = 0.0604.

the form of stochastic differential equations (SDEs). Since nonlinear filtering tools generally work with discrete maps, there is a need to reduce the system and observation SDEs into such maps. Use of stochastic calculus, and, in particular, explicit forms of stochastic Taylor expansions, has presently been made to derive the discrete maps with a high formal order of accuracy. Certain useful variations in these derivations have also been proposed to account for a class of non-smooth (i.e., discontinuous) terms in the drift fields. In particular, the study considers the problem of identifying the parameters in two oscillators, namely, the Duffing and Coulomb oscillators using dynamic state estimation techniques, with the latter having a discontinuous signum function (accounting for dry friction) in the drift field. The parameters

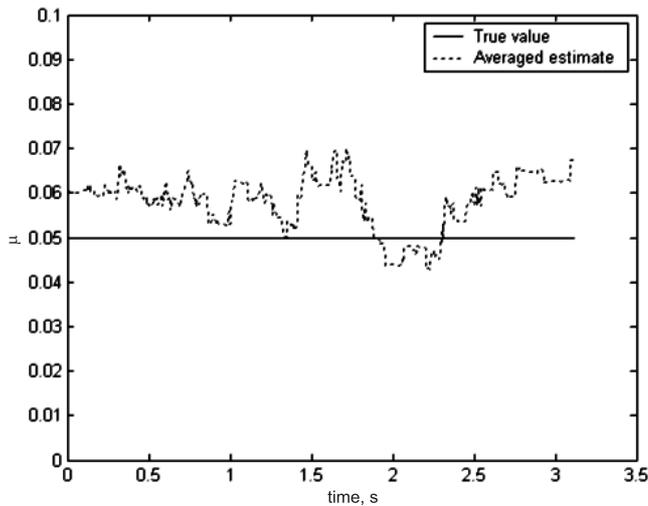


Figure 11. Studies on the Coulomb oscillator; filter 1; estimation of the nonlinearity parameter; results on the mean estimate obtained by ensemble averaging across 10 realizations of the estimate; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$; time average of the mean = 0.0582.

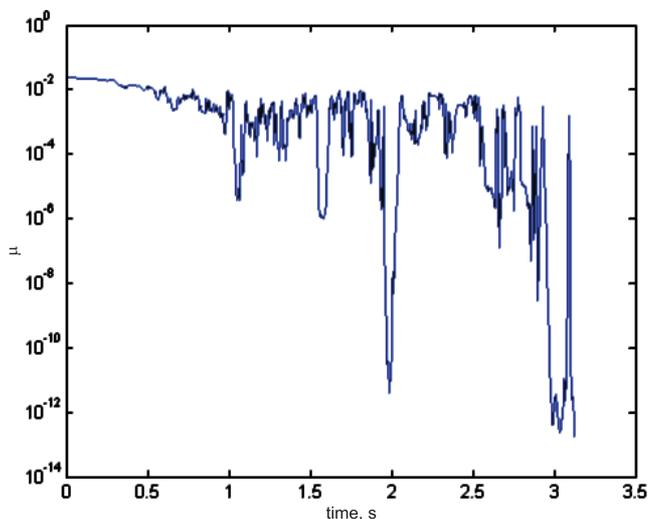


Figure 12. Studies on the Coulomb oscillator; filter 1; estimation of the nonlinearity parameter; results on the estimate of the standard deviation obtained by ensemble averaging across 10 realizations of the estimate; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$.

to be identified are treated as additional dynamical state variables and an augmented system of process equations is formed. The basic objective of the study has been to construct the posterior PDF of this augmented state vector based on all available information. This PDF is represented by a set of random samples with associated weights and they are in turn used to compute the desired estimates. The focus has been on investigating the utility of three Monte Carlo simulation based filters, namely, a density-based Monte Carlo filter, a bootstrap filter and a sequential importance-sampling filter. The parameters to be identified have presently been treated as additional dynamic state variables. An alternative to this approach would be

