Parameter identification of a differentiable Bouc-Wen model using constrained extended Kalman filter

Dan Li [1], Yang Wang [1, 2, *]

[1] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

[2] School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA

*yang.wang@ce.gatech.edu

Abstract: Hysteresis is of critical importance to structural safety under severe dynamic loading conditions. One of the widely used hysteretic models for civil structures is the Bouc-Wen model, whose effectiveness depends on suitable model parameters. The locally non-differentiable governing equation of the conventional Bouc-Wen model poses difficulty on existing identification algorithms, especially the extended Kalman filter (EKF), which relies on linearized system equations to propagate state estimates and covariance. In addition, the standard EKF usually does not incorporate parameter constraints, and therefore may result in unreasonable estimates. In this paper, a modified and differentiable Bouc-Wen model, together with a constrained extended Kalman filter (CEKF), is proposed to identify the hysteretic model parameters in a reliable way. The partial derivatives of the differentiable Bouc-Wen model with respect to hysteretic parameters can be easily calculated for implementing the identification algorithm. CEKF restricts the Kalman gain to ensure that the estimates of parameters satisfy constraints from physical laws. Parameter identification using simulated and experimental data collected from a four-story structure demonstrates that CEKF can achieve more reliable identification results than the standard EKF.

Keywords: parameter identification, hysteresis, nonlinear system, Bouc-Wen model, constrained extended Kalman filter

1 Introduction

Nonlinearity and inelasticity are usually encountered in structures under severe dynamic loading, such as earthquake and strong wind. Reliable modeling of structures subject to such extreme events is of critical importance to structural damage assessment and post event maintenance. These nonlinear and inelastic structural behaviors due to dynamic loading usually exhibit in a form of hysteresis, which refers to the path-dependent relationship between restoring force and deformation. The area of the hysteresis loop represents the dissipated energy during the loading duration. Detailed modeling of hysteretic behaviors of structures is usually too complicated for engineering application [1]. In addition, the obtained models are always problem dependent and difficult to be extended for general usage. For these reasons, phenomenological models are developed to characterize the hysteretic features. A survey on the phenomenological models for hysteresis can be found in [2]. Among these models, the Bouc-Wen model has been extensively used in civil engineering to describe the hysteresis phenomenon of, for example, magnetorheological (MR) dampers [3-5], beam-column joints [6, 7], and soil-structure interaction [8].

The governing equation of the Bouc-Wen model is a first-order nonlinear differential equation, the parameters of which control the shape and size of the hysteresis loop. In order to accurately describe the hysteretic properties of structures, the parameters in the Bouc-Wen model need to be identified from the output and/or the input of the structural system. During past decade, various system identification algorithms have been proposed and investigated to tune the hysteretic parameters. These algorithms include evolutionary algorithms (EAs) [9-12], least-square estimation (LSE) [13-15], sequential Monte Carlo methods [16-19], etc. Inspired by biological evolution, EAs search the solution for an optimization problem stochastically. EAs can be used for

complex optimization problems but the solution is sensitive to initial population and affected by the problem size. In addition, EAs are metaheuristic optimization algorithms and cannot guarantee optimality [20]. LSE methods estimate the model parameters according to measurements and structural responses including acceleration, velocity, and displacement. The requirement on velocity information limits the application of LSE methods. Sequential Monte Carlo methods rely on a large number of sampling points to estimate parameters together with system states. Although the methods perform well on nonlinear systems with various noise models, the computation complexity limits the application on small or midsize problems.

The nonlinear versions of Kalman filter, especially extended Kalman filter (EKF) and unscented Kalman filter (UKF), are also commonly used identification algorithms for hysteretic systems. Kalman filter produces *a posteriori* probabilistic estimates of unknown state variables based on system equations and noisy measurements. Parameter identification is performed by treating the model parameters as augmented system states and thus, estimated through the measurement data. EKF linearizes the system equation and measurement equation around the current optimal estimate, and updates the estimate based on the linearized equations. It has been shown that EKF works well for system with mild nonlinearity but often provides unreliable estimates for highly nonlinear systems due to the large linearization error. A powerful alternative to EKF is UKF which relies on the unscented transformation for estimating system states and parameters. UKF is designed based on the intuition that it should be easier to approximate a given distribution than to approximate an arbitrary nonlinear function. At each iteration, UKF generates a sample distribution by a set of sampling points called sigma points, which capture the mean and covariance of the *a posteriori* distribution of system states. These sigma points can be easily propagated through the nonlinear system equation and used for updating the estimate. UKF has arisen in many engineering

applications [21-23]. Extensive research has also been conducted to evaluate the performance of EKF and UKF on parameter identification of the Bouc-Wen model. Wu and Smyth compared EKF and UKF on updating parameters of a single degree of freedom (SDOF) Bouc-Wen system, and showed that UKF outperformed EKF in estimation accuracy for this highly nonlinear system [24].

Although the large linearization error limits the application of the standard EKF on parameter identification for highly nonlinear systems, some researchers have investigated techniques to improve the accuracy of EKF. Applying constraints on parameters during the estimation process is one effective technique. Yang and Ma proposed a constrained EKF, in which the constrained parameters are replaced by parameters in the entire solution space through specific functions, for example sinusoidal functions and square function [25]. In this way, the original problem has been converted to a problem of estimating new parameters using the standard EKF. However, these functions are not one-to-one functions and thus multiple values of a new parameter could correspond to the same value of the original parameter. This fact increases the difficulty of accurate parameter estimation. Sen and Bhattacharya investigated a constrained EKF which restricts the Kalman gain to ensure that the updated estimate lies within the constrained space [26]. This restricted Kalman gain is obtained by solving a constrained optimization problem, which can be computational expensive. Gupta and Hauser [27] discussed two methods to incorporate constraints in Kalman filter, among which one method projects the updated state estimate onto the constrained domain, and the other method restricts the Kalman gain so that the updated estimate satisfies the constraints. Although developed for Kalman filter, these two methods can be conveniently implemented in EKF. This paper adopts the idea of restricting the Kalman gain and derives the analytical solution of the Kalman gain when there are inequality constraints. With the explicit expression of the Kalman gain, the estimation process can be significantly accelerated. The

proposed constrained EKF (CEKF) is validated through both numerical simulation and laboratory experiment.

Another difficulty of applying EKF on hysteretic parameter identification is that the governing equation of the Bouc-Wen model is non-differentiable. As EKF requires to linearize system equation to update state estimate and covariance, strictly speaking, EKF is not applicable for parameter identification of the Bouc-Wen model. To circumvent this challenge, most of the existing algorithms use numerical approximation to differentiate the Bouc-Wen system equations [24]. This paper analyzes factors contributing to the non-differentiability of the widely used Bouc-Wen model and proposes a modified and differentiable Bouc-Wen model. Using the proposed model, partial derivatives of the system equations with respect to hysteretic parameters can be analytically and explicitly calculated for the implementation of EKF and CEKF.

The rest of paper is organized as follows. Section 2 discusses the contributing factors for the non-differentiability of the widely used Bouc-Wen model, and proposes a differentiable Bouc-Wen model accordingly. Section 3 briefly reviews EKF algorithm and presents CEKF algorithm for hysteretic parameter identification. Section 4 shows numerical simulations and a laboratory experiment validating the proposed parameter identification method. In the end, Section 5 provides a summary and future work.

2 Hysteresis and the Bouc-Wen model

A vector variable $\boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}}$ is used to represent various structural model parameters, including mass, stiffness, and damping coefficient. The equation of motion for an *N*-DOF structure with nonlinear hysteresis can be expressed as:

$$\mathbf{M}(\boldsymbol{\theta})\ddot{\mathbf{q}} + \mathbf{C}(\boldsymbol{\theta})\dot{\mathbf{q}} + \mathbf{r}(\boldsymbol{\theta}, \mathbf{q}, \dot{\mathbf{q}}) = \mathbf{u} + \mathbf{w} \tag{1}$$

where $\mathbf{q} \in \mathbb{R}^N$ is the displacement vector, $\mathbf{M} \in \mathbb{R}^{N\times N}$ is the mass matrix, $\mathbf{C} \in \mathbb{R}^{N\times N}$ is the damping matrix, $\mathbf{r} \in \mathbb{R}^N$ is the nonlinear restoring force vector describing hysteresis, $\mathbf{u} \in \mathbb{R}^N$ is the external load vector, and $\mathbf{w} \in \mathbb{R}^N$ is the input uncertainty.

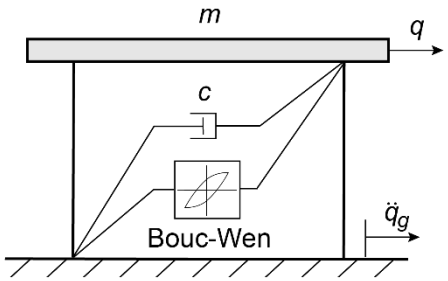


Fig. 1 Bouc-Wen hysteretic system with viscous damping

Various types of hysteretic models have been developed. One of the most widely used is a differential model originally proposed by Bouc and later developed by Wen and other researchers. In this Bouc-Wen model, the restoring force **r** is associated with a nonlinear first order differential equation. Take an SDOF Bouc-Wen hysteretic model (Fig. 1) subject to earthquake as an example. The governing equation of the system with mass m, damping coefficient c, stiffness k and ground excitation $\ddot{q}_g$ is shown as:

$$m\ddot{q}(t) + c\dot{q}(t) + kz(t) = -m\big(\ddot{q}_g(t) + w\big) \tag{2}$$

Here the excitation to the system is $u = -m\ddot{q}_g(t)$, and the ground acceleration input $\ddot{q}_g(t)$ is contaminated with uncertainty w. The nonlinear restoring force is $r = kz(t)$, and z is a hidden hysteretic displacement. A first-order differential equation describes the hysteretic displacement:

$$\dot{z} = \dot{q} - \beta|\dot{q}||z|^{n-1}z - \gamma\dot{q}|z|^n = \dot{q}\left(1 - |z|^n\big(\gamma + \beta\mathrm{sgn}(z\dot{q})\big)\right) \tag{3}$$

Here β, γ, and n are dimensionless parameters controlling the shape and magnitude of the hysteresis loop; $\mathrm{sgn}(\cdot)$ is the signum function. This differential model has many advantages in describing nonlinear hysteresis. By adjusting the parameters, this model is capable of generating a large variety of hysteretic loops. In order to identify proper values for the Bouc-Wen parameters, researchers have studied online recursive techniques that can search parameter values using real-time dynamic response data. EKF, as one of parameter identification techniques, has been widely applied for this purpose [28, 29].

Notice that the model equation Eq. (3) is not differentiable at $\dot{q} = 0$ or $z = 0$, and this singularity is not ideal for the linearization in EKF. Three reasons causing the non-differentiability are discussed as follows.

(i) Derivative of the signum function $\mathrm{sgn}(a)$ with respect to a and derivative of the absolute value function $|a|$ with respect to a

As the absolute function $|a| = \mathrm{sgn}(a)\, a$, it suffices to only discuss the derivative of the signum function $\mathrm{sgn}(a)$. This derivative is not defined at $a = 0$. In order to address this problem, the hyperbolic tangent function $\tanh(\cdot)$ can be adopted to approximate the signum function $\mathrm{sgn}(\cdot)$.

$$\mathrm{sgn}(a) \approx \tanh(\rho a) \tag{4}$$

Here $\rho > 0$ is a factor controlling the curvature. Fig. 2 shows the plot of $\tanh(\rho a)$ with different values of ρ. When the value of ρ is larger, the differentiable function $\tanh(\rho a)$ better approximates $\mathrm{sgn}(a)$.

