

# A sequential importance sampling filter with a new proposal distribution for state and parameter estimation of nonlinear dynamical systems

BY SHUVA J. GHOSH, C. S. MANOHAR\* AND D. ROY

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

The problem of estimating parameters of nonlinear dynamical systems based on incomplete noisy measurements is considered within the framework of Bayesian filtering using Monte Carlo simulations. The measurement noise and unmodelled dynamics are represented through additive and/or multiplicative Gaussian white noise processes. Truncated Ito–Taylor expansions are used to discretize these equations leading to discrete maps containing a set of multiple stochastic integrals. These integrals, in general, constitute a set of non-Gaussian random variables. The system parameters to be determined are declared as additional state variables. The parameter identification problem is solved through a new sequential importance sampling filter. This involves Ito–Taylor expansions of nonlinear terms in the measurement equation and the development of an ideal proposal density function while accounting for the non-Gaussian terms appearing in the governing equations. Numerical illustrations on parameter identification of a few nonlinear oscillators and a geometrically nonlinear Euler–Bernoulli beam reveal a remarkably improved performance of the proposed methods over one of the best known algorithms, i.e. the unscented particle filter.

**Keywords:** dynamic state estimation; particle filter; structural system identification; stochastic Taylor expansion

## 1. Introduction

Dynamic state estimation of dynamical systems is of central importance in the areas of structural vibration control and system parameter identification. The pioneering work of Kalman (1960) provides the exact solution to the problem of state estimation in linear dynamical systems with Gaussian additive noises. The filter is formulated in the time domain, is recursive in nature, implemented in real time, and the filter estimate minimizes the mean square error. The filter is extensively studied and has found widespread applications in the field of signal

\* Author for correspondence (manohar@civil.iisc.ernet.in).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspa.2007.0075> or via <http://www.journals.royalsoc.ac.uk>.

processing. Nevertheless, its inability to treat nonlinearity and/or non-Gaussian noises impairs its applicability to most problems of practical importance. In the context of mechanical systems, such problems are often characterized by their large dimensionality (possibly owing to a finite-dimensional projected semi-discrete form of an infinite-dimensional mathematical model) and diverse forms of nonlinearity (possibly resulting from deformation geometry or material inelasticity). Further complications may arise through the possible emergence of non-Gaussian terms in the temporal discretization of the semi-discrete form, which mostly appears as a system of stochastic differential equations. There are essentially two alternative strategies in the existing literature to treat at least some of these difficulties. The first strategy employs some form of linearization to approximate the given nonlinear state-space model and subsequently apply the Kalman filter to the linearized equations (e.g. Saito & Hoyshiya 1984; Imai *et al.* 1989; Ghanem & Shinozuka 1995; Brown & Hwang 1997; Ghosh *et al.* 2007). The second strategy, on the other hand, aims at numerically treating the posterior non-Gaussian probability density functions (pdfs) of the system states. Interestingly, recursive formulae for the evolution of such posterior pdfs conditioned on available measurements are derivable in an exact form. These are valid, in general, for nonlinear systems and Gaussian or non-Gaussian noises that could be additive or multiplicative in nature (Ristic *et al.* 2004). These formulae contain a large number of multidimensional integrals that are often computationally intractable. Consequently, the formulae remain essentially formal in nature. On the other hand, such recursive relations may form the basis of numerical strategies to tackle the problem approximately. This, indeed, has been the focus of a wide range of numerical techniques styled variously as particle filters, Monte Carlo filters, population Monte Carlo algorithms and sequential Monte Carlo sampling methods. In these methods, the posterior pdf is represented by a set of sample realizations of the states (obtainable using Monte Carlo simulations) along with a set of associated weights, and the algorithm ensures that these weights evolve recursively as the measurements are assimilated sequentially (Gordon *et al.* 1993; Tanizaki 1996; Merwe *et al.* 2000; Doucet *et al.* 2001; Liu 2001; Ristic *et al.* 2004; Cappé *et al.* 2007). A particle filtering technique of specific relevance to the present study is the sequential importance sampling (SIS) filter (Doucet *et al.* 2000). This filter is based on importance sampling from a proposal density. The choice of this density function itself constitutes an important step in the implementation of the method. For the special class of filtering problems involving nonlinear process equations and linear measurement models (with only additive Gaussian noises), it has been shown that the ideal importance sampling pdf (ispdf), which minimizes the variance of weights conditioned on previous states and measurements up to the current time, has a Gaussian form.

Parameter identification of nonlinear structural dynamical systems has attracted extensive research attention. A wide variety of identification strategies have been formulated in time and frequency domains. Worden & Tomlinson (2001) and Kerschen *et al.* (2006) provide extensive reviews of these developments. However, these reviews do not cover the methods based on dynamic state estimation techniques. Indeed, the application of particle filters to problems of relevance in structural dynamics is not yet widely researched. A few studies in the recent years have been done by Ching *et al.* (2006), Manohar & Roy (2006), Namdeo & Manohar

(2007) and Sajeeb *et al.* (2007, submitted). The application of the Bayesian inference approach for model selection in the context of nonlinear dynamical system identification has been considered by Kerschen *et al.* (2003).

We presently consider the problem of state estimation and parameter identification in nonlinear dynamical systems when the measurements are noisy and nonlinear functions of the system states. The unmodelled dynamics in the formulation of process equation is represented as additive and/or multiplicative white noise vector processes (i.e. formal derivatives of Weiner vector processes). The process and measurement equations are discretized into maps via Ito–Taylor expansions. The discretization involves a set of multiple stochastic integrals (MSIs) that are, in general, mutually dependent, zero-mean, non-Gaussian random variables (local martingales). The system parameters to be identified are treated as auxiliary state variables, thereby converting the parameter estimation problem into one of state estimation. Measurements are allowed to be sufficiently smooth nonlinear functions of the system states. Thus, for instance, the reaction transferred to the support in a nonlinear oscillator is a nonlinear function of displacement and velocity processes; similarly, strains measured on a beam or a plate, undergoing large dynamic deformations, would be nonlinear functions of the discretized displacement and velocity fields. The problem at hand is then characterized by the presence of non-Gaussian random variables in the discretized equations, and a nonlinear functional relationship between the measured quantities and the state variables. These features pose difficulties in developing sequential Monte Carlo schemes for the filtering problem, and the present study aims at addressing these difficulties. Specifically, we propose a strategy to select an ideal importance sampling density function for this class of problems. This function is given by a weighted sum of multidimensional Gaussian density functions and is shown to be asymptotically exact as the nonlinearity in the measurement model tends to vanish. Furthermore, subject to nonlinear terms being sufficiently smooth, the formal accuracy of the expansion improves as the number of terms in the truncated Ito–Taylor expansion increases. By way of illustration, we consider a few examples from low-dimensional dynamical systems and the problem of state and parameter estimation of an Euler–Bernoulli beam undergoing large-amplitude (geometrically nonlinear) oscillations, with measurements being made on bending strains at a set of points on the beam. The proposed method is shown to perform significantly better than the unscented particle method (Merwe *et al.* 2000; Ristic *et al.* 2004) in terms of both the numerical accuracy and the computational time.

## 2. Governing equations

We consider dynamical systems governed by a set of stochastic differential equation (SDE) of the generic form

$$d\mathbf{u}(t) = \tilde{\mathbf{a}}(\mathbf{u}(t), \mathbf{u}_d, \boldsymbol{\phi}, t)dt + \tilde{\mathbf{b}}(\mathbf{u}(t), \boldsymbol{\phi}, t)d\mathbf{B}_1(t); \quad \mathbf{u}(0) = \mathbf{u}_0, \quad (2.1)$$

where  $\mathbf{u}(t) \in \mathbf{R}^n$  is the state vector;  $\mathbf{u}_d \in \mathbf{R}^f$  is the deterministic forcing function;  $\boldsymbol{\phi} \in \mathbf{R}^p$  is the unknown vector of system parameters to be identified;  $\tilde{\mathbf{a}} : \mathbf{R}^n \times \mathbf{R}^f \times \mathbf{R}^p \times \mathbf{R} \rightarrow \mathbf{R}^n$  is the drift vector or the state transition function;  $\tilde{\mathbf{b}} : \mathbf{R}^n \times \mathbf{R}^p \times \mathbf{R} \rightarrow \mathbf{R}^n \times \mathbf{R}^r$  is the diffusion coefficient matrix;  $\mathbf{B}_1(t) \in \mathbf{R}^r$  is a

vector of standard Brownian motion processes; and  $\mathbf{u}_0 \in \mathbf{R}^n$  is the vector of initial conditions (possibly random, but assumed to be independent of  $d\mathbf{B}_I(t)$ ).  $\mathbf{B}_I(t)$  is taken to account for any unmodelled dynamics (while being questionable, this is standard in the particle filters literature). To facilitate determining the vector  $\boldsymbol{\phi}$ , we declare the elements of  $\boldsymbol{\phi}$  as additional states evolving according to the equation

$$d\boldsymbol{\phi}(t) = \sigma_\phi d\mathbf{B}_\Pi(t), \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{2.2}$$

where  $\sigma_\phi \in \mathbf{R}^p \times \mathbf{R}^p$  is a diagonal matrix of the noise intensities;  $\mathbf{B}_\Pi(t) \in \mathbf{R}^p$  is a vector of increments of standard Brownian processes independent of  $\mathbf{B}_I(t)$ ; and  $\boldsymbol{\phi}_0 \in \mathbf{R}^p$  is modelled as random variables independent of  $\mathbf{u}_0$ ,  $d\mathbf{B}_I(t)$  and  $d\mathbf{B}_\Pi(t)$ . The time dependence of  $\boldsymbol{\phi}$ , as in equation (2.2), is an artefact to convert the problem of parameter estimation to one of state estimation. Now, we define the augmented state vector  $\mathbf{x}(t) = [\mathbf{u}(t)^\top, \boldsymbol{\phi}(t)^\top]^\top$  and combine equations (2.1) and (2.2) to obtain

$$d\mathbf{x}(t) = \mathbf{a}(\mathbf{x}(t), \mathbf{u}_d, t)dt + b(\mathbf{x}(t), t)d\mathbf{B}(t); \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{2.3}$$

where  $\mathbf{x}(t) \in \mathbf{R}^s$ , with  $s = n + p$  is the augmented state vector;  $\mathbf{a} : \mathbf{R}^s \times \mathbf{R}^f \times \mathbf{R} \rightarrow \mathbf{R}^s$  is the augmented drift function;  $\mathbf{B}(t) = \{\mathbf{B}_I^\top, \mathbf{B}_\Pi^\top\}^\top \in \mathbf{R}^{n_1}$ , with  $n_1 = r + p$ , is the augmented vector of standard Brownian motion processes;  $b : \mathbf{R}^s \times \mathbf{R} \rightarrow \mathbf{R}^s \times \mathbf{R}^{n_1}$  is the augmented diffusion coefficient matrix; and  $\mathbf{x}_0 \in \mathbf{R}^s$  is the augmented initial condition vector (assumed to be independent of  $\mathbf{B}(t)$ ). The implementation of the state estimation algorithm generally requires that the governing state equation (2.3) be discretized in  $t$ . We presently accomplish this by using the Ito–Taylor expansion (Kloeden & Platen 1992). Thus, we consider discretization times  $t_0 = 0 < t_1 < t_2 < \dots < t_j < \dots$  with  $\Delta_j = t_{j+1} - t_j$ ;  $j = 0, 1, 2, \dots$ . In order to stay focused on the basic elements of the method, we presently adopt a uniform step size  $\Delta_j = \Delta$ . A 1.5 local-order strong Ito–Taylor expansion of the multidimensional SDE (2.3) has its  $i$ th component given by

