


Article

Parameter Estimation of Nonlinear Structural Systems Using Bayesian Filtering Methods

Kalil Erazo 

Department of Civil and Environmental Engineering, Rice University, Houston, TX 77005, USA; ke11@rice.edu

Abstract: This paper examines the performance of Bayesian filtering system identification in the context of nonlinear structural and mechanical systems. The objective is to assess the accuracy and limitations of the four most well-established filtering-based parameter estimation methods: the extended Kalman filter, the unscented Kalman filter, the ensemble Kalman filter, and the particle filter. The four methods are applied to estimate the parameters and the response of benchmark dynamical systems used in structural mechanics, including a Duffing oscillator, a hysteretic Bouc–Wen oscillator, and a hysteretic Bouc–Wen chain system. Based on the performance, accuracy, and computational efficiency of the methods under different operating conditions, it is concluded that the unscented Kalman filter is the most effective filtering system identification method for the systems considered, with the other filters showing large estimation errors or divergence, high computational cost, and/or curse of dimensionality as the dimension of the system and the number of uncertain parameters increased.

Keywords: system identification; extended Kalman filter; unscented Kalman filter; ensemble Kalman filter; particle filter

1. Introduction

To design and evaluate the performance of structural and mechanical systems it is essential to reliably predict features of the system behavior and response characteristics. Physics-based mechanics models are typically employed to predict the response of complex large-scale structural systems, and the features of the response that can be captured by parameterized mechanics-based models depend on their functional form and the adopted parameters. Moreover, a set of parameters rather than point values might properly characterize structural behavior due to the presence of inherent modeling uncertainty, unmeasured input excitations, and/or unknown initial conditions [1].

System identification (SI) has emerged as a primary engineering analysis tool aimed at improving the predicting capability of models from measured response data [2]. SI can be defined as the inverse problem of using measurements to infer the state and parameters of models that best fit the observed data. The objective is to maximize the predicting accuracy and capabilities of a model by minimizing modeling errors and uncertainty. Applications of SI in structural and mechanical systems include structural condition assessment and management, structural damage diagnosis/prognosis, response control, improvement of computer-aided design methods, and enhancement of experimental testing techniques, among others. In the context of linear systems, an extensive number of SI algorithms have been developed over the past decades [3]. In some applications, the features of the dynamic response of the system of interest cannot be captured by a linear model. Sources of nonlinearity include nonlinear kinematic or geometric response features,



Academic Editor: Atef Mohany

Received: 22 October 2024

Revised: 11 December 2024

Accepted: 26 December 2024

Published: 31 December 2024

Citation: Erazo, K. Parameter Estimation of Nonlinear Structural Systems Using Bayesian Filtering Methods. *Vibration* **2025**, *8*, 1. <https://doi.org/10.3390/vibration8010001>

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

nonlinear material or constitutive behavior, nonlinear boundary conditions, nonlinear modeling of energy dissipation mechanisms or devices, and modeling of actuators, among others [4]. Moreover, nonlinear behavior and phenomena can be exploited to design systems with enhanced performance and energy harvesting capabilities. Estimation of the model parameters of nonlinear systems presents an increased challenge with respect to the linear systems counterpart, in part because of the complexity of modeling nonlinear phenomena and the lack of a closed-form general input–output mapping that applies to all nonlinear systems.

SI methods can be broadly classified as deterministic or probabilistic. Deterministic methods rely on an optimization-based strategy to minimize an objective function defined as a measure of the discrepancy between measured responses and model predictions [5,6]. On the other hand, probabilistic methods use a stochastic model of the system treating the uncertain parameters as random variables and using statistical inference (frequentist or Bayesian) to estimate the parameters; in particular Bayesian methods have received significant attention due to their ability to rigorously handle a broad class of estimation problems in linear and nonlinear systems. In Bayesian SI, the information about the parameters contained in measurements is integrated using conditional probability distributions obtained using Bayes' theorem; the estimate of the parameters and their uncertainty are encapsulated in a probability distribution conditional on the measured data [7–10].

In addition to the estimation approach adopted, SI can be performed either considering a set of response measurements (batch estimation) or in real-time with the estimation performed every time step that a new measurement is available (recursively/sequential estimation). In dynamical systems, when the estimation process is performed recursively the resulting procedure is known as filtering [11]. In the filtering setting the estimation is performed in a predictor–corrector fashion that involves a forward projection of the estimate of the response and the parameters using a model (prediction), and subsequently integrating the measurements (correction). Filtering was formalized using probabilistic methods in the context of optimal linear filtering [12]. The most celebrated method is the Kalman filter, a recursive estimation algorithm that allows for optimal estimation (unbiased, minimum variance, and minimum mean-squared error) of the state of a linear system subjected to disturbing inputs modeled as a stochastic process by combining model predictions and response measurements.

Filtering can be applied to system identification by including uncertain parameters in the state, and the resulting procedure is known as joint state-parameter estimation or augmented state estimation [13]. When the underlying system is linear including the parameters in the state results in a nonlinear estimation problem, and nonlinear filtering methods are needed; thus, the parameter estimation problem is nonlinear irrespective of the underlying dynamics being linear or nonlinear, but with significantly stronger nonlinearities in the latter case. In the Bayesian solution of the nonlinear filtering problem, a complete functional description of the conditional probability density describing the evolution of the state is needed. Despite the possibility, in theory, to compute this evolving density function, computational limitations both in algorithmic running time and storage limits prevent the Bayesian nonlinear filtering problem from being solved in closed form. Under these conditions the problem admits a wider family of sub-optimal solutions.

The first extension of linear filtering theory to nonlinear problems proposed in the literature is the extended Kalman filter (EKF), which became the de facto accepted solution to the problem for several decades. The EKF is a recursive estimation algorithm that follows from approximating nonlinear transformations of the state distribution by linearizing the

nonlinear model and using the Kalman filter on the linearized model [13]. The EKF is neither an unbiased minimum variance, nor the minimum mean-squared error estimate of the state, and its two main drawbacks are the computational issues related to the propagation of the covariance matrix, and errors resulting from the linearization performed where high-order nonlinear terms are neglected. The result is a loss of accuracy and potential filter divergence. The EKF also requires determining the Jacobian of the model, a task that is generally computationally intensive.

To overcome the main limitations of the EKF, recent efforts have focused on the development of more advanced nonlinear filtering methods that show improved performance. The three most well-established new methods are the unscented Kalman filter (UKF) [14], the ensemble Kalman filter (EnKF) [15], and the particle filter (PF) [16]. The Kalman filtering-based nonlinear methods approximate the distributions involved as Gaussian, with their mean and covariance estimated using a deterministic sample (for the UKF) or a random sample (for the EnKF). On the other hand, particle filters use a random sample to directly approximate the filtering distribution by a probability mass function. This implies that Kalman filtering methods are suitable only when the distributions involved are unimodal, while particle filters can handle distributions with non-Gaussian features, such as high skewness, heavy tails, or multimodality. The aforementioned methods and some of their variations have found applications in structural mechanics, earthquake engineering and structural dynamics, structural damage assessment, condition monitoring and performance prediction [17–44].

The objective of this paper is to present a comprehensive case study of the performance of Bayesian filtering methods in structural and mechanical systems, studying the accuracy and efficiency of the EKF, UKF, EnKF, and PF under the same operating conditions. Despite a significant body of work in analyzing the performance of the individual methods, there is no current guideline regarding which method would have more promising performance under different identification constraints. Such guidelines depend on a performance assessment of the four aforementioned algorithms under the same conditions in the context of nonlinear structural and mechanical systems.

2. Modeling of Nonlinear Dynamical Systems and State-Space Identification Model

This paper focuses on nonlinear structural and mechanical systems whose dynamic behavior is modeled by the following equation:

$$M\ddot{q}(t) + C(\theta)\dot{q}(t) + F_R(q(t), \dot{q}(t), \theta) = u(t) \quad (1)$$

where $q(t)$ is the displacement vector at time t , M is the mass matrix, C is the damping matrix, and F_R is the restoring force; a dot on top of a variable indicates differentiation with respect to time. The damping matrix and the restoring force are completely or partially parameterized by the uncertain parameters vector θ . The forcing input u can be deterministic or stochastic, and, in the latter case, Equation (1) becomes a stochastic differential equation to be interpreted using Itô's definition of a stochastic differential.

The class of system identification methods studied herein operate using state-space models. Moreover, when the uncertain parameters vector θ needs to be estimated the dynamic state is augmented to include these parameters. To perform joint state-parameter estimation the model defined by Equation (1) can be written in state-space form by defining the augmented state as $x = [q \ \dot{q} \ \theta] \in \mathbb{R}^n$ resulting in a model of the following form:

$$\dot{x}(t) = f(x(t)) + Du(t) \quad (2)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defines the combined system dynamics and parameter estimation model, and D maps the input to the augmented state-space. The objective of joint state-parameter estimation is to infer the augmented state (which includes the uncertain parameters) by integrating the identification model with noise-contaminated response measurements modeled as follows:

$$y(t) = h(x(t)) + v(t) \quad (3)$$

where $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ maps the augmented state to the measurements space, and it is defined depending on the type of response measurement; $v(t)$ is the measurement noise modeled as a zero-mean Gaussian white process.

3. Parameter Estimation Using Nonlinear Bayesian Filtering

The objective of parameter estimation is to infer the uncertain vector θ that (completely or partially) parameterizes the model from response measurements/data. Bayesian inference provides a consistent and robust framework to tackle estimation rigorously in a probabilistic setting for a broad class of problems. In the Bayesian framework the estimate of the parameters is characterized by a probability density function conditional on the available data. Bayesian estimators known as nonlinear filtering methods incorporate an array of discrete measurements $Y_k = \{y_0 \dots y_k\}$ where $y_i = y(t_i)$ and extract the information contained in the measurements by processing the data as it becomes available. The framework employs a predictor-corrector framework where the prediction and correction follow from the Theorem of Total Probability and Bayes' Theorem, given respectively by the following:

$$p(x_{k+1}|Y_k) = \int p(x_{k+1}|x_k, Y_k)p(x_k|Y_k)dx_k = \int p(x_{k+1}|x_k)p(x_k|Y_k)dx_k \quad (4)$$

and

$$p(x_{k+1}|Y_{k+1}) = \frac{p(y_{k+1}|x_{k+1}, Y_k)p(x_{k+1}|Y_k)}{p(y_{k+1}|Y_k)} = \frac{p(y_{k+1}|x_{k+1})p(x_{k+1}|Y_k)}{p(y_{k+1}|Y_k)} \quad (5)$$

where $x_{k+1} = x(t_{k+1})$, the prior/predictive probability density function $p(x_{k+1}|Y_k)$ follows from projecting the state using the dynamics model, and the likelihood function $p(y_{k+1}|x_{k+1})$ is found using the response measurements model. The challenge with recursive Bayesian estimation is that the posterior projection defined by Equation (4) and used to obtain the predictive distribution and used as the prior distribution in the update step in Equation (5) requires either solving a Fokker–Planck equation (when the stochastic model is in continuous form) or solving the high-dimensional integral in Equation (4), both of which are analytically intractable and computationally prohibitive for most applications.

The main approach to address this issue has been assuming that all the distributions are Gaussian and using a linearized model to estimate the mean and covariance of the predictive probability distribution, resulting in the extended Kalman filter (EKF) [13]. This approach provides acceptable results in differentiable systems that do not show significant excursions to nonlinear response regimes, but the estimates tend to show divergence in cases of nonlinear problems that involve a parameter space of from low to moderate dimensions. Two approaches have been sought in the literature to address the limitations of the EKF: (i) improved approximate parameterizations of the predictive distribution that do not linearize the model, and (ii) approximating directly the posterior distribution using a discrete sample. In the first class of methods the predictive distribution is generally approximated by a Gaussian distribution, reducing the problem of estimating a complete probability distribution to the estimation of only the first two moments either from a deterministic or a random sample. A recently proposed deterministic approach to estimate

these moments is based on using the Unscented Transform (UT), a sampling-based method that guarantees second-order accuracy for any functional form of the model or type of nonlinearity; the application of the UT in filtering problems is known as the Unscented Kalman Filter (UKF). Another approach of this class uses a random sample and a Monte Carlo estimate of the mean and the covariance of the predictive distribution, resulting in the Ensemble Kalman Filter (EnKF). The second class of methods, known as particle filters, attempt to directly approximate/estimate the posterior distribution using a weighted sample. In contrast to the first class of methods that assume a Gaussian distribution for all the distributions involved, in particle filters the posterior distribution is directly estimated without assuming a functional form.

