An offline approach for output-only Bayesian identification of stochastic nonlinear systems using unscented Kalman filtering

Kalil Erazo[a,*], Satish Nagarajaiah[a]

[a]*Department of Civil and Environmental Engineering, Rice University, Houston, TX 77005, USA*

Abstract

In this paper an offline approach for output-only Bayesian identification of stochastic nonlinear systems is presented. The approach is based on a re-parameterization of the joint posterior distribution of the parameters that define a postulated state-space stochastic model class. In the re-parameterization the state predictive distribution is included, marginalized, and estimated recursively in a state estimation step using an unscented Kalman filter, bypassing state augmentation as required by existing online methods. In applications expectations of functions of the parameters are of interest, which requires the evaluation of potentially high-dimensional integrals; Markov chain Monte Carlo is adopted to sample the posterior distribution and estimate the expectations. The proposed approach is suitable for nonlinear systems subjected to non-stationary inputs whose realization is unknown, and that are modeled as stochastic processes. Numerical verification and experimental validation examples illustrate the effectiveness and advantages of the approach, including: i) an increased numerical stability with respect to augmented-state unscented Kalman filtering, avoiding divergence of the estimates when the forcing input is unmeasured; ii) the ability to handle arbitrary prior and posterior distributions. The experimental validation of the approach is conducted using data from a large-scale structure tested on a shake table. It is shown that the approach is robust to inherent modeling errors in the description of the system and forcing input, providing accurate prediction of the dynamic response when the excitation history is unknown.

Keywords: Nonlinear System Identification, Bayesian Parameter Estimation, Stochastic Systems, Unscented Kalman Filter, Markov Chain Monte Carlo

1. Introduction

System identification (SI) is the process of using measurements of the response of a system to infer the parameters that define its mathematical models [1, 2]. Several applications of SI have been presented in structural mechanics, including operational condition assessment and management, improvement of design methods, structural control, and structural reliability applications. For structural and mechanical systems whose dynamic behavior can be captured by linear models an extensive literature of SI algorithms has been developed over the past decades [1, 2]. However, in some applications a linear model cannot capture features of the dynamic response of a system of interest. Typical sources of nonlinearity include large displacements, large deformations, material nonlinearity, boundary conditions, energy dissipation devices for vibration suppression, actuators, among others [3]. Identification of nonlinear systems presents a significantly increased challenge with respect to its linear counterpart, mainly because of the lack of a general input-output map for nonlinear operators [4].

In the identification of dynamic systems it is often the case that within a proposed model class a set of uncertain parameters is consistent with the available data, rather than a single point value. Furthermore, the system may be subjected to forcing inputs that are difficult to measure, as occurs for example in structural systems subjected to wind-induced effects, offshore structures and bridges with submerged foundations. Sensors malfunction is also a motivation

*Corresponding author
Email addresses: kalilerazo@gmail.com (Kalil Erazo), Satish.Nagarajaiah@rice.edu (Satish Nagarajaiah)

for the development of approaches that do not rely on knowledge of the input time history [5]. In such applications a probabilistic or stochastic model can be adopted to quantify and propagate the uncertainties involved in the modeling of dynamic systems [6]. In this context Bayesian inference provides a logically consistent, robust and rigorous theory that can be applied to characterize modeling uncertainty and system identification for a wide class of problems [7–12]. In Bayesian system identification the uncertainty in the parameters is characterized by a probability density function (PDF) conditional in the available data, known as the posterior distribution.

Bayesian SI methods can be broadly classified as online (also known as real-time, recursive or sequential estimation) and offline (also known as batch estimation). Online approaches aim to estimate the parameters sequentially as the data becomes available, while offline approaches consider a fixed observation record. Popular methods for online Bayesian SI include the Kalman filtering based algorithms (extended, unscented and ensemble Kalman filters) [13, 14] and stochastic simulation based approaches (particle filters) [15, 16]. Among the various Bayesian filters in the literature the unscented Kalman filter (UKF) has received notable attention in structural mechanics applications, mainly due to its increased accuracy with respect to the extended Kalman filter at a less computational cost than sampling based filters (ensemble Kalman and particle filters) [17–23]. Recent work in Bayesian SI in nonlinear systems has focused in (augmented-state) joint state-parameter estimation using online methods in applications where the time history of the main component of the forcing input is known (measured) up to an additive noise [15–17, 24, 25]. Applications where the main component of the input is unmeasured and modeled as a stochastic process are sought herein; the identification under such conditions will be referred to as output-only identification [26]. Effective probabilistic approaches for output-only parameter identification in stationary and/or linear systems have been previously proposed in the literature [26–28]. However, the development of robust and efficient output-only Bayesian SI approaches for non-stationary nonlinear structural and mechanical systems has been limited [5].

In this paper an offline approach for Bayesian identification of stochastic nonlinear systems is presented. In contrast to existing output-only identification approaches that typically assume a white noise input model, non-stationary nonlinear systems are studied herein. In the proposed approach the augmented-state parameter estimation problem is decoupled and performed in two sequential stages: an offline Bayesian SI stage where only model parameters are estimated, and a state estimation stage where model predictions are performed using all plausible models of the model class. For this purpose the parameters posterior PDF is re-parameterized in order to include the state predictive distribution as shown in a further section of the paper; the correlation between the measurements, the state and the parameters is provided by the postulated stochastic model class. The state predictive distribution is recursively estimated using an unscented Kalman filter in a state estimation step, bypassing the state augmentation performed by existing online techniques. The decoupling of the problem renders an algorithm with increased numerical stability. Moreover, the proposed approach has the capability to handle non-Gaussian prior and posterior distributions, in contrast to standard augmented-state unscented Kalman filtering estimation that inherits from Kalman filtering the need to use only second-order statistics. The proposed approach is not suitable for applications where an estimate of the state and the parameters is needed near real-time, such as control applications. Offline Bayesian identification algorithms of this kind have become increasingly popular, mainly because of their enhanced stability and robustness with respect to augmented-state parameter estimation approaches when the dimension of the parameter space is relatively large, the number of response measurements is limited and input time history is unknown [29, 30]. In particular it is well-known that for nonlinear systems online algorithms perform poorly in high-dimensional parameter spaces, an issue that is further exacerbated when the forcing input history is not available during the estimation.

In offline Bayesian SI the evaluation of potentially high-dimensional integrals needs to be performed to obtain expectations of functions of the parameters. For example, the parameters marginal and bivariate distributions reveal non-unique optimal estimates and correlations between the parameters. Markov chain Monte Carlo (MCMC) is adopted as an efficient method to sample the joint posterior distribution and approximate the expectations using sampling statistics. In MCMC a Markov chain is constructed such that the stationary distribution of the chain is a target distribution (in this case the parameters posterior distribution). To construct the Markov chain the re-parameterized posterior is used to evaluate the likelihood function using state estimation-based unscented Kalman filtering. The resulting approach will be referred to as the UKF-MCMC approach. Related approaches based on the application of ensemble Kalman and particle filters have been previously proposed in the literature [30]; these approaches have the capability to handle stronger nonlinearities at the expense of a significant increase in the required computational effort. The UKF-MCMC approach uses the unscented Kalman filter to efficiently compute the state predictive distribution, reducing the computational resources demanded by sampling-based stochastic filters.

The proposed UKF-MCMC approach is numerically verified using synthetic data in a bilinear hysteretic oscillator subjected to Gaussian white noise, and a four degree of freedom nonlinear chain subjected to a base acceleration consisting of a modulated Kanai-Tajimi process. The examples show that the approach has the capability to estimate the model parameters correctly, providing confidence intervals to characterize the uncertainty in the estimates. Moreover, the second numerical example is used to numerically illustrate the increased stability of the proposed approach with respect to augmented-state UKF parameter estimation. The enhanced stability is attributed to the improved observability/identifiability conditions of the decoupled problem with respect to the augmented-state problem, due to the increased nonlinearity of the latter. In particular the likelihood function of the augmented-state problem is typically a non-convex multi-modal function with several local minima, which tends to result in rank-deficiency related issues and divergence of the estimates when the forcing input time history is unknown.

The proposed approach is experimentally validated using data from a large-scale structure tested on the NEES-UCSD shake table [31]. The data from this experimental program has been previously employed for (linear) modal identification and damage identification using state-of-the-art input-output and output-only algorithms [32], and for dynamic response estimation using state observers [33, 34]. The experimental validation shows that the approach is robust to inherent modeling errors in the description of the system and forcing input, providing consistent estimates of the parameters and accurate prediction of the dynamic response of the structure.

2. Bayesian System Identification

Consider the class of stochastic nonlinear systems modeled by the following Itô differential equation

$$\mathrm{d}\mathbf{x}(t) = f(\mathbf{x}(t), \boldsymbol{\Phi})\mathrm{d}t + \mathbf{B}_1\mathbf{u}(t)\mathrm{d}t + \mathbf{B}_2\mathrm{d}\boldsymbol{\beta}(t) \tag{1}$$

where the state $\mathbf{x}(t)$ is an $\mathbb{R}^n$-valued stochastic process, and the function $f : \mathbb{R}^n \times \mathbb{R}^\omega \to \mathbb{R}^n$ is a Borel measurable function that defines the dynamic model. The stochastic model is completely or partially parameterized by the uncertain parameters vector $\boldsymbol{\Phi} \in \mathbb{R}^\omega$. The matrix $\mathbf{B_1} \in \mathbb{R}^{n\times r}$ defines the spatial distribution of the known inputs vector $\mathbf{u}(t) \in \mathbb{R}^r$, and the vector $\boldsymbol{\beta}(t)$ is an $\mathbb{R}^s$-valued independent Brownian motions process, used to model non-parametric model errors, unmeasured inputs and/or noise in measured inputs. The matrix $\mathbf{B_2} \in \mathbb{R}^{n\times s}$ maps the stochastic input to the state-space.

We are interested in application where noise contaminated measurements are available and modeled by

$$\mathbf{y}_k = h(\mathbf{x}(t_k), \mathbf{u}(t_k), \boldsymbol{\Phi}) + \boldsymbol{\nu}_k \qquad k = 1, \ldots, T \tag{2}$$

where T is the total number of measurements and t_k indicates the time at which measurement k is recorded. The measurements vector is $\mathbf{y}_k \in \mathbb{R}^m$, the function $h : \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^\omega \to \mathbb{R}^m$ maps to the output space, and $\boldsymbol{\nu}_k \in \mathbb{R}^m$ is the measurement noise, assumed to be a Gaussian sequence with zero mean and covariance matrix $\mathbf{R}$. The model defined by Eq. (1) can be written in discrete time as [35]

$$\mathbf{x}_{i+1} = f_d(\mathbf{x}_i, \boldsymbol{\Phi}) + g(\mathbf{U}_i, \mathbf{W}_i) \tag{3}$$

where $\mathbf{x}_i$ is the discrete approximation of $\mathbf{x}(t_i)$, $\mathbf{U}_i = [\mathbf{u}(t_i)\ \mathbf{u}(t_{i-1}) \ldots \mathbf{u}(t_{i-z})]$ and $\mathbf{W}_i = [\mathbf{w}(t_i)\ \mathbf{w}(t_{i-1}) \ldots \mathbf{w}(t_{i-e})]$ is a vector of independent Gaussian random variables; the functions f_d and g, and the indexes z and e depend on the discretization scheme employed.