$$\begin{aligned} x_{k+1}^i &= x_k^i + a^i \Delta + \frac{1}{2} L^0 a^i \Delta^2 + \sum_{j=1}^s (b^{i,j} I_j + L^0 b^{i,j} I_{(0,j)} + L^j a^i I_{(j,0)}) \\ &+ \sum_{j_1, j_2=1}^s L^{j_1} b^{k, j_2} I_{(j_1, j_2)} + \sum_{j_1, j_2, j_3=1}^s L^{j_1} L^{j_2} b^{k, j_3} I_{(j_1, j_2, j_3)} \quad k = 0, 1, 2, \dots, \end{aligned} \tag{2.4}$$

where the operators are given by

$$L^r = \sum_{j=1}^n b^{j,r} \frac{\partial}{\partial x^j} \quad \text{and}, \tag{2.5a}$$

$$L^0 = \frac{\partial}{\partial t} + \sum_{j=1}^n a^j \frac{\partial}{\partial x^j} + \frac{1}{2} \sum_{r=1}^s \sum_{i,j=1}^n b^{r,i} b^{r,j} \frac{\partial^2}{\partial x^i \partial x^j}. \tag{2.5b}$$

The quantities  $I_j, I_{(j,0)}, I_{(0,j)}, I_{(j_1, j_2)}$  and  $I_{(j_1, j_2, j_3)}$  appearing in equation (2.4) are known as MSIs. They are given by

$$I_{(j)} = \int_{t_k}^{t_{k+1}} d\mathbf{B}_\tau^j, \quad I_{(0,j)} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{\tau_1} ds d\mathbf{B}_{\tau_1}^j \quad \text{and} \quad I_{(j,0)} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{s_1} d\mathbf{B}_s^j ds_1, \tag{2.6}$$

$$I_{(j_1, j_2)} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{\tau_1} d\mathbf{B}_{\tau_2}^{j_1} d\mathbf{B}_{\tau_1}^{j_2} \quad \text{and} \quad I_{(j_1, j_2, j_3)} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{\tau_1} \int_{t_k}^{\tau_2} d\mathbf{B}_{\tau_3}^{j_1} d\mathbf{B}_{\tau_2}^{j_2} d\mathbf{B}_{\tau_1}^{j_3}. \tag{2.7}$$

For a given  $k$ , the MSIs constitute a set of zero-mean random variables that are mutually dependent and non-Gaussian in general. The lower-order MSIs, *viz.*  $I_j$ ,  $I_{(j,0)}$  and  $I_{(0,j)}$ , may be shown to be jointly Gaussian while the higher-order MSIs,  $I_{(j_1,j_2)}$  and  $I_{(j_1,j_2,j_3)}$ , are non-Gaussian. We refer to Kloeden & Platen (1992, pp. 351–356) for the details of the evaluation of MSIs. Thus, the discretized version of equation (2.3) may be cast in the form

$$\mathbf{x}_{k+1} = \mathbf{a}_k(\mathbf{x}_k, \mathbf{u}_{d_k}) + b_k(\mathbf{x}_k, \mathbf{u}_{d_k})\mathbf{w}_k + c_k(\mathbf{x}_k, \mathbf{u}_{d_k})\boldsymbol{\psi}_k, \quad (2.8)$$

where  $\mathbf{a}_k \in \mathbf{R}^s \times \mathbf{R}^f \rightarrow \mathbf{R}^s$ ;  $b_k \in \mathbf{R}^s \times \mathbf{R}^f \rightarrow \mathbf{R}^s \times \mathbf{R}^{n_2}$ ;  $c_k \in \mathbf{R}^s \times \mathbf{R}^f \rightarrow \mathbf{R}^s \times \mathbf{R}^{n_3}$ ;  $\mathbf{w}_k \in \mathbf{R}^{n_2}$ ; and  $\boldsymbol{\psi}_k \in \mathbf{R}^{n_3}$ . The explicit dependence on time has been suppressed in equation (2.8) for notational conciseness. While  $w_k$  contain lower-order MSIs that are strictly Gaussian (with  $\mathbf{w}_k \in \mathbf{R}^{n_2} \sim N[M_1, C_1], \mathbf{M}_1 = \{0\}$ ),  $\boldsymbol{\psi}_k$  is the set of higher-order non-Gaussian MSIs. It may be shown that elements of  $\boldsymbol{\psi}_k$  are expressible as nonlinear functions of a set of Gaussian random variables, denoted by the vector  $\boldsymbol{\alpha}_k$ . The joint distribution of the random variables  $\mathbf{w}_k$  and  $\boldsymbol{\alpha}_k$  is determinable, and we illustrate the relevant details through specific examples later in this paper.

The identification of  $\boldsymbol{\phi}$  is based on a set of measurements made on functions of the system states. Such measurements are invariably via the sampling of the system response at a set of discrete time instants. We assume that the sampling instants coincide with the discretization times  $0 < t_1 < t_2 < \dots < t_j < \dots$  used in discretizing the process equation (2.3). Accordingly, we postulate a model for the measurement given by

$$\mathbf{y}(t_k) = \mathbf{h}[\mathbf{x}(t_k), t_k] + q[\mathbf{x}(t_k), t_k]\mathbf{m}_k \quad \text{or} \quad \mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + q_k(\mathbf{x}_k)\mathbf{m}_k; \quad k = 1, 2, 3, \dots, \quad (2.9)$$

where  $\mathbf{y}_k \in \mathbf{R}^m$ ;  $\mathbf{h}_k : \mathbf{R}^s \times \mathbf{R} \rightarrow \mathbf{R}^m$ ;  $q_k : \mathbf{R}^s \times \mathbf{R} \in \mathbf{R}^m \times \mathbf{R}^{n_4}$ ; and  $\mathbf{m}_k \in \mathbf{R}^{n_4}$ . Moreover,  $\mathbf{m}_k, k=1, 2, \dots$ , is a sequence of independent vector-valued Gaussian random variables with zero mean and given covariance. These random variables represent the effect of measurement noise and unmodelled mechanics of the problem in relating measurement  $\mathbf{y}_k$  to the system state vector  $\mathbf{x}_k$ . The functions  $\mathbf{h}_k(\mathbf{x}_k)$  and  $q_k(\mathbf{x}_k)$  are, in general, nonlinear.

Equations (2.8) and (2.9), taken together, constitute the governing equations for the estimation of the states. We now introduce the vector sequences  $\mathbf{x}_{0:k} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k\}$  and  $\mathbf{y}_{1:k} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ . The problem is then to determine the multidimensional posterior pdf  $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$  and the filtering density  $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ . Using  $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ , one may readily determine the moments of interest (e.g. the mean and the covariance of  $\mathbf{x}_k$  conditioned on  $\mathbf{y}_{1:k}$ ). A formal solution to this problem is available (see Ristic *et al.* 2004). It consists of the following pair of equations corresponding to the prediction and updating steps:

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1} \quad \text{and} \\ p(\mathbf{x}_k|\mathbf{y}_{1:k}) &= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_k}. \end{aligned} \quad (2.10)$$

The implementation of these equations is, however, not feasible in most cases as they involve a large number of multidimensional integrals with non-Gaussian pdfs in the integrand. Consequently, one resorts to either approximate analytical schemes or numerical solutions based on Monte Carlo simulations.

### 3. State estimation using the sequential importance sampling filter

The SIS filter is a widely used method for dynamic state estimation via Monte Carlo simulations. We first present the salient features of this method and indicate its limitations, which the present study offers to overcome. Consider the problem of evaluation of the expectation

$$I[\Psi_k] = E_p[\Psi_k(\mathbf{x}_{0:k})] = \int \Psi_k(\mathbf{x}_{0:k}) p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) d\mathbf{x}_{0:k}, \quad (3.1)$$

using Monte Carlo simulations. Here,  $E_p[\cdot]$  denotes the expectation with respect to the pdf  $p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})$ . As there is no straightforward way of drawing samples from  $p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})$ , we rewrite equation (3.1) as

$$I[\Psi_k] = \int \Psi_k(\mathbf{x}_{0:k}) \frac{p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})}{\pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})} \pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) d\mathbf{x}_{0:k}. \quad (3.2)$$

Here,  $\pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})$  is a valid pdf to be selected, so that  $p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) > 0 \Rightarrow \pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) > 0$  (i.e.  $\pi$  is absolutely continuous with respect to  $p$ ). While  $p|\pi$  is the Radon–Nikodym derivative,  $\pi$  is the ispdf or the proposal density function. It follows that:

$$I[\Psi_k] = E_\pi \left[ \Psi_k(\mathbf{x}_{0:k}) \frac{p(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})}{\pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})} \right] \quad (3.3)$$

and an estimate of this integral is given by

$$\hat{I}_N[\Psi_k] = \frac{1}{N} \sum_{i=1}^N \Psi_k(\mathbf{x}_{i,0:k}) \frac{p(\mathbf{x}_{i,0:k} | \mathbf{y}_{1:k})}{\pi(\mathbf{x}_{i,0:k} | \mathbf{y}_{1:k})}, \quad (3.4)$$

where  $\mathbf{x}_{i,0:k}$  denotes the  $i$ th sample drawn from the ispdf  $\pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k})$ . Equation (3.4) can be recast as

$$\hat{I}_N[\Psi_k] = \frac{1}{N} \sum_{i=1}^N \Psi_k(\mathbf{x}_{i,0:k}) \tilde{W}(\mathbf{x}_{i,0:k}); \quad \text{with} \quad \tilde{W}(\mathbf{x}_{i,0:k}) = \frac{W(\mathbf{x}_{i,0:k})}{\sum_{i=1}^N W(\mathbf{x}_{i,0:k})} \quad (3.5)$$

$$\text{and} \quad W(\mathbf{x}_{i,0:k}) = \frac{p(\mathbf{y}_{1:k} | \mathbf{x}_{i,0:k}) p(\mathbf{x}_{i,0:k})}{\pi(\mathbf{x}_{i,0:k} | \mathbf{y}_{1:k})}.$$

In addition, by choosing the ispdf in the form

$$\pi(\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) = \pi(\mathbf{x}_{0:k-1} | \mathbf{y}_{1:k-1}) \pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}), \quad (3.6)$$

the weights  $W(\mathbf{x}_{i,0:k})$  may be shown to be recursively computable as

$$W(\mathbf{x}_{i,0:k}) = W(\mathbf{x}_{i,0:k-1}) \frac{p(\mathbf{y}_k | \mathbf{x}_{i,k}) p(\mathbf{x}_{i,k} | \mathbf{x}_{i,k-1})}{\pi(\mathbf{x}_{i,k} | \mathbf{x}_{i,0:k-1}, \mathbf{y}_{1:k})}. \quad (3.7)$$