3.1. Extended Kalman Filter (EKF)

The EKF has been the established method for filtering in nonlinear systems. The method follows from approximating the projected mean \hat{x}_k^- and covariance $\hat{P}_{x_k x_k}^-$ of nonlinear transformations of the state using a linearized model based on a truncated Taylor series expansion. The linearized projection is used in conjunction with the Kalman filter as follows:

$$\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-) \tag{6}$$

$$\hat{P}_{x_k x_k} = \hat{P}_{x_k x_k}^- - K_k \hat{P}_{y_k y_k}^- K_k^T \tag{7}$$

$$K_k = \hat{P}_{x_k y_k}^- \left(\hat{P}_{y_k y_k}^- \right)^{-1} \tag{8}$$

where \hat{x}_k and $\hat{P}_{x_k x_k}$ are, respectively, estimates of the mean and covariance of the state posterior distribution defined by Equation (5), y_k is the measured response, and K_k is the Kalman gain, which serves as a weight between model predictions and measured responses and leverages process noise and measurement noise. The Kalman filtering equations result in the optimal linear estimator for the mean and covariance of the state posterior distribution for both linear and nonlinear systems if only the mean and covariance of the distributions involved are considered. The difficulty in applying the Kalman filtering approach with nonlinear models is that computing the projected/prior mean \hat{x}_k^- and covariance $\hat{P}_{x_k x_k}^-$ of the state prior defined by Equation (4) requires computing the moments of nonlinear transformations of random variables, which, in general, cannot be computed in closed-form, particularly when the state dimension is relatively large. To estimate the prior mean and covariance the EKF uses a first-order truncated Taylor series, resulting in the following [13]:

$$\hat{x}_k^- \approx f_d(\hat{x}_{k-1}) \tag{9}$$

$$\hat{P}_{x_k x_k}^- \approx \nabla f_d |_{\hat{x}_{k-1}} \hat{P}_{x_{k-1} x_{k-1}} \nabla f_d^T |_{\hat{x}_{k-1}} + B_d Q_{d_{k-1}} B_d^T \tag{10}$$

where f_d is a numerical discretization of the dynamic model, and ∇ is the gradient operator. The term $B_d Q_{d_{k-1}} B_d^T$ in Equation (10) is the projected forcing input covariance, where $Q_{d_{k-1}}$ is the input noise covariance and R is the measurement noise covariance. The limitations of this approach include it being a first-order method, which can lead to large estimation errors and divergence of the filter, and the computation of the gradient which can be computationally prohibitive in high-dimensional problems. The following pseudocode shows the overall steps in the implementation of the EKF (Algorithm 1):

Algorithm 1: EKF

1. Define state initial/prior mean \hat{x}_0 and covariance $\hat{P}_{x_0x_0}$
 2. Compute the model gradient and evaluate it in the initial/prior estimate $\nabla f_d|_{\hat{x}_0}$
 3. Compute the measurement gradient (if nonlinear) and evaluate it in the initial/prior estimate $\nabla h|_{\hat{x}_0}$; note that if h is linear $h(x) = Hx$
 4. For $k = 1, \dots, N$
 - 4.1 Perform the projection step

$$\hat{x}_k^- = f_d(\hat{x}_{k-1})$$

$$\hat{P}_{x_kx_k}^- = \nabla f_d|_{\hat{x}_{k-1}} \hat{P}_{x_{k-1}x_{k-1}} \nabla f_d^T|_{\hat{x}_{k-1}} + B_d Q_{d_{k-1}} B_d^T$$
 - 4.2 Perform the update step

$$K_k = (\hat{P}_{x_kx_k}^-)^{-1} H^T (H \hat{P}_{x_kx_k}^- H^T + R)^{-1}$$

$$\hat{P}_{x_kx_k} = \hat{P}_{x_kx_k}^- - K_k \hat{P}_{y_ky_k}^- K_k^T$$

$$\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-)$$
-

3.2. Unscented Kalman Filter (UKF)

The UKF is a recursive estimation method based on the use of the Unscented Transform (UT) to estimate the first two moments of the predictive distribution [45]. The UT uses a deterministic sample known as sigma points to estimate moments of transformations of a random variable [14]. Let $p(x_k|Y_k)$ represent the probability distribution of the state conditional on all measurements available up to time step t_k and let \hat{x}_k and $\hat{P}_{x_kx_k}$ denote, respectively, the mean and covariance of $p(x_k|Y_k)$. To estimate the mean and covariance of the predictive distribution $p(x_{k+1}|Y_k)$ the UT uses the following deterministic sample:

$$\chi_i = \begin{cases} \hat{x}_k & i = 0 \\ \hat{x}_k + \left(\sqrt{(N + \lambda) \hat{P}_{x_kx_k}} \right)_i & i = 1, \dots, n \\ \hat{x}_k - \left(\sqrt{(N + \lambda) \hat{P}_{x_kx_k}} \right)_i & i = n + 1, \dots, 2n \end{cases} \quad (11)$$

where n is the dimension of the state and $\left(\sqrt{\hat{P}} \right)_i$ denotes the i -th column of the matrix square root of \hat{P} . The parameter λ is used to adjust the scaling of the sample and it is tuned off-line using the analytical model; choosing $\lambda = 3 - n$ ensures second-order accuracy in the first and second moment estimates. The sigma vectors are projected to the next time step using the model, and the mean and covariance of $p(x_{k+1}|Y_k)$ are estimated as a weighted sum of the projected vectors, defined by the following:

$$\hat{x}_k^- = \sum_{i=0}^{2n} W_i \chi_i^- \text{ where } \chi_i^- = f_d(\chi_i) \quad (12)$$

$$\hat{P}_{x_kx_k}^- = \sum_{i=0}^{2n} W_i (\chi_i^- - \hat{x}_k^-) (\chi_i^- - \hat{x}_k^-)^T \quad (13)$$

where the weights are given by $W_0 = \lambda / (n + \lambda)$ and $W_i = 1 / 2(n + \lambda)$ for $i \neq 0$. The sample defined by the sigma points in Equation (11) might result in a non-positive definite covariance when the dimension of the state is larger than 3. To address this issue the scaled unscented transform was proposed. In the modified algorithm a new set of sigma points is obtained using an auxiliary nonlinear transformation on the original points which guarantees positive definiteness of the covariance. In the scaled unscented transform approach the new sigma points and weights are computed as $\chi_i^* = \chi_0 + \alpha(\chi_i - \chi_0)$ with

the associated weights given by $W_0^* = (1/\alpha^2)W_0 + (1 - 1/\alpha^2)$ and $W_i^* = (1/\alpha^2)W_i$ for $i \neq 0$.

The implementation of the UKF can be summarized in the following pseudocode (Algorithm 2):

Algorithm 2: UKF

1. Define state initial/prior mean \hat{x}_0 and covariance $\hat{P}_{x_0x_0}$
 2. Select the scaled UT parameter α , and compute the weights $\{W_i^*\}$
 3. Compute the initial set of sigma points $\{\chi_i\}$ based on Steps 1–2
 4. For $k = 1, \dots, N$
 - 4.1 Perform the projection step

$$\chi_i^- = f_d(\chi_i) \text{ and } y_i^- = h(\chi_i)$$

$$\hat{x}_k^- = \sum_{i=0}^{2n} W_i^* \chi_i^- \text{ and } \hat{y}_k^- = \sum_{i=0}^{2n} W_i^* y_i^-$$

$$\hat{P}_{x_k x_k}^- = \sum_{i=0}^{2n} W_i^* (\chi_i^- - \hat{x}_k^-)(\chi_i^- - \hat{x}_k^-)^T + B_d Q_{d_{k-1}} B_d^T$$

$$\hat{P}_{y_k y_k}^- = \sum_{i=0}^{2n} W_i^* (y_i^- - \hat{y}_k^-)(y_i^- - \hat{y}_k^-)^T$$

$$\hat{P}_{x_k y_k}^- = \sum_{i=0}^{2n} W_i^* (\chi_i^- - \hat{x}_k^-)(y_i^- - \hat{y}_k^-)^T$$
 - 4.2 Perform the update step

$$K_k = \hat{P}_{x_k y_k}^- (\hat{P}_{y_k y_k}^-)^{-1}$$

$$\hat{P}_{x_k x_k} = \hat{P}_{x_k x_k}^- - K_k \hat{P}_{y_k y_k}^- K_k^T$$

$$\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-)$$

Compute a new set of sigma points $\{\chi_i\}$ using \hat{x}_k and $\hat{P}_{x_k x_k}$
-

3.3. Ensemble Kalman Filter (EnKF)

The EnKF is a filtering method that addresses two issues in state estimation for nonlinear systems in high-dimensional spaces: (i) the computational issues related to the propagation of the covariance matrix, and (ii) the closure problem resulting from neglecting higher order terms in the state error covariance matrix propagation [15]. The EnKF uses the Kalman filter framework with the mean and covariance of the predictive distribution computed using sample-based statistical estimates based on an ensemble/random sample of states.

Let $X_{k|k-1} = \{x_{k|k-1}^{(1)}, x_{k|k-1}^{(2)}, \dots, x_{k|k-1}^{(N)}\}$ denote a prior ensemble of N states obtained by random sampling the posterior at step k and projecting each individual sample using the dynamic model. The ensemble covariance matrix of the projected distribution is obtained using the sample-based estimate as follows:

$$\bar{P}_{k|k-1} = \sum_{i=1}^N \frac{(x_{k|k-1}^{(i)} - \bar{x}_{k|k-1})(x_{k|k-1}^{(i)} - \bar{x}_{k|k-1})^T}{N - 1} \tag{14}$$

where $\bar{x}_{k|k-1}$ is the sample mean. Using the measurement y_k at time step k , a perturbed measurements matrix is defined by the following:

$$Y_k = \{y_k + \varepsilon_k^1, y_k + \varepsilon_k^2, \dots, y_k + \varepsilon_k^N\} \tag{15}$$

where ε_k^i is the i^{th} realization of a zero mean Gaussian white computed based on the noise process parameters. The ensemble measurement error covariance matrix is computed using the sample-based estimate as follows:

$$\bar{R}_k = \sum_{i=1}^N \frac{\varepsilon_k^{(i)} \varepsilon_k^{(i)T}}{N-1} \tag{16}$$

The posterior ensemble is computed using the prior sample adjusted by the prediction error using Kalman filter type estimate given by the following:

$$X_{k|k} = X_{k|k-1} + \bar{P}_{k|k-1} H^T (H \bar{P}_{k|k-1} H^T + \bar{R}_k)^{-1} (Y_k - H X_{k|k-1}) \tag{17}$$

where the matrix H is defined based on a linear measurements model of the form $h(x) = Hx$. Equation (17) provides an ensemble/sample estimate of the posterior distribution from which the mean and covariance matrix are estimated. The method is computationally demanding since the size of the ensemble needed is generally large.

For a sufficiently large sample the EnKF is expected to outperform all other Kalman filtering based methods (including the EKF and UKF) since unbiased sample-based estimates converge to the exact parameters values in the limit $N \rightarrow \infty$ with a convergence rate of the order $\frac{1}{\sqrt{N}}$, although, in practice, there are computational limitations to the sample size that can be employed particularly for systems with a large augmented state-space. Similarly to the EKF and UKF, the approach is not appropriate for problems with highly non-Gaussian features, such as distributions with significant skewness and/or heavy tails.

