Nonlinear Dyn (2009) 55: 113–137
DOI 10.1007/s11071-008-9349-z

ORIGINAL PAPER

Nonlinear filters for chaotic oscillatory systems

M. Khalil · A. Sarkar · S. Adhikari

Received: 5 October 2007 / Accepted: 7 March 2008 / Published online: 28 March 2008
© Springer Science+Business Media B.V. 2008

Abstract This paper examines and contrasts the feasibility of joint state and parameter estimation of noise-driven chaotic systems using the extended Kalman filter (EKF), ensemble Kalman filter (EnKF), and particle filter (PF). In particular, we consider the chaotic vibration of a noisy Duffing oscillator perturbed by combined harmonic and random inputs ensuing a transition probability density function (pdf) of motion which displays strongly non-Gaussian features. This system offers computational simplicity while exhibiting a kaleidoscope of dynamical behavior with a slight change of input and system parameters. An extensive numerical study is undertaken to contrast the performance of various nonlinear filtering algorithms with respect to sparsity of observational data and strength of model and measurement noise. In general, the performance of EnKF is better than PF for smaller ensemble size, while for larger ensembles PF outperforms EnKF. For moderate measurement noise and frequent measurement data, EKF is able to correctly track the dynamics of the system. However, EKF performance is unsatisfactory in the presence of sparse observational data or strong measurement noise.

Keywords Chaos · Duffing oscillator · Ensemble Kalman filter · Extended Kalman filter · Latin hypercube sampling · Non-Gaussian transition probability density function · Nonlinear dynamics · Nonlinear filters · Parameter and state estimation · Particle filter

Abbreviations

EKF	extended Kalman filter
KF	Kalman filter
EnKF	ensemble Kalman filter
LHS	Latin hypercube sampling
pdf	probability density function
pde	partial differential equation
PF	particle filter
UKF	unscented Kalman filter

1 Introduction

A wide variety of methods have been developed to deal with the joint state and parameter estimation of dynamical systems in the data assimilation research community [1–3]. Kalman filter (KF) has become a popular tool for state estimation problems in linear systems (i.e., linear model and measurement operators) due to its mathematical simplicity [4, 5]. KF results in a recursive analytical solution of the *posterior* distribution of the system state conditional upon

M. Khalil · A. Sarkar (✉)
Department of Civil and Environmental Engineering, Carleton University, Mackenzie Building, Colonel by Drive, Ottawa, Ontario K1S 5B6, Canada
e-mail: abhijit_sarkar@carleton.ca

S. Adhikari
School of Engineering, University of Wales, Singleton Park, Swansea SA2 8PP, UK

measurement data. The filter optimally estimates the state of the system in the presence of additive Gaussian model and measurement noise. If both model and observation noise are zero-mean and uncorrelated, KF is the *best linear unbiased estimator* [1]. In addition, if the errors are also Gaussian, KF is also the *maximum a posteriori estimator* or the *maximum likelihood estimator* [1, 2].

In the framework of filtering theory applied to parameter estimation, one or more unknown system parameters are appended to the original state vector [1, 2, 6]. In general, the augmented state vector evolves nonlinearly regardless of whether the original system model is linear or nonlinear. This fact precludes the direct application of KF. Under certain conditions, however, KF can be extended to the nonlinear case by linearizing the model and measurement operators around the current estimate of the state vector, leading to the popular (but no longer optimal) extended Kalman filter (EKF) [1, 7]. EKF has been widely applied to parameter identification problems (e.g., [8–13]). The major limitation of KF and EKF stems from the fact that the state evolution is assumed to be Gaussian or approximately Gaussian, permitting a statistical closure at the second order moments. Evensen [14] and Gauthier et al. [15] showed that the Gaussian assumption may produce instabilities and even divergence, when applied to strongly nonlinear systems. To alleviate such problem, the unscented transform (UT) [16] is used to propagate the statistical moments of the state vector through the full nonlinear model operator leading to the so-called unscented Kalman filter (UKF) [17–20]. Using a set of carefully chosen deterministic samples (so-called sigma points), UKF estimates the first and second order moments of the state vector correct to the second order Taylor series expansion of the model operator.

In KF and EKF, the analysis of the measurements are performed sequentially leading to so-called *sequential* data assimilation techniques. Variational data assimilation methods, on the other hand, strive to produce an optimal analysis which fits a set of measurement data taken over a period of time [3, 21–24]. The optimality criterion is formulated as the minimization of a cost functional that incorporates the least-squares distance to (1) a prior estimate of the initial and boundary state (in the context of partial differential questions), (2) time distributed observations, and (3) modeling residual in an analysis interval. For a linear and Gaussian model, variational method leads to the same estimate as KF [3]. Variational data assimilation with *adjoint* models [3, 25] leads to difficulties for chaotic systems due to (i) their sensitive dependence on initial conditions; and (ii) its inability to tackle hypersensitivity of state trajectories to small perturbations [26]. Lea et al. [27] and Kohl et al. [28] made some attempts to alleviate these issues.

Recently, Monte Carlo based sequential filtering algorithms have been developed to tackle the general case of nonlinear systems perturbed by non-Gaussian noise. Monte Carlo based approach involves representing the probability density of the state and parameter estimate by a finite number of randomly generated samples. In the context of data assimilation, these methods have been proved to be successful [1]. These algorithms circumvent the need for labor-intensive development of the *tangent linear model* and its adjoint, as required for variational schemes. For a comparative study between Monte Carlo assimilation algorithms and variational techniques, the reader may refer to [1, 29]. The most widely used Monte Carlo based filtering algorithms are (i) ensemble Kalman filter (EnKF) [30–32]; and (ii) particle filter (PF) [33–41].

The usefulness of EnKF [30, 42, 43] has been successfully demonstrated for different applications in data assimilation problem (e.g., [1, 31, 32, 40]). EnKF inherits the same analysis step from the KF (or EKF). Subsequently, the ensemble members are propagated from one time step to another using the original model operator (no tangent linear system operator is required). Additionally, the measurement error covariance matrix in EnKF is estimated from the statistics of the ensemble of state vectors [30]. An artificially generated ensemble of measurements is also required by perturbing the original measurements [42, 43]. In numerous applications, it has been demonstrated that EnKF resolves the issue of poor error covariance evolution commonly encountered in EKF [30]. In the context of structural dynamics, EnKF has recently been introduced as a nonparametric identification tool by Ghanem and Ferro [32].

EnKF is a particularly appealing ensemble-based data assimilation method because it assumes a Gaussian prior pdf as in KF. The Gaussian assumption greatly simplifies the analysis step. However, the nonlinear and non-Gaussian effect is fully accounted for in the state evolution (*forecast*) in contrast to KF. Although this approach has been widely used for some time, some publications reported rather

poor performance by EnKF in certain applications. For example, Kivman [40] demonstrated that EnKF performed rather poorly when applied to simultaneous state and parameter estimation in the Lorenz model. For an atmospheric system, Miller et al. [44] also reported the poor performance of EnKF. In the aforementioned cases, highly non-Gaussian posterior pdfs lead to difficulties in EnKF. To alleviate such problem, a non-Gaussian extension of EnKF is reported that utilizes a weighted sum of Gaussian pdfs (instead of a single Gaussian pdf) to represent the prior distribution [35, 45, 46]. Mandel et al. [47] proposed another extension that combines EnKF with *nonparametric* density estimation techniques based on the distance in Sobolev spaces.

A more general approach, namely PF, is devised in the framework of Bayesian inference [36, 41, 48, 49]. This methodology relaxes the assumption of Gaussian posterior pdf and linearity of model and measurement operators as assumed in KF, EKF and EnKF. In this approach, the ensemble members (so-called *particles*) are properly weighted based on the measurements. However, in contrast to EnKF, the trajectories of the particles evolve independently without any analysis step. This fact partly undermines the efficacy of PF demanding larger ensembles compared to EnKF leading to higher computational costs. This problem is generally tackled using the so-called *resampling* technique [36, 37, 50]. Perhaps the earliest applications of PF for statistical inference are given by Handschin [51], Akashi, and Kumamoto [52], and Zaritskii et al. [53]. However, the formal particle filter approach was established by Doucet et al. [49] and Ristic et al. [48] through the introduction of a novel resampling technique. The approach is also sometimes known as the bootstrap filter, condensation algorithm, interacting particle approximation, or survival of the fittest algorithm [36, 48, 50, 54–56]. In the context of nonlinear dynamics, Manohar et al. applied PF to identify a nonlinear stiffness parameter of a Duffing oscillator [37–39]. Kivman compared the applicability of EnKF and PF for parameter estimation in the Lorenz model [40]. Ching, Beck, and Porter compared the performance of EKF and PF for the state and parameter estimation of various nonlinear systems [57]. To improve the performance of PF, hybrid algorithms which combine EKF or unscented Kalman filter (UKF) with PF have also been proposed leading to the so-called EKF and UKF particle filters (e.g., [20]).

In the literature of data assimilation (most common in climate research community), the Lorenz system describing the three state variable model for chaotic atmospheric turbulence is extensively used as a demonstration tool to study various filtering algorithms (e.g., [1, 40, 44]). Such comparative studies for nonlinear filters are, however, not widely reported in the literature of mechanical vibration. In this paper, we conduct a detailed numerical investigation to compare the performance of various Gaussian and non-Gaussian filters for chaotic mechanical vibration. In particular, we investigate the chaotic vibration of a Duffing oscillator[1] which may, for instance, represent a single mode Galerkin approximation of the motion of a buckled elastic beam [59]. In particular, we consider a *double-well* potential Duffing oscillator whereby the unforced system is characterized by two stable and one unstable *fixed points* [59]. This system offers computational simplicity while exhibiting a kaleidoscope of dynamical behavior with a slight change of its input and system parameters.

This investigation compares the feasibility of using EKF, EnKF and PF for joint state and parameter estimation of a noisy Duffing oscillator undergoing chaotic motion. The system is perturbed by combined harmonic and random inputs ensuing a transition pdf of motion which displays strongly non-Gaussian features. An extensive numerical study is undertaken to contrast the performance of various filtering algorithms with respect to sparsity of observational data and strength of model and measurement noise. The paper is organized as follows: Sect. 2 provides a description of noisy dynamical systems in the framework of Itô stochastic differential equations. A brief background of KF is given in Sect. 3. The application of EKF to nonlinear systems is briefly presented in Sect. 4. A brief introduction to EnKF is presented in Sect. 5. In Sect. 6, the formulation of PF is reviewed. Section 7 reports the results of numerical investigations elucidating the capabilities and limitations of various filtering techniques. The paper concludes in Sect. 8 where a summary and findings of the current investigation are provided.

[1] Although the Lorenz and Duffing models describe entirely different physical phenomena relating to turbulence and mechanical vibration, respectively, a connection can be drawn with these two systems whereby the Lorenz model can be compared to a *controlled* Duffing oscillator [58].

2 Background to noisy dynamical systems

The temporal evolution of a nonlinear dynamical system can be represented by the vector Itô stochastic differential equation [7, 60–62]

$$d\mathbf{x} = \mathbf{g}(\mathbf{x}, t)\, dt + \mathbf{H}(\mathbf{x}, t)\, d\mathbf{w}, \tag{1}$$

where $\mathbf{x}$ is an n-dimensional state vector, $\mathbf{g}$ is an n-dimensional random vector function, $\mathbf{H}$ is a $n \times m$ matrix-valued function and $\mathbf{w}$ is an m-dimensional vector Wiener process. It is a well-known fact that the interpretation of the solutions of (1) is problematical in the classical sense as the solution $\mathbf{x}(t)$ is continuous everywhere but not differentiable anywhere with probability 1.

The transition pdf $p(\mathbf{x}, t|\mathbf{x}_0, t_0)$ of the state vector $\mathbf{x}(t)$ given the initial pdf of state vector $p(\mathbf{x}_0, t_0)$ satisfies the Fokker–Planck equation [7, 60–62] given by:

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\{p g_i\} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left\{p\left(\mathbf{H}\mathbf{H}^T\right)_{ij}\right\}. \tag{2}$$

Note that the Fokker–Planck equation is an advection-diffusion partial differential equation (pde) having spatial dimension equal to the order of the state vector $\mathbf{x}$. The closed-form solution of the Fokker–Plank equation is available only for a restrictive class of systems [60, 61, 63]. In the framework of data assimilation, one is interested in obtaining the pdf of the state vector conditional upon the observational data $\mathbf{d}_k$ available at time t_k. Applying Bayes rule, the conditional pdf is given by:

$$p(\mathbf{x}_k, t_k|\mathbf{d}_k) = \frac{p(\mathbf{d}_k|\mathbf{x}_k, t_k)\, p(\mathbf{x}_k, t_k|\mathbf{d}_k^-)}{\int p(\mathbf{d}_k|\mathbf{x}_k, t_k)\, p(\mathbf{x}_k, t_k|\mathbf{d}_k^-) \mathrm{d}\mathbf{x}_k}. \tag{3}$$

where $p(\mathbf{x}_k, t_k|\mathbf{d}_k^-)$ is the pdf given by the Fokker–Planck equation up to but not including the new observation $\mathbf{d}_k$ and $p(\mathbf{d}_k|\mathbf{x}_k, t_k)$ is the pdf of the observation $\mathbf{d}_k$ given the current state vector $\mathbf{x}_k$. The analytical expression for the conditional pdf of the state vector $p(\mathbf{x}_k, t_k|\mathbf{d}_k)$ is not available in general, therefore, numerical solutions are sought using filtering techniques, such as EKF, EnKF and PF [1, 7]. For a problem described by linear model and measurement operators, perturbed by additive Gaussian model and measurement noise, KF solves (3) analytically [1, 7].

3 Kalman filter

KF assimilates data into linear systems with linear measurement operator subject to additive Gaussian noise. The filter has the advantage of being a sequential data assimilation method in which only the state variables at the previous time step are required. In addition, KF also serves as the mathematical foundation of other filters, such as EKF and EnKF. This section presents a brief overview of KF for completeness based on the references [1, 7, 48, 64, 65].