The objective of system identification is to infer the vector $\boldsymbol{\Phi}$ that completely or partially parameterize the dynamic model, from an array of noise contaminated measurements $\mathbf{Y}_T = \{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$. Bayesian inference provides a consistent, robust and rigorous theory to tackle this inverse problem for a wide range of systems in a probabilistic setting [7]. In Bayesian SI the uncertainty in the parameters is characterized by a joint probability density function (PDF) conditional on the available information. The updated parameters PDF (also known as the posterior distribution) is denoted as $p(\boldsymbol{\Phi}|\mathbf{Y}_T)$, and is obtained in principle using Bayes' theorem

$$p(\boldsymbol{\Phi}|\mathbf{Y}_T) = \frac{p(\mathbf{Y}_T|\boldsymbol{\Phi})p(\boldsymbol{\Phi})}{p(\mathbf{Y}_T)} \tag{4}$$

where $p(\boldsymbol{\Phi})$ is the prior distribution, used to incorporate information available about the parameters before the data is gathered, $p(\mathbf{Y}_T|\boldsymbol{\Phi})$ as a function of $\boldsymbol{\Phi}$ is the likelihood function, and $p(\mathbf{Y}_T)$ is a normalizing constant.

The posterior PDF embodies all the information about the parameters contained in the data. From this distribution a point estimate can be computed, usually the mean or the mode. The posterior mode (or maximum a posteriori estimate) is

$$\boldsymbol{\Phi}_{MAP} = \arg\max_{\boldsymbol{\Phi}} \; p(\boldsymbol{\Phi}|\mathbf{Y}_T) \tag{5}$$

In applications the parameters marginal distributions are also of interest, and given by

$$p(\Phi_i|\mathbf{Y}_T) = \int p(\boldsymbol{\Phi}|\mathbf{Y}_T)\mathrm{d}(\boldsymbol{\Phi} \setminus \Phi_i) \tag{6}$$

where the integral is about the product measure of the probability space excluding coordinate Φ_i. To quantify the uncertainty in the estimates confidence intervals (CI) are used. Bayesian confidence intervals provide the probability that the parameters that better describe the data lie in a region of the parameter space, given the observed data. The CI for parameter Φ_i is defined as

$$CI_i = [\phi_i^a, \phi_i^b] \qquad such\ that \qquad P(\Phi_i \in CI_i|\mathbf{Y}_T) = \int_{\phi_i^a}^{\phi_i^b} p(\Phi_i|\mathbf{Y}_T)\mathrm{d}\Phi_i \tag{7}$$

where $P(\Phi_i \in CI_i|\mathbf{Y}_T)$ is a target probability. In practice the 95% mean-centered and mode-centered confidence intervals are typically used.

Although the posterior is given in principle by straightforward application of Bayes' theorem, finding it in closed-form poses significant theoretical and computational challenges. In particular in applications involving stochastic dynamic models of the form in Eq. (1) it is a challenging task to compute $p(\boldsymbol{\Phi}|\mathbf{Y}_T)$ directly, mainly because of the lack of a general input-output map for nonlinear operators [4]. A popular strategy to overcome this issue consists in augmenting the state to include the parameters to be estimated, and employ Bayesian filtering to estimate the augmented state online, an strategy referred to as joint state-parameter estimation or augmented-state estimation [36].

A drawback of this standard approach is that by augmenting the state the computational load required by Bayesian filters is significantly increased. In particular filters based on Kalman filtering become susceptible to diverge, an issue attributed to observability/identifiability and rank deficiency related issues [37]. A rank deficient covariance matrix results from the ill-posedness of the inverse problem, which is fundamentally connected to the complex topology of the nonlinear model when the parameters are treated as variables. This issue is exacerbated when the realization of the forcing input is unknown (output-only identification). To overcome some of these difficulties an alternative approach that does not require state augmentation is presented in the following section.

3. Offline approach for Bayesian identification of stochastic nonlinear systems

The most direct application of Bayesian inference to the identification of nonlinear dynamic systems consists in augmenting the state to include the parameters, and recursively estimating the augmented state. Although this strategy have shown to be effective in input-output identification of nonlinear systems [15, 17, 24, 25], its application to output-only identification of nonlinear systems have shown to be a challenging task.

In this section an offline approach for output-only identification of nonlinear dynamic systems is presented. The proposed approach is based on a re-parameterization of the likelihood function where the state predictive distribution is included, marginalized and estimated recursively using the unscented Kalman filter without state augmentation. This way the joint state-parameter estimation problem is decoupled and performed in two sequential and independent stages: estimation of only the parameters posterior PDF, followed by estimation of the state (without model parameters) using the complete model class weighted by the posterior parameters PDF. Decoupling of the problem increases the numerical stability of the algorithm allowing to increase the dimension of the parameter space to be estimated.

As mentioned before the objective in Bayesian identification is to compute the parameters posterior $p(\boldsymbol{\Phi}|\mathbf{Y}_T)$, where the vector of parameters $\boldsymbol{\Phi}$ partially or completely parameterize the model f in Eq. (1). To this end the likelihood function $p(\mathbf{Y}_T|\boldsymbol{\Phi})$ will first be re-parameterized as follows: $p(\mathbf{Y}_k|\boldsymbol{\Phi})$ based on the first k measurements

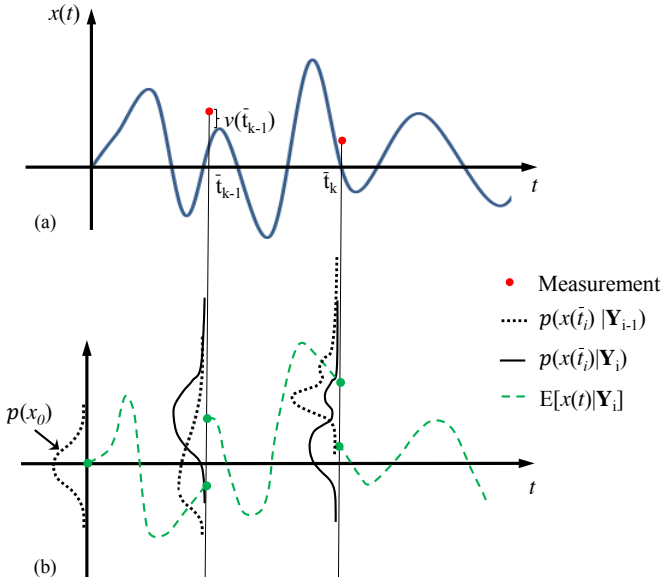

Figure 1: Bayesian state estimation with discrete measurements [21]. (a) system response and noise contaminated measurements. (b) filtering distribution, $p(x(\bar{t}_i)|\mathbf{Y}_i,\Phi)$, and the evolution of the conditional mean, $\mathbb{E}[x(t)|\mathbf{Y}_i]$. The function $p(x(t)|\mathbf{Y}_{i-1})$ for $\bar{t}_{i-1} < t \leq \bar{t}_i$ denotes the state predictive distribution.

($k > 1$) can be written as [30]

$$p(\mathbf{Y}_k|\mathbf{\Phi}) = p(\mathbf{Y}_{k-1},\mathbf{y}_k|\mathbf{\Phi}) = p(\mathbf{y}_k|\mathbf{Y}_{k-1},\mathbf{\Phi})p(\mathbf{Y}_{k-1}|\mathbf{\Phi}) \tag{8}$$

$$= p(\mathbf{Y}_{k-1}|\mathbf{\Phi}) \int p(\mathbf{x}_k,\mathbf{y}_k|\mathbf{Y}_{k-1},\mathbf{\Phi})\mathrm{d}\mathbf{x}_k \tag{9}$$

$$= p(\mathbf{Y}_{k-1}|\mathbf{\Phi}) \int p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{\Phi})p(\mathbf{x}_k|\mathbf{Y}_{k-1},\mathbf{\Phi})\mathrm{d}\mathbf{x}_k \tag{10}$$

where in the last equality the fact that from Eq. (2) it follows that the current measurement conditional in the state is stochastically independent of past measurements was used. The PDF $p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{\Phi})$ is given by the measurement model, while the distribution $p(\mathbf{x}_k|\mathbf{Y}_{k-1},\mathbf{\Phi})$ is recognized as the state prior or state predictive distribution (at step k) in Bayesian state estimation.

Using the recursion in Eq. (10) the likelihood function is given by

$$p(\mathbf{Y}_T|\mathbf{\Phi}) = \prod_{i=1}^{T} \int p(\mathbf{y}_i|\mathbf{x}_i,\mathbf{\Phi})p(\mathbf{x}_i|\mathbf{Y}_{i-1},\mathbf{\Phi})\mathrm{d}\mathbf{x}_i \tag{11}$$

with $p(\mathbf{x}_1|\mathbf{Y_0},\mathbf{\Phi}) = p(\mathbf{x}_1|\mathbf{\Phi})$. Substitution of Eq. (11) in Eq. (4) yields (after removing the normalization constant)

$$p(\mathbf{\Phi}|\mathbf{Y}_T) \propto p(\mathbf{\Phi}) \prod_{i=1}^{T} \int p(\mathbf{y}_i|\mathbf{x}_i,\mathbf{\Phi})p(\mathbf{x}_i|\mathbf{Y}_{i-1},\mathbf{\Phi})\mathrm{d}\mathbf{x}_i \tag{12}$$

To compute $p(\mathbf{x}_i|\mathbf{Y}_{i-1},\mathbf{\Phi})$ state estimation is employed. In Bayesian state estimation the objective is to recursively compute the state posterior $p(\mathbf{x}_i|\mathbf{Y}_i,\mathbf{\Phi})$ (with a fixed parameters vector)

$$p(\mathbf{x}_i|\mathbf{Y}_i,\mathbf{\Phi}) = \frac{p(\mathbf{x}_i|\mathbf{Y}_{i-1},\mathbf{\Phi})p(\mathbf{y}_i|\mathbf{x}_i,\mathbf{\Phi})}{p(\mathbf{y}_i|\mathbf{Y}_{i-1},\mathbf{\Phi})} \tag{13}$$

Note that this is fundamentally different to augmented-state estimation, since the distribution is conditional on

the parameters, instead of the parameters being included in the state. To obtain the state estimation predictive distribution, $p(\mathbf{x}_i|\mathbf{Y}_{i-1},\boldsymbol{\Phi})$, the posterior at the previous step $p(\mathbf{x}_{i-1}|\mathbf{Y}_{i-1},\boldsymbol{\Phi})$ is projected forward in time using the stochastic model as depicted in Fig. 1. There are two fundamental approaches to perform this projection: i) solving the Fokker-Planck equation associated with the continuous model in Eq. (1), which consists in solving a partial differential equation whose solution provides the transition PDF between measurements; or ii) performing nonlinear transformations of random variables using the associated discrete model in Eq. (3). The first approach cannot be employed in applications where the state dimension is relatively large because of the computational effort involved in solving the Fokker-Planck equation. The alternative is to rely on the second approach, in which case the projection is given in principle by

$$p(\mathbf{x}_i|\mathbf{Y}_{i-1},\boldsymbol{\Phi}) = \int p(\mathbf{x}_i|\mathbf{x}_{i-1},\boldsymbol{\Phi})p(\mathbf{x}_{i-1}|\mathbf{Y}_{i-1},\boldsymbol{\Phi})\mathrm{d}\mathbf{x}_{i-1} \tag{14}$$

In general, with the exception of special cases, this integral is intractable. Sub-optimal approximations used to perform the state projection include first-order linearization, Monte Carlo estimation and the unscented transform, resulting respectively in the extended, ensemble and unscented Kalman filters [13, 14]. In this paper the unscented transform and the unscented Kalman filter (discussed in Appendix A) are adopted to estimate $p(\mathbf{x}_i|\mathbf{Y}_{i-1},\boldsymbol{\Phi})$ in a prediction-update fashion, resulting in an estimate that captures the mean and covariance of the nonlinear transformations at least to second order terms in Taylor series expansions.

The main feature of this strategy is the avoidance of state augmentation, which renders a problem with increased numerical stability and reduced ill-posedness. This is due to the increased nonlinearity and complex topology of the augmented-state estimation problem with respect to the parameters-only estimation problem. The decoupling of the augmented-state results in a problem with improved identifiability conditions which allows to increase the dimension of the parameters vector to be estimated. This is numerically illustrated in a further section of the paper.