The algorithm implementation of the EnKF can be summarized as follows (Algorithm 3):

Algorithm 3: EnKF

1. Define state initial/prior mean \hat{x}_0 and covariance $\hat{P}_{x_0x_0}$
 2. Select the sample size M, and generate the initial ensemble $\{x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(M)}\}$
 3. For $k = 1, \dots, N$
 - 3.1 Perform the projection step

$$x_i^- = f_d(x_k^{(i)}) \text{ and } y_i^- = h(x_k^{(i)})$$

$$\hat{x}_k^- = \frac{1}{M} \sum_{i=0}^M x_i^- \text{ and } \hat{y}_k^- = \frac{1}{M} \sum_{i=0}^M y_i^-$$

$$\hat{P}_{x_k x_k}^- = \frac{1}{M-1} \sum_{i=0}^M (x_i^- - \hat{x}_k^-)(x_i^- - \hat{x}_k^-)^T + B_d Q_{d_{k-1}} B_d^T$$

$$\hat{P}_{y_k y_k}^- = \sum_{i=0}^M (y_i^- - \hat{y}_k^-)(y_i^- - \hat{y}_k^-)^T$$

$$\hat{P}_{x_k y_k}^- = \frac{1}{M} \sum_{i=0}^M (x_i^- - \hat{x}_k^-)(y_i^- - \hat{y}_k^-)^T$$
 - 3.2 Perform the update step

$$K_k = \hat{P}_{x_k y_k}^- (\hat{P}_{y_k y_k}^-)^{-1}$$

$$\hat{P}_{x_k x_k} = \hat{P}_{x_k x_k}^- - K_k \hat{P}_{y_k y_k}^- K_k^T$$

$$\hat{x}_k^{(i)} = x_i^- + K_k (y_k - \hat{y}_k^-)$$

$$x_{k+1}^{(i)} = \hat{x}_k^{(i)}$$
-

3.4. Particle Filter (PF)

The PF is a filtering algorithm that attempts to estimate the state posterior distribution from at each time step from discrete random sample. The approach does not make any assumptions about the parametric form of the posterior such as the Gaussian assumption made by the Kalman filtering-based methods. This allows for the treatment of problems

where the distributions are multi-modal, heavily skewed, or have heavy tails. Let X_k denote the set/sample of states from the posterior distribution $X_k = \{x_0, x_1, \dots, x_k\}$. Using Bayes theorem, the Markovian nature of the model, and the output model, the filtering distribution is given by the following [46]:

$$p(X_k|Y_k) = \frac{p(x_0)}{p(Y_k)} \prod_{i=1}^k p(y_i|x_i)p(x_i|x_{i-1}) \quad (18)$$

To obtain features of the joint posterior distribution, such as marginal distributions, high-dimensional integrals have to be computed with respect to sub-spaces of the augmented state space, a task that is prohibitive for state spaces of large dimensions. To overcome this issue stochastic simulation methods based on sampling are employed. An approach to obtain a sample from the posterior is sampling from an auxiliary distribution using importance sampling (IS), where expectations of the state

$$E[g(X_k)] = \int g(X)p(X_k|Y_k)dX_k \quad (19)$$

are estimated as

$$\hat{E}[g(X_k)] = \sum_{i=1}^N g(X_k^{(i)}) \tilde{w}_n^{(i)} \quad (20)$$

where

$$\tilde{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{j=1}^M w_n^{(j)}} \quad w_n^{(i)} = \frac{p(Y_k|X_k)p(x_k)}{\pi(X_k|Y_k)} \quad (21)$$

and the independent samples are drawn from π , the importance distribution function. Under the premise that the support of π includes the support of $p(X_k|Y_k)$ and $E[g(X_k)] < \infty$, the estimator $\hat{E}[g(X_k)]$ is guaranteed to converge almost surely to the true $E[g(X_k)]$, with the rate of convergence strongly depending on the choice of importance function. The selection of the importance function of prime importance for the PF algorithm to be accurate. Because of computational efficiency constraints, it is desired to use importance functions that result in a recursive algorithm, allowing samples from a time step to be used in the following step without having to re-sample on each step, which would be practically unfeasible for the sampling rate of interest in systems with fast dynamics. Thus, for convenience, the sampling distribution is defined by the (unconditional) forward model $p(x_i|x_{i-1})$, resulting in a recursive algorithm.

The main limitation of this approach is that the importance functions that allow the algorithm to be implemented in a recursive fashion, do so at the expense of increasing the variance of the weights [16]. This implies that after a number of analysis steps, which in practice is relatively low when compared to both the model discretization time step and measurements sampling rate, the effective sample size decreases and ultimately degenerates to a single point eliminating the ability of the sample to represent the distribution. This issue is exacerbated when the measurement noise is small, resulting in a likelihood function that decays fast. Some strategies have been proposed to alleviate this issue, but not to completely eliminate it. The most popular approach is to perform a re-sampling step when the weights variance exceeds a threshold, replicating samples with high weights and discarding samples with low weights. The drawback of the re-sampling step is that the samples become correlated and can no longer be considered an independent sample, increasing the variance of the estimates.

Despite using a re-sampling strategy and other proposed approaches discussed in Ref. [16], for dynamical systems with sampling rates higher than 100 Hz (which is the case in most practical applications), the variance of the weights increases rapidly and the sample collapse limiting the accuracy of the PF. The development of approaches to improve the

performance of the PF remains an open research area. The parallelized PF with resampling (“bootstrap filter”) can be implemented using the following algorithm (Algorithm 4):

Algorithm 4: PF with resampling

1. Define state initial/prior mean \hat{x}_0 and covariance $\hat{P}_{x_0x_0}$
2. Select the number of particle filters L and the sample size M for each filter, and generate the initial sample $\{x_1^{(1,1)}, x_1^{(1,2)}, \dots, x_1^{(1,M)}, \dots, x_1^{(L,M)}\}$
3. For $j = 1, \dots, L$

Set the weights of filter j to $w_0^{(i)} = 1/M$ where $i = 1, \dots, M$

For $k = 1, \dots, N$

$$x_i^- = f_d(x_k^{(j,i)}) \text{ and } y_i^- = h(x_k^{(j,i)})$$

$$\hat{x}_k^- = \sum_{i=0}^M x_i^- \left(\frac{w_{k-1}^{(i)}}{\sum_{s=1}^M w_{k-1}^{(s)}} \right)$$

$$\hat{P}_{x_k x_k}^- = \sum_{i=0}^M (x_i^-)(x_i^-)^T \left(\frac{w_{k-1}^{(i)}}{\sum_{s=1}^M w_{k-1}^{(s)}} \right) + B_d Q_{d_{k-1}} B_d^T$$

Evaluate the likelihood function for each sample: $p_{y|x}^{(i)} = N(y_i^-, y_k, R)$

Note: $N(a, \mu, \Sigma)$ is a Normal distribution with parameters (μ, Σ) evaluated at a

$$w_k^{(i)} = w_{k-1}^{(i)} \times p_{y|x}^{(i)}$$

$$\hat{x}_k = \sum_{i=0}^M x_i^- \left(\frac{w_k^{(i)}}{\sum_{s=1}^M w_k^{(s)}} \right)$$

$$\hat{P}_{x_k x_k} = \sum_{i=0}^M (x_i^-)(x_i^-)^T \left(\frac{w_k^{(i)}}{\sum_{s=1}^M w_k^{(s)}} \right)$$

$$x_{k+1}^{(j,i)} = x_i^-$$

COV = coefficient of variation of the sample $\{w_k^{(i)}\}$

If COV > 2 (resample step)

Obtain the sample $\{x_{k+1}^{(j,i)}\}$ by random sampling with replacement the

set $\{x_i^-\}$ with probabilities defined by $\{w_k^{(i)}\}$

Set $w_k^{(i)} = 1/M$

Compute the average of \hat{x}_k and $\hat{P}_{x_k x_k}$ across the L filters

4. Numerical Examples

The performance of the EKF, UKF, EnKF, and PF is assessed when the filters are applied to system identification in the context of three systems: a Duffing oscillator, a nonlinear hysteretic oscillator, and a nonlinear hysteretic chain system. The objective is to study the accuracy and computational effort required by the methods under the same operating conditions. The scalability of the methods is also studied, that is, their ability to handle systems of increased complexity with parameter spaces of increasing dimension.

4.1. Duffing Oscillator

The Duffing oscillator dynamics are modeled by the following:

$$m\ddot{q}(t) + c\dot{q}(t) + k_1q(t) + k_2q^3(t) = u(t) \tag{22}$$

where m is the oscillator mass, c is the damping coefficient, and k_1 and k_2 are, respectively, the linear and cubic stiffness coefficients. The Duffing oscillator shows a wide variety of nonlinear response features, and it is applied to model the behavior of many mechanical, structural, and electrical systems, such as the restoring characteristics of deformable solids, buckling of slender members, nonlinearity in circuits, superconductors, and the chaotic behavior of dynamical systems in different domains, among others [47].

The exact/true system parameters used to generate the data are (in consistent SI units) $m = 1$ kg, $c = 0.3$ Ns/m, $k_1 = -1$ N/m and $k_2 = 1$ N/m³. For these parameter values under free vibration, the unforced system has three fixed points, namely, an unstable fixed point at 0 and two stable fixed points at 1 and -1 . We consider the case where the system is driven by the deterministic harmonic input $u(t) = 0.3\cos(1.25t)$ N; under this condition, the system exhibits two period-two subharmonic oscillations [30]. The initial conditions are selected as $q(0) = 1$ and $\dot{q}(0) = 0$. To solve the forward problem and to perform the estimation, the system is discretized using the fourth-order Runge–Kutta method with time step $\Delta t = 0.005$ s.

The objective is to estimate the parameters vector $\theta = \{c \ k_1 \ k_2\}$ assumed to be uncertain/unknown. To perform Bayesian system identification, an initial prior distribution for the parameters needs to be specified. Figure 1 shows the prior distributions adopted and the system's exact/true parameters. Significant uncertainty about the true values of the parameters has been assumed, with the true parameters falling on the tails of the prior distributions and initial errors between 30% and 50%; the algorithms were initialized with different sets of initial estimates, and no significant variation in the results was observed. The measurement data used for estimation consists of a noise-contaminated displacement response, with the noise modeled as Gaussian white noise with a 10% noise-to-signal root-mean-square (RMS). The noise model is consistent with models for this type of sensor [48]. It was observed that the performance of the filters did not change significantly with RMS values of up to 20%.

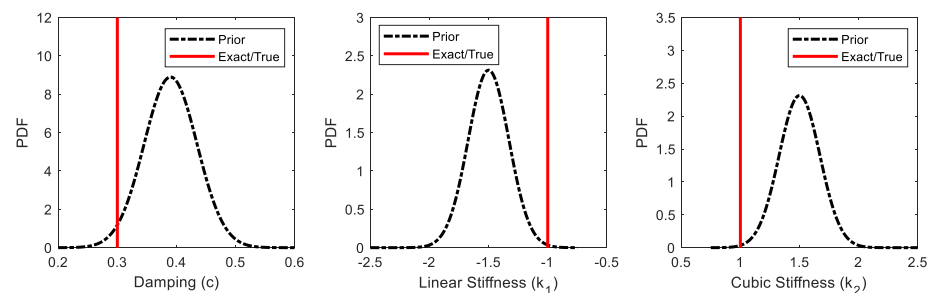


Figure 1. Parameters prior distributions for the Duffing oscillator example (in SI units).

The UKF was implemented using the scaled unscented transform algorithm with the scale parameter selected as $\alpha = 10^{-3}$ based on previous filter performance assessments [45]. The sample size for the EnKF was selected as $N = 2000$; the sample size was selected based on a convergence analysis where the sample size was increased until no significant variation in the estimates was observed. The PF was applied using the sequential importance sampling (SIS) algorithm with resampling, also known as the bootstrap filter [46], combined with the parallelized implementation proposed in Ref. [20] to reduce and delay sample degeneracy. The weight coefficient of the variation threshold for resampling was selected as 3, with 20 parallel particle filters, each running with 4000 particles. For the PF, increasing the sample size reduces sample degeneracy and improves the estimation; however, there are computational limitations to the sample size due to limited computer memory and the required computational time; the sample size of 80,000 total particles selected is close to the maximum that a standard desktop computer memory allows. The use of high-

performance computing systems is needed to further increase the sample size at the expense of a high computational effort. For most practical applications, the availability of such high-performance computing systems is limited.

Figures 2–4 show the estimation results for the oscillator damping, linear stiffness, and cubic stiffness parameters, respectively; the figures show the mean estimate and the uncertainty bounds defined as ± 3 standard deviations, with ‘SYS’ indicating the exact/true system response. The root-mean-square error (RMSE) is used as a measure of the estimation accuracy of the filters; the RMSE of the estimate $\hat{\varphi}$ of the true/exact parameter φ is defined as $\text{RMSE} = \sqrt{\mathbb{E}[(\hat{\varphi} - \varphi)^2]}$ where $\mathbb{E}[\cdot]$ is the expectation operator, and the estimate $\hat{\varphi}$ is a random variable. The RMSE allows us to quantify the estimation accuracy by accounting for both estimation bias and uncertainty, and, for this reason, it is widely adopted as an estimation accuracy metric; if two estimators are unbiased, then the estimator with larger uncertainty would result in a larger RMSE. Figures 2–4 show that the EKF, UKF, and EnKF successfully estimate all the oscillator parameters, with the PF not showing convergence to the exact parameters’ values. It is interesting to note that the EKF provided good estimation results for this example, in agreement with the results reported in Ref. [30]. This is attributed to the time step $\Delta t = 0.005$ used for the simulations, which is considered small for the characteristic time scale of the system’s dynamic behavior, resulting in improved accuracy of the linearization employed by the EKF.

Based on the particle’s coefficient of variation, the PF showed particle degeneracy at approximately $t = 5$ s. This is confirmed by Figures 2–4, and particularly the RMSE, which show that all the filters provide comparable estimates, uncertainty, and accuracy up to $t = 5$ s; after this time, all the particles of the PF collapsed into a single particle, and the sample was no longer able to properly approximate the parameters posterior distribution. The PF RMSE is seen to remain large when compared to the RMSE of the other filters. It is worth pointing out that the main factor defining the onset of particle degeneracy is the number of analysis time steps. In the case of the present example, degeneracy started after 1000 steps. A strategy to attempt to delay the onset of particle degeneracy is to downsample the measurements by not updating the posterior distribution at every time step, that is, not using all the available measurements; this approach was successfully employed in Ref. [30] to improve the accuracy of the PF. This strategy is adopted at the expense of significantly decreasing the sampling frequency, which is not ideal in applications since it is desirable to fully exploit and extract the information available on the measurements. Moreover, downsampling the measurements has an impact on the accuracy and significantly increases the estimate’s uncertainty. Another strategy to delay particle degeneracy is to include a fictitious process noise that inflates the particles’ variance. Although this strategy, in general, improves the estimation accuracy of filtering methods by increasing the parameter uncertainty, it does not resolve the particle degeneracy issue nor significantly delay the particle’s collapse. Including a fictitious noise is considered a heuristic, and there are no rigorous methods to define the variance of the noise or how much fictitious uncertainty should be introduced to the problem.