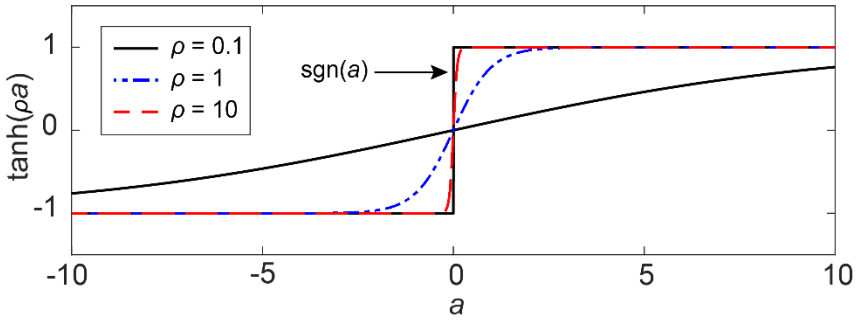


Fig. 2 Plot of $\tanh(\rho a)$

Using this approximation, the Bouc-Wen equation is then modified as:

$$\dot{z} = \dot{q}\left(1 - (\tanh(\rho z)\, z)^n(\gamma + \beta \tanh(\rho z \dot{q}))\right) \tag{5}$$

(ii) Derivative of the exponential function a^b with respect to a ($b < 1$)

For the term $f = (\tanh(\rho z)\, z)^n$ in Eq. (5), $\frac{\partial f}{\partial z}$ requires taking derivative with respect to the base. The derivative of the exponential function a^b with respect to a is calculated as $\frac{\partial}{\partial a} a^b = b a^{b-1}$. When $b < 1$, the derivative is undefined at $a = 0$. To avoid this singular point, we can simply require $b \geq 1$. Applying this constraint on the state equation of the Bouc-Wen model requires $n \geq 1$, which satisfies most of engineering applications.

(iii) Derivative of the exponential function a^b with respect to b ($a = 0$)

For the term $f = (\tanh(\rho z)\, z)^n$ in Eq. (5), $\frac{\partial f}{\partial n}$ requires taking derivative with respect to the power. The derivative of the exponential function a^b with respect to b is calculated as $\frac{\partial}{\partial b} a^b = a^b \ln a$, which is not defined for $a \leq 0$. In the system equation Eq. (5), the base of exponential functions is $\tanh(\rho z)\, z$ which is always nonnegative. Therefore, we only need to consider the case $a = 0$. Note that in (ii), we require $b \geq 1$. Using L'Hospital's rule, the limit of the function $a^b \ln a$, as a approaches 0^+ can be calculated as:

$$\lim_{a\to 0^+} a^b \ln a = \lim_{a\to 0^+} \frac{\ln a}{a^{-b}} = \lim_{a\to 0^+} \frac{a^{-1}}{-ba^{-b-1}} = \lim_{a\to 0^+} \frac{a^b}{-b} = 0 \tag{6}$$

In the application of EKF, we define the derivative $\frac{\partial}{\partial b} a^b = 0$ at $a = 0$.

3 Extended Kalman filter (EKF)

EKF is an extension of the standard Kalman filter for optimally estimating the state of a nonlinear system from measurement output. Section 3.1 briefly introduces the standard EKF algorithm for state and parameter estimation. Section 3.2 describes the proposed constrained EKF (CEKF) algorithm which can prevent the estimates from being unrealistic.

3.1 The standard EKF

Consider a general dynamical system governed by nonlinear state-space equation as:

$$\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}, \mathbf{u}, \mathbf{w}) \tag{7}$$

where $\mathbf{u}$ is known excitation applied on the system and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{w}})$ is a zero-mean white Gaussian process noise with covariance matrix $\boldsymbol{\Sigma}_{\mathbf{w}}$. At time $t = k\Delta t$, the measurement $\mathbf{y}_k$ is given as:

$$\mathbf{y}_k = \boldsymbol{h}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k) \tag{8}$$

where $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{v}})$ is the zero-mean white Gaussian measurement noise with covariance matrix $\boldsymbol{\Sigma}_{\mathbf{v}}$.

The EKF estimation is separated into two main steps, i.e. measurement update step and time update step. In the measurement update step, the *a priori* estimate $\hat{\mathbf{x}}_{k|k-1}$ of the state is available. The predicted measurement $\hat{\mathbf{y}}_{k|k-1}$ of $\hat{\mathbf{x}}_{k|k-1}$ is estimated as:

$$\hat{\mathbf{y}}_{k|k-1} = \boldsymbol{h}\left(\hat{\mathbf{x}}_{k|k-1}, \mathbf{u}_k, \mathbf{0}\right) \tag{9}$$

The Kalman gain matrix is calculated to minimize the trace of the covariance matrix for the *a posteriori* estimate:

$$\mathbf{L}_k = \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} \left(\mathbf{H}_k^{\mathbf{x}} \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + \mathbf{H}_k^{\mathbf{v}} \boldsymbol{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} \right)^{-1} \tag{10}$$

Here $\mathbf{H}_k^{\mathbf{x}}$ and $\mathbf{H}_k^{\mathbf{v}}$ are the linearized matrices of the measurement equation $\boldsymbol{h}$:

$$\mathbf{H}_k^{\mathbf{x}} = \left. \frac{\partial \boldsymbol{h}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} \tag{11}$$

$$\mathbf{H}_k^{\mathbf{v}} = \left. \frac{\partial \boldsymbol{h}}{\partial \mathbf{v}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} \tag{12}$$

After measurement $\mathbf{y}_k$ is available, the *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ is calculated using the Kalman gain matrix as:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k \left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} \right) \tag{13}$$

Along with the measurement update of the state, the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}$ for the *a posteriori* estimate can be evaluated as:

$$\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}} = (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k^{\mathbf{x}}) \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + \mathbf{L}_k \mathbf{H}_k^{\mathbf{v}} \boldsymbol{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} \mathbf{L}_k^{\mathrm{T}} \tag{14}$$

In the time update step, the *a priori* estimate $\hat{\mathbf{x}}_{k+1|k}$ of the state is predicted based on the system model:

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{k|k} + \int_{k\Delta t}^{(k+1)\Delta t} f(\hat{\mathbf{x}}, \mathbf{u}, \mathbf{0}) dt \tag{15}$$

Along with the time update of the state, the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}_{k+1|k}}$ for the *a priori* estimate can be evaluated as [30]:

$$\boldsymbol{\Sigma}_{\mathbf{x}_{k+1|k}} = \boldsymbol{\Phi}_k^{\mathbf{x}} \boldsymbol{\Sigma}_{\mathbf{x}_{k|k}} (\boldsymbol{\Phi}_k^{\mathbf{x}})^{\mathrm{T}} + \boldsymbol{\Phi}_k^{\mathbf{w}} \boldsymbol{\Sigma}_{\mathbf{w}} (\boldsymbol{\Phi}_k^{\mathbf{w}})^{\mathrm{T}} \tag{16}$$

Here $\boldsymbol{\Phi}_k^{\mathbf{x}}$ and $\boldsymbol{\Phi}_k^{\mathbf{w}}$ are the state transition matrices and can be calculated by linearization of state-space equation $\boldsymbol{f}$:

$$\boldsymbol{\Phi}_k^{\mathbf{x}} = \mathbf{I} + \left.\frac{\partial \boldsymbol{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} \Delta t \tag{17}$$

$$\boldsymbol{\Phi}_k^{\mathbf{w}} = \left.\frac{\partial \boldsymbol{f}}{\partial \mathbf{w}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} \Delta t \tag{18}$$

Repeating Eq. (9) ~ Eq. (18), EKF can recursively update the system states for a nonlinear system.

3.2 The constrained EKF

The standard EKF on parameter identification problems finds the estimate of parameters through the entire unconstrained solution space. However, in structural applications, some model parameters must satisfy equality or inequality constraints from physics. Without incorporating those constraints in the estimation process, the standard EKF may lead to infeasible solutions. This section discusses an efficient approach of constrained EKF (CEKF).

To lighten notations, denote the measurement innovation $\mathbf{r}_k$ and innovation covariance $\boldsymbol{\Sigma}_{\mathbf{y}_k}$ at time $t = k\Delta t$ as:

$$\mathbf{r}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} \tag{19}$$

$$\mathbf{\Sigma}_{\mathbf{y}_k} = \mathbf{H}_k^{\mathbf{x}} \mathbf{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + \mathbf{H}_k^{\mathbf{v}} \mathbf{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} \tag{20}$$

To distinguish between the solutions without and with constraints, hereinafter we rename the Kalman gain matrix and the *a posteriori* estimate of the unconstrainted EKF, from Eq. (10) and Eq. (13), with a tilde sign.

$$\tilde{\mathbf{L}}_k = \mathbf{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} \mathbf{\Sigma}_{\mathbf{y}_k}^{-1} \tag{21}$$

$$\tilde{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{L}}_k \mathbf{r}_k \tag{22}$$

EKF is the minimum-mean-square-error (MMSE) estimator for linearized dynamical systems. The Kalman gain $\tilde{\mathbf{L}}_k$ can be analytically derived as the solution of the following optimization problem, which minimizes the trace of *a posteriori* state covariance matrix without any constraints to the optimization variable **L**.

$$\underset{\mathbf{L}}{\text{minimize}} \ \text{Trace}\left((\mathbf{I} - \mathbf{L}\mathbf{H}_k^{\mathbf{x}}) \mathbf{\Sigma}_{\mathbf{x}_{k|k-1}} (\mathbf{I} - \mathbf{L}\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + \mathbf{L}\mathbf{H}_k^{\mathbf{v}} \mathbf{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} \mathbf{L}^{\mathrm{T}} \right) \tag{23}$$

Suppose the constraint vector function is denoted as $\boldsymbol{g}(\cdot): \mathbb{R}^{n_{\mathbf{x}}} \to \mathbb{R}^{n_c}$, where n_c represents the number of scalar constraints. When a general constraint $\boldsymbol{g}(\cdot) \geq 0$ is imposed on the system states, the closed-form solution of the Kalman gain matrix is usually difficult to obtain, if not impossible. In this situation, an obvious way of constraining the EKF is to numerically calculate the constrained Kalman gain by solving the following optimization problem. The objective function still minimizes the trace of the *a posteriori* state covariance matrix, while the constraint $\boldsymbol{g}(\cdot) \geq 0$ is applied to the measurement-updated state estimation from Eq. (22).

$$\begin{aligned} &\underset{\mathbf{L}}{\text{minimize}}\ \text{Trace}\left((\mathbf{I}-\mathbf{L}\mathbf{H}_k^{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{I}-\mathbf{L}\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}+\mathbf{L}\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\right) \\ &\text{subject to}\ \boldsymbol{g}\left(\hat{\mathbf{x}}_{k|k-1}+\mathbf{L}\mathbf{r}_k\right)\geq 0 \end{aligned} \tag{24}$$

Using the optimal Kalman gain matrix $\mathbf{L}^*$ solved from Eq. (24), the updated *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}=\hat{\mathbf{x}}_{k|k-1}+\mathbf{L}^*\mathbf{r}_k$ is guaranteed to satisfy the constraint $\boldsymbol{g}\left(\hat{\mathbf{x}}_{k|k}\right)\geq 0$. The covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}$ corresponding to the updated *a posteriori* estimates can also be calculated using $\mathbf{L}^*$. The disadvantage of this approach of constraining the EKF is that the computational cost may increase significantly from solving the optimization problem in Eq. (24), depending on the exact form of the vector function $\boldsymbol{g}(\cdot)\colon \mathbb{R}^{n_{\mathbf{x}}}\to\mathbb{R}^{n_c}$.

