Journal of Sound and Vibration 344 (2015) 81–100

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

The estimation of time-invariant parameters of noisy nonlinear oscillatory systems

Mohammad Khalil [a,1], Abhijit Sarkar [a,*], Sondipon Adhikari [b], Dominique Poirel [c]

[a] Carleton University, Department of Civil and Environmental Engineering, Mackenzie Building, Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada
[b] Swansea University, College of Engineering, Singleton Park, Swansea SA2 8PP, UK
[c] Department of Mechanical and Aerospace Engineering, Royal Military College of Canada, Kingston, Ontario K7K 7B4, Canada

ARTICLE INFO

Article history:
Received 1 November 2013
Received in revised form
2 September 2014
Accepted 2 October 2014
Handling Editor: K. Worden
Available online 16 February 2015

ABSTRACT

The inverse problem of estimating time-invariant (static) parameters of a nonlinear system exhibiting noisy oscillation is considered in this paper. Firstly, a Markov Chain Monte Carlo (MCMC) simulation is used for the time-invariant parameter estimation which exploits a non-Gaussian filter, namely the Ensemble Kalman Filter (EnKF) for state estimation required to compute the likelihood function. Secondly, a recently proposed Particle Filter (PF) (that uses the EnKF for its proposal density for the state estimation) has been adapted for combined state and parameter estimation. Numerical illustrations highlight the strengths and limitations of the MCMC, EnKF and PF algorithms for time-invariant parameter estimation. For low measurement noise and dense measurement data, the performances of the MCMC, EnKF and PF algorithms are comparable. For high measurement noise and sparse observational data, the EnKF fails to provide accurate parameter estimates. Hence the adapted PF algorithm becomes necessary in order to obtain parameter estimates comparable in accuracy to the MCMC simulation with EnKF. It highlights the fact that the augmented state space model for the combined state and parameter estimation contains stronger nonlinearity than the original state space model. Hence the EnKF effectively handles the state estimation of the original state space model, but it fails for the combined state and parameter estimation using the augmented system. The effectiveness of the EnKF for the state estimation is therefore leveraged in the MCMC simulation for the time-invariant parameter estimation. In order to obtain accurate parameter estimates using the augmented system, the adapted PF becomes necessary to match the parameter estimates obtained using the MCMC simulation complemented by EnKF for likelihood function computation.

1. Introduction

In engineering systems, noisy oscillation arises in numerous aero-elastic and hydro-elastic problems (e.g. [1–3]). The noisy oscillation phenomenon received widespread attention in the structural dynamics research community as it may lead to large amplitude response and fatigue leading to system failure. The state and parameter estimates of systems exhibiting

* Corresponding author. Tel.: +1 613 520 2600x6320; fax: +1 613 520 3951.
E-mail address: abhijit_sarkar@carleton.ca (A. Sarkar).
[1] Currently at Sandia National Laboratories, P.O. Box 969 MS 9051, Livermore, CA 94551-9051, USA.

http://dx.doi.org/10.1016/j.jsv.2014.10.002
0022-460X/

noisy oscillation obtained from noisy observational data provide valuable information for assessing the safety and reliability of the engineering system in its operational state. In this paper, the problem of time-invariant (static) parameter estimation of a nonlinear system displaying noisy oscillation is considered blending informations contained in the observational data with the predictive model.

The parameter estimation problem of noisy oscillation falls in the general category of system identification [4]. In the context of nonlinear structural modelling, one of the earliest attempts of system identification was made by Ibanez [5] and Masri and Caughey [6]. Various techniques have been subsequently developed to tackle nonlinearity. For a thorough overview, the reader may refer to the book by Worden and Tomlinson [4] and review article by Kerschen et al. [7]. For a Duffing oscillator, Aguilar-Ibanez et al. [8] recently utilized an algebraic approach to the identification of the system parameters. Narayanan et al. [9] applied a hybrid time/frequency-domain-based Fourier series identification method to estimate the Duffing system parameters. The book by Bendat [10] provides a detailed exposition of Volterra series based nonlinear system identification techniques. The spectral identification procedures for nonlinear systems are reported by Zeldin and Spanos [11] and Spanos and Lu [12]. Kougioumtzoglou and Spanos [13] developed identification techniques for nonlinear systems based on harmonic wavelets. Manohar and Roy [14] estimated the state and nonlinear stiffness parameter of the Duffing oscillator from noisy observations using the PF with a Gaussian proposal distribution. In this paper, we continue our previous investigations [15,16] by applying EnKF, and PF for combined state and parameter estimation of a Duffing system.

Kalman Filter (KF) is generally applied to estimate the mean and covariance of the state based on observational data [17]. KF provides an optimal estimate for linear systems with additive Gaussian noise. For weakly nonlinear systems, KF may still provide reasonable estimates using linearization techniques leading to the so-called Extended Kalman Filter (EKF) [18–20]. The major limitation of EKF is due to the third- and higher-order moments in the error covariance evolution equation being discarded, leading to its poor performance [20]. While dealing with strongly nonlinear systems, Monte Carlo based sequential filtering algorithms have gained popularity due to their superiority over EKF. These sampling-based methods represent the probability density function (pdf) of the state vector using a finite number of randomly generated states. Typical examples of Monte Carlo based filters include EnKF [20] and PF [21–23]. EnKF effectively resolves some major problems encountered in EKF, including poor error covariance evolution [24]. Using a non-parametric approach, the nonlinear system identification technique is developed by using EnKF in structural dynamics [25]. EnKF has been shown to perform poorly in some applications involving highly non-Gaussian system behaviour [26,27]. As a remedy to this problem, a filter has been proposed by Anderson [28] that utilizes a weighted sum of Gaussian pdfs to represent the pdf of the system state. For time-invariant parameter estimation using PF for real time applications, Storvik [29] developed a methodology to marginalize the static parameters from the posterior when the pdf of the unknown parameters lies on some low-dimensional sufficient statistics. Storvik's approach [29] avoids the sample impoverishment in PF while estimating time-invariant parameters. For static parameter estimation, Vrugt et al. [30] used the Particle Markov Chain Monte Carlo method [31] which uses PF to design efficient proposal density for MCMC.

PF [21–23] is a Bayesian data assimilation technique that makes neither Gaussian assumption nor linearization of the model and measurement operators. For highly nonlinear systems, PF may provide better estimates compared to EnKF (e.g. [15,16]). For joint state and parameter estimation, PF usually requires an extremely large ensemble of particles (or Monte Carlo samples) compared to EnKF in order to avoid filter divergence in PF [20,32]. The required ensemble size may be reduced in PF through a resampling step [14,23,33,34] in conjunction with an efficient sampling technique, such as Latin Hypercube Sampling (LHS) [35,36], as shown in previous investigations by the authors [15,16]. To further alleviate the requirement of large ensembles, an artificial dynamic model of the unknown parameters is introduced by which the parameters are modelled as Wiener processes [37,38], but the performance of this approach is rather poor [39]. More recently, several investigators [40,33,39] have proposed regularization of the distribution of the state vector. However, the regularization step utilizes convolution kernels which have an effect similar to that caused by the introduction of artificial dynamics [39,38]. In addition to these proposed strategies, a number of PF algorithms have been proposed that utilize other nonlinear filters in providing more efficient proposal distributions, leading to the so-called EKF–PF and Unscented Kalman Filter (UKF)–PF (e.g. [41]) and the EnKF–PF [42,43].

In [42,43], the (non-Gaussian) proposal is obtained with kernel density estimation (KDE) [44] applied to the updated EnKF ensemble using the distance in Sobolev spaces. More specifically, the norm of the Cameron–Martin space [45,46] is used for deriving the density estimates. For the joint state and parameter estimation using nonlinear filters, the state vector is augmented by the unknown parameters to be estimated. The norm associated with the Cameron–Martin space is no longer applicable to this augmented vector since it can no longer be considered as a discrete representation of smooth functions. This PF algorithm will be extended in this paper to deal with the estimation of time-invariant parameters. The authors propose the application of a generalization of Scott's rule for multivariate KDE as in [44,47]. This step consists of applying a Mahalanobis transformation [48] to the augmented state vector to transform the estimated covariance matrix of the augmented state vector to identity. KDE will subsequently be performed using this transformed vector using a generalization of Scott's rule [44,47] for the multivariate case. Finally, the estimated pdf will be transformed back to the original coordinate. The proposed PF circumvents the need to artificially inflate the variances of the parameters in order to avoid filter divergence, thus making it well-suited for the time-invariant (i.e. static or fixed) parameter estimation.

For the parameter estimation of dynamical systems using nonlinear filters, the unknown system parameters are concatenated to the state vector leading to a higher dimensional state space model than the original system. Even if the

original state space model is linear, the augmented system generally becomes nonlinear. For the filter convergence, the static parameters are treated as time-varying quantities being perturbed by the artificial noise for the optimal filter performance (e.g. [15,38,49]). Although an appropriate choice of the joint prior pdf may alleviate this difficulty, it becomes difficult for high-dimensional parameter spaces.

In this paper, we adopt a MCMC simulation for the time-invariant parameter estimation complemented by a non-Gaussian filter, namely EnKF for the state estimation. Due to the presence of stronger nonlinearity in the augmented system, the combined state and parameter estimation method fails to provide accurate estimates using EnKF. Consequently, a recently proposed PF algorithm by Mandel and Beezley [42,43] has been extended for the combined state and parameter estimation in order to obtain results which are comparable in accuracy to the MCMC simulation. As alluded to previously, the augmented state space model, constructed for the combined state and parameter estimation, inherits stronger nonlinearity than the original state space model used for the state estimation. This fact is exploited in the MCMC based parameter estimation procedure whereby the state estimation problem is adequately handled by EKF (e.g. [50–53]) as the conditional pdf of the state is close to Gaussian. In this paper, we consider the cases when the conditional pdf of the state vector can no longer be approximated to be Gaussian. To this end, the contributions of the paper are as follows:

(1) Development of a time-invariant (static) parameter estimation algorithm for strongly nonlinear oscillatory systems using MCMC simulation whereby the likelihood function in the parameter posterior is computed by EnKF based non-Gaussian filter. To the authors' best knowledge, such nested two-level sampling (i.e. the first level for EnKF and second level for MCMC) scheme, that is capable of handling strong nonlinearity in mechanical oscillators, has not been reported in the literature.
(2) We extend the PF algorithm by Mandel and Beezley [37,38], originally proposed for just state estimation in atmospheric science applications, to handle combined state and parameter estimation of strongly nonlinear oscillatory systems.
(3) The application of these algorithms for the time-invariant parameter estimation of a noisy nonlinear oscillatory system which highlights the merits and limitations of (a) MCMC simulation complemented by EnKF for likelihood function computation, (b) EnKF and (c) the adapted PF algorithms. Such comparative analysis has not been reported in the literature. This specific application clearly illustrates the strengths and limitations of these algorithms which is not apparent from mathematical formulations alone.

In this investigation, we conduct extensive numerical investigations using the Duffing oscillator model to study the effects of sparsity of observational data and strength of measurement noise on the parameter estimates obtained using MCMC, EnKF and an extension of the newly developed PF by Mandel and Beezley [42,43]. As the Duffing oscillator model displays a wide range of dynamical behaviour with small perturbations to the system parameters [54,55], it is chosen for numerical investigation.

2. Mathematical formulation for parameter estimation

The model and measurement equations for a discrete state-space representation of a nonlinear system are given by [18–20,22,23,33,38,41,56]

$$\mathbf{u}_{k+1} = \mathfrak{g}_k(\mathbf{u}_k, \boldsymbol{\phi}, \mathbf{f}_k, \mathbf{q}_k^u), \tag{1}$$

$$\mathbf{d}_j = \mathfrak{h}_j(\mathbf{u}_{k(j)}, \boldsymbol{\phi}, \boldsymbol{\epsilon}_j). \tag{2}$$

Here $\mathbf{u} \in \mathbb{R}^n$ is the state vector, $\mathfrak{g} \in \mathbb{R}^n$ is the discrete nonlinear model operator, $\boldsymbol{\phi} \in \mathbb{R}^{n_\phi}$ is the static (time-invariant) parameter vector, $\mathbf{f} \in \mathbb{R}^p$ is a deterministic input, $\mathbf{q}_k^u \in \mathbb{R}^s$ is a Gaussian random vector with the mean $\overline{\mathbf{q}_k^u} \in \mathbb{R}^s$ and covariance matrix $\mathbf{Q}^{uu} \in \mathbb{R}^{s\times s}$ respectively and $\mathbf{d} \in \mathbb{R}^m$, the measurement vector, maps the true state by the nonlinear measurement operator $\mathbf{h} \in \mathbb{R}^m$; $\boldsymbol{\epsilon}_k \in \mathbb{R}^r$ is a Gaussian random vector with the mean $\overline{\boldsymbol{\epsilon}_k} \in \mathbb{R}^r$ and covariance matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{r\times r}$ respectively. Here $\mathbf{q}_k^u$ and $\boldsymbol{\epsilon}_k$ are assumed to be independent. Note that the indexes k and $k(j)$ denote the time steps for the model integration and the arrival of measurement data (e.g. [20]).