From a theoretical point of view, the only approach to effectively delay particle degeneracy is to modify the proposal distribution using improved estimates obtained, for example, using a Gaussian mixture model (Gaussian mixture sigma-point particle filter), a model linearization (Extended particle filter), or the unscented transform (Unscented particle filter). Although the more refined versions of the particle filter delay the onset of particle degeneracy, the issue still persists, with the particles eventually concentrating on small regions of the probability space, albeit a few time steps after the standard PF [24].

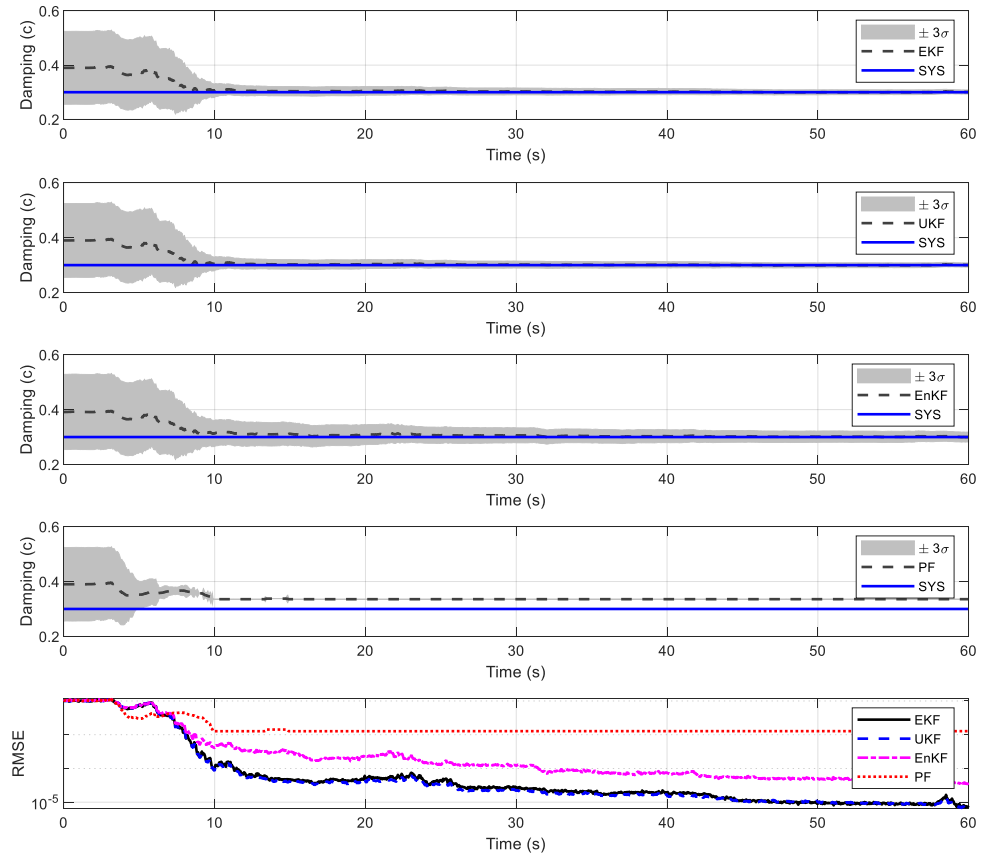


Figure 2. Damping coefficient (c) estimates (in Ns/m).

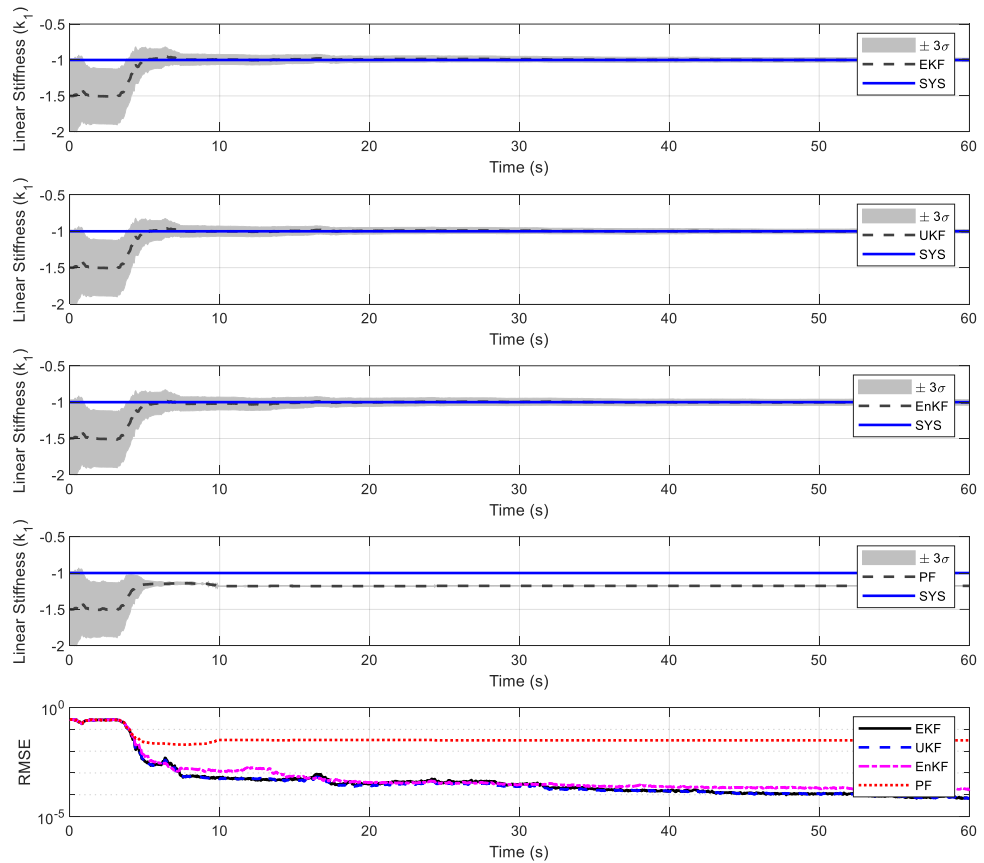


Figure 3. Linear stiffness coefficient (k_1) estimates (in N/m).

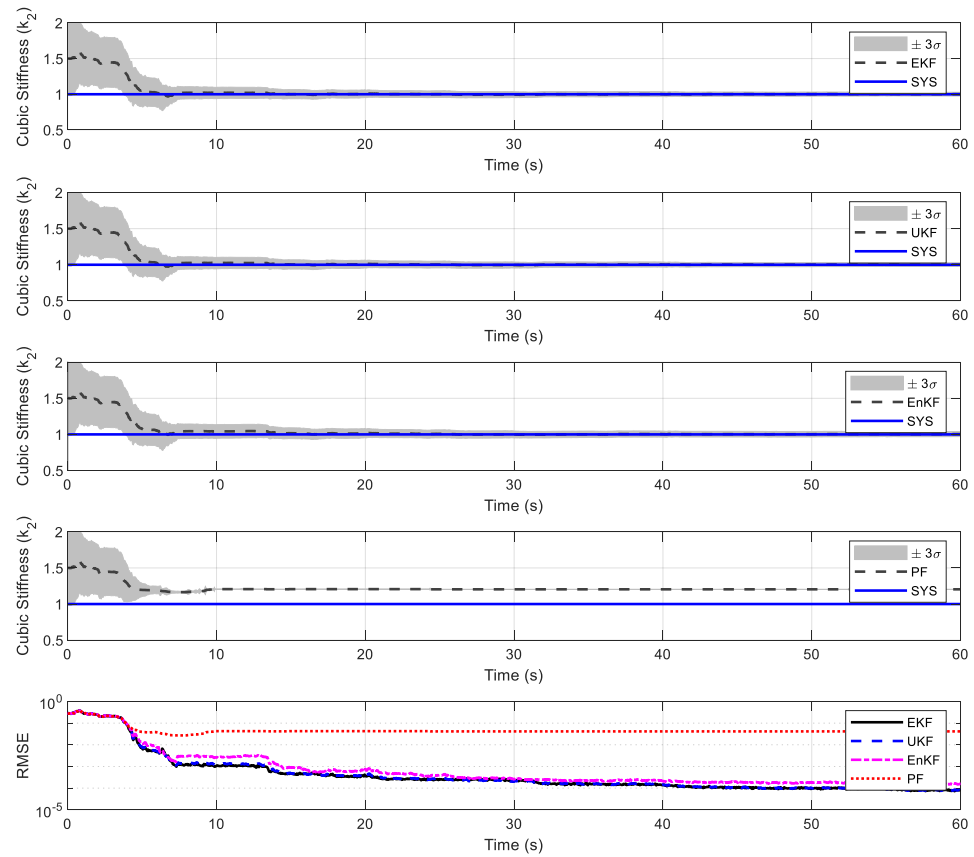


Figure 4. Cubic stiffness coefficient (k_2) estimates (in N/m^3).

Figures 5 and 6 show the estimates of the oscillator displacement and velocity, respectively. All the filters show a state estimation accuracy comparable to the parameter estimation, with the PF showing degradation due to particle degeneracy. Similarly to the parameter estimation results, the PF dynamic response estimates and uncertainty are in agreement with the other filters up to $t = 5$ s, the approximate time instant at which the particle sample collapses to a single particle. In addition to the accuracy of the filters, the computational effort and efficiency are also of importance for implementation in practice. The computational time (in seconds) required by the filters was, for the EKF, 500 s, for the UKF, 2 s, for the EnKF, 90 s, and, for the PF, 5370 s. The results show that the UKF is the most efficient estimation algorithm with good accuracy at an affordable computational cost. Most of the computational time required for the EKF is used in the computation of the gradient required to linearize the model. The EnKF and PF are computationally intensive and not suitable for online or real-time applications due to the effort required to project/update the large set of sample points/particles employed by the algorithms. The RMSE of the estimates providing a measure of the accuracy of the four filters is summarized in Table 1.

Table 1. RMSE summary for the Duffing oscillator example.

	c (Ns/m)	k_1 (N/m)	k_2 (N/m^3)
EKF	7.4×10^{-6}	6.7×10^{-5}	7.9×10^{-5}
UKF	7.4×10^{-6}	6.7×10^{-5}	7.9×10^{-5}
EnKF	4.8×10^{-5}	1.5×10^{-4}	1.5×10^{-4}
PF	0.0012	0.03	0.04

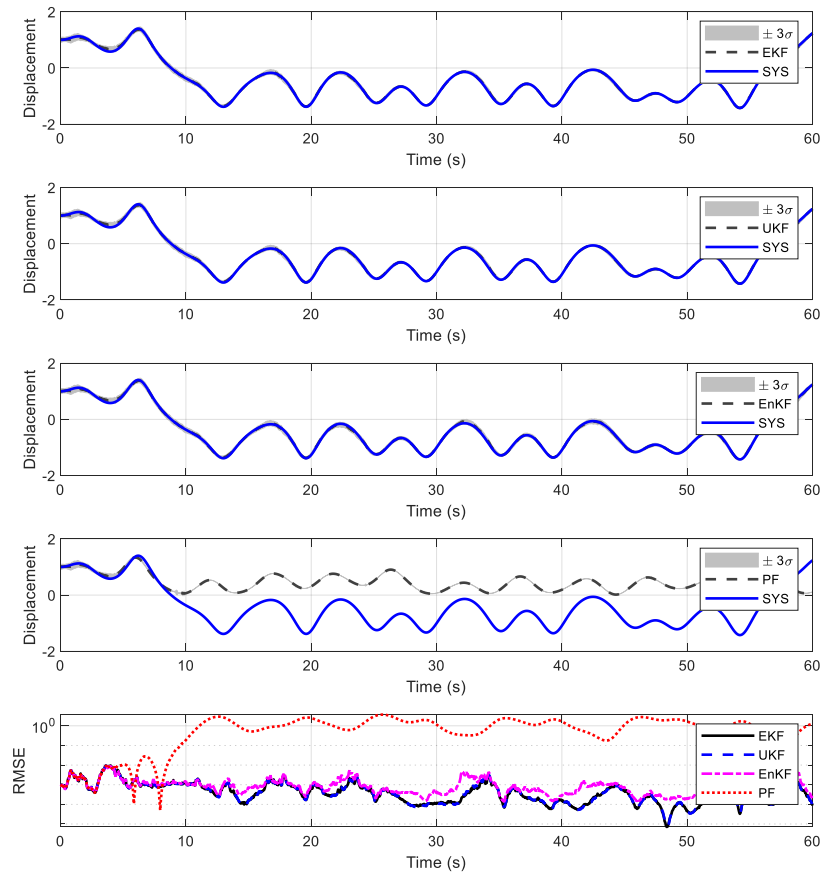


Figure 5. Oscillator displacement estimates (in m).

