A Discontinuous Extended Kalman Filter for Non-Smooth Dynamic Problems

M.N. Chatzis [a], **E.N. Chatzi** [b] and **S.P. Triantafyllou** [c]

[a] Associate Professor in Dept. of Engineering Science, the University of Oxford, manolis.chatzis@eng.ox.ac.uk

[b] Assistant Professor in Dept. of Civil, Environmental and Geomatic Engineering, ETH Zurich, chatzi@ibk.baug.ethz.ch

[c] Assistant Professor in Dept. of Civil Engineering, the University of Nottingham, Savvas.Triantafyllou@nottingham.ac.uk

1. Abstract

Problems that result into locally non-differentiable and hence non-smooth state-space equations are often encountered in engineering. Examples include problems involving material laws pertaining to plasticity, impact and highly non-linear phenomena. Estimating the parameters of such systems poses a challenge, particularly since the majority of system identification algorithms are formulated on the basis of smooth systems under the assumption of observability, identifiability and time invariance. For a smooth system, an observable state remains observable throughout the system evolution with the exception of few selected realizations of the state vector. However, for a non-smooth system the observable set of states and parameters may vary during the evolution of the system throughout a dynamic analysis. This may cause standard identification (ID) methods, such as the Extended Kalman Filter, to temporarily diverge and ultimately fail in accurately identifying the parameters of the system. In this work, the influence of observability of non-smooth systems to the performance of the Extended and Unscented Kalman Filters is discussed and a novel algorithm particularly suited for this purpose, termed the Discontinuous Extended Kalman Filter (DEKF), is proposed.

2. Introduction

Systems with pronounced non-linearities are often encountered in engineering. The task of accurately identifying the parameters of such systems is often challenging. For one, it is well known that the convergence of commonly employed methods, such as the Extended Kalman Filter, i.e., the most widely employed extension of the Kalman Filter ([1]) to non-linear systems, depends on the initial values assumed for the states, the parameters and the covariance matrix. An improvement of the EKF, namely the Unscented Kalman Filter, was suggested by Julier and Uhlmann in [2]. This variant achieves rapid convergence by additionally alleviating the need to evaluate derivative quantities and Jacobians.

An implied assumption of any system identification method is however that the dynamic states of the system and the time-invariant parameters are observable ([1, 3]) and identifiable ([4, 5]) respectively. In other words, the augmented state vector created by the underlying dynamic states and the parameters is observable ([6, 7]). While a non-linear system with smooth state-space and measurement equations may either be observable or unobservable for a specific measurement setup, the same does not apply for systems with non-differentiable state-space equations. In fact, it was shown in [8] that non-smooth systems that can be separated into smooth branches may result into some of the parameters being identifiable within some branches and unidentifiable in others. This work also demonstrated how, despite the local unidentifiability of certain parameters at a given time interval, the parameters of the overall system may still be identified.

However, as noted in [9], the Kalman-Filter is expected to diverge for unobservable states or parameters and the same would apply for its non-linear alternatives, the EKF and UKF, for the case of unobservable non-linear problems. Modifications of the Kalman filter that may allow for the simultaneous identification of the input force ([10]) and methods based on observers of similar nature ([11, 12]) are also liable to such effects. In the case of non-smooth sys-

tems in particular, the fact that a parameter may be unidentifiable over some time interval, may also result in the divergence of the predicted values when employing these methods during this interval. Since these methods have been developed under the assumption of observability for all states and hence identifiability of the parameters, the overall convergence of the algorithms is inevitably adversely affected. It is further noted that within the context of engineering, non-smooth systems are often associated with plastic response, impact or sliding and phenomena pertaining to damage propagation and failure. Identifying the latter is the topic of interest of several recent works, e.g., [13, 14, 15].

In this work, the effect of the observability properties of non-smooth systems in the convergence of the EKF and UKF is studied. Moreover, a modified version of the EKF is suggested, which is able to take the piecewise notion of observability of these systems into consideration. Based on this approach, the filter operates exclusively on observable states within respective intervals, while the parameters that are unidentifiable during these intervals are maintained time invariant. The method is termed the Discontinuous Extended Kalman Filter, $DEKF$.

The proposed method is compared against the EKF and UKF for selected non-smooth problems that involve material plasticity and impact. The examples demonstrate that the suggested approach substantially outperforms the standard EKF in such problems, further illustrating the key role of observability for non-smooth problems. Useful conclusions on why standard methods, such as the EKF and UKF may diverge in such problems are drawn.

3. Non-Smooth Dynamical Systems

A non-linear system with state variables $\mathbf{x}_t$, time-invariant parameters $\boldsymbol{\theta}$, known input vector $\mathbf{u}$, and measurement vector $\mathbf{y}$ can in general be described by the following system of equations:

$$\dot{\mathbf{x}}_t = E(\mathbf{x}_t, \boldsymbol{\theta}, \mathbf{u}), \qquad \dot{\boldsymbol{\theta}} = 0, \qquad \mathbf{y} = G(\mathbf{x}_t, \boldsymbol{\theta}, \mathbf{u}) \tag{1}$$

where E and G designate the non-linear state-space and measurement functions respectively. For the purposes of System Identification, the state-space and measurement equations shown in equation (1) can be written in an augmented form by introducing the state vector $\mathbf{x} = [\mathbf{x}_t, \boldsymbol{\theta}]$:

$$\dot{\mathbf{x}} = e(\mathbf{x}, \mathbf{u}), \qquad \mathbf{y} = g(\mathbf{x}, \mathbf{u}) \tag{2}$$

In the latter representation one treats both the dynamic states and the parameters of the system as states of the augmented system. A dynamical system is further characterized as analytic, or smooth, when the state-space equations (2) are continuous and infinitely differentiable. Very often however the state-space equations of physical models may not be analytic, either due to discontinuities in the state-space equation or in their derivatives. In this paper, we deal with models for which the state-space equations are continuous, but not differentiable, and whose state-space equations can be separated into smooth, i.e., continuous and infinitely differentiable, branches of the form:

$$\begin{aligned} \dot{\mathbf{x}} &= e_1(\mathbf{x}), \text{ when } \mathbf{x} \in R_1^n \\ &\vdots \\ \dot{\mathbf{x}} &= e_l(\mathbf{x}), \text{ when } \mathbf{x} \in R_l^n \end{aligned} \tag{3}$$

where $e_i(\mathbf{x})$ is an analytic set of functions within R_i^n. It should be noted that at a specific time instance the state has a given realization corresponding to a single branch of equation (3). As the system evolves dynamically over time, it is expected to shift between the individual branches. This transition between branches will be referred to as a dynamic event, and the corresponding time instance as the time of the event.

3.1. Observability of Non-Smooth Dynamical Systems

The augmented representation of equation (2) admits the implementation of observability assessment tools ([3, 16]) on the augmented system ([17, 7, 6]) in

order to deduce the observability of both the dynamic state $\mathbf{x}_t$ and parameter vector $\boldsymbol{\theta}$. As discussed in [8], for a smooth system that is observable all the states are observable and the time-invariant parameters in $\boldsymbol{\theta}$ are identifiable. On the other hand, if a parameter is unobservable, it is unidentifiable, and may not be identified via a system identification procedure. It is reminded that the terms observability and identifiability refer to the states and parameters being at least locally observable and identifiable, while the term unobservability and unidentifiability signify that the states or parameters do not have the corresponding properties locally, as more thoroughly explained in [8]. Furthermore, the property of identifiability considered in this paper guarantees finiteness of solutions for that parameter, but not uniqueness (i.e., global identifiability) and does not attempt to enumerate the number of finite solutions, as for example is performed in the work of [18, 19].

The previous remarks however are directly applicable to the case where the state-space and measurement equations of the system are at least analytical, i.e., infinitely differentiable. For the systems examined herein this condition is not satisfied. The observability of such systems has been discussed in [8]. The method proposed in that work, involves the study of the observability of each of the smooth subsystems. Since each subsystem is analytic within that branch, geometric observability algorithms can be used to deduce their observability, as for example the *Observability Rank Condition* (*ORC* [3]). The algorithm results into characterizing the system corresponding to each branch either as observable, for which all the states are observable, and hence the parameters are identifiable, or as unobservable, which means that not all states are observable, and hence not all parameters are necessarily identifiable. In general, separating the states of an analytic system into observable and unobservable sets requires a non-linear transformation ([20]). However, for the systems examined herein it is further assumed that for each of the subsystems i, we can further separate the state vector $\mathbf{x}$ into observable and unobservable components, denoted as $\mathbf{x}^{oi}$ and $\mathbf{x}^{ui}$, in a straightforward manner.

If the union of the observable components from all subsystems is a strict sub-

set of the state vector $\mathbf{x}$ ($\cup_{i=1}^{l}\mathbf{x}^{oi} \subset \mathbf{x}$), i.e., does not contain at least one of the components of $\mathbf{x}$, then it may be concluded that these uncontained components of $\mathbf{x}$ are unobservable and cannot be identified via a System Identification algorithm. If on the other hand the union of the observable components is the state vector $\mathbf{x}$, $\cup_{i=1}^{l}\mathbf{x}^{oi} = \mathbf{x}$, then each component of the state vector $\mathbf{x}$ could potentially be identified within the corresponding smooth branch within which it is observable. Hence, if the response of the system includes at least one branch for which a parameter is identifiable, then a system identification algorithm could potentially succeed in identifying the value of that parameter. In this paper, the latter case of systems is studied, i.e., systems for which the parameters of the model may be inferred via an appropriate system identification method.

4. Extended Kalman Filter

The Extended Kalman Filter, EKF, algorithm is an extension of the standard Kalman Filter ([1]) to non-linear systems. Let us assume a dynamical system whose discrete state-space and measurement equations are written as:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1}, \qquad \mathbf{y}_k = h(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{v}_k \tag{4}$$

where $\mathbf{w}_k$ is the process noise and $\mathbf{v}_k$ is the observation noise, both of which are considered to be white Gaussian noise processes with covariance matrices $\mathbf{Q}$ and $\mathbf{P}$ respectively. The filter then involves the steps included in Table 1.