To achieve a more efficient and heuristic way of constraining the EKF, consider that most engineering applications only encounter affine constraints which can be expressed as $\boldsymbol{g}(\mathbf{x})=\mathbf{A}\mathbf{x}-\mathbf{b}\geq\mathbf{0}$. Note the "≥" symbol represents entry-wise inequality between two vectors. We denote $\mathbf{A}=\begin{pmatrix}\mathbf{a}_1^{\mathrm{T}}\\ \vdots\\ \mathbf{a}_{n_c}^{\mathrm{T}}\end{pmatrix}\in\mathbb{R}^{n_c\times n_{\mathbf{x}}}$ and $\mathbf{b}=\begin{pmatrix}b_1\\ \vdots\\ b_{n_c}\end{pmatrix}\in\mathbb{R}^{n_c}$, such that the *i*-th constraint is simply a scalar inequality $g_i(\mathbf{x})=\mathbf{a}_i^{\mathrm{T}}\mathbf{x}-b_i\geq 0, i=1,2,\dots,n_c$. For a feasible $\mathbf{x}$, the constraint is said to be *active* if the equal sign holds, i.e. $\mathbf{a}_i^{\mathrm{T}}\mathbf{x}=b_i$; the constraint is said to be *inactive* if instead we have the strict greater-than relationship $\mathbf{a}_i^{\mathrm{T}}\mathbf{x}-b_i>0$ [31].

Our suggested way of constraining the EKF works as follows for the affine constraint $\boldsymbol{g}(\mathbf{x})=\mathbf{A}\mathbf{x}-\mathbf{b}\geq\mathbf{0}$. Suppose that at time $t=k\Delta t$, the unconstrained *a posteriori* estimate $\tilde{\mathbf{x}}_{k|k}$ calculated from Eq. (22) violates n_{ac} number of the n_c inequality constraints ($n_{ac}\leq n_c$). In order to obtain a feasible solution, these n_{ac} number of constraints are set as active. Accordingly, we aggregate all the rows in $\mathbf{A}$ and $\mathbf{b}$ that correspond to these constraints (that were violated by $\tilde{\mathbf{x}}_{k|k}=\hat{\mathbf{x}}_{k|k-1}+$

$\tilde{\mathbf{L}}_k\mathbf{r}_k$) as $\mathbf{A}_\mathrm{a} \in \mathbb{R}^{n_{ac}\times n_\mathbf{x}}$ and $\mathbf{b}_\mathrm{a} \in \mathbb{R}^{n_{ac}}$ respectively, *i.e.* where we have $\mathbf{A}_\mathrm{a}\tilde{\mathbf{x}}_{k|k} - \mathbf{b}_\mathrm{a} < \mathbf{0}$. So that the constrained *a posteriori* estimate may satisfy the constraints, a new (constrained) Kalman gain matrix needs to be found by solving the optimization problem below. Note that to make the new estimate $\hat{\mathbf{x}}_{k|k-1} + \mathbf{L}\mathbf{r}_k$ "barely" satisfy these previously violated constraints concerning $\mathbf{A}_\mathrm{a}$ and $\mathbf{b}_\mathrm{a}$, these constrains are set active as $\mathbf{A}_\mathrm{a}\left(\hat{\mathbf{x}}_{k|k-1} + \mathbf{L}\mathbf{r}_k\right) - \mathbf{b}_\mathrm{a} = \mathbf{0}$.

$$\begin{aligned} &\underset{\mathbf{L}}{\text{minimize}}\ \mathrm{Trace}\left((\mathbf{I} - \mathbf{L}\mathbf{H}_k^{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{I} - \mathbf{L}\mathbf{H}_k^{\mathbf{x}})^\mathrm{T} + \mathbf{L}\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^\mathrm{T}\mathbf{L}^\mathrm{T}\right) \\ &\text{subject to}\ \mathbf{A}_\mathrm{a}\left(\hat{\mathbf{x}}_{k|k-1} + \mathbf{L}\mathbf{r}_k\right) - \mathbf{b}_\mathrm{a} = \mathbf{0} \end{aligned} \tag{25}$$

Solution to Eq. (25) is highly efficient because closed-form analytical solution usually exists. Under mild assumption, the solution is given as follows and Appendix 2 details the derivation. The same solution is obtained by [27], where the derivation is more complicated than the process in Appendix 2.

$$\mathbf{L}_k = \tilde{\mathbf{L}}_k - \mathbf{A}_\mathrm{a}^\mathrm{T}(\mathbf{A}_\mathrm{a}\mathbf{A}_\mathrm{a}^\mathrm{T})^{-1}\left(\mathbf{A}_\mathrm{a}\tilde{\mathbf{x}}_{k|k-1} - \mathbf{b}_\mathrm{a}\right)\left(\mathbf{r}_k^\mathrm{T}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k\right)^{-1}\mathbf{r}_k^\mathrm{T}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1} \tag{26}$$

In addition, using this constrained optimal Kalman gain, the *a posteriori* estimate of CEKF, $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k\mathbf{r}_k$, is found to be related to the unconstrained EKF estimation $\tilde{\mathbf{x}}_{k|k}$ as follows through the active/equality constraints.

$$\hat{\mathbf{x}}_{k|k} = \tilde{\mathbf{x}}_{k|k} - \mathbf{A}_\mathrm{a}^\mathrm{T}(\mathbf{A}_\mathrm{a}\mathbf{A}_\mathrm{a}^\mathrm{T})^{-1}\left(\mathbf{A}_\mathrm{a}\tilde{\mathbf{x}}_{k|k} - \mathbf{b}_\mathrm{a}\right) \tag{27}$$

It should be noted that this heuristic approach assumes the constrained *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ also satisfies the other $(n_c - n_{ac})$ number of inactive constraints that are not required by Eq. (25). In other words, altogether the entire $\boldsymbol{g}\left(\hat{\mathbf{x}}_{k|k}\right) = \mathbf{A}\hat{\mathbf{x}}_{k|k} - \mathbf{b} \geq \mathbf{0}$ is satisfied. Although there is no

guarantee the assumption always holds, this is found to be the case for the examples studied in this paper. Finally, the CEKF algorithm is summarized in Table 1.

Table 1 CEKF algorithm

Initial Estimate		
State estimate	$\hat{\mathbf{x}}_{0\|-1} = \mathrm{E}[\mathbf{x}_0]$	
State covariance	$\boldsymbol{\Sigma}_{\mathbf{x}_{0\|-1}} = \mathrm{E}\left[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0\|-1})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0\|-1})^{\mathrm{T}}\right]$	
`for` $k = 0, 1, \cdots, n$		
Measurement Update		
Measurement innovation	$\mathbf{r}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k\|k-1}$	Rept. (22)
Innovation covariance matrix	$\boldsymbol{\Sigma}_{\mathbf{y}_k} = \mathbf{H}_k^{\mathrm{x}} \boldsymbol{\Sigma}_{\mathbf{x}_{k\|k-1}} (\mathbf{H}_k^{\mathrm{x}})^{\mathrm{T}} + \mathbf{H}_k^{\mathrm{v}} \boldsymbol{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathrm{v}})^{\mathrm{T}}$	Rept. (23)
Kalman gain	$\tilde{\mathbf{L}}_k = \boldsymbol{\Sigma}_{\mathbf{x}_{k\|k-1}} (\mathbf{H}_k^{\mathrm{x}})^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}$	Rept. (24)
State estimate	$\tilde{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \tilde{\mathbf{L}}_k \mathbf{r}_k$	Rept. (25)
Check constraints		
`if (` $\mathbf{A}\tilde{\mathbf{x}}_{k\|k} - \mathbf{b} > \mathbf{0}$ is satisfied `)`		
	$\mathbf{L}_k = \tilde{\mathbf{L}}_k$	
	$\hat{\mathbf{x}}_{k\|k} = \tilde{\mathbf{x}}_{k\|k}$	
`else`		
According to violations in $\mathbf{A}\tilde{\mathbf{x}}_{k\|k} - \mathbf{b} > \mathbf{0}$, identify $\mathbf{A}_{\mathrm{a}} \in \mathbb{R}^{n_{ac} \times n_{\mathbf{x}}}$ and $\mathbf{b}_{\mathrm{a}} \in \mathbb{R}^{n_{ac}}$ for the n_{ac} number of active constraints: $\mathbf{A}_{\mathrm{a}}\mathbf{x} - \mathbf{b}_{\mathrm{a}} = \mathbf{0}$ in Eq. (21).		
	$\mathbf{L}_k = \tilde{\mathbf{L}}_k - \mathbf{A}_{\mathrm{a}}^{\mathrm{T}}(\mathbf{A}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}}^{\mathrm{T}})^{-1}(\mathbf{A}_{\mathrm{a}}\tilde{\mathbf{x}}_{k\|k-1} - \mathbf{b}_{\mathrm{a}})(\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k)^{-1}\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}$	Rept. (26)
	$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k \mathbf{r}_k$	
`end`		
State covariance	$\boldsymbol{\Sigma}_{\mathbf{x}_{k\|k}} = (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k^{\mathrm{x}}) \boldsymbol{\Sigma}_{\mathbf{x}_{k\|k-1}} (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k^{\mathrm{x}})^{\mathrm{T}} + \mathbf{L}_k \mathbf{H}_k^{\mathrm{v}} \boldsymbol{\Sigma}_{\mathbf{v}} (\mathbf{H}_k^{\mathrm{v}})^{\mathrm{T}} \mathbf{L}_k^{\mathrm{T}}$	Rept. (14)
Time Update		
State estimate	$\hat{\mathbf{x}}_{k+1\|k} = \hat{\mathbf{x}}_{k\|k} + \int_{k\Delta t}^{(k+1)\Delta t} f(\hat{\mathbf{x}}, \mathbf{u}, \mathbf{0}) dt$	Rept. (15)
State covariance	$\boldsymbol{\Sigma}_{\mathbf{x}_{k+1\|k}} = \boldsymbol{\Phi}_k^{\mathbf{x}} \boldsymbol{\Sigma}_{\mathbf{x}_{k\|k}} (\boldsymbol{\Phi}_k^{\mathbf{x}})^{\mathrm{T}} + \boldsymbol{\Phi}_k^{\mathbf{w}} \boldsymbol{\Sigma}_{\mathbf{w}} (\boldsymbol{\Phi}_k^{\mathrm{w}})^{\mathrm{T}}$	Rept. (16)
`end loop`		

4 Application

To validate the proposed methods, model updating using simulated and experimental data is conducted. Section 4.1 introduces parameter identification of an SDOF nonlinear hysteretic system. The proposed differentiable Bouc-Wen model is compared with the original Bouc-Wen model. Based on the differentiable Bouc-Wen model, parameters of an SDOF nonlinear hysteretic system are updated through EKF and CEKF using analytically and numerically evaluated partial derivatives, respectively. Section 4.2 presents parameter identification of a four-story shear structure using simulated data. To further investigate the performance of the proposed methods, parameter identification of a four-story shear structure using experimental data is conducted and presented in Section 4.3.

4.1 Numerical simulation – an SDOF nonlinear hysteretic system

4.1.1 Structural response simulation

To demonstrate the effectiveness of the proposed differentiable Bouc-Wen model, simulation on an SDOF nonlinear hysteretic system (Fig. 1) is conducted. In this simulation example, system parameters are set as $m = 1$ kg, $c = 0.3$ Ns/m, $k = 12$ N/m, $\beta = 2$, $\gamma = 1$, and $n = 2$. In the differentiable Bouc-Wen model, the curvature controlling parameter is set as $\rho = 100$. A scaled El Centro earthquake excitation of 40 s duration is applied to excite the system. The system responses are obtained by numerical integration using zero-order hold. Fig. 3 plots the displacement q, velocity $\dot{q}$, hysteretic displacement z, and the hysteretic loops using the original Bouc-Wen model (Eq. (3)) and the differentiable Bouc-Wen model (Eq. (5)). The simulated results indicate that the structural responses of both models are close to each other and the proposed differentiable Bouc-Wen model is capable of capturing the hysteretic behaviors of the system with

acceptable accuracy. In the discussion afterwards, the proposed differentiable Bouc-Wen model will be adopted for parameter identification.