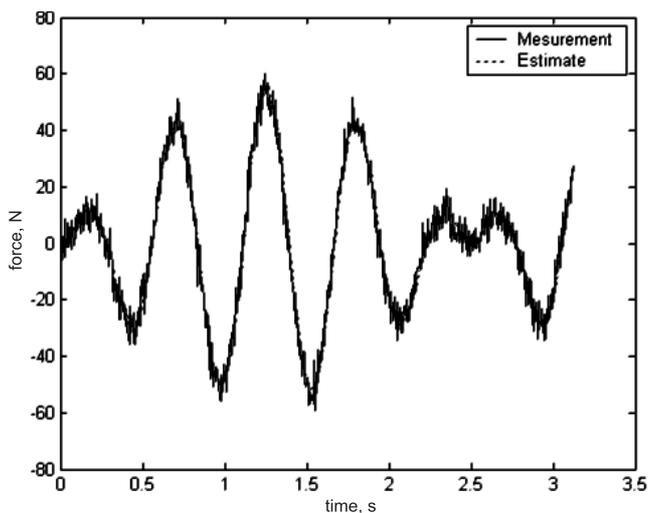


Figure 13. Study on Coulomb oscillator; filter 2; measured and estimated reaction transferred to the support.

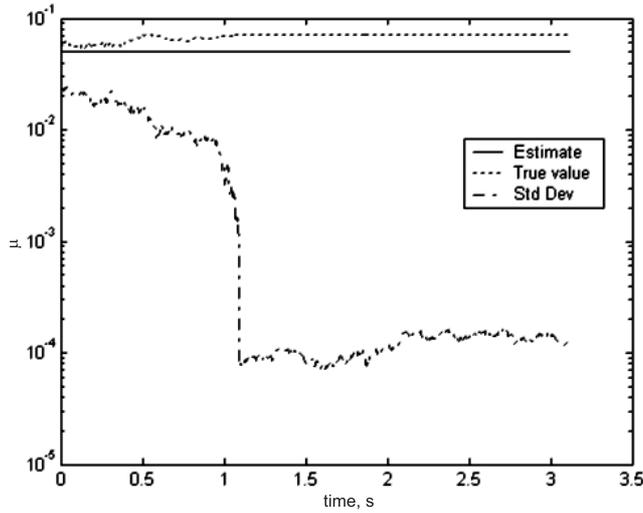


Figure 14. Studies on the Coulomb oscillator; filter 2; estimation of the nonlinearity parameter; results on the estimate of the mean and standard deviation; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$; time average of the mean = 0.0682.

to employ a maximization of the likelihood function to obtain the estimates of system parameters. While only explicit forms of time-integration strategies have been employed in this paper, explorations of implicit techniques could be useful in maintaining the numerical accuracy and stability of the estimation procedure under a larger variation of time-step sizes. The issue of the effect of the ensemble size on the quality and accuracy of the estimation algorithm also needs to be understood better. Such problems constitute interesting elements of future research.