2.1. Nonlinear filtering approach

In the framework of nonlinear filtering, the unknown time-invariant parameter vector $\boldsymbol{\phi}$ is treated as a time-varying quantity (e.g. [15,38]). Additionally, the parameter vector is artificially perturbed randomly for filter convergence (e.g. [15,38]). Therefore, the revised state space model is written as (e.g. [15,38])

$$\mathbf{u}_{k+1} = \mathfrak{g}_k(\mathbf{u}_k, \boldsymbol{\phi}_k, \mathbf{f}_k, \mathbf{q}_k^u), \tag{3}$$

$$\boldsymbol{\phi}_{k+1} = \boldsymbol{\phi}_k + \mathbf{q}_k^\phi, \tag{4}$$

$$\mathbf{d}_j = \mathfrak{h}_j(\mathbf{u}_{k(j)}, \boldsymbol{\phi}_{k(j)}, \boldsymbol{\epsilon}_j) \tag{5}$$

where $\mathbf{q}_k^{\phi}$ is the artificial random noise vector which may require finetuning for optimal filter performance (e.g. [15,38]). Next the new state vector is formed by appending the unknown parameter vector to the original state vector leading to the following augmented state space model:

$$\mathbf{x}_{k+1} = \mathbf{g}_k(\mathbf{x}_k, \mathbf{f}_k, \mathbf{q}_k), \tag{6}$$

$$\mathbf{d}_j = \mathbf{h}_j(\mathbf{x}_{k(j)}, \boldsymbol{\epsilon}_j), \tag{7}$$

where

$$\mathbf{x}_k = \begin{Bmatrix} \mathbf{u}_k \\ \boldsymbol{\phi}_k \end{Bmatrix}; \quad \mathbf{q}_k = \begin{Bmatrix} \mathbf{q}_k^u \\ \mathbf{q}_k^{\phi} \end{Bmatrix} \tag{8}$$

Consequently, any nonlinear filter can be used to tackle the combined state and parameter estimation from which the relevant parameter vector is easily extracted. In this paper, we use EnKF and PF to obtain the joint state and parameter estimates. These results are compared to those obtained using MCMC approach as described next.

2.2. Markov chain Monte Carlo simulation approach

In the general Bayesian framework (e.g. [20]) , the joint pdf of the state and parameter vector can be written as [57]

$$\begin{aligned} p(\mathbf{u}_1, \ldots, \mathbf{u}_{k(J)}, \ldots, \mathbf{u}_k, \boldsymbol{\phi} | \mathbf{d}_1, \ldots, \mathbf{d}_J) \propto p(\boldsymbol{\phi}) \left[\prod_{k=1}^{k(1)} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \right] p(d_1 | \mathbf{u}_{k(1)}, \boldsymbol{\phi}) \\ \vdots \\ \left[\prod_{k=k(J-1)+1}^{k(J)} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \right] p(d_J | \mathbf{u}_{k(J)}, \boldsymbol{\phi}) \\ \left[\prod_{k=k(J)+1}^{K} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \right] \end{aligned} \tag{9}$$

Consequently, the marginal posterior of the parameter vector can be obtained as [57]

$$\begin{aligned} p(\boldsymbol{\phi} | \mathbf{d}_1, \ldots, \mathbf{d}_J) \propto p(\boldsymbol{\phi}) \left[\prod_{k=1}^{k(1)-1} \int_{\infty}^{\infty} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_k \right] \left[\int_{\infty}^{\infty} p(\mathbf{u}_{k(1)} | \mathbf{u}_{k(1)-1}, \boldsymbol{\phi}) p(d_1 | \mathbf{u}_{k(1)}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_{k(1)} \right] \\ \vdots \\ \left[\prod_{k=k(J-1)+1}^{k(J)-1} \int_{\infty}^{\infty} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_k \right] \left[\int_{\infty}^{\infty} p(\mathbf{u}_{k(J)} | \mathbf{u}_{k(J)-1}, \boldsymbol{\phi}) p(d_J | \mathbf{u}_{k(J)}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_{k(J)} \right] \\ \left[\prod_{k=k(J)+1}^{K} \int_{\infty}^{\infty} p(\mathbf{u}_k | \mathbf{u}_{k-1}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_k \right] \end{aligned} \tag{10}$$

The above equation can be concisely written as [57]

$$p(\boldsymbol{\phi} | \mathbf{d}_1, \ldots, \mathbf{d}_J) \propto p(\boldsymbol{\phi}) \underbrace{\prod_{j=1}^{J} \left[\int_{\infty}^{\infty} p(\mathbf{u}_{k(j)} | \mathbf{u}_{k(j)-1}, \boldsymbol{\phi}) p(d_j | \mathbf{u}_{k(j)}, \boldsymbol{\phi}) \, \mathrm{d}\mathbf{u}_{k(j)} \right]}_{\textit{likelihood function}} \tag{11}$$

where the computation of $p(\mathbf{u}_{k(j)} | \mathbf{u}_{k(j)-1}, \boldsymbol{\phi})$ involves a state estimation problem. In our previous work, the likelihood function in Eq. (11) is computed semi-analytically using the Extended Kalman Filter (EKF) [51]. The EKF based approach provides acceptable results when the state vector is close to Gaussian; and this methodology is computationally attractive as it avoids Monte Carlo sampling for state estimation. As the state becomes non-Gaussian due to nonlinearities in the system, EKF based likelihood computation becomes highly inaccurate. In the current investigation, EnKF handles the state estimation for the likelihood computation which involves Monte Carlo sampling, but accurately handles non-Gaussian state estimation.

MCMC method [58,59] can be used to generate samples of the parameter vector from $p(\boldsymbol{\phi} | \mathbf{d}_1, \ldots, \mathbf{d}_J)$ (known only up to a proportionality) as detailed in [51,60–62]. In this investigation, a new sample $\boldsymbol{\phi}'$ of the parameter vector is generated from the current sample ϕ using a random-walk Metropolis–Hastings (MH) algorithm [59] as described in [51,60–62]. In this approach, the acceptance probability $\alpha(\boldsymbol{\phi}, \boldsymbol{\phi}')$ is (e.g. [58,59])

$$\alpha(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min\left(1, \frac{p(\boldsymbol{\phi}' | \mathbf{d}_1, \ldots, \mathbf{d}_J) \; q(\boldsymbol{\phi}', \boldsymbol{\phi})}{p(\boldsymbol{\phi} | \mathbf{d}_1, \ldots, \mathbf{d}_J) \; q(\boldsymbol{\phi}, \boldsymbol{\phi}')}\right) \tag{12}$$

where the proposal density q is assumed to be uniformly distributed in $[-w, w]^{n_\phi}$ [51]. The appropriate choice of the proposal width w is critical to the efficiency of MH MCMC and chosen to have 25 percent acceptance ratio [51].

3. Ensemble Kalman filter

In this section, we generalize the analysis step of EnKF for nonlinear measurements, closely following the derivation of the Unscented Kalman filter (e.g. [63]). For strong nonlinearity, EKF may provide erroneous conditional mean and covariance estimates due to linearization of the model and measurement operators. In EnKF, proposed by Evensen [20], a finite number of Monte Carlo samples of the state vector $\mathbf{x}_k$ are propagated forward in time using the original model operator. One can estimate the pdf of $\mathbf{x}_k$ using ensemble averaging. However, EnKF performs a linear analysis step which involves the linearization of nonlinear measurement operator. Furthermore, the analysis step assumes a Gaussian state and measurement noise which offers computational efficiency but introduces errors in the estimated conditional pdf of $\mathbf{x}_k$.

Furthermore the prior pdf of $\mathbf{x}$ is assumed to be $\mathbf{x}_k \sim p(\mathbf{x}_k^f)$. A prior pdf is the pdf of the state prior to assimilating the available data. In the analysis step, EnKF estimates the posterior pdf $\mathbf{x}_{k(j)} \sim p(\mathbf{x}_{k(j)}^a)$ based on the measurement vector $\mathbf{d}_j$.

Considering again the model and measurement equations as described by Eqs. (6)–(7), EnKF provides estimates of the conditional mean and covariance of the state vector as follows [20]:

(1) Create an initial ensemble $\{\mathbf{x}_{0,i}^f\}$ of size N with $i=1,\ldots,N$, using the prior pdf of $\mathbf{x}_0$.
(2) Compute perturbed measurements as

$$\mathbf{d}_{j,i} = \mathbf{h}_j\left(\mathbf{x}_{k(j),i}^f, \boldsymbol{\epsilon}_{j,i}\right). \tag{13}$$

(3) Analysis step:

$$\overline{\mathbf{d}}_j = \frac{1}{N}\sum_{i=1}^{N} \mathbf{d}_{j,i}, \tag{14}$$

$$\overline{\mathbf{x}}_{k(j)}^f = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_{k(j),i}^f, \tag{15}$$

$$\mathbf{P}_{xd} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{x}_{k(j),i}^f - \overline{\mathbf{x}}_{k(j)}^f\right)(\mathbf{d}_{j,i} - \overline{\mathbf{d}}_j)^T, \tag{16}$$

$$\mathbf{P}_{dd} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{d}_{j,i} - \overline{\mathbf{d}}_j)(\mathbf{d}_{j,i} - \overline{\mathbf{d}}_j)^T, \tag{17}$$

$$\mathbf{K}_{k(j)} = \mathbf{P}_{xd}\mathbf{P}_{dd}^{-1}, \tag{18}$$

$$\overline{\boldsymbol{\epsilon}_j} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\epsilon}_{j,i}, \tag{19}$$

$$\boldsymbol{\Gamma}_j = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{\epsilon}_{j,i} - \overline{\boldsymbol{\epsilon}_j})(\boldsymbol{\epsilon}_{j,i} - \overline{\boldsymbol{\epsilon}_j})^T \tag{20}$$

$$\mathbf{x}_{k(j),i}^a = \mathbf{x}_{k(j),i}^f + \mathbf{K}_{k(j)}\left(\mathbf{d}_{j,i} - \mathbf{h}_j(\mathbf{x}_{k(j),i}^f, \boldsymbol{\epsilon}_{j,i})\right) \tag{21}$$

(4) Forecast step:

$$\mathbf{x}_{k+1,i}^f = \mathbf{g}_k(\mathbf{x}_{k,i}^a, \mathbf{f}_k, \mathbf{q}_{k,i}) \tag{22}$$

Here we generalize the analysis step of EnKF closely following the derivation of the Unscented Kalman filter (e.g. [63]). For strongly nonlinear models, the analysis step in EnKF is the major limitation due to the implicit Gaussian assumption of the state vector. For the case of *additive* measurement noise $\mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\epsilon}_k) = \mathbf{h}_k(\mathbf{x}_k) + \boldsymbol{\epsilon}_k$, one can however augment the original state vector with $\mathbf{h}_k(\mathbf{x}_k)$ for full nonlinear analysis [64]. The perturbation of measurements can be avoided using the square-root algorithm [20]. In contrast to EKF, EnKF avoids the memory-intensive storage of the covariance matrix $\mathbf{P}_k$ [20]. In EnKF, the statistics of the state vector can be obtained by ensemble averaging which introduces statistical sampling errors [20].

4. Particle filter

In the section, we extend the PF algorithm proposed by Mandel and Beezley [42,43], originally used for the state estimation, to handle combined state and parameter estimation exploiting a general multivariate kernel density estimator. PF [21–23,18] can handle the most general forms of nonlinearities in measurement and model operators and non-Gaussian

model and measurement noise. In contrast to EKF and EnKF, no Gaussian assumption on the state is necessary. Therefore, PF is superior to EnKF and EKF for strongly non-Gaussian models (e.g. [26,27]). The formulation of PF presented next closely follows Tanizaki [21] and Ristic et al. [23].