The discrete state-space representation of a linear dynamical system with a linear measurement operator is given by

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{f}_k + \mathbf{q}_k, \tag{4}$$

$$\mathbf{d}_k = \mathbf{C}_k \mathbf{x}_k + \boldsymbol{\epsilon}_k, \tag{5}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$ is the discrete model operator, $\mathbf{f} \in \mathbb{R}^n$ is a deterministic external input, $\mathbf{d} \in \mathbb{R}^m$ is the measurement vector which relates to the true state by the measurement matrix $\mathbf{C} \in \mathbb{R}^{m\times n}$. $\mathbf{q} \in \mathbb{R}^n$ and $\boldsymbol{\epsilon} \in \mathbb{R}^m$ are independent zero-mean Gaussian vector random processes with covariance matrices $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and $\mathbf{\Gamma} \in \mathbb{R}^{m\times m}$, denoting model and measurement errors, respectively. It is also assumed that $\mathbf{x}$ has a Gaussian prior pdf given by $\mathbf{x}_k \sim \mathcal{N}(\mathbf{x}_k^f, \mathbf{P}_k)$. KF first estimates the conditional mean and covariance of $\mathbf{x}_k$ given the measurement vector $\mathbf{d}_k$, denoted by $\mathbf{x}_k^a$ and $\widehat{\boldsymbol{P}}_k$, respectively. This constitutes the *analysis* step of the filter:

Analysis step:

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}_k^T \left[\mathbf{\Gamma}_k + \mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T\right]^{-1}, \tag{6}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k\left(\mathbf{d}_k - \mathbf{C}_k \mathbf{x}_k^f\right), \tag{7}$$

$$\widehat{\boldsymbol{P}}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{C}_k]\mathbf{P}_k. \tag{8}$$

$\mathbf{K}_k \in \mathbb{R}^{n\times m}$ is known as the *Kalman gain matrix*. The estimate $\mathbf{x}_k^a$ given in (7) is the *optimal linear unbiased estimate*. The optimality is achieved in the sense that the error covariance of the state conditional upon an observation $\mathbf{d}_k$ given in (8) is minimum. In (6–8), the superscript f indicates the *best* estimate of the current state denoting the *forecast* obtained from the previous state. The superscript a denotes the *analysis* step providing the best estimate of the current state conditional upon available observation $\mathbf{d}_k$.

The next step in KF is to estimate the mean and covariance of the state vector for the next time step, given by

Forecast step:

$$\mathbf{x}_{k+1}^{f} = \mathbf{\Phi}_k \mathbf{x}_k^a + \mathbf{f}_k, \tag{9}$$

$$\mathbf{P}_{k+1} = \mathbf{\Phi}_k \widehat{\boldsymbol{P}}_k \mathbf{\Phi}_k^T + \mathbf{Q}_k. \tag{10}$$

The underlying assumptions in KF involves linearity in the model and measurement operators and Gaussian model and measurement noise. These two assumptions are the limitations of this technique to tackle nonlinear systems. EKF is an extension of KF to deal with weakly nonlinear evolution equations as described next.

4 Extended Kalman filter

For nonlinear systems, the pdf of the state vector $\mathbf{x}$ is generally non-Gaussian even if the model noise is additive and Gaussian. For weakly non-Gaussian behavior, one can reasonably approximate the pdf of $\mathbf{x}$ by a Gaussian process through linearization.

For general nonlinear dynamical systems with additive model noise, the state evolution equation is given by

$$\mathbf{x}_{k+1} = \mathbf{\Psi}_k(\mathbf{x}_k, \mathbf{f}_k) + \mathbf{q}_k, \tag{11}$$

where $\mathbf{\Psi}$ is the discrete nonlinear model operator, $\mathbf{f}$ and $\mathbf{q}$ are the deterministic and random input terms as described in the previous section.

Given some measurement vector $\mathbf{d}_k$, we obtain a Gaussian conditional pdf of the state vector with mean $\mathbf{x}_k^a$ and covariance matrix $\widehat{\boldsymbol{P}}_k$ using the same analysis step as in KF given by (6)–(8). Linearizing $\mathbf{\Psi}_k$ about $\mathbf{x}_k^a$, we obtain

$$\mathbf{\Psi}_k(\mathbf{x}_k, \mathbf{f}_k) \approx \mathbf{\Psi}_k\big(\mathbf{x}_k^a, \mathbf{f}_k\big) + \mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big(\mathbf{x}_k - \mathbf{x}_k^a\big) \tag{12}$$

where $\mathbf{\Psi}_k'(\mathbf{x}_k, \mathbf{f}_k)$ is the tangent linear operator which denotes the Jacobian matrix of $\mathbf{\Psi}_k(\mathbf{x}_k, \mathbf{f}_k)$ with respect to $\mathbf{x}_k$.

From (11) and (12), we obtain

$$\mathbf{x}_{k+1} = \mathbf{\Psi}_k\big(\mathbf{x}_k^a, \mathbf{f}_k\big) + \mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big(\mathbf{x}_k - \mathbf{x}_k^a\big) + \mathbf{q}_k. \tag{13}$$

From (13), the posterior mean and error covariance [1, 48, 64, 65] can be expressed as:

$$\mathbf{x}_{k+1}^{f} = \mathbf{\Psi}_k\big(\mathbf{x}_k^a, \mathbf{f}_k\big), \tag{14}$$

$$\mathbf{P}_{k+1} = \big[\mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big]\widehat{\boldsymbol{P}}_k\big[\mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big]^T + \mathbf{Q}_k. \tag{15}$$

Note that the Jacobian matrix is being used in (15) instead of $\mathbf{\Phi}_k$ in (9) in KF. Therefore, the major steps in EKF are:

1. Analysis step:

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}_k^T \big[\mathbf{\Gamma}_k + \mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T\big]^{-1}, \tag{16}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k\big(\mathbf{d}_k - \mathbf{C}_k \mathbf{x}_k^f\big), \tag{17}$$

$$\widehat{\boldsymbol{P}}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{C}_k]\mathbf{P}_k. \tag{18}$$

2. Forecast step:

$$\mathbf{x}_{k+1}^{f} = \mathbf{\Psi}_k\big(\mathbf{x}_k^a, \mathbf{f}_k\big), \tag{19}$$

$$\mathbf{P}_{k+1} = \big[\mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big]\widehat{\boldsymbol{P}}_k\big[\mathbf{\Psi}_k'\big(\mathbf{x}_k^a, \mathbf{f}_k\big)\big]^T + \mathbf{Q}_k. \tag{20}$$

5 Ensemble Kalman filter

The linearization step in EKF may lead to poor error covariance evolution for strongly nonlinear systems [1, 48]. EnKF partly alleviates this issue. EnKF is based on the representation of the probability density of the state estimate by a finite number N of randomly generated system states, known as ensemble members. Each sample is integrated forward in time independently using the full nonlinear evolution model defined in (11). However, the linear analysis step performed in KF and EKF remains the same for EnKF. EnKF algorithm can be summarized as follows [66]:

1. Create an ensemble $\{\mathbf{x}_{0,i}^f\}$ of size N with $i = 1, \ldots, N$, using the prior pdf of $\mathbf{x}_0$.
2. For each subsequent step, obtain perturbed measurements and estimated measurement error covariance matrix:

$$\mathbf{d}_{k,i} = \mathbf{d}_k + \boldsymbol{\epsilon}_{k,i}, \tag{21}$$

$$\mathbf{\Gamma}_k = \frac{1}{N-1}\sum_{j=1}^{N} \boldsymbol{\epsilon}_{k,j} \boldsymbol{\epsilon}_{k,j}^T. \tag{22}$$

3. Analysis step:

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}_k^T \left[\mathbf{\Gamma}_k + \mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T \right]^{-1}, \tag{23}$$

$$\mathbf{x}_{k,i}^a = \mathbf{x}_{k,i}^f + \mathbf{K}_k \left(\mathbf{d}_{k,i} - \mathbf{C}_k \mathbf{x}_{k,i}^f \right), \tag{24}$$

$$\widehat{\boldsymbol{P}}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{C}_k] \mathbf{P}_k. \tag{25}$$

4. Forecast step:

$$\mathbf{x}_{k+1,i}^f = \mathbf{\Psi}_k \left(\mathbf{x}_{k,i}^a, \mathbf{f}_k \right) + \mathbf{q}_{k,i}, \tag{26}$$

$$\widehat{\mathbf{x}}_{k+1}^f = \frac{1}{N} \sum_{j=1}^{N} \mathbf{x}_{k+1,j}^f, \tag{27}$$

$$\mathbf{P}_{k+1} = \frac{1}{N-1} \sum_{j=1}^{N} \left(\mathbf{x}_{k+1,j}^f - \widehat{\mathbf{x}}_{k+1}^f \right) \times \left(\mathbf{x}_{k+1,j}^f - \widehat{\mathbf{x}}_{k+1}^f \right)^T. \tag{28}$$

It is worthwhile at this stage to point out the following features of EnKF:

- The linear analysis step performed in KF and EKF remains the same for EnKF. This step provides an analytical expression for the conditional mean and covariance of the state vector at the analysis step. However, this is also the major limitation of EnKF for strongly nonlinear models, due to the statistical closure at the second order moments in the analysis step. This closure leads to saturation of the performance of EnKF at a certain ensemble size. In other words, a further increase in ensemble size beyond a critical value does not improve the estimate [40]. The inevitable loss of information (regarding higher-order statistics) emerging from this closure may have a significant impact on the estimate for strongly non-Gaussian models.
- In the forecast step, the full nonlinear model is integrated forward in time which, in contrast to EKF, captures the non-Gaussian effect in the state vector introduced by the nonlinear evolution. This eliminates the errors introduced in EKF in the forecast step due to linearization. Furthermore, the computational overhead associated with the tangent linear operator $\mathbf{\Psi}_k'(\mathbf{x}_k^a, \mathbf{f}_k)$ required in EKF is also obviated.
- In EKF, the computation of the forecast covariance matrix $\mathbf{P}_{k+1}$ explicitly requires the analysis covariance matrix $\widehat{\boldsymbol{P}}_k$. The storage of $\widehat{\boldsymbol{P}}_k$ may be memory intensive when the dimension of the state vector n is large. In EnKF, the forecast covariance matrix $\mathbf{P}_{k+1}$ is estimated using statistical averaging, circumventing the need to store the analysis covariance matrix.
- The need to perturb the measurements as in (21) is pointed out by Burgers et al. [42] and Whitaker and Hamill [67]. Such perturbation of the observational data is necessary to prevent underestimation of the error covariance matrix. It may be conjectured that the measurement perturbation induces a regularization effect on the estimates of EnKF.

6 Particle filter

Consider a more general representation of the evolution and measurement equations:

$$\mathbf{x}_{k+1} = \mathbf{g}_k(\mathbf{x}_k, \mathbf{f}_k, \mathbf{q}_k), \tag{29}$$

$$\mathbf{d}_k = \mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\epsilon}_k), \tag{30}$$

where $\mathbf{x}$ is the system state vector, $\mathbf{q}$ and $\boldsymbol{\epsilon}$ are independent zero-mean random vectors describing model and measurement errors, respectively, $\mathbf{f}$ is a deterministic input vector and $\mathbf{d}$ is a vector of the measurements. Let us define the state and the measurement matrices as

$$\mathbf{X}_k = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k\}, \tag{31}$$

$$\mathbf{D}_k = \{\mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_k\}, \tag{32}$$

where the state matrix $\mathbf{X}_k$ and observation matrix $\mathbf{D}_k$ denote the collection of state and observation vectors up to and including time instance t_k, respectively.

By applying Bayes' theorem [1, 37, 48], we obtain

$$p(\mathbf{X}_k|\mathbf{D}_k) = \frac{p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k)}{\int p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k) \mathrm{d}\mathbf{X}_k}. \tag{33}$$

Thus,

$$p(\mathbf{x}_k|\mathbf{D}_k) = \frac{\int p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k) \mathrm{d}\mathbf{X}_{k-1}}{\int p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k) \mathrm{d}\mathbf{X}_k}. \tag{34}$$

One obtains the conditional mean of the state vector by

$$\hat{\mathbf{x}}_k = \int \mathbf{x}_k \, p(\mathbf{x}_k|\mathbf{D}_k) \mathrm{d}\mathbf{x}_k = \frac{\int \mathbf{x}_k \, p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k) \mathrm{d}\mathbf{X}_k}{\int p(\mathbf{D}_k|\mathbf{X}_k) p(\mathbf{X}_k) \mathrm{d}\mathbf{X}_k}. \tag{35}$$

Note from (30), $\mathbf{d}_k$ depends only on $\mathbf{x}_k$. We, therefore, obtain

$$p(\mathbf{D}_k|\mathbf{X}_k) = \prod_{s=0}^{k} p(\mathbf{d}_s|\mathbf{x}_s). \tag{36}$$

Furthermore, $\mathbf{x}_{k+1}$ depends on $\mathbf{x}_k$, as evident from (29), leading to

$$p(\mathbf{X}_k) = p(\mathbf{x}_0) \prod_{s=1}^{k} p(\mathbf{x}_s|\mathbf{x}_{s-1}). \tag{37}$$

N random samples of $\mathbf{X}_k$ are generated from $p(\mathbf{X}_k)$ using (37). The estimate of the mean in (35) can be statistically approximated by (e.g., [48, 49])

$$\hat{\mathbf{x}}_k \approx \frac{\frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_{k,i}\, p(\mathbf{D}_k|\mathbf{X}_{k,i})}{\frac{1}{N}\sum_{j=1}^{N} p(\mathbf{D}_k|\mathbf{X}_{k,j})} = \sum_{i=1}^{N} w_{k,i}\mathbf{x}_{k,i} \tag{38}$$

where

$$w_{k,i} = \frac{p(\mathbf{d}_k|\mathbf{x}_{k,i})w_{k-1,i}}{\sum_{j=1}^{N} p(\mathbf{d}_k|\mathbf{x}_{k,j})w_{k-1,j}}. \tag{39}$$