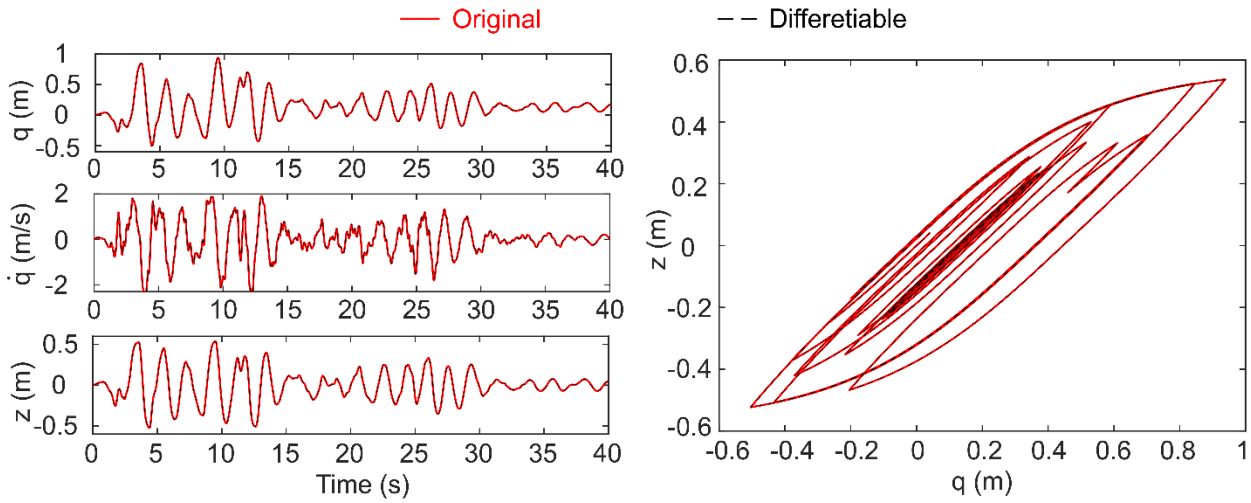


Fig. 3 Structural responses of the SDOF Bouc-Wen hysteretic system

4.1.2 Parameter identification using analytically evaluated partial derivatives

In this parameter identification example, the values of system parameters and excitation are the same as those in Section 4.1.1. The mass m is treated as accurate and other parameters are chosen for identification. The initial estimates are set as $c_{0|-1} = 0.15$ Ns/m, $k_{0|-1} = 6$ N/m, $\beta_{0|--1} = 0.5$, $\gamma_{0|-1} = 0.5$, and $n_{0|-1} = 4$. The state-space system equation for parameter identification can be formulated as:

$$\mathbf{x} = \begin{pmatrix} q \\ \dot{q} \\ z \\ c \\ k \\ \beta \\ \gamma \\ n \end{pmatrix} \qquad \dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}, \ddot{q}_g) = \begin{pmatrix} \dot{q} \\ -(\ddot{q}_g + w) - (c\dot{q} + kz)/m \\ \dot{q}\left(1 - (\tanh(\rho z)\, z)^n (\gamma + \beta \tanh(\rho z \dot{q}))\right) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{28}$$

The absolute acceleration of the mass is measured at $t = k\Delta t$ and the measurement equation is given as:

$$y_k = -(c\dot{q}_k + kz_k)/m + v_k \tag{29}$$

The covariance of the process noise is set as $\Sigma_w = (10^{-2}\ \mathrm{m/s^2})^2$, and the covariance of sensor noise is set as $\Sigma_v = (10^{-2}\ \mathrm{m/s^2})^2$.

Two parameter identification methods are adopted and compared using this SDOF hysteretic system. The first method is the standard EKF introduced in Section 3.1. The second method is the proposed CEKF introduced in Section 3.2. In the identification process, the linearization of system equation $\boldsymbol{f}$ and measurement equation $\boldsymbol{h}$ uses analytically derived partial derivatives.

For CEKF, inequality constraints applied on the parameters are listed as follows [32]:

$$c \geq 0, \qquad k \geq 0, \qquad \beta + \gamma \geq 0, \qquad \beta - \gamma \geq 0, \qquad n \geq 1 \tag{30}$$

Using the EKF and CEKF methods, the parameters of the nonlinear Bouc-Wen hysteretic system are identified together with the original system states, including displacement, velocity, and hysteretic displacement. Fig. 4 shows the time histories of the *a posteriori* estimates of the parameters and the system states. Except for the damping parameter c and the stiffness parameter k, all the other parameters cannot be updated correctly by EKF. On the other hand, the proposed CEKF can recursively update all the parameters from their initial values to the corresponding true values. The bound constraints can effectively prevent the estimates from being unreasonable values. The estimates of stiffness parameter k and damping coefficient c converge faster than the estimates of hysteretic parameters, which remain not updated and change rapidly after about 2

seconds. This is because the structure has not exhibited nonlinear behavior within the first 2 seconds of the estimation process.

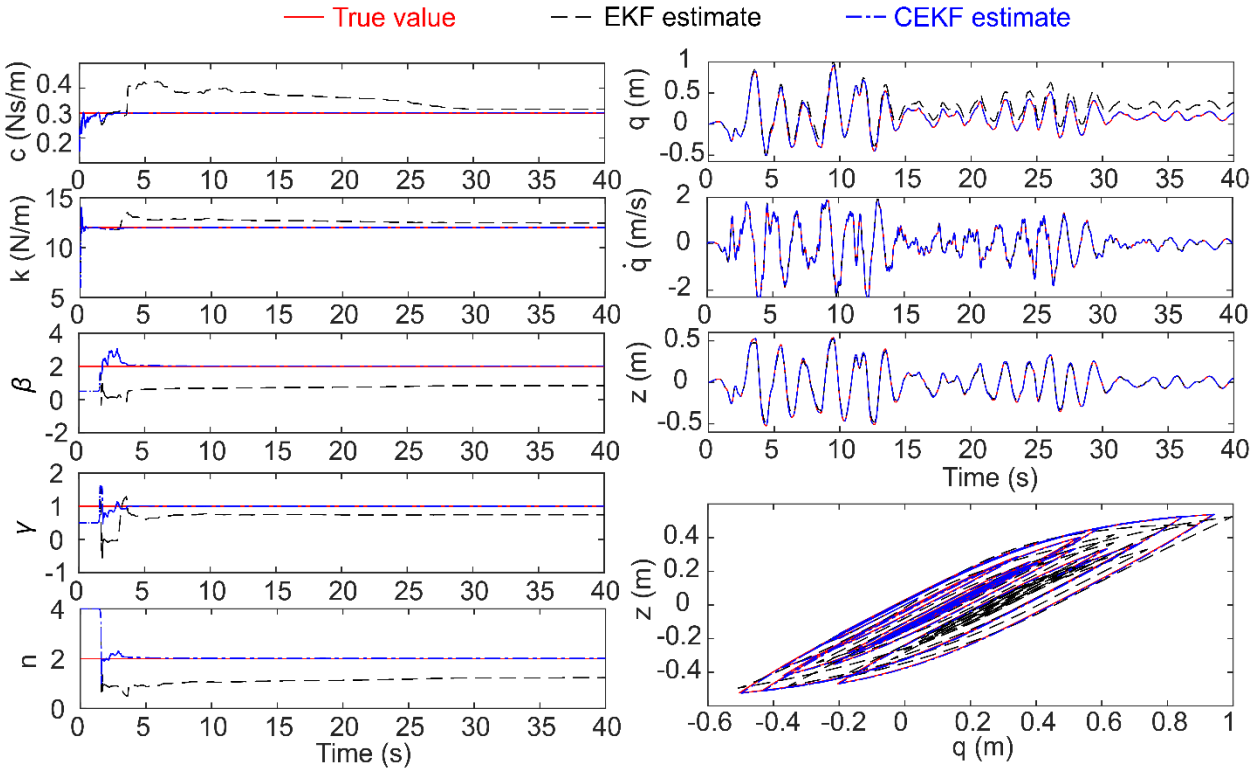


Fig. 4 Updating results on the SDOF Bouc-Wen hysteretic model using analytically linearized system and measurement equations

Similar to the estimation of parameters, the estimated states and hysteretic loop from the proposed CEKF match well with the actual states and hysteretic loop, respectively. On the other hand, EKF can provide accurate estimates for velocity $\dot{q}$ and hysteretic displacement z, while the estimate of displacement q suffers slow drift over time during the model updating process, and results in an inaccurate hysteretic loop.

A comparison of the final estimated values using different identification algorithms is summarized in Table 2. EKF estimation errors of all the hysteretic parameters are greater than 25%. The

difficulty of EKF identifying parameters of highly nonlinear systems has also been reported by other researchers [24]. On the other hand, the proposed CEKF is capable of accurately identifying model parameter values; all estimation errors are within $\pm 1\%$. This simulation example shows that the proposed CEKF outperforms EKF when applied for parameter identification of highly nonlinear systems.

Table 2 Comparison of estimation results on the SDOF Bouc-Wen hysteretic model using analytically evaluated partial derivatives

Parameters	Actual values	EKF		CEKF	
		Values	Errors (%)	Values	Errors (%)
c (Ns/m)	0.3	0.3149	4.9603	0.3007	0.2185
k (N/m)	12	12.4596	3.8297	11.9958	−0.0353
β	2	0.8424	−57.8792	2.0109	0.5464
γ	1	0.7389	−26.1072	1.0026	0.2634
n	2	1.2313	−38.4326	2.0080	0.4011

4.1.3 Parameter identification using numerically evaluated partial derivatives

Although partial derivatives of the modified Bouc-Wen model can be calculated analytically, sometimes it is necessary to calculate the partial derivatives using numerical methods, such as finite difference method. The two parameter identification methods are repeated using numerically evaluated partial derivatives for linearizing system and measurement equations.

Fig. 5 shows the time histories of the *a posteriori* estimates of the parameters and the system states. The estimation results can be found are almost the same as those using analytically linearized system and measurement equations (Fig. 4). The time histories show that it is difficult for EKF to accurately identify hysteretic parameters and displacement, while the proposed CEKF can provide accurate estimates for both system parameters and system states.