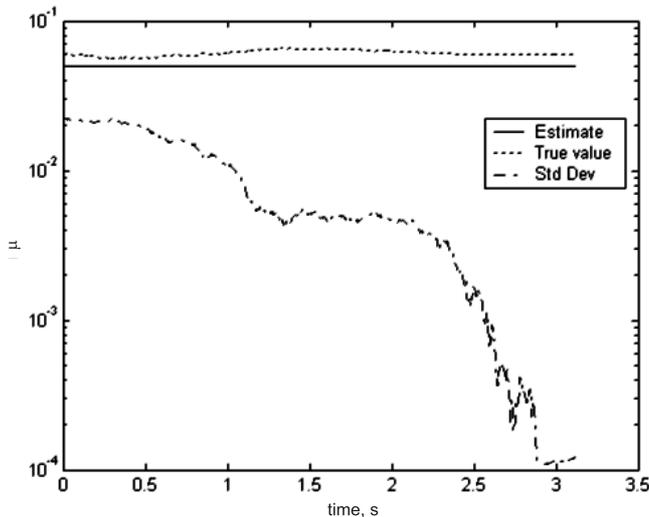


Figure 15. Studies on the Coulomb oscillator; filter 2; estimation of the nonlinearity parameter; results on the estimate of the mean and standard deviation by ensemble averaging across 10 time histories of estimates; $\mu = 0.05$; $\mu_0 \equiv U(0.02, 0.1)$; time averaged mean of estimate of $\mu = 0.063$.

The work reported in this paper has been partly supported by funds from the ISRO–IISc Space Technology Cell.

Appendix A. A brief introduction to Ito's formula

A.1 Wiener processes

A q -dimensional Wiener process $\{B_r(t) | r = 1, \dots, q\}$ is a vector stochastic process with the following properties:

- (1) (Normal increments) $B_r(t) - B_r(s)$ has normal distribution with mean zero and variance $(t - s)$ for every $r \in [1, q]$.
- (2) (Independence of increments) $B_r(t) - B_r(s)$ is independent of the past, i.e., of $B_r(u)$, $u \leq s < t$ for every $r \in [1, q]$.
- (3) (Continuity of paths) There exists a version of $B_r(t, \omega)$ which is a continuous function of time t for almost every realization ω (w.p. 1) and for every $r \in [1, q]$ (ω is an element of the event space Ω).
- (4) Any two Wiener processes $B_r(t)$ and $B_l(t)$ are mutually independent stochastic processes for $r \neq l, r, l \in [1, q]$.

A.2 Ito integral

Let $0 = t_0^N < t_1^N < \dots < t_N^N = T$ be a partition of the interval $[0, T]$ such that $\delta_n = \max_{i \in [0, N-1]} (t_{i+1}^N - t_i^N) \rightarrow 0$ as $N \rightarrow \infty$ and $X(t)$ be a stochastic process. Then the Ito integral $\int_0^T X(t) dW_r(t)$ is defined as:

$$\int_0^T X(t) dB_r(t) = \lim_{\substack{\delta_N \rightarrow 0 \\ N \rightarrow \infty}} \sum_{i=0}^{N-1} X(t_i^N) (B_r(t_{i+1}^N) - B_r(t_i^N)). \quad (\text{A1})$$

For instance, one may readily show that

$$\int_0^T B_r(t) dB_r(t) = \frac{1}{2} (B_r^2(T) - T). \quad (\text{A2})$$

In defining the Ito integral in (A1), attention is restricted to the class of predictable stochastic processes $X(t)$, $t \in [0, T]$ such that:

$$\int_0^T X^2(t) dt < \infty, \text{ almost surely (a.s.).} \quad (\text{A3})$$

Moreover, the process $X(t)$ is assumed to be F_t measurable, where F_t is the increasing family of σ -algebras generated by $\{B_r(s) | s \in [0, t]; r \in [1, q]\}$. It may be shown (Kloeden & Platen 1999) that the Ito integral is an $N(0, \sigma)$ random variable such that:

$$\sigma^2 = \text{Var} \left[\int_0^T X(t) dB_r(t) \right] = \int_0^T E[X^2(t)] dt. \quad (\text{A4})$$

A.3 Quadratic variation

Given a partition $\{t_i^N | i \in [0, N]\}$ of the interval $[0, t]$ as before, the quadratic variation of a process $X(t)$ is defined as:

$$[X, X](t) = \lim_{\substack{N \rightarrow \infty \\ \delta_N \rightarrow 0}} \sum_{i=0}^{N-1} |X(t_{i+1}^N) - X(t_i^N)|^2, \quad (\text{A5})$$

where the above limit is taken in probability. For instance, one may show that

$$[B_r, B_r](t) = t \quad (\text{A6a})$$

and,

$$\left[\int_0^t X(s) dB_r(s), \int_0^t X(s) dB_r(s) \right] (t) = \int_0^t X^2(s) ds \quad (\text{A6b})$$