The implementation of the filter begins by assigning the value of  $1/N$  to all the weights at  $k=0$ . However, given that weights are generated through a multiplicative scheme (3.7) and that they must be bounded, most weights, excepting one, go to zero as  $k$  increases. There is thus a gradual depletion in the number of samples with distinct values, and this compromises the ability

to assimilate further measurements. Although using a very large  $N$  could address this issue in principle, it is impracticable. An alternative strategy involving a resampling step is implemented in which, at every  $k$ , an effective sample size, given by

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^N \{W(\mathbf{x}_{i,k})\}^2}, \quad (3.8)$$

is evaluated. Note that  $N_{\text{eff}}=N$  if all weights are equal and  $N_{\text{eff}}=1$  if all weights excepting one are zero. Thus, when the effective sample at any time step falls below a threshold  $N_{\text{thres}}$ , a step involving resampling is implemented with a view to arresting the depletion of samples. The performance of the SIS filter crucially depends upon the choice of the ispdf  $\pi(\mathbf{x}_k|\mathbf{x}_{i,k-1}, \mathbf{y}_k)$ . Doucet *et al.* (2000) argue that the ispdf given by  $\pi(\mathbf{x}_k|\mathbf{y}_{1:k}, \mathbf{x}_{i,0:k-1}) = p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{i,k-1})$  minimizes the variance of the weights conditioned on  $\mathbf{x}_{0:k-1}$  and  $\mathbf{y}_{1:k}$ . This optimal ispdf is, in general, not determinable. However, it can be shown that, when the process equation has additive Gaussian noise terms and the measured quantities are linear functions of the system states with the measurement noise being additive Gaussian, the optimal ispdf is also Gaussian. Thus, when the process and measurement equations are given, respectively, by

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{v}_k; & \mathbf{v}_k &\sim N[\{0\}_{n_v \times 1}, \Sigma_v] \quad \text{and} \\ \mathbf{y}_k &= H\mathbf{x}_k + \mathbf{w}_k; & \mathbf{w}_k &\sim N[\{0\}_{n_w \times 1}, \Sigma_w], \end{aligned} \quad (3.9)$$

the ideal ispdf is Gaussian, with mean  $\mathbf{M}_k$  and covariance matrix  $\Sigma$  given by

$$\begin{aligned} \mathbf{M}_k &= \Sigma(\Sigma_v^{-1}\mathbf{f}(\mathbf{x}_{k-1}) + H^T\Sigma_w^{-1}\mathbf{y}_k) \quad \text{and} \\ \Sigma^{-1} &= \Sigma_v^{-1}\mathbf{f}(\mathbf{x}_{k-1}) + H^T\Sigma_w^{-1}H. \end{aligned} \quad (3.10)$$

Owing to the non-availability of an analytical form of the ispdf in most cases, several approximate methods have been proposed (Doucet *et al.* 2000). However, each of them has specific disadvantages, e.g. a fixed pdf does not account for either process or measurement dynamics, the prior pdf does not account for the current observation, local linearization of the observation function relies on the accuracy of linearization, extended Kalman filter–particle filter and unscented particle filters (UPFs) are computationally expensive and depend on the accuracy of the extended and unscented Kalman filters, respectively.

In the present study, while the process equation (2.8) is allowed to have multiplicative Gaussian noise terms as well as non-Gaussian random noise terms, the measurement equation (2.9) may have terms that are nonlinear in the system states and multiplicative Gaussian noise terms. Clearly, these equations do not conform to the form given in equation (3.9) and, therefore, the problem of determining the optimal ispdf for this case requires the development of newer strategies.

#### 4. An optimal ispdf based on the Ito–Taylor expansion

In order to implement the SIS filtering with the ideal ispdf for the problem at hand, we need to accomplish two tasks: (i) to determine the functional form of the ispdf and (ii) to develop a strategy to draw samples from this ispdf. In order to determine the ideal ispdf, we note that

$$\pi(\mathbf{x}_k | \mathbf{y}_k, \mathbf{x}_{i,k-1}) = p(\mathbf{x}_k | \mathbf{y}_k, \mathbf{x}_{i,k-1}) = \frac{p(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{i,k-1})}{p(\mathbf{y}_k | \mathbf{x}_{i,k-1})}. \quad (4.1)$$

For the process equation (2.8) and the measurement equation (2.9), the pdfs  $p(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{i,k-1})$  and  $p(\mathbf{y}_k | \mathbf{x}_{i,k-1})$  are non-Gaussian and their functional form is not readily determinable. Consequently, drawing samples from the ideal ispdf would not be straightforward. We propose herein a strategy to overcome these difficulties in an approximate manner. We consider the nonlinear functions  $\mathbf{h}[\mathbf{x}(t), t]$  and  $q[\mathbf{x}(t), t]$  in the measurement equation over  $t \in (t_{k-1}, t_k]$  and propose that these terms be Ito–Taylor expanded around the state  $\mathbf{x}_{k-1} := \mathbf{x}(t_{k-1})$ . Accordingly, we obtain

$$\begin{aligned} \mathbf{h}_k(t, \mathbf{x}(t)) &= \tilde{\mathbf{h}}_k(t, t_{k-1}, \mathbf{x}(t_{k-1})) + \tilde{B}_k(t, t_{k-1}, \mathbf{x}(t_{k-1})) \tilde{\mathbf{m}}_k^1 + \boldsymbol{\rho}_1 \\ &\approx \tilde{\mathbf{h}}_k(t, t_{k-1}, \mathbf{x}_{k-1}) + \tilde{B}_k(t, t_{k-1}, \mathbf{x}_{k-1}) \tilde{\mathbf{m}}_k^1, \end{aligned} \quad (4.2a)$$

$$\begin{aligned} q_k^{(i,j)}(t, \mathbf{x}(t)) &= \tilde{q}_k^{(i,j)}(t, t_{k-1}, \mathbf{x}(t_{k-1})) + \tilde{Q}_k^{(i,j)}(t, t_{k-1}, \mathbf{x}(t_{k-1})) \tilde{\mathbf{m}}_k^{2(i,j)} + \rho_2^{(i,j)} \\ &\approx \tilde{q}_k^{(i,j)}(t, t_{k-1}, \mathbf{x}_{k-1}) + \tilde{Q}_k^{(i,j)}(t, t_{k-1}, \mathbf{x}_{k-1}) \tilde{\mathbf{m}}_k^{2(i,j)} \\ &\quad i = 1, 2, \dots, m \quad \text{and} \quad j = 1, 2, \dots, s_1, \end{aligned} \quad (4.2b)$$

where  $\tilde{\mathbf{m}}_k^1$  and  $\tilde{\mathbf{m}}_k^2$  are a vector and a matrix of MSIs, respectively; and  $\boldsymbol{\rho}_1$  and  $\rho_2$  are the associated vector-valued and matrix-valued remainder terms, respectively. In order to simplify notations, we drop the time indices and, based on the above approximations, the measurement equation can be approximated as

$$\mathbf{y}_k \approx \bar{\mathbf{h}}_k(\mathbf{x}_{k-1}) + \bar{B}_k(\mathbf{x}_{k-1}, \bar{\boldsymbol{\psi}}_k) \bar{\mathbf{m}}_k + \bar{B}_{1_k}(\mathbf{x}_{k-1}) \bar{\boldsymbol{\psi}}_k, \quad (4.3)$$

where  $\bar{\mathbf{h}}_k : \mathbf{R}^s \times \mathbf{R} \rightarrow \mathbf{R}^m$ . The vectors  $\bar{\mathbf{m}}_k \in \mathbf{R}^{n_5}$  and  $\bar{\boldsymbol{\psi}}_k \in \mathbf{R}^{n_6}$  represent the resulting Gaussian and non-Gaussian terms as a result of the approximation,  $\bar{B}_k : \mathbf{R}^s \times \mathbf{R}^{n_6} \times \mathbf{R} \rightarrow \mathbf{R}^m \times \mathbf{R}^{n_5}$  and  $\bar{B}_{1_k} : \mathbf{R}^s \times \mathbf{R} \rightarrow \mathbf{R}^m \times \mathbf{R}^{n_6}$ .

From equations (2.8) and (4.3), it follows that the process equation and the approximated measurement equation have non-Gaussian noise terms, and we introduce the notation  $\boldsymbol{\theta}_k = \left\{ \boldsymbol{\psi}_{k-1}^\top, \bar{\boldsymbol{\psi}}_k^\top \right\}^\top$  to collectively denote all these non-Gaussian random variables. It follows from equation (4.3) that  $p(\mathbf{y}_k | \mathbf{x}_{i,k-1}, \bar{\boldsymbol{\psi}}_k)$  is Gaussian in nature and is given by

$$p(\mathbf{y}_k | \mathbf{x}_{k-1}, \boldsymbol{\theta}_k) = p(\mathbf{y}_k | \mathbf{x}_{k-1}, \bar{\boldsymbol{\psi}}_k) = N(\mathbf{M}_2, C_2), \quad (4.4)$$

where  $\mathbf{M}_2 = \bar{\mathbf{h}}_k(\mathbf{x}_{k-1}) + \bar{B}_{1_k}(\mathbf{x}_{k-1}) \bar{\boldsymbol{\psi}}_k$ ;  $C_2 = \bar{B}_k(\mathbf{x}_{k-1}, \bar{\boldsymbol{\psi}}_k) \bar{R} \bar{B}_k^\top(\mathbf{x}_{k-1}, \bar{\boldsymbol{\psi}}_k)$ ; and  $\bar{R}$  is the covariance matrix of the resultant Gaussian noise vector  $\bar{\mathbf{m}}_k$  in equation (4.3). Considering equations (2.8) and (4.3), it follows that  $\mathbf{x}_k$  and  $\mathbf{y}_k$  are jointly Gaussian given  $\mathbf{x}_{k-1}, \boldsymbol{\theta}_k$ , i.e.

$$p(\mathbf{y}_k, \mathbf{x}_k | \mathbf{x}_{k-1}, \boldsymbol{\theta}_k) = N(\mathbf{M}_3, C_3), \quad (4.5)$$



with  $\mathbf{M}_3 = \{\mathbf{a}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}}) + c_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}})\boldsymbol{\psi}_{k-1}, \bar{\mathbf{h}}_k(\mathbf{x}_{k-1}) + \bar{\mathbf{B}}_{1_k}(\mathbf{x}_{k-1})\bar{\boldsymbol{\psi}}_k\}^T$ . If  $\boldsymbol{\Theta}$  denotes the vector of Gaussian random variables consisting of the Gaussian vectors  $\mathbf{w}_{k-1}$  and  $\bar{\mathbf{m}}_k$  and  $\Xi(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k)$  denotes the corresponding coefficient matrix to be obtained from the joint distribution of  $\mathbf{x}_k$  and  $\mathbf{y}_k$  given  $\mathbf{x}_{k-1}, \boldsymbol{\theta}_k$ , then

$$C_3 = \Xi(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k)E[\{\boldsymbol{\Theta}\}\{\boldsymbol{\Theta}\}^T]\Xi(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k)^T.$$

From the discrete map (2.8), it follows that

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \boldsymbol{\theta}_k) = p(\mathbf{x}_k|\mathbf{x}_{k-1}, \boldsymbol{\psi}_{k-1}) = N(\mathbf{M}_4, C_4), \tag{4.6}$$

where  $\mathbf{M}_4 = \mathbf{a}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}}) + c_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}})\boldsymbol{\psi}_{k-1}$ ;  $C_4 = b_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}})Qb_{k-1}^T(\mathbf{x}_{k-1}, \mathbf{u}_{d_{k-1}})$ ; and  $Q$  is the covariance matrix of  $\mathbf{w}_{k-1}$ .  $Q$  has the same functional form as  $C_1$ . Based on the theory of vector Gaussian random variables, it may be shown that

$$p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{k-1}, \boldsymbol{\theta}_k) = N(\mathbf{M}_I, C_1), \tag{4.7}$$

with  $\mathbf{M}_I = \mathbf{M}_4 + \tilde{C}_3 C_2^{-1}(\mathbf{y}_k - \mathbf{M}_2)$  and  $C_1 = C_4 - \tilde{C}_3 C_2^{-1} \tilde{C}_3^T$ . Here,  $\tilde{C}_3$  denotes the cross terms of the covariance matrix  $C_3$ .