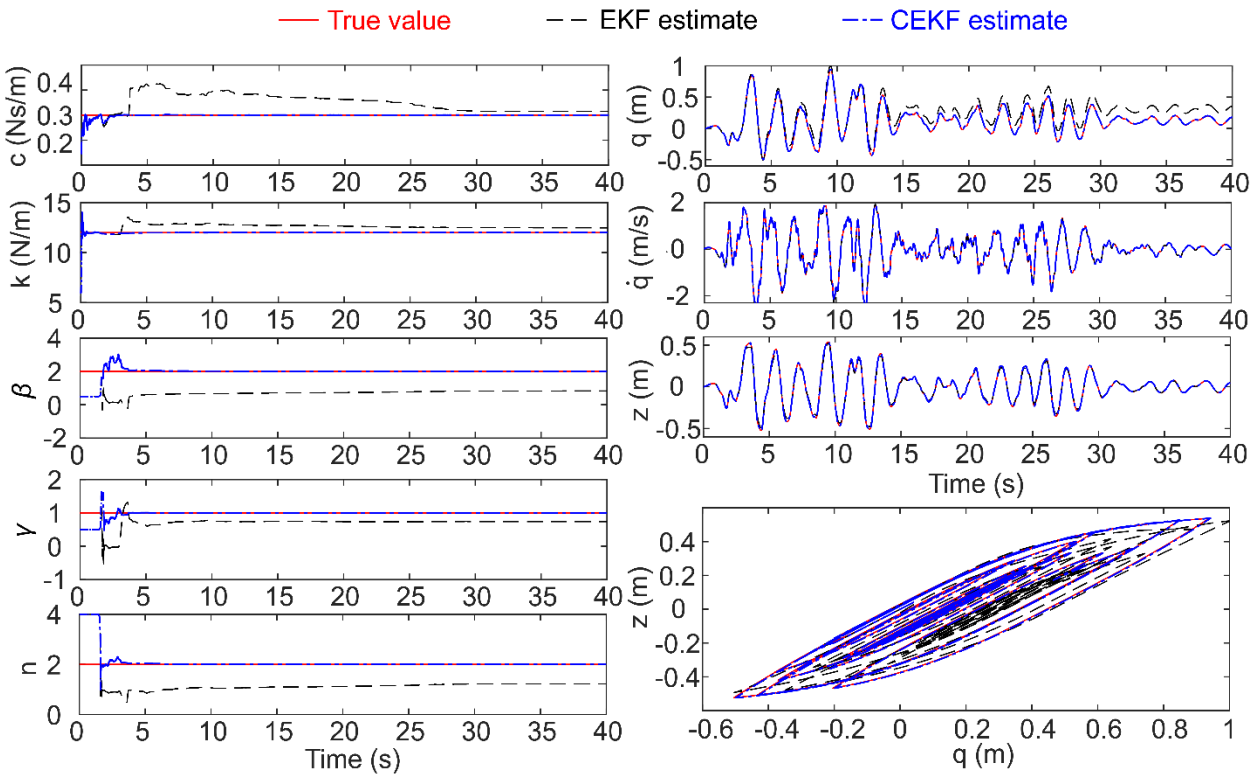


Fig. 5 Updating results on the SDOF Bouc-Wen hysteretic model using numerically linearized system and measurement equations

Table 3 summarizes the final estimated values of system parameters using different identification algorithms. It is shown that the proposed CEKF can provide more accurate estimation results than EKF. Comparing to the estimation results in Table 2, the difference between the final estimated values using numerically linearized equations and analytically linearized equations is within $\pm 1\%$. In terms of simulation time, it takes 2.36 s using the analytical derivative, while it takes 4.36 s using the numerical derivative (about 85% longer time than the analytical approach). Both simulations are conducted on a laptop PC with Intel® Core™ i7-8750H (2.20 GHz) and 8 GB RAM memory.

Table 3 Comparison of estimation results on the SDOF Bouc-Wen hysteretic model using numerically evaluated partial derivatives

Parameters	Actual values	EKF		CEKF	
		Values	Errors (%)	Values	Errors (%)
c (Ns/m)	0.3	0.3153	5.1039	0.3007	0.2211
k (N/m)	12	12.4706	3.9213	11.9959	−0.0345
β	2	0.8291	−58.5466	2.0111	0.5530
γ	1	0.7401	−25.9944	1.0030	0.3021
n	2	1.2218	−38.9081	2.0081	0.4049

4.1.4 Comparison with unscented Kalman filter (UKF)

The performance of CEKF on parameter identification is further investigated through comparison with unscented Kalman filter (UKF). The same SDOF Bouc-Wen hysteretic model is updated using UKF and the final estimated values of system parameters is compared with the results of CEKF using analytically evaluated partial derivatives. Table 4 summarizes the comparison results, where the largest estimation error by CEKF is 0.5464% (for parameter β), and the largest error by UKF is 0.2177% (for the same parameter β). It is expected that UKF provides lower estimation errors than CEKF, as UKF achieves higher order accuracy with the unscented transformation. At the same time, it is found that the CEKF with analytical derivative computes much faster than UKF. On a laptop PC with Intel® Core™ i7-8750H (2.20 GHz) and 8 GB RAM memory, the CEKF estimation takes 2.36 s to finish, while the UKF takes 8.05 s, which is about 2.41 times longer than CEKF.

Table 4 Comparison of estimation results on the SDOF Bouc-Wen hysteretic model using UKF and CEKF

Parameters	Actual values	UKF		CEKF	
		Values	Errors (%)	Values	Errors (%)
c (Ns/m)	0.3	0.3003	0.1092	0.3007	0.2185
k (N/m)	12	11.9987	−0.0111	11.9958	−0.0353
β	2	2.0044	0.2177	2.0109	0.5464
γ	1	1.0024	0.2352	1.0026	0.2634
n	2	2.0034	0.1707	2.0080	0.4011

4.2 Numerical simulation – four-story shear structure

The second validating example is a four-story shear structure, as shown in Fig. 6. In this example, each inter-story element consists of viscous damping and hysteresis. The hysteretic force within the i-th story is $k_i z_i$, where the hysteretic displacement z_i is described by the differentiable Bouc-Wen model.

$$\dot{z}_i = \begin{cases} \dot{q}_i\left(1 - (\tanh(\rho z_i) z_i)^n (\gamma + \beta \tanh(\rho z_i \dot{q}_i))\right) & i = 1 \\ (\dot{q}_i - \dot{q}_{i-1})\left(1 - (\tanh(\rho z_i) z_i)^n \left(\gamma + \beta \tanh\left(\rho z_i (\dot{q}_i - \dot{q}_{i-1})\right)\right)\right) & i = 2,3,4 \end{cases} \tag{31}$$

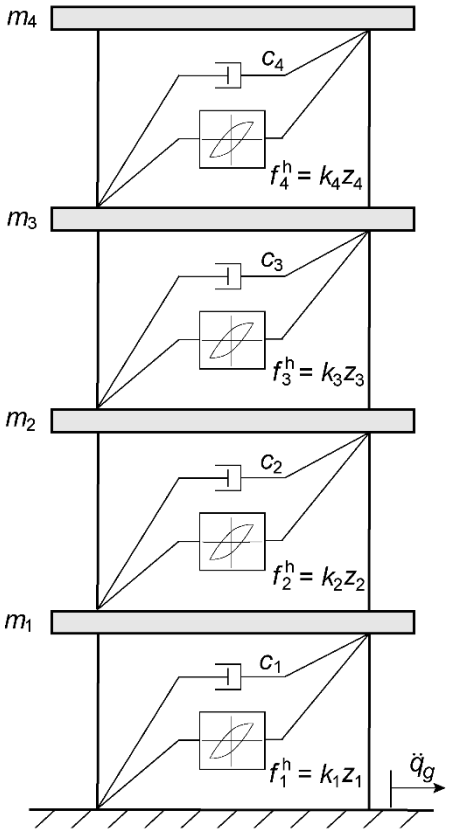


Fig. 6 Four-story shear structure

The equation of motion for this four-story structure with nonlinear hysteresis can be expressed as:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{z} = -\mathbf{M}\boldsymbol{\iota}\left(\ddot{q}_g + w\right) \tag{32}$$

where $\mathbf{q} \in \mathbb{R}^4$ is the displacement vector, $\mathbf{z} \in \mathbb{R}^4$ is the hysteretic displacement vector with each z_i described by Eq. (31), $\mathbf{M} \in \mathbb{R}^{4\times4}$ is the mass matrix, $\mathbf{C} \in \mathbb{R}^{4\times4}$ is the damping matrix, $\boldsymbol{\iota} = \{\mathbf{1}\} \in \mathbb{R}^4$ is the influence vector, $\ddot{q}_g \in \mathbb{R}$ is the ground acceleration, $w \in \mathbb{R}$ is the input uncertainty, and $\mathbf{K}$ contains the inter-story stiffness values:

$$\mathbf{K} = \begin{pmatrix} k_1 & -k_2 & & \\ & k_2 & -k_3 & \\ & & k_3 & -k_4 \\ & & & k_4 \end{pmatrix} \tag{33}$$

The state-space system equation for parameter identification can be formulated as:

$$\mathbf{x} = \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \\ \mathbf{z} \\ \mathbf{c} \\ \mathbf{k} \\ \beta \\ \gamma \\ n \end{pmatrix} \qquad \dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}, \ddot{q}_g, w) = \begin{pmatrix} \dot{\mathbf{q}} \\ -\mathbf{M}^{-1}(\mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{z}) - \boldsymbol{\iota}(\ddot{q}_g + w) \\ \dot{\mathbf{z}} \\ \mathbf{0} \\ \mathbf{0} \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{34}$$

Here $\dot{\mathbf{z}}$ is determined by Eq. (31). Table 5 lists values of model parameters of the four-story shear structure. Except for mass, all the other parameters are identified by EKF and CEKF.

Table 5 Parameter values

Parameters	Initial value	Actual value
$m_1 \sim m_4$ (kg)	5	5
c_1 (Ns/m)	5	9
c_2 (Ns/m)	5	8
c_3 (Ns/m)	5	7
c_4 (Ns/m)	5	6
k_1 (kN/m)	1.5	1.0
k_2 (kN/m)	1.5	1.2
k_3 (kN/m)	1.5	1.6
k_4 (kN/m)	1.5	2.0
β	0.5	2
γ	0.5	1
n	4	2

The same scaled El Centro earthquake excitation used in Section 4.1 is applied to excite the structure. It is assumed that all the four floors are instrumented with sensors measuring acceleration at sampling frequency of 200 Hz. At $t = k\Delta t$, the absolute accelerations of all floors are measured:

$$\mathbf{y}_k = -\mathbf{M}^{-1}(\mathbf{C}\dot{\mathbf{q}}_k + \mathbf{K}\mathbf{z}_k) + \mathbf{v}_k \tag{35}$$

The covariance of the process noise is set as $\Sigma_w = (10^{-2}\ \mathrm{m/s^2})^2$, and the covariance of sensor noise is set as $\mathbf{\Sigma_v} = (10^{-2}\ \mathrm{m/s^2})^2\mathbf{I}$.

To identify the model parameters, the standard EKF and the proposed CEKF are conducted. In the identification process, the system equation and measurement equation are linearized analytically. For CEKF, inequality constraints applied on the parameters are listed as follows:

$$c_i \geq 0, \qquad k_i \geq 0, \qquad \beta + \gamma \geq 0, \qquad \beta - \gamma \geq 0, \qquad n \geq 1 \tag{36}$$

Fig. 7 shows the time histories of the *a posteriori* estimates of the model parameters. With constraints applied during the estimation process, CEKF can recursively update all the parameters from their initial values to the corresponding true values. On the other hand, without applying constraints, EKF cannot guarantee that the estimates of model parameters, especially the hysteretic parameters, converge to their true values.