The above two equations provide an amazing and counter-intuitive result in the theory of stochastic processes. For instance, from an inspection of the left hand side of equation (A6a), we would normally expect it to be a random variable (for a given t). However, it turns out that it is indeed a deterministic number. It is also interesting to note that the quadratic variation of a differentiable and bounded process (as generally encountered in deterministic dynamics of many engineering systems) over any finite time interval is identically zero. For instance, the quadratic variation of each (complex valued) scalar function $X_m(t) = \exp(\sqrt{-1}mt)$ of the Fourier basis set $\{X_m(t), m \in \mathbb{Z}\}$ in the Hilbert space of continuous and bounded functions is zero. This is probably the most conspicuous difference between a stochastic process and a deterministic (non-stochastic) one.

A.4 Stochastic differential equations (SDEs) and Ito's formula

Consider the vector stochastic process $X(t) \in \mathbb{R}^n$ governed by the SDE:

$$dX(t) = a(t, X)dt + \sigma(t, X)dB(t), \quad (\text{A7})$$

where $B(t) \in \mathbb{R}^q$ is a q -dimensional vector Wiener process, $\sigma(t, X) \in \mathbb{R}^{n \times q}$ is the diffusion matrix and $a(t, X) \in \mathbb{R}^n$ is the drift vector. The SDE (A7) is expressed in an incremental (and not in differentiated) form as $B(t)$ is not a differentiable vector stochastic process (with respect to time). Denoting the i th and j th scalar component processes of the vector process $X(t) \in \mathbb{R}^n$ as $X_i(t)$ and $X_j(t)$ respectively, the differential quadratic variation may be shown to be:

$$d[X_i, X_j](t) = C_{ij}dt; i, j \in [1, n], \quad (\text{A8})$$

$$\text{where } \{C\}_{n \times n} = [\sigma]_{n \times q} [\sigma]_{q \times n}^T, \quad (\text{A9})$$

with the superscript T denoting matrix transposition.

Let $f(X_1, X_2, \dots, X_n)$ be a C^2 function. Then $f(X_1, X_2, \dots, X_n)$ is also an Ito process, whose stochastic differential is given by the Ito formula:

$$\begin{aligned} df(X_1(t), X_2(t), \dots, X_n(t)) &= \sum_{i=1}^n \frac{\partial}{\partial X_i} f(X_1(t), X_2(t), \dots, X_n(t)) dX_i(t) \\ &+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial X_i \partial X_j} f(X_1(t), X_2(t), \dots, X_n(t)) d[X_i, X_j](t) \end{aligned} \quad (\text{A10})$$

It is of interest to observe that for a deterministic and differentiable vector process $X(t)$, we have $d[X_i, X_j](t) = 0$ and thus Ito's formula reduces to the well-known relation for the total derivative:

$$\frac{df}{dt} = \sum_i \frac{\partial f}{\partial X_i} \dot{X}_i. \quad (\text{A11})$$

Appendix B. The stochastic Taylor expansion

To start with, it is instructive to derive the deterministic form of Taylor's expansion. Let $X(t) \in R^n$ be a deterministic and sufficiently differentiable process governed by the ODE-s:

$$\frac{dX(t)}{dt} = \dot{X}(t) = a(t, X(t)) \quad (\text{B1})$$

Let $f(t, X)$ be a scalar or a vector function (assumed to be sufficiently smooth with respect to its arguments), so that one may write:

$$\frac{d}{dt} f(t, X(t)) = \frac{\partial f}{\partial t} + \sum_{i=1}^n \frac{\partial f}{\partial X_i} a_i(t, X) \quad (\text{B2})$$