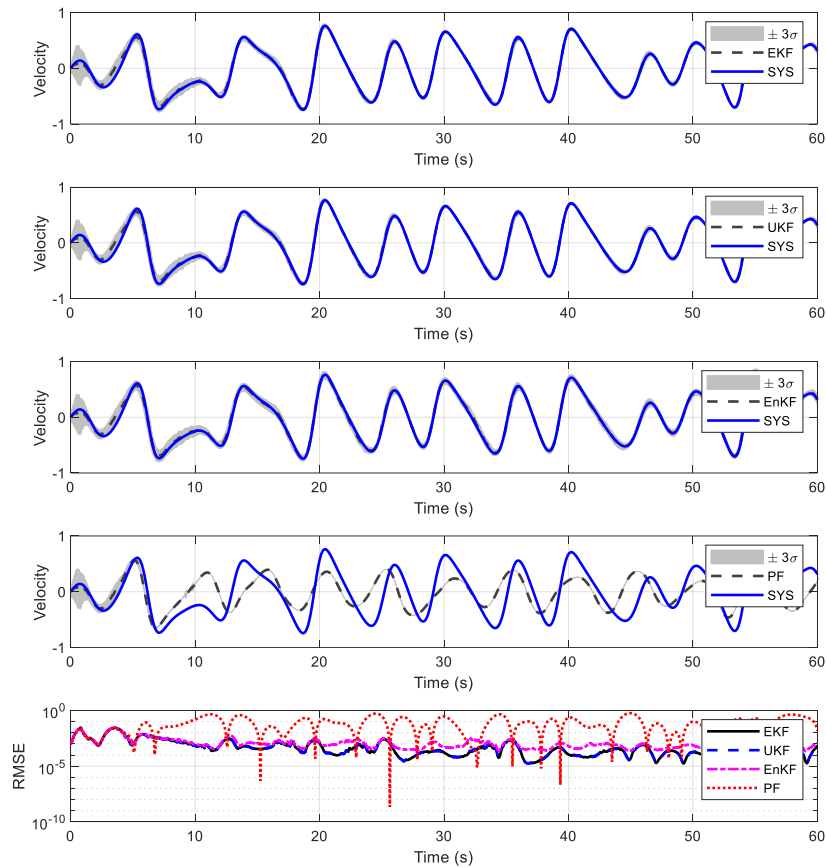


Figure 6. Oscillator velocity estimates (in m/s).

4.2. Bouc–Wen Hysteretic Oscillator

In this section, we consider a nonlinear hysteretic oscillator of the Bouc–Wen type [49]. The development of response and parameter estimation methods in the context of hysteretic models has received notable attention in recent decades due to their application in the modeling of material nonlinearity and inelastic behavior in structural mechanics, including low-cycle damage due to severe loading and high-cycle accumulated fatigue [50–55]. The dynamic behavior of an oscillator with a Bouc–Wen model for hysteresis is governed by the following equation:

$$m\ddot{q}(t) + c\dot{q}(t) + D_y k z(t) = u(t) \tag{23}$$

where k is the initial stiffness and D_y is the yielding displacement. The normalized hysteretic force variable z is governed by the following:

$$\dot{z}(t) = \frac{1}{D_y} \left[\dot{q}(t) - \beta |\dot{q}(t)| |z(t)|^{\nu-1} z(t) - \gamma \dot{q}(t) |z(t)|^\nu \right] \tag{24}$$

where β , γ , and ν define the shape and geometry of the hysteresis loops including the transition between elastic and plastic behavior regimes. The parameters β and γ are not unique when treated independently, and the constraint $\beta + \gamma = 1$ is used to enforce uniqueness; this implies that only one of the parameters needs to be estimated [49]. Without loss of generality, the parameter β is estimated, and the parameter γ is readily computed from the constraint equation.

The true/exact system parameters used to generate the data are as follows (in consistent SI units): $m = 1$ kg, $k = 9$ N/m, $D_y = 0.3$ m, $\beta = 0.2$, $\nu = 2$, and a 5% damping ratio. The oscillator is excited by a base acceleration, \ddot{u}_g , in which case the forcing input is $u = -m\ddot{u}_g$. To solve the forward problem and to perform the estimation, the system is time discretized using the fourth-order Runge–Kutta method with time step $\Delta t = 0.005$ s. The objective is to estimate the nonlinear parameters vector $\theta = \{ \beta, D_y, \nu \}$, which is assumed to be uncertain/unknown. The measurement data used for estimation consists of noise-contaminated absolute acceleration response with a 10% noise-to-signal RMS. The nonlinear Bayesian filters are implemented using the filter parameters of the previous section. The prior distributions adopted for the parameters are depicted in Figure 7, which shows that initial errors between 60 and 100% are assumed for all the parameters. The figure also shows that significant uncertainty about the true values of the parameters is accounted for in the analysis. The level of uncertainty in the prior distributions adopted are in agreement with the expected level of prior knowledge about structural nonlinear parameters where nonlinear parameters show significant uncertainty and variability.

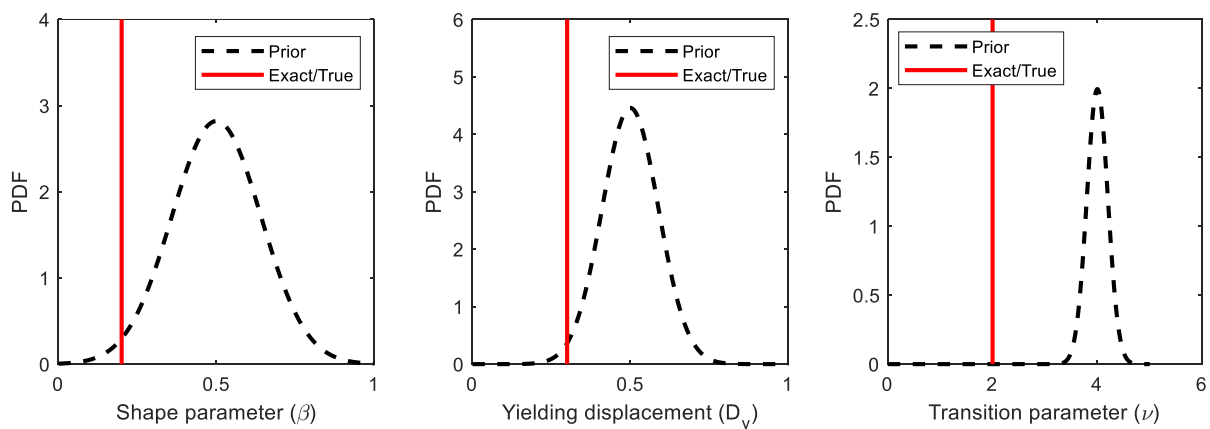


Figure 7. Parameters prior distributions for the hysteretic oscillator example (in SI units).

Figures 8–10 show the estimation results for the nonlinear hysteresis geometry parameters; the figures show the mean estimate and the uncertainty bounds defined as ± 3 standard deviations. The figures show that the EKF, UKF, and EnKF are able to successfully estimate the oscillator nonlinear parameters, while the PF shows convergence difficulties. Similarly to the previous example, based on the particle’s coefficient of variation, the particle degeneracy was detected at $t = 2$ s, which can be confirmed from the estimated time histories by noting the rapid collapse of the uncertainty bounds at this time when compared to the other filters. The reduction in the PF uncertainty bounds at $t = 2$ s is not related to the updating of the posterior distribution as more data becomes available, and instead is a limitation of the algorithm due to the importance distribution adopted that results in a rapid collapse of the particles; the uncertainty reduction is a fictitious artifact due to the collapse and not representative of the actual distribution uncertainty, which can be verified by comparing the uncertainty bounds provided by the other filters. The RMSE of the estimates providing a measure of the accuracy of the four filters is summarized in Table 2.

Figure 11 shows the displacement estimates provided by all the filters, while Figure 12 shows the force–displacement histories, which depict the hysteretic behavior. The UKF shows the best hysteretic response tracking capabilities, followed by the EnKF and EKF.

Table 2. RMSE summary for the Bouc–Wen oscillator example.

	β	ν	D_y (m)
EKF	2.6×10^{-5}	1.1×10^{-4}	7.2×10^{-5}
UKF	1.9×10^{-5}	1.3×10^{-2}	1.2×10^{-3}
EnKF	2.0×10^{-5}	1.2×10^{-2}	1.2×10^{-3}
PF	0.25	3.12	0.15

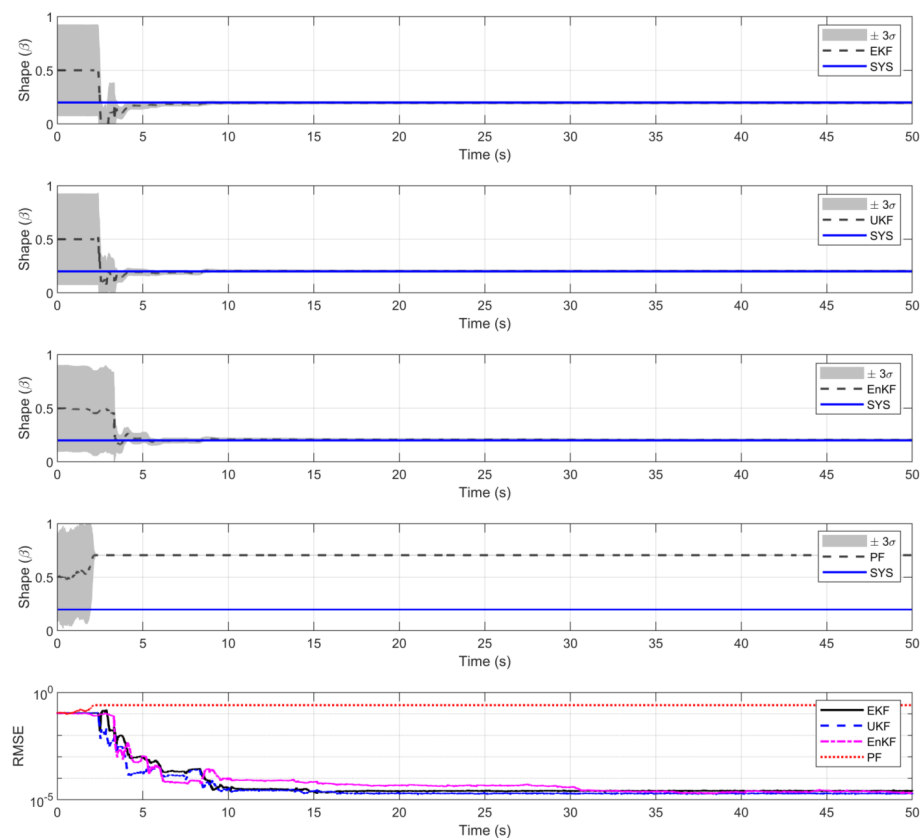


Figure 8. Hysteresis shape parameter (β) estimates (non-dimensional).

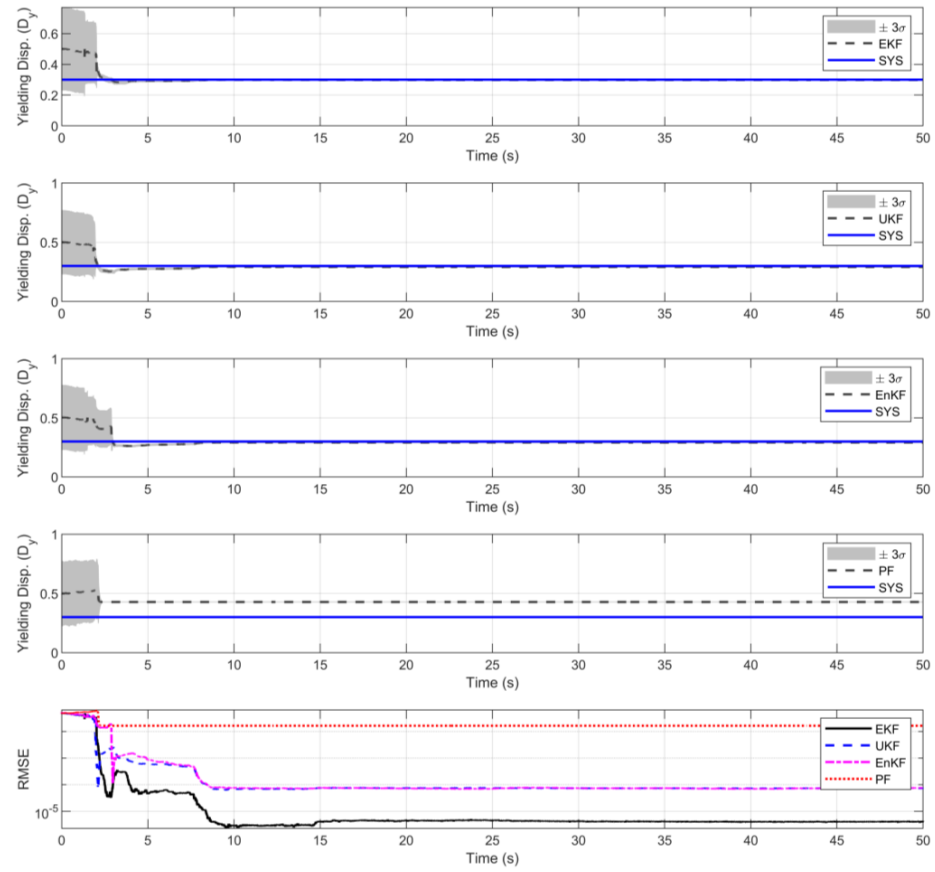


Figure 9. Yielding displacement parameter (D_y) estimates (in m).