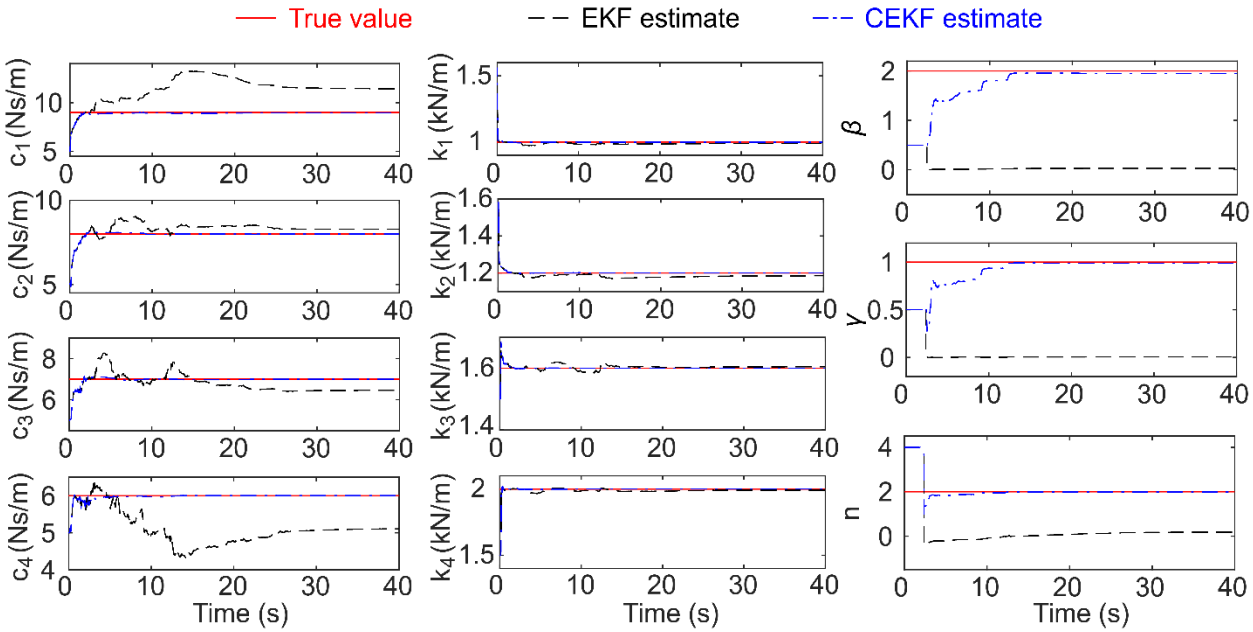


Fig. 7 Updating results using EKF and CEKF on the four-story shear structure

Table 6 shows the comparison of the final estimated values using EKF and CEKF algorithms. Similar to the updating results of the SDOF Bouc-Wen hysteretic model, it shows that CEKF can accurately identify the model parameters, while EKF performs poorly on damping and hysteretic parameters. All CEKF estimation errors are within $\pm 2.5\%$. EKF estimation errors of hysteretic parameters are larger than 90% and the largest estimation error is close to 100% for parameter γ.

Table 6 Comparison of estimated results using EKF and CEKF on the four-story shear structure

Parameters	Actual values	EKF		CEKF	
		Values	Errors (%)	Values	Errors (%)
c_1 (Ns/m)	9	11.4241	26.9347	8.9564	−0.4850
c_2 (Ns/m)	8	8.2740	3.4250	7.9921	−0.0983
c_3 (Ns/m)	7	6.4602	−7.7120	6.9949	−0.0728
c_4 (Ns/m)	6	5.1075	−14.8746	6.0043	0.0712
k_1 (kN/m)	1.0	0.9900	−0.9956	1.0000	−0.0040
k_2 (kN/m)	1.2	1.1851	1.2415	1.2006	0.0479
k_3 (kN/m)	1.6	1.6050	0.3102	1.6004	0.0219
k_4 (kN/m)	2.0	1.9925	−0.3742	2.0001	0.0073
β	2	0.0332	−98.3396	1.9526	−2.3707
γ	1	0.0072	−99.2777	0.9888	−1.1231
n	2	0.1812	−90.9397	1.9906	−0.4713

4.3 Experimental test – four-story shear structure

To further validate the proposed CEKF algorithm, a four-story laboratory structure is tested (Fig. 8). The structure is made of aluminum alloy with the total height of 1.182 m. The weight of each floor is measured by scale. The structure is mounted on a shake table which provides horizontal base excitation. To measure both the response of the structure and the base excitation, an accelerometer (Crossbow CXL01LF1) and a displacement sensor are instrumented on each floor and the shake table, as illustrated in the figure. During the estimation process, only the acceleration data are used. The displacement data are used later to further evaluate the estimation performance. The sampling frequency is set as 200 Hz.

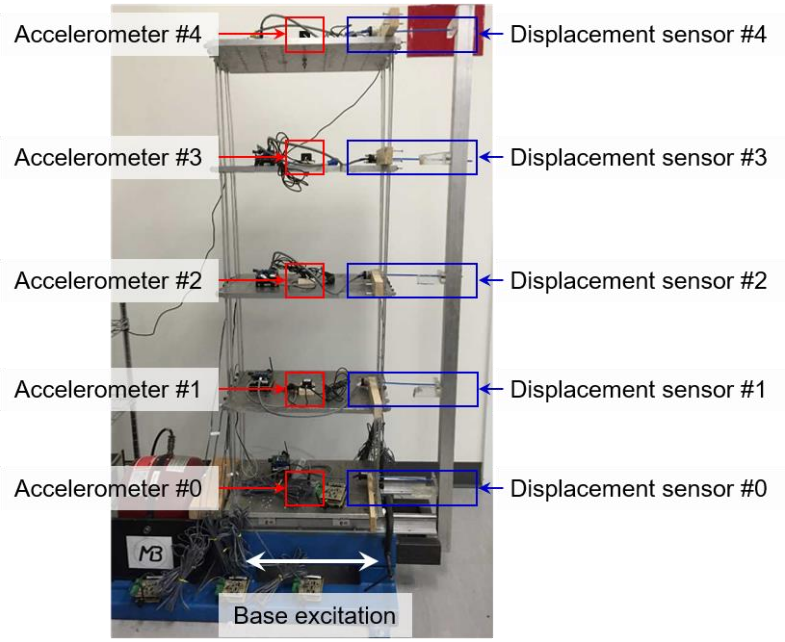
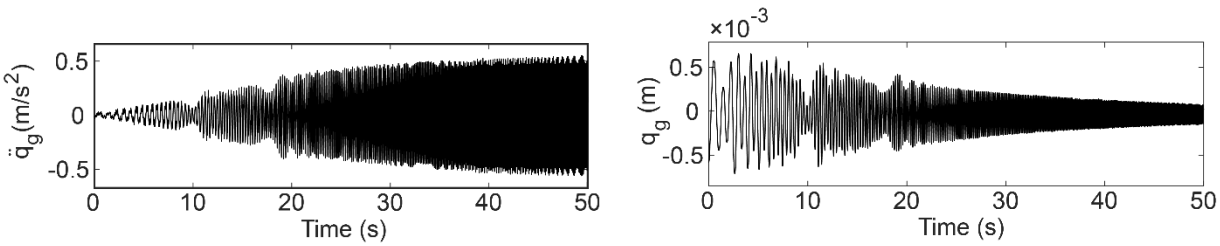


Fig. 8 Experimental setup

To excite the structure, a scaled chirp signal from 0 Hz to 10 Hz is generated as ground excitation. The measured ground acceleration and displacement are shown in Fig. 9.



(a) Ground acceleration (b) Ground displacement

Fig. 9 Ground acceleration and displacement

To identify the model parameters together with the system states, EKF and CEKF are conducted. The acceleration data from all the accelerometers are used for parameter identification. The initial estimates are set as $c_{i,0|-1} = 5$ Ns/m, $k_{i,0|-1} = 1.5$ kN/m, $\beta_{0|-1} = 0.5$, $\gamma_{0|-1} = 0.5$, and $n_{0|-1} = 4$. The covariance of the process noise is set as $\Sigma_w = (3.5 \times 10^{-3}\ \mathrm{m/s^2})^2$, and the covariance of sensor noise is set as $\mathbf{\Sigma_v} = (3.5 \times 10^{-3}\ \mathrm{m/s^2})^2\mathbf{I}$. For CEKF, the same inequality constraints in Eq. (36) are incorporated during the estimation process.

The time histories of the *a posteriori* estimates of the model parameters of the four-story shear structure using experimental data are plotted in Fig. 10. It is observed that the updating results of EKF and CEKF are slightly different. It should be noted that estimates generated from CEKF always stay within the feasible domain, while some estimates from EKF fail to satisfy the constrains during the identification process. For example, EKF estimate of c_3 become negative at around 20 s as shown in the figure.

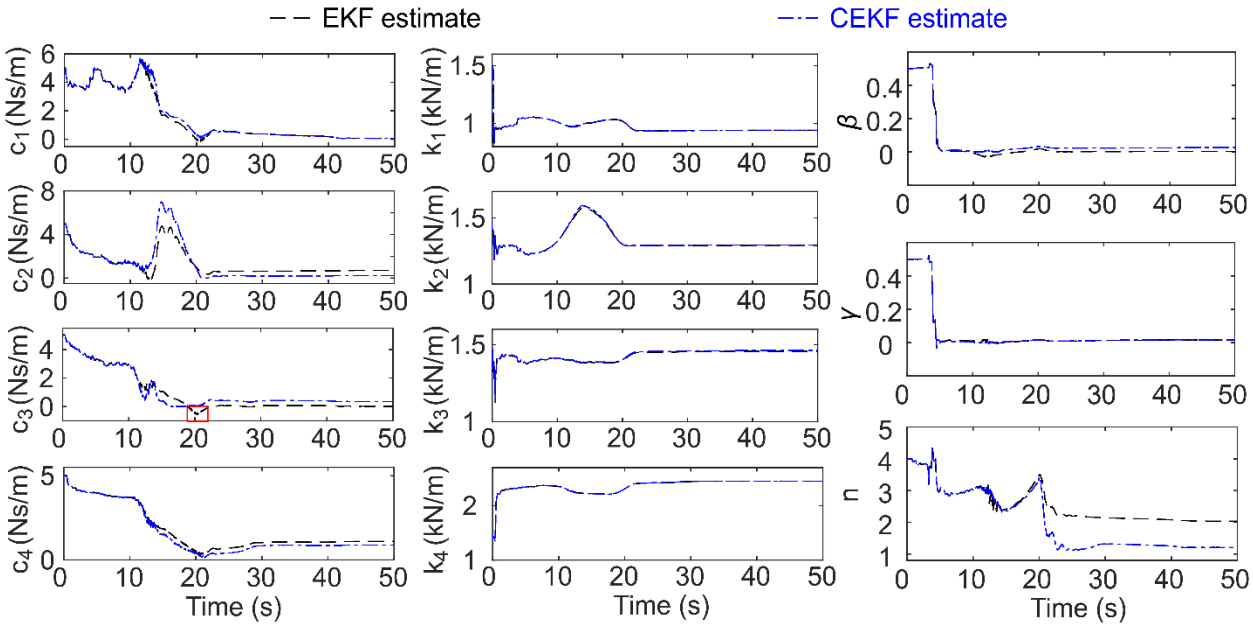


Fig. 10 Updating results using EKF and CEKF on the four-story shear structure using experimental data

Table 7 summarizes the estimated results provided by EKF and CEKF for the four-story shear structure using experimental data. Both EKF and CEKF results show that lower stories demonstrate much less inter-story stiffness due to significant *P*-Δ effect of the lab structure. The stiffness values estimated by EKF and CEKF values are similar to those obtained from previous frequency domain model updating approaches (assuming linear structure) [33, 34]. It should be noted that the hysteretic parameters estimated by EKF are not reasonable, as β is smaller than γ. On the other hand, CEKF ensures the hysteretic parameters remain within realistic bounds.

Table 7 Estimated results using EKF and CEKF on the four-story shear structure

Parameters	EKF	CEKF
c_1 (Ns/m)	0.0594	0.0635
c_2 (Ns/m)	0.7010	0.2309
c_3 (Ns/m)	0.0219	0.3562
c_4 (Ns/m)	1.1048	0.8935
k_1 (kN/m)	0.9418	0.9418
k_2 (kN/m)	1.2932	1.2958
k_3 (kN/m)	1.4560	1.4612
k_4 (kN/m)	2.4484	2.4520
β	0.0032	0.0257
γ	0.0152	0.0163
n	2.0319	1.2079

To investigate the performance of parameter identification algorithms, structural responses are simulated using the initial parameter values and the updated parameter values from EKF and CEKF. Fig. 11 plots the acceleration responses of the shear structure simulated using initial, EKF, and CEKF updated model parameters. Acceleration responses of entire time span from 0 s to 50 s are plotted in Fig. 11(a). The close-up plots of 5 s to 8 s are shown in Fig. 11(b). Both EKF and CEKF updated parameters can provide acceleration responses close to measurement data, while initial model parameters cannot generate accurate acceleration response at the beginning. The close-up plots of 30 s to 33 s are shown in Fig. 11(c). It can be observed that acceleration responses provided by all the model parameters are similar to each other and close to the measurement data as time increases.