Once again, Eqs. (6) and (7) describe the model and measurement respectively. Let us define the collection of the state and measurement vectors using the following matrices [21,23]:

$$\mathbf{X}_k = \{\mathbf{x}_1, \ldots, \mathbf{x}_{k(1)}, \ldots, \mathbf{x}_{k(j)}, \ldots, \mathbf{x}_k\}, \tag{23}$$

$$\mathbf{D}_j = \{\mathbf{d}_1, \ldots, \mathbf{d}_j\}. \tag{24}$$

By Bayes' Theorem, we have (e.g. [21–23])

$$p(\mathbf{X}_k|\mathbf{D}_j) = \frac{p(\mathbf{D}_j|\mathbf{X}_k)p(\mathbf{X}_k)}{\int p(\mathbf{D}_j|\mathbf{X}_k)p(\mathbf{X}_k)\mathrm{d}\mathbf{X}_k}. \tag{25}$$

Under the assumptions that (1) the successive states of the system form a Markov chain of order one and (2) the measurement vectors are statistically independent at different time instants are independent, a recursive relationship between the posterior pdf at time t_k and time t_{k-1} can be obtained as (e.g. [21,23,65,66])

$$p(\mathbf{X}_{k(j)}|\mathbf{D}_j) = p(\mathbf{X}_{k(j)-1}|\mathbf{D}_{j-1})\frac{p(\mathbf{d}_j|\mathbf{x}_{k(j)})p(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1})}{p(\mathbf{d}_j|\mathbf{D}_{j-1})}. \tag{26}$$

In filtering applications, quantities of interest relating to the conditional state vector, including the conditional mean and variance, can be estimated as follows:

$$\mathrm{E}[g(\mathbf{X}_k)|\mathbf{D}_j] = \int g(\mathbf{X}_k)p(\mathbf{X}_k|\mathbf{D}_j)\,\mathrm{d}\mathbf{X}_k. \tag{27}$$

Although there are methods for generating samples from an arbitrary distribution, the computational cost is generally high, especially when the probability density is multidimensional [59,67,68]. Thus, it is assumed that samples can be easily generated according to another pdf $q(\mathbf{X}_k|\mathbf{D}_j)$ (i.e. the proposal distribution, or importance density function) and that one could evaluate $p(\mathbf{X}_k|\mathbf{D}_j)$ easily. It is then possible to rewrite the integral in Eq. (27) as

$$\mathrm{E}[g(\mathbf{X}_k)|\mathbf{D}_j] = \int g(\mathbf{X}_k)\frac{p(\mathbf{X}_k|\mathbf{D}_j)}{q(\mathbf{X}_k|\mathbf{D}_j)}q(\mathbf{X}_k|\mathbf{D}_j)\,\mathrm{d}\mathbf{X}_k \tag{28}$$

which can be approximated using independent samples $\{\mathbf{X}_{k,i}\}$, $i = 1, \ldots, N$, distributed according to $q(\mathbf{X}_k|\mathbf{D}_j)$ by (e.g. [21–23])

$$\hat{\mathrm{E}}[g(\mathbf{X}_k)|\mathbf{D}_j] = \sum_{i=1}^{N} g(\mathbf{X}_{k,i})w_{k,i} \tag{29}$$

where the weights are given by

$$w_{k,i} = \frac{p(\mathbf{X}_{k,i}|\mathbf{D}_j)}{q(\mathbf{X}_{k,i}|\mathbf{D}_j)} \tag{30}$$

In practice one does not know the analytical expression of $p(\mathbf{X}_k|\mathbf{D}_j)$ as it is precisely the density that one seeks to approximate. However, the calculation of the weights admits a recursive expression. Using the identity in Eq. (26) and a proposal distribution that permits the following factorization:

$$q(\mathbf{X}_{k(j)}|\mathbf{D}_j) = q(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1}, \mathbf{D}_j)\; q(\mathbf{X}_{k(j)-1}|\mathbf{D}_{j-1}), \tag{31}$$

the recursive expression of the weights becomes (e.g. [21,23,65,66])

$$w_{k(j),i} \propto w_{k(j)-1,i}\,\frac{p(\mathbf{d}_j|\mathbf{x}_{k(j),i})p(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i})}{q(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i}, \mathbf{D}_j)}. \tag{32}$$

In the above expression, $q(\mathbf{x}_{k,i}|\mathbf{x}_{k-1,i}, \mathbf{D}_j)$ is the proposal density used to sample the kth point of the trajectory, conditional on the previous states and measurements. The particle filtering algorithm consists of two steps at each iteration:

(1) Forecast step: Obtain samples from a proposal distribution

$$\mathbf{x}_{k,i} \sim q(\mathbf{x}_{k,i}|\mathbf{x}_{k-1,i}, \mathbf{D}_j), \quad i = 1, \ldots, N \tag{33}$$

(2) Analysis step: Update the weights of all particles based on the observation $\mathbf{d}_j$ according to

$$\tilde{w}_{k(j),i} = w_{k(j)-1,i}\,\frac{p(\mathbf{d}_j|\mathbf{x}_{k(j),i})p(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i})}{q(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i}, \mathbf{D}_j)} \tag{34}$$

$$w_{k(j),i} = \frac{\tilde{w}_{k(j),i}}{\sum_{i=1}^{N} \tilde{w}_{k(j),i}}. \tag{35}$$

4.1. Necessity of a resampling step

In the classical problem of recursive estimation, uncertainty of the state is initially high which gradually decreases with the sequential assimilation of observations. In most cases, the support of $p(\mathbf{x}_k|\mathbf{D}_j)$ is initially large and decreases over time. The particles are scattered across the state space according to some proposal distribution. Depending on the choice of proposal distribution, in some cases one ends up with the majority of particles being dispersed in areas where $p(\mathbf{x}_k|\mathbf{D}_j)$ is very small in value. Therefore, these particles have a very low weight. The severity of this so-called degeneracy phenomenon depends on the problem and the choice of the proposal distribution.

To alleviate the degeneracy of the particles, a resampling step will be undertaken to remove particles of negligible weight and duplicate the particles of high weight. There are different methods for resampling including multinomial sampling [34,69–71]), residual resampling [72,73] and stratified sampling [72]. Resampling can be performed at every step of the filter, or only performed if degeneracy is detected. The resulting filter is commonly known as a *Sampling Importance Resampling* (SIR) filter. The multinomial resampling method used in this investigation is summarized as follows:

(1) Draw N independent and identically distributed samples from a uniform distribution on $[0, 1]$

$$u_i \sim \mathcal{U}(0,1), \quad i = 1, \ldots, N \tag{36}$$

(2) Obtain the cumulative density function v of the weights:

$$v_j = \sum_{l=1}^{j} w_{k,l} \tag{37}$$

(3) For $m = 1, \ldots, N$, determine n such that $v_n \le u_m < v_{n+1}$ and copy the nth particle $\mathbf{x}_{k,n}$ into the new ensemble
(4) Reset the particle weights $w_{k,i} = 1/N$, for $i = 1, \ldots, N$.

An indicator of the presence of degeneracy the effective ensemble size defined by [22,73–75]

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} (w_{k,i})^2}. \tag{38}$$

N_{eff} is equal to the ensemble size N if all weights are equal and decreases as the extent of degeneracy increases. Thus, resampling can be performed when the effective ensemble size falls below a certain threshold value, denoted by N_{thr} [74,73,22,75]. A common threshold for resampling is $N_{\text{eff}} = N/3$ [22], which will be used herein.

It is important to note that although the resampling step is essential to the effective functioning of the particle filter, it may have adverse effects on the operation of the filter and the quality of the particle representation of the posterior pdf $p(\mathbf{x}_k|\mathbf{D}_j)$. Indeed, particles with relatively high weight are duplicated numerous times leading to loss of diversity among the ensemble. If the model noise term is relatively weak, the ensemble members will be close to each other during the subsequent iterations and the space state is insufficiently explored. This is the motivation behind performing the resampling step when degeneracy is detected.

4.2. Choice of proposal distribution

The resampling step aims at reducing the effect of the degeneracy phenomenon whereby the majority of particles are clustered in areas where $p(\mathbf{x}_k|\mathbf{D}_k)$ is very small, due to a poor choice of proposal distribution. Thus the judicious selection of proposal distribution minimizes the occurrence of degeneracy and reduce the overall number of resampling steps.

Utilizing the prior pdf as the proposal distribution, i.e.

$$q(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1}, \mathbf{D}_j) = p(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1}) \tag{39}$$

leads to the simplest method, often referred to as the *bootstrap filter* [34]. Such proposal pdf ignores the most recent observation $\mathbf{d}_k$. The particle trajectories are obtained by simply applying the stochastic model equation to update the state vector of each particle. The most recent observation is used in the weight-update equation which simplifies to [23,34]

$$\tilde{w}_{k(j),i} = w_{k(j)-1,i}\ p(\mathbf{d}_j|\mathbf{x}_{k(j),i}). \tag{40}$$

This choice of proposal distribution ignores the most recent observation and thus is not optimal (it does not significantly minimize the occurrence rate of degeneracy). The optimal proposal distribution (or importance function) is the distribution that minimizes the variance of the particle weights (and thus also minimizes the occurrence of the degeneracy

phenomenon) [22]. The optimal importance function is given by [22]

$$q(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1},\mathbf{D}_j)=p(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1,i},\mathbf{d}_j) \tag{41}$$

which would in turn give the following weight update equation [22]:

$$\tilde{w}_{k(j),i}=w_{k(j)-1,i}\ p(\mathbf{d}_j|\mathbf{x}_{k(j)-1,i}). \tag{42}$$

The optimal proposal suffers from two major drawbacks. It requires the ability to sample from $p(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1,i},\mathbf{d}_j)$ and to evaluate $p(\mathbf{d}_j|\mathbf{x}_{k(j)-1,i})$, which cannot be achieved in general cases. On the other hand, a Gaussian proposal distribution of the form

$$q(\mathbf{x}_{k(j)}|\mathbf{x}_{k(j)-1},\mathbf{D}_j)=\mathcal{N}(\mathbf{x}_{k(j)};\hat{\mathbf{x}}_{k(j)},\mathbf{P}_{k(j)}) \tag{43}$$

could be easily sampled. Furthermore, $p(\mathbf{d}_j|\mathbf{x}_{k(j)-1,i})$ would be easy to evaluate. This proposal distribution can take the form of the posterior or forecast pdf obtained by EKF or UKF. For instance, Van der Merwe and Wan [76] developed a particle filter with importance function given by a bank of unscented Kalman filters. A separate UKF is associated with each particle providing a unique proposal for each of the particles. Since the proposal distribution is not unique for all particles, this proposal results in a sub-optimal filter [77].

4.2.1. Proposal distribution as forecast of EnKF

For strongly nonlinear systems, the proposal based on EnKF is perhaps most appropriate as it maintains some non-Gaussian characteristics in the proposal. Unlike EKF and UKF based proposal, EnKF proposal does not differ from particle to particle. In other words, a single proposal is used to update the particle weights as explained next. The (non-Gaussian) EnKF based proposal $q(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i},\mathbf{D}_j)$ is obtained by the kernel density estimation (KDE) [44]. The forecast pdf $p(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i})$ of the particle filter is also estimated by KDE. This proposal is proposed by Mandel and Beezley and termed as the predictor–corrector filter [42,43], in which the following weight update equation was proposed:

$$\begin{aligned}\tilde{w}_{k(j),i}&=w_{k(j)-1,i}\frac{p(\mathbf{d}_j|\mathbf{x}_{k(j),i})p(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i})}{q(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i},\mathbf{D}_j)}\\&=w_{k(j)-1,i}p\left(\mathbf{d}_j|\mathbf{x}^a_{k(j),i}\right)\frac{\sum_{l=1}^{N}w_{k(j)-1,l}\mathbf{1}_{\|\mathbf{x}^f_{k(j),l}-\mathbf{x}^a_{k(j),i}\|_H\le h(\mathbf{x}^a_{k(j),i})}}{\sum_{l=1}^{N}\frac{1}{N}\mathbf{1}_{\|\mathbf{x}^a_{k(j),l}-\mathbf{x}^a_{k(j),i}\|_H\le h(\mathbf{x}^a_{k(j),i})}},\end{aligned} \tag{44}$$

with $\mathbf{1}_{(\cdot)}$ being the indicator function and $h(\mathbf{x}^a_{k(j),i})$ the bandwidth, chosen to be the distance to the $\sqrt{N}$−th nearest sample (data point) in the $\|\cdot\|_H$ norm. Mandel and Beezley were concerned with the state (and not parameter) estimation problem in which the vector $\mathbf{x}$ consists of a discrete representation of smooth functions, for which the norm associated with the so-called Cameron–Martin space is suitable [45,46].