Once the posterior of the parameters $p(\boldsymbol{\Phi}|\mathbf{Y}_T)$ is computed, the state prediction is achieved using the complete model class, weighted by the parameters posterior PDF using the Theorem of Total Probability

$$p(\mathbf{x}_i|\mathbf{Y}_T) = \int p(\mathbf{x}_i|\mathbf{Y}_T,\boldsymbol{\Phi})p(\boldsymbol{\Phi}|\mathbf{Y}_T)\mathrm{d}\boldsymbol{\Phi} \tag{15}$$

When $i < T$ the distribution $p(\mathbf{x}_i|\mathbf{Y}_T,\boldsymbol{\Phi})$ is equivalent to the distribution obtained using a Bayesian smoother, such as an unscented smoother (with fixed parameters) [36].

In summary, the joint state-parameter estimation problem has been decoupled into two steps: the posterior PDF of the uncertain parameters is first estimated using Eq. (12) and the unscented Kalman filter (discussed in Appendix A), followed by the computation of the posterior PDF of the state using Eq. (15). The main drawback of the proposed approach is that, as an offline algorithm, it cannot be employed in applications where the time evolution of the state and the parameters are needed near real-time, such as control applications.

A potential numerical difficulty in offline Bayesian identification is the need to perform high-dimensional integrals to obtain the marginal distributions of the parameters. An efficient approach consists in approximating the integrals using stochastic simulation methods based on Markov chain Monte Carlo (MCMC) [38, 39]. In MCMC a Markov chain is constructed such that the (unique) stationary distribution of the chain is a target distribution (in this case the parameters joint posterior distribution). MCMC is adopted herein to explore the parameters space and obtain a sample of the posterior distribution, while the unscented Kalman filter is used in a state estimation step to evaluate the likelihood function needed for the progression of the Markov chain. The resulting approach will be referred in the following as the UKF-MCMC approach. Within this approach any MCMC algorithm can be adopted.

4. UKF - MCMC Bayesian Identification Algorithm

An algorithm for the application of the UKF-MCMC approach for system identification is presented next. The MCMC Metropolis-Hastings algorithm is adopted to sample the parameters posterior distribution (other MCMC algorithms can be employed with minor modifications to the algorithm below) [38]. For completeness of the exposition a brief discussion of the Metropolis-Hastings algorithm is included in Appendix B.

UKF - MCMC Algorithm

Select the parameters prior distribution $p(\boldsymbol{\Phi})$

Select the number of samples N and proposal distribution q for MCMC (See Appendix B)

Select the initial state of the Markov chain $\boldsymbol{\Phi} = \boldsymbol{\phi}_0$

Compute $p(\mathbf{x}_j|\mathbf{Y}_{j-1}, \boldsymbol{\phi}_0)$ for all j using Eqs. (A.6)-(A.13)

Compute $\log(p(\boldsymbol{\Phi} = \boldsymbol{\phi}_0|\mathbf{Y}_T))$ using Eq. (12)

$i \rightarrow 1$

While $i < N$

  Propose a move to $\boldsymbol{\Phi} = \boldsymbol{\phi}_i$ using the proposal distribution q (See Appendix B)

  Compute $p(\mathbf{x}_j|\mathbf{Y}_{j-1}, \boldsymbol{\phi}_i)$ for all j using Eqs. (A.6)-(A.13)

  Compute $\log(p(\boldsymbol{\phi}_i|\mathbf{Y}_T))$ using Eq. (12)

  Compute $\log(\alpha_j^i)$ using Eq. (B.1)

  Generate a variate $u \sim uniform[0,1]$

    If $log(u) < \log(\alpha_j^i)$

      Accept the move to $\boldsymbol{\Phi} = \boldsymbol{\phi}_i$

      $i \rightarrow i+1$

    Else

      Propose a different move $\boldsymbol{\Phi} = \boldsymbol{\phi}_i$ using the proposal distribution q (See Appendix B)

    End If

End While

With this algorithm a set of N samples of the posterior $p(\boldsymbol{\Phi}|\mathbf{T}_T)$ is obtained. Using the samples expectations of functions of the parameters are readily obtained using sampling estimates. In this algorithm MCMC is used to explore the parameter space; at each step of the Markov chain a state estimation step is performed using the UKF in order to evaluate the likelihood function and for the progression of the chain. Note that the UKF is used in a state estimation step without state augmentation since the state predictive distribution is conditional in the current point in the parameter space of the Markov chain.

5. Numerical Verification

In this section numerical results aimed at verifying the effectiveness of the proposed approach for Bayesian identification are presented. For this purpose two examples are considered:

i) A bilinear inelastic oscillator subjected to Gaussian white noise
ii) A four degree-of-freedom nonlinear (bilinear-inelastic) chain subjected to a base acceleration consisting of a modulated Kanai-Tajimi stochastic process.

The realization of the input used to generate the synthetic data is assumed to be unknown for identification purposes.

5.1. Example 1: Bilinear inelastic single degree of freedom system

The first numerical verification example consists of a bilinear hysteretic single degree of freedom system driven by Gaussian white noise. Defining the state as $\mathbf{x}(t) = (q(t), \dot{q}(t), z(t))$ the stochastic model is given by

$$\begin{bmatrix} \mathrm{d}q(t) \\ \mathrm{d}\dot{q}(t) \\ \mathrm{d}z(t) \end{bmatrix} = \begin{bmatrix} \dot{q}(t) \\ \frac{c}{m}\dot{q}(t) - f_r(q(t), z(t), D_y, k) \\ f_h(z(t), \dot{q}(t), D_y, \alpha) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mathrm{d}\beta(t) \tag{16}$$

where $q(t) \in \mathbb{R}$ is the displacement, m is the mass, and c is the damping. The function f_r is the restoring force, dependent on the hysteretic variable $z(t) \in \mathbb{R}$ and parameterized by $\{D_y, k\}$, where k is the initial stiffness, D_y is the yield displacement and the post-yield stiffness is $k_y = \alpha k$. The nonlinear hysteretic function f_h defines the springs bilinear hysteretic model; this function can be projected forward in time using a numerical solver, such as the Newmark methods [40]. The system parameters were selected as follows: $m = 1\mathrm{kg}$, $k = 6.14\mathrm{Nm}^{-1}$, $D_y = 0.4\mathrm{m}$ and $\alpha = 0.1$. The damping ratio was selected as $\xi = 0.05$. Previous efforts to estimate the parameters of nonlinear hysteretic models have been reported in the literature using an augmented-state parameter estimation strategy with measured input (see [17, 41, 42] and references therein). In the approach proposed herein state augmentation is not required. The input realization is assumed to be unknown and modeled as a (non-stationary) filtered stochastic process, and thus colored and non-stationary excitations can be treated in a straightforward fashion.

A realization of the input was used to generate the synthetic data; the displacement response history and the force-displacement plots are shown in Fig. 2. The measurement data consists of the oscillator velocity, contaminated by a white Gaussian sequence used to model noise, with a noise to signal root-mean square (RMS) ratio of 0.10.

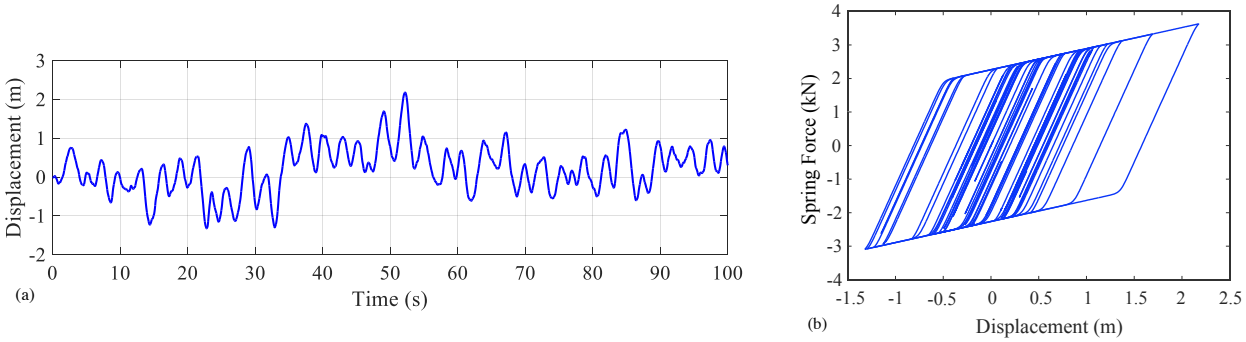

Figure 2: Oscillator response. (a) Displacement history; (b) restoring force-displacement history.

The parameters to be estimated are $\mathbf{\Phi} = [D_y \ \ \alpha \ \ k]$. The prior distributions for the parameters were selected as the following independent distributions: for D_y and α a uniform prior in $[0, 1]$; for k a Gaussian distribution with mean $\mu_k = 10\mathrm{Nm}^{-1}$ and a coefficient of variation of 20%. The prior distributions are depicted in Fig. 3, where the vertical lines denote the true (system) parameters used to generate the data.

The UKF-MCMC approach was implemented as described in section 4. The number of samples used was $N = 1000$, and the initial state of the Markov chain was selected as the mean of the prior distributions. The proposals q for MCMC simulation were selected as independent Gaussian distributions, with mean at the current state of the chain and the following standard deviation: $\sigma_k = 1\mathrm{Nm}^{-1}$, $\sigma_{D_y} = 6 \times 10^{-4}\mathrm{m}$, $\sigma_\alpha = 2 \times 10^{-3}$. For monitoring the convergence and mixing of the chain only the second half of the chain samples were used, and the jumping distribution q was selected so that the acceptance rate is close to 40% [38]. The number of samples was gradually increased until no significant variation in the results was observed. The computational time required to run the algorithm was approximately 3 minutes in a standard desktop computer.

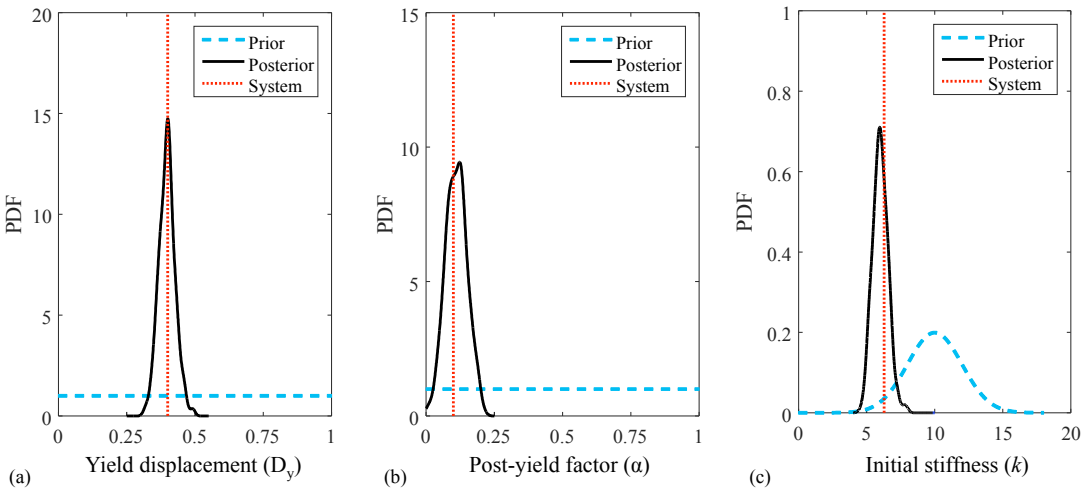

Figure 3: Parameters marginal posterior distributions; (a) Yield displacement, (b) Post-yield factor, (c) Initial stiffness.

The parameters marginal posterior PDF are shown in Fig. 3. As can be seen the updated distributions are consistent with the true (system) parameters. The estimation results are summarized in Table 1, where the mean and 95% mean-centered confidence intervals (C.I.) are shown. For comparison purpose the standard augmented-state UKF was employed in a joint state-input-parameter setting, with the input variables included in the state and the input model included as part of the dynamic model. The mean estimates and confidence intervals given by the UKF are also summarized in Table 1. It can be seen that the proposed UKF-MCMC approach shows an increased accuracy with respect to the augmented-state UKF.