As discussed in [2], the EKF algorithm propagates the mean value and covariance of the Gaussian random vector $\mathbf{x}$ by linearizing the system around the mean at a specific time step. Thus, at a specific step of the algorithm, the time and measurement update steps are based on a single realization of the state vector, i.e., the estimated mean value of the distribution $\hat{\mathbf{x}}_k$. The real realization of the state vector at that step, $\mathbf{x}_k$, lies in a specific subspace R_i^n, of R^n, and hence the corresponding smooth state-space equations are the ones corresponding to subsystem i of equation (3). It is now assumed that $\mathbf{x} \in R_i^n$ for

Table 1: The steps of the EKF algorithm.

EKF
Initialization at time t_0: $\hat{\mathbf{x}}_0 = \mathbb{E}[\mathbf{x}_0]$
• Time-Update: 1. Predicted mean and covariance: $$\hat{\mathbf{x}}_{k\|k-1} = f(\hat{\mathbf{x}}_{k-1\|k-1}, \mathbf{u}_{k-1})$$ $$\mathbf{P}_{k\|k-1} = F_{k-1}\,\mathbf{P}_{k-1\|k-1}\,F_{k-1}^T + \mathbf{Q} \quad (5)$$ where $F_{k-1} = \frac{\partial f}{\partial \mathbf{x}}\|_{(\hat{\mathbf{x}}_{k\|k-1}, \mathbf{u}_{k-1})}$ and $\mathbf{Q}$ is the process noise matrix
• Measurement Update: 2. Calculation of Kalman Gain: $\mathbf{S}_k = H_k\,\mathbf{P}_{k\|k-1}\,H_k^T + \mathbf{R}$ $\mathbf{K}_k = \mathbf{P}_{k\|k-1}\,H_k^T\,(\mathbf{S}_k)^{-1}$ where $H_k = \frac{\partial h}{\partial \mathbf{x}}\|_{(\hat{\mathbf{x}}_{k\|k-1}, \mathbf{u}_{k-1})}$ and $\mathbf{R}$ is the observation noise matrix 3. Improve predictions of the state and covariance using the latest observations: $$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{K}_k\,(\mathbf{z}_k - h(\mathbf{x}_{k\|k-1}))$$ $$\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{K}_k\,H_k)\,\mathbf{P}_{k\|k-1} \quad (6)$$

a series of consecutive time steps defined in the time window $[t_1,\, t_2]$. During this time interval, the states can be separated into the observable and unobservable part $\mathbf{x}^{oi}$ and $\mathbf{x}^{ui}$. Moreover, during this interval the EKF cannot be expected to converge towards an accurate estimation of the $\mathbf{x}^{ui}$. This has already been noticed in [9] for the case of smooth unobservable systems. Hence, during such intervals we can at best expect for the observable part of the state, $\mathbf{x}^{oi}$, to converge.

This however raises the question of how to efficiently treat the unobservable part during such an interval. The focus of this paper is on systems for which the unobservable states are a subset of the model parameters, which hence are unidentifiable, and it is argued that the best option is to update only the identifiable parameters via the EKF, while retaining the estimates for the mean values of $\mathbf{x}^{ui}$ constant. This calls for the implementation of a modified version of the EKF for the non-smooth systems examined here.

5. Unscented Kalman Filter

The UKF succeeds in simulating non-linear behavior by approximating the state as a Gaussian random variable (GRV), represented by a set of carefully chosen deterministic points known as the Sigma Points. This section only provides a basic overview of the filter equations; more details can be found in [2, 21] and previous work of the authors ([22, 23, 24]).

Consider the general dynamical system described by equations (4). Given the state vector at step $k-1$ and assuming that this has a mean value of $\hat{\boldsymbol{x}}_{k-1}$ and covariance $\mathbf{P}_{k-1}$, we can calculate the statistics of $\boldsymbol{x}_k$ by using the Unscented Transformation, or in other words by computing the set of $2L+1$ sigma points $\boldsymbol{\chi}_k^i$ with associated weights W_i. The steps of the method are summarized in Table 5

At this point it should be noted that in comparison to the EKF, the UKF calculates the mean and standard deviation without the need to linearize the state-space or measurement equations. This results in a more accurate propagation of these properties and usually in a faster convergence rate of the method in comparison to the EKF. However, the unobservable states $\mathbf{x}^{ui}$ may still diverge during the corresponding intervals. Unlike the EKF, the sigma points used by the UKF do not necessarily lie in a single system branch at a given time step and hence the observability properties might differ for the subsystem corresponding to each sigma point. However, it should be reminded that the real dynamic system lies at that time within a single smooth branch i of corresponding unobservable states $\mathbf{x}^{oi}$. The overall convergence of the method is ensured only when a parameter converges faster during identifiable time steps, than it diverges during unidentifiable steps.

6. The Discontinuous Extended Kalman Filter

As noted in the previous sections, during a specific time instance only part of the state vector may be observable and therefore the EKF algorithm is expected to converge only for that observable part $\mathbf{x}^{oi}$. The predictions furnished during

UKF
Initialization at time t_0: $\hat{\mathbf{x}}_0 = \mathbb{E}[\mathbf{x}_0]$
• The Unscented Transform 1. Augment the state vector to include the noise parameters: $\boldsymbol{x}^{\alpha}_{k-1} = [\boldsymbol{x}^T_{k-1}\ \boldsymbol{w}^T_{k-1}\ \mathbf{v}^T_{k-1}]^T$ 2. Formulation of the Sigma Point vector: $\boldsymbol{\chi}^{\alpha}_{k-1} = [\hat{\boldsymbol{x}}^{\alpha}_{k-1} \quad \hat{\boldsymbol{x}}^{\alpha}_{k-1} + \sqrt{(L+\lambda)\mathbf{P}^{\alpha}_k} \quad \hat{x}^{\alpha}_{k-1} - \sqrt{(L+\lambda)\mathbf{P}^{\alpha}_k}]$ where λ is a UKF parameter, L is the dimension of the state vector x and $\mathbf{P}^{\alpha} = diag(\mathbf{P}, \mathbf{Q}, \mathbf{R})$
• Time-Update: 3. Propagation of the Sigma points through the system model: $\boldsymbol{\chi}^{i}_{k|k-1} = f(\boldsymbol{\chi}^{i}_{k-1}, \boldsymbol{\chi}^{w,i}_{k-1}), i = 0, .., 2L$ 4. Predicted mean and covariance: $\hat{\boldsymbol{x}}_{k|k-1} = \sum_{i=0}^{2L} W^m_i \boldsymbol{\chi}^{i}_{k|k-1}$ and $\mathbf{P}_{k|k-1} = \sum_{i=0}^{2L} W^c_i [\boldsymbol{\chi}^{i}_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}][\boldsymbol{\chi}^{i}_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}]^T$
• Measurement Update: 5. Measurement Mean: $\hat{\boldsymbol{y}}_{k|k-1} = \sum_{i=0}^{2L} W^m_i \mathscr{Y}^{i}_{k|k-1}$ and $\mathscr{Y}_{k|k-1} = h(\boldsymbol{\chi}^{i}_{k|k-1}, \boldsymbol{\chi}^{\eta,i}_{k-1})$ 6. Calculation of Kalman Gain: $\mathbf{K}_k = \mathbf{P}^{xy}_k (\mathbf{P}^{yy}_k)^{-1}$ where: $\mathbf{P}^{yy}_k = \sum_{i=0}^{2L} W^c_i [\mathscr{Y}^{i}_{k|k-1} - \hat{\boldsymbol{y}}_{k|k-1}][\mathscr{Y}^{i}_{k|k-1} - \hat{\mathbf{y}}_{k|k-1}]^T$ and $\mathbf{P}^{xy}_k = \sum_{i=0}^{2L} W^c_i [\boldsymbol{\chi}^{i}_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}][\mathscr{Y}^{i}_{k|k-1} - \hat{\boldsymbol{y}}_{k|k-1}]^T$ 7. Improve predictions of the state and covariance using the latest observations: $\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \mathbf{K}_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})$ $\mathbf{P}_k = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}^{yy}_k \mathbf{K}^T_k$

Table 2: The steps of the UKF algorithm.

this interval by the EKF for the unobservable part $\mathbf{x}^{ui}$, which in this work is assumed to be the unidentifiable parameters, are non-optimal and it is also quite likely that during these time intervals the values of $\mathbf{x}^{ui}$ may very well diverge from the real solutions. In fact, these are expected to be less optimal than the initial value of that parameter in the beginning of the interval. Hence, during such intervals it is argued that the optimal choice would be to update only the observable part of the state. To do so, equation (3) is rewritten as:

$$\begin{aligned} M_1 : \dot{\mathbf{x}}^{o1} &= e_1(\mathbf{x}^{o1}, \mathbf{u}) \quad y_1 = h_1(\mathbf{x}^{o1}, \mathbf{u}) \\ &\vdots \\ M_l : \dot{\mathbf{x}}^{ol} &= e_l(\mathbf{x}^{ol}, \mathbf{u}) \quad y_l = h_l(\mathbf{x}^{ol}, \mathbf{u}) \end{aligned} \tag{7}$$

where models $M_1, \cdots, M_l$ are the smooth observable models that occur using only the observable states $\mathbf{x}^{oi}$ for each branch in (3). Each of these models are hence observable and the original state vector comprises of the union $\mathbf{x} = \cup_{i=1}^{l} \mathbf{x}^{oi}$. It should be noted that the size of these state vectors for different models is not the same and that the dimension of the states of model M_i, $dim(M_i)$ will in general be equal to $dim(\mathbf{x}) - dim(\mathbf{x}^{ui})$.

Each model is then accompanied by an event condition, i.e., an equation for the states that defines the transition from model M_i to one of the neighboring models M_j defined as $g_{i\to j}(\mathbf{x}) = 0 \Rightarrow g_{i\to j}(\mathbf{x}^{oi}, \mathbf{x}^{ui}) = 0$. It should also be noted that while the unobservable states $\mathbf{x}^{ui}$ do not appear in equations (7) they might appear in the transition equations between the models. As mentioned earlier, transitions between models are herein denoted as events.