Equation (39) implies the need for initial values $w_{0,i}$, $i = 1, \ldots, N$. One can start with $w_{0,i} = 1/N$. This leads to the following algorithm for the particle filter:

1. Draw N samples $\mathbf{x}_{0,i}$ using $p(\mathbf{x}_0)$, $i = 1, \ldots, N$
2. Set $w_{0,i} = 1/N$
3. Perform the following steps recursively:
   (a) Obtain $\mathbf{x}_{k,i}$ from $\mathbf{x}_{k-1,i}$ for each value of i using (29)
   (b) Obtain $w_{k,i}$ for each value of i using (39)
   (c) Compute the estimate of $\mathbf{x}_k$ using (38)

The following points should be noted which contrast PF to EnKF:

- PF is a fully non-Gaussian filter that handles general forms of nonlinearities in measurement and model operators and non-Gaussian model and measurement errors. No Gaussian approximation is resorted to, in contrast to EnKF and EKF whereby a Gaussian assumption is inherent in the analysis step. Therefore, PF may outperform EnKF and EKF in dealing with strongly nonlinear systems.
- PF involves a fully Bayesian paradigm which relies on nonparametric statistics and directly deals with the estimation of the pdf of state vector. In the limit of infinitely large ensembles, PF offers an asymptotically exact estimate to the conditional pdf in (3). This is in stark contrast to EnKF whereby a saturation of the filter estimate is encountered at a relatively small ensemble size. Excluding the non-Gaussian effect in the analysis step leads to erroneous state estimates even with infinitely large ensembles.
- Note that in EKF and EnKF, the analysis step induces a shock (or discontinuity) in the dynamics of the state by modifying the current estimate of the state vector using the observational data (see (24)). Evidently, such shock is absent in PF, thereby a natural temporal evolution of the state vector is maintained.
- Given a small ensemble size, EnKF generally provides better estimates compared to PF due to the presence of an analysis step. This may be attributed to the fact that the analysis step *nudges* the state trajectories toward the observational data by implicitly performing Tikhonov regularization [68]. When large ensemble sizes are necessary, PF may become impractical due to intensive computational requirements. In that case, EnKF may provide more practical means for data assimilation.
- The requirement for larger ensembles can be partly alleviated by adopting stratified sampling methods, such as Latin hypercube sampling (LHS) [69, 70], which is adopted in this paper. From the experience gathered by the authors through numerical investigations, the application of LHS, as opposed to simple Monte Carlo simulation, may dramatically reduce the required ensemble size PF. In particular, the application of LHS crucially improves the resampling step in PF.

6.1 Resampling

In most practical applications of PF, all but one particle (sample) will have negligible weights $w_{k,i}$ after a certain number of recursive steps. This is known as the degeneracy phenomenon [36, 71]. As a result of degeneracy, a large computational effort is wasted in updating particles which make little contribution to the state vector estimate. A suitable measure of degeneracy is the effective sample size [36, 48] given by

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N}(w_{k,i})^2}. \tag{40}$$

When $w_{k,i} = 1/N$, we obtain $N_{\text{eff}} = N$. On the other hand, when all but one weight is zero, we have $N_{\text{eff}} = 1$ indicating degeneracy.

Thus, the degeneracy phenomenon can be detected when $N_{\text{eff}} < N_{\text{thr}}$, where N_{thr} is the threshold value. When such condition is encountered, a resampling step is introduced in the following manner:

1. Draw N particles from the current particle set with probabilities proportional to their weights $w_{k,i}$, replacing the current particle set with the new one.
2. Set $w_{k,i} = 1/N$ for $i = 1, \ldots, N$.

Note that the resampling step may artificially reduce the estimated variance of the state vector, which may be misleading while interpreting the filter estimates.

7 Application to nonlinear dynamical systems

In this section, we consider a Duffing oscillator for joint parameter and state estimation using the aforementioned nonlinear filters. This simple model is well suited to demonstrate the capabilities of the various filtering methods without undue computational complexity.

7.1 Original system model

The stochastic analogue of the deterministic Duffing equation has the following form [59]:

$$\ddot{u}(t) + c\dot{u}(t) + k_1 u(t) + k_2 u^3(t) = \mathcal{T}\cos(\omega t) + \sigma_1 \xi_1(t), \tag{41}$$

where c is the damping coefficient, k_1 and k_2 are the stiffness coefficients, $u(t)$ is the displacement, $\mathcal{T}$ and ω are the amplitude and frequency of the harmonic input, respectively, $\xi_1(t)$ is a Gaussian white noise describing modeling error and σ_1 is some constant describing the strength of the random input. Our aim is to estimate the displacement u as well as the stiffness coefficients k_1 and k_2 from some noisy measurements d_k obtained at specific times t_k expressed by

$$d_k = u(t_k) + \epsilon_k. \tag{42}$$

The state-space representation of the above equation is

$$\dot{x}_1 = x_2, \tag{43}$$

$$\dot{x}_2 = -\left[cx_2 + k_1 x_1 + k_2 x_1^3\right] + \mathcal{T}\cos(\omega t) + \sigma_1 \xi_1(t), \tag{44}$$

where $x_1 = u$ and $x_2 = \dot{u}$. The above equation can be recast in the Itô stochastic differential equation form as in (1) given by

$$dx_1 = x_2 dt, \tag{45}$$

$$dx_2 = -\left[cx_2 + k_1 x_1 + k_2 x_1^3 - \mathcal{T}\cos(\omega t)\right] dt + \sigma_1 \xi_1(t)\, dt. \tag{46}$$

In (46), $\xi_1(t)\, dt = dW_1 = W_1(t_{k+1}) - W_1(t_k)$ where dW_1 is a Brownian path increment. Temporal discretization with time step Δt leads to the discrete state-evolution equations

$$\{x_1\}_{k+1} = \{x_1\}_k + \Delta t \{x_2\}_k, \tag{47}$$

$$\{x_2\}_{k+1} = \{x_2\}_k - \Delta t\left[c\{x_2\}_k + k_1\{x_1\}_k + k_2\{x_1\}_k^3 - \mathcal{T}\cos(\omega t_k)\right] + \sigma_1\sqrt{\Delta t}\,\varepsilon_{1,k}, \tag{48}$$

where $\varepsilon_{1,k}$ is a standard (zero-mean and unit standard deviation) Gaussian random variable.

Figure 1 displays the system response u under purely deterministic loading when $c = 0.3$, $k_1 = -1$, $k_2 = 1$, $\mathcal{T} = 0.5$, $\omega = 1.25$, $\sigma_1 = 0$ and $\Delta t = 5 \times 10^{-4}$. The phase-space diagram for the system indicating chaotic behavior is given in Fig. 2. The Poincaré map of the system is shown in Fig. 3, in which a strange attractor is displayed demonstrating some underlying structure.

Figures 4–5 present the transition pdf of the system state x_1 under purely random excitation as a function of time, with $\sigma_1 = 0.1$ and $\sigma_1 = 0.2$, respectively. These results illustrate the highly non-Gaussian nature of the system even in the absence of harmonic forcing. The results are shown for $c = 0.3$, $k_1 = -1$, $k_2 = 1$ and $\mathcal{T} = 0$. The transition pdfs are obtained using 1.5×10^6 realizations. Initial conditions are $x_1 \sim \mathrm{U}(-2, 2)$ and $x_2 = 0$, where U denotes the uniform probability density function. The transition pdfs evolves from a uniform initial pdf to a bimodal stationary (equilibrium) distributions shown in both cases. It is interesting to note that for the case of weaker random input shown in Fig. 4, the probability of the state variable x_1 being zero is almost zero. This is in contrast for the case of stronger random input shown in Fig. 5. Note that the underlying autonomous system