As already noted, elements of the vector  $\boldsymbol{\theta}_k$  are non-Gaussian random variables that are described by nonlinear functions of Gaussian random variables. This means that, even though it is difficult to determine the pdf  $p(\boldsymbol{\theta}_k)$ , drawing samples from this pdf is a simpler exercise. We use this in obtaining an approximation to the ispdf. Note that

$$p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{k-1}, \boldsymbol{\theta}_k)p(\boldsymbol{\theta}_k)d\boldsymbol{\theta}_k. \tag{4.8}$$

Furthermore, we adopt a weighted Monte Carlo representation of the pdf  $p(\boldsymbol{\theta}_k) \approx \sum_{i=1}^{N_1} p_i \delta(\boldsymbol{\theta}_k - \boldsymbol{\theta}_k^i)$ , where  $\boldsymbol{\theta}_k^i$  are the sampled values and  $p_i$  are the associated weights. It then follows from equation (4.8) that

$$p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{k-1}) = \sum_{i=1}^{N_1} p_i p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{x}_{k-1}, \boldsymbol{\theta}_k^i) = \sum_{i=1}^{N_1} p_i N[\mathbf{M}_I(\boldsymbol{\theta}_k^i), C_1(\boldsymbol{\theta}_k^i)]. \tag{4.9}$$

This is the required optimal ispdf. This function, thus, turns out to be a weighted mixture of Gaussian pdfs. The strategy for drawing samples from such pdfs is well established (Liu 2001).

We also need to determine the terms  $p(\mathbf{y}_k|\mathbf{x}_{i,k})$  and  $p(\mathbf{x}_{i,k}|\mathbf{x}_{i,k-1})$  appearing in the equation for the weights (equation (3.7)). The determination of  $p(\mathbf{y}_k|\mathbf{x}_{i,k})$  is straightforward since this density function may be shown to be given by  $p(\mathbf{y}_k|\mathbf{x}_{i,k}) \sim N[\mathbf{h}_k(\mathbf{x}_{i,k}), C_5(\mathbf{x}_{i,k})]$ , with  $C_5(\mathbf{x}_{i,k}) = q_k(\mathbf{x}_{i,k})\bar{R}q_k(\mathbf{x}_{i,k})^T$  (where  $\bar{R}$  is the covariance matrix of the measurement noise  $\mathbf{m}_k$  in equation (2.9)). On the other hand, owing to the presence of non-Gaussian terms in equation (2.8), determining  $p(\mathbf{x}_{i,k}|\mathbf{x}_{i,k-1})$  is not so simple. To circumvent this difficulty, we follow the same strategy as has been used in deriving equation (4.9). Thus, we write

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}, \boldsymbol{\psi}_{k-1})p(\boldsymbol{\psi}_{k-1})d\boldsymbol{\psi}_{k-1} = \sum_{i=1}^{N_1} q_i p(\mathbf{x}_k|\mathbf{x}_{k-1}, \boldsymbol{\psi}_{k-1}^i), \tag{4.10}$$

with  $p(\boldsymbol{\psi}_k) \approx \sum_{i=1}^{N_1} q_i \delta(\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^i)$  being the weighted Monte Carlo representation for  $p(\boldsymbol{\psi}_k)$ . The above pdf is again a weighted mixture of Gaussian densities, and it can easily be computed in order to derive the weights in a recursive manner via equation (3.7). A pseudo code of the proposed algorithm is given in appendix B in the electronic supplementary material.

### Remarks

- (i) In the case where the process and measurement equations conform to the format given in equation (3.9), the optimal non-Gaussian ispdf (equation (4.9)) may be shown to reduce to the ideal Gaussian ispdf valid for the system in equation (3.9). Details are available in appendix D in the electronic supplementary material.
- (ii) Non-Gaussian noise terms in the present study essentially arise due to the application of Ito–Taylor expansion while discretizing the process equation or in the treatment of the nonlinear terms appearing in the measurement equation. The extent to which these terms appear in the formulation depends upon the nature of the process and measurement equations and, of course, the number of terms retained in the truncated Ito–Taylor expansion.
- (iii) In the cases where the discretization of the process SDE or the treatment of the measurement equation does not lead to non-Gaussian MSIs, it is verified that the optimal ispdf in equation (4.9) reduces to a Gaussian pdf. In that case, drawing samples and calculating the corresponding weights become simpler. For further details, refer to appendix C in the electronic supplementary material.
- (iv) It is well known that the Ito–Taylor expansion is a convergent series for sufficiently smooth drift and diffusion functions. Hence, for general nonlinear cases, the formal accuracy of the estimated states by the proposed method increases as the number of terms retained in the expansion increases (or as the time-step size decreases for a given number of terms in the truncated expansions).
- (v) In the context of the treatment of the non-Gaussian MSIs, it may be useful to recall the representation of continuous local martingales, vanishing at  $t=0$ , as time-changed Brownian motions (Revuz & Yor 1991). In particular, considering any such non-Gaussian MSI  $I_\alpha(t)$  ( $\alpha$  being an appropriate multi-index), there is a time-changed Brownian representation for  $I_\alpha(t)$  via the identity  $\mathbf{B}_t = I_\alpha(\tau)$ , where  $\tau = \inf\{s : \langle I_\alpha, I_\alpha \rangle_s \geq t\}$  ( $\langle \cdot, \cdot \rangle$  denotes the quadratic variation operator). The use of such representation should then enable one to use Gaussian characterizations of the MSIs, thereby simplifying the filter algorithm to a great extent. This point of view will however be explored elsewhere.

## 5. Numerical illustrations

We illustrate the formulations presented through three examples. First, we consider a nonlinear dynamical system governed by a scalar first-order SDE with parametric and external excitations. Second, we consider the problem of state

and parameter estimation in an externally driven single degree-of-freedom Duffing oscillator, wherein measurements are made on the reaction transferred to the support. Finally, we consider state and parameter estimation in an Euler–Bernoulli beam undergoing geometrically nonlinear oscillations under external harmonic excitations and measurements are made on bending strains at a set of points on the beam surface. In all the examples, measurement time histories are synthetically generated using numerical simulations. Since the ‘true’ states are known in such cases, a metric of performance of the filter can be formulated. The implementation of the filter requires the sample size to be chosen and the estimates obtained are clearly subject to sampling fluctuations. To gain insights into the performance of the filter for a given sample size, we repeat the filtering process  $N_f$  times. We employ a measure of root mean square error (r.m.s.e.) given by

$$\chi_{z_j}(i) = \frac{1}{M} \sqrt{\sum_{k=1}^M [x_k^j(i) - \hat{x}_k^j(i)]^2}; \quad i = 1, 2, \dots, N_f, \quad (5.1)$$

where  $x_k^j(i)$  is the true state of  $j$ th state in the  $i$ th performance of the experiment at  $t = t_k$  and  $\hat{x}_k^j(i)$  is the corresponding estimate obtained by filtering. The three examples taken together bring out the various features of the formulation, including the occurrence of non-Gaussian MSIs in process equations and the treatment of nonlinear measurement models. Limited comparisons with the results via the UPF are also made to highlight the conceptual and computational superiority of the proposed method.

(a) *Example 5.1: a nonlinear first-order system*

We consider the process equation given by a first-order scalar SDE,

$$dx(t) = (-\alpha x^3(t) + \tilde{q} \cos(\lambda t))dt - Kx(t)(dt + \varepsilon_1 x(t)dB_1(t)) + \sigma dB_2(t). \quad (5.2)$$

The measurement model is given by

$$y_k = x_k(1 + \varepsilon_2 \eta_{1k}) + \sigma_m \eta_{2k}, \quad (5.3)$$

with  $\eta_{1k}, \eta_{2k} \sim N(0, 1)$  being mutually independent. The Ito–Taylor expansion for the process equation with an accuracy of order 1 (Milstein 1995) leads to the discrete map

$$\begin{aligned} x_k &= x_{k-1} + [-Kx_{k-1} - \alpha(x_{k-1})^3 + \tilde{q} \cos(\lambda t_{k-1})]\Delta - K\varepsilon_1 x_{k-1} I_1 + \sigma I_2 \\ &+ (K\varepsilon_1)^2 x_{k-1} I_{11} - \sigma K\varepsilon_1 I_{21}. \end{aligned} \quad (5.4)$$

From equations (5.3) and (5.4), we get

$$\begin{aligned} y_k &= (x_{k-1} + (-Kx_{k-1} - \alpha(x_{k-1})^3 + \tilde{q} \cos(\lambda t_{k-1}))\Delta - K\varepsilon_1 x_{k-1} I_1 + \sigma I_2 \\ &+ (K\varepsilon_1)^2 x_{k-1} I_{11} - \sigma K\varepsilon_1 I_{21})(1 + \varepsilon_2 \eta_{1k}) + \sigma_m \eta_{2k}. \end{aligned} \quad (5.5)$$

The vector of non-Gaussian random variables is given by

$$\begin{aligned} \boldsymbol{\theta}_k &= \{ \theta_{1k} \quad \theta_{2k} \}^T \\ &= \left\{ (K\varepsilon_1)^2 x_{k-1} I_{11} - \sigma K \varepsilon_1 I_{21} \quad \left( (K\varepsilon_1)^2 x_{k-1} I_{11} - \sigma K \varepsilon_1 I_{21} - K \varepsilon_1 x_{k-1} I_1 + \sigma I_2 \right) \right\}^T \end{aligned} \quad (5.6)$$