The following modified version of the EKF algorithm, termed the *Discontinuous Extended Kalman Filter*, $DEKF$, is now formulated for such systems:

Let us assume that at a given time instance t_s the estimated value for the states is $\hat{\mathbf{x}}$, and that according to that mean realization for the states, the corresponding model that describes the behavior of the body is M_i. The observable part of the states then has a realization $\hat{\mathbf{x}}^{oi}$ and the unobservable part $\hat{\mathbf{x}}^{ui}$. Hence, the covariance matrix of the states $\mathbf{x}$ may be brought in the form $\mathbf{P} = \left[\begin{array}{c|c} \mathbf{P}^{oo} & (\mathbf{P}^{uo})^T \\ \hline \mathbf{P}^{uo} & \mathbf{P}^{uu} \end{array}\right]$.

The state-space equations of (7) are rewritten in discrete form:

$$M_i: \quad \mathbf{x}_k^{oi} = f_i(\mathbf{x}_{k-1}^{oi}, u_{k-1}) + \mathbf{w}_{k-1} \quad \mathbf{y}_k = h_i(\mathbf{x}_k^{oi}, \mathbf{u}_k) + \mathbf{v}_k \tag{8}$$

It is also assumed that model M_i is the observable model in the interval $[t_s, t_f]$, which generally comprises a subset of the the sampling interval $[t_k, t_{k-1}]$. The equations for the time and measurement updates of the observable components are obtained by applying the EKF:

Time Update of the observable components:

$$\begin{aligned}\hat{\mathbf{x}}^{oi}_{k_f|k-1} &= f_i(\hat{\mathbf{x}}^{oi}_{k_s|k-1}, \mathbf{u}_{k_s}) \\ \mathbf{P}^{oo}_{k_f|k-1} &= F_{k-1}\,\mathbf{P}^{oo}_{k_s|k-1}\,F^T_{k-1} + \mathbf{Q}_{k_f|k_s}\end{aligned} \tag{9}$$

where $F_{k-1} = \frac{\partial f_i}{\partial \mathbf{x}^{oi}}|_{(\hat{\mathbf{x}}^{oi}_{k_s|k-1}, \mathbf{u}_{k_s})}$ and $\mathbf{Q}_{k_f|k_s}$ is the process noise having taking into account the time increment $t_s - t_f$. The time update (9) is applied until t_f becomes equal to t_k, at which point the measurement update is applied. Assuming that at that time instance, t_k, the observable model is M_i this step becomes:

Measurement Update of the observable components:

$$\begin{aligned}\mathbf{S}_k &= H_k\,\mathbf{P}^{oo}_{k|k-1}\,H^T_k + \mathbf{R} \\ \mathbf{K}_k &= \mathbf{P}^{oo}_{k|k-1}\,H^T_k\,(\mathbf{S}_k)^{-1} \\ \hat{\mathbf{x}}^{oi}_{k|k} &= \hat{\mathbf{x}}^{oi}_{k|k-1} + \mathbf{K}_k\,(\mathbf{z}_k - h(\mathbf{x}^{oi}_{k|k-1})) \\ \mathbf{P}^{oo}_{k|k} &= (\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k)\,\mathbf{P}^{oo}_{k|k-1}\end{aligned} \tag{10}$$

where $H_k = \frac{\partial h_i}{\partial \mathbf{x}^{oi}}|_{(\hat{\mathbf{x}}^{oi}_{k|k-1}, \mathbf{u}_{k-1})}$.

During the interval $[t_s, t_f]$ the estimates of the unobservable states $\hat{\mathbf{x}}^{ui}$ are maintained unaltered, since in this work these correspond to the unidentifiable subset of the parameters $\boldsymbol{\theta}$. The corresponding terms in the covariance $[\mathbf{P}^{uu}]$ will also remain constant. However, the cross-covariance terms $[\mathbf{P}^{uo}]$ will change, due to the change of the observable variables $\mathbf{x}^{oi}$. In order to evaluate this cross-covariance matrix the Schmidt-Kalman Filter is applied ([25, 26]). The Schmidt-Kalman Filter provisions for the presence of parameters that are purposely maintained unaltered during both the time and measurement update steps. While this is done in the original method ([25]) so as not to increase the computational intensity of the problem, in the suggested method this aims at preventing the divergence of the unobservable parameters. Further noting that the measurement equations for any model M_i in (8) do not include the unobservable terms, it is apparent that these terms do not affect $\hat{\mathbf{x}}^{oi}$ and $\mathbf{P}^{oo}$ but only $\mathbf{P}^{uo}$, which evolves according to the equations:

Time-Update of $\mathbf{P}^{uo}$:

$$\mathbf{P}^{uo}_{k_f|k-1} = \mathbf{P}^{uo}_{k_s|k-1} F_{k-1} \tag{11}$$

Measurement-Update of $\mathbf{P}^{uo}$:

$$\mathbf{P}^{uo}_{k|k} = ((\mathbf{P}^{uo}_{k|k-1})^T - \mathbf{K}_k \, H_k \, (\mathbf{P}^{uo}_{k|k-1})^T)^T \tag{12}$$

using the definitions of $\mathbf{F}_{k-1}$, $\mathbf{H}_k$ and $\mathbf{K}_k$ from equation (10). Hence, during any time interval $[t_s, t_f]$ all components of $\hat{\mathbf{x}}$ and $\mathbf{P}$ may be defined. Following the assumption that model M_i is valid until time instance t_f, the value of $\hat{\mathbf{x}}$ at that time instance will define a transition from M_i to M_j (event). Two different cases may be distinguished:

1. The event occurs during the Time Update step (9).

This corresponds to a dynamic event, describing a transition between models due to the predicted dynamics of the system. The time of this event is determined herein using the event function of the Matlab ode solvers ([27, 28]). The event function is able to accurately determine the time instance t_f at which the zero crossing of the event function $g_{i \to j}(\hat{\mathbf{x}}) = 0$ occurs. In this paper the Runge-Kutta 4-5 pair solver (ode 45 [29]) is employed. When such an event is detected, the modified EKF solver temporarily halts at that time instance t_f, so that the model is switched. The time-update equations (9) and (11) are employed, without applying the measurement-update equations (10) and (12), since the measurement becomes available in the future time t_k. The temporary output of the algorithm is hence: $\hat{\mathbf{x}}_{k_f|k-1}$ and $\mathbf{P}_{k_f|k-1}$.

2. The event occurs during the Measurement Update step (10).

This implies that a transition from model M_i to M_j occurred when applying the measurement-update equation (10), indicating that this event occurred at a sampling step ($t_f = t_k$). The $DEKF$ temporarily halts at that instance in order to perform a model switch. The output of the algorithm is $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{k|k}$ and the covariance is evaluated after the measurement update, i.e., using $\mathbf{P}^{oo}_{k|k}$ and $\mathbf{P}^{uo}_{k|k}$.

In both previous cases, the algorithm will subsequently need to switch from model M_i to model M_j and to re-enter the $DEKF$ after setting $t_s = t_f$. Regardless of the nature of the event, i.e., whether it occurred during the time- or measurement-update step, a switch from one model to another does occur. It is assumed again that a switch is performed from M_i to M_j and hence the observable and unobservable states switch from $\mathbf{x}^{oi}$ and $\mathbf{x}^{ui}$ to $\mathbf{x}^{oj}$ and $\mathbf{x}^{uj}$ respectively. Hence, one needs to select the elements of $\hat{\mathbf{x}}$ and $\mathbf{P}$ that correspond to the observable components, which will be updated according to equations (9) and (10). The unobservable states and corresponding covariance terms are held constant, while the $\mathbf{P}^{uo}$ terms are updated according to equations (11) and (12). Table 3 summarizes the method used for the unobservable and observable parts of $\hat{\mathbf{x}}$ and $\mathbf{P}$. A schematic representation of the $DEKF$ is presented in Figure 1.

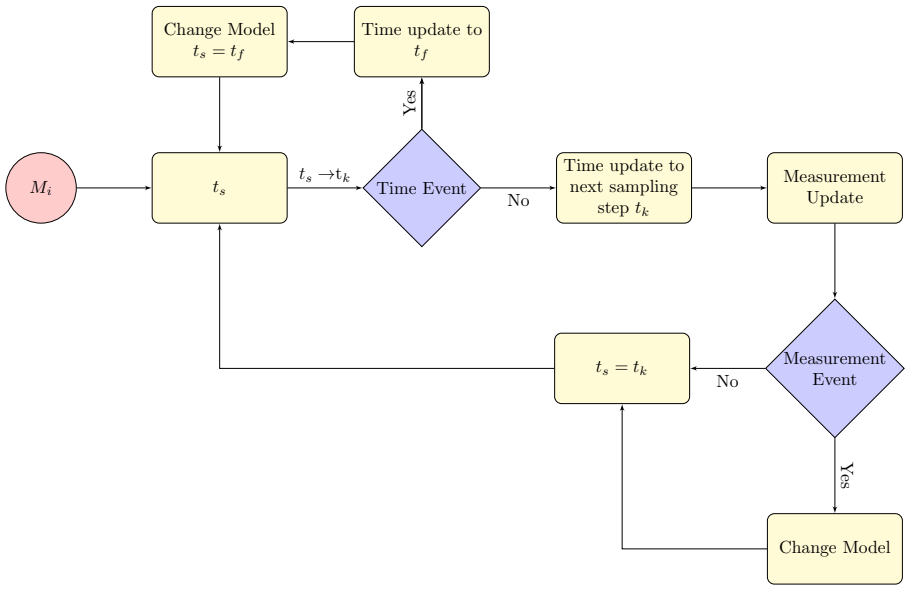


Figure 1: Schematic representation of the $DEKF$.