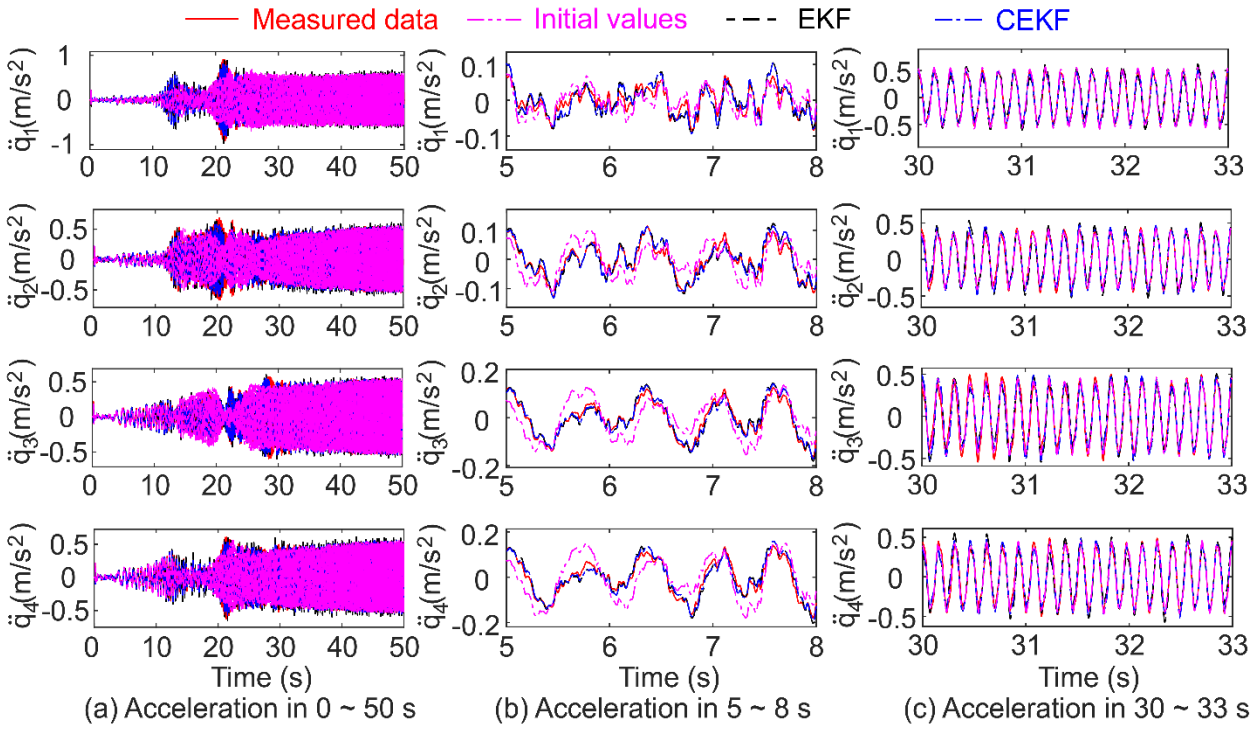


Fig. 11 Simulated acceleration responses of the shear structure

Fig. 12 plots the displacement responses of the shear structure simulated using initial, EKF, and CEKF updated model parameters. Displacement responses of entire time span from 0 s to 50 s are plotted in Fig. 12(a). The close-up plots of 5 s to 8 s are shown in Fig. 12(b). The simulation results show that both EKF and CEKF updated parameters provide similar displacement responses at the beginning. Comparing to the displacement responses generated from initial model parameters, the EKF and CEKF responses are much closer to the measurement data. The close-up plots of 30 s to 33 s are shown in Fig. 12(c). It can be seen from the figure that CEKF updated parameters performs consistently well over time. On the other hand, the displacements of EKF updated parameters show obvious oscillation and differ from the measurement data. In addition, initial parameters can generate relatively accurate displacement responses in the later part of simulation.

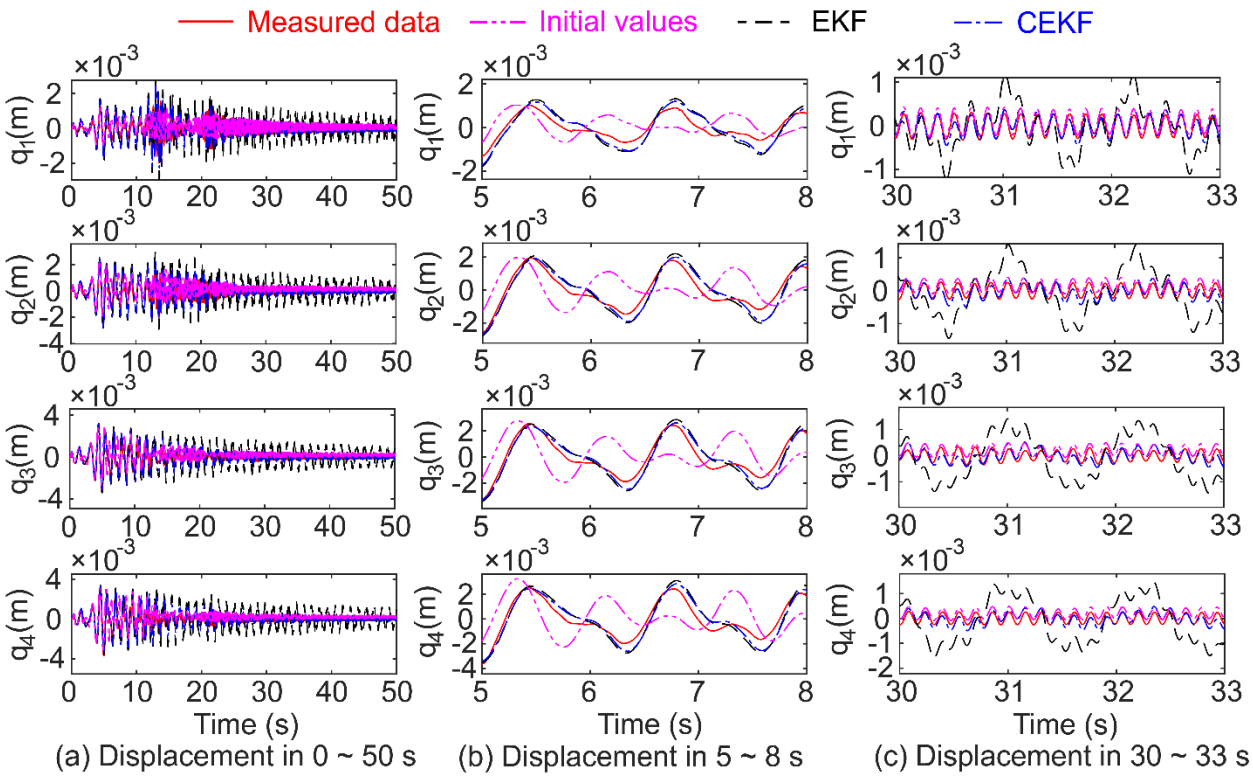


Fig. 12 Simulated displacement responses of the shear structure

The root mean square (RMS) errors between the simulated responses and all experimental measurements are summarized in Table 8. Both the EKF and CEKF updated parameter values can reduce the simulation error in terms of acceleration response. CEKF performs better than EKF. In addition, the reliable parameter values updated using CEKF achieve less simulation error for the displacement responses. On the other hand, the displacement responses calculated using EKF parameter values have larger RMS errors than those simulated using initial parameters.

Table 8 Simulated responses RMS error comparison using full measurements

RMS error	Initial	EKF	CEKF
$\ddot{q}_1$ $(\times 10^{-2}\ \mathrm{m/s^2})$	10.7044	7.4750	3.5921
$\ddot{q}_2$ $(\times 10^{-2}\ \mathrm{m/s^2})$	12.0739	5.7188	4.3160
$\ddot{q}_3$ $(\times 10^{-2}\ \mathrm{m/s^2})$	6.9700	4.8263	3.8885
$\ddot{q}_4$ $(\times 10^{-2}\ \mathrm{m/s^2})$	7.1973	5.8907	3.4961
q_1 $(\times 10^{-4}\ \mathrm{m})$	3.0880	4.9317	2.3347
q_2 $(\times 10^{-4}\ \mathrm{m})$	4.6322	6.6412	3.1165
q_3 $(\times 10^{-4}\ \mathrm{m})$	5.5168	7.9809	3.6720
q_4 $(\times 10^{-4}\ \mathrm{m})$	5.9176	8.7002	4.0522

To further investigate the performance of CEKF, parameters of the four-story laboratory structure are updated using partial measurement data from accelerometers #3 and #4 only (Fig. 8). Using the updated parameters, structural responses are again simulated and compared with the measured responses. Table 9 summarizes the RMS errors between the simulated responses and all experimental measurements. The results demonstrate that overall CEKF performs better than EKF. Comparing to the RMS errors shown in Table 8, simulation using the EKF-updated parameters happens to provide fairly accurate acceleration responses, but the displacement responses have large errors from the experimental measurements. Without incorporating constraints, EKF achieves accurate acceleration responses at the expense that some updated parameters violate reasonable constraints. These unreliable parameters result in the large errors of displacement

responses. On the other hand, CEKF performance is consistent in all simulated responses. Compared with previous simulation using CEKF estimates from full measurements (in Table 8), the RMS errors of $\ddot{q}_1$, $\ddot{q}_2$, and $\ddot{q}_4$ increase slightly while the RMS errors of other responses decrease.

Table 9 Simulated responses RMS error comparison using partial measurements (accelerometers #3 and #4 only)

RMS error	Initial	EKF	CEKF
$\ddot{q}_1$ $(\times 10^{-2}\ \mathrm{m/s^2})$	10.7044	3.4186	4.0514
$\ddot{q}_2$ $(\times 10^{-2}\ \mathrm{m/s^2})$	12.0739	3.6888	4.6096
$\ddot{q}_3$ $(\times 10^{-2}\ \mathrm{m/s^2})$	6.9700	4.5028	3.0547
$\ddot{q}_4$ $(\times 10^{-2}\ \mathrm{m/s^2})$	7.1973	4.3323	3.6840
q_1 $(\times 10^{-4}\ \mathrm{m})$	3.0880	6.0123	2.2866
q_2 $(\times 10^{-4}\ \mathrm{m})$	4.6322	9.7278	3.0068
q_3 $(\times 10^{-4}\ \mathrm{m})$	5.5168	12.1781	3.4511
q_4 $(\times 10^{-4}\ \mathrm{m})$	5.9176	12.9802	3.8775

5 Summary and Future Work

This paper investigates nonlinear parameter identification of hysteretic systems using the constrained extended Kalman filter (CEKF). A differentiable Bouc-Wen model is proposed for capturing the hysteretic characteristics of the structural system. This differentiable Bouc-Wen model enables parameter identification using EKF and CEKF which require to linearize system equation to propagate state estimate and covariance. Numerical simulation and experimental test have shown that EKF can easily result in unreliable estimates of model parameters due to large linearization error for the highly nonlinear system. Comparing to EKF, the proposed CEKF can effectively prevent the estimates of model parameters from being unrealistic and finally provide reasonable estimates by applying constraints on parameters during the estimation process.