In this investigation, the authors propose to extend this methodology to tackle the problem of joint state and parameter estimation problem. The augmented state vector can no longer be considered as a discrete representation of smooth functions, and thus a different approach to KDE must be taken. The general multivariate kernel estimator [78] will be used resulting in a weight update equation given by [57]

$$\begin{aligned}\tilde{w}_{k(j),i}&=w_{k(j)-1,i}\frac{p(\mathbf{d}_j|\mathbf{x}_{k(j),i})p(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i})}{q(\mathbf{x}_{k(j),i}|\mathbf{x}_{k(j)-1,i},\mathbf{D}_j)}\\&=w_{k(j)-1,i}p\left(\mathbf{d}_j|\mathbf{x}^a_{k(j),i}\right)\frac{\sum_{l=1}^{N}w_{k(j)-1,l}\frac{1}{|H_f|}K\left(H_f^{-1}\left(\mathbf{x}^f_{k(j),l}-\mathbf{x}^a_{k(j),i}\right)\right)}{\sum_{l=1}^{N}\frac{1}{N}\frac{1}{|H_a|}K\left(H_a^{-1}\left(\mathbf{x}^a_{k(j),l}-\mathbf{x}^a_{k(j),i}\right)\right)},\end{aligned} \tag{45}$$

with a Gaussian kernel $K(\mathbf{y})=(2\pi)^{-n/2}\exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{y}\right)$ and a bandwidth matrix H proportional to the square root of the ensemble covariance matrix, as suggested in [47]. The constant of proportionality is obtained by generalizing Scott's rule for the univariate case to the multivariate one [44], in which case one obtains

$$\begin{aligned}H_f&=N_{\text{eff}}^{-1/(n+4)}\Sigma_f^{1/2}\\&=\left(\frac{1}{\sum_{i=1}^{N}(w_{k(j),i})^2}\right)^{-1/(n+4)}\\&\times\left[\left(\frac{1}{1-\sum_{i=1}^{N}{w^f_{k(j),i}}^2}\right)\sum_{l=1}^{N}w^f_{k(j),l}\left(\mathbf{x}^f_{k(j),l}-\overline{\mathbf{x}}^f_{k(j)}\right)\left(\mathbf{x}^f_{k(j),l}-\overline{\mathbf{x}}^f_{k(j)}\right)^T\right]^{1/2}\\H_a&=N^{-1/(n+4)}\Sigma_a^{1/2}\end{aligned} \tag{46}$$

$$= N^{-1/(n+4)}\left[\left(\frac{1}{N-1}\right)\sum_{l=1}^{N}\left(\mathbf{x}^a_{k(j),l}-\overline{\mathbf{x}}^a_{k(j)}\right)\left(\mathbf{x}^a_{k(j),l}-\overline{\mathbf{x}}^a_{k(j)}\right)^T\right]^{1/2} \tag{47}$$

The algorithm of the particle filter with EnKF proposal is summarized next:

(1) Using the prior pdf of $\mathbf{x}_0$, generate an ensemble $\{\mathbf{x}^f_{0,i}\}$ having size N with $i=1,\ldots,N$.
(2) Analysis step:

$$\mathbf{d}_{j,i} = \mathbf{h}_j\left(\mathbf{x}^f_{k(j),i}, \boldsymbol{\epsilon}_{j,i}\right), \tag{48}$$

$$\overline{\mathbf{d}}_j = \sum_{i=1}^{N} w_{k(j)-1,i}\,\mathbf{d}_{j,i}, \tag{49}$$

$$\overline{\mathbf{x}}^f_{k(j)} = \sum_{i=1}^{N} w_{k(j)-1,i}\,\mathbf{x}^f_{k(j),i}, \tag{50}$$

$$\mathbf{P}_{xd} = \frac{1}{1-\sum_{i=1}^{N} w^2_{k(j)-1,i}}\sum_{i=1}^{N} w_{k(j)-1,i}\left(\mathbf{x}^f_{k(j),i}-\overline{\mathbf{x}}^f_{k(j)}\right)(\mathbf{d}_{j,i}-\overline{\mathbf{d}}_j)^T, \tag{51}$$

$$\mathbf{P}_{dd} = \frac{1}{1-\sum_{i=1}^{N} w^2_{k(j)-1,i}}\sum_{i=1}^{N} w_{k(j)-1,i}\left(\mathbf{d}_{j,i}-\overline{\mathbf{d}}_j\right)(\mathbf{d}_{j,i}-\overline{\mathbf{d}}_j)^T, \tag{52}$$

$$\mathbf{K}_{k(j)} = \mathbf{P}_{xd}\mathbf{P}^{-1}_{dd}, \tag{53}$$

$$\mathbf{x}^a_{k(j),i} = \mathbf{x}^f_{k(j),i} + \mathbf{K}_{k(j)}\left(\mathbf{d}_j - \mathbf{h}_k(\mathbf{x}^f_{k(j),i}, \boldsymbol{\epsilon}_{j,i})\right) \tag{54}$$

$$H_a = N^{-1/(n+4)}\left[\left(\frac{1}{N-1}\right)\sum_{l=1}^{N}\left(\mathbf{x}^a_{k(j),l}-\overline{\mathbf{x}}^a_{k(j)}\right)\left(\mathbf{x}^a_{k(j),l}-\overline{\mathbf{x}}^a_{k(j)}\right)^T\right]^{1/2} \tag{55}$$

$$H_f = \left(\frac{1}{\sum_{i=1}^{N}(w_{k(j)-1,i})^2}\right)^{-1/(n+4)} \times\left[\left(\frac{1}{1-\sum_{i=1}^{N} w^2_{k(j)-1,i}}\right)\sum_{l=1}^{N} w_{k(j)-1,l}\left(\mathbf{x}^f_{k(j),l}-\overline{\mathbf{x}}^f_{k(j)}\right)\left(\mathbf{x}^f_{k(j),l}-\overline{\mathbf{x}}^f_{k(j)}\right)^T\right]^{1/2} \tag{56}$$

$$\tilde{w}_{k(j),i} = w_{k(j)-1,i}\,p\left(\mathbf{d}_j|\mathbf{x}^a_{k(j),i}\right)\frac{\sum_{l=1}^{N} w_{k(j)-1,l}\frac{1}{|H_f|}K\left(H_f^{-1}\left(\mathbf{x}^f_{k(j),l}-\mathbf{x}^a_{k(j),i}\right)\right)}{\sum_{l=1}^{N}\frac{1}{N}\frac{1}{|H_a|}K\left(H_a^{-1}\left(\mathbf{x}^a_{k(j),l}-\mathbf{x}^a_{k(j),i}\right)\right)} \tag{57}$$

$$w_{k(j),i} = \frac{\tilde{w}_{k(j),i}}{\sum_{i=1}^{N}\tilde{w}_{k(j),i}}. \tag{58}$$

(3) Forecast step:

$$\mathbf{x}^f_{k+1,i} = \mathbf{g}_k(\mathbf{x}^a_{k,i}, \mathbf{f}_k, \mathbf{q}_{k,i}). \tag{59}$$

5. Noisy oscillation of a Duffing system

In this section, we perform several numerical experiments for joint state and parameter estimation of a Duffing oscillator [54,55] excited by combined harmonic and random input. For the parameter set chosen in the numerical experiments, the system exhibits a period-two subharmonic oscillation in the absence of random input [55]. EnKF, PF and MCMC sampling technique are then used to estimate the stiffness parameters in the presence of both model and measurement noise as described in detail next.

5.1. Duffing oscillator model

This section illustrates the usefulness of the aforementioned theoretical formulations using a Duffing oscillator model. The Duffing system forced by combined harmonic and random inputs is described by the following nonlinear differential equation (e.g. [54,55]):

$$m\ddot{u}(t)+c\dot{u}(t)+k_1u(t)+k_3u^3(t)=\mathcal{T}\ \cos(\omega t)+\sigma\xi(t) \tag{60}$$

with m [kg] being the mass, c [N s m^{-1}] is the damping coefficient, k_1 [N m^{-1}] and k_3 [N m^{-3}] are the linear and cubic stiffness coefficients, respectively, $u(t)$ [m] is the displacement, $\mathcal{T}$ [N] and ω [rad s^{-1}] denote the amplitude and frequency of the harmonic excitation, respectively, $\xi(t)$ represents a Gaussian white noise random input (additive modelling error) and σ[N] represents its strength. The state-space model of Eq. (60) becomes

$$\dot{x}_1=x_2 \tag{61}$$

$$\dot{x}_2=-\frac{1}{m}[cx_2+k_1x_1+k_3x_1^3]+\frac{1}{m}\mathcal{T}\ \cos(\omega t)+\frac{1}{m}\sigma\xi(t), \tag{62}$$

where $x_1=u$ and $x_2=\dot{u}$.

We are interested in estimating the stiffness coefficients k_1 and k_3 and the damping coefficient c. For joint state and parameter estimation using EnKF and PF, the original state vector $\mathbf{x}=\{x_1,x_2\}^T$ is augmented with the coefficients c, k_1, and k_3 as three new state variables $x_3=c$, $x_4=k_1$ and $x_5=k_3$ which satisfy the following equations:

$$\dot{x}_3=0, \tag{63}$$

$$\dot{x}_4=0, \tag{64}$$

$$\dot{x}_5=0, \tag{65}$$

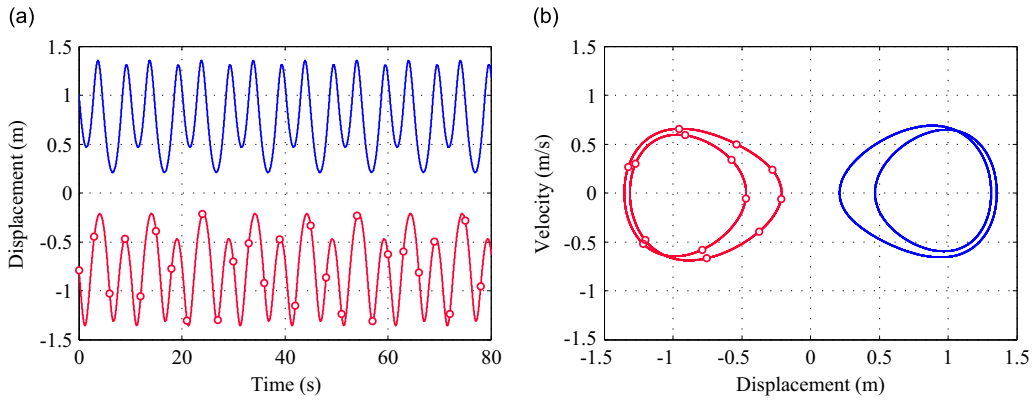

Fig. 1. Response of the Duffing oscillator under purely deterministic input with $\mathcal{T}=0.3$ N, $\sigma=0$ N: (a) two steady-state trajectories, (b) the associated phase-space diagrams (adapted from [16]).

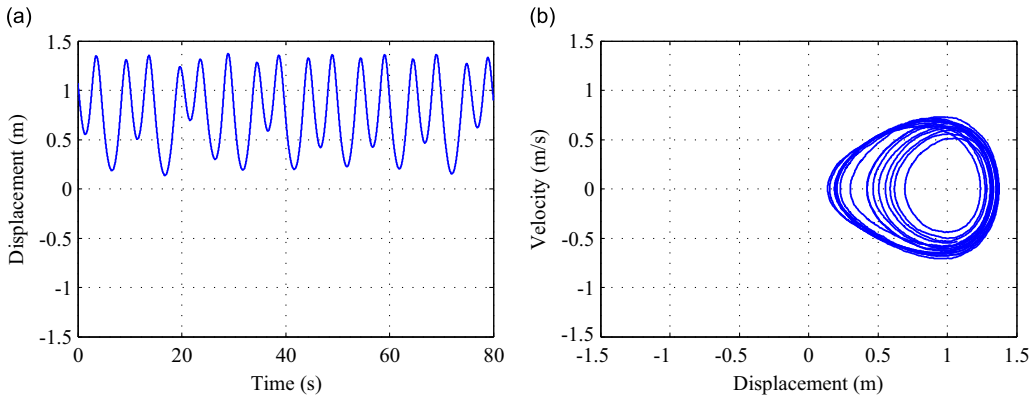

Fig. 2. Response of the Duffing oscillator under combined deterministic and random input with $\mathcal{T}=0.3$ N, $\sigma_1=0.015$ N: (a) steady-state trajectory, (b) the associated phase-space diagram.

i.e. they are modelled as time-invariant quantities. In a previous investigation by the authors [15], the unknown system parameters were modelled as Wiener processes resulting in an artificial inflation the variance of the estimates in order to avoid filter divergence (e.g. [79]). The need for this artificial perturbation was to provide a range of possible parameter values and select the best value(s) based on the partial observations of the system state. The use of a good proposal obtained using an EnKF analysis step circumvents the need to artificially inflate the variances of the unknown parameters for filter convergence and is thus avoided in this investigation.

The above set of equations have the following Ito Stochastic Differential Equation (SDE) representation

$$\mathrm{d}x_1 = x_2\,\mathrm{d}t \tag{66}$$

$$\mathrm{d}x_2 = -\frac{1}{m}\left[x_3x_2 + x_4x_1 + x_5x_1^3 - \mathcal{T}\,\cos(\omega t)\right]\mathrm{d}t + \frac{1}{m}\sigma\xi(t)\,\mathrm{d}t. \tag{67}$$

$$\mathrm{d}x_3 = 0 \tag{68}$$

$$\mathrm{d}x_4 = 0 \tag{69}$$

$$\mathrm{d}x_5 = 0 \tag{70}$$

where $\xi(t)\,\mathrm{d}t = \mathrm{d}W = W(t_{k+1}) - W(t_k)$ denotes a Brownian path increment.