In order to treat the MSIs appearing in equation (5.6), we introduce a set of mutually independent Gaussian random variables with zero mean and unit standard deviation as

$$\{ \xi_j, \zeta_{j,1}, \dots, \zeta_{j,p}, \eta_{j,1}, \dots, \eta_{j,p}, \mu_{j,p}, \phi_{j,p} \}_{j=1}^2.$$

Following Kloeden & Platen (1992), we obtain

$$I_1 = \sqrt{\Delta} \xi_1, \quad (5.7a)$$

$$I_2 = \sqrt{\Delta} \xi_2, \quad (5.7b)$$

$$I_{11} = \frac{1}{2} [I_1^2 - \Delta] \quad \text{and}, \quad (5.7c)$$

$$I_{2,1}^{\bar{p}} = \frac{1}{2} \Delta \xi_1 \xi_2 - \frac{1}{2} \sqrt{\Delta} [\xi_2 a_{1,0} - \xi_1 a_{2,0}] + A_{2,1}^{\bar{p}} \Delta, \quad (5.7d)$$

where  $\bar{p}$  denotes the number of terms retained in the series expansion of  $A_{2,1}$ , given by

$$A_{2,1}^{\bar{p}} = \frac{1}{2\pi} \sum_{r=1}^{\bar{p}} [\zeta_{2,r} \eta_{1,r} - \eta_{2,r} \zeta_{1,r}]; \quad a_{1,0} = -\frac{\sqrt{2\Delta}}{\pi} \sum_{r=1}^{\bar{p}} \frac{1}{r} \zeta_{1,r} - 2\sqrt{\Delta \rho_{\bar{p}}} \mu_{1,\bar{p}};$$

$$\rho_{\bar{p}} = \frac{1}{12} - \frac{1}{2\pi^2} \sum_{r=1}^{\bar{p}} \frac{1}{r^2}; \quad a_{2,0} = -\frac{\sqrt{2\Delta}}{\pi} \sum_{r=1}^{\bar{p}} \frac{1}{r} \zeta_{2,r} - 2\sqrt{\Delta \rho_{\bar{p}}} \mu_{2,\bar{p}}.$$

Process and the measurement equations are written in terms of  $\boldsymbol{\theta}$  as

$$x_k = b^{k-1} + \theta_{1k} - K \varepsilon_1 x_{k-1} I_1 + \sigma I_2, \quad (5.8)$$

$$y_k = b^{k-1} + \theta_{1k} + \varepsilon_2 \{ b^{k-1} + \theta_{2k} \} \eta_{1k} - K \varepsilon_1 x_{k-1} I_1 + \sigma I_2 + \sigma_m \eta_{2k}, \quad (5.9)$$

with  $b^{k-1} = x_{k-1} + (-Kx_{k-1} - \alpha(x_{k-1})^3 + \tilde{q} \cos(\lambda t_{k-1})) \Delta$ . Equations (5.8) and (5.9) are in the format as in equations (2.8) and (4.3), and, consequently, the formulation outlined in §4 becomes applicable to solve the state estimation problem. The measurement data are generated with  $K=3 \text{ N m}^{-1}$ ,  $\alpha=9 \text{ N m}^{-3}$ ,  $\tilde{q}=15 \text{ N}$  and  $\lambda=2\pi \text{ rad s}^{-1}$ . The process noise intensity is fixed at 1% of the r.m.s. value of the force,  $\sigma=0.11 \text{ N}$ . The simulation is carried out for 3 s with a step size of  $\Delta=0.01 \text{ s}$ . The additive measurement noise level is taken to be 5% of the noise-free displacement amplitude ( $\sigma_m=0.056 \text{ m}$ ). Multiplicative noise intensities in the process and observation equations are taken as  $\varepsilon_1=0.1$  and  $\varepsilon_2=0.4$ , respectively. The sample size is fixed at 50. Number of samples from  $\boldsymbol{\theta}_k$  is taken to be 10. The threshold sample size below which resampling is needed is fixed at two-thirds of the original sample size.  $\bar{p}=3$  is used while evaluating MSIs using equation (5.7a)–(5.7d).

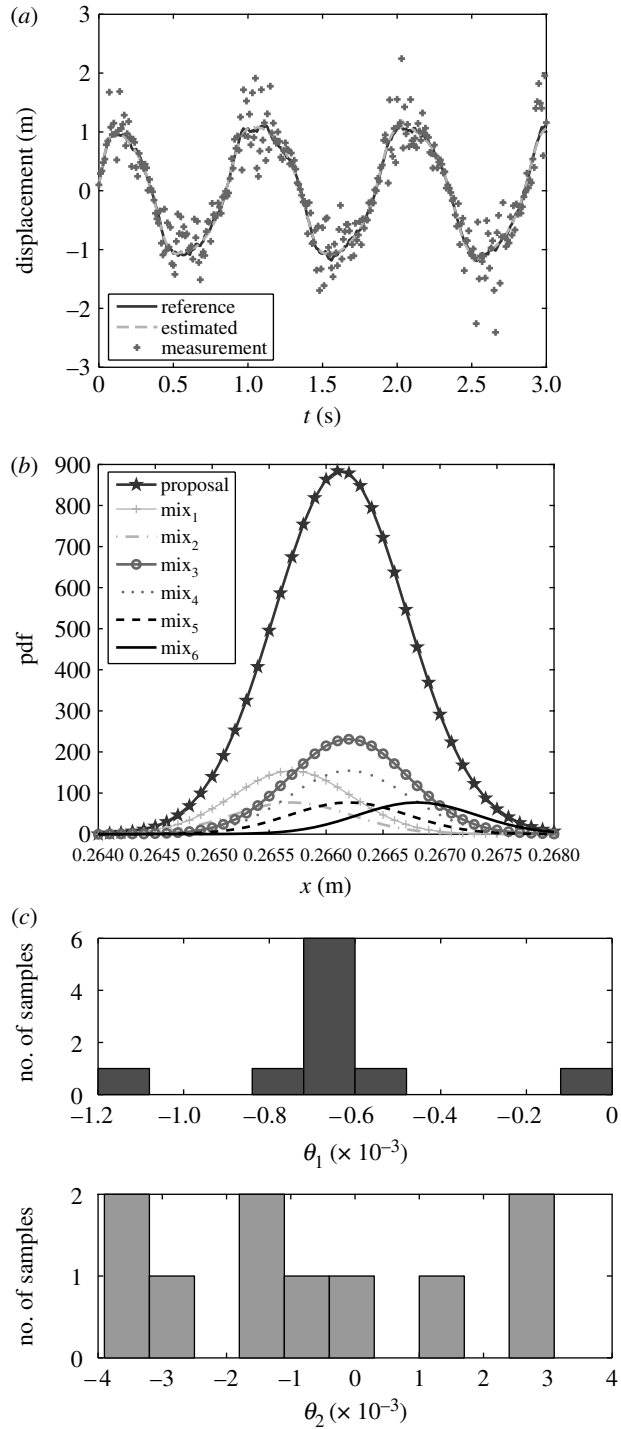


Figure 1. Results for example 5.1: (a) state estimation; (b) proposal density and weighted mixture components,  $t_k=0.3$  s; (c) marginal histograms of  $\theta_1$  and  $\theta_2$ ,  $t_k=0.3$  s.

Figure 1a compares the conditional expected value of  $x_k$  with the reference trajectory along with the noisy observation  $y_k$ . The estimated trajectory appears to closely follow the reference trajectory. The optimal ispdf in this case is non-Gaussian and is obtained as weighted-sum Gaussian pdfs. Figure 1b provides the details of these pdfs. The corresponding plots of histograms of the non-Gaussian random variables  $\theta_1$  and  $\theta_2$  are shown in figure 1c.

(b) *Example 5.2: a hardening Duffing oscillator*

We consider a Duffing oscillator driven by harmonic excitations. The process equation in this case is

$$\begin{aligned} dx^1 &= x^2 dt \quad \text{and} \\ dx^2 &= \frac{1}{m} \left[ -cx^2 - Kx^1 + \alpha(x^1)^3 + \sum_{i=1}^{n_f} p_i \cos \lambda_i t \right] dt + \sigma dB_1(t). \end{aligned} \quad (5.10)$$

Here,  $(c, K, \alpha)$  are the system parameters to be identified. We augment the above equations with additional state equations

$$dx^3 = \sigma_2 dB_2(t); \quad dx^4 = \sigma_3 dB_3(t); \quad dx^5 = \sigma_4 dB_4(t).$$

We have  $(c, k, \alpha) = (x_3, x_4, x_5)$ . The discretized equations based on a 1.5 local-order Ito-Taylor expansion are obtained as

$$\begin{aligned} x_k^1 &= x_{k-1}^1 + x_{k-1}^2 \Delta + \frac{1}{2} \left( -x_{k-1}^5 x_{k-1}^1 - x_{k-1}^4 x_{k-1}^2 - x_{k-1}^3 (x_{k-1}^1)^3 \right. \\ &\quad \left. + \sum_{i=1}^{n_f} p_i \cos(\lambda_i t_{k-1}) \right) \Delta^2 + \sigma_1 I_{1,0}, \\ x_k^2 &= x_{k-1}^2 + \left( -x_{k-1}^5 x_{k-1}^1 - x_{k-1}^4 x_{k-1}^2 - x_{k-1}^3 (x_{k-1}^1)^3 + \sum_{i=1}^{n_f} p_i \cos(\lambda_i t_{k-1}) \right) \Delta \\ &\quad + \frac{1}{2} \left( -\sum_{i=1}^{n_f} p_i \lambda_i \sin(\lambda_i t_{k-1}) + x_{k-1}^2 (-x_{k-1}^5 - 3x_{k-1}^3 (x_{k-1}^1)^2) \right. \\ &\quad \left. - \left( -x_{k-1}^5 x_{k-1}^1 - x_{k-1}^4 x_{k-1}^2 - x_{k-1}^3 (x_{k-1}^1)^3 + \sum_{i=1}^{n_f} p_i \cos(\lambda_i t_{k-1}) \right) \right) \Delta^2 \\ &\quad + \sigma_1 I_1 - \sigma_1 x_{k-1}^4 I_{1,0} - \sigma_2 (x_{k-1}^1)^3 I_{2,0} - \sigma_3 x_{k-1}^2 I_{3,0} - \sigma_4 x_{k-1}^1 I_{4,0} \\ x_k^3 &= x_{k-1}^3 + \sigma_2 I_2 \\ x_k^4 &= x_{k-1}^4 + \sigma_3 I_3 \\ x_k^5 &= x_{k-1}^5 + \sigma_4 I_4. \end{aligned} \quad (5.11)$$

The measurements are presently on the reaction transferred to the support and, accordingly, the measurement equation is obtainable as

$$y_k = x_k^4 x_k^2 + x_k^5 x_k^1 + x_k^3 (x_k^1)^3 + \sigma_m m_k, \quad m_k = N(0, 1). \quad (5.12)$$