	Algorithm Used	Time Update/Measurement Update Equation
$\hat{\mathbf{x}}^{oi}$, $\mathbf{P}^{oo}$	EKF	(9) & (10)
$\hat{\mathbf{x}}^{ui}$, $\mathbf{P}^{uu}$	Retain Invariant	-
$\mathbf{P}^{uo}$	Schmidt-Kalman	(11)& (12)

Table 3: Equations used to update observable and unobservable parts of $\hat{\mathbf{x}}$ and $\mathbf{P}$.

6.1. Switching Condition

At any given time the real dynamics of the system lay within a specific smooth branch and are hence fully described by the corresponding observable model M_i. Since the switching condition of the $DEKF$ $g_{i\to j}(\hat{\mathbf{x}}) = 0$ is a function of the estimated value of the state vector $\hat{\mathbf{x}}$, the model used at a given time by the $DEKF$ is also an estimate $\hat{M}_i$ of the actual model M_i. Hence, the following two cases should be distinguished:

1. $M_i \equiv \hat{M}_i$. In this case the $DEKF$ is optimal, as the estimated model uses the smooth branch that generated the data at that time instance.

2. $M_i \neq \hat{M}_i$. Since the real and estimated models are different the estimated as observable states $\hat{\mathbf{x}}^{oi}$ will not converge during such intervals towards their real values. This is owed to the use of a smooth subsystem, which is different to the one generating the data. This is however a problem for both the $DEKF$ and EKF, as in both methods the subsystem which is used is based on the estimated values of $\hat{\mathbf{x}}$. The effective difference in the two approaches lies in the treatment of the unidentifiable parameters $\hat{\mathbf{x}}^{ui}$. However, in the case of the EKF although taken into account, these do not affect the measurement equation (6). Therefore, even though some of them would be identifiable, if the estimated state vector where to lie within the correct smooth branch, their estimation would still not converge even when employing the whole state-vector as in the original EKF. Hence, during such intervals although sub-optimal, the $DEKF$ does not perform inferior to the original EKF method.

The suggested $DEKF$ method can also be related to switching Kalman Filters ([30]) for which the the effect of choosing an estimated model $\hat{M}_i$ has

been investigated in greater depth. For the purposes of this paper, it should be kept in mind that the $DEKF$ is still an improvement over the EKF regardless of the estimated model $\hat{M}_i$ as explained previously.

7. Applications

7.1. The Impact Problem

The first example investigates a drop weight problem. The ground is simulated by means of vertical springs and dampers, with mass normalized stiffness and damping k and c respectively, which are active only when the body is in contact with the ground, i.e., when the relative position of the body with respect to the undeformed surface of the ground, as defined in Figure 2 is positive. It is also assumed that the ground and its undeformed surface have a common vertical acceleration $\ddot{x}_g$ and that gravity $g = 9.81 m/sec^2$ acts on the body.

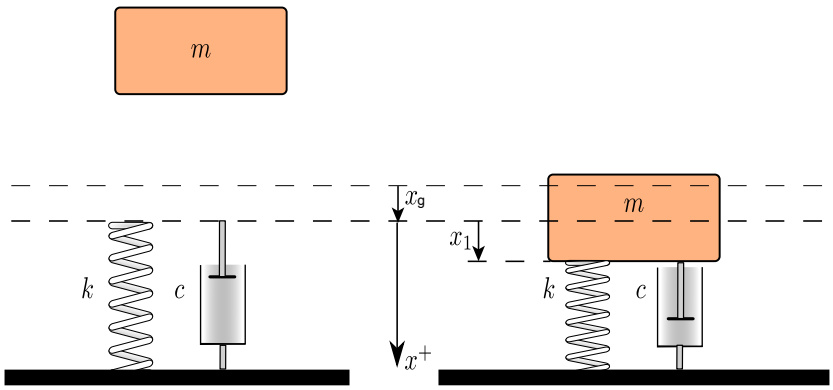
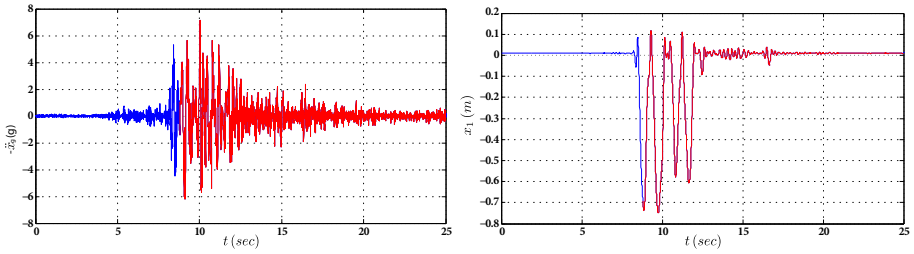


Figure 2: A body of mass m falling on a ground simulated via springs k and damper elements c that are only active during contact.

Defining x_1 to be the relative position of the body with respect to the undeformed ground surface, the state-space and measurement equations describing this problem become:

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= \begin{cases} -(k\,x_1 + c\,x_2) + g - \ddot{x}_g\,, \text{ when } x_1 >= 0 \\ g - \ddot{x}_g\,, \text{ when } x_1 < 0 \end{cases} \\
\dot{k} &= 0 \\
\dot{c} &= 0 \\
y &= x_1
\end{aligned}
\tag{13}
$$

The resulting system for $k = 1000\,[1/sec^2]$ and $c = 10\,[1/sec]$, is modeled for a ground acceleration input and part of the input and simulated output is then employed for identification of the system properties, as indicated in the following Figure 3. Note that while in the beginning of the simulation the system lies at rest, for the used segment of the measured data the actual initial-conditions are in fact non-zero. The input and measurement vectors are contaminated with zero mean Gaussian white noise vectors. The noise to signal rms ratio for the input is 1%. The measurement noise rms corresponds to 1.2% of the rms of the positive part of the measurement signal. The negative part of the signal, corresponding to the free-flight response of the body, is several times larger than the response of the body when the springs and dampers are active and is therefore not accounted for in the rms calculation.



(a) $-\ddot{x}_g$ (b) x_1

Figure 3: (a) Ground acceleration $\ddot{x}_g$ and (b) relative displacement x_1. Red color denotes the part of the input and output vectors that were considered as measured for the identification algorithms.

The state-space equations (13) are not smooth and can be separated into two smooth branches ([8]) depending on the value of x_1 . When, $x_1 >= 0$, i.e., when there is contact between the body and the deformable ground, it can be shown, using the ORC, that all states and parameters (x_1, x_2, k, c) are observable. On the other hand, when $x_1 < 0$, i.e., when the body experiences free-flight, (x_1, x_2) are observable and (k, c) are unobservable and hence unidentifiable. This, not unexpectedly, implies that one cannot obtain useful information regarding the spring and the damper when the body experiences free-flight. For use with the $DEKF$ two models are determined:

$$M_1 : \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\ddot{x}_g + g \\ y &= x_1 \\ M_1 &\rightarrow M_2 : x_1 = 0\,(- \rightarrow +) \end{aligned} \tag{14}$$

$$M_2 : \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -(k\,x_1 + c\,x_2) - \ddot{x}_g + g \\ \dot{k} &= 0 \\ \dot{c} &= 0 \\ y &= x_1 \\ M_2 &\rightarrow M_1 : x_1 = 0(+ \rightarrow -) \end{aligned} \tag{15}$$

Hence, $\mathbf{x}^{o1} = [x_1, x_2]$, $\mathbf{x}^{u1} = [k, c]$, $\mathbf{x}^{o2} = \mathbf{x}$ and $\mathbf{x}^{u2} = \emptyset$. The system will be identified using the normal EKF, the UKF and the $DEKF$. All three algorithms operate under the assumption of the correct process and observation noise, although this is not a requirement of any of the three methods. The following initial conditions are assumed: $x_1(0) = 9.81/2000$, $x_2(0) = 0$, $k(0) = 2000$, $c(0) = 20$ and $\mathbf{P}(0) = 2 \times D(\mathbf{x}(0)) + 10^{-8}\, I$, where $D(\mathbf{x}(0))$ is the diagonal matrix created by the initial realization of vector $\mathbf{x}$.

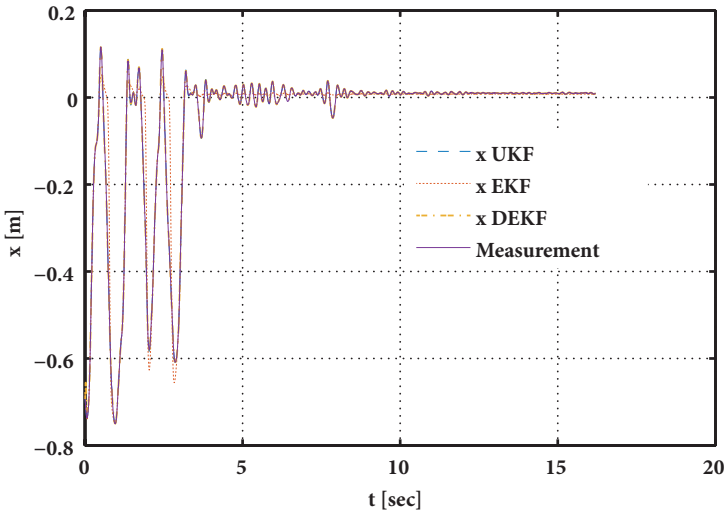
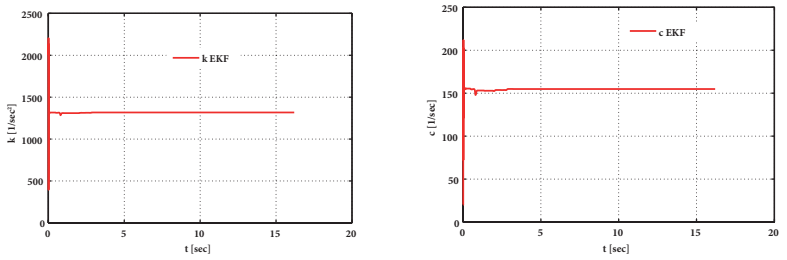
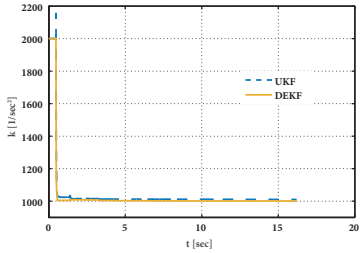


Figure 4: Identified displacement x_1 from the three methods versus measured signal.