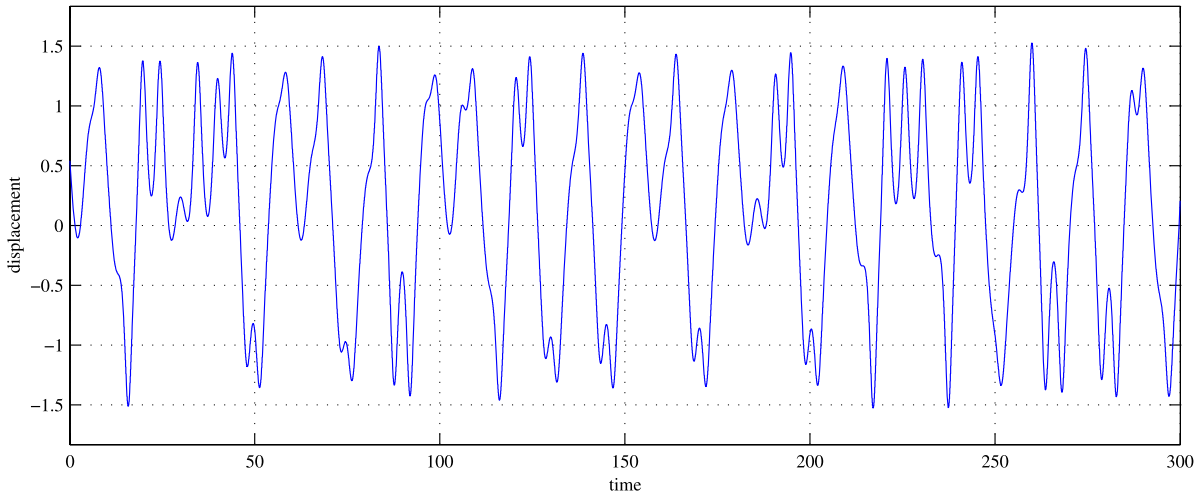

Fig. 1 Chaotic trajectory $u(t)$ of the Duffing oscillator

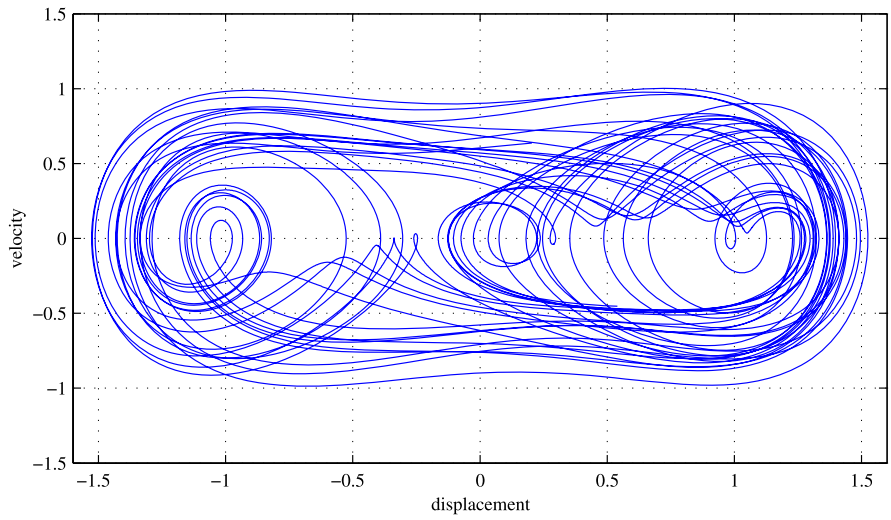

Fig. 2 Phase-space diagram of the chaotic Duffing oscillator

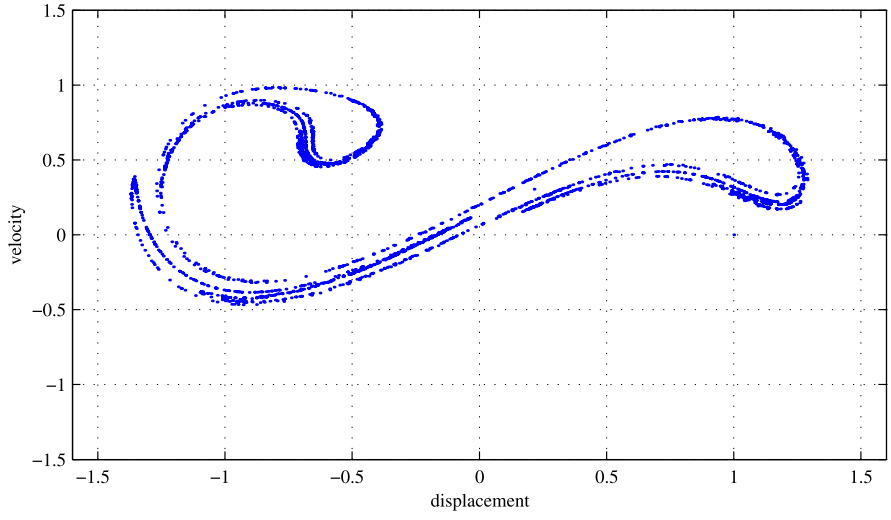

Fig. 3 Poincaré map of the chaotic Duffing oscillator

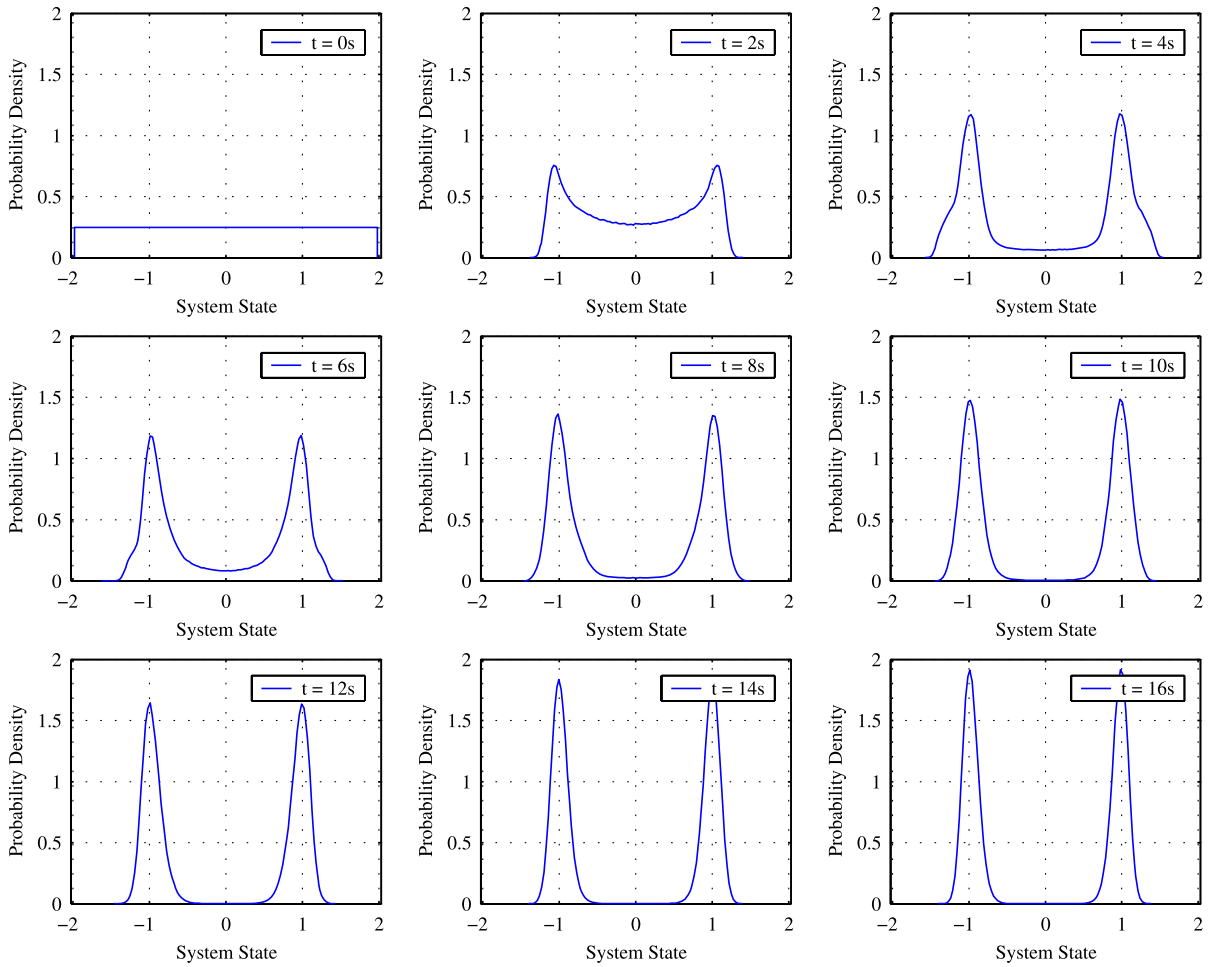

Fig. 4 Transition pdf of the system under purely random input, with $\sigma_1 = 0.1$ and $\mathcal{T} = 0$

has three *fixed points* at $u = 0$ and ± 1. The two dominant peaks of the pdfs are centered around the stable fixed points $u = \pm 1$.

Figures 6–7 show the transition pdf of the system state x_1 under combined deterministic and random input, with $\sigma_1 = 0.1$ and $\sigma_1 = 0.2$, respectively. The pdf results are obtained using 1.5×10^6 realizations. Initial conditions are $x_1 \sim \mathrm{U}(-2, 2)$ and $x_2 = 0$. It is clear that the pdfs are highly non-Gaussian, with multimodal characteristics. It is also interesting to see that increasing the strength of modeling error flattens the transition pdf inducing (loosely speaking) more randomness in the response.

Figure 8 shows the prior and posterior pdfs of the state x_1 with uniform initial conditions and experimental data available at a single time instant $t_k = 12$ s. For simplicity, the posterior pdfs are plotted by assimilating only a single measurement data for demonstration purposes. Two posterior pdfs (see (3)) are shown, one for the measurement data $d_k = -0.2$ and the other for the measurement $d_k = 0.2$, assuming zero mean Gaussian measurement errors with variance equal to 0.2. It is clear that the posterior pdfs in Fig. 8 are strongly non-Gaussian, again identifying the need for non-Gaussian filters for data assimilation.

7.2 State estimation

Figure 9 shows the displacement of the oscillator under the harmonic force $f(t) = 0.5\cos(1.25t)$ with a noise perturbation term whose amplitude is $\sigma_1 = 0.1$. The measured displacement d_k contaminated by Gaussian noise $\epsilon_k \sim \mathcal{N}(0, 7.8 \times 10^{-3})$ is also shown

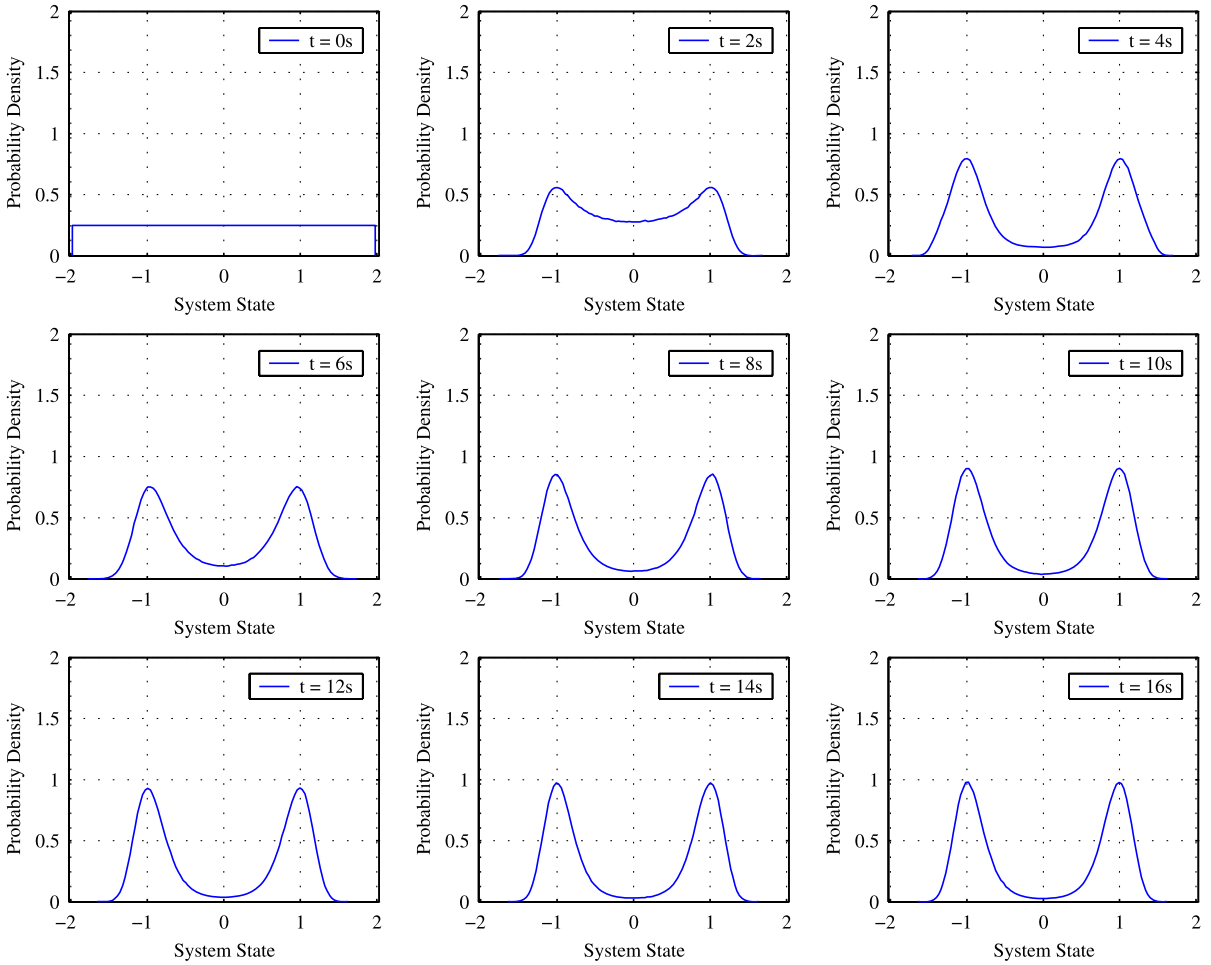

Fig. 5 Transition pdf of the system under purely random input, with $\sigma_1 = 0.2$ and $\mathcal{T} = 0$

in Fig. 9. The standard deviation of the measurement noise is 10% of that of the true displacement. We consider the observations obtained at time intervals of 1 second.

The results of the state estimation using the three filtering techniques are plotted in Fig. 9. For EnKF and PF, firstly, one hundred independent experiments are performed. In each of these experiments, an ensemble size $N = 10$ is used. Secondly, another twenty independent experiments are conducted with an ensemble size $N = 50$. The mean of the estimates of these experiments is plotted in Fig. 9 to reduce statistical errors. As the results from individual experiments may be misleading, these are not shown. To minimize sampling errors, LHS [69, 70] is used as an efficient sampling scheme for EnKF and PF. Initial conditions are $x_1 \sim \mathcal{N}(0.5, 0.1)$ and $x_2 \sim \mathcal{N}(0, 0.1)$. This figure presents the state estimation results using EKF (second panel from the top), EnKF and PF using 10 samples (third and fifth panels from the top) and 50 samples (fourth and sixth panels from the top). The moving average of the normalized root-mean-square (RMS) error of the estimates is also plotted in the bottom-most panel. The error is normalized by the variance of the true displacement and the average is taken over the last 10 seconds. The threshold effective sample size for PF is chosen to be 75% of the ensemble size (i.e., $N_{\text{thr}} = 0.75N$), which is found to be adequate for effective resampling. Using only 10 samples, EnKF provides a better estimate than PF (as observed in the RMS error plot). EKF gives the least accurate estimates. When comparing the estimates of EnKF and

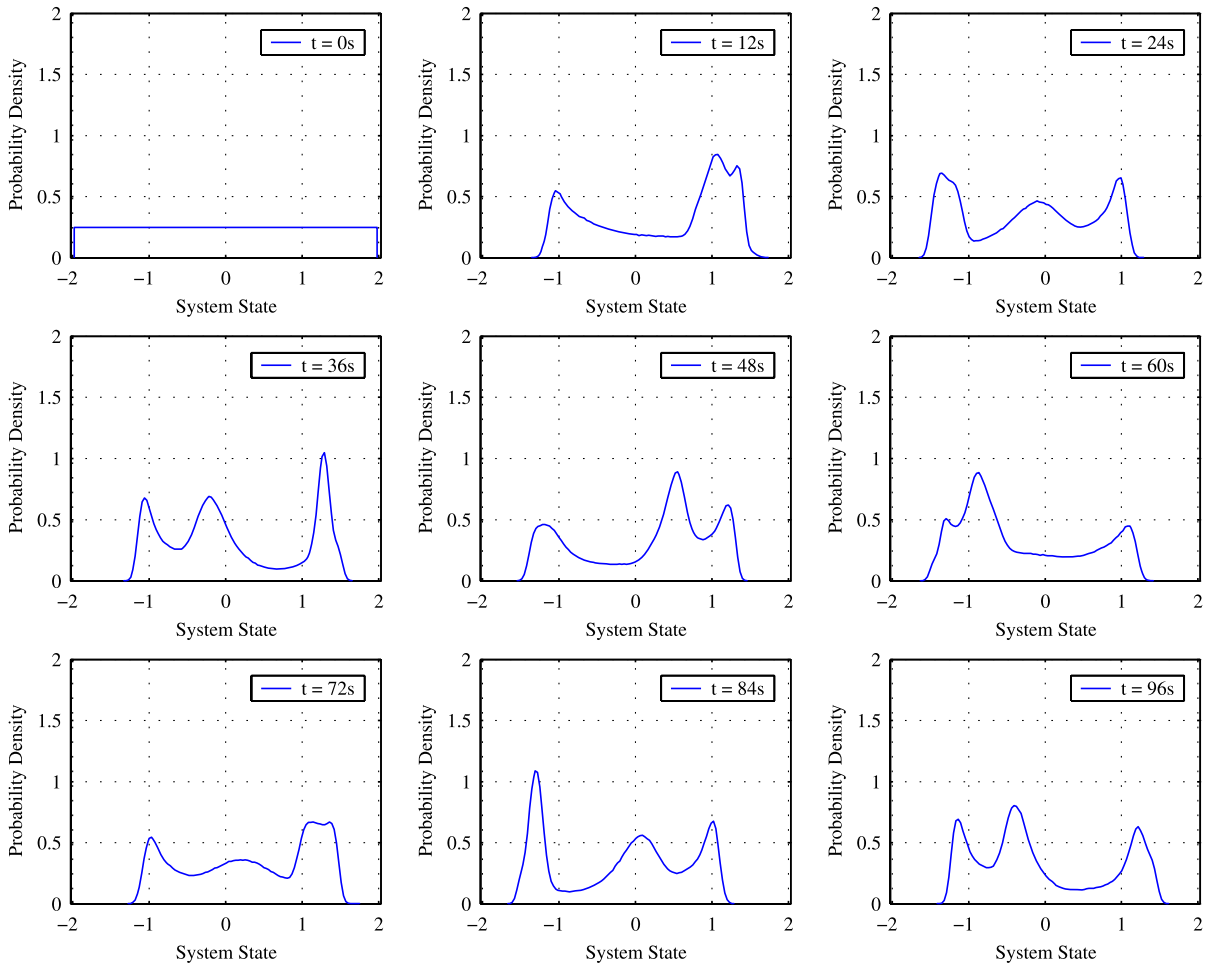

Fig. 6 Transition pdf of the system under combined deterministic and random input, with $\sigma_1 = 0.1$ and $\mathcal{T} = 0.5$

PF using 10 samples and 50 samples, it is apparent that PF improves in accuracy with an increase of ensemble size, whereas EnKF does not improve significantly. The authors conjecture that the saturation in performance of EnKF is due to the Gaussian closure inherent to the analysis step. Therefore, the further increase in the ensemble size does not offer any improvement in the state estimation. On the other hand, PF increasingly benefits from larger ensembles providing better estimates by capturing non-Gaussian features in the state vector.