Thus we have:

$$f(t+h, X(t+h)) = f(t, X) + \int_t^{t+h} Lf(s, X(s)) ds, \quad (\text{B3})$$

where the operator L is defined as:

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^n a_i(t, X) \frac{\partial}{\partial X_i}. \quad (\text{B4})$$

Now, letting $f(t, X(t)) = X(t)$, we immediately have from (B3) and (B4):

$$X(t+h) = X(t) + \int_t^{t+h} a(s, X(s)) ds. \quad (\text{B5})$$

Further, using (B3) for $a(s, X(s))$ in (B5), we obtain:

$$\begin{aligned} X(t+h) &= X(t) + \int_t^{t+h} [a(t, X(t)) + \int_t^s La(s_1, X(s_1))ds_1]ds \\ &= X(t) + a(t, X(t))h + \int_t^{t+h} (t+h-s)La(s, X(s))ds \end{aligned} \quad (\text{B6})$$

Applying (B3) once more to $La(s, X(s))$ in (B6), the following expression is readily derived:

$$\begin{aligned} X(t+h) &= X(t) + a(t, X(t))h + La(t, X(t))\frac{h^2}{2} \\ &\quad + \int_t^{t+h} \frac{(t+h-s)^2}{2} L^2a(s, X(s))ds \end{aligned} \quad (\text{B7})$$

Continuing in this way m times (i.e., by iterating m times with the formula (B3) on the integral expansion (B5)), we arrive at the well-known Taylor expansion of $X(t+h)$ in powers of h in the neighbourhood of t :

$$\begin{aligned} X(t+h) &= X(t) + a(t, X(t))h + La(t, X(t))\frac{h^2}{2} + \dots + L^{m-1}a(t, X(t))\frac{h^m}{m!} \\ &\quad + \int_t^{t+h} \frac{(t+h-s)^m}{m!} L^m a(s, X(s))ds \end{aligned} \quad (\text{B8})$$

Given a system of SDEs, the corresponding expansion of the stochastic process $X(t)$ in a neighbourhood of t is referred to as the stochastic Taylor expansion. Alternatively, such an expansion is also called the Ito–Taylor (or Stratonovich–Taylor) expansion depending upon the type interpretation of stochastic integral being employed. In this paper, use has only been made of the Ito–Taylor expansion and thus the derivation of such an expansion is briefly described here. Now consider the solution $X(t) \in R^n$ of the system of SDEs (A7), which may be written in the following alternative form:

$$dX(t) = a(t, X)dt + \sum_{r=1}^q \sigma_r(t, X)dB_r(t), \quad (\text{B9})$$

where the vector function $\sigma_r(t, X) : R \times R^n \rightarrow R^n$ is the r th column of the diffusion matrix $[\sigma]_{n \times q}$. In what follows, the scalar diffusion function $\sigma_r^{(j)}(t, X)$ would denote the $[j, r]$ th element of the matrix $[\sigma]_{n \times q}$. If $f(t, X)$ is a sufficiently smooth scalar or vector function, then by Ito's formula (A10) and the formula for the differential quadratic variation (A8), we immediately get for $t_0 \leq t \leq s$:

$$f(s, X(s)) = f(t, X) + \sum_{r=1}^q \int_t^s \Lambda_r f(s_1, X(s_1))dB_r(s_1) + \int_t^s Lf(s_1, X(s_1))ds_1. \quad (\text{B10})$$

The above equation may therefore be thought of as an integral form of Ito's formula. Similar to the deterministic case, the above equation provides a formula for the stochastic increment

$\Delta f = f(s, X(s)) - f(t, X(t))$ of the function f , where the operators Λ_r and L are given by:

$$\begin{aligned}\Lambda_r &= \sum_{j=1}^n \sigma_r^{(j)} \frac{\partial}{\partial X_j}, \\ L &= \frac{\partial}{\partial t} + \sum_{j=1}^n a_j(t, X) \frac{\partial}{\partial X_j} + \frac{1}{2} \sum_{r=1}^q \sum_i^n \sum_{j=1}^n \sigma_r^{(i)} \sigma_r^{(j)} \frac{\partial^2}{\partial X_i \partial X_j}.\end{aligned}\quad (\text{B11})$$