Table 1: Estimation results summary.

		UKF - MCMC			UKF		
	System	Mean	95% C.I.	Error	Mean	95% C.I.	Error
Yield displacement - D_y	0.40	0.40	$[0.34, 0.46]$	0%	0.41	$[0.35, 0.47]$	2.5%
Post-yield factor - α	0.10	0.11	$[0.04, 0.19]$	10%	0.17	$[0.10, 0.24]$	70%
Initial stiffness - k	6.14	6.01	$[5.00, 7.20]$	2.1%	5.85	$[5.19, 6.51]$	4.7%

The estimates of the displacement response and force-displacement hysteresis are shown, respectively, in Fig. 4 and Fig. 5. As can be seen, the estimates of the displacement provided by both approaches are in agreement with the true displacement, which is due to the robustness of the UKF in state estimation [21]. However, the error in the estimated nonlinear model parameters translates to errors in the hysteresis loops.

As mentioned before the standard UKF was formulated as a joint state-input-parameter estimation problem, similarly to the implementation in Ref. [26] but without transforming the state to modal coordinates. For this purpose the input variables are included in the state and the input model is included as part of the dynamic model.The decline in accuracy of the UKF with respect to other studies in the literature stems from the fact that in general the UKF is implemented either in an output-only fashion in *linear* systems [26], or in an input-output fashion in nonlinear systems where it is assumed that the major component of the input causing the nonlinearity is known (up to an additive noise) [17, 24]. In this paper we are interested in output-only identification of nonlinear systems subjected to colored excitations. The decline in accuracy and implementation issues of the standard UKF in an output-only setting are exacerbated when the state dimension increases, as shown in the following example where a four degree-of-freedom chain-type system is considered.

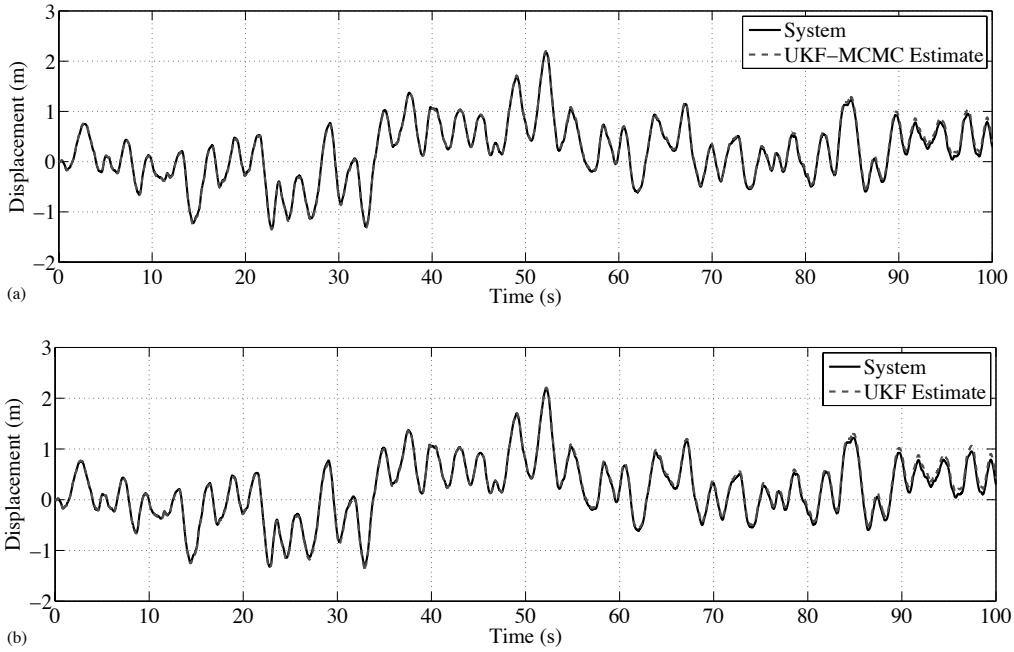

Figure 4: Displacement response estimates. (a) UKF-MCMC approach; (b) UKF approach.

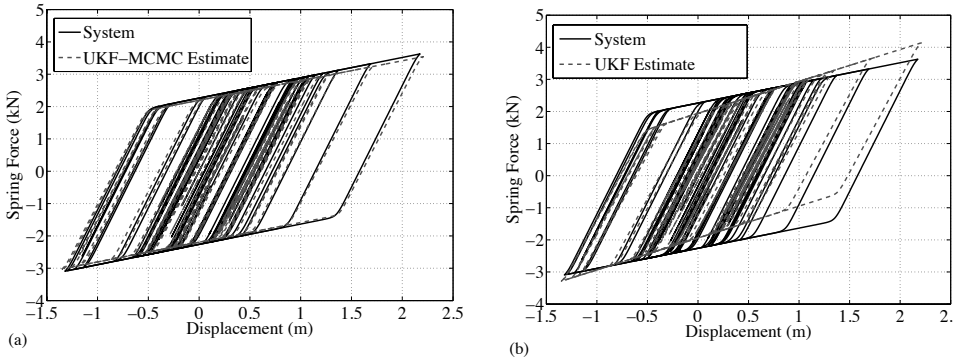

Figure 5: Force-displacement hysteresis estimates. (a) UKF-MCMC approach; (b) UKF approach.

5.2. Example 2: Bilinear-inelastic four degree of freedom chain system

The second numerical example consists of a four degree of freedom nonlinear chain system with a bilinear model of hysteresis for each spring. The model is depicted in Fig. 6; for spring i the initial stiffness is k_i, the yield displacement is D_{y_i}, and the post-yield stiffness is $k_{y_i} = \alpha_i k_i$. The chain is excited by an uncertain input ground motion consisting of a modulated Kanai-Tajimi process. In this model the input excitation is given by

$$\ddot{u}_{gm}(t) = I(t)\left[\omega_g^2 u_g(t) + 2\xi_g \omega_g \dot{u}_g(t)\right] \tag{17}$$

where $u_g(t)$ is the solution of

$$\ddot{u}_g(t) + 2\xi_g \omega_g \dot{u}_g(t) + \omega_g^2 u_g(t) = w(t) \tag{18}$$

where w is formally a white noise process with spectral density $S_{ww}(\omega) = G_o$, and $I(t)$ is an amplitude modulating

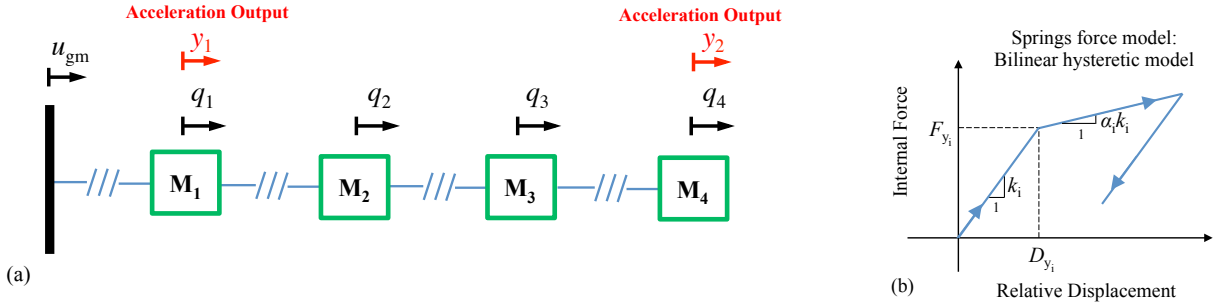

Figure 6: Model with restoring force characteristics; (a) four degree-of-freedom nonlinear chain system, (b) springs force model.

function. The process $\ddot{u}^*(t) = \ddot{u}_{gm}(t)/I(t)$ has a Kanai-Tajimi power spectral density given by

$$S_{\ddot{u}^*\ddot{u}^*}(\omega) = G_o \frac{1+4\xi_g^2\left(\frac{\omega}{\omega_g}\right)^2}{\left[1-\left(\frac{\omega}{\omega_g}\right)^2\right]^2+4\xi_g^2\left(\frac{\omega}{\omega_g}\right)^2} \qquad -\infty<\omega<\infty \tag{19}$$

Defining the state as $\mathbf{x}(t) = (\mathbf{q}(t), \dot{\mathbf{q}}(t), \mathbf{z}(t), \mathbf{u}_g(t), \dot{\mathbf{u}}_g(t))$ the stochastic model is given by

$$\begin{bmatrix} \mathrm{d}\mathbf{q}(t) \\ \mathrm{d}\dot{\mathbf{q}}(t) \\ \mathrm{d}\mathbf{z}(t) \\ \mathrm{d}u_g(t) \\ \mathrm{d}\dot{u}_g(t) \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{q}}(t) \\ -\mathbf{M}^{-1}\mathbf{C}\dot{\mathbf{q}}(t) - \mathbf{M}^{-1} f_r(\mathbf{q}(t), \mathbf{z}(t), \mathbf{D}_y, \mathbf{k}) - \mathbf{r}I(t)\left[\omega_g^2 u_g(t) + 2\xi_g\omega_g\dot{u}_g(t)\right] \\ f_h(\mathbf{z}(t), \dot{\mathbf{q}}(t), \mathbf{D}_y, \boldsymbol{\alpha}) \\ \dot{u}_g(t) \\ -2\xi_g\omega_g\dot{u}_g(t) - \omega_g^2 u_g(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \mathrm{d}\beta(t) \tag{20}$$

where $\mathbf{q}(t) \in \mathbb{R}^4$ is the displacement vector, $\mathbf{M}$ is the mass matrix, $\mathbf{C}$ is the damping matrix and f_r is the restoring force vector parameterized by $\{\mathbf{D}_y, \mathbf{k}\}$; the vector $\mathbf{r}$ is an influence vector. The hysteretic variable $\mathbf{z}(t) \in \mathbb{R}^4$ is used to define the nonlinear hysteretic model f_h. The true (system) parameters were selected as follows: $\mathbf{M} = diag([1\ 1\ 1\ 1])$kg, $\mathbf{k} = [1000\ \ 950\ \ 850\ \ 750]\mathrm{Nm}^{-1}$, $\mathbf{D}_y = [0.12\ \ 0.10\ \ 0.09\ \ 0.07]$m, $\boldsymbol{\alpha} = [0.1\ \ 0.1\ \ 0.1\ \ 0.1]$. The input model parameters were chosen as $\xi_g = 0.35$ and $\omega_g = 10\mathrm{rad\,s}^{-1}$, selected based on calibration results with past recorded ground motions (see e.g., [43], and references therein). A modal damping matrix was used with a damping ratio of 5% in all modes. The response measurements used for system identification consists of accelerations at degrees of freedom 1 and 4, contaminated by a Gaussian white sequence to model noise, with a noise to signal RMS of 0.10.

The parameters to be estimated are $\boldsymbol{\Phi} = \{\mathbf{D}_y, \mathbf{k}, \boldsymbol{\alpha}\}$. The prior distribution for the parameters were selected as the following independent distributions: for $\mathbf{D}_y$ and $\boldsymbol{\alpha}$ a uniform prior in $[0,1]$; for $\mathbf{k}$ a Gaussian distribution with mean $\boldsymbol{\mu}_k = [1200\ \ 1200\ \ 900\ \ 900]\mathrm{Nm}^{-1}$ and a coefficient of variation of 10%. The prior distributions for the stiffness parameters are depicted in Fig. 7, where the red lines denote the true (system) parameters used to generate the data; since the posterior distributions of $\mathbf{D}_y$ are highly peaked the priors were omitted for ease of visualization due to the scale of the plot.