7.2.1 Effect of measurement noise amplitude

To demonstrate the effect of noise on the state estimates, we consider the observations obtained at time intervals of 1 second, as in the previous experiment, but with stronger measurement noise in relation to the previous experiment. Figure 10 shows the true response of the oscillator and the measured response d_k polluted by $\epsilon_k \sim \mathcal{N}(0, 7.0 \times 10^{-2})$. The standard deviation of the measurement noise is taken to be 30% of that of the true displacement. The results of the state estimation using various filtering techniques are also plotted in Fig. 10. The same experimental setup was used from the previous section, with the same initial conditions. It is clear that an increase in the strength of measurement noise diminishes the accuracy of the estimates for all filters, in comparison to the previous experiment shown in Fig. 9. EKF is most severely affected leading to strongly biased estimates. Again, it is clear that EnKF does not benefit from a larger ensemble size $N = 50$, when com-

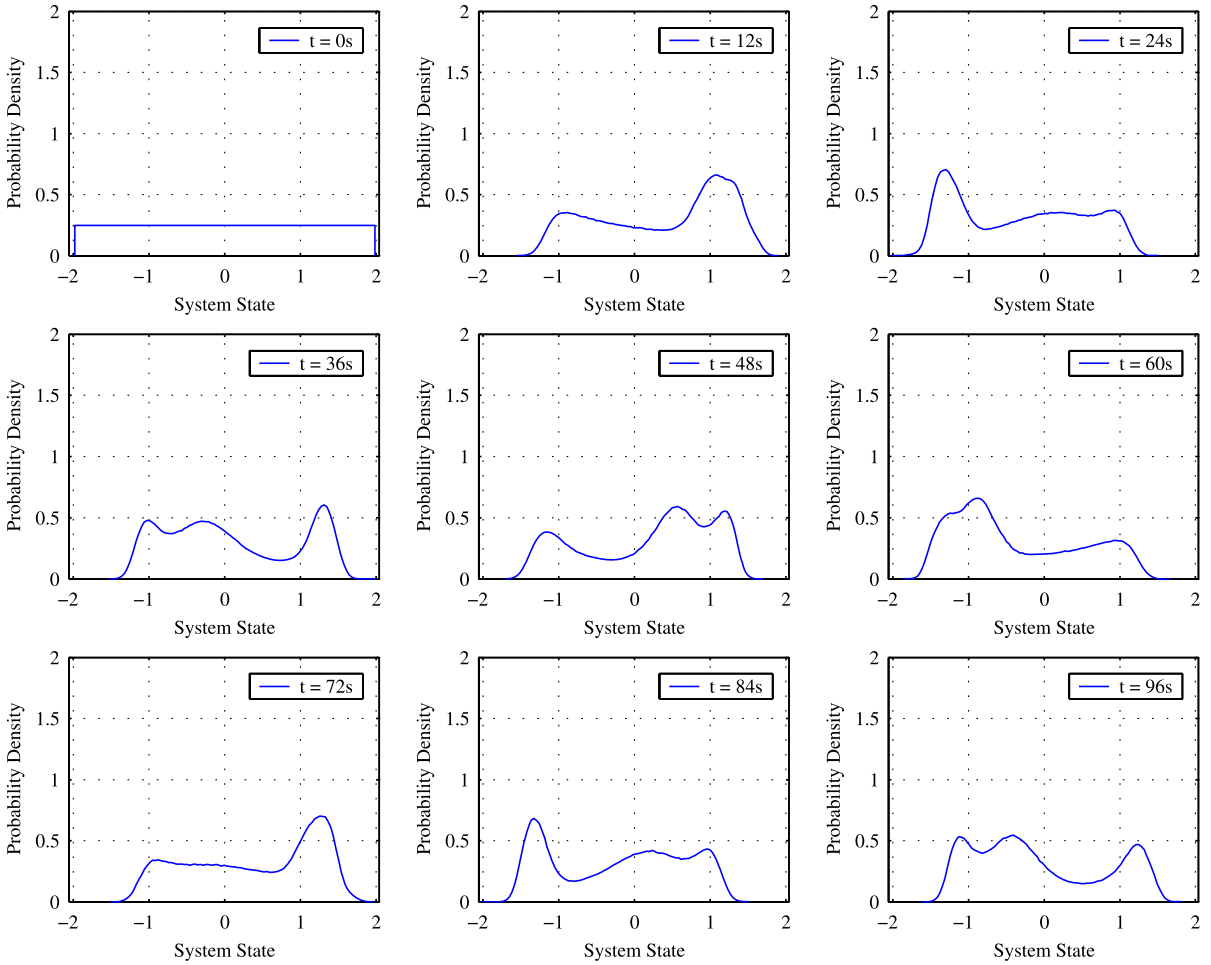

Fig. 7 Transition pdf of the system under combined deterministic and random input, with $\sigma_1 = 0.2$ and $\mathcal{T} = 0.5$

pared to $N = 10$. PF provides less accurate results than EnKF with 10 samples. PF however, outperforms EnKF for the case with $N = 50$ samples. It is believed that the slightly superior performance of PF in this experiment is due to stronger non-Gaussian features in the conditional pdf compared to the previous case. This maybe inferred, for instance, observing the conditional pdf plotted in Fig. 8 whereby the increase in measurement noise more effectively retains the non-Gaussian features in the posterior pdf dominant in the prior pdf.

7.2.2 Effect of measurement sparsity

To demonstrate the effect of measurement sparsity on the state estimates, we consider the case of sparse observations obtained at time intervals of 3 seconds, instead of 1 second used in the previous experiments. The noise signal is taken to have standard deviation equal to 10% of that of the true displacement. Figure 11 shows the true response of the oscillator and the measured response d_k corrupted by $\epsilon_k \sim \mathcal{N}(0, 7.8 \times 10^{-3})$. The results of the state estimation are plotted in Fig. 11. Similar to the last experiment, EKF fails spectacularly in estimating the true state as the data sparsity is increased. It is also observed that EnKF does not benefit significantly from a larger ensemble. PF provides more accurate results with a larger ensemble size ($N = 50$). The infrequent assimilation of data due to increased data sparsity permits the pdf of the state to increasingly gain the non-Gaussian features of the transition pdf in Figs. 6–7. In contrast, the frequent assimilation of measurement data with

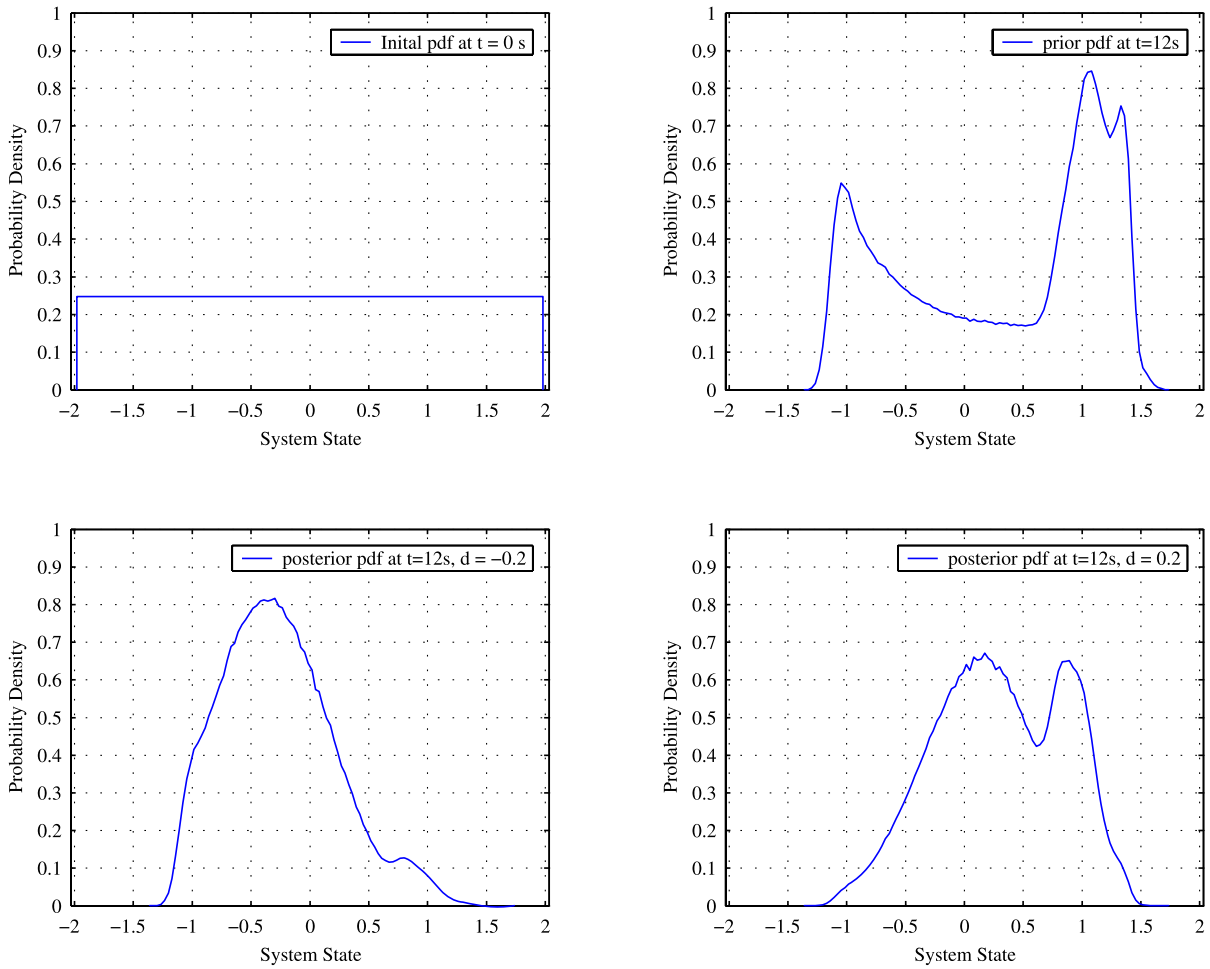

Fig. 8 Prior and posterior pdf of the system

Gaussian error tends to make conditional pdfs more Gaussian.

7.2.3 Effect of modeling uncertainty

To demonstrate the effect of modeling uncertainty on the filter estimates, we consider the case with modeling noise strength $\sigma_1 = 0.2$. Figure 12 presents the true response of the oscillator and the noisy measured response d_k with $\epsilon_k \sim \mathcal{N}(0, 7.8 \times 10^{-3})$. The strength of the measurement noise is 10% of that of the true displacement. The results of the state estimation using various filters are also plotted in Fig. 12. We present the state estimates using $N = 10$ and $N = 50$ samples for EnKF and PF. As in the previous cases, using 10 samples, EnKF provides good estimates when compared to the case when modeling error was significantly smaller (refer to Fig. 9). When 50 samples are used, PF matches the accuracy of EnKF in the state estimate.

7.3 Joint state and parameter estimation

We now investigate the performance of EKF, EnKF and PF for combined state and parameter estimation. The unknown parameters to be estimated are the stiffness coefficients k_1 and k_2. We augment the state vector by appending the coefficients k_1 and k_2 as two new state variables $x_3 = k_1$ and $x_4 = k_2$. The new variables are assumed to evolve using the following model