Now applying (B10) recursively to $\Lambda_r f(s_1, X(s_1))$ and $Lf(s_1, X(s_1))$, we have:

$$\begin{aligned}f(t+h, X(t+h)) &= f(t, X) + \sum_{r=1}^q \Lambda_r f \int_t^{t+h} dB_r(s) + Lf \int_t^{t+h} ds \\ &\quad + \sum_{r=1}^q \int_t^{t+h} \left(\sum_{p=1}^q \int_t^s \Lambda_p \Lambda_r f(s_1, X(s_1)) dB_p(s_1) dB_r(s) \right. \\ &\quad + \sum_{r=1}^q \int_t^{t+h} \left(\int_t^s L \Lambda_r f(s_1, X(s_1)) ds_1 \right) dB_r(s) \quad (\text{B12}) \\ &\quad + \sum_{r=1}^q \int_t^{t+h} \left(\int_t^s \Lambda_r Lf(s_1, X(s_1)) dB_r(s_1) \right) ds \\ &\quad + \int_t^{t+h} \left(\int_t^s L^2 f(s_1, X(s_1)) ds_1 \right) ds.\end{aligned}$$

Just as in deterministic Taylor expansion, if Ito's formula (B10) is applied to expand $\Lambda_p \Lambda_r f(s_1, X(s_1))$, $L \Lambda_r f(s_1, X(s_1))$, and $\Lambda_r Lf(s_1, X(s_1))$, in (B12), then the following expression results:

$$\begin{aligned}f(t+h, X(t+h)) &= f(t, X) + \sum_{r=1}^q \Lambda_r f \int_t^{t+h} dB_r(s) + Lf \int_t^{t+h} ds \\ &\quad + \sum_{r=1}^q \sum_{i=1}^q \Lambda_i \Lambda_r \int_t^{t+h} \int_t^s dB_r(s_1) dB_i(s) \\ &\quad + \sum_{r=1}^q \sum_{i=1}^q \sum_{j=1}^q \Lambda_j \Lambda_i \Lambda_r \int_t^{t+h} \int_t^s \int_t^{s_1} dB_j(s_2) dB_i(s_1) dB_r(s) \quad (\text{B13}) \\ &\quad + \sum_{r=1}^q \Lambda_r Lf \int_t^{t+h} \int_t^s dB_r(s_1) ds + \sum_{r=1}^q L \Lambda_r f \int_t^{t+h} \int_t^s ds_1 dB_r(s) \\ &\quad + L^2 f \int_t^{t+h} \int_t^s ds_1 ds + \rho,\end{aligned}$$

where ρ is the set of remainder terms. When written out in long hand, these terms contain expressions involving multiple integrals of third- and fourth-order compositions of operators

Λ_r and L . Still higher order expansions (as in the higher order Newmark method described in this paper) may be derived by iterating the error terms ρ using (B10), i.e., the integral form of Ito's formula. It is now clear that a distinguishing feature of the Ito–Taylor expansion *vis-à-vis* the deterministic Taylor expansion is that the former has multiple integrals of the forms $I_{ri} = \int_t^{t+h} \int_t^s dB_r(s_1)dB_i(s)$, $I_{rij} = \int_t^{t+h} \int_t^s \int_t^{s_1} dB_r(s_{12})dB_i(s_1)dB_j(s)$, $I_{r0} = \int_t^{t+h} \int_t^s dB_r(s_1)ds$ etc. These integrals involve increments of scalar Wiener components $dW_r(t)$ and hence are referred to as multiple stochastic integrals (MSIs). Note that these integrals are denoted by I followed by a set of integer subscripts indicating the scalar components of the Wiener increments in the same order as they appear within the integrals. Moreover, a zero subscript is taken to indicate integration with respect to time, t .