The UKF-MCMC approach was implemented as described in section 4, using the same parameters of the previous

example. For monitoring the convergence and mixing of the chain only the second half of the chain samples were used, and the jumping distribution q was selected so that the acceptance rate is close to 40% [38]. The number of samples was gradually increased until no significant variation in the results was observed. The computational time required to run the algorithm was approximately 8 minutes in a standard desktop computer. The computational effort can be considerably reduced by employing more efficient MCMC methods. This will be studied in further work.

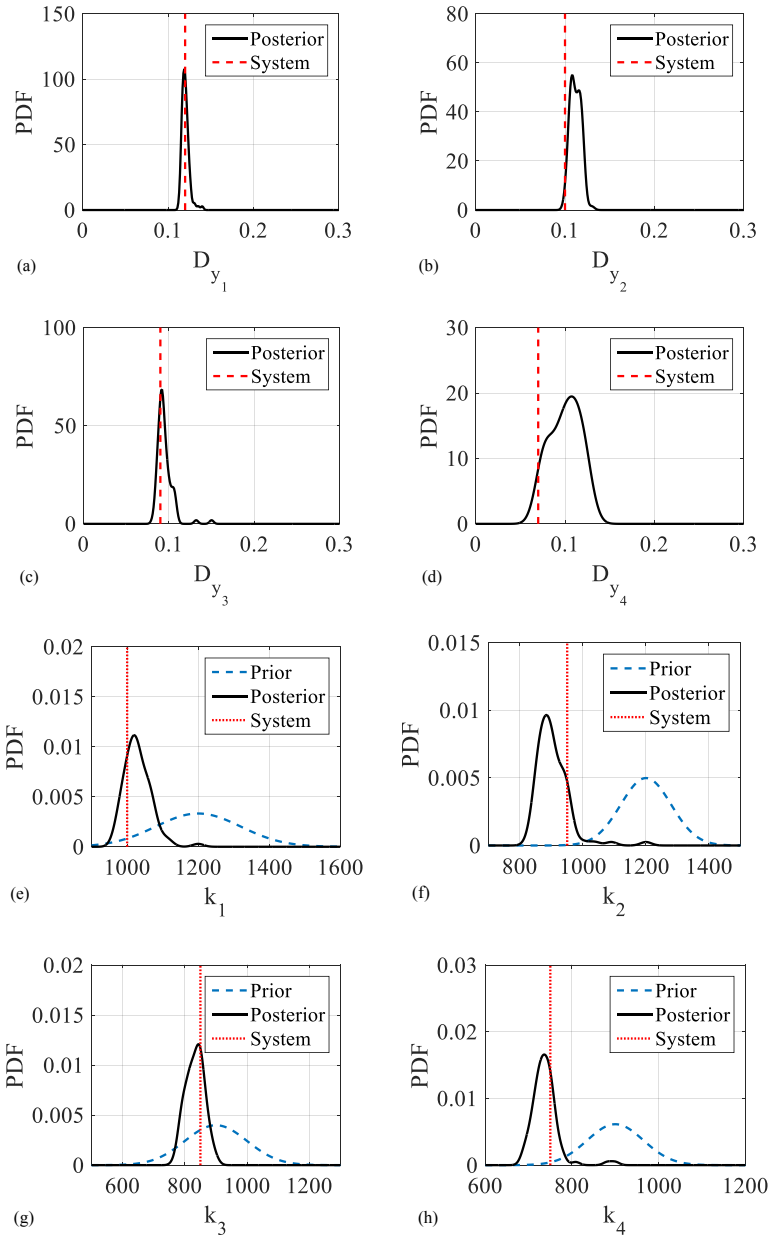

Figure 7: UKF-MCMC estimate of the parameters marginal posterior distributions. (a)-(d) Yield displacements, (e)-(h) Initial stiffnesses.

The parameters marginal distributions estimated using the UKF-MCMC algorithm are depicted in Fig. 7, where it can be seen that the estimates are consistent with the 'true' system parameters. The estimation results are summarized in Table 2, where the mean and 95% confidence intervals are shown. The joint bivariate distributions of the yield

Table 2: UKF-MCMC estimates statistics.

		UKF-MCMC		
	System	Mean	95% C.I.	Error
D_{y_1}	0.12	0.12	$[0.11, 0.13]$	0%
D_{y_2}	0.10	0.11	$[0.10, 0.12]$	10%
D_{y_3}	0.09	0.09	$[0.08, 0.11]$	0%
D_{y_4}	0.07	0.10	$[0.06, 0.14]$	42%
k_1	1000	1030	$[982, 1078]$	3%
k_2	950	890	$[821, 958]$	6%
k_3	850	833	$[777, 890]$	2%
k_4	750	734	$[692, 777]$	2%

displacements are depicted in Fig. 8. The figure shows the complex topologies characterizing the posterior distribution of the nonlinear chain system. No evident strong correlations are observed between the yield displacements. The estimates of the dynamic response (displacements, velocities and spring forces) at measured and unmeasured locations are depicted in Fig. 9, where it can be observed that the estimates are in good agreement with the system response. To study the robustness of the approach to parametric errors in the input model, an error factor in the range $[0.75, 1.25]$ was applied to the 'true' system parameters for the estimation. No significant variations in the results were observed for factors in this range. The effect of parametric model errors will be addressed in the following section where a recorded ground motion is used in an experimental setup.

The standard augmented-state UKF was also employed in this example as a joint state-input-parameter estimation problem, however, the estimates diverged. Specifically the covariance matrix became non-positive definite and the algorithm collapsed. As mentioned before this issue is attributed to identifiability of the parameters and the nature of the model in the augmented-state formulation where the parameters are considered as variables. The high accuracy of the UKF approach in nonlinear systems has been demonstrated in applications where the main component of the input generating the nonlinearity is known [17, 24]. However, to the best knowledge of the authors, this has not been demonstrated in the output-only identification of non-stationary nonlinear systems of the type considered herein.

The proposed UKF-MCMC approach does not show this divergence because the problem is not formulated as a joint state-parameter estimation problem. Instead, we rely on a strategy based on state estimation using unscented Kalman filtering without state augmentation to compute the parameters posterior, which is sampled using a Markov chain Monte Carlo method. The algorithm shows an increased robustness and stability with respect to the standard UKF because of the reduced ill-posedness of the decoupled state estimation inverse problem.

6. Experimental Validation

In this section the UKF-MCMC approach is validated using vibration data from an experiment performed by a team of researchers from the University of California at San Diego (UCSD). The validation seeks to study the robustness of the approach in the presence of inherent modeling errors in the description of the system and forcing input. The experiment consisted of a densely instrumented full-scale reinforced concrete shear wall structure tested in the NEES-UCSD shake table [31, 44].

6.1. Test Structure and Model Description

The test structure consisted of a seven story full-scale reinforced concrete cantilever shear wall, resembling a portion of a typical building with a shear wall lateral force-resisting system [44]. The structure was designed using the displacement-based capacity approach resulting in smaller design forces than those given by current force-based design codes.

The gravity load system consisted of a reinforced concrete slab supported by high-strength columns. The structure was supported by a reinforced concrete foundation, which was connected to the shake table using post-tensioned cables. The test structure on the shake table is shown in the left panel of Figure 10. The structure total height was

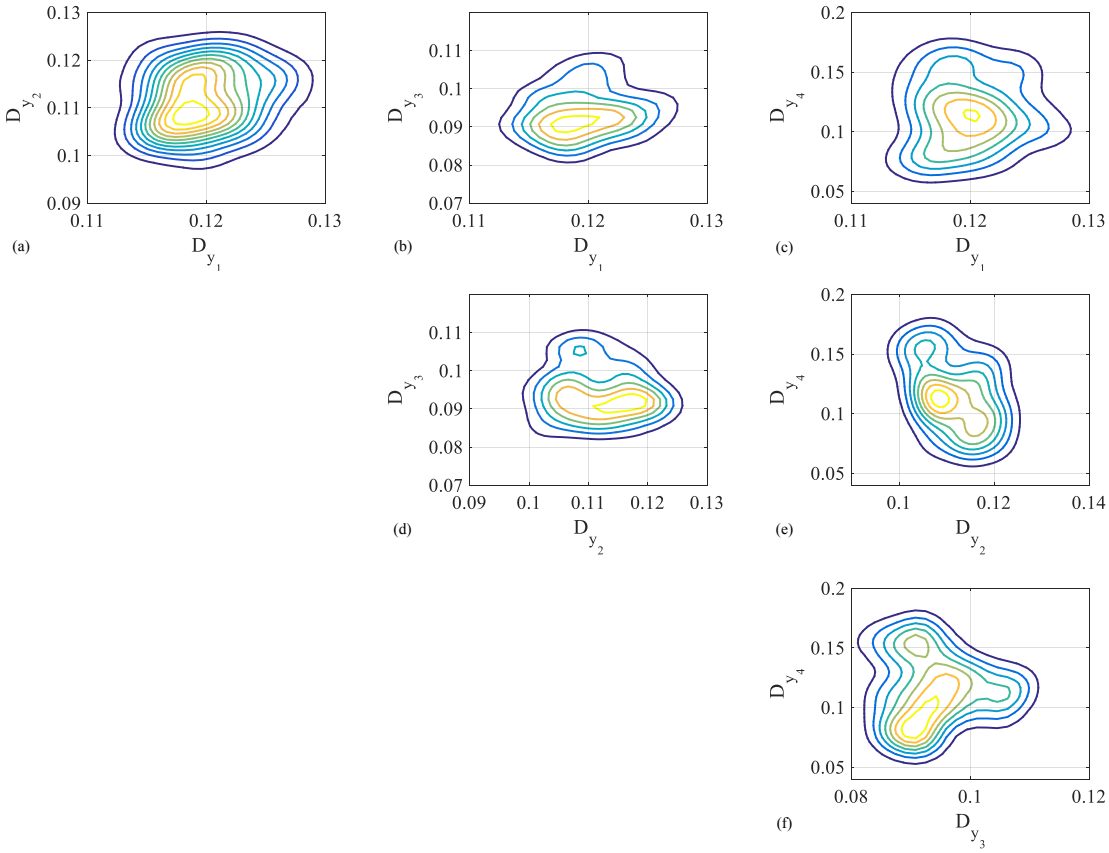

Figure 8: UKF-MCMC estimates of the yield displacement bivariate posterior distributions.

approximately 20m. The web wall was 0.30m wide, with a thickness of 0.20m in levels 1 and 7, and 0.15m in levels 2-6. The web wall was connected in one end to a 5m wide flange wall with a thickness of 0.20m in level 1 and 0.15m in levels 2-7. The web wall was connected to the flange wall using a slotted connection to prevent coupling between the two walls. On the other end the shear wall was connected to a post-tensioned gravity column by pinned braces (both the column and braces were designed to remain elastic during the test); the purpose of this column was to increase the structure torsional rigidity to reduce out-of-plane effects with respect to the loading direction.