The discretization of the above Ito SDEs using the Euler–Maruyama scheme [80–82] with time step Δt leads to

$$\{x_1\}_{k+1} = \{x_1\}_k + \Delta t\{x_2\}_k \tag{71}$$

$$\{x_2\}_{k+1} = \{x_2\}_k - \frac{1}{m}\Delta t\left[\{x_3\}_k\{x_2\}_k + \{x_4\}_k\{x_1\}_k + \{x_5\}_k\{x_1\}_k^3 - \mathcal{T}\cos(\omega t_k)\right] + \frac{1}{m}\sigma\sqrt{\Delta t}\varepsilon_k \tag{72}$$

$$\{x_3\}_{k+1} = \{x_3\}_k \tag{73}$$

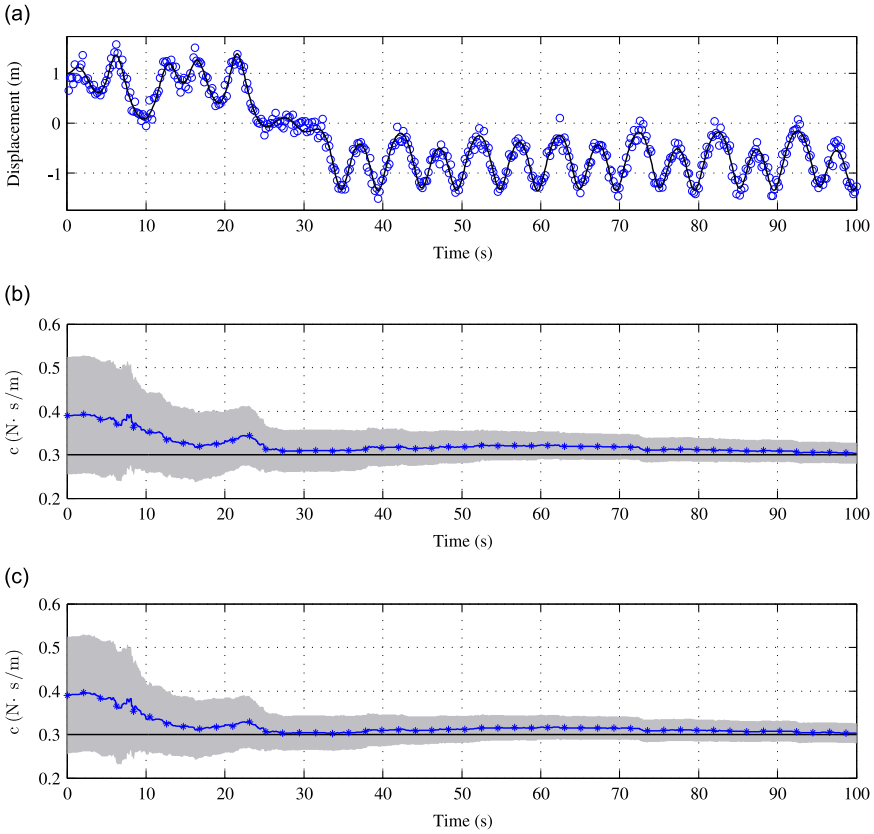

Fig. 3. c parameter estimates: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

$$\{x_4\}_{k+1} = \{x_4\}_k \tag{74}$$

$$\{x_5\}_{k+1} = \{x_5\}_k \tag{75}$$

where the variables $\{\varepsilon_k\}$ denote independent and identically distributed unit standard Gaussian random variables.

5.2. Noisy oscillation of the Duffing oscillator

The following numerical values are chosen [54,55,83]: $m=1$ kg, $c=0.3$ N s m^{-1}, $k_1=-1$ N m^{-1}, $k_3=1$ N m^{-3}, $\omega=1.25$ rad s^{-1} and $\Delta t=5\times10^{-3}$ s. The unforced system has three fixed points: an unstable fixed point at $u=0$ m and two stable fixed points at ±1 m. Fig. 1 shows two period-two subharmonic oscillation with period $4\pi/\omega$ under purely deterministic loading ($\sigma=0$ N) for $\mathcal{T}=0.3$ N [55,83]. Each steady-state trajectory presented in the figure arises from two different initial conditions.

For modelling noise with $\sigma=0.015$ N, a sample of the steady-state noisy oscillation is shown in Fig. 2. Note that the quasi-periodic behaviour of the phase-space curve arises due to the modelling noise.

5.3. Parameter estimation

In this section, we address the combined state and parameter estimation problem of the Duffing system using EnKF and PF. We also use MCMC sampling technique to estimate the parameters with EnKF providing the state estimates. In particular, we are interested in estimating the damping parameter c and stiffness parameters k_1 and k_3. For simplicity, although the strength of model noise σ is assumed to be known in this investigation, it can also be estimated along with the system parameters as described in [51].

We assume that the observational data is modelled as

$$d_k = u_k + \epsilon_k. \tag{76}$$

For this experiment, the measurement error is assumed to be Gaussian: $\epsilon_k \sim \mathcal{N}(0, 1.3\times10^{-2})$. The measurement arrives at an interval of 0.2 time unit. The measurement error standard deviation is assumed to be 15 percent of the

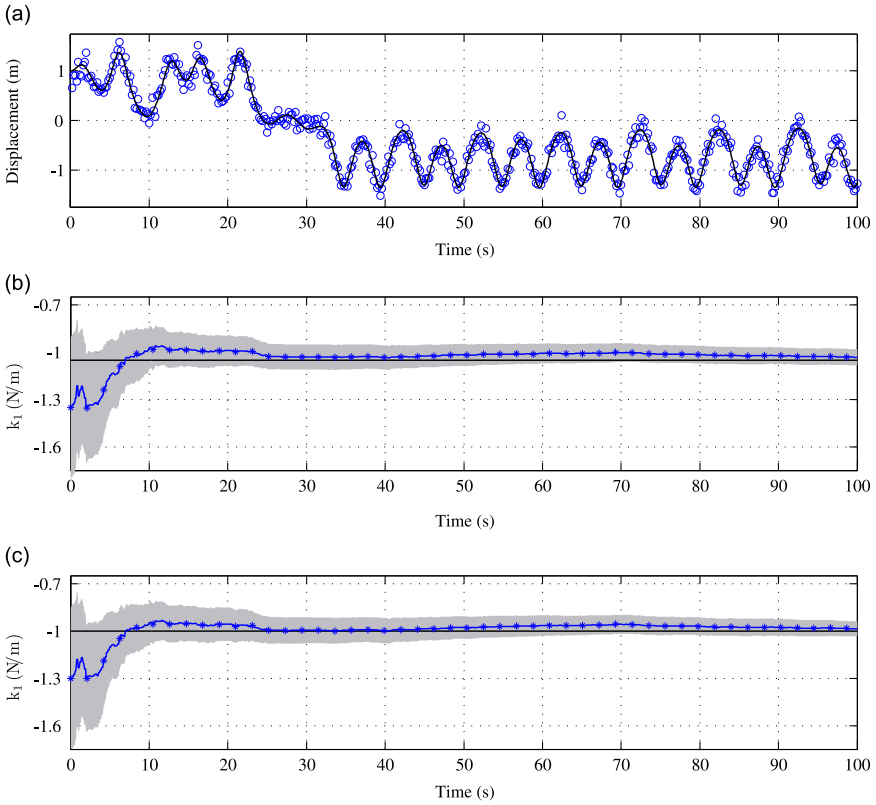

Fig. 4. k_1 parameter estimates: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ±3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ±3 error standard deviations (shaded area).

root-mean-square (RMS) value of the true displacement of the oscillator. The modelling error amplitude is chosen to be $\sigma = 0.015$ N. Fig. 3a shows the true displacement and the measurement d_k. For this parameter estimation experiment, the initial conditions are $u \sim \mathcal{N}(1,0)$, $\dot{u} \sim \mathcal{N}(0,0)$, $c \sim \mathcal{N}(0.39, 0.002025)$, $k_1 \sim \mathcal{N}(-1.3, 0.0225)$, $k_3 \sim \mathcal{N}(1.3, 0.0225)$. An ensemble size of $N = 25{,}000$ is used for EnKF and PF. Furthermore, 450,000 MCMC samples from the joint posterior parameter pdf are obtained with 5000 EnKF samples used for each MCMC point. The estimates of the pdfs with 450,000 MCMC samples are smooth. The resampling step is performed in PF at the threshold value $N_{\text{thr}} = 0.5$ N. From numerical investigations, it turns out that this threshold value yields the best performance for PF. To minimize sampling errors, an efficient sampling scheme

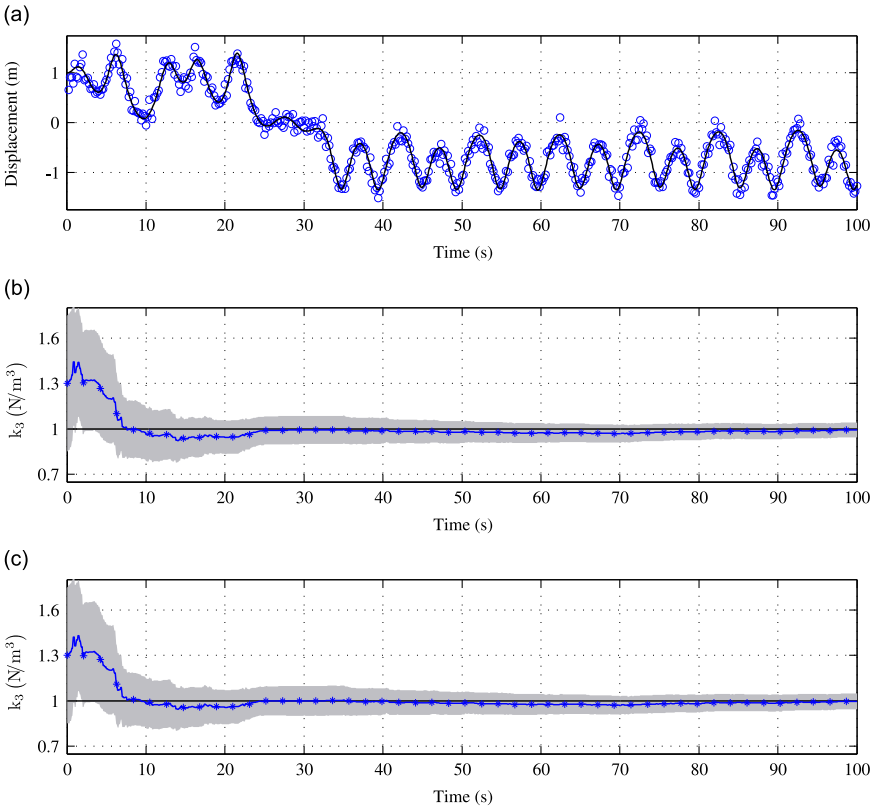

Fig. 5. k_3 parameter estimates: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

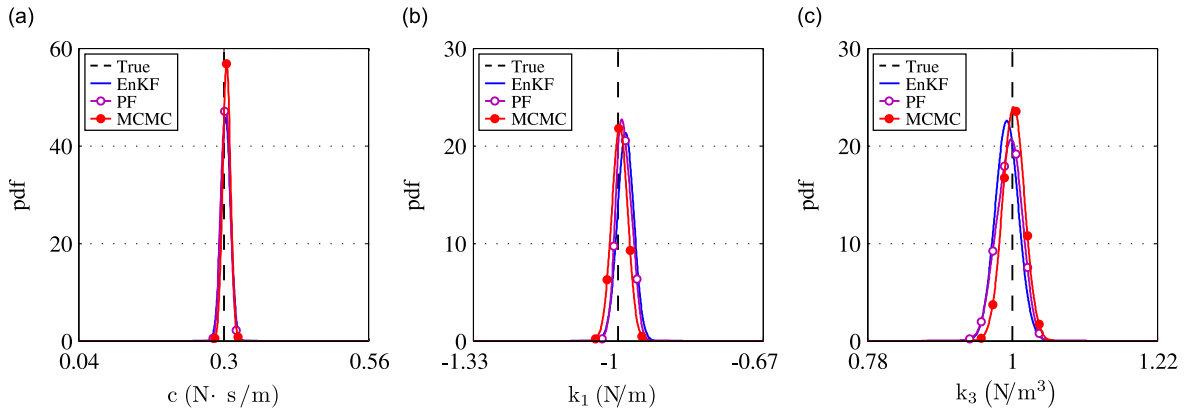

Fig. 6. Posterior marginal pdfs of parameters: (a) damping coefficient c, (b) linear stiffness coefficient k_1, (c) nonlinear stiffness coefficient k_3. Dashed vertical line indicates the true parameter value. Solid curve represents the EnKF estimate. Solid curve with hollow circular markers represents the PF estimate. Solid curve with solid circular markers represents the MCMC estimate.