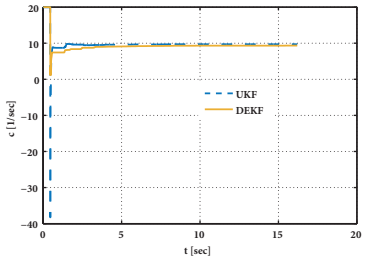


(a) $EKF\ k$

(b) $EKF\ c$



(c) $DEKF, UKF\ k$



(d) $DEKF, UKF\ c$

Figure 5: Identified k and c from the three methods.

As observed in Figures 4 and 5 the traditional EKF method fails to provide a reasonable result. This is expected since in the first time steps, during which in reality the body experiences free-flight, the algorithm delivers a divergent prediction for parameters k and c. Even when contact with the ground is re-engaged the algorithm, which during this interval could potentially start converging towards the correct solution, fails to do so. This behavior is generally expected for the EKF, firstly because convergence during the observable time steps is not guaranteed to be faster than divergence during the unobservable steps, and secondly because even during an observable step the convergence of the EKF depends on the initial guess adopted for both $\hat{\mathbf{x}}$ and $\mathbf{P}$.

On the other hand, as observed in Figures 4 and 5 the $DEKF$ converges towards the true solution. Unlike the EKF, the algorithm does not shift the values of k and c when the body experiences uplift, but only when contact is estimated to occur between the body and the ground according to the values of $\hat{\mathbf{x}}$. At this point, it should be reminded that a model is chosen according to the values of $\hat{\mathbf{x}}$, hence there are time instances during which the body might experience uplift while the model used is that for contact and vice versa. While these periods of miss-match between the real dynamics and the $DEKF$ estimate are not contributing towards the convergence of the algorithm, it should be noted that these are generally short and hence do not lead to divergence. Finally, when comparing the $DEKF$ to the traditional EKF, the former is able to 'correct' the values of k and c for a longer period than the EKF and hence benefits from more time intervals during which these parameters are observable.

Finally, it need be mentioned that the UKF algorithm also succeeds in identifying the correct solution. It is reminded here that the unidentifiable parameters $\mathbf{x}^{ui}$ may also diverge during the corresponding intervals when employing the UKF. Moreover, both the EKF and UKF algorithms are designed on the basis of observability for all states. Hence, the rate of divergence of the unidentifiable parameters is not a known or well-studied property of the methods. On the other hand, the rate of convergence of the identifiable parameters is commonly faster for the UKF as opposed to the EKF. The fact that the UKF converges

overall for the specific problem studied and the input and measurements used, indicates that the rate of convergence for the parameters during identifiable intervals happens to be faster than the corresponding rate of divergence during unidentifiable intervals.

This first example demonstrates one of the main points of this paper, which is the potentially suboptimal performance of the EKF in non-smooth problems due to its divergence during unobservable intervals. The proposed method $DEKF$ remedies this by switching between observable models and achieves an accurate estimate for all the parameters. This point will be further illustrated in the following examples.

7.2. Non-linear hysteretic Bouc-Wen model

In this example the hysteretic system illustrated in Figure 6 comprising a Bouc-Wen spring of mass normalized stiffness k and linear damping c is examined.

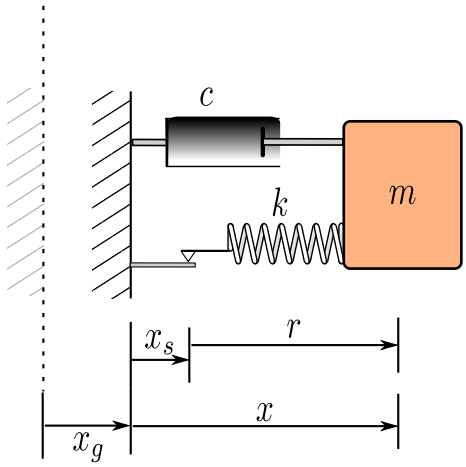


Figure 6: Mass on a Bouc-Wen Spring.

The relative displacement x of the body with respect to the ground is considered as the measured quantity. The observability of this system was examined

in [8]. The equations of motion are formulated as:

$$\begin{aligned} &\ddot{x} + k\,r + c\dot{x} = -\ddot{x}_g \\ &\dot{r} = \dot{x} - \beta\,|\dot{x}|\,|r|^{\nu-1}\,r - \gamma\,\dot{x}\,|r|^{\nu} \end{aligned} \tag{16}$$

where k is the stiffness of the spring, c the damping coefficient, and β, γ and ν are the parameters of the Bouc-Wen model. The term $\dot{r}$ can be re-written as $\dot{r} = \dot{x} - \dot{x}_s$, where x_s is the displacement of the slider and $\dot{x}_s = \beta\,|\dot{x}|\,|r|^{\nu-1}\,r - \gamma\,\dot{x}\,|r|^{\nu}$. Hence, r can be thought of as the displacement of the elastic spring. As stated in that paper the dynamic equations of motion of the system can be separated into four smooth branches:

$$\begin{aligned} &(A):\ \dot{r} = \dot{x} - \beta\,\dot{x}\,r^{\nu} - \gamma\,\dot{x}\,r^{\nu}\,, \text{for } \dot{x} > 0\,\&\,r > 0 \\ &(B):\ \dot{r} = \dot{x} + \beta\,\dot{x}\,r^{\nu} - \gamma\,\dot{x}\,r^{\nu}\,, \text{for } \dot{x} < 0\,\&\,r > 0 \\ &(C):\ \dot{r} = \dot{x} + \beta\,\dot{x}\,(-r)^{\nu} - \gamma\,\dot{x}\,(-r)^{\nu}\,, \text{for } \dot{x} > 0\,\&\,r < 0 \\ &(D):\ \dot{r} = \dot{x} - \beta\,\dot{x}\,(-r)^{\nu} - \gamma\,\dot{x}\,(-r)^{\nu}\,, \text{for } \dot{x} < 0\,\&\,r < 0 \end{aligned} \tag{17}$$

within these branches the system is not fully observable but can be rewritten in the form:

$$\begin{aligned} &(A):\ \dot{r} = \dot{x} - \Delta_1\,\dot{x}\,r^{\nu}\,, \text{for } \dot{x} > 0\,\&\,r > 0 \\ &(B):\ \dot{r} = \dot{x} + \Delta_2\,\dot{x}\,r^{\nu}\,, \text{for } \dot{x} < 0\,\&\,r > 0 \\ &(C):\ \dot{r} = \dot{x} + \Delta_2\,\dot{x}\,(-r)^{\nu}\,, \text{for } \dot{x} > 0\,\&\,r < 0 \\ &(D):\ \dot{r} = \dot{x} - \Delta_1\,\dot{x}\,(-r)^{\nu}\,, \text{for } \dot{x} < 0\,\&\,r < 0 \end{aligned} \tag{18}$$

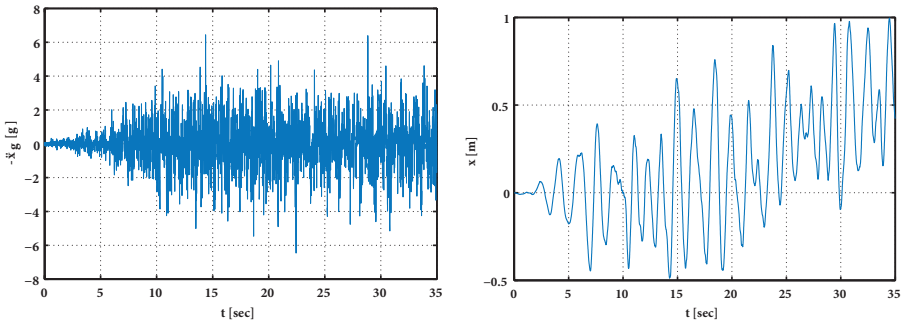
where $\Delta_1 = \beta + \gamma$ and $\Delta_2 = \beta - \gamma$. In this new representation, within each branch all of the appearing states $(x, \dot{x}, r, k, c)$ and either Δ_1 or Δ_2 are observable and hence the parameters are identifiable. Two models are defined for use in the $DEKF$:

$$M1: \quad \begin{aligned} &\dot{r} = \dot{x} - \Delta_1\,\dot{x}\,|r|^{\nu} \\ &M_1 \rightarrow M_2 : sign(\dot{x}\,r) = 0\,(+ \rightarrow -) \end{aligned} \tag{19}$$

$$M2: \quad \begin{array}{c} \dot{r} = \dot{x} + \Delta_2 \, \dot{x} \, |r|^{\nu} \\ M_2 \rightarrow M_1 : sign(\dot{x}\, r) = 0 \, (- \rightarrow +) \end{array} \tag{20}$$

where of course the models are presented only in terms of the $\dot{r}$ equation, as equation $\ddot{x} + k\, r + c\dot{x} = -\ddot{x}_g$ is common for both of them. It should also be noted that for each of the two models M_1 and M_2 all corresponding states are observable, as shown when these are separated into the corresponding smooth branches in terms of the sign of r as in equation (18). In implementing the $DEKF$, smoothness of the models does not pose a requirement, as long as the models themselves are observable within all of the implied smooth sub-systems. To completely define the models in terms of the $DEKF$ it is further noted that: $\mathbf{x}^{o1} = [x, \dot{x}, k, c, \nu, \Delta_1]$, $x^{u1} = \Delta_2$, $\mathbf{x}^{o2} = [x, \dot{x}, k, c, \nu, \Delta_2]$ and $\mathbf{x}^{u2} = \Delta_1$.