$$\dot{x}_3 = \sigma_2 \xi_2(t), \tag{49}$$

$$\dot{x}_4 = \sigma_3 \xi_3(t), \tag{50}$$

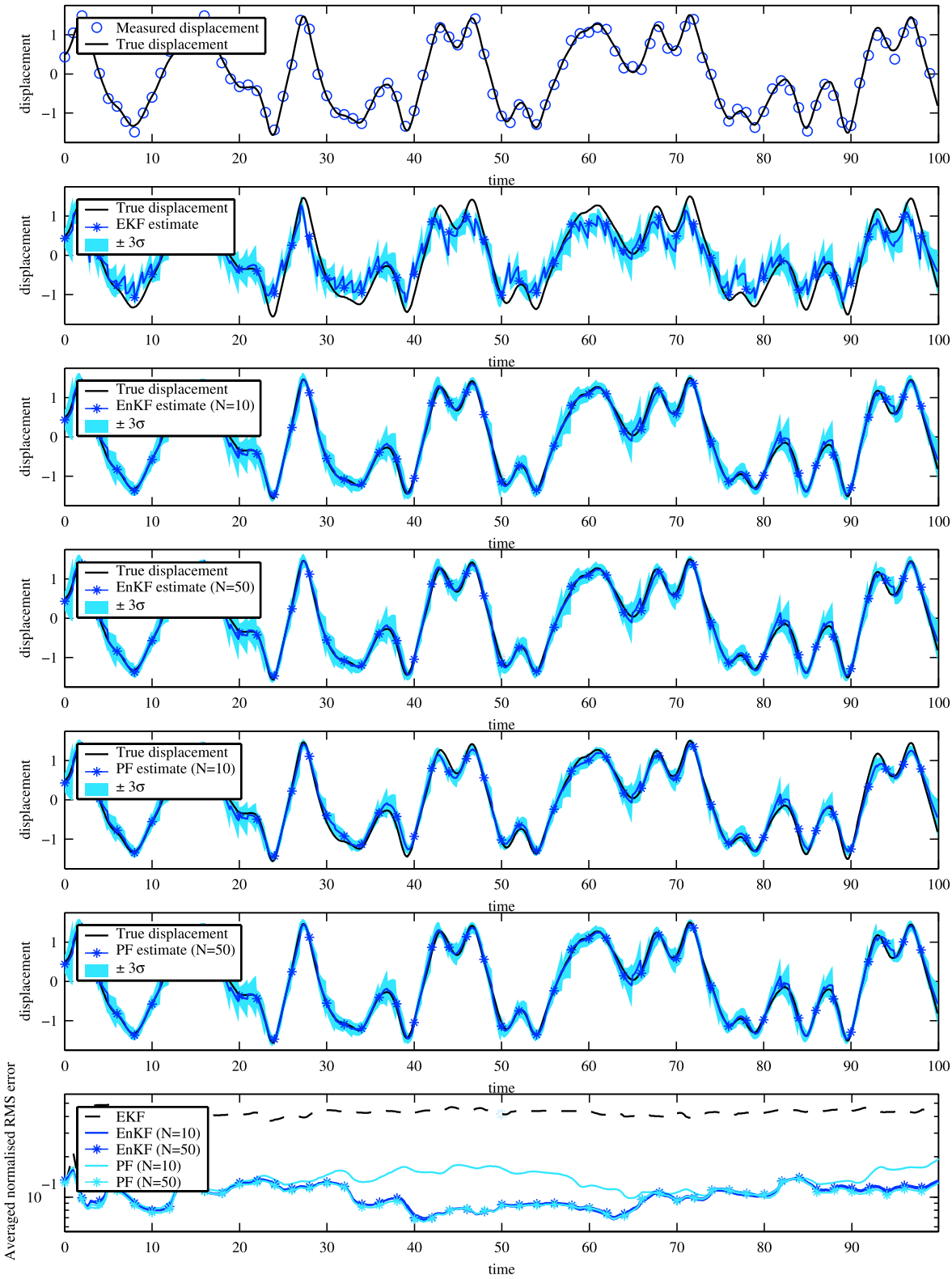

Fig. 9 True, measured and estimated displacement of the Duffing oscillator

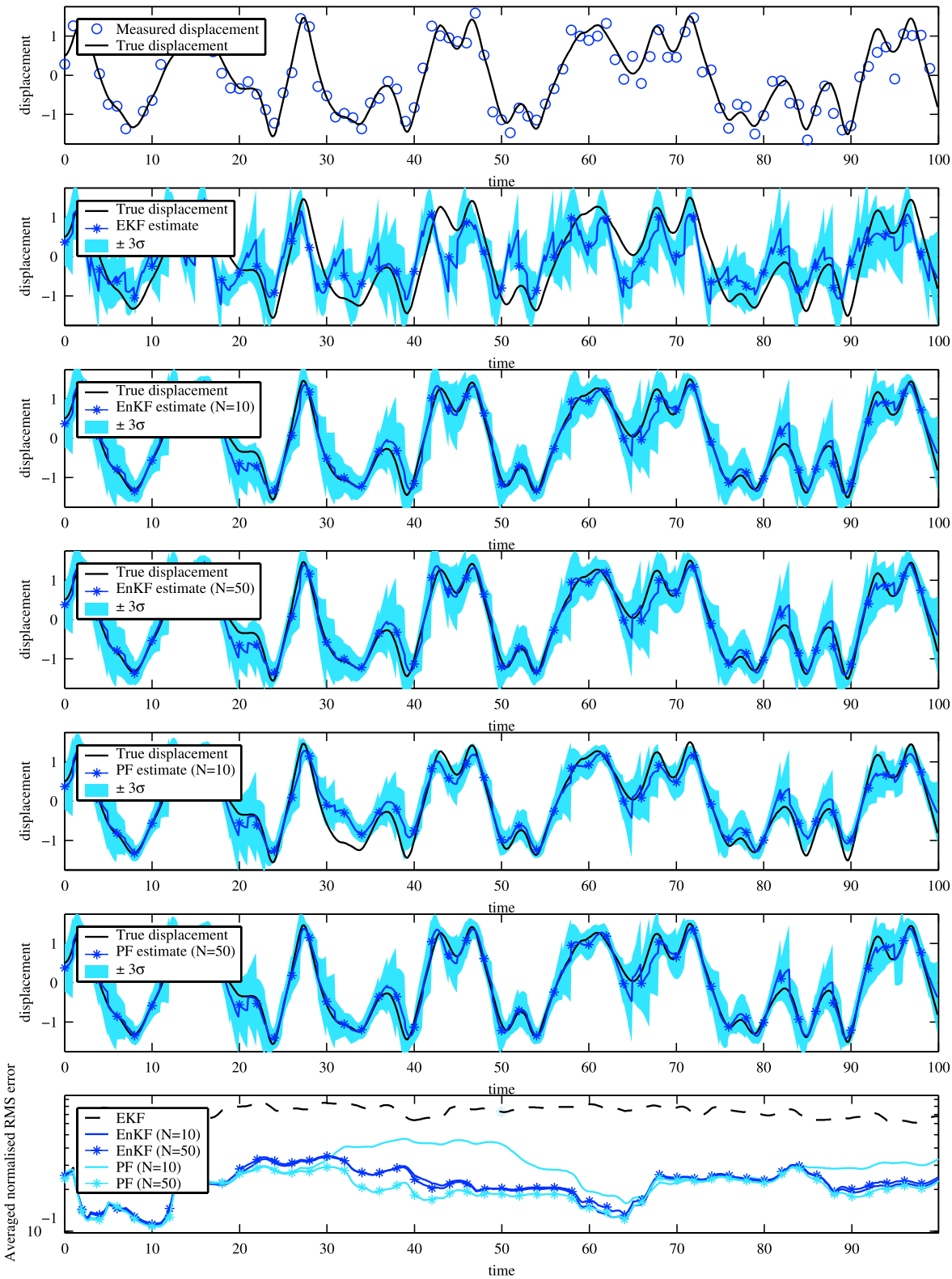

Fig. 10 True, measured and estimated displacement of the Duffing oscillator under stronger measurement noise

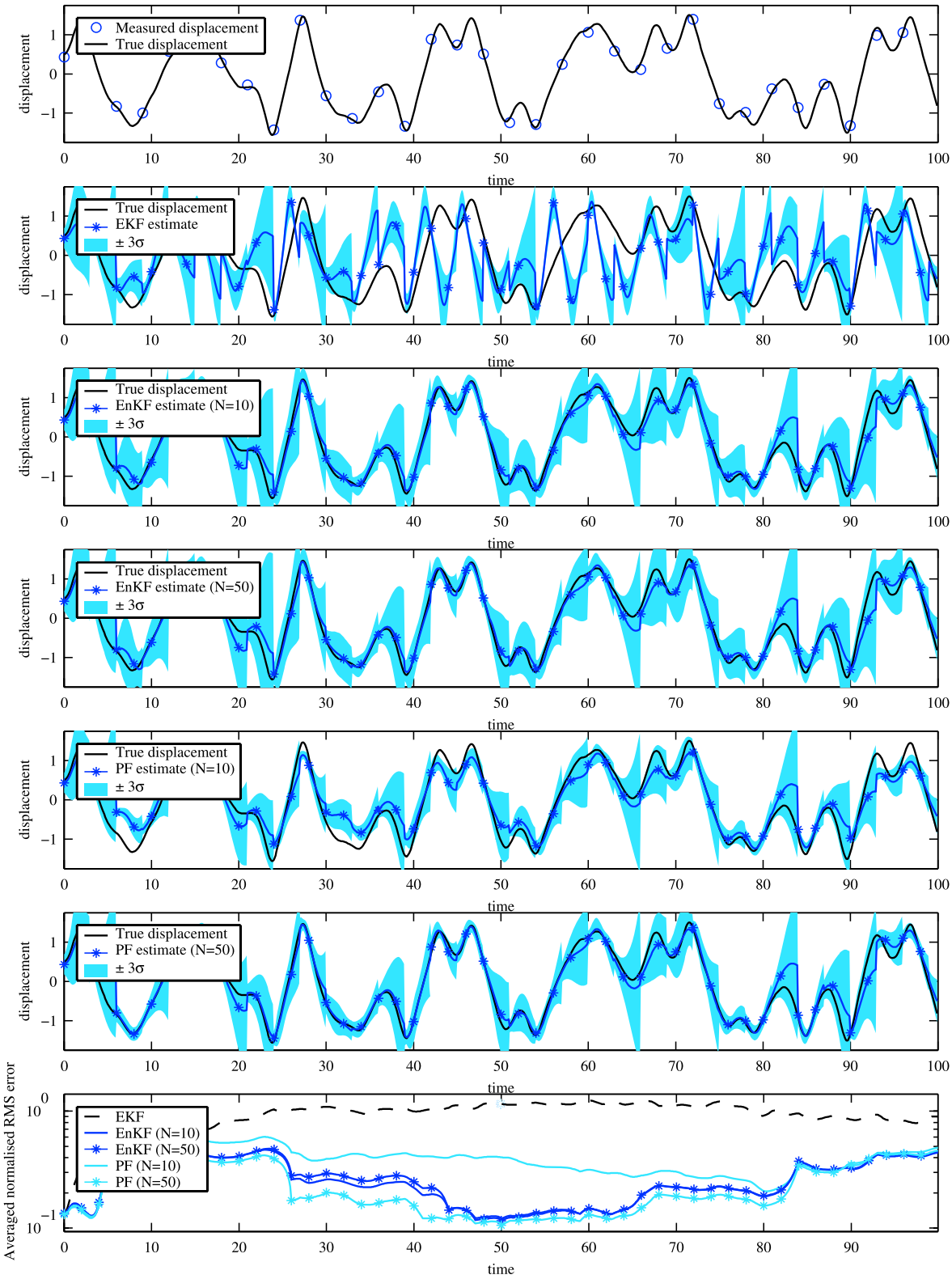

Fig. 11 True, measured and estimated displacement of the Duffing oscillator with sparse measurements

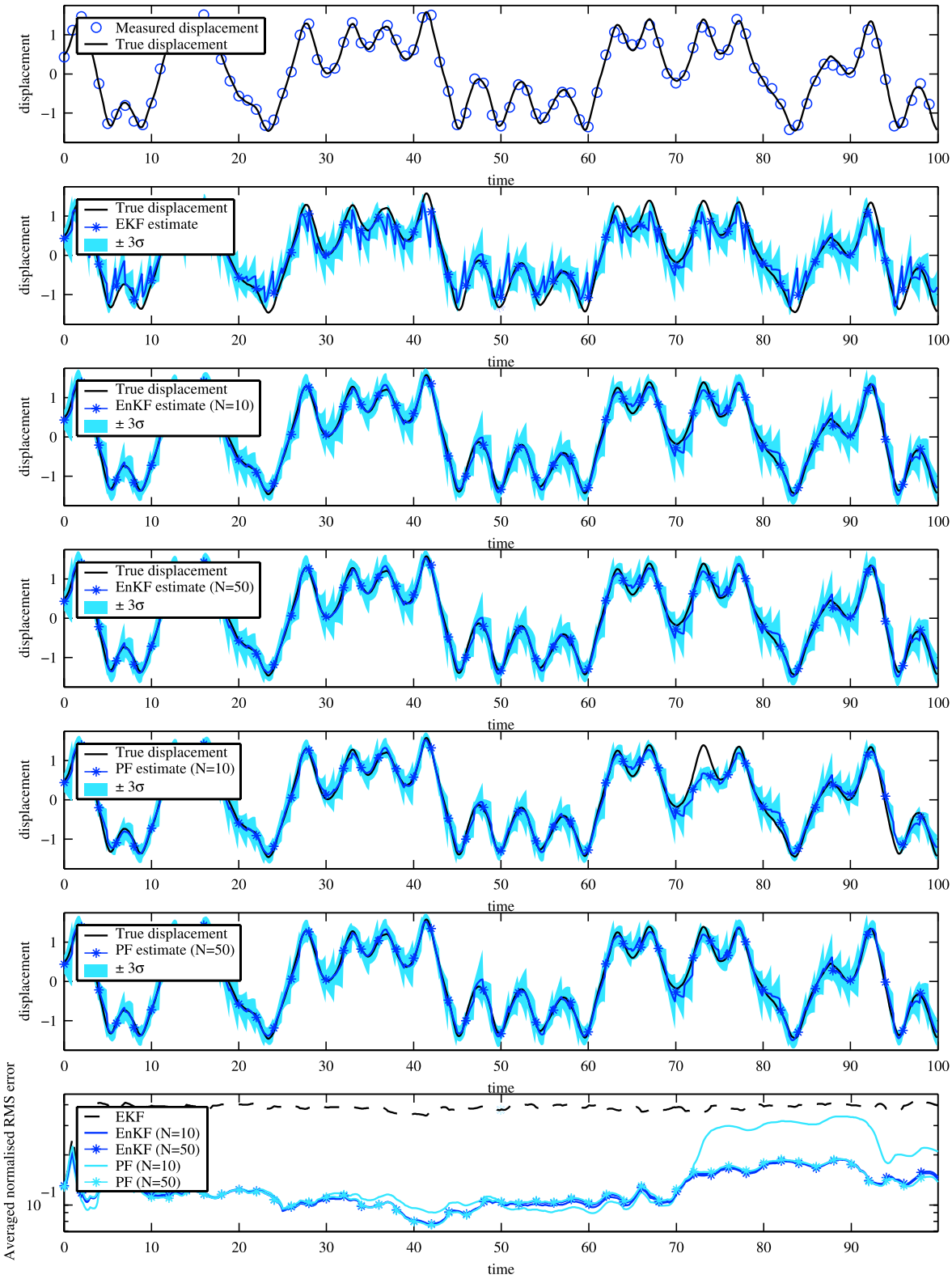

Fig. 12 True, measured and estimated displacement of the Duffing oscillator under greater modeling uncertainty

leading to the following augmented state-evolution equations

$$\{x_1\}_{k+1} = \{x_1\}_k + \Delta t\{x_2\}_k, \tag{51}$$

$$\begin{aligned}\{x_2\}_{k+1} = \{x_2\}_k - \Delta t\big[&c\{x_2\}_k + \{x_3\}_k\{x_1\}_k \\ &+ \{x_4\}_k\{x_1\}_k^3 - \mathcal{T}\cos(\omega t_k)\big] \\ &+ \sigma_1\sqrt{\Delta t}\varepsilon_{1,k},\end{aligned} \tag{52}$$

$$\{x_3\}_{k+1} = \{x_3\}_k + \sigma_2\sqrt{\Delta t}\varepsilon_{2,k}, \tag{53}$$

$$\{x_4\}_{k+1} = \{x_4\}_k + \sigma_3\sqrt{\Delta t}\varepsilon_{3,k}, \tag{54}$$

where $\varepsilon_{1,k}$, $\varepsilon_{2,k}$ and $\varepsilon_{3,k}$ are independent standard Gaussian random variables. The measurement equation remains the same as in (42).