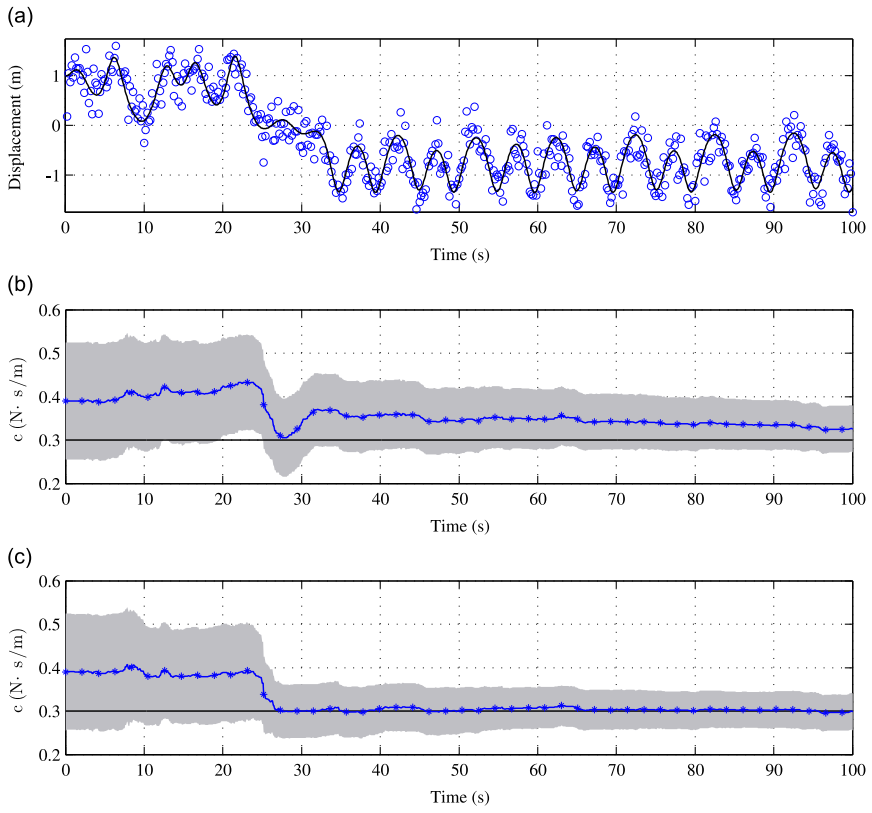

Fig. 7. c parameter estimates with stronger measurement noise: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

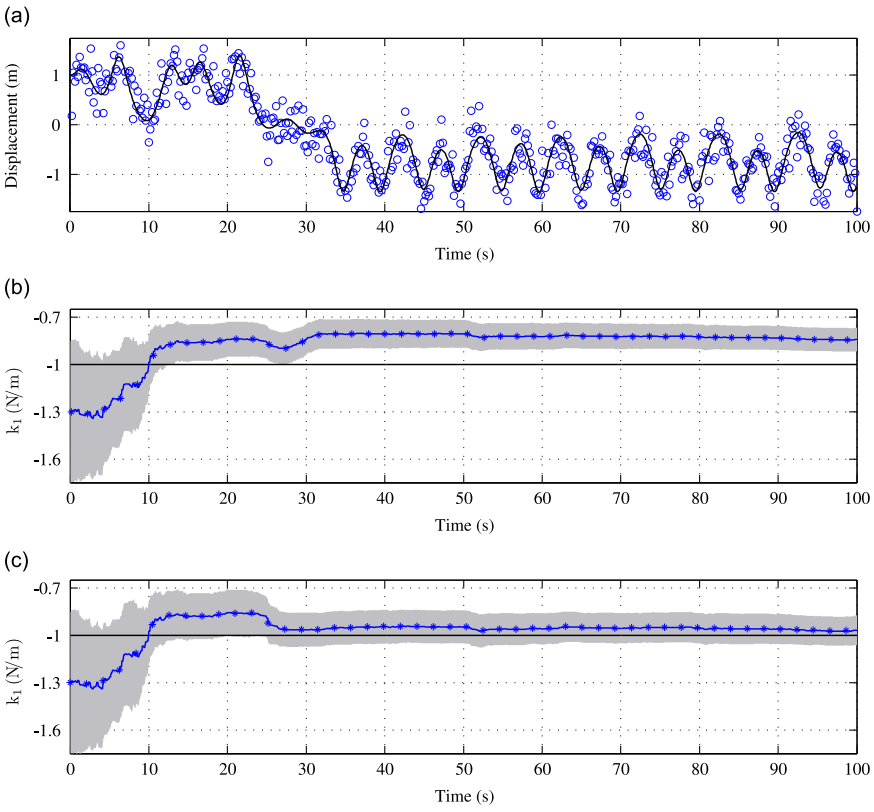

Fig. 8. k_1 parameter estimates with stronger measurement noise: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

based on LHS [35,36] is adopted for EnKF and PF. The typical execution times for the algorithms running on a Linux cluster with 22 nodes, each node having 2 Quad-core 3.0 GHz Intel Xeon processors and 32 GB of memory, are as follows: (a) EnKF with 25,000 samples: 2 min; (b) PF having EnKF proposal with 25,000 samples: 8 h; and (c) MCMC (having EnKF used for the likelihood computation) with 450,000 samples using 150 parallel chains: 2 h.

For the joint state and parameter estimation experiments, Figs. 3–5 show the estimates of the damping and stiffness coefficients c, k_1, and k_3 using EnKF in subplot (b) and PF in subplot (c). The marginal pdfs of the parameters at the final time step of simulation are provided in Fig. 6. The marginal pdfs of the parameters using MCMC sampling are also provided in Fig. 6. All three methods yield parameter estimates of similar accuracy.

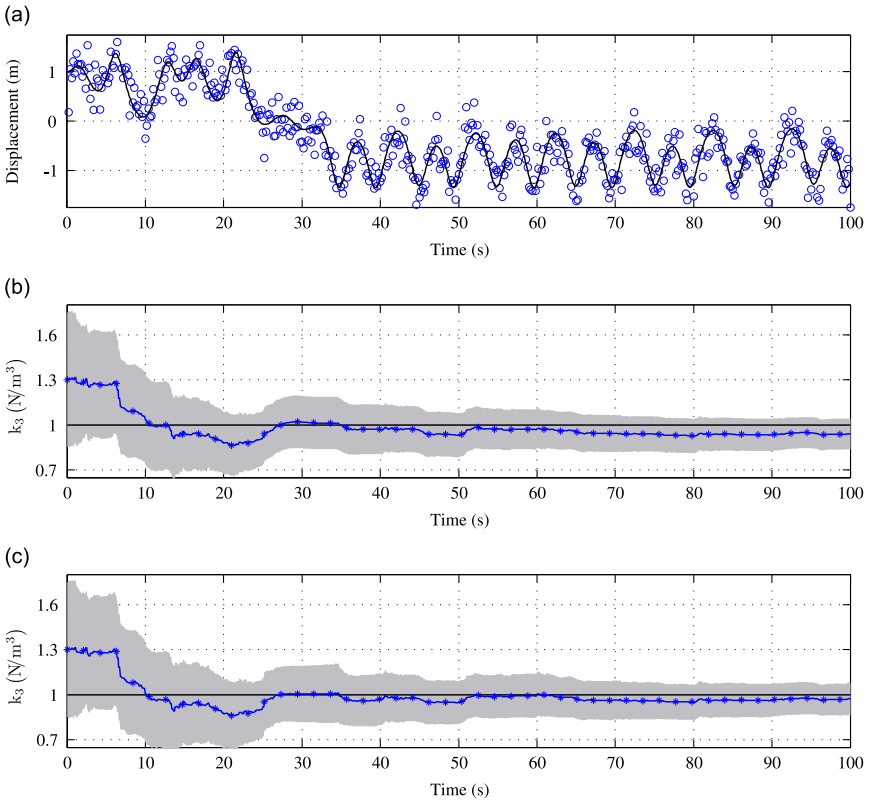

Fig. 9. k_3 parameter estimates with stronger measurement noise: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

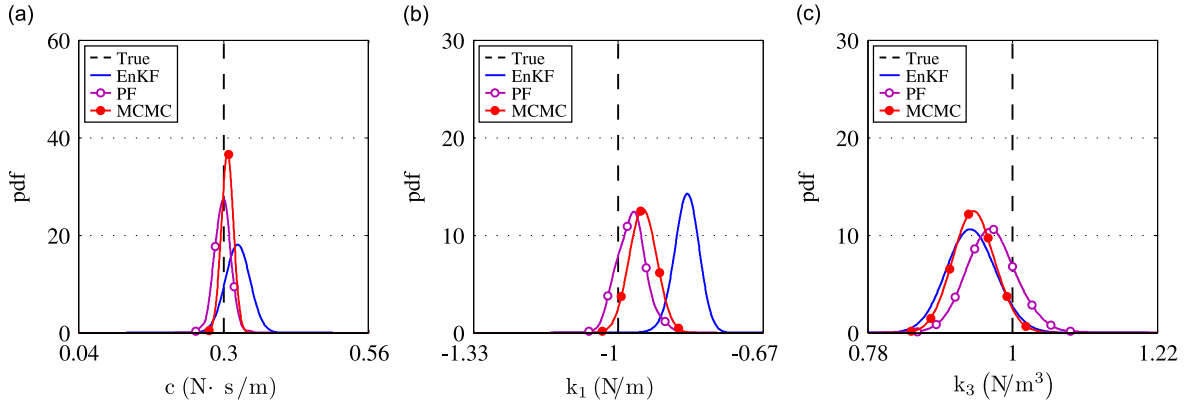

Fig. 10. Posterior marginal pdfs of parameters with stronger measurement noise: (a) damping coefficient c, (b) linear stiffness coefficient k_1, (c) nonlinear stiffness coefficient k_3. Dashed vertical line indicates the true parameter value. Solid curve represents the EnKF estimate. Solid curve with hollow circular markers represents the PF estimate. Solid curve with solid circular markers represents the MCMC estimate.

5.3.1. Effect of measurement noise

In this experiment, we study the effect of measurement noise intensity on the filter estimates. The measurement noise is now taken to be Gaussian given by $\epsilon_k \sim \mathcal{N}(0, 7.1 \times 10^{-2})$. In contrast to the previous experiment, the standard deviation of the measurement error is now assumed to be 35 percent of the RMS value of the true displacement. Similar to the previous experiment, the time interval between the observational data is taken to be 0.2 time unit. Fig. 7a shows the true displacement of the system and the associated synthetic observational data d_k.

For this joint state and parameter estimation experiment, the same initial conditions, ensemble size (for PF and EnKF) and number of MCMC samples are used as in the previous experiment.

Figs. 7–9 show the estimates of the damping and stiffness coefficients and Fig. 10 provides the posterior marginal pdfs once all the data has been assimilated. Evidently, the parameter estimates obtained using PF and MCMC are equivalent in accuracy. EnKF was only able to estimate the nonlinear stiffness coefficient k_3 with reasonable accuracy. The superior performance of PF over EnKF is attributed to its ability to handle strongly non-Gaussian systems. Although MCMC sampling techniques relies on EnKF for state estimation, it still provides accurate estimates. In contrast to the extended state vector in joint state and parameter estimation, the state vector in pure state estimation carries less non-Gaussian characteristics. Thus EnKF is able to provide accurate state estimates for MCMC sampling, but encounters difficulties in joint state and parameter estimation. Of course, the error in the estimates is much larger for all three methods compared to the previous experiment due to the presence of larger measurement error.

5.3.2. Effect of observational data sparsity

In this last experiment, the effect of observational data sparsity is examined. The measurement noise is taken to be Gaussian given by $\epsilon_k \sim \mathcal{N}(0, 1.3 \times 10^{-2})$. The measurement time interval is taken to be 1.0 time unit, instead of 0.2 time unit used in the first experiment. Fig. 11a plots the true (actual) displacement of the system and the measurement d_k. Apart from the smaller sampling rate for the observations, all experimental parameters are unchanged from the first experiment (Section 5.3).