The nonlinear terms appearing in equation (5.12) are expanded using Ito-Taylor's expansion and, retaining terms up to 1.5 local order, we get

$$\begin{aligned} x_k^5 x_k^1 &= x_{k-1}^5 x_{k-1}^1 + x_{k-1}^5 x_{k-1}^2 \Delta + a_1 x_{k-1}^5 \frac{\Delta^2}{2} + \sigma_4 x_{k-1}^2 I_{4,0} + \sigma_1 x_{k-1}^5 I_{1,0} \\ &+ \sigma_4 x_{k-1}^1 I_4 + \sigma_4 x_{k-1}^2 I_{04} \end{aligned} \quad (5.13a)$$

$$\begin{aligned} x_k^4 x_k^2 &= x_{k-1}^4 x_{k-1}^2 - \{x_{k-1}^4 (x_{k-1}^1)^3 + x_{k-1}^4 x_{k-1}^5 x_{k-1}^1 + (x_{k-1}^4)^2 x_{k-1}^2\} \Delta \\ &- \{3(x_{k-1}^1)^2 x_{k-1}^2 x_{k-1}^3 x_{k-1}^4 + x_{k-1}^4 x_{k-1}^5 x_{k-1}^2 + (x_{k-1}^4)^2 a_1 + \sigma_3^2 x_{k-1}^2\} \frac{\Delta^2}{2} \\ &+ \sigma_3 x_{k-1}^2 I_3 + \sigma_3 a_1 I_{03} + \sigma_3 \sigma_1 I_{13} + \sigma_1 x_{k-1}^4 I_1 + \sigma_3 \sigma_1 I_{31} \\ &+ \sigma_2 x_{k-1}^4 (x_{k-1}^1)^3 I_{20} + \sigma_3 x_{k-1}^3 (x_{k-1}^1)^3 I_{30} + \sigma_4 x_{k-1}^4 x_{k-1}^1 I_{40} \\ &+ \sigma_3 x_{k-1}^5 x_{k-1}^1 I_{30} + 2\sigma_3 x_{k-1}^4 x_{k-1}^2 I_{30} + \sigma_1 (x_{k-1}^4)^2 I_{10} \end{aligned} \quad (5.13b)$$

$$\begin{aligned} x_k^3 (x_k^1)^3 &= x_{k-1}^3 (x_{k-1}^1)^3 + \{3x_{k-1}^3 (x_{k-1}^1)^2 x_{k-1}^2\} \Delta \\ &+ \{6x_{k-1}^3 (x_{k-1}^2)^2 x_{k-1}^1 + 3x_{k-1}^3 (x_{k-1}^1)^2 a_1\} \frac{\Delta^2}{2} + \sigma_2 (x_{k-1}^1)^3 I_2 \\ &+ 3\sigma_2 x_{k-1}^2 (x_{k-1}^1)^2 I_{20} + 3\sigma_1 x_{k-1}^3 (x_{k-1}^1)^2 I_{10} + 3\sigma_2 x_{k-1}^2 (x_{k-1}^1)^2 I_{02} \end{aligned} \quad (5.13c)$$

where  $a_1 = -x_{k-1}^5 x_{k-1}^1 - x_{k-1}^4 x_{k-1}^2 - x_{k-1}^3 (x_{k-1}^1)^3 + \sum_{i=1}^{n_f} p_i \cos(\lambda_i t_{k-1})$ . Note that the approximation of the measurement equation presently results in two non-Gaussian MSIs ( $I_{31}$  and  $I_{13}$ ) in the second term and the filtering problem is handled within the framework outlined in §4. In the numerical work to synthetically generate the measurements, we take  $m=1$  kg,  $K=17$  N m<sup>-1</sup>,  $\alpha=120$  N m<sup>-3</sup>,  $c=2.47$  N s m<sup>-1</sup>,  $n_f=3$ ,  $p_i=15$  N ( $i=1, 2, 3$ ),  $\lambda_1=\pi$  rad s<sup>-1</sup>,  $\lambda_2=2\pi$  rad s<sup>-1</sup> and  $\lambda_3=4\pi$  rad s<sup>-1</sup>. The process noise intensity is fixed at 1% of the r.m.s. value of the force,  $\sigma=0.11$  N. The simulation is carried out for 5 s at an interval of 0.01 s. The measurement noise level is taken to be 5% of the r.m.s. value of the true (noise-free) reaction ( $\sigma_m=4.21$  N). The filtering problem is solved with 2000 samples and  $N_{\text{thresh}}=1330$ . The system parameters ( $c, K, \alpha$ ) at  $k=0$  are taken as distributed uniformly:  $\alpha_{\text{ini}} \sim U(80, 150)$ ;  $K_{\text{ini}} \sim U(2, 20)$ ;  $c_{\text{ini}} \sim U(1, 5)$ . **Figure 2a** compares the estimate of the reaction transferred to the support along with the corresponding reference trajectory. Also shown in **figure 2a** is the contribution from only the linear terms to the reference trajectory. The influence of nonlinearity on the reaction is amply evident. The conditionally expected values of the parameters ( $c, K, \alpha$ ) are shown, respectively, in **figure 2b–d**. The estimates are observed to approach their reference values as measurements are increasingly assimilated during the filtering. The evolution of the pdf of one of the system parameters, *viz.*  $\alpha$ , is illustrated in **figure 2e**.

The state estimation problems being studied in this work may also be tackled using suboptimal filtering strategies currently available in the existing literature. Among these, the UPF is recognized to be one of the successful filtering methods.

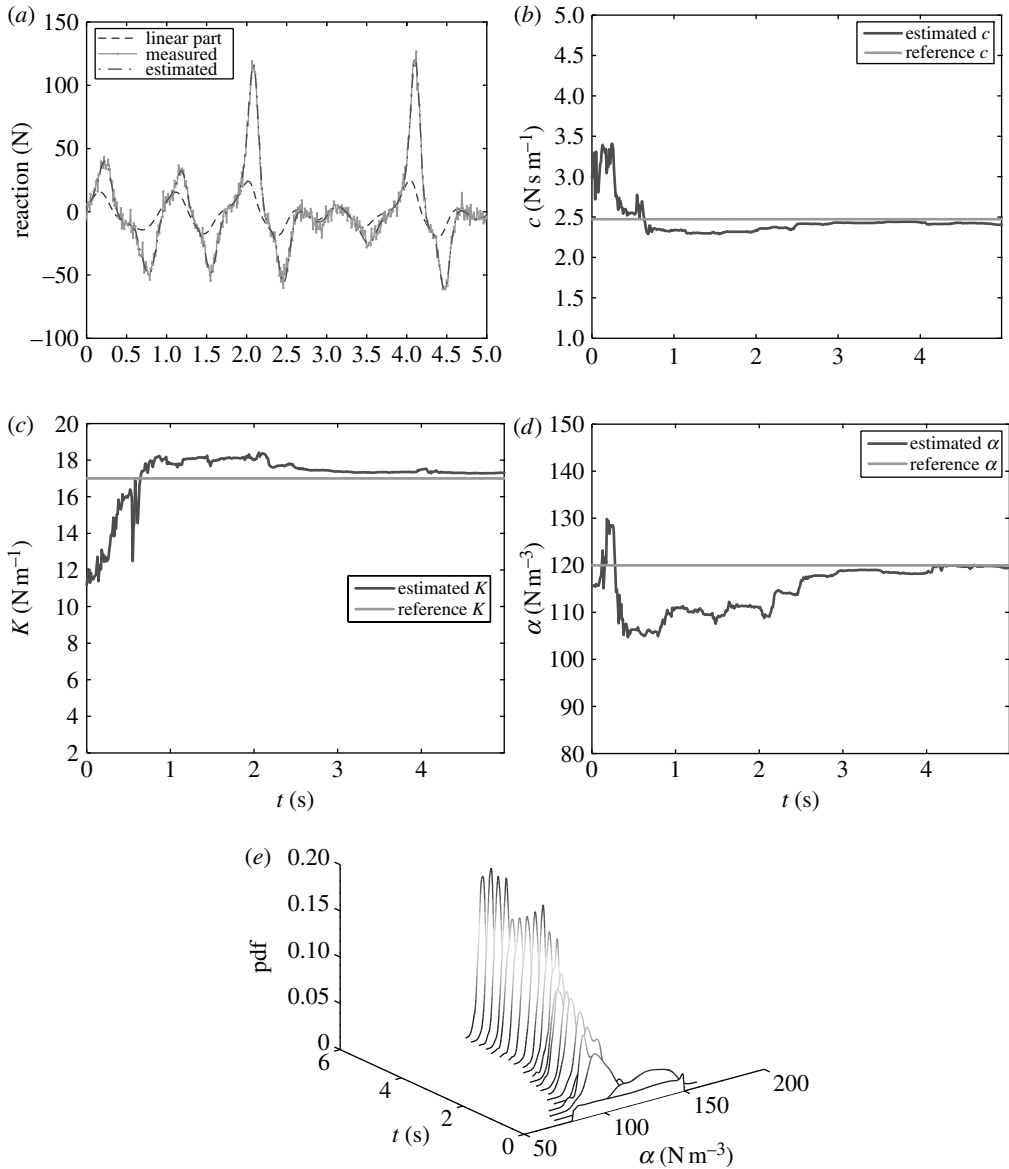


Figure 2. Results for example 5.2: (a) reaction transferred to the support; (b) estimate of conditional expected  $c$ ; (c) estimate of conditional expected  $K$ ; (d) estimate of conditional expected  $\alpha$ ; (e) time evolution of the estimated pdf of the nonlinearity parameter.

Thus, it is of interest to examine the performance of the present method vis-à-vis the existing UPF. For this purpose, we restrict ourselves to the problem of state estimation of Duffing's oscillator when all the system parameters are known. Accordingly, we derive the reduced process equation involving only  $x^1$  and  $x^2$ , with the measurement still being made on the reaction transferred to the support. For the purpose of illustration, we take  $m=1$  kg,  $K=17$   $\text{N m}^{-1}$ ,  $\alpha=120$   $\text{N m}^{-3}$ ,  $c=6.5$   $\text{N s m}^{-1}$ ,  $n_f=1$ ,  $p_1=p=15$  N and  $\lambda=2\pi$   $\text{rad s}^{-1}$ . The process noise intensity is at 1% of the r.m.s. value of the force,  $\sigma=0.12$  N. The



state estimation is carried out for 3 s using a step size of 0.01 s. To gain an insight into the influence of sampling fluctuations and robustness of state estimates as experiments are repeated, we consider the following two cases: (i) case I corresponds to repeated runs of the filter using different sets of particles with the same measurement time history and a fixed sample size, and (ii) case II corresponds to repeated runs of the filter, again with a fixed sample size but with different sets of particles and different measurement time histories. In each case, we consider two different measurement noise levels, namely 1% (sub-cases I(a) and II(a)) and 5% (sub-cases I(b) and II(b)) of noise-free reaction magnitude. While the Ito–Taylor expansion for the measurement equation is presently of (local) order 2.5, the process equation is expanded with 1.5 local order. Standard deviations of noise processes are fixed as  $\sigma_m=0.18$  and 0.85 N. The sample sizes in these two cases are fixed as 90 (high measurement noise) and 300 (low measurement noise), respectively. In both the cases, the threshold sample size (below which resampling is needed) is two-thirds of the sample size. The number of repetitions of filtering is taken to be 20. In implementing the UPF, we use the procedure as outlined by Merwe *et al.* (2000) and take the scaling parameters  $\alpha'=1$ ,  $\beta=2$  and  $\kappa=0$ . Results on the performance of the proposed filter and the UPF are summarized in table 1. It emerges that the proposed strategy is computationally more efficient (with the UPF consuming approx. five times more CPU time) and more accurate (as discerned from the values of r.m.s.e. for the two methods). From the sampling variance of the estimates of r.m.s.e., it is also observed that the present filter shows lesser sampling fluctuations than the UPF. Further studies are clearly needed to understand the nature of sampling fluctuations and convergence characteristics of the proposed method.