The purpose of introducing the perturbation terms whose amplitudes are σ_2 and σ_3 is to inflate the variance of the parameter estimates and thus avoid filter divergence. Several methods are reported in the literature for inflating the estimates (see [72] for an overview): (1) Additive inflation in which noise is added to the estimates; (2) Multiplicative inflation where the estimate covariance matrix is multiplied by a constant factor, usually greater than one; (3) Model-specific inflation where only a subset of the model parameters are perturbed.

In this investigation, we employ model-specific additive inflation for which several methods exist to set the values σ_2 and σ_3: (a) Set the parameters initially to a fixed value, and possibly reducing the initial value gradually toward zero as the estimates converge in time; (b) Set their values proportional to the current error standard deviation of the parameter estimates; (c) Set the values proportional to the mismatch between the measurement data and the observed state estimate. We adopted the first method as this simple approach leads to satisfactory estimates of the stiffness parameters for the specific cases investigated in the paper. From extensive numerical experiments, the values $\sigma_2 = 0.03$ and $\sigma_3 = 0.03$ lead to rapid convergence of the filter estimates. Setting these values too large lead to divergence of the estimates, whereas the convergence of the estimates is slower for smaller values of these parameters.

The result of the joint state and parameter estimation are shown in Figs. 13–15. The system was exited by $f(t) = 0.5\cos(1.25t)$. Furthermore, the measurement error is given by $\epsilon_k \sim \mathcal{N}(0, 7.5 \times 10^{-3})$. The strength of the measurement noise is taken to be 10% of standard deviation of the true displacement. Initial conditions are $x_1 \sim \mathcal{N}(0, 0.01)$, $x_2 \sim \mathcal{N}(0, 0.01)$, $k_1 = x_3 \sim \mathcal{N}(-0.5, 0.01)$ and $k_2 = x_4 \sim \mathcal{N}(0.5, 0.01)$. The choice of initial conditions is another factor that influences the convergence of the filters. This set of initial conditions is chosen to illustrate the ability of the filtering algorithms to successfully track the system even with inaccurate initial estimates of the parameters. If the initial conditions are grossly inaccurate, the filters may diverge in which case one has to rerun the filters with different initial conditions. Normally, the prior knowledge, whenever available, should influence the choice of initial conditions. Figure 13 shows the true response of the oscillator and the noisy measured response. The filter estimates of the displacement of the oscillator plotted in Fig. 13. Comparing the results for the state estimation alone (see Fig. 9), the combined state and parameter estimation provides significant improvement in the state estimates as observed from Fig. 13. All three filters give satisfactory results (albeit the need for larger ensemble ($N = 50$) for PF as evident in Fig. 13).

The estimates for the stiffness coefficients k_1 and k_2 are plotted in Figs. 14–15, respectively. Note that EKF parameter estimates are surprisingly accurate, when compared to those of EnKF and PF. The superior performance of EKF may be specific to the particular experiment adopted here in which measurement data is relatively dense and measurement noise strength is relatively small. In general, this conclusion relating to the performance of EKF may not be true for other cases.

Using $N = 10$ samples, EnKF provides better estimates than PF. Increasing the number of samples to $N = 50$ improves the estimates for PF. While comparing the standard deviation of the error in the estimates, it is clear that PF leads to smaller error standard deviation. This is partly attributed to the resampling step undertaken whenever degeneracy is encountered in PF.

8 Conclusion

This paper explored the capabilities of EKF, EnKF and PF for joint state and parameter estimation for a noisy Duffing oscillator undergoing chaotic motion. Such methods can simultaneously estimate the state and parameters of a nonlinear system even in the presence

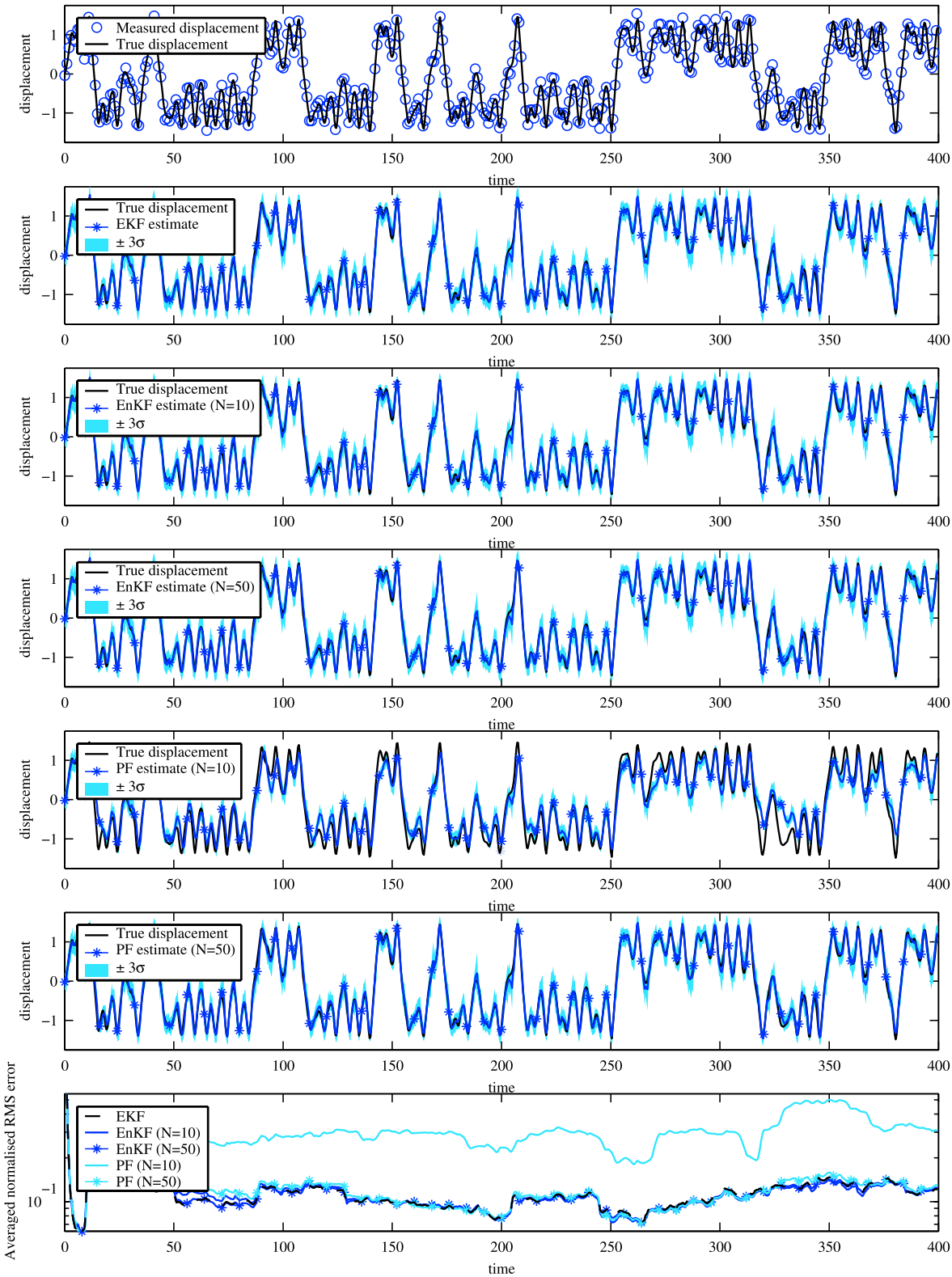

Fig. 13 True, measured and jointly-estimated displacement of the Duffing oscillator

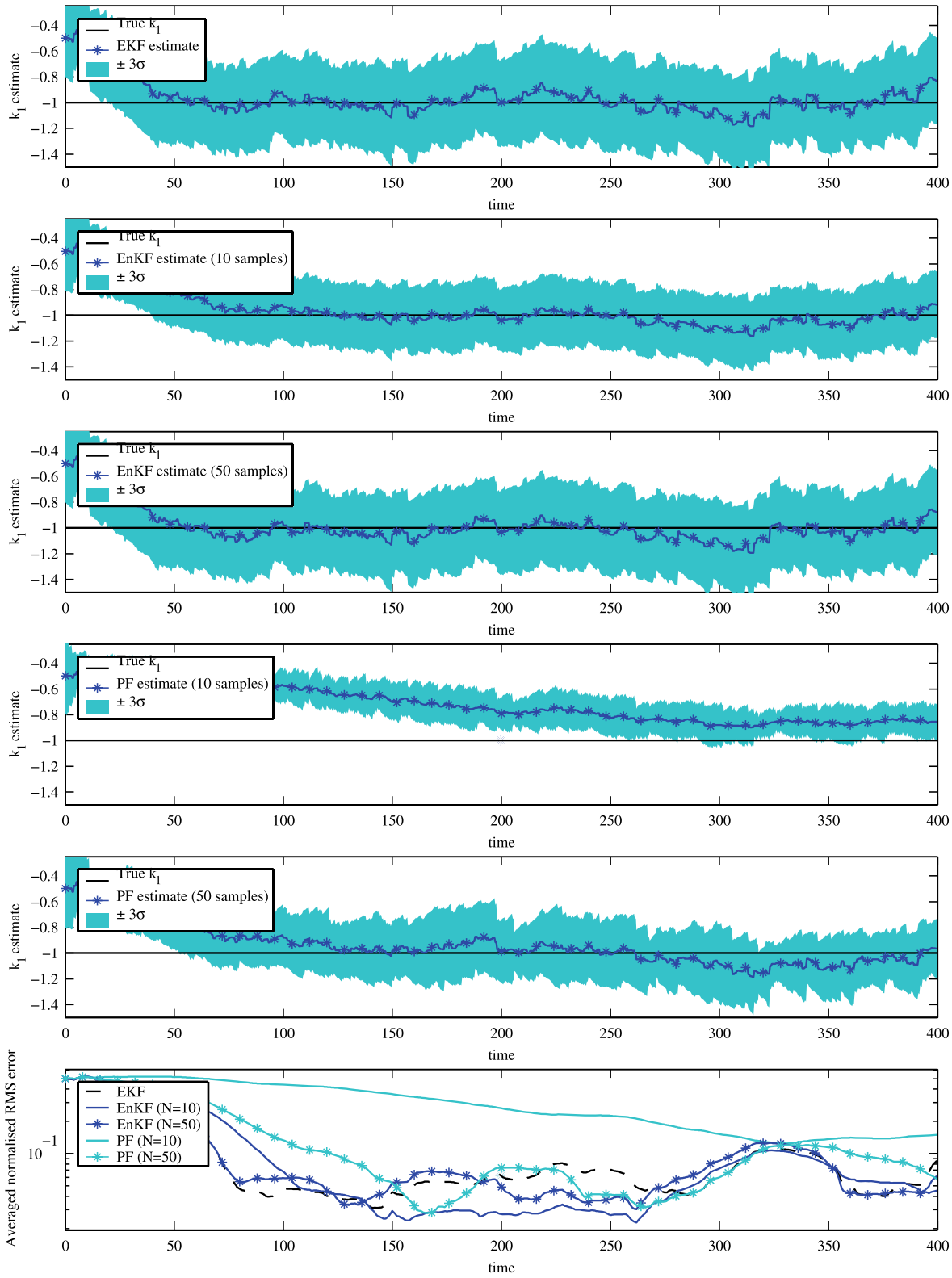

Fig. 14 k_1 parameter estimates

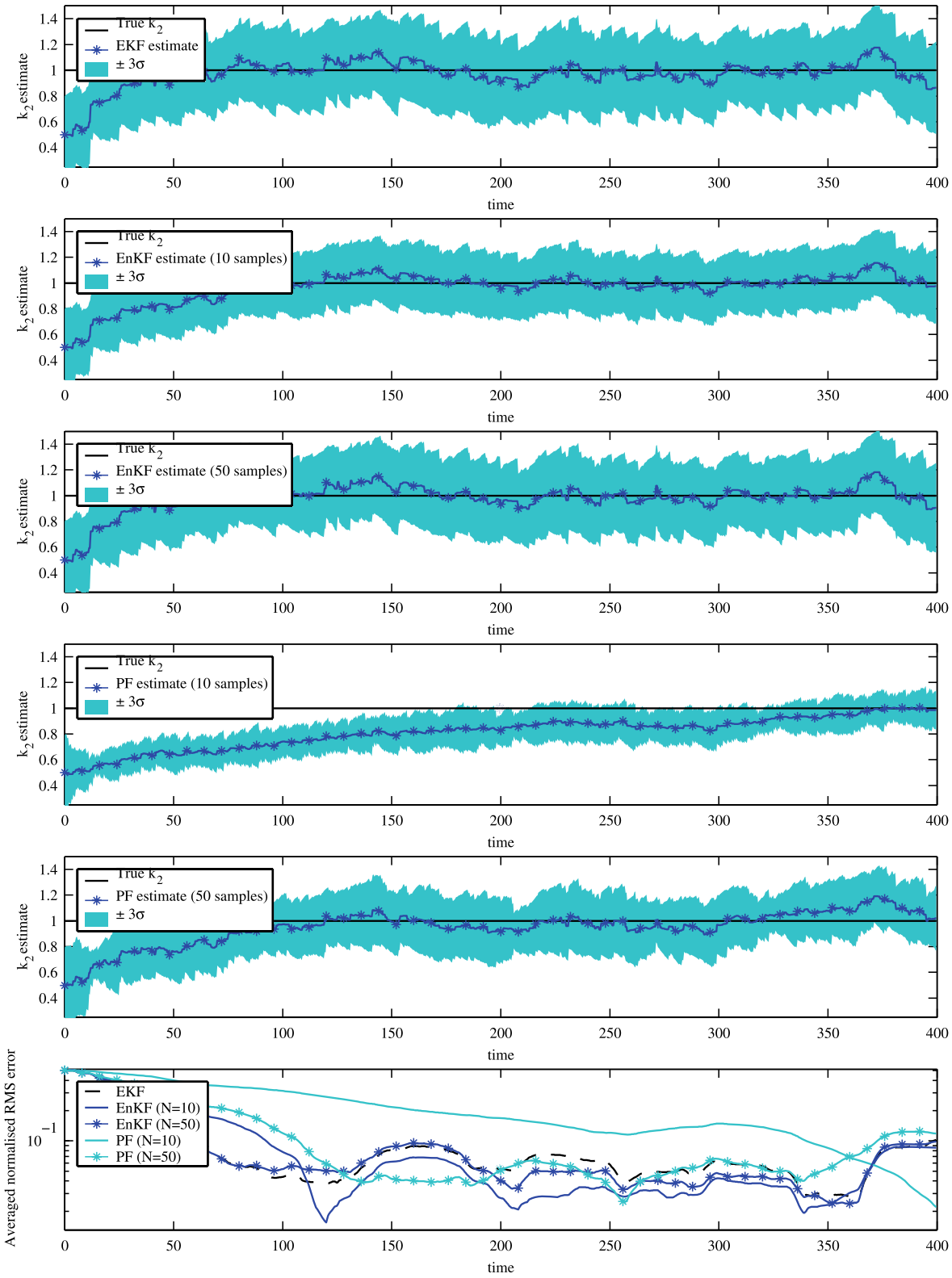

Fig. 15 k_2 parameter estimates

of model and measurement errors and accommodate non-Gaussian signals. To the authors' best knowledge, such comparative study in the context of mechanical vibration is not widely reported in the literature. The following features were brought out from the current investigation:

1. It is demonstrated that EnKF and PF perform better than EKF in tracking the true state of the system in the presence of large measurement noise or infrequent measurement data. This is attributed to non-Gaussian nature of the state variables, induced by strong nonlinearities.
2. In relation to PF, the performance of EnKF saturates beyond a certain ensemble size due to Gaussian closure inherent to the analysis step. On the other hand, the performance of PF increases steadily with larger ensemble sizes.
3. Even in the presence of significant model and measurement noise, the nonlinear stiffness coefficients of the Duffing system can be estimated with reasonable accuracy by EKF, EnKF and PF for system identification purposes, albeit the need for a larger ensemble size for PF is pointed out.

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. The second author acknowledges the support of a Discovery Grant from National 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 computing infrastructure is supported by the Canadian Foundation of Innovation (CFI) and the Ontario Innovation Trust (OIT). The authors would like to thank two anonymous reviewers for their comments which improved the manuscript.

References

1. Evensen, G.: Data Assimilation: The Ensemble Kalman Filter. Springer, Berlin (2006)
2. Kaipio, J., Somersalo, E.: Statistical and Computational Inverse Problems. Springer, New York (2005)
3. Bennett, A.F.: Inverse Modeling of the Ocean and the Atmosphere. Cambridge University Press, Cambridge (2002)
4. Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng. **82**, 35–45 (1960)
5. Kalman, R.E., Bucy, R.C.: New results in linear filtering and prediction theory. J. Basic Eng. **83**, 95–108 (1961)
6. Cohn, S.E.: An introduction to estimation theory. J. Meteorol. Soc. Jpn. **75**(1B), 257–288 (1997)
7. Jazwinski, A.H.: Stochastic Processes and Filtering Theory. Academic, San Diego (1970)
8. Hoshiya, M., Saito, E.: Structural identification by extended Kalman filter. ASCE J. Eng. Mech. **110**(12), 1757–1770 (1984)
9. Koh, C.G., See, L.M.: Identification and uncertainty estimation of structural parameters. ASCE J. Eng. Mech. **120**(6), 1219–1236 (1991)
10. Miller, R.N., Ghill, M., Gauthiez, F.: Advanced data assimilation in strongly nonlinear dynamical systems. J. Atmos. Sciences **51**(8), 1037–1056 (1994)
11. Loh, C.H., Tou, I.C.: A system-identification approach to the detection of changes in both linear and nonlinear structural parameters. Earthquake Eng. Struct. Dyn. **24**(1), 85–97 (1995)
12. Corigliano, A., Mariani, S.: Parameter identification of a time-dependent elastic-damage interface model for the simulation of debonding in composites. Compos. Sci. Technol. **61**(2), 191–203 (2001)
13. Corigliano, A., Mariani, S.: Parameter identification in explicit structural dynamics: performance of the extended Kalman filter. Comput. Methods Appl. Mech. Eng. **193**(36–38), 3807–3835 (2004)
14. Evensen, G.: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. J. Geophys. Res. Oceans **97**(C11), 17905–17924 (1992)
15. Gauthier, P., Courtier, P., Moll, P.: Assimilation of simulated wind lidar data with a Kalman filter. Mon. Weather Rev. **121**(6), 1803–1820 (1993)
16. Uhlmann, J.K.: Dynamic map building and localization: New theoretical foundations. PhD thesis, University of Oxford, Oxford, U.K. (1995)
17. Julier, S.J., Uhlmann, J.K.: A new extension of the Kalman filter to nonlinear systems. In: Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Multi Sensor Fusion, Tracking and Resource Management, Orlando, FL (1997)
18. Sitz, A., Schwarz, U., Kurths, J., Voss, H.U.: Estimation of parameters and unobserved components for nonlinear systems from noisy time series. Phys. Rev. E **66**, 016210 (2002)
19. Mariani, S., Ghisi, A.: Unscented Kalman filtering for nonlinear structural dynamics. Nonlinear Dyn. **49**, 131–150 (2007)
20. Wan, E.A., van der Merwe, R.: The unscented Kalman filter. In: Kalman Filtering and Neural Networks. Wiley, New York (2001), chapter 7
21. LeDimet, F.X., Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations—theoretical aspects. Tellus Ser. A Dyn. Meteorol. Oceanogr. **38**(2), 97–110 (1986)
22. Gauthier, P.: Chaos and quadri-dimensional data assimilation—a study based on the Lorenz model. Tellus Ser. A Dyn. Meteorol. Oceanogr. **44A**(1), 2–17 (1992)
23. Fisher, M., Leutbecher, M., Kelly, G.A.: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. Q. J. R. Meteorol. Soc. **131**(613), 3235–3246 (2005)

24. van Scheltinga, A.D.T., Dijkstra, H.A.: Nonlinear data-assimilation using implicit models. Nonlinear Process. Geophys. **12**(4), 515–525 (2005)
25. Errico, R.M.: What is an adjoint model? Bull. Am. Meteorol. Soc. **78**(11), 2577–2591 (1997)
26. Annan, J.D., Lunt, D.J., Hargreaves, J.C., Valdes, P.J.: Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. Nonlinear Process. Geophys. **12**(3), 363–371 (2005)
27. Lea, D.J., Allen, M.R., Haine, T.W.N.: Sensitivity analysis of the climate of a chaotic system. Tellus Ser. A Dyn. Meteorol. Oceanogr. **52**(5), 523–532 (2000)
28. Kohl, A., Willebrand, J.: An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models. Tellus Ser. A Dyn. Meteorol. Oceanogr. **54**(4), 406–425 (2002)
29. Lorenc, A.C.: The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. Q. J. R. Meteorol. Soc. **129**(595), 3183–3203 (2003)
30. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. **99**(C5), 10143–10162 (1994)
31. Annan, J.D.: Parameter estimation using chaotic time series. Tellus Ser. A Dyn. Meteorol. Oceanogr. **57**(5), 709–714 (2005)
32. Ghanem, R., Ferro, G.: Health monitoring for strongly nonlinear systems using the ensemble Kalman filter. Struct. Control Health Monit. **13**(1), 245–259 (2006)
33. Pham, D.T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems. Mon. Weather Rev. **129**(5), 1194–1207 (2001)
34. van Leeuwen, P.J.: A variance-minimizing filter for large-scale applications. Mon. Weather Rev. **131**(9), 2071–2084 (1998)
35. Kim, S., Eyink, G.L., Restrepo, J.M., Alexander, F.J., Johnson, G.: Ensemble filtering for nonlinear dynamics. Mon. Weather Rev. **131**(11), 2586–2594 (2003)
36. Doucet, A., Godsill, S.J., Andrieu, C.: On sequential Monte Carlo sampling methods for Bayesian filtering. Stat. Comput. **10**(3), 197–208 (2000)
37. Manohar, C.S., Roy, D.: Monte Carlo filters for identification of nonlinear structural dynamical systems. Sadhana Acad. Proc. Eng. Sci. **31**, 399–427 (2006), Part 4
38. Namdeo, V., Manohar, C.S.: Nonlinear structural dynamical system identification using adaptive particle filters. J. Sound Vib. **306**, 524–563 (2007)
39. Ghosh, S.J., Manohar, C.S., Roy, D.: A sequential importance sampling filter with a new proposal distribution for state and parameter estimation of nonlinear dynamical systems. Proc. R. Soc. Lond. Ser. A **464**, 25–47 (2007)
40. Kivman, G.A.: Sequential parameter estimation for stochastic systems. Nonlinear Process. Geophys. **10**(3), 253–259 (2003)
41. Tanizaki, H.: Nonlinear Filters: Estimation and Applications, 2nd edn. Springer, Berlin (1996)
42. Burgers, G., van Leeuwen, P.J., Evensen, G.A.: Analysis scheme in the ensemble Kalman filter. Mon. Weather Rev. **126**(6), 1719–1724 (1998)
43. van Leeuwen, P.J.: Comment on data assimilation using an ensemble Kalman filter technique. Mon. Weather Rev. **127**(6), 1374–1377 (1999)
44. Miller, R.N., Carter, E.F., Blue, S.T.: Data assimilation into nonlinear stochastic models. Tellus Ser. A Dyn. Meteorol. Oceanogr. **51**(2), 167–194 (1999)
45. Anderson, J.L., Anderson, S.L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev. **127**(12), 2741–2758 (1999)
46. Bengtsson, T., Snyder, C., Nychka, D.: Toward a nonlinear ensemble filter for high-dimensional systems. J. Geophys. Res. Atmos. **108**(D24), 8775 (2003)
47. Mandel, J., Beezley, J.D.: Predictor-corrector and morphing ensemble filters for the assimilation of sparse data into high-dimensional nonlinear systems. In: Proceedings of the IOAS-AOLS, 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, p. 4.12, San Antonio, TX (2007)
48. Ristic, B., Arulampalam, S., Gordon, N.: Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House, Boston (2004)
49. Doucet, A., de Freitas, N., Gordon, N. (eds.): Sequential Monte Carlo Methods in Practice. Springer, New York (2001)
50. Gordon, N.J., Salmond, D.J., Smith, A.F.M.: Novel approach to nonlinear non-Gaussian Bayesian state estimation. IEE Proc. F Radar Signal Process. **140**(2), 107–113 (1993)
51. Handschin, J.E.: Monte Carlo techniques for prediction and filtering of non-linear stochastic processes. Automatica **6**, 555–563 (1970)
52. Akashi, H., Kumamoto, H., Nose, K.: Application of Monte Carlo method to optimal control for linear-systems under measurement noise with Markov dependent statistical property. Int. J. Control **22**(6), 821–836 (1975)
53. Zaritskii, V.S., Svetnik, V.B., Smihelevich, L.I.: Monte Carlo techniques in problems of optimal information processing. Autom. Remote Control **12**, 95–103 (1975)
54. Carpenter, J., Clifford, P., Fearnhead, P.: Improved particle filter for nonlinear problems. IEE Proc. Radar Sonar Navig. **146**(1), 2–7 (1999)
55. Crisan, D., Del Moral, P., Lyons, T.: Discrete filtering using branching and interacting particle systems. Markov Process. Relat. Fields **5**(3), 293–318 (1999)
56. Del Moral, P.: Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications. Springer, Berlin (2004)
57. Ching, J., Beck, J.L., Porter, K.A.: Bayesian state and parameter estimation of uncertain dynamical systems. Probab. Eng. Mech. **21**, 81–96 (2006)
58. Kunin, I., Chen, G.: Controlling the Duffing oscillator to the Lorenz system and generalizations. In: Proceedings of the COC, 2nd International Conference on Control of Oscillations and Chaos, pp. 229–231, St. Petersburg, Russia (2000)
59. Guchenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Field. Springer, New York (1983)
60. Lin, Y.K., Cai, G.Q.: Probabilistic Structural Dynamics. McGraw–Hill, New York (2004)
61. Lin, Y.K.: Probabilistic Theory of Structural Dynamics. McGraw–Hill, New York (1967)

62. Gardiner, C.W.: Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Springer, Berlin (1985)
63. Soize, C.: The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions
64. Haykin, S. (ed.): Kalman Filtering and Neural Networks. Wiley, New York (2001)
65. Chui, C.K., Chen, G.: Kalman Filtering with Real-time Applications, 3rd edn. Springer, Berlin (1999)
66. Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. **53**(4), 343–367 (2003)
67. Whitaker, J.S., Hamill, T.M.: Ensemble data assimilation without perturbed observations. Mon. Weather Rev. **130**(7), 1913–1924 (2002)
68. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems. Wiley, Winston (1977)
69. Mckay, M., Beckman, R., Conover, W.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **21**, 239–245 (1979)
70. Olsson, A., Sandberg, G.: Latin hypercube sampling for stochastic finite element analysis. J. Eng. Mech. **128**, 121–125 (2002)
71. Kong, A., Liu, J.S., Wong, W.H.: Sequential imputations and Bayesian missing data problems. J. Am. Stat. Assoc. **89**, 278–288 (1994)
72. Constantinescu, E.M., Sandu, A., Chaib, T., Carmichael, G.R.: Ensemble-based chemical data assimilation. I: General approach. Q. J. R. Meteorol. Soc. **133**(626), 1229–1243 (2007)