References

- Brown R G, Hwang P Y C 1992 *Introduction to random signals and applied Kalman filtering* 2nd edn (New York: John Wiley and Sons)
- Chui C K, Chen G 1998 *Kalman filtering with real time applications* (New York: Springer)
- Crisan D, Doucet A 2002 A survey of convergence results on particle filtering methods for practioners. *IEEE Trans. Signal Process.* 50: 736–746
- Doucet A 1998 On sequential simulation-based methods for Bayesian filtering. Technical report CUED/F-INFENG/TR.310 (1998), Department of Engineering, University of Cambridge, UK
- Doucet A, Godsill S, Andrieu C 2000 On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* 10: 197–208
- Ewins D J 2000 *Modal testing: theory, procedures and applications* (Badlock: Research Studies Press)
- Gard T C 1988 *Introduction to stochastic differential equations* (New York: Marcel Dekker)
- Ghanem R, Shinozuka M 1995 Structural system identification I: Theory, *J. Eng. Mech., ASCE* 121: 255–264.
- Gordon N J, Salmond D J, Smith A F M 1993 Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *Inst. Elec. Eng., Proc.* F140: 107–113
- Grewal M S, Andrews A P 2001 *Kalman filtering: Theory and practice using Matlab* (New York: John Wiley and Sons)
- Heylen W, Lammens S, Sas P 1997 *Modal analysis: theory and testing* (Leuven, Heverlee: Katholieke Universiteit)
- Iba Y 2001 Population Monte Carlo algorithms. *Trans. Jap. Soc. Artif. Intell.* 16: 279–286
- Imai H, Yun C B, Maruyama O, Shinozuka M 1989 Fundamentals of system identification in structural dynamics. *Probab. Eng. Mech.* 4: 162–173
- Kailath T 1981 *Lectures on Wiener and Kalman filtering* (New York: Springer Verlag)
- Kalman R E 1960 A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.* D82: 35–45
- Kloeden P E, Platen E 1999 *Numerical solution of stochastic differential equations* (Berlin: Springer)
- Lieven N A L, Ewins D J 2001 Experimental modal analysis. (Theme issue) *Philos. Trans. R. Soc. London* A359: 1–219
- Liu J S, Chen R 1998 Sequential Monte Carlo methods for dynamical systems. *J. Am. Stat. Assoc.* 93: 1032–1044
- Ljung L 1997 *System identification: theory for the user* (Englewood Cliffs, NJ: Prentice-Hall)
- Milstein G N 1995 *Numerical integration of stochastic differential equations* (Dordrecht: Kluwer Academic)
- Peeters B, Roeck G D 2001 Stochastic system identification for operational modal analysis. *ASME J. Dyn. Syst., Measure. Control* 123: 1–9
- Pintelon R, Schoukens J 2001 *System identification: a frequency domain approach* (New York: IEEE Press)

- Pitt M K, Shephard N 1999 Filtering via simulation: auxiliary particle filter. *J. Am. Stat. Assoc.* 94: 59–599
- Ristic B, Arulampallam S, Gordon N 2004 *Beyond the Kalman filter: Particle filters for tracking applications* (Boston: Artech House)
- Shinozuka M, Ghanem R 1995 Structural system identification II: experimental verification. *J. Eng. Mech., ASCE* 121: 265–273
- Smith A F M, Gelfand A E 1992 Bayesian statistics without tears: a sampling-resampling perspective. *Am. Stat.* 46(2): 84–88
- Tanizaki H 1996 *Nonlinear filters: estimation and applications* 2nd edn (Berlin: Springer Verlag)
- Worden K, Manson G 2001 Nonlinearity in experimental modal analysis. *Philos. Trans. R. Soc. (London)* A359: 113–130