The model used for structural identification consists of a nonlinear chain connected to a linear elastic cantilever. A bilinear inelastic (hysteretic) model is adopted for the chain springs internal force - displacement behavior, while the the cantilever (assumed to have no mass) has flexural stiffness k_f; the model with the stiffness characteristics is depicted in the right panel of Fig. 10. Damping is modeled as viscous modal damping, constructed as the superposition of modal damping matrices computed using the initial stiffness matrix; this model is usually preferred to avoid the spurious forces introduced by the Rayleigh damping model [45]. The analytic model equation is given by Eq. (20), with the vectors adjusted to the adequate dimension and an added term to account for the cantilever stiffness. Similar models have been previously proposed to estimate the dynamic response of building structures in earthquake engineering applications [33, 46, 47]. The motivation for employing this class of models is that the lateral response in buildings is characterized by a combination of flexural and shear deformations. It has been shown that using only shear-type models may result in significant errors in building response estimates [48]. This has also been shown to be the case in structures in which shear walls are used as the main lateral deformation resisting mechanism [46]. The use of an adequate stochastic identification model class that is able to accurately capture the response of the structure of interest has been found to be of prime importance for the successful application of Bayesian identification approaches

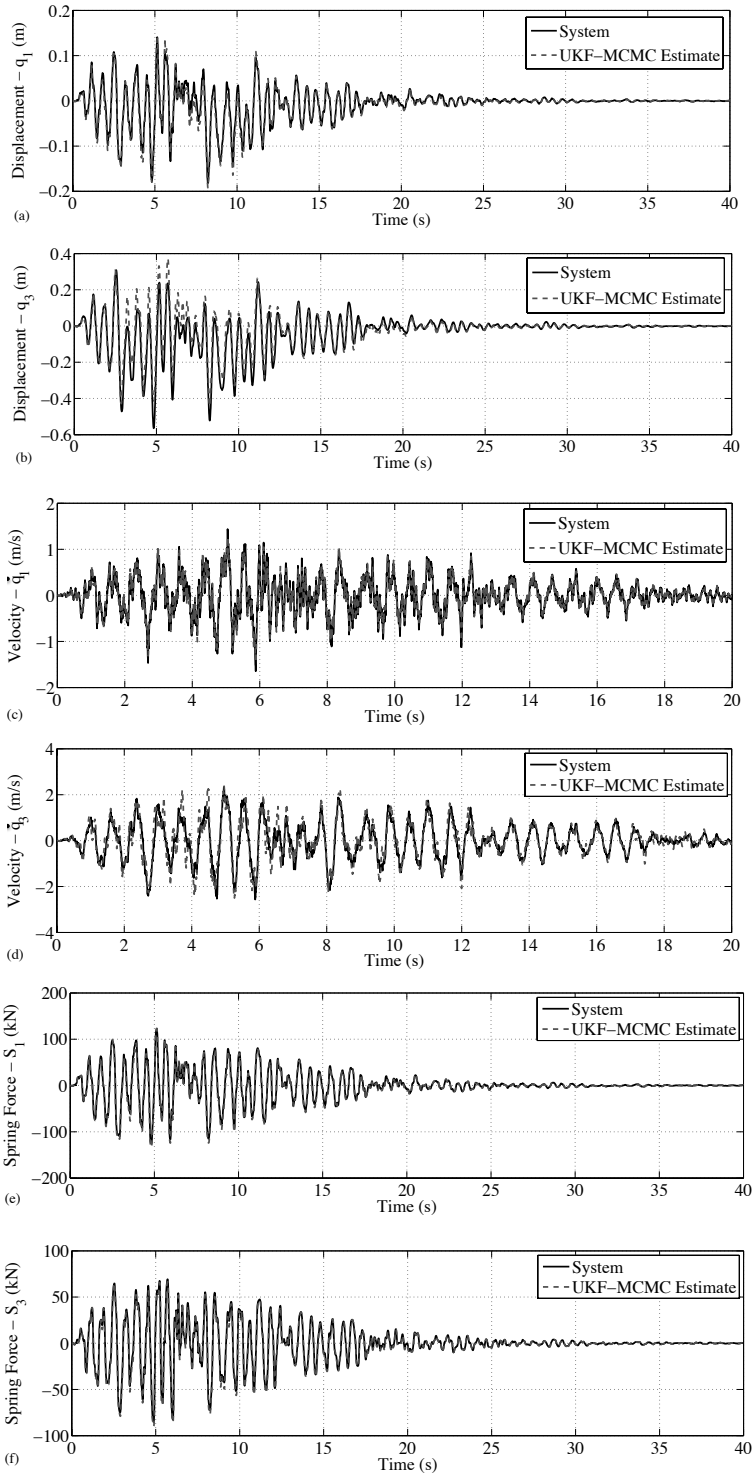

Figure 9: UKF-MCMC dynamic response estimates. (a) Displacement - DOF 1; (b) Displacement - DOF 3; (c) Velocity - DOF 1; (d) Velocity - DOF 3; (e) Spring 1 Force; (f) Spring 3 Force.

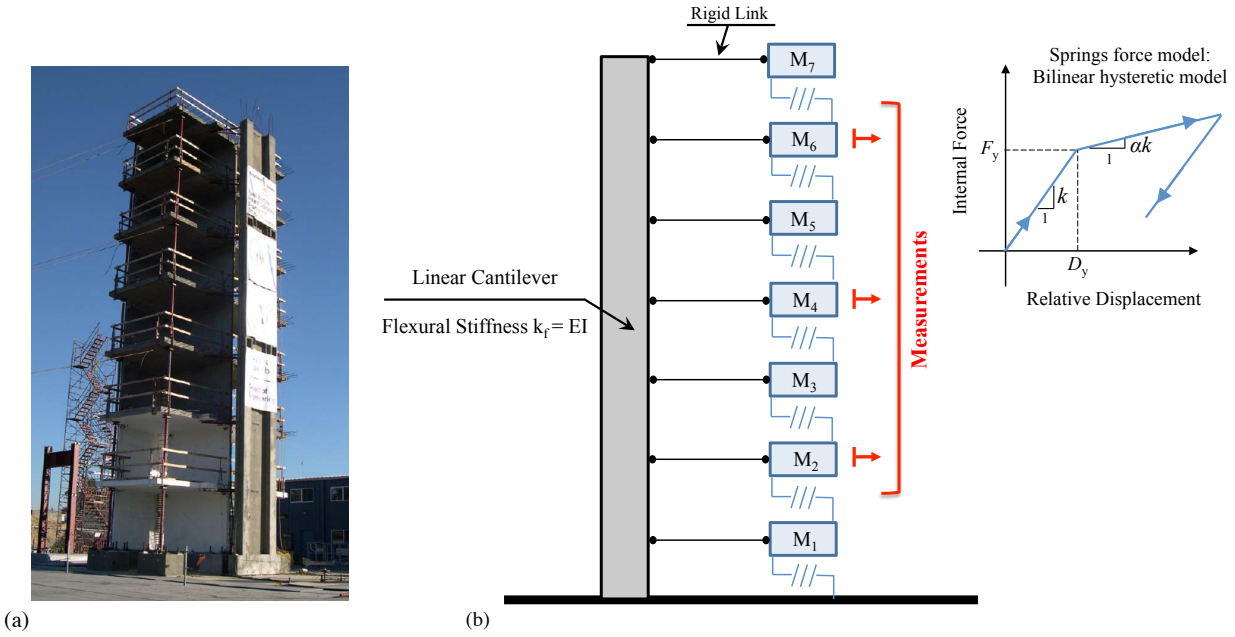

Figure 10: Experimental validation. (a) Test shear wall structure [31]; (b) Model consisting of a nonlinear inelastic chain coupled to a linear cantilever.

[15, 21].

6.1.1. Instrumentation and Testing Program

Strong-motion data was obtained from a dense array of sensors deployed throughout the structure. The instrumentation included 139 accelerometers, 88 displacement transducers (LVDT), 314 strain gages and 23 pressure transducers sampled at 240 Hz [44].

During the testing program the structure was sequentially subjected to four recorded earthquake ground motion time histories: i) the Van Nuys longitudinal component record of 1971 San Fernando earthquake, ii) Van Nuys transverse component record of 1971 San Fernando earthquake, iii) Oxnard Boulevard in Woodland Hill longitudinal component of 1994 Northridge earthquake, and iv) the Sylmar Olive View Med 360 component record from 1994 Northridge earthquake. Before and after each earthquake the structure was subjected to banded (0.25-25 Hz) white noise excitations with varying RMS of 0.02g, 0.03g and 0.05g. In addition, ambient vibration measurements were recorded at the different stages.

6.2. Identification Results

In this section the UKF-MCMC approach is applied for the identification of the structure depicted in Fig. 10. To obtain the data used herein the structure was subjected to the Van Nuys transverse component record of 1971 San Fernando earthquake depicted in Fig. 11; the record is assumed to be unknown for identification purposes. The ground motion model adopted is the Kanai-Tajimi model discussed in the previous section, with the same parameters values. Recall that the input model parameters were selected based on the calibration results with past recorded ground motions performed by Pires et al. [43], which did not include the record in Fig. 11. The output measurements used for identification consist of the acceleration response at degrees of freedom 2, 4 and 6 (see Fig. 10).

The coupled nonlinear chain-cantilever model depicted in the right panel of Fig. 10 is used for structural identification. The mass at each degree of freedom is $M_i = 3.1 \times 10^4 \mathrm{Ns}^2\mathrm{m}^{-1}$, found using tributary mass and the structural drawings. The parameters vector to be estimated is $\boldsymbol{\Phi} = \left[k_f, k, \xi, D_y\right]$, where k_f is the flexural stiffness of the cantilever, k is the initial stiffness of the nonlinear springs, D_y is the yield displacement of the springs, and ξ is the

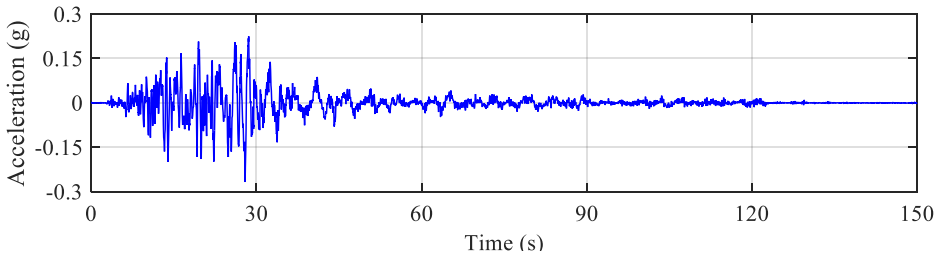

Figure 11: Van Nuys transverse component record of 1971 San Fernando earthquake.

damping ratio of all modes. These are the parameters that dominate the structure behavior and whose uncertainty might result in the inadequate prediction of the dynamic response.

The parameters prior distributions were selected based on structural drawings, design specifications, common engineering practice and previous work by the authors as the following independent distributions: for k_f a Gaussian distribution with mean $\mu_{k_f} = 5 \times 10^7 \mathrm{Nm}^2$ and a coefficient of variation of 30%; for k a Gaussian distribution with mean $\mu_k = 1.3 \times 10^7 \mathrm{Nm}^{-1}$ and a coefficient of variation of 10%; for ξ a uniform distribution in $[0, 0.20]$; and for D_y a uniform distribution in $[0, 0.15]$ m. As mentioned before an advantage of the proposed approach is that it allows the use of any desired probability distributions, in contrast to standard augmented-state unscented Kalman filtering estimation that inherits from Kalman filtering the need to use only Gaussian distributions. The prior distributions are depicted in Fig. 12.

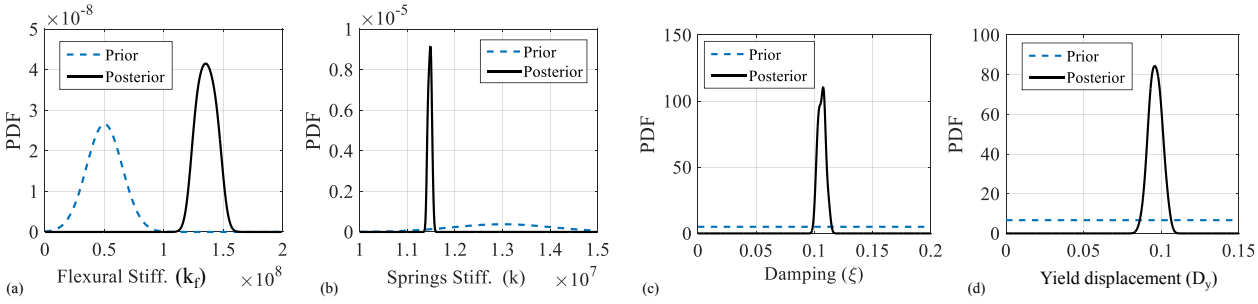

Figure 12: UKF-MCMC estimate of the parameters marginal posterior distributions; (a) Flexural stiffness, (b) Springs stiffness, (c) Damping ratio, (d) Yield displacement.