A system with mass normalized stiffness and damping terms $k = 9 \, \frac{1}{sec^2}$ and $c = 0.25 \, \frac{1}{sec}$, respectively and Bouc-Wen parameters $\Delta_1 = 3$, $\Delta_2 = 1$ (or equivalently $\beta = 2$, $\gamma = 1$) and $\nu = 2$ initially at rest is subjected to the input ground motion shown in Figure 7. The measured signal is assumed to be the displacement of the system x. Both the measurement and input signals are contaminated with a zero mean Gaussian White noise each having a noise to signal rms ratio of 1%.



(a) $-\ddot{x}_g$ (b) x

Figure 7: (a) Ground acceleration (b) Relative displacement.

Initially the EKF method is used. The state vector to be identified is:

$[x, \dot{x}, k, c, \nu, \Delta_1, \Delta_2]$. The initial estimates of the parameters are $k_0 = 18\ \frac{1}{sec^2}$, $c_0 = 0.5\ \frac{1}{sec}$, $\Delta_{1_0} = 5$, $\Delta_{2_0} = 2$, $\nu_0 = 2.8$. It should be noted that the model is not fully observable. In fact at any single time instance one of the parameters Δ_1 or Δ_2 is unidentifiable. The results are presented in the following Figure 8:

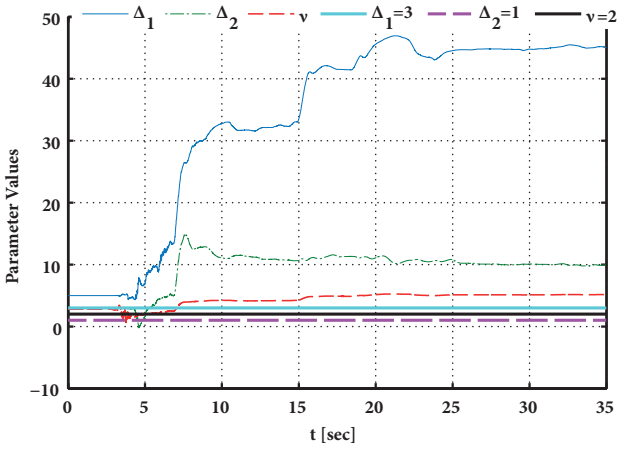
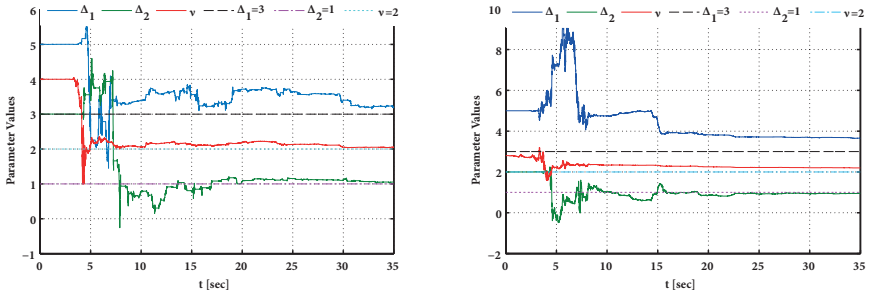


Figure 8: Predictions of the EKF model for the corresponding parameters of the Bouc-Wen.

It is apparent from Figure 8 that the EKF fails to converge to the correct solution for the parameters. The final predictions of the method for the parameters correspond to: $\Delta_1 = 45$, $\Delta_2 = 10$ and $\nu = 5$, in severe contrast to the real values $\Delta_1 = 3$, $\Delta_2 = 1$ and $\nu = 2$.

Next the UKF is compared against the $DEKF$ using models M_1 and M_2 defined in equations (19) and (20) respectively.

A comparison of Figures 8 and 9 reveals that the $DEKF$ and UKF methods do not diverge. In fact, the $DEKF$ identifies the values of the parameters very efficiently. Hence, unlike the original EKF method, the $DEKF$ is not affected by the unidentifiability of either Δ_1 or Δ_2 at a given interval. Additionally, it indicates that the EKF divergence may be attributed not only to the ill-posing of the Jacobian at the transition points, as noticed in [22], but primarily to the divergence of the unobservable parameters $\mathbf{x}^{ui}$. By retaining the unobservable parameters invariant during the corresponding intervals, the $DEKF$ remedies



(a) $DEKF$ (b) UKF

Figure 9: Predictions of $DEKF$ and UKF for the corresponding parameters of the Bouc-Wen model.

this problem and demonstrates robustness in these non-smooth problems often related to material plasticity models. It should be highlighted that the initial estimates of the parameters used for the $DEKF$ in Figure 14(a) ($\Delta_{1_0} = 5$, $\Delta_{2_0} = 3$, $\nu_0 = 4$) are less favorable than those used for the EKF and UKF ($\Delta_{1_0} = 5$, $\Delta_{2_0} = 2$, $\nu_0 = 2.8$), yet the method provides a better final estimate of the parameters in this studied case.

Regarding the UKF, it can be noted that while it clearly performs better than the EKF, more favorable initial estimates than those used for the $DEKF$ have to be used to allow for its convergence. However, it should be noted that this convergence is dependent on a number of factors including, the initial conditions used, the input and the measured response signal due to the non-linear nature of the problem. This is again linked to the parameter convergence rate versus the corresponding rate of divergence during time intervals at which it is unidentifiable. These rates inevitably depend on the initial conditions used.

This example demonstrates that the proposed $DEKF$ method comprises a viable option for identification tasks involving material plasticity performing on par with, if not better for some cases to the very robust UKF method. It also provides an explanation of what is often noted in the literature, i.e., the fact that the EKF ([31, 32]) may encounter difficulties in determining the values of parameters β and γ (or Δ_1 and Δ_2). This is herein attributed to their

identifiability properties, which may in turn affect the successful identification of parameters that are always identifiable like ν.

7.3. Elasto-plastic system

In this third example the identification of a mass on a perfect elasto-plastic spring is examined. The displacement of the mass is assumed as the measured quantity. The mass normalized elastic stiffness and yield limit of the spring are $k = 1000 sec^{-2}$ and $F_y = 50 (m\, sec^2)$. The mass is also connected to a mass normalized linear damper with $c = 2\sqrt{1000}\, 0.05 sec^{-1}$. The behavior of the spring is shown in Figure 10.

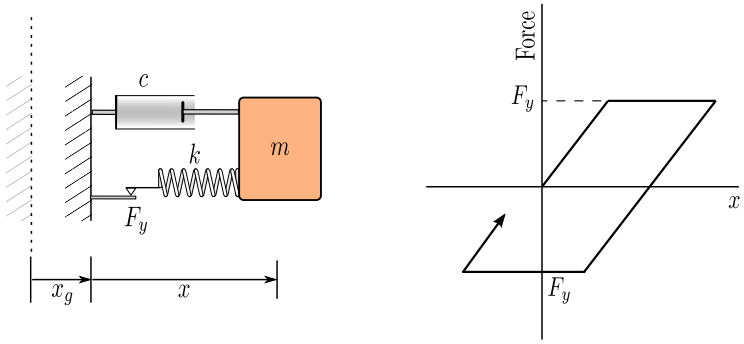


Figure 10: Behavior of elasto-plastic spring.

Denoting $x_1 = x$ and $x_2 = \dot{x}$, The equations of motion of this system for the elastic branch are written as:

$$\begin{aligned} \dot{x}_1 &= x_2, & \dot{x}_2 &= -k\, x_{el} - c\, x_2 - \ddot{x}_g \\ \dot{x}_{el} &= x_2, & \dot{k} &= 0 \\ \dot{c} &= 0, & \dot{F}_y &= 0 \\ \text{switch to plastic: } k\, |x_{el}| &= F_y \end{aligned} \tag{21}$$

while the equations describing the plastic branch are as follows:

$$
\begin{aligned}
\dot{x}_1 &= x_2, & \dot{x}_2 &= -F_y\, sign(x_2) - c\, x_2 - \ddot{x}_g \\
\dot{x}_{el} &= 0, & \dot{k} &= 0 \\
\dot{c} &= 0, & \dot{F}_y &= 0 \\
\text{switch to elastic: } x_2 &= 0
\end{aligned}
\tag{22}
$$

where x_{el} is the elastic deformation of the spring. Note that, in the elastic branch F_y is unobservable. In the plastic branch k and x_{el} are both constants, and in fact an implied constraint automatically satisfiable in forward simulations is $k\ |x_{el}| = F_y$. Hence, during the plastic branch all states are observable with the exception of k, x_{el} for which only their non-linear product $k\, x_{el}$ would have been observable. It is also worth noting that equations (21) and (22) require detection of the transition event even in forward simulation, since otherwise the states could shift in a region lying outside the elasto-plastic curve, in which case a return-mapping scheme would be required. Moreover, a second implied constraint satisfied exactly at the transition to the plastic branch is: $sign(x_{el}) = sign(x_2)$. While this is automatically satisfied in a forward simulation, it is not necessarily satisfied herein due to the measurement-update step of the identification algorithms. Hence, the transition from the elastic branch to the plastic branch is re-written as:

$$
\text{if } k\ |x_{el}| = F_y, \text{then if: } \begin{cases} x_{el}\, x_2 \geq 0 & \rightarrow \text{switch to plastic} \\ x_{el}\, x_2 < 0 & \rightarrow \text{remain in the elastic branch} \end{cases}
\tag{23}
$$

After the measurement update, the constraint $k\ |x_{el}| \leq F_y = 0$ has to be imposed, if violated. For the case of the EKF this is carried out by linearizing this constraint and imposing it after the measurement-update ([33]):

$$
\begin{aligned}
D\,\mathbf{x} &= d \\
D &= [\ 0 \quad 0 \quad k\,sign(x_{el}) \quad |x_{el}| \quad 0 \quad -1\ ] \\
d &= -k\,|x_{el}| + F_y + D\,\mathbf{x} \\
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k} - \mathbf{P}_{k|k}\,D^T\,(D\,\mathbf{P}_{k|k}\,D^T)^{-1}\,(D\,\mathbf{x} - d) \\
Y &= \mathbf{P}_{k|k}\,D^T\,(D\,\mathbf{P}_{k|k}\,D^T)^{-1} \\
\mathbf{P}_{k|k} &= (I - Y\,D)\,\mathbf{P}_{k|k}\,(I - Y\,D)^T
\end{aligned}
\tag{24}
$$