The estimates of the damping and stiffness coefficients are plotted in Figs. 11–13 and the marginal posterior pdfs in Fig. 14. Clearly, the performances of PF and MCMC are superior to EnKF which provides an accurate parameter estimate for only the nonlinear stiffness parameter. The infrequent assimilation of data allows the augmented state vector to regain its

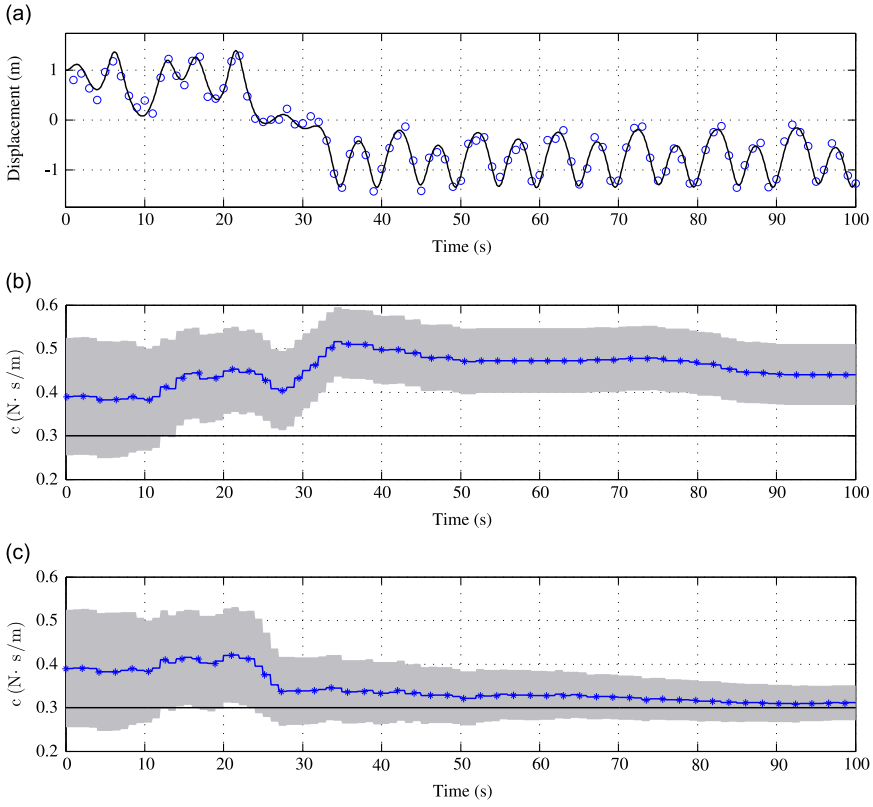

Fig. 11. c parameter estimates with sparse observational data: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

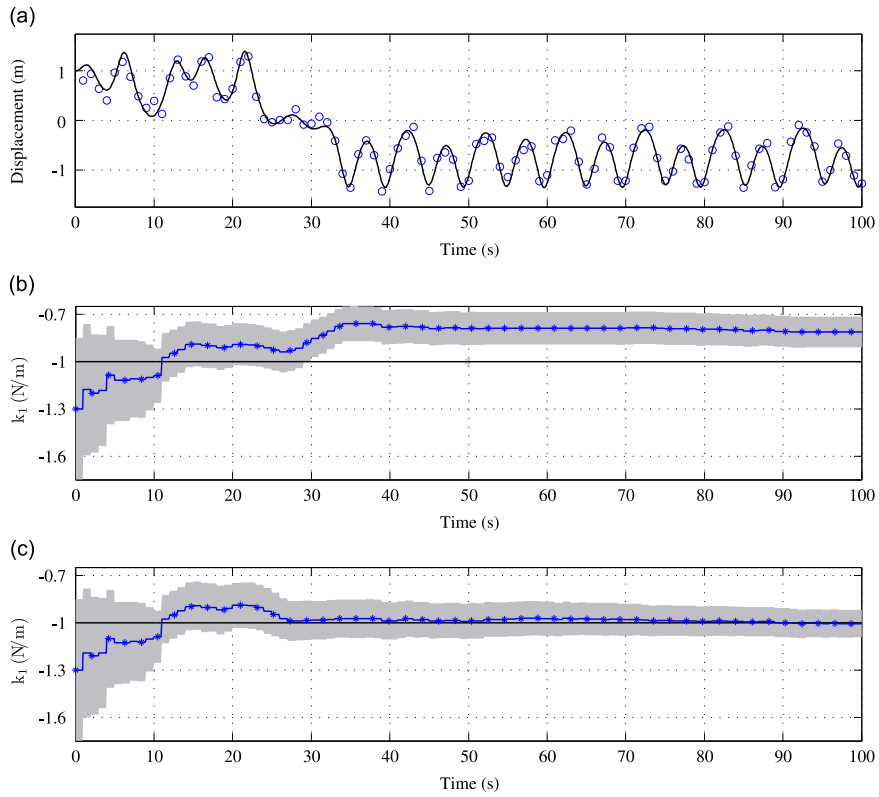

Fig. 12. k_1 parameter estimates with sparse observational data: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

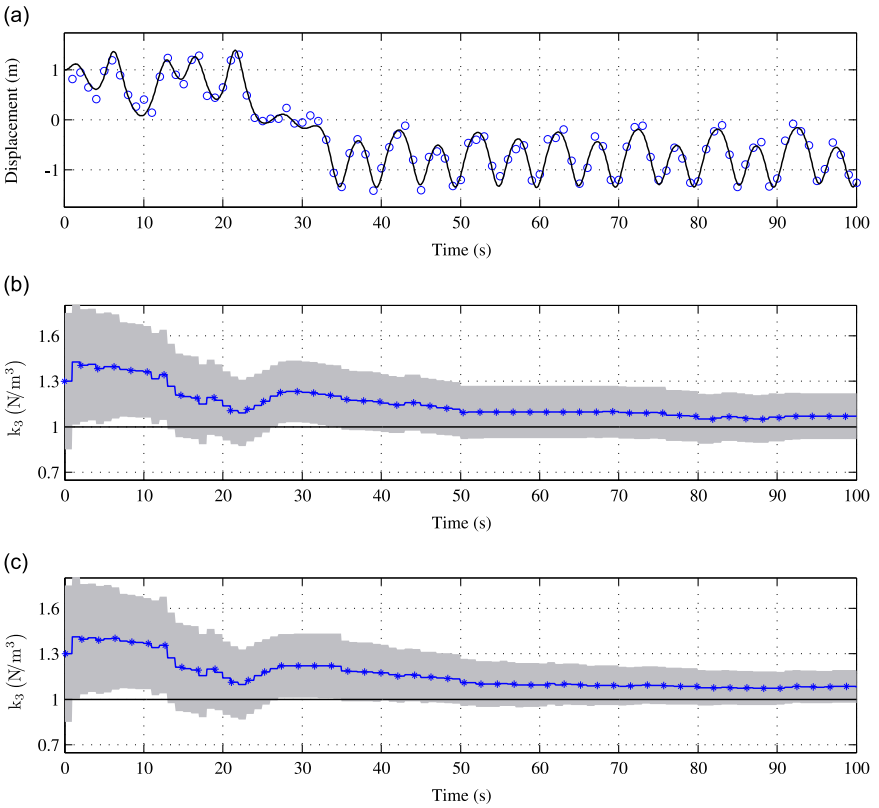

Fig. 13. k_3 parameter estimates with sparse observational data: (a) true (solid line) and measured (circles) displacement, (b) EnKF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area), (c) PF estimate (line with stars) and estimated mean ± 3 error standard deviations (shaded area).

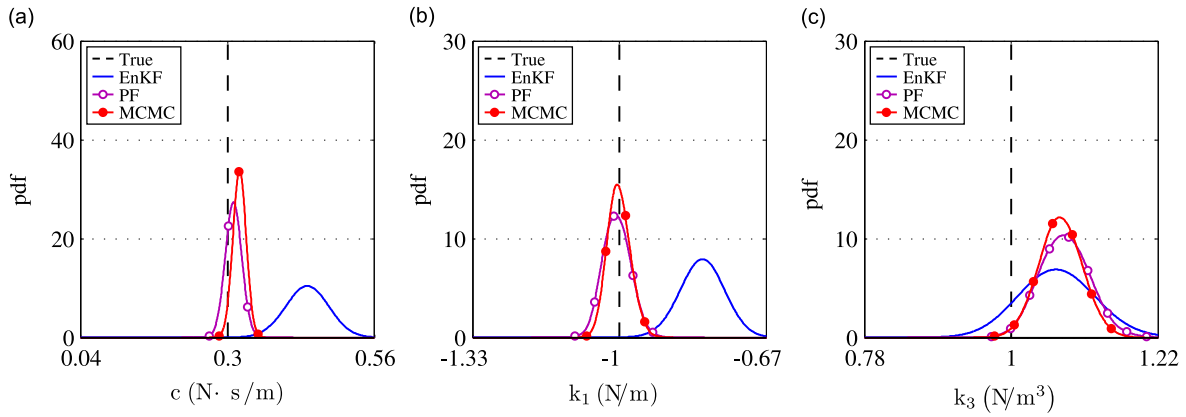

Fig. 14. Posterior marginal pdfs of parameters with sparse observational data: (a) damping coefficient c, (b) linear stiffness coefficient k_1, (c) nonlinear stiffness coefficient k_3. Dashed vertical line indicates the true parameter value. Solid curve represents the EnKF estimate. Solid curve with hollow circular markers represents the PF estimate. Solid curve with solid circular markers represents the MCMC estimate.

non-Gaussian features due to the nonlinear dynamics, leading to a severe degradation in the performance of EnKF for combined state and parameter estimation.

6. Conclusion

In the framework of nonlinear filtering, the combined state and time-invariant parameter estimation of a dynamical system involves appending the unknown parameters to the state vector, leading to an augmented state space model. The nonlinearity in the augmented system is generally stronger than the original state space model. The weakly non-Gaussian filter, such as EnKF, may adequately handle the state estimation of the original state space model, but it fails to tackle the combined state and parameter estimation problem using the augmented state space model. This fact is exploited in this paper in order to use the MCMC simulation for the time-invariant parameter estimation of nonlinear oscillatory systems. It is also pointed out that the combined state and parameter estimation approach using EnKF may become highly inaccurate for the time-invariant parameter estimates. It is demonstrated that the MCMC simulation, complemented by EnKF for just the state estimation, provides accurate parameter estimates for the Duffing system considered for numerical investigations. In order to achieve the same level of accuracy of the MCMC method, the PF algorithm by Mandel and Beezley has been extended for the combined state and parameter estimation. Although some of the results in this paper may be intuitively obvious, to the authors best knowledge, they are not widely available in the open literature.

In the context of time-invariant parameter estimation, this paper demonstrates the difficulties encountered in the combined state and parameter estimation using nonlinear filtering as the modified state space model constructed by augmenting the state by the parameter vector introduces stronger nonlinearities compared to the original state space model. In order to alleviate these difficulties in estimating time-invariant parameters, the paper then provides basic mathematical expositions of a two-level (nested and combined) MCMC and EnKF algorithm; and a joint state and parameter estimation approach using a modified PF based on the proposal density obtained from EnKF. At this initial stage, the primary objective of this investigation is focussed on highlighting the benefits and limitations of the MCMC, EnKF and PF algorithms from statistical perspectives (i.e. effects of measurement noise and observational data sparsity) and hence the use of only *noisy* period-two subharmonic oscillations of a Duffing oscillator for numerical experiments. Further research is needed to assess the usefulness of these parameter estimation algorithms for a broad range of dynamical systems. For instance, the methodologies can be applied to the parameter estimation of low dimensional dynamical systems such as the Duffing systems, exhibiting different dynamical behaviours ranging from period-one, period-four to chaotic motions. On the other hand, it will be worthwhile to investigate the performances of the methodologies for time-invariant parameter estimation of high-dimensional state space models (for instance, arising from the discretization of partial differential equations), particularly focussing on the parallel implementations of these algorithms to enhance their computational efficiency by exploiting high performance computing systems. These aspects are the subjects of current investigations.

Acknowledgements

The first author acknowledges the support of the Natural Sciences and Engineering Research Council of Canada through the award of a Canada Graduate Scholarship and the Canadian Department of National Defence. The second author acknowledges the support of a Discovery Grant from Natural Sciences and Engineering Research Council of Canada and the Canada Research Chair Program. The third author acknowledges the support of the UK Engineering and Physical Sciences

Research Council (EPSRC) through the award of an Advanced Research Fellowship and the Royal Society of London for the award of a visiting fellowship at Carleton University, Canada. The fourth author acknowledges the support of the Canadian Department of National Defence, through the DSRI-TIF program, and a Discovery Grant from Natural Sciences and Engineering Research Council of Canada. The computing infrastructure is supported by the Canada Foundation for Innovation (CFI) and the Ontario Innovation Trust (OIT).