(c) *Example 5.3: state and parameter estimation in a nonlinear Euler–Bernoulli beam*

We consider the dynamics of a simply supported, one-span, Euler–Bernoulli beam undergoing large-amplitude oscillations under the action of harmonic excitations. One of the simplified models for this system, incorporating nonlinear strain–displacement relations, is given by (Nayfeh & Mook 1979)

$$m(x) \frac{\partial^2 v_0}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left( EI(x) \frac{\partial^2 v_0}{\partial x^2} \right) - \left( \frac{EA(x)}{2l} \int_0^l \left( \frac{\partial v_0}{\partial x} \right)^2 dx \right) \frac{\partial^2 v_0}{\partial x^2} = f(x, t), \quad (5.14)$$

where  $x \in [0, l]$  is the independent spatial variable;  $v_0(x, t)$  is the transverse displacement field;  $EI(x)$  is the flexural rigidity;  $EA(x)$  is the axial rigidity;  $m(x)$  is the mass per unit length;  $l$  is the beam span; and  $f(x, t)$  is the externally applied load. Measurements are presently made on the strain component  $\varepsilon_{xx}(x, t)$  evaluated at a set of points on the beam surface. This strain component may be approximately related to the transverse displacement  $v_0(x, t)$  and axial displacement  $\tilde{u}_0(x, t)$  through the equation

$$\varepsilon_{xx}(x, y, t) \cong \frac{\partial \tilde{u}_0}{\partial x} - y \frac{\partial^2 v_0}{\partial x^2} + \frac{1}{2} \left( \frac{\partial v_0}{\partial x} \right)^2 \quad (5.15)$$

$$\text{with } \tilde{u}_0(x, t) = -\frac{1}{2} \int_0^x \left[ \frac{\partial v_0(s, t)}{\partial s} \right]^2 ds + \frac{x}{2l} \int_0^l \left[ \frac{\partial v_0(s, t)}{\partial s} \right]^2 ds.$$

Table 1. State estimation of Duffing's oscillator: example 5.2.

metrics of performance	case I(a)		case I(b)		case II(a)		case II(b)	
	proposed method	UPF	proposed method	UPF	proposed method	UPF	proposed method	UPF
mean execution time (s)	12.87	67.89	39.14	213.67	10.64	61.42	35.1	197.25
mean of r.m.s.e. displacement (m)	0.002426	0.034411	0.001929	0.002156	0.002669	0.047771	0.001769	0.002873
mean of r.m.s.e. velocity ( $\text{m s}^{-1}$ )	0.022637	0.192477	0.019772	0.025404	0.023414	0.244306	0.019358	0.027104
variance of r.m.s.e. displacement ( $\text{m}^2$ )	0.000000	0.000049	0.000000	0.000002	0.000000	0.000375	0.000000	0.000006
variance of r.m.s.e. velocity ( $\text{m}^2 \text{s}^{-2}$ )	0.000000	0.002027	0.000000	0.000025	0.000004	0.007156	0.000002	0.000103

To proceed further, we take the displacement field  $v_0(x, t)$  to be expressible by the series  $v_0(x, t) = \sum_{j=1}^{\infty} u_j(t)\varphi_j(x)$ , with  $\varphi_j(x)$  being the  $j$ th linear beam mode shape and  $u_j(t)$  being the  $j$ th generalized coordinate. Applying a Galerkin projection, the governing equations for the generalized coordinates (after including viscous damping) are

$$m_n \frac{d^2 u_n}{dt^2} + K_n u_n + c_n \frac{du_n}{dt} = -\frac{EA}{2l} \left( \frac{n^2 \pi^4}{4l^2} \right) u_n \sum_{p=1}^{\infty} p^2 u_p^2 + f_n(t); \quad n = 1, 2, \dots, \quad (5.16)$$

with  $m_n = \int_0^l m(\varphi_n(x))^2 dx$ ;  $K_n = m_n(n^2 \pi^2 / l^2) \sqrt{EI/m}$  and  $f_n = \int_0^l f(x, t)\varphi_n(x)dx$ . We truncate the series for  $v_0(x, t)$  at the third term and account for the effect of unmodelled dynamics by a set of Gaussian white noise processes. Now, we define the augmented set of states as  $\mathbf{U}(t) = \{U^1, U^2, U^3, U^4, U^5, U^6, U^7\}^T := \{u_1, u_2, u_3, \dot{u}_1, \dot{u}_2, \dot{u}_3, E\}^T$ . Thus, Young's modulus  $E$  is declared as an additional state. This leads to the process equations

$$\begin{aligned} dU^i &= U^{i+3} dt \\ dU^{i+3} &= \left[ -\frac{I\pi^4}{ml^4} U^7 U^i - \frac{c_i}{m_i} U^{i+3} - \frac{A\pi^4}{8m_i l^3} U^7 U^i ((U^1)^2 + 4(U^2)^2 + 9(U^3)^2) + f_n(t) \right] dt \\ &\quad + \sigma_{i+3} dB_{i+3} \\ dU^7 &= \sigma_7 dB_7 \quad i = 1, 2, 3, \dots \end{aligned} \quad (5.17)$$

Assume that two (noisy) strain measurements are available at two different beam locations  $\{(x_1, y_1), (x_2, y_2)\}$  and that the channel noises are independent. The measurement equation is then given by

$$\mathbf{Y}(t_k) = \begin{Bmatrix} \varepsilon_1(t_k) \\ \varepsilon_2(t_k) \end{Bmatrix} = H\mathbf{U}(t_k) + \mathbf{U}(t_k)^T C \mathbf{U}(t_k) + B_k \mathbf{m}_k, \quad (5.18)$$

with

$$\begin{aligned} H &= \begin{bmatrix} -y_1 \boldsymbol{\Phi}_{dd}(x_1) & [0]_{1 \times 4} \\ -y_2 \boldsymbol{\Phi}_{dd}(x_2) & [0]_{1 \times 4} \end{bmatrix}; \quad \boldsymbol{\Phi}_{dd}(x) = \begin{bmatrix} \frac{\partial^2 \phi_1(x)}{\partial x^2} & \frac{\partial^2 \phi_2(x)}{\partial x^2} & \frac{\partial^2 \phi_3(x)}{\partial x^2} \end{bmatrix}; \\ B_k &= \begin{bmatrix} \sigma_{m1} & 0 \\ 0 & \sigma_{m2} \end{bmatrix}; \quad C = \frac{1}{2} \boldsymbol{\Phi}_d^T \boldsymbol{\Phi}_d + \frac{1}{2l} \int_0^l \boldsymbol{\Phi}_d^T \boldsymbol{\Phi}_d dx - \frac{d}{dx} \left[ \frac{1}{2} \int_0^x \boldsymbol{\Phi}_d^T \boldsymbol{\Phi}_d dx \right]; \\ \boldsymbol{\Phi}_d &= \begin{bmatrix} \frac{\partial \phi_1}{\partial x} & \frac{\partial \phi_2}{\partial x} & \frac{\partial \phi_3}{\partial x} \end{bmatrix}; \quad \mathbf{m}_k = \{\xi_{m1} \quad \xi_{m2}\}^T, \end{aligned}$$

where  $\xi_{m1}$  and  $\xi_{m2}$  are mutually independent Gaussian random variables with zero mean and unit standard deviation. The quadratic terms appearing in the measurement equation are further expanded in the Ito–Taylor series with 1.5

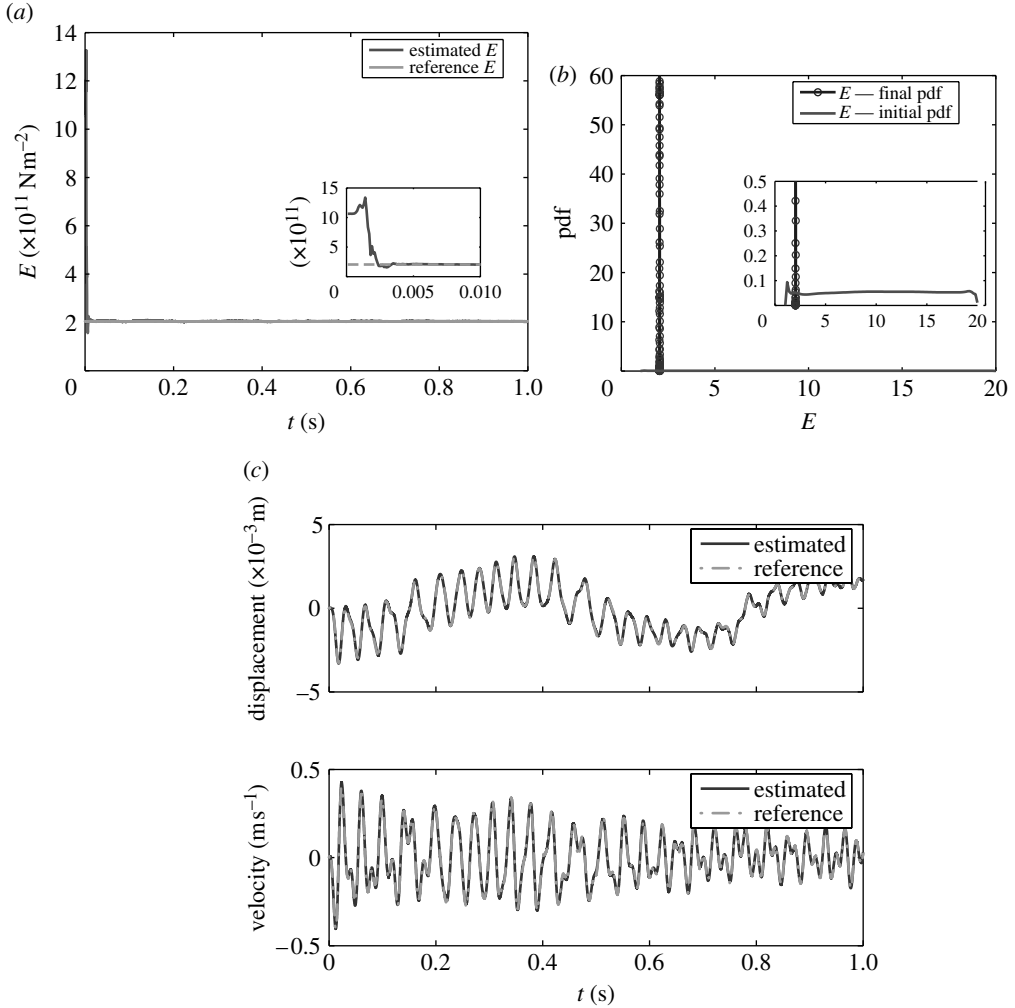


Figure 3. Results for example 5.3: (a) estimate of conditional expected  $E$  (inset: zoomed part up to 0.01 s); (b) estimate of initial and final pdfs of  $E$  (inset: zoomed part up to 0.5 on the  $y$ -axis); (c) estimated displacement and velocity at  $x = l/4$ .