For the UKF the following modification is applied to each sigma point that violates the constraint:

$$
x_{el} = F_y/k * sign(x_{el}) \tag{25}
$$

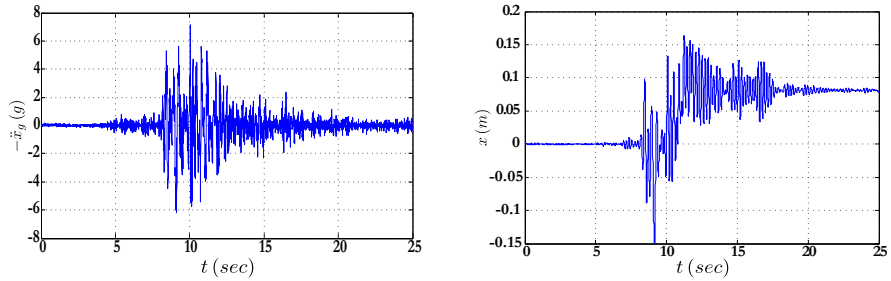
This leads to the setup of the necessary equations for the UKF and EKF algorithms. In order to set up the models for the $DEKF$ method the state vector used is defined as: $\mathbf{X} = [x_1,\, x_2,\, c,\, kx_{el},\, k]$, where a new state kx_{el} is introduced as the product $k \times x_{el}$. The two observable models used for the $DEKF$ then result as:

$$
M1: \quad
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -kx_{el} - c\,x_2 - \ddot{x}_g \\
\dot{kx}_{el} &= k\,x_2 \\
\dot{k} &= 0 \\
\dot{c} &= 0 \\
M_1 \rightarrow M_2 &: \text{switch to plastic as in equation (23)}
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
M2: \quad & \dot{x}_1 = x_2 \\
& \dot{x}_2 = -k x_{el} - c\, x_2 - \ddot{x}_g \\
& \dot{k x_{el}} = 0 \\
& \dot{c} = 0 \\
& M_2 \rightarrow M_1 : x_2 = 0
\end{aligned} \tag{27}
$$

where M_1 is the elastic loading and unloading model and M_2 is the plastic model. It should be noted that all of the states in $\mathbf{X}$ are observable for M_1, i.e., $\mathbf{x}^{o1} = \mathbf{X}$ and $\mathbf{x}^{u1} = \emptyset$, while for M_2, $\mathbf{x}^{o2} = x_1$, x_2, c, kx_{el}, and $\mathbf{x}^{u2} = k$. It should be observed that the assumption of retaining the unobservable variable k invariable for M_2 is equivalent to the assumption that as the values of kx_{el} change, that change would only affect the values of $x_e l$. Moreover, it should be noted that F_y does not appear as an observable or unobservable state for any of the two models. The parameter F_y appears in the switching equation from $M_1 \rightarrow M_2$ and is updated only during the intervals for which M_2 is applicable, through the equation $F_y = kx_{el}\, sign(x_2)$.

Having defined the models used for the EKF, UKF and $DEKF$ the identification of the system via each method is presented next. The used input and measured displacement of the system are shown in Figure 11.

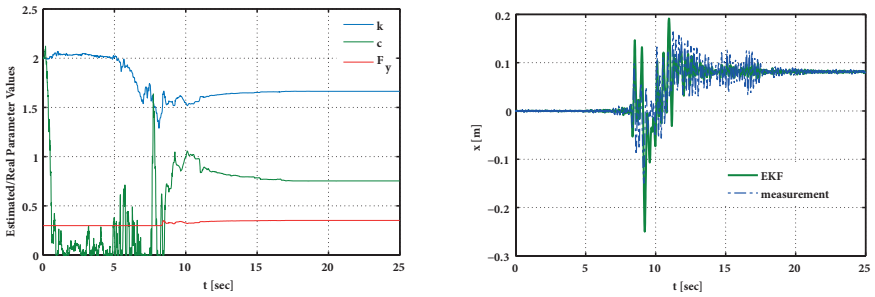


(a) $-\ddot{x}_g$ (b) x

Figure 11: Used input and measured displacement.

In all ID methods the initial guess is $X_0 = [0, 0, 2000, 2\sqrt{(1000)}\, 10/100, 0]$ and $F_y = 15$. The noise to signal rms ratio for the input and measurement noise

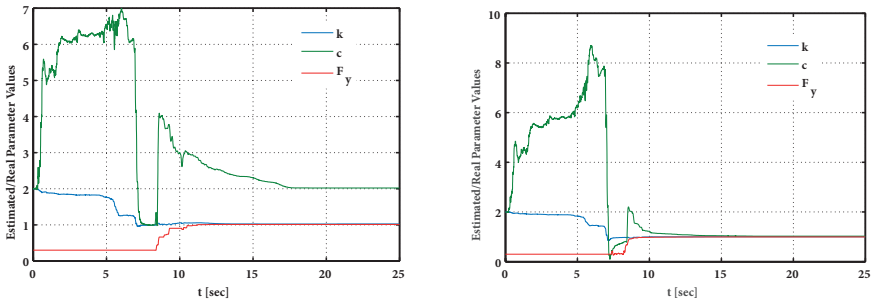
vectors is 1%. The results are presented first for the EKF method in Figure 12.



(a) EKF (b) x

Figure 12: EKF, ratio of estimated to real parameters and estimated x versus measured.

As observed in Figure 12, the EKF fails to converge to the true parameter values and it practically fails to update the values of F_y at all. The EKF once again under-performs, yielding diverging estimates for the parameters. The identified parameters according to the $DEKF$ and UKF are shown in Figure 13.



(a) $DEKF$ (b) UKF

Figure 13: Predictions of $DEKF$ and UKF for the corresponding parameters of the Elasto-Plastic model.

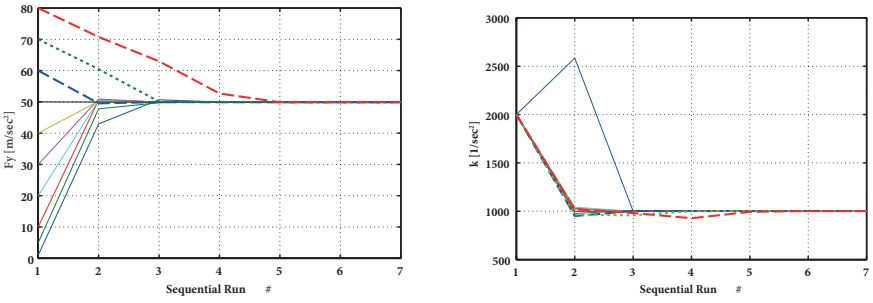
As evidenced in Figure 13, the two methods do not diverge and succeed in updating all the involved parameters. The UKF achieves an excellent estimate of parameters due to the fact that convergence in observable intervals is faster

than divergence in unobservable intervals. Once again however, it should be reminded that in the standard UKF there is no control over this phenomenon, which depends on the underlying dynamics.

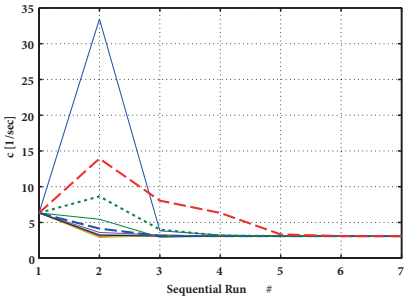
The $DEKF$ also provides a very good estimate of k and F_y and but its estimate of c is not optimal. This may be attributed to the fact that during the last 12 seconds of the identification process the response of the body is mainly elastic and, as a result, the method is practically not updating F_y. The small difference between the real and estimated value of F_y results into a lower loss of energy which the method tries to compensate for using a higher value of c. This sub-optimal convergence of the $DEKF$ results also as a consequence of the fact that the criterion for switching between models depends on the estimated values for the state kx_{el} and the indirect parameter F_y. In particular, it is observed that if the initial assumption for the value of F_y used is big enough to prevent the constraint $kx_{el} < |F_y|$ from ever being violated, then the $DEKF$ will never switch to the plastic model unlike the real system dynamics.

Hence, this sub-optimal convergence depends on the initial estimate of X_0 and F_y used. However, by using the method sequentially on the same set of data, i.e., using the $DEKF$ on the data and then using the final estimates for the parameters as initial conditions for the next run using again the same set of data, then as can be shown in the following Figure 14, the algorithm converges for a substantial range of initial values for F_y.

Despite the fact that for the specific example the $DEKF$ is not as robust for online purposes as the UKF, it still provides an acceptable solution. The sub-optimality of the method can be remedied at the price of its online nature, by using the method sequentially. As suggested in Figure 14, this offline procedure can provide an excellent estimate of the parameters even for assumed initial values of the parameters that are far from the real values. This is an important feature delivered by the proposed $DEKF$ approach.



(a) F_y (b) k



(c) c

Figure 14: Using the $DEKF$ sequentially on the same set of data. The figure shows the initial values used for the parameters at each run. The last prediction of the method for the parameters is used as initial condition for the next run.

8. Discussion and Conclusions

In this paper a modified version of the EKF, termed the *Discontinuous Extended Kalman Filter* $DEKF$, is suggested for non-smooth dynamical systems whose state-space equations can be separated into smooth branches. For each branch, the observability of the subsystem is deduced and the states are separated into observable and unobservable sets. Subsequently for each branch a model containing only the observable states is formed together with so-called event equations that describe the transition from one model to the other. The method then applies the EKF updating steps only to the observable states of each model, retaining the unobservable part invariant during that time interval.