References

[1] G.A. Vio, J.E. Cooper, Limit cycle oscillation prediction for aeroelastic systems with discrete bilinear stiffness, *International Journal of Applied Mathematics and Mechanics* 3 (2005) 100–119.
[2] D. Poirel, S.J. Price, Bifurcation characteristics of a two-dimensional structurally non-linear airfoil in turbulent flow, *Journal of Nonlinear Dynamics* 48 (4) (2007) 423–435.
[3] M.P. Païdoussis, *Fluid-Structure Interactions*: *Slender Structures and Axial Flow*, second ed. Academic Press, San Diego, CA, 1998.
[4] K. Worden, G.R. Tomlinson (Eds.), *Nonlinearity in Structural Dynamics*: Detection, Identification and Modelling, Institute of Physics Publishing, Philadelphia, 2001.
[5] P. Ibanez, Identification of dynamic parameters of linear and non-linear structural models from experimental data, *Nuclear Engineering and Design* 25 (1973) 30–41.
[6] S.F. Masri, T.K. Caughey, A nonparametric identification technique for nonlinear dynamic problems, *Journal of Applied Mechanics* 46 (1979) 433–447.
[7] G. Kerschen, K. Worden, A.F. Vakakis, J.-C. Golinval, Past, present and future of nonlinear system identification in structural dynamics, *Mechanical Systems and Signal Processing* 20 (2006) 505–592.
[8] C. Aguilar-Ibanez, J. Sanchez, M.S. Suarez, J.C. Martinez, On the algebraic reconstruction of the Duffing's mechanical system, *Physics Letters A* 372 (25) (2008) 4569–4573.
[9] M.D. Narayanan, S. Narayanan, C. Padmanabhan, Multiharmonic excitation for nonlinear system identification, *Journal of Sound and Vibration* 311 (3–5) (2008) 707–728.
[10] J.S. Bendat, *Nonlinear Systems Techniques and Applications*, John Wiley & Sons, New York, 1998.
[11] B.A. Zeldin, P.D. Spanos, Spectral identification of nonlinear structural systems, *Journal of Engineering Mechanics* 124 (7) (1998) 728–733.
[12] P.D. Spanos, R. Lu, Nonlinear system identification in offshore structural reliability, *Journal of Offshore Mechanics and Arctic Engineering* 117 (3) (1995) 171–177.
[13] I.A. Kougioumtzoglou, P.D. Spanos, An identification approach for linear and nonlinear time-variant structural systems via harmonic wavelets, *Mechanical Systems and Signal Processing* 37 (1–2) (2013) 338–352.
[14] C.S. Manohar, D. Roy, Monte Carlo filters for identification of nonlinear structural dynamical systems, *Sadhana – Academy Proceedings in Engineering Sciences* 31 (2006) 399–427. Part 4.
[15] M. Khalil, A. Sarkar, S. Adhikari, Nonlinear filters for chaotic oscillatory systems, *Journal of Nonlinear Dynamics* 55 (1–2) (2009) 113–137.
[16] M. Khalil, A. Sarkar, S. Adhikari, Tracking noisy limit cycle oscillations with nonlinear filters, *Journal of Sound and Vibration* 392 (2) (2010) 150–170.
[17] R.E. Kalman, A new approach to linear filtering and prediction problems, *Journal of Basic Engineering* 82 (1960) 35–45.
[18] J. Kaipio, E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, New York, 2005.
[19] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, San Diego, CA, 1970.
[20] G. Evensen, *Data Assimilation*: *The Ensemble Kalman Filter*, Springer, Berlin, 2006.
[21] H. Tanizaki, *Nonlinear Filters*: *Estimation and Applications*, second ed. Springer, Berlin, 1996.
[22] A. Doucet, S.J. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, *Statistics and Computing* 10 (3) (2000) 197–208.
[23] B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman Filter*: *Particle Filters for Tracking Applications*, Artech House, Boston, 2004.
[24] G. Evensen, Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research* 99 (C5) (1994) 10143–10162.
[25] R. Ghanem, G. Ferro, Health monitoring for strongly non-linear systems using the ensemble Kalman filter, *Structural Control and Health Monitoring* 13 (1) (2006) 245–259.
[26] G.A. Kivman, Sequential parameter estimation for stochastic systems, *Nonlinear Processes in Geophysics* 10 (3) (2003) 253–259.
[27] R.N. Miller, E.F. Carter, S.T. Blue, Data assimilation into nonlinear stochastic models, *Tellus Series A – Dynamic Meteorology and Oceanography* 51 (2) (1999) 167–194.
[28] J.L. Anderson, S.L. Anderson, A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review* 127 (12) (1999) 2741–2758.
[29] G. Storvik, Particle filters in state space models with the presence of unknown static parameters, *IEEE Transactions on Signal Processing* 50 (2) (2002) 281–289.
[30] J.A. Vrugt, C.J.F. ter Braak, C.G.H. Diks, G. Schoups, Hydrologic data assimilation using particle Markov Chain Monte Carlo simulation: *theory, concepts and applications*, *Advances in Water Resources* 51 (2013) 457–478.
[31] C. Andrieu, A. Doucet, R. Holenstein, Particle Markov Chain Monte Carlo methods, *Journal of Royal Statistical Society, Series B* 72 (3) (2010) 269–342.
[32] C. Andrieu, A. Doucet, Online expectation–maximization type algorithms for parameter estimation in general state space models, *Proceedings of ICASSP '03, the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, 2003.
[33] A. Doucet, N. de Freitas, N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001.
[34] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear non-Gaussian Bayesian state estimation, *IEE Proceedings – F Radar and Signal Processing* 140 (2) (1993) 107–113.
[35] M. Mckay, R. Beckman, W. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* 21 (1979) 239–245.
[36] A. Olsson, G. Sandberg, Latin hypercube sampling for stochastic finite element analysis, *Journal of Engineering Mechanics* 128 (2002) 121–125.
[37] G. Kitagawa, Non-Gaussian state-space modeling of nonstationary time series, *Journal of the American Statistical Association* 82 (400) (1987) 1032–1063.
[38] C.K. Chui, G. Chen, *Kalman Filtering with Real-time Applications*, third ed. Springer, Berlin, 1999.
[39] F. Campillo, V. Rossi, Convolution particle filter for parameter estimation in general state-space models, *IEEE Transactions on Aerospace and Electronic Systems* 45 (3) (2009) 1063–1072.
[40] C. Musso, N. Oudjane, F. LeGland, Improving regularized particle filters, in: A. Doucet, N. de Freitas, N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001, pp. 247–271.
[41] E.A. Wan, R. van der Merwe, The unscented Kalman filter, in: S. Haykin (Ed.), *Kalman Filtering and Neural Networks*, Wiley, New York, 2001. (Chapter 7).
[42] J. Mandel, J.D. Beezley, Predictor–corrector and morphing ensemble filters for the assimilation of sparse data into high-dimensional nonlinear systems, *Proceedings of the IOAS-AOLS, 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, San Antonio, TX, USA, 2007, pp. 4–12.
[43] J. Mandel, J.D. Beezley, An ensemble Kalman-particle predictor–corrector filter for non-Gaussian data assimilation, *ICCS 2009, Lecture Notes in Computer Science*, Vol. 5545, Springer, New York 2009, pp. 470–478.
[44] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, New York, 1986.

[45] R.H. Cameron, W.T. Martin, The orthogonal development of non-linear functionals in series of Fourier–Hermite functionals, *Annals of Mathematics, Second Series* 48 (2) (1947) 385–392.
[46] H.H. Kuo, Gaussian measures in Banach spaces, A. Dold, B. Eckmann (Eds.), *Lecture Notes in Mathematics*, Vol. 463, Springer-Verlag, New York, 1975, pp. 1–109.
[47] W. Häardie, M. Müller, Multivariate and semiparametric kernel regression, in: M.G. Schimek (Ed.), *Smoothing and Regression: Approaches, Computation, and Application*, John Wiley & Sons Inc., New York, 2000, pp. 357–391.
[48] W.K. Härdle, L. Simar, *Applied Multivariate Statistical Analysis*, Springer, New York, 2012.
[49] Mohammad Khalil, Abhijit Sarkar, Sondipon Adhikari, Ensemble Kalman and particle filter for noise-driven oscillatory systems, *Non-Deterministic Approaches Forum*, *49th AIAA/ASME/ASCE/AHS/ASC Structures*, *Structural Dynamics, and Materials Conference*, Proceeding Paper no: AIAA-2008-1989, Schaumburg, IL, USA, 2008.
[50] I. Mbalawata, S. Särkkä, H. Haario, Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering, *Computational Statistics* (2012) 1–29.
[51] M. Khalil, D. Poirel, A. Sarkar, Parameter estimation of a fluttering aeroelastic system in the transitional Reynolds number regime, *Journal of Sound and Vibration* 332 (2013) 3670–3691.
[52] M. Khalil, D. Poirel, A. Sarkar, Parameter estimation of a fluttering aeroelastic system in the transitional Reynolds number regime, *Proceedings of IFASD 2011*, *the International Forum on Aeroelasticity and Structural Dynamics*, Paris, France, 2011.
[53] M. Khalil, D. Poirel, A. Sarkar, Parameter estimation of a fluttering aeroelastic system in the transitional Reynolds number regime, *Proceedings of the ASME 3rd US-European Fluids Engineering Summer Meeting*, Montreal, Canada, 2010.
[54] J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Field*, Springer-Verlag, New York, 1983.
[55] S. Lynch, *Dynamical Systems with Applications using MATLAB*, Birkhäuser, Boston, 2004.
[56] S. Haykin (Ed.), *Kalman Filtering and Neural Networks*, Wiley, New York, 2001.
[57] M. Khalil, Bayesian Inference for Complex and Large-Scale Engineering Systems, PhD Thesis, Carleton University, Ottawa, ON, Canada, 2013.
[58] W.R. Gilks, S. Richardson, D.J. Speigelhalter (Eds.), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, New York, 1996.
[59] J.S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer-Verlag, New York, 2001.
[60] R. Sandhu, M. Khalil, A. Poirel, D. Sarkar, Model selection methods for nonlinear aeroelastic systems using wind-tunnel data, *54th AIAA Non-Deterministic Approaches Conference*, 2013.
[61] P. Bisaillon, R. Sandhu, M. Khalil, D. Poirel, A. Sarkar, Model selection for strongly nonlinear systems, *54th AIAA/ASME/ASCE/AHS/ASC Structures*, *Structural Dynamics and Materials Conference*, Boston, MA, 2013.
[62] R. Sandhu, M. Khalil, A. Sarkar, D. Poirel, Bayesian model selection for nonlinear aeroelastic systems using wind-tunnel data, *Computer Methods in Applied Mechanics and Engineering* 282 (2014) 161–183.
[63] R. van der Merwe, Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models, PhD Thesis, Oregon Health & Science University, Portland, OR, USA, 2004.
[64] G. Evensen, The ensemble Kalman filter: *theoretical formulation and practical implementation*, *Ocean Dynamics* 53 (4) (2003) 343–367.
[65] N. Kantas, A. Doucet, S.S. Singh, J.M. Maciejowski, Overview of sequential Monte Carlo methods for parameter estimation on general state space models, *Proceedings of the 15th IFAC Symposium on System Identification* (SYSID), Saint-Malo, France, 2009.
[66] N. Papadakis, E. Memin, A. Cuzol, N. Gengembre, Data assimilation with the weighted ensemble Kalman filter, *Tellus – Series A: Dynamic Meteorology and Oceanography* 62 (5) (2010) 673–697.
[67] W.R. Gilks, S. Richardson, D.J. Speigelhalter (Eds.), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, 1996.
[68] D. Gamerman, *Markov Chain Monte Carlo*: *Stochastic Simulation for Bayesian Inference*, Chapman & Hall, London, 1997.
[69] A. Doucet, On Sequential Simulation-based Methods for Bayesian Filtering, Technical Report CUED/FINFENG/TR 310, Department of Engineering, Cambridge University, 1998.
[70] M.K. Pitt, N. Shephard, Filtering via simulation: *auxiliary particle filters*, *Journal of the American Statistical Association* 94 (446) (1999) 590–599.
[71] R. Douc, O. Cappe, E. Moulines, Comparison of resampling schemes for particle filtering, *Proceedings of ISPA 2005*, *the 4th International Symposium on Image and Signal Processing and Analysis*, Zagreb, Croatia, 2005, pp. 64–69.
[72] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of Computational and Graphical Statistics* 5 (1996) 1–25.
[73] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamic systems, *Journal of the American Statistical Association* 93 (1998) 1032–1044.
[74] A. Kong, J.S. Liu, W.H. Wong, Sequential imputations and Bayesian missing data problems, *Journal of the American Statistical Association* 89 (1994) 278–288.
[75] A. Doucet, N. de Freitas, K. P. Murphy, S. J. Russell, Rao–Blackwellised particle filtering for dynamic Bayesian networks, *Proceedings of the Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, Stanford, CA, USA, 2000.
[76] R. van der Merwe, E. Wan, The square-root unscented Kalman filter for state and parameter estimation, *Proceedings of ICASSP '01*, *the IEEE International Conference on Acoustics*, *Speech*, *and Signal Processing*, Salt Lake City, UT, USA, 2001.
[77] M. Bocquet, C.A. Pires, L. Wu, Beyond Gaussian statistical modeling in geophysical data assimilation, *Monthly Weather Review* 138 (2010) 2997–3023.
[78] D.W. Scott, *Multivariate Density Estimation*: *Theory Practice and Visualization*, John Wiley & Sons, New York, 1992.
[79] E.M. Constantinescu, A. Sandu, T. Chaib, G.R. Carmichael, Ensemble-based chemical data assimilation. I: *general approach*, *Quarterly Journal of the Royal Meteorological Society* 133 (626) (2007) 1229–1243.
[80] G. Maruyama, Continuous Markov processes and stochastic equations, *Rendiconti del Circolo Matematico di Palermo* 4 (1955) 48–90.
[81] P.E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, New York, 1999.
[82] D.J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations, *SIAM Review, Education Section* 43 (2001) 525–546.
[83] Stephen Lynch, *Dynamical Systems with Applications using Mathematica*, Springer, New York, 2007.