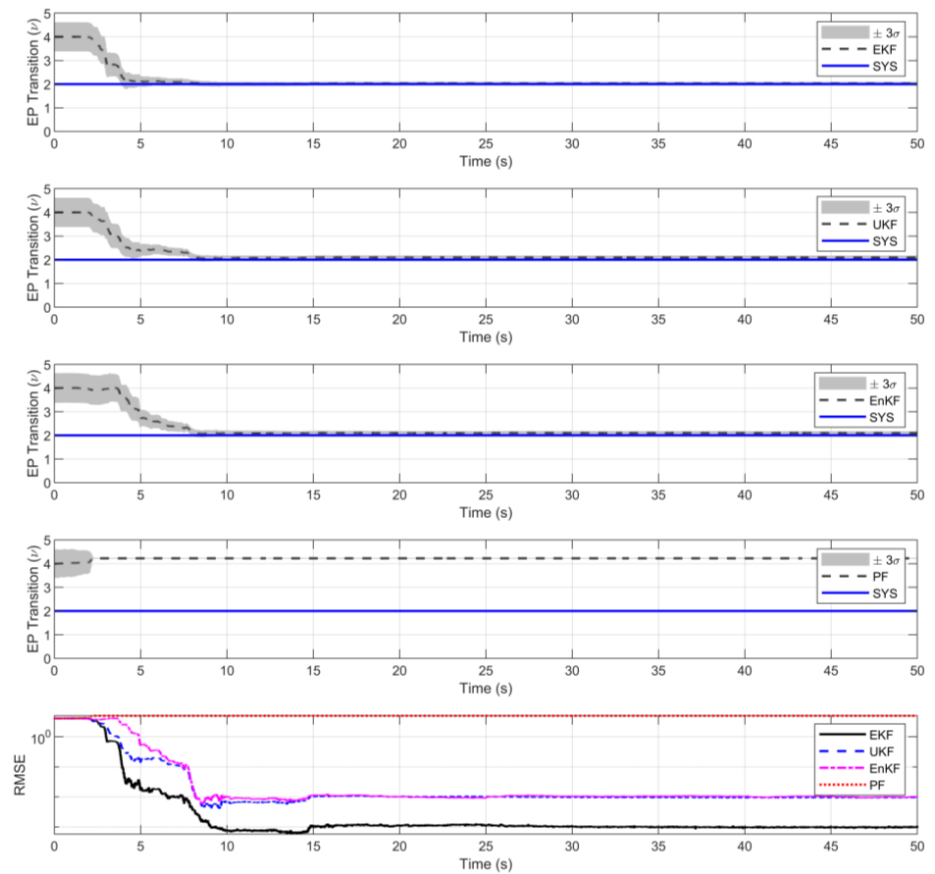


Figure 10. Elastic-to-Plastic transition parameter (ν) estimates (non-dimensional).

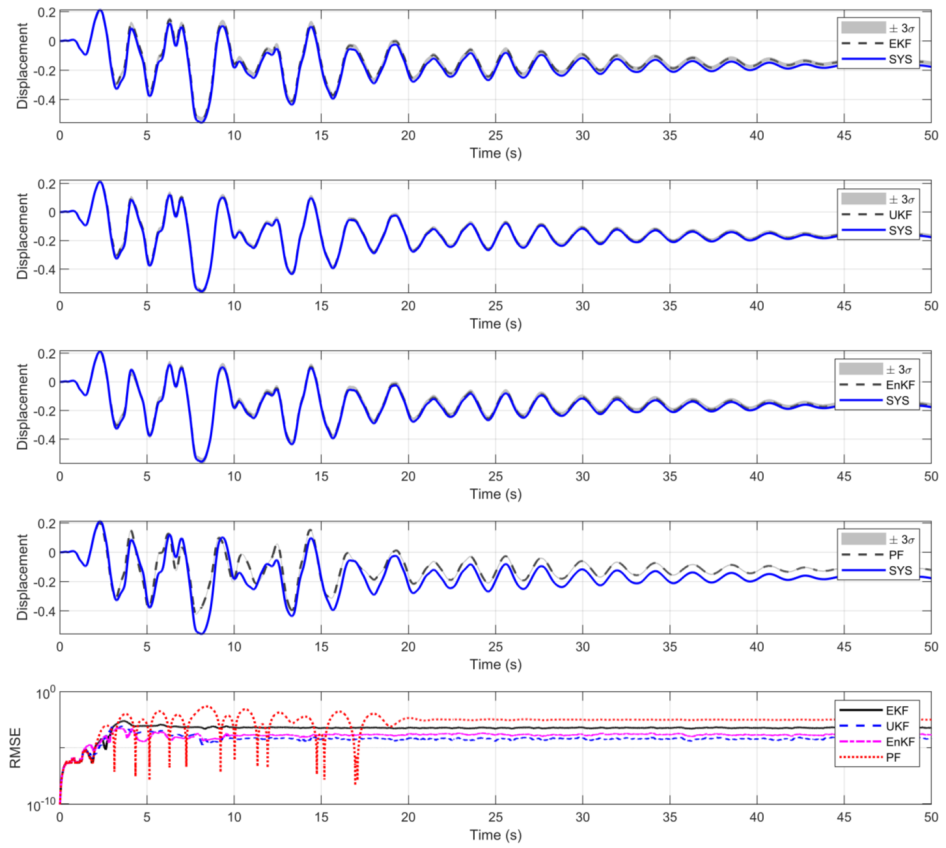


Figure 11. Oscillator displacement estimates (in m).

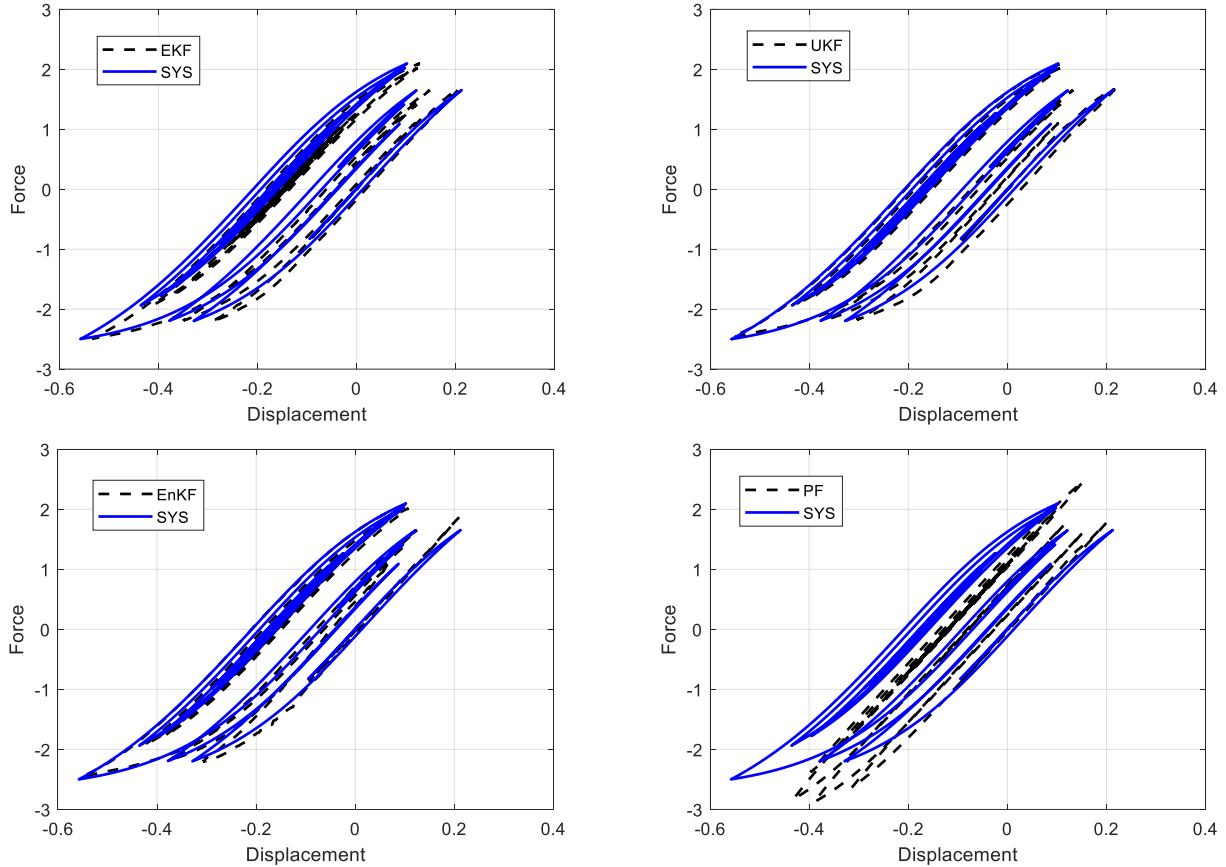


Figure 12. Oscillator force–displacement hysteretic response estimates (in SI units).

4.3. Bouc–Wen Hysteretic Chain System

In this section, we study the scalability of the methods for a system with an increased number of parameters and limited response measurements. For this purpose, we consider a five-degrees-of-freedom chain-type system, where each spring of the chain is governed by the hysteretic Bouc–Wen model discussed in the previous section. The mass and restoring force characteristics of the system are shown in the left panel of Figure 13. Damping is modeled as modal damping with a 5% damping ratio for all modes, computed based on the initial linear properties of the system.

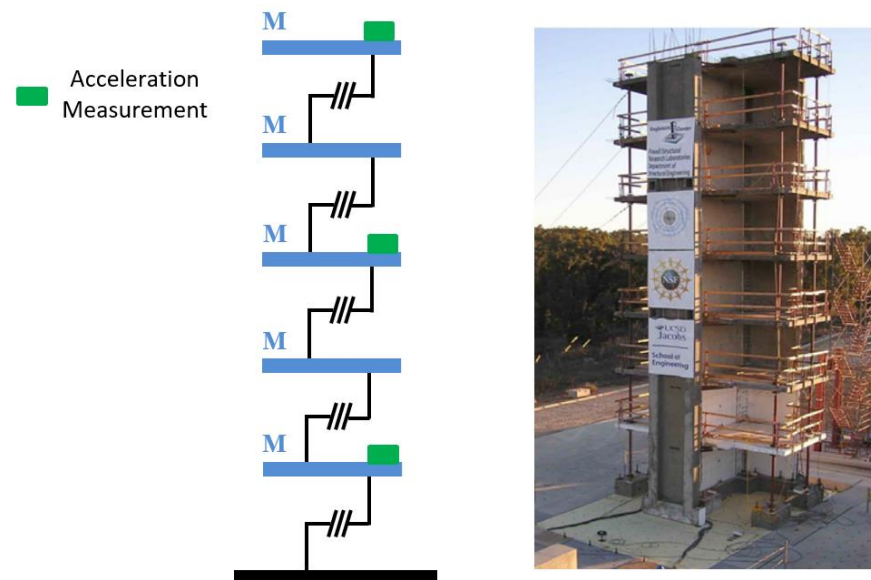


Figure 13. Hysteretic five-degrees-of-freedom chain system.

The true/exact system parameters are selected as follows (in consistent SI units). The mass and initial linear stiffness (for all springs) are $M = 4.3 \times 10^4$ and $k = 2 \times 10^7$. The hysteretic parameters for all the springs are selected as $D_y = 0.10$, $\beta = 0.20$, and $\nu = 2$. The parameters are selected to be consistent with the parameters of the reinforced concrete structure tested on a shake table shown in Figure 13 [56,57]. The system is excited by a base acceleration consisting of a record of the 1994 Northridge earthquake. The objective is to estimate the parameters vector $\theta = \{\beta D_y \nu\}$ for each of the five springs of the model. The measurement data used for estimation consist of the noise-contaminated absolute acceleration response of degrees-of-freedom 1, 3, and 5, with a 10% noise-to-signal RMS (see Figure 13). Note that, in this study, we focus on the parameters of the system; for a related study focused on seismic response reconstruction, the interested reader can consult Ref. [57].