Additionally, it was shown that for this type of non-smooth problems, associated with plasticity and impact problems, the time intervals during which a parameter is unobservable may affect the results of methods that do not incorporate observability considerations into the analysis. Specifically, it was demonstrated that the divergence of unobservable parameters is the primary reason for the failure of the EKF method in delivering a successful parameter estimate in problems of this type. Although the UKF suffers from the same issue, its faster convergence properties during observable intervals allow it to overcome the divergence rate of the same parameters during unobservable intervals.

This property is however not derived from the design of the method; it is rather a bi-product of its algorithmic robustness. Hence it is not guaranteed that this will indeed be true for any problem, and will greatly depend on factors such as the initial state guess, or more the amount of noise in the input and measurement signals. Indeed despite the overall very satisfactory performance of the UKF, the second example illustrates how the presence of locally unidentifiable parameters can adversely affect the performance of the method for an inappropriate initial guess for the values of the parameters. It should also be noted that this is a novel justification for the superior behavior of the UKF over the EKF for problems involving non-differentiable state-space equations. So far predominantly in the literature ([34]) this has been attributed to the

inability of the EKF to accurately calculate the derivatives around the points at which the state-space equations are non-differentiable .

Unlike the EKF and UKF, the $DEKF$ takes into account the observability properties of the system at each time instance. It thus ensures that the unidentifiable parameters will not deviate, maintaining these as invariant over such intervals. The presented examples illustrate the superior performance of the method compared to that of the EKF for non-smooth problems. In fact, the method performs on par with or in some cases better than the UKF, as shown in the second example. The third example illustrates the use of the method for a problem with a constraint equation originating from the law of perfectly plastic behavior. For such a problem, the standard EKF is incapable of producing results, however it is shown that the $DEKF$ is able to furnish accurate parameter estimates. It is further demonstrated that if the method is used sequentially in an offline manner, a highly accurate parameter estimation is attained for an initial guess that is substantially far from the true parameter value. This is often very useful in practice, in problems where a poor initial estimate is inevitable due to lack of a-priori knowledge of the system.

This work introduces an enhanced version of the EKF method, capable of handling problems of non-smooth dynamics. It additionally offers further insight, based on the concepts of observability and identifiability, as to the reasons behind the divergence of the standard EKF method in such problems. Via the proposed analysis, a better understanding regarding the good performance of the UKF in these types of problems is attained. At the same time it is highlighted that the convergence of the latter may depend on the underlying dynamics, the initial estimates and the amount of noise in the input and measurement signals. Hence, a next direction for this research would lie in coupling the superior convergence properties of the UKF, together with the robust handling of unobservable parameters proposed in this work.

9. Acknowledgment

The first author would like to acknowledge the Marie Curie FP7 Integration Grant N^o 618359 within the 7th European Union Framework Programme, for the support of this research. The second author would like to gratefully acknowledge the support of the Albert Lck Stiftung.

[1] R. Kalman, Mathematical description of linear dynamical systems, Journal of the Society for Industrial and Applied Mathematics Series A Control 1 (2) (1963) 152–192. arXiv:http://epubs.siam.org/doi/pdf/10.1137/0301010.

[2] S. J. Julier, J. K. Uhlmann, A new extension of the Kalman filter to nonlinear systems., Proceedings of AeroSense: The 11th Int. Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando.

[3] R. Hermann, A. Krener, Nonlinear controllability and observability, Automatic Control, IEEE Transactions on 22 (5) (1977) 728–740.

[4] E. Walter, Identifiability of State Space Models, Springer, Berlin, 1982.

[5] S. Mukhopadhyay, H. Luş, R. Betti, Modal parameter based structural identification using input and output data: Minimal instrumentation and global identifiability issues, Mechanical Systems and Signal Processing 45 (2) (2014) 283 – 301. doi:http://dx.doi.org/10.1016/j.ymssp.2013.11.005.
URL http://www.sciencedirect.com/science/article/pii/S0888327013005736

[6] S. Diop, M. Fliess, Nonlinear observability, identifiability, and persistent trajectories, in: Decision and Control, 1991., Proceedings of the 30th IEEE Conference on, IEEE, 1991, pp. 714–719.

[7] L. Ljung, T. Glad, On global identifiability for arbitrary model parametrizations, Automatica 30 (2) (1994) 265 – 276.

[8] M. N. Chatzis, E. N. Chatzi, A. W. Smyth, On the observability and identifiability of nonlinear structural and mechanical systems, Structural Control and Health Monitoring.

[9] P.-T. Liu, F. Li, H. Xiao, A state decoupling approach to estimate unobservable tracking systems, Oceanic Engineering, IEEE Journal of 21 (3) (1996) 256–259.

[10] E. Lourens, E. Reynders, G. D. Roeck, G. Degrande, G. Lombaert, An augmented kalman filter for force identification in structural dynamics, Mechanical Systems and Signal Processing 27 (0) (2012) 446 – 460. doi:http://dx.doi.org/10.1016/j.ymssp.2011.09.025.

[11] E. M. Hernandez, A natural observer for optimal state estimation in second order linear structural systems, Mechanical Systems and Signal Processing 25 (8) (2011) 2938 – 2947. doi:http://dx.doi.org/10.1016/j.ymssp.2011.06.003.
URL http://www.sciencedirect.com/science/article/pii/S0888327011002251

[12] H. Luş, R. Betti, R. W. Longman, Obtaining refined first-order predictive models of linear structural systems, Earthquake Engineering & Structural Dynamics 31 (7) (2002) 1413–1440. doi:10.1002/eqe.169.

[13] S. S. Bisht, M. P. Singh, An adaptive unscented kalman filter for tracking sudden stiffness changes, Mechanical Systems and Signal Processing 49 (12) (2014) 181 – 195. doi:http://dx.doi.org/10.1016/j.ymssp.2014.04.009.
URL http://www.sciencedirect.com/science/article/pii/S0888327014001150

[14] R. Astroza, H. Ebrahimian, J. P. 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. arXiv:http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0000851, doi:10.1061/(ASCE)EM.1943-7889.0000851.
URL http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0000851

[15] Z. Xie, J. Feng, Real-time nonlinear structural system identification via iterated unscented kalman filter, Mechanical Systems and Signal Processing 28 (2012) 309 – 322, interdisciplinary and Integration Aspects in Structural Health Monitoring. doi:http://dx.doi.org/10.1016/j.ymssp.2011.02.005.
URL http://www.sciencedirect.com/science/article/pii/S0888327011000690

[16] A. Isidori, Nonlinear Control Systems, 3rd Edition, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1995.

[17] S. T. Glad, L. Ljung, Model structure identifiability and persistence of excitation, in: Decision and Control, 1990., Proceedings of the 29th IEEE Conference on, IEEE, 1990, pp. 3236–3240.

[18] J. L. Beck, L. S. Katafygiotis, Updating models and their uncertainties. i: Bayesian statistical framework, Journal of Engineering Mechanics 124 (4) (1998) 455–461.

[19] L. S. Katafygiotis, J. L. Beck, Updating models and their uncertainties. ii: Model identifiability, Journal of Engineering Mechanics 124 (4) (1998) 463–467.

[20] C. D. Persis, A. Isidori, On the observability codistributions of a nonlinear system, Systems & control letters 40 (5) (2000) 297–304.

[21] E. Wan, R. Van Der Merwe, The unscented Kalman filter for nonlinear estimation, in: Adaptive Systems for Signal Processing, Communications, and Control Symposium, AS-SPCC, IEEE, Lake Louise, Alberta, Canada, Oct,2000, pp. 153–158.

[22] E. Chatzi, A. Smyth, The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated het-

erogeneous sensing, Structural Control and Health Monitoring 16 (1) (2009) 99–123.

[23] E. N. Chatzi, A. W. Smyth, S. F. Masri, Experimental application of on-line parametric identification for nonlinear hysteretic systems with model uncertainty, Journal of Structural Safety 32 (5) (2010) 326 – 337.

[24] M. Chatzis, E. Chatzi, A. Smyth, An experimental validation of time domain system identification methods with fusion of heterogeneous data, Earthquake Engineering & Structural Dynamics 44 (4) (2015) 523–547. doi:10.1002/eqe.2528.

[25] S. F. Schmidt, Applications of state space methods to navigation problems, Advances in Control Systems 3 (1966) 293–340.

[26] R. Y. Novoselov, S. M. Herman, S. M. Gadaleta, A. B. Poore, Mitigating the effects of residual biases with Schmidt-Kalman filtering, in: Information Fusion, 2005 8th International Conference on, Vol. 1, IEEE, 2005, pp. 8–pp.

[27] MATLAB, version 8.1.064 (R2013a), The MathWorks Inc., Natick, Massachusetts, 2013.

[28] L. F. Shampine, M. W. Reichelt, The Matlab ode suite, SIAM journal on scientific computing 18 (1) (1997) 1–22.

[29] L. F. Shampine, M. K. Gordon, Computer solution of ordinary differential equations: the initial value problem, WH Freeman San Francisco, 1975.

[30] K. P. Murphy, Switching Kalman filters, Tech. rep., U.C. Berkeley (1998).

[31] M. Wu, A. W. Smyth, Application of the unscented kalman filter for real-time nonlinear structural system identification, Journal of Structural Control and Monitoring 14, No. 7 (November. 2007) 971–990.

[32] J. Ching, J. L. Beck, K. A. Porter, Bayesian state and parameter estimation of uncertain dynamical systems, Probabilistic Engineering Mechanics 21 (1)

(2006) 81 – 96. doi:http://dx.doi.org/10.1016/j.probengmech.2005.08.003.
URL http://www.sciencedirect.com/science/article/pii/S0266892005000597

[33] D. Simon, Kalman filtering with state constraints: a survey of linear and nonlinear algorithms, IET Control Theory & Applications 4 (8) (2010) 1303–1318.

[34] S. Mariani, A. Ghisi, Unscented kalman filtering for nonlinear structural dynamics, Nonlinear Dynamics 49 (1) (2007) 131–150. doi:10.1007/s11071-006-9118-9.
URL http://dx.doi.org/10.1007/s11071-006-9118-9