The UKF-MCMC approach was employed to estimate the parameters vector $\mathbf{\Phi}$ as described in section 4, with $N = 1000$ samples, and the initial state of the Markov chain selected as the mean of the prior distributions. The proposals for MCMC were selected as independent Gaussian distributions, with mean at the current state of the chain and standard deviations tuned such that the acceptance rate is close to 40%; for monitoring the convergence and mixing of the chain only the second half of the chain samples were used [38]. The number of samples was gradually increased until no significant variation in the results was observed. The parameters for the UKF state estimation step were selected as follows: the process noise spectral density (related directly to the covariance matrix) is $G_0 = 3.15$ which results from the input model adopted, while the measurement noise covariance matrix was selected based on a noise-to-signal RMS (root-mean-square) ratio of 0.01, selected based on the measurement confidence level provided by the sensors manufacturer. The computational time required to run the algorithm was approximately 7 minutes in a standard desktop computer. The UKF-MCMC estimates of the parameters posterior distribution are depicted in Fig. 12. The estimation results are summarized in Table 3. As can be seen, the respective priors and posteriors differ substantially, indicative of gain in information from the response measurements. Indeed, some information gain measures depend on the variation of these two distributions [7].

Table 3: Structural identification summary.

Parameter	Posterior	
	Mean	C.I.
Flexural Stiffness - $k_f = EI\,(\mathrm{Nm}^2)$	1.357×10^8	$[1.215 \times 10^8, 1.498 \times 10^8]$
Springs Stiffness - $k\,(\mathrm{Nm}^{-1})$	1.147×10^7	$[1.140 \times 10^7, 1.154 \times 10^7]$
Damping ratio - ξ	0.106	$[0.100, 0.112]$
Yield Displacement - $D_y\,(\mathrm{m})$	0.096	$[0.088, 0.104]$

The identified damping mean $\xi = 0.106$, which is larger than typical values observed in concrete structures, is attributed to the high level of cracking at the stage of the testing program that this test was conducted, resulting in an increase of energy dissipation due to crack opening.

The identified model was used to estimate the dynamic response of the structure, with the input ground motion record still assumed to be unknown, and only using the Kanai-Tajimi stochastic model described before for the estimation. The estimate of the displacement and acceleration at the top of the structure are depicted in Fig. 13. Measurements of these quantities are shown for comparison purposes; the displacement was measured using a GPS. As can be seen the estimates are in well agreement with the measurements, indicative of the consistency with the data of the identified parameters. Similarly to the second numerical example, the standard augmented-state UKF failed to estimate the parameters and the algorithm collapsed. As discussed before, nonlinear identification using the augmented-state UKF approach has major limitations when the input is unmeasured.

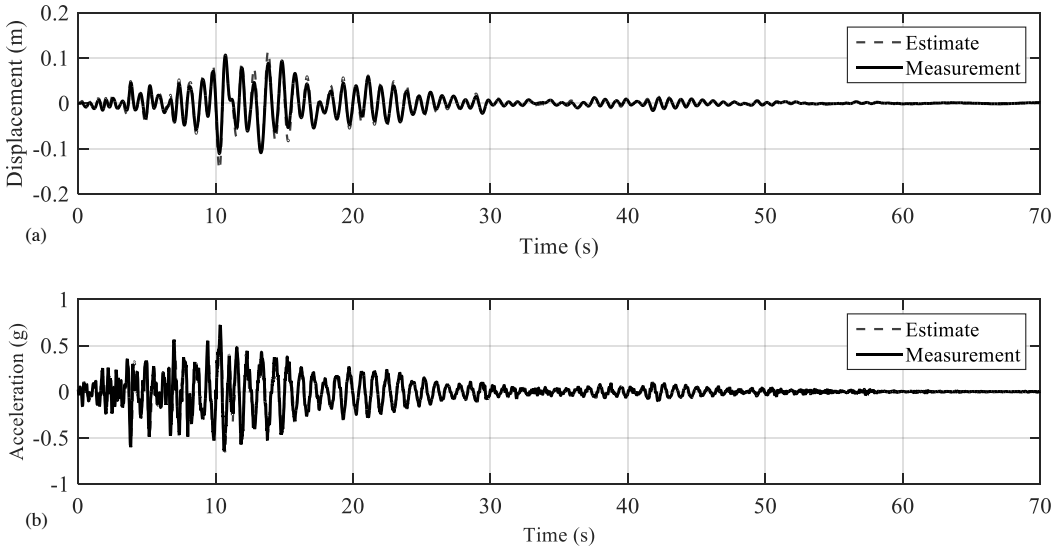

Figure 13: Response estimates at the top of the structure. (a) Displacement response; (b) Acceleration response.

7. Conclusion

In this paper an output-only approach for Bayesian identification of stochastic nonlinear systems subjected to non-stationary inputs was presented. The approach is based on a re-parameterization of the parameters joint posterior distribution. In the re-parameterization the state predictive distribution is included, marginalized, and estimated recursively in a state estimation step using an unscented Kalman filter, bypassing the state augmentation required

by existing on-line methods. To compute expectations of functions of the parameters the posterior is sampled using Markov chain Monte Carlo (MCMC). The main advantages of the proposed approach are:

- The estimation is performed using only output measurements, i.e., the realization of the input generating the nonlinearity is assumed to be unknown (unmeasured).
- Capability to handle non-Gaussian prior and posterior distributions, in contrast to standard augmented-state unscented Kalman filtering where only the mean and covariance of the posterior are estimated.
- Increased numerical stability with respect to augmented-state unscented Kalman filtering when the number of response measurements is limited and the forcing input is unknown. This is because in the augmented-state problem the parameters are treated as variables, which typically results in a likelihood function with increased nonlinearity, several local minima, and a complex topology. Under these conditions the estimates of the covariance matrix may become rank-deficient due to observability/identifiability related issues. Avoidance of state augmentation renders a less ill-conditioned inverse problem.

The proposed approach was verified using two numerical examples: i) a bilinear hysteretic single degree of freedom system subjected to Gaussian white noise; and ii) a four degree of freedom chain with a bilinear model of hysteresis subjected to a base excitation consisting of a modulated Kanai-Tajimi process. In both examples the forcing input used to generate the synthetic data was assumed to be unknown. It was shown that the proposed UKF-MCMC approach has the capability to consistently estimate the parameters correctly, providing confidence intervals to characterize the uncertainty in the parameters. The approach was also validated experimentally using vibration data from a full-scale shake table experiment performed by a team of researchers from the University of California at San Diego. It was shown that the approach is robust to modeling errors in the description of the structure and the forcing input, providing consistent estimates of the parameters and accurate prediction of the dynamic response of the structure. In future work the approach will be implemented using high-fidelity nonlinear finite element models and different MCMC methods to study and improve its computational efficiency.

8. Acknowledgment

The authors acknowledge access to the "Shake Table Response of Full Scale Reinforced Concrete Wall Building Slice" project database (DOI:10.4231/D35T3G04T) of the George E. Brown Jr Network for Earthquake Engineering Simulation (NEES) Project Warehouse, which is available to the research community under the ODC Attribution License. Access to the database is gratefully acknowledged.

Appendix A. State Estimation using Unscented Kalman filtering

The unscented Kalman filter (UKF) is a Bayesian filtering algorithm for nonlinear systems based on the application of the unscented transform (UT) to obtain the prior distribution or state predictive distribution. The UT approach uses a set of deterministic sampled points (known as sigma points) to parameterize the mean and covariance of nonlinear transformations of random variables [14]. The sigma points are propagated by the nonlinear transformation and the resulting vectors are used to estimate the mean and covariance of the transformed variable. This approach is able to capture the mean and covariance (at least) to the second order terms of the Taylor series expansion for any nonlinear function [14].

Let $\mathbf{x}_{k-1}|\mathbf{Y}_{k-1}$ denote the state posterior at $t = t_{k-1}$, with mean and covariance $\hat{\mathbf{x}}_{k-1}$ and $\widehat{\mathbf{P}}_{x_{k-1}x_{k-1}}$ respectively. To estimate the statistics of propagating this distribution and the statistics of the measurement prediction the following set of $2n+1$ sigma vectors $\boldsymbol{\chi}_i$ is employed

$$\boldsymbol{\chi}_0 = \hat{\mathbf{x}}_{k-1} \tag{A.1}$$

$$\boldsymbol{\chi}_i = \hat{\mathbf{x}}_{k-1} + \left(\sqrt{(n+\lambda)\widehat{\mathbf{P}}_{x_{k-1}x_{k-1}}}\right)_i \qquad i = 1,\ldots,n \tag{A.2}$$

$$\boldsymbol{\chi}_i = \hat{\mathbf{x}}_{k-1} - \left(\sqrt{(n+\lambda)\widehat{\mathbf{P}}_{x_{k-1}x_{k-1}}}\right)_{i-n} \qquad i = n+1,\ldots,2n \tag{A.3}$$

with corresponding weights given by

$$W_0 = \lambda/(n+\lambda) \tag{A.4}$$

$$W_i = 1/2(n+\lambda) \qquad i = 1,\ldots,2n \tag{A.5}$$

where n is the dimension of the state and λ a parameter controlling the spread of the vectors in an n-dimensional sphere. The terms in parenthesis are the columns of the matrix square root of the scaled covariance matrix. Choosing $\lambda = 3-n$, at least second order accuracy is achieved in both the mean and covariance estimate.

The sigma points are projected by the nonlinear transformation and the state prior mean and covariance estimates are given by

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n} W_i \boldsymbol{\chi}_i^- \qquad \text{where} \qquad \boldsymbol{\chi}_i^- = f_d(\boldsymbol{\chi}_i, \boldsymbol{\Phi}) + \mathbf{B}_{d_1}\mathbf{u}_{k-1} \tag{A.6}$$

$$\widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{x}_k}^- = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\chi}_i^- - \hat{\mathbf{x}}_k^-\right)\left(\boldsymbol{\chi}_i^- - \hat{\mathbf{x}}_k^-\right)^T + \mathbf{B}_{d_2}\mathbf{Q}_{d_{k-1}}\mathbf{B}_{d_2}^T \tag{A.7}$$

$$\hat{\mathbf{y}}_k^- = \sum_{i=0}^{2n} W_i \boldsymbol{\Upsilon}_i \qquad \text{where} \qquad \boldsymbol{\Upsilon}_i = h(\boldsymbol{\chi}_i^-, \mathbf{u}_k, \boldsymbol{\Phi}) \tag{A.8}$$

$$\widehat{\mathbf{P}}_{\mathbf{y}_k\mathbf{y}_k}^- = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\Upsilon}_i - \hat{\mathbf{y}}_k^-\right)\left(\boldsymbol{\Upsilon}_i - \hat{\mathbf{y}}_k^-\right)^T + \mathbf{R} \tag{A.9}$$

$$\widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{y}_k}^- = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\chi}_i^- - \hat{\mathbf{x}}_k^-\right)\left(\boldsymbol{\Upsilon}_i - \hat{\mathbf{y}}_k^-\right)^T \tag{A.10}$$

where $\boldsymbol{\Phi}$ is fixed, and the matrices $\mathbf{B}_{d_1}$ and $\mathbf{B}_{d_2}$ depend on the discretization scheme employed. To estimate the mean and covariance of the posterior, $\hat{\mathbf{x}}_k$ and $\widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{x}_k}$, the Kalman filter equations are used

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k^-) \tag{A.11}$$

$$\widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{x}_k} = \widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{x}_k}^- - \mathbf{K}_k\widehat{\mathbf{P}}_{\mathbf{y}_k\mathbf{y}_k}^-\mathbf{K}_k^T \tag{A.12}$$

$$\mathbf{K}_k = \widehat{\mathbf{P}}_{\mathbf{x}_k\mathbf{y}_k}^-\left(\widehat{\mathbf{P}}_{\mathbf{y}_k\mathbf{y}_k}^-\right)^{-1} \tag{A.13}$$

The Kalman filter is the optimal (i.e., the minimum mean-square error) linear estimator for nonlinear estimation problems [14]. When the state dimension is greater than three, the covariance estimate given by the previous algorithm might be non-positive semi-definite. The scaled unscented transformation was developed to address this issue by using a new set of sigma points obtained after applying the original ones to an auxiliary nonlinear transformation [14]. The resulting algorithm has an additional parameter to control the scaling of the points and avoids the possibly non-positive semi-definite covariance. The new sigma points (and their corresponding weights) are given by

$$\boldsymbol{\chi}_i^* = \boldsymbol{\chi}_0 + \alpha(\boldsymbol{\chi}_i - \boldsymbol{\chi}_0) \tag{A.14}$$

$$W_0^* = (1/\alpha^2)W_0 + (1 - 1/\alpha^2) \tag{A.15}$$

$$W_i^* = (1/\alpha^2)W_i \tag{A.16}$$

where $i = 1,\ldots,2n$ and $\alpha \in [0,1]$ is the new scaling parameter which depends on the nonlinear functions involved. The computation of the original and modified sigma points can be combined in a single step to reduce the number of computations.