local order, and thus we obtain

$$\begin{aligned}
 (U^p(t_k) U^q(t_k)) &= U_k^p U_k^q = U_{k-1}^p U_{k-1}^q + \left( U_{k-1}^p U_{k-1}^{q+3} + U_{k-1}^q U_{k-1}^{p+3} \right) \Delta \\
 &+ \left( U_{k-1}^p a_{q+3}^{k-1} + a_p^{k-1} U_{k-1}^{q+3} + a_{p+3}^{k-1} U_{k-1}^q + a_q^{k-1} U_{k-1}^{p+3} \right) \frac{\Delta^2}{2} \\
 &+ \sigma_{q+3} U_{k-1}^p I_{q+3,0} + \sigma_{p+3} U_{k-1}^q I_{p+3,0}; \quad p, q = 1, 2, 3, \dots
 \end{aligned}
 \tag{5.19}$$

Here,  $a_j$  denotes the  $j$ th drift term in the governing SDE (5.17). Based on this, the measurement equation is obtained in the form of equation (4.3). It may be noted that all the MSIs that appear in the discretized process and approximate measurement equations are Gaussian in nature (see Saha & Roy (2007) for further details). This simplifies the filtering strategy considerably.

In the numerical work, we take the beam to have a rectangular cross section ( $0.025 \text{ m} \times 0.003 \text{ m}$ ),  $l=0.75 \text{ m}$ , mass density  $= 7850 \text{ kg m}^{-3}$ ,  $c_i=0.1 \text{ N s m}^{-1}$  ( $i=1, 2, 3$ ) and  $E=2.05 \times 10^{11} \text{ N m}^{-2}$ . The external load is given by  $f(x, t) = f_0 \cos[\lambda t + \theta_i]$ , for  $(i-1)(l/5) < x < (il/5)$ ;  $i=1, 2, \dots, 5$  with  $f_0=70 \text{ N m}^{-1}$ ,  $\lambda=10 \text{ rad s}^{-1}$  and  $\{\theta_i\}_{i=1}^5 = \{5.97, 1.45, 3.81, 3.05 \text{ and } 5.60\}^T$ . The process noise intensity is fixed at 1% of the r.m.s. value, with  $\sigma_i=0.1 \text{ N}$ , for  $i=1, 2, 3$ . The simulation has been carried out for 1 s at an interval of 0.0001 s. The measurement noise level is 6% of the r.m.s. value of the maximum strain time history ( $\sigma_m=9.28 \times 10^{-5}$ ). The sample size of 500 with  $N_{\text{thresh}}=167$  is used. As the initial guess, we assume  $E$  to be uniformly distributed, i.e.  $E_{\text{ini}} \sim U(1 \times 10^{11}, 20 \times 10^{11}) \text{ N m}^{-2}$ . Axial strains on the bottom fibre of the beam at  $x=l/4$  and  $l/2$  are measured.

In the numerical work, it is verified that nonlinearity in the strain–displacement equation indeed makes significant contributions to the strain. Figure 3a,b shows the estimates and pdf of Young’s modulus of the beam. While the initial guess of the pdf of  $E$  has a wide support ( $E_{\text{ini}} \sim U(1 \times 10^{11}, 20 \times 10^{11}) \text{ N m}^{-2}$ ), the conditional expectation of  $E$  begins to converge to the reference value,  $E=2.05 \times 10^{11} \text{ N m}^{-2}$ , within a short time of assimilating the measured strains. Estimates of displacement and velocity at  $x=l/4$  are observed to compare well with the corresponding reference trajectories in figure 3c.

## 6. Conclusions

The present work deals with state and parameter estimation in dynamical systems with nonlinear and non-Gaussian state-space models. The emphasis is on dynamical systems of interest in structural mechanics. Nonlinearity in these systems typically arises in relating strains with displacements and/or stresses with strains. Moreover, in problems of parameter identification, the parameters to be identified are declared as additional state variables and the governing equations are thus nonlinear even though the mechanical system is strictly linear for a given set of parameters. In addition, measurements made on quantities, such as forces transferred to the supports and strains, are often nonlinear functions of system displacements and velocities. In the present study, we have treated the noise terms appearing in the process and the measurement equations as white noise processes and integrations of the resulting stochastic differential equations are performed within Ito’s theory. The study reveals that, even as noise terms in the governing SDEs are Gaussian in nature, discretizations of the SDEs through Ito–Taylor series lead to maps that contain both Gaussian and non-Gaussian sequences of random variables. The study addresses the complications in state and parameter estimation problems arising out of the coexistence of nonlinear and non-Gaussian terms in the system equations. Specifically, we develop a scheme to formulate the optimal ispdf applicable to this general class of problems, and this ispdf is embedded into an SIS filtering strategy. We show that the optimal ispdf is non-Gaussian and admits a representation as a weighted sum of Gaussian pdfs, and that we can effectively sample from this representation. While selecting the optimal ispdf, the proposed method incorporates knowledge of both the process dynamics and the current observation. A crucial aspect in the formulation is the use of Ito–Taylor

expansions for the nonlinear terms in the measurement equation. Such expansions permit (at least formally) an improvement of the accuracy of representation by retaining higher-order terms. Limited comparisons of the performance of the proposed method with the UPF indicate that the present method offers substantial computational and numerical advantages.

C.S.M. thanks the Aeronautical R&D Board, Government of India, and D.R. thanks the Naval Science and Technological Laboratory, Government of India, for the financial support provided to conduct this work.

## References

- Brown, R. G. & Hwang, P. Y. C. 1997 *Introduction to random signals and applied Kalman filtering*. New York, NY: Wiley.
- Cappé, O., Godsill, S. J. & Moulines, E. 2007 An overview of existing methods and recent advances in sequential Monte Carlo. *IEEE Proc.* **95**, 899–924. (doi:10.1109/JPROC.2007.893250)
- Ching, J., Beck, J. L. & Porter, K. A. 2006 Bayesian state and parameter estimation of uncertain dynamical systems. *Probab. Eng. Mech.* **21**, 81–96. (doi:10.1016/j.probengmech.2005.08.003)
- Doucet, A., Godsill, S. & Andrieu, C. 2000 On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **10**, 197–208. (doi:10.1023/A:1008935410038)
- Doucet, A., de Freitas, N. & Gordon, N. 2001 *Sequential Monte Carlo methods in practice*. New York, NY: Springer.
- Ghanem, R. & Shinozuka, M. 1995 Structural system identification I: theory. *J. Eng. Mech. ASCE*. **121**, 255–264. (doi:10.1061/(ASCE)0733-9399(1995)121:2(255))
- Ghosh, S. J., Roy, D. & Manohar, C. S. 2007 New forms of extended Kalman filter via transversal linearization and applications to structural system identification. *Comput. Methods Appl. Mech. Eng.* **196**, 5063–5083. (doi:10.1016/j.cma.2007.07.004)
- Gordon, N. J., Salmond, D. J. & Smith, A. F. M. 1993 Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Trans. Radar Signal Process.* **140**, 107–113.
- Imai, H., Yun, B., Maruyama, O. & Shinozuka, M. 1989 Fundamentals of system identification in structural dynamics. *Probab. Eng. Mech.* **4**, 162–173. (doi:10.1016/0266-8920(89)90022-2)
- Kalman, R. E. 1960 A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.* **82**, 35–45.
- Kerschen, G., Golinval, J. C. & Hemez, F. 2003 Bayesian model screening for the identification of non-linear mechanical structures. *J. Vib. Acoust.* **125**, 389–397. (doi:10.1115/1.1569947)
- Kerschen, G., Worden, G. K., Vakakis, A. F. & Golinval, J. C. 2006 Past, present and future of nonlinear system identification in structural dynamics. *Mech. Syst. Signal Process.* **20**, 505–592. (doi:10.1016/j.ymssp.2005.04.008)
- Kloeden, P. E. & Platen, E. 1992 *Numerical solution of stochastic differential equations*. Berlin, Germany: Springer.
- Liu, J. S. 2001 *Monte Carlo strategies in scientific computing*. Springer texts in statistics. New York, NY: Springer.
- Manohar, C. S. & Roy, D. 2006 Monte Carlo filters for identification of nonlinear structural dynamical systems. *Sadhana* **31**, 399–427.
- Merwe, R. V., Doucet, A., de Freitas N. & Wan, E. 2000 The unscented particle filter, Technical report, CUED/F-INFENG/TR.380, Department of Engineering, University of Cambridge, UK.
- Milstein, G. N. 1995 *Numerical integration of stochastic differential equations*. Dordrecht, The Netherlands: Kluwer Academic Press.
- Nomeo, V. & Manohar, C. S. 2007 Non-linear structural dynamical system identification using adaptive particle filters. *J. Sound Vib.* **306**, 524–563. (doi:10.1016/j.jsv.2007.05.040)
- Nayfeh, A. H. & Mook, D. T. 1979 *Nonlinear oscillations*. New York, NY: Wiley International.
- Revuz, D. & Yor, M. 1991 *Continuous martingales and Brownian motion*. Berlin, Germany: Springer.

- Ristic, B., Arulampalam, S. & Gordon, N. 2004 *Beyond the Kalman filter: particle filters for tracking applications*. Boston, MA: Artech House.
- Saha, N. & Roy, D. 2007 Higher order weak linearizations of stochastically driven nonlinear oscillators. *Proc. R. Soc. A* **463**, 1827–1856. (doi:10.1098/rspa.2007.1852)
- Saito, E. & Hoyshiya, M. 1984 Structural identification by extended Kalman filter. *J. Eng. Mech. ASCE* **110**, 1757–1770.
- Sajeeb, R., Manohar, C. S. & Roy, D. 2007 Use of particle filters in an active control algorithm for noisy nonlinear structural dynamical systems. *J. Sound Vib.* **306**, 111–135. (doi:10.1016/j.jsv.2007.05.043)
- Sajeeb, R., Manohar, C. S. & Roy, D. Submitted. Rao–Blackwellization with substructuring for state and parameter estimations of a class of dynamical systems.
- Tanizaki, H. 1996 *Nonlinear filters: estimation and applications*, 2nd edn. Berlin, Germany: Springer.
- Worden, K. & Tomlinson, G. R. 2001 *Nonlinearity in structural dynamics: detection, identification and modeling*. Bristol, UK: Institute of Physics.