The nonlinear Bayesian filters are implemented using the filter parameters discussed in the previous section. The parameters prior distributions are selected with random initial errors between 25% and 75% for all the parameters and a coefficient of variation of 5%. Note that the dimension of the augmented state for this example is $x \in \mathbb{R}^{30}$, with three dynamic response variables and three parameters for each degree of freedom. The relationship between the dimension of the state and the number, type, and location of measured outputs is the main factor determining the accuracy and capability of the estimators to successfully infer the parameters (identifiability), with scalability and identifiability limitations arising as the dimension of the augmented state increases, an issue known as the curse of dimensionality. Moreover, the use of absolute acceleration measurements, usually used in practice due to the affordability of accelerometer sensors, presents an

increased challenge in the estimation of nonlinear parameters with reduced identifiability with respect to other types of measurements, such as displacements [58].

Figures 14–16 show the estimation results for the hysteresis shape parameter, the yielding displacement, and the linear–inelastic transition parameter, respectively; the figures show the mean estimates for the springs of each degree of freedom. The figures show that only the UKF and EnKF are able to successfully estimate the nonlinear parameters of the system. The PF shows the same particle degeneracy issue of previous examples, although the issue is exacerbated by the dimension of the augmented state. It is interesting to note the degradation of the estimation accuracy of the EKF with respect to the oscillator counterpart. As shown in Section 4.2, the EKF was able to estimate the single-degree-of-freedom hysteretic oscillator parameters, while, in this section, the EKF estimates are not accurate under an increased dimension of the augmented state and the use of limited acceleration measurements. On the other hand, the UKF and EnKF were able to maintain a comparable level of accuracy.

Finally, it is interesting to note that the computational time required by the filters for this example were EKF—4 h, UKF—19 min, EnKF—5 h, and PF—50 h. Thus, the UKF is the most efficient estimation algorithm with an accuracy comparable to the EnKF at a significantly reduced computational cost.

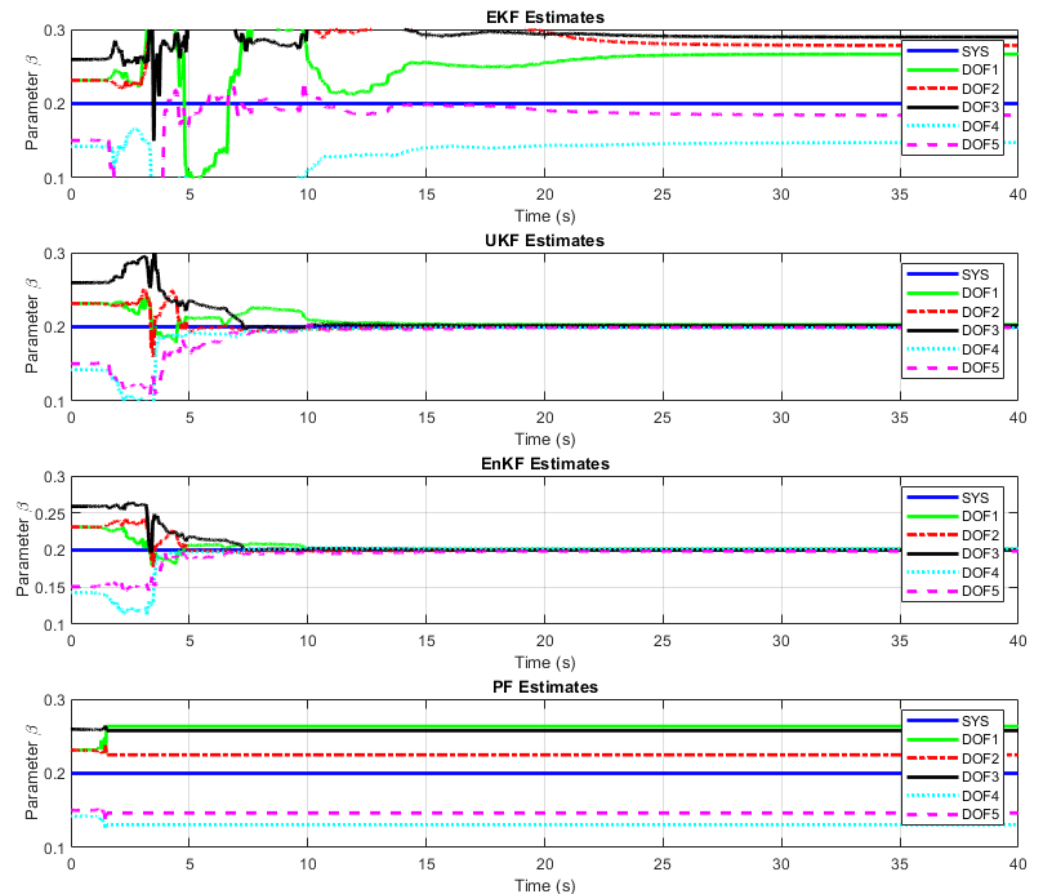


Figure 14. Hysteresis parameter (β) mean estimates for all degrees of freedom (non-dimensional).

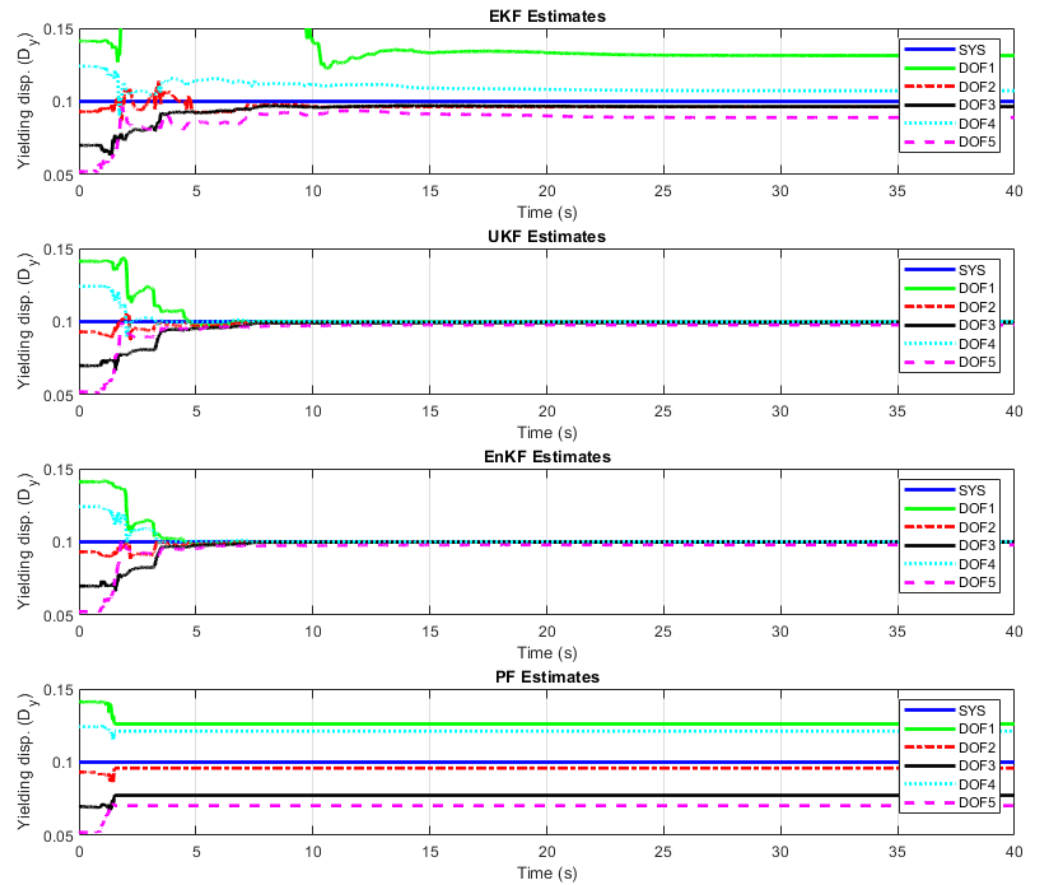


Figure 15. Yielding displacement parameter (D_y) mean estimates for all degrees of freedom (in m).

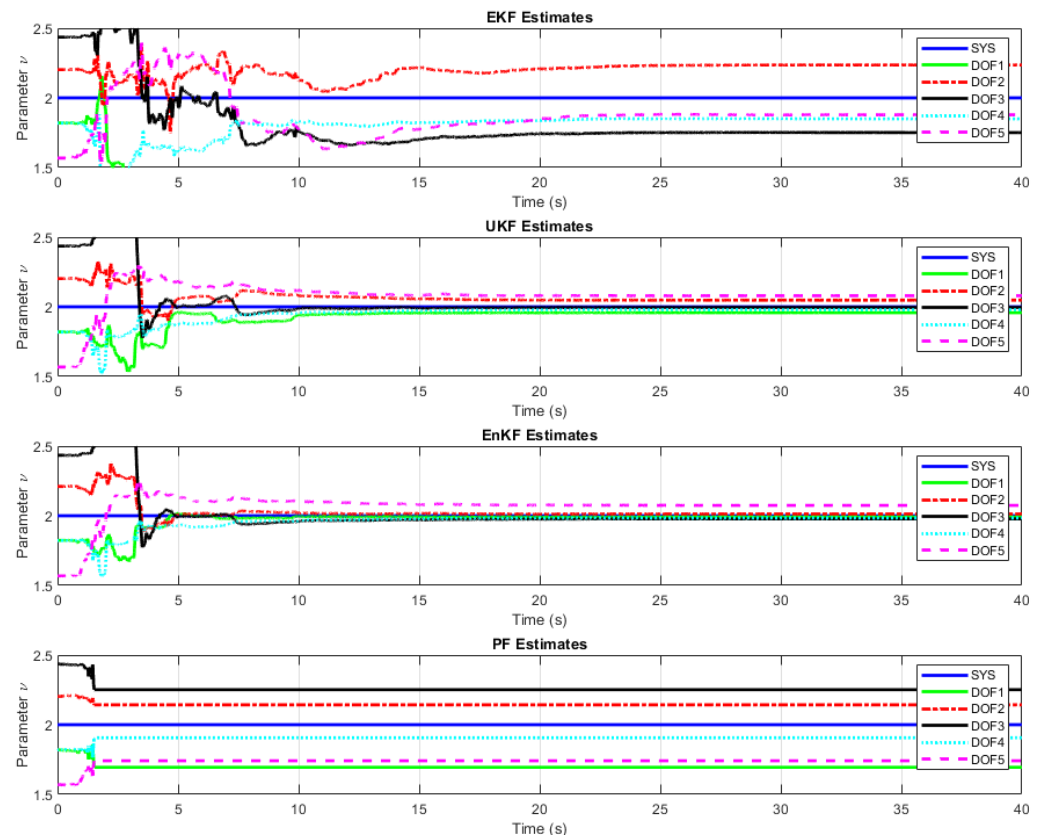


Figure 16. Elastic-to-Plastic transition parameter (ν) mean estimates for all degrees of freedom (non-dimensional).

5. Conclusions

This paper examined the performance of nonlinear Bayesian filtering when applied to the identification of the parameters of structural and mechanical systems. This study employed metrics to assess the accuracy and limitations of the four most well-established nonlinear filtering methods: the extended Kalman filter (EKF), the unscented Kalman filter (UKF), the ensemble Kalman filter (EnKF), and the particle filter (PF). The filtering methods were applied to estimate the parameters and the response of three benchmark dynamical systems used in structural mechanics, namely, a Duffing oscillator, a hysteretic oscillator, and a hysteretic chain-type system.