Appendix B. Markov chain Monte Carlo using the Metropolis-Hastings algorithm

The solution to the Bayesian system identification problem is given by the parameters posterior PDF. The parameters marginal distributions of this joint PDF are of interest in applications to characterize the parameters and their

uncertainty. This requires the evaluation of multidimensional integrals, usually unfeasible to solve in closed-form or using numerical integration. Recent developments in stochastic simulation have pointed to the use of Markov chain Monte Carlo (MCMC) to perform this task. In this section a brief review of MCMC using the Metropolis-Hastings algorithm is presented.

The objective of MCMC is to construct a Markov chain whose stationary distribution matches a distribution of interest (in this case the parameters posterior PDF). It can be shown that the stationary distribution is an eigenvector of the Markov chain transition matrix with unit eigenvalue. A sufficient condition to achieve this is that the Markov chain is reversible (satisfies the *detailed balance condition*). Some side conditions need to be satisfied, namely, irreducibility, aperiodicity and recurrency (these conditions are usually satisfied in practical applications). For further discussions of Markov chains the interested reader can consult [6] and references therein. The following algorithm yields a transition matrix that satisfies the detailed balance condition [38]:

Metropolis-Hastings algorithm

- Start at an arbitrary point $\boldsymbol{\Phi} = \boldsymbol{\phi}_i$ in the parameter space
- Generate a variate $\boldsymbol{\Phi} = \boldsymbol{\phi}_j$ from a proposal distribution $q(j|i)$ representing a move from state i to state j
- Compute
$$\log(\alpha_j^i) = \min\left\{0, \log\left(\frac{q(i|j)p(\boldsymbol{\phi}_j|\mathbf{Y}_T)}{q(j|i)p(\boldsymbol{\phi}_i|\mathbf{Y}_T)}\right)\right\} \tag{B.1}$$
- Generate a variate $u \sim uniform[0,1]$
- If $\log(u) < \log(\alpha_j^i)$ accept the proposed move to $\boldsymbol{\Phi} = \boldsymbol{\phi}_j$
  Otherwise propose a move to a different state $\boldsymbol{\Phi} = \boldsymbol{\phi}_i$

The process is repeated until the target number of samples N is obtained. Since the chain is started at an arbitrary point in the parameter space, it takes a number of iterations before the chain starts exploring the high probability region of the posterior ("burn-in period"). Thus, the convergence of the chain must be monitored. Some recommendations for applications include discarding the first half of the samples and tuning the jumping distribution q so that the acceptance rate is close to 40% [38].

References

[1] L. Ljung, System Identification: Theory for the user, Prentice Hall, New Jersey, USA, 1999.
[2] T. Söderström, P. Stoica, System Identification, Prentice Hall International, United Kingdom, 2001.
[3] J.-P. Noël, G. Kerschen, Past, present and future of nonlinear system identification in structural dynamics, Mechanical Systems and Signal Processing 20 (2006) 505–592.
[4] G. Kerschen, K. Worden, A. Vakakis, J.-C. Golinval, Past, present and future of nonlinear system identification in structural dynamics, Mechanical Systems and Signal Processing 20 (2006) 505–592.
[5] K. Yuen, J. Beck, Updating properties of nonlinear dynamical systems with uncertain input, Journal of Engineering Mechanics 129 (1) (2003) 9–20.
[6] M. Grigoriu, Stochastic systems: Uncertainty quantification and propagation, Springer, London, 2012.
[7] J. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring 17 (2010) 825–847.
[8] K. Worden, J. Hensman, Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference, Mechanical Systems and Signal Processing 32 (2012) 153–169.
[9] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, Mechanical Systems and Signal Processing.
[10] B. Goller, G. Schueller, Investigation of model uncertainties in Bayesian structural model updating, Journal of Sound and VIbration 330(25) (2011) 6122–6136.
[11] H. Sun, O. Buyukozturk, Probabilistic updating of building models using incomplete modal data, Mechanical Systems and Signal Processing 75 (2016) 27–40.
[12] H. Sun, A. Mordret, G. Prieto, M. Nafi, O. Buyukozturk, Bayesian characterization of buildings using seismic interferometry on ambient vibrations, Mechanical Systems and Signal Processing 85 (2017) 468–486.
[13] G. Evensen, The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean Dynamics 53 (2003) 343–367.
[14] S. Julier, J. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92 (3).

[15] J. Ching, J. Beck, K. Porter, R. Shaikhutdinov, Bayesian state estimation method for nonlinear systems and its application to recorded seismic response, Journal of Engineering Mechanics 132 (4) (2006) 396–410.
[16] V. Namdeo, C. Manohar, Nonlinear structural dynamical system identification using adaptive particle filters, Journal of Sound and Vibration 306 (2007) 524–563.
[17] E. Chatzi, A. Smyth, The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing, Structural Control and Health Monitoring 16 (2009) 99–123.
[18] Z. Xie, J. Feng, Real-time nonlinear structural system identification via iterated unscented Kalman filter, Mechanical Systems and Signal Processing 28 (2012) 309–322.
[19] S. Azam, A. Ghisi, S. Mariani, Parallelized sigma-point Kalman filtering for structural dynamics, Computers and Structures 92-93 (2012) 193–205.
[20] R. Astroza, H. Ebrahimian, J. Conte, Material parameter identification in distributed plasticity FE models of frame-type structures using nonlinear stochastic filtering, Journal of Engineering Mechanics 141 (5) (2015) 04014149.
[21] K. Erazo, E. Hernandez, Uncertainty quantification of state estimation in nonlinear structural systems with application to seismic response in buildings, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems (2015) B5015001.
[22] K. Erazo, E. Hernandez, Bayesian model-data fusion for mechanistic postearthquake assessment of building structures, Journal of Engineering Mechanics (04016062) (2016) 10.1061/(ASCE)EM.1943–7889.0001114.
[23] K. Erazo, E. Hernandez, State estimation in nonlinear structural systems, Nonlinear Dynamics. Proceedings of the 32nd IMAC, A Conference and Exposition on Structural Dynamics 2 (2014) 249–257.
[24] M. Wu, A. Smyth, Application of the unscented Kalman filter for real-time nonlinear structural system identification, Structural Control and Health Monitoring 14 (2007) 971–990.
[25] R. Ghanem, G. Ferro, Health monitoring for strongly non-linear systems using the ensemble Kalman filter, Structural Control and Health Monitoring 13 (2006) 245–259.
[26] S. Azam, E. Chatzi, C. Papadimitriou, A dual Kalman filter approach for state estimation via output-only acceleration measurements, Mechanical Systems and Signal Processing 60-61 (2015) 866–886.
[27] K. Maes, A. Smyth, G. DeRoeck, G. Lombaert, Joint input-state estimation in structural dynamics, Mechanical Systems and Signal Processing 70-71 (2016) 445–466.
[28] E. Lourens, C. Papadimitriou, S. Gillijns, E. Reynders, G. DeRoeck, G. Lombaert, Joint input-response estimation for structural systems based on reduced-order models and vibration data from a limited number of sensors, Mechanical Systems and Signal Processing 29 (2012) 310–327.
[29] C. Andrieu, A. Doucet, R. Holenstein, Particle Markov chain Monte Carlo methods, Journal of the Royal Statistical Society 72 (3) (2010) 269–342.
[30] M. Khalil, A. Sarkar, S. Adhikari, D. Poirel, The estimation of time-invariant parameters of noisy nonlinear oscillatory systems, Journal of Sound and Vibration 344 (2015) 81–100.
[31] M. Panagiotou, J. Restrepo, J. Conte, Shake table response of 7-story rc bearing wall building, Network for Earthquake Engineering Simulation (distributor), Dataset, DOI:10.4231/D35T3G04T.
[32] B. Moaveni, X. He, J. Conte, J. Restrepo, M. Panagiotou, System identification study of a seven-story full-scale building slice tested on the ucsd-nees shake table, Journal of Structural Engineering 137 (6) (2011) 705–717.
[33] K. Erazo, E. Hernandez, High-resolution seismic monitoring of instrumented buildings using a model-based state observer, Earthquake Engineering and Structural Dynamics (2016) DOI 10.1002/eqe.2781.
[34] K. Erazo, Bayesian filtering in nonlinear structural systems with application to structural health monitoring, University of Vermont, Graduate College Dissertations and Theses (2015) 3688814.
[35] P. Kloeden, E. Platen, Numerical solution of stochastic differential equations, Springer, 1999.
[36] S. Särkkä, Bayesian filtering and smoothing, Cambridge University Press, England, 2013.
[37] M. Chatzis, E. Chatzi, A. Smyth, On the observability and identifiability of nonlinear structural and mechanical systems, Structural Control and Health Monitoring (2014) DOI: 10.1002/stc.1690.
[38] A. Gelman, J. Carlin, H. Stern, D. Rubin, Bayesian Data Analysis, Chapman and Hall CRC, Boca Raton, Florida, USA, 2004.
[39] J. Beck, S.-K. Au, Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, Journal of Engineering Mechanics 128 (4) (2002) 380–391.
[40] A. Chopra, Dynamics of Structures, Prentice-Hall International Series, 2016.
[41] S. Li, Y. Suzuki, M. Noori, Identification of hysteretic systems with slip using a bootstrap filter, Mechanical Systems and Signal Processing 18 (2004) 781–795.
[42] M. Hoshiya, E. Saito, Structural identification by extended Kalman filter, Journal of Engineering Mechanics 110 (12) (1984) 1757–1770.
[43] J. Pires, Y. Wen, A. Ang, Stochastic analysis of liquefaction under earthquake loading, Civil Engineering Studies UIUC.
[44] M. Panagiotou, J. Restrepo, J. Conte, Shake table test of a 7-story full scale reinforced concrete structural wall building slice phase i: Rectangular wall section, UCSD Report No. SSRP-07-07.
[45] A. Chopra, F. Mckenna, Modeling viscous damping in nonlinear response history analysis of buildings for earthquake excitation, Earthquake Engineering and Structural Dynamics 45 (2016) 193–211.
[46] F. Khan, J. Sbarounis, Interaction of shear walls and frames, Journal of the Structural Division 90 (3) (1964) 285–335.
[47] Y. Park, A. Ang, Y. Wen, Seismic damage analysis and damage-limiting design of R.C. buildings, Technical Report of Research, Civil Engineering Studies No. 516 UIUC.
[48] A. Chopra, E. Cruz, Evaluation of building code formulas for earthquake forces, ASCE Journal of Structural Engineering 112 (1986) 1881–1899.