For the single-degree-of-freedom oscillators, the EKF and UKF converged to the true system parameters with a comparable estimation accuracy, with the EKF showing a higher computational cost. In the case of multiple degrees-of-freedom systems, the EKF was not able to scale appropriately to large parameter spaces, while the UKF preserved the estimation accuracy observed for the oscillators. In the case of the EnKF, the algorithm provided estimates of the same order of accuracy as the UKF, albeit at a significantly higher computational cost that was on the order of several hours. The PF was not able to converge to the true parameters for any of the systems studied, with the particles showing severe degeneracy, collapsing to a reduced subset of the parameter space after a few implementation steps. Based on the performance, accuracy, and computational efficiency of the methods under different operating conditions, it is concluded that the UKF is the most effective filtering-based system identification method for the type of systems considered herein, showing a stable estimation accuracy for parameter spaces considered of high dimension in applications.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Grigoriu, M. *Stochastic Systems: Uncertainty Quantification and Propagation*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
2. Ljung, L. System Identification. In *Signal Analysis and Prediction*; Procházka, A., Uhlř, J., Rayner, P.W.J., Kingsbury, N.G., Eds.; Birkhäuser Verlag: Basel, Switzerland, 1998.
3. Maia, N.M.M.; Montalvão Silva e Silva, J.M. *Theoretical and Experimental Modal Analysis*; Wiley: Hoboken, NJ, USA, 1997.
4. Kerschen, G.; Worden, K.; Vakakis, A.F.; Golinval, J.-C. Past, present and future of nonlinear system identification in structural dynamics. *Mech. Syst. Signal Process.* **2005**, *20*, 505–592. [[CrossRef](#)]
5. Moaveni, B.; He, X.; Conte, J.P.; Restrepo, J.I. Damage identification study of a seven-story full-scale building slice tested on the UCSD-NEES shake table. *Struct. Saf.* **2010**, *32*, 347–356. [[CrossRef](#)]
6. Ortiz, G.A.; Alvarez, D.A.; Bedoya-Ruiz, D. Identification of Bouc–Wen type models using multi-objective optimization algorithms. *Comput. Struct.* **2013**, *114–115*, 121–132. [[CrossRef](#)]
7. Beck, J.L. Bayesian system identification based on probability logic. *Struct. Control Health Monit.* **2010**, *17*, 825–847. [[CrossRef](#)]
8. Yuen, K.V.; Mu, H.Q. Real-time system identification: An algorithm for simultaneous model class selection and parametric identification. *Comput.-Aided Civ. Infrastruct. Eng.* **2015**, *30*, 785–801. [[CrossRef](#)]
9. Behmanesh, I.; Moaveni, B.; Lombaert, G.; Papadimitriou, C. Hierarchical Bayesian model updating for structural identification. *Mech. Syst. Signal Process.* **2015**, *64*, 360–376. [[CrossRef](#)]
10. Yuen, K.V.; Kuok, S.C.; Dong, L. Self-calibrating Bayesian real-time system identification. *Comput.-Aided Civ. Infrastruct. Eng.* **2019**, *34*, 806–821. [[CrossRef](#)]
11. Jazwinski, A.H. *Stochastic Processes and Filtering Theory*; Courier Corporation: Chelmsford, MA, USA, 2007.
12. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **1960**, *82*, 35–45. [[CrossRef](#)]
13. Simon, D. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*; John Wiley and Sons: Hoboken, NJ, USA, 2006.

14. Julier, S.; Uhlmann, J. A new extension of the Kalman filter to nonlinear systems. In *The Robotic Research Group Report*; The University of Oxford: Oxford, UK, 1997.
15. Evensen, G. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean. Dyn.* **2003**, *53*, 343–367. [[CrossRef](#)]
16. Doucet, A.; Godsill, S.; Andrieu, C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **2000**, *10*, 197–208. [[CrossRef](#)]
17. Corigliano, A.; Mariani, S. Parameter identification in explicit structural dynamics: Performance of the extended Kalman filter. *Comput. Methods Appl. Mech. Eng.* **2004**, *193*, 3807–3835. [[CrossRef](#)]
18. Lin, J.W.; Betti, R. On-line identification and damage detection in non-linear structural systems using a variable forgetting factor approach. *Earthq. Eng. Struct. Dyn.* **2004**, *33*, 419–444. [[CrossRef](#)]
19. Ghanem, R.; Ferro, G. Health monitoring for strongly non-linear systems using the Ensemble Kalman filter. *Struct. Control Health Monit.* **2005**, *13*, 245–259. [[CrossRef](#)]
20. Ching, J.; Beck, J.; Porter, K.; Shaikhutdinov, R. Bayesian state estimation method for nonlinear systems and its application to recorded seismic response. *ASCE J. Eng. Mech.* **2006**, *132*, 396–410. [[CrossRef](#)]
21. Yang, J.N.; Lin, S.; Huang, H.; Zhou, L. An adaptive extended Kalman filter for structural damage identification. *Struct. Control Health Monit.* **2006**, *13*, 849–867. [[CrossRef](#)]
22. Namdeo, V.; Manohar, C. Nonlinear structural dynamical system identification using adaptive particle filters. *J. Sound Vib.* **2007**, *306*, 524–563. [[CrossRef](#)]
23. Wu, M.; Smyth, A.W. Application of the unscented Kalman filter for real-time nonlinear structural system identification. *Struct. Control Health Monit.* **2007**, *14*, 971–990. [[CrossRef](#)]
24. Chatzi, E.N.; Smyth, A.W. The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-located heterogeneous sensing. *Struct. Control Health Monit.* **2009**, *16*, 99–123. [[CrossRef](#)]
25. Lourens, E.; Papadimitriou, C.; Gillijns, S.; Reynders, E.; De Roeck, G.; Lombaert, G. Joint input-response estimation for structural systems based on reduced-order models and vibration data from a limited number of sensors. *Mech. Syst. Signal Process.* **2012**, *29*, 310–327. [[CrossRef](#)]
26. Xu, B.; He, J.; Rovekamp, R.; Dyke, S.J. Structural parameters and dynamic loading identification from incomplete measurements: Approach and validation. *Mech. Syst. Signal Process.* **2012**, *28*, 244–257. [[CrossRef](#)]
27. Chatzis, M.; Chatzi, E.; Smyth, A. An experimental validation of time domain system identification methods with fusion of heterogeneous data. *Earthq. Eng. Struct. Dyn.* **2014**, *44*, 523–547. [[CrossRef](#)]
28. Sun, H.; Betti, R. Simultaneous identification of structural parameters and dynamic input with incomplete output-only measurements. *Struct. Control Health Monit.* **2014**, *21*, 868–889. [[CrossRef](#)]
29. Azam, S.E.; Chatzi, E.; Papadimitriou, C. A dual Kalman filter approach for state estimation via output-only acceleration measurements. *Mech. Syst. Signal Process.* **2015**, *60*, 866–886. [[CrossRef](#)]
30. Khalil, M.; Sarkar, A.; Adhikari, S.; Poirel, D. The estimation of time-invariant parameters of noisy nonlinear oscillatory systems. *J. Sound Vib.* **2015**, *344*, 81–100. [[CrossRef](#)]
31. Liu, L.; Su, Y.; Zhu, J.; Lei, Y. Data fusion based EKF-UI for real-time simultaneous identification of structural systems and unknown external inputs. *Measurement* **2016**, *88*, 456–467. [[CrossRef](#)]
32. Maes, K.; Smyth, A.; De Roeck, G.; Lombaert, G. Joint input-state estimation in structural dynamics. *Mech. Syst. Signal Process.* **2016**, *70–71*, 445–466. [[CrossRef](#)]
33. Yan, G.; Sun, H.; Büyükoztürk, O. Impact load identification for composite structures using Bayesian regularization and unscented Kalman filter. *Struct. Control Health Monit.* **2016**, *24*, e1910. [[CrossRef](#)]
34. Erazo, K.; Nagarajaiah, S. An offline approach for output-only Bayesian identification of stochastic nonlinear systems using unscented Kalman filtering. *J. Sound Vib.* **2017**, *397*, 222–240. [[CrossRef](#)]
35. Erazo, K.; Nagarajaiah, S. Bayesian structural identification of a hysteretic negative stiffness earthquake protection system using unscented Kalman filtering. *Struct. Control Health Monit.* **2018**, *25*, e2203. [[CrossRef](#)]
36. Erazo, K.; Sen, D.; Nagarajaiah, S.; Sun, L. Vibration-based structural health monitoring under changing environmental conditions using Kalman filtering. *Mech. Syst. Signal Process.* **2018**, *117*, 1–15. [[CrossRef](#)]
37. Lei, Y.; Xia, D.; Erazo, K.; Nagarajaiah, S. A novel unscented Kalman filter for recursive state-input-system identification of nonlinear systems. *Mech. Syst. Signal Process.* **2019**, *127*, 120–135. [[CrossRef](#)]
38. Kammouh, O.; Gardoni, P.; Cimellaro, G.P. Probabilistic framework to evaluate the resilience of engineering systems using Bayesian and dynamic Bayesian networks. *Reliab. Eng. Syst. Saf.* **2020**, *198*, 106813. [[CrossRef](#)]
39. Ierimonti, L.; Cavalagli, N.; Venanzi, I.; García-Macías, E.; Ubertini, F. A transfer Bayesian learning methodology for structural health monitoring of monumental structures. *Eng. Struct.* **2021**, *247*, 113089. [[CrossRef](#)]
40. Diaz, M.; Charbonnel, P.; Chamoin, L. A new Kalman filter approach for structural parameter tracking: Application to the monitoring of damaging structures tested on shaking-tables. *Mech. Syst. Signal Process.* **2022**, *182*, 109529. [[CrossRef](#)]

41. Yu, X.; Li, X.; Bai, Y. Evaluating maximum inter-story drift ratios of building structures using time-varying models and Bayesian filters. *Soil Dyn. Earthq. Eng.* **2022**, *162*, 107496. [[CrossRef](#)]
42. Hu, L.; Bao, Y.; Sun, Z.; Meng, X.; Tang, C.; Zhang, D. Outlier Detection Based on Nelder-Mead Simplex Robust Kalman Filtering for Trustworthy Bridge Structural Health Monitoring. *Remote Sens.* **2023**, *15*, 2385. [[CrossRef](#)]
43. Liu, W.; Lai, Z.; Bacsá, K.; Chatzi, E. Neural extended Kalman filters for learning and predicting dynamics of structural systems. *Struct. Health Monit.* **2023**, *23*, 1037–1052. [[CrossRef](#)]
44. Erazo, K.; Di Matteo, A.; Spanos, P. Parameter Estimation of Stochastic Fractional Dynamic Systems Using Nonlinear Bayesian Filtering System Identification Methods. *J. Eng. Mech.* **2024**, *150*, 04023117. [[CrossRef](#)]
45. Julier, S.; Uhlmann, J.; Durrant-Whyte, H.F. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Autom. Control.* **2001**, *45*, 477–482. [[CrossRef](#)]
46. Särkkä, S. *Bayesian Filtering and Smoothing*; Cambridge University Press: Cambridge, UK, 2013.
47. Kudryashov, N.A. The generalized Duffing oscillator. *Commun. Nonlinear Sci. Numer. Simul.* **2021**, *93*, 105526. [[CrossRef](#)]
48. Jerath, K.; Brennan, S.; Lagoa, C. Bridging the gap between sensor noise modeling and sensor characterization. *Measurement* **2018**, *116*, 350–366. [[CrossRef](#)]
49. Ikhouane, F.; Rodellar, J. *Systems with Hysteresis: Analysis, Identification and Control Using the Bouc-Wen Model*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
50. Charalampakis, A.E.; Koumousis, V.K. Identification of Bouc–Wen hysteretic systems by a hybrid evolutionary algorithm. *J. Sound Vib.* **2008**, *314*, 571–585. [[CrossRef](#)]
51. Worden, K.; Hensman, J. Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference. *Mech. Syst. Signal Process.* **2012**, *32*, 153–169. [[CrossRef](#)]
52. Omrani, R.; Hudson, R.E.; Taciroglu, E. Parametric identification of nondegrading hysteresis in a laterally and torsionally coupled building using an unscented Kalman filter. *J. Eng. Mech.* **2013**, *139*, 452–468. [[CrossRef](#)]
53. Erazo, K.; Hernandez, E.M. State estimation in nonlinear structural systems. In *Nonlinear Dynamics, Volume 2, Proceedings of the 32nd IMAC, A Conference and Exposition on Structural Dynamics, Orlando, FL, USA, 3–6 February 2014*; Allen, M., Catbas, F.N., Foss, G., Kerschen, G., Mayes, R., Niezrecki, C., Rixen, D., Atamturktur, H.S., De Clerck, J., Allemang, R.J., et al., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2014; pp. 249–257.
54. Erazo, K.; Hernandez, E.M. Uncertainty quantification of state estimation in nonlinear structural systems with application to seismic response in buildings. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2016**, *2*, B5015001. [[CrossRef](#)]
55. Roohi, M.; Erazo, K.; Rosowsky, D.; Hernandez, E.M. An extended model-based observer for state estimation in nonlinear hysteretic structural systems. *Mech. Syst. Signal Process.* **2020**, *146*, 107015. [[CrossRef](#)]
56. Panagiotou, M.; Restrepo, J.I.; Conte, J.P. Shake-table test of a full-scale 7-story building slice. Phase I: Rectangular wall. *J. Struct. Eng.* **2011**, *137*, 691–704. [[CrossRef](#)]
57. Erazo, K.; Moaveni, B.; Nagarajaiah, S. Bayesian seismic strong-motion response and damage estimation with application to a full-scale seven story shear wall structure. *Eng. Struct.* **2019**, *186*, 146–160. [[CrossRef](#)]
58. Chatzis, M.N.; Chatzi, E.N.; Smyth, A.W. On the observability and identifiability of nonlinear structural and mechanical systems. *Struct. Control Health Monit.* **2014**, *22*, 574–593. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.