While CEKF has demonstrated more reliable performance than EKF in parameter identification, comprehensive comparison between CEKF and other algorithms, such as UKF and particle filter, can be conducted to further identify their pros and cons. In addition, future research will be needed to investigate the CEKF performance using partial measurements on larger structures, which presents more challenging conditions for the estimator.

Reference

1. Ismail, M., F. Ikhouane, and J. Rodellar, *The hysteresis Bouc-Wen model, a survey*. Archives of Computational Methods in Engineering, 2009. **16**(2): p. 161-188.
2. Hassani, V., T. Tjahjowidodo, and T.N. Do, *A survey on hysteresis modeling, identification and control.* Mechanical systems and signal processing, 2014. **49**(1-2): p. 209-233.
3. Spencer Jr, B., et al., *Phenomenological model for magnetorheological dampers.* Journal of engineering mechanics, 1997. **123**(3): p. 230-238.
4. Yang, G., et al., *Dynamic modeling of large-scale magnetorheological damper systems for civil engineering applications.* Journal of Engineering Mechanics, 2004. **130**(9): p. 1107-1114.
5. Miah, M.S., et al., *Nonlinear modeling of a rotational MR damper via an enhanced Bouc–Wen model.* Smart Materials and Structures, 2015. **24**(10): p. 105020.
6. Sengupta, P. and B. Li, *Modified Bouc–Wen model for hysteresis behavior of RC beam–column joints with limited transverse reinforcement.* Engineering Structures, 2013. **46**: p. 392-406.
7. Ning, C.-L., B. Yu, and B. Li, *Beam-column joint model for nonlinear analysis of non-seismically detailed reinforced concrete frame.* Journal of Earthquake Engineering, 2016. **20**(3): p. 476-502.
8. Soneji, B. and R. Jangid, *Influence of soil–structure interaction on the response of seismically isolated cable-stayed bridge.* Soil Dynamics and Earthquake Engineering, 2008. **28**(4): p. 245-257.
9. Kwok, N., et al., *Bouc–Wen model parameter identification for a MR fluid damper using computationally efficient GA.* ISA transactions, 2007. **46**(2): p. 167-179.
10. Charalampakis, A. and V. Koumousis, *Identification of Bouc–Wen hysteretic systems by a hybrid evolutionary algorithm.* Journal of Sound and Vibration, 2008. **314**(3-5): p. 571-585.

11. Monti, G., G. Quaranta, and G.C. Marano, *Genetic-algorithm-based strategies for dynamic identification of nonlinear systems with noise-corrupted response.* Journal of Computing in Civil Engineering, 2009. **24**(2): p. 173-187.

12. Charalampakis, A. and C. Dimou, *Comparison of evolutionary algorithms for the identification of Bouc-Wen Hysteretic Systems.* Journal of Computing in Civil Engineering, 2013. **29**(3): p. 04014053.

13. Smyth, A., et al., *On-line parametric identification of MDOF nonlinear hysteretic systems.* Journal of Engineering Mechanics, 1999. **125**(2): p. 133-142.

14. Lin, J.W., et al., *On-line identification of non-linear hysteretic structural systems using a variable trace approach.* Earthquake Engineering & Structural Dynamics, 2001. **30**(9): p. 1279-1303.

15. Yang, J.N. and S. Lin, *On-line identification of non-linear hysteretic structures using an adaptive tracking technique.* International Journal of Non-Linear Mechanics, 2004. **39**(9): p. 1481-1491.

16. Ching, J., J.L. Beck, and K.A. Porter, *Bayesian state and parameter estimation of uncertain dynamical systems.* Probabilistic engineering mechanics, 2006. **21**(1): p. 81-96.

17. Li, S., Y. Suzuki, and M. Noori, *Improvement of parameter estimation for non-linear hysteretic systems with slip by a fast Bayesian bootstrap filter.* International journal of non-linear mechanics, 2004. **39**(9): p. 1435-1445.

18. Chatzi, E.N. and A.W. Smyth, *Particle filter scheme with mutation for the estimation of time-invariant parameters in structural health monitoring applications.* Structural Control and Health Monitoring, 2013. **20**(7): p. 1081-1095.

19. Olivier, A. and A.W. Smyth, *Particle filtering and marginalization for parameter identification in structural systems.* Structural Control and Health Monitoring, 2017. **24**(3): p. e1874.

20. Gandomi, A.H., et al., *Metaheuristic applications in structures and infrastructures*. 2013: Newnes.

21. Hu, G., S. Gao, and Y. Zhong, *A derivative UKF for tightly coupled INS/GPS integrated navigation.* ISA Transactions, 2015. **56**: p. 135-144.

22. He, W., et al., *State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation.* International Journal of Electrical Power & Energy Systems, 2014. **62**: p. 783-791.

23. Olivier, A. and A.W. Smyth, *A marginalized unscented Kalman filter for efficient parameter estimation with applications to finite element models.* Computer Methods in Applied Mechanics and Engineering, 2018. **339**: p. 615-643.

24. Wu, M. and A.W. Smyth, *Application of the unscented Kalman filter for real-time nonlinear structural system identification.* Structural Control and Health Monitoring, 2007. **14**(7): p. 971-990.

25. Yang, Y. and F. Ma, *Constrained Kalman filter for nonlinear structural identification.* Modal Analysis, 2003. **9**(12): p. 1343-1357.

26. Sen, S. and B. Bhattacharya, *Online structural damage identification technique using constrained dual extended Kalman filter.* Structural Control and Health Monitoring, 2017. **24**(9): p. e1961.

27. Gupta, N. and R. Hauser, *Kalman filtering with equality and inequality state constraints.* arXiv preprint arXiv: 0709.2791, 2007.

28. Jeen-Shang, L. and Z. Yigong, *Nonlinear structural identification using extended Kalman filter.* Computers & structures, 1994. **52**(4): p. 757-764.

29. Yang, J.N., et al., *An adaptive extended Kalman filter for structural damage identification.* Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures, 2006. **13**(4): p. 849-867.

30. Simon, D., *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. 2006: John Wiley & Sons.

31. Boyd, S.P. and L. Vandenberghe, *Convex Optimization*. 2004, Cambridge, UK ; New York: Cambridge University Press. xiii, 716 p.

32. Wang, L. and Z. Lu, *Identification of Bouc-Wen hysteretic parameters based on enhanced response sensitivity approach.* Journal of physics: conference series, 2017. **842**(1): p. 012021.

33. Li, D., X. Dong, and Y. Wang, *Model updating using sum of squares (SOS) optimization to minimize modal dynamic residuals.* Structural Control and Health Monitoring, 2018. **25**(12): p. e2263.

34. Hong, Y., et al., *Model-updating with experimental frequency response function considering general damping.* Advances in Structural Engineering, 2018. **21**(1): p. 82-92.

Appendix

1. Matrix differentiation

Suppose that $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a matrix and $f(\mathbf{A})$ is a scalar function. The partial derivative of a scalar function with respect to a matrix is defined as:

$$\frac{\partial f}{\partial \mathbf{A}} = \begin{pmatrix} \frac{\partial f}{\partial A_{11}} & \frac{\partial f}{\partial A_{12}} & \cdots & \frac{\partial f}{\partial A_{1n}} \\ \frac{\partial f}{\partial A_{21}} & \frac{\partial f}{\partial A_{22}} & \cdots & \frac{\partial f}{\partial A_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial A_{m1}} & \frac{\partial f}{\partial A_{m2}} & \cdots & \frac{\partial f}{\partial A_{mn}} \end{pmatrix} \tag{A.1}$$

Consider matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times n}$, and a scalar function $f(\mathbf{A}) = \text{Trace}(\mathbf{ABA}^{\text{T}})$. The partial derivative of $f(\mathbf{A})$ with respect to $\mathbf{A}$ is found as [30]:

$$\frac{\partial f}{\partial \mathbf{A}} = \mathbf{AB}^{\text{T}} + \mathbf{AB} \tag{A.2}$$

If $\mathbf{B} = \mathbf{B}^{\text{T}}$, the partial derivative is simplified as:

$$\frac{\partial f}{\partial \mathbf{A}} = 2\mathbf{AB} \tag{A.3}$$

Consider matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, vectors $\mathbf{x} \in \mathbb{R}^{m}$ and $\mathbf{y} \in \mathbb{R}^{n}$, and a scalar function $f(\mathbf{A}) = \mathbf{x}^{\text{T}}\mathbf{Ay}$. The scalar function $f(\mathbf{A})$ can be expanded as:

$$f(\mathbf{A}) = \mathbf{x}^{\text{T}}\mathbf{Ay} = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum\nolimits_{i,j} x_i A_{ij} y_j \tag{A.4}$$

The partial derivative of $f(\mathbf{A})$ with respect to $\mathbf{A}$ is then calculated as:

$$\frac{\partial f}{\partial \mathbf{A}} = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_my_1 & x_my_2 & \cdots & x_my_n \end{pmatrix} = \mathbf{x}\mathbf{y}^{\mathrm{T}} \tag{A.5}$$

2. Derivation of Kalman gain for CEKF

Consider the equality constrained optimization problem:

$$\begin{aligned} &\underset{\mathbf{L}}{\text{minimize}}\ \text{Trace}\left((\mathbf{I}-\mathbf{L}_k\mathbf{H}_k^{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{I}-\mathbf{L}_k\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}+\mathbf{L}_k\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}}\mathbf{L}_k^{\mathrm{T}}\right) \\ &\text{subject to}\ \mathbf{A}_{\mathrm{a}}\left(\hat{\mathbf{x}}_{k|k-1}+\mathbf{L}\left(\mathbf{y}_k-\hat{\mathbf{y}}_{k|k-1}\right)\right)-\mathbf{b}_{\mathrm{a}}=\mathbf{0} \end{aligned} \tag{A.6}$$

For brevity, denote the measurement innovation $\mathbf{r}_k$ and innovation covariance $\boldsymbol{\Sigma}_{\mathbf{y}_k}$ at time $t = k\Delta t$ as:

$$\mathbf{r}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} \tag{A.7}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}_k} = \mathbf{H}_k^{\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + \mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} \tag{A.8}$$

Using Lagrange multiplier $\boldsymbol{\nu} \in \mathbb{R}^{n_{ac}}$, the Lagrangian for the problem is:

$$\begin{aligned} \mathcal{L}(\mathbf{L},\boldsymbol{\nu}) = &\ \text{Trace}\left((\mathbf{I}-\mathbf{L}_k\mathbf{H}_k^{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{I}-\mathbf{L}_k\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}+\mathbf{L}_k\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}}\mathbf{L}_k^{\mathrm{T}}\right) \\ &+\boldsymbol{\nu}^{\mathrm{T}}\left(\mathbf{A}_{\mathrm{a}}\left(\hat{\mathbf{x}}_{k|k-1}+\mathbf{L}\mathbf{r}_k\right)-\mathbf{b}_{\mathrm{a}}\right) \end{aligned} \tag{A.9}$$

Note that $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} = \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}^{\mathrm{T}}$ and $\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} = (\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}})^{\mathrm{T}}$. Based on Eq. (A.3) and Eq. (A.5), the partial derivatives of $\mathcal{L}(\mathbf{L},\boldsymbol{\nu})$ with respect to $\mathbf{L}$ and $\boldsymbol{\nu}$, respectively, are:

$$\frac{\partial}{\partial \mathbf{L}}\mathcal{L}(\mathbf{L},\boldsymbol{\nu}) = -2(\mathbf{I}-\mathbf{L}\mathbf{H}_k^{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}} + 2\mathbf{L}\mathbf{H}_k^{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}(\mathbf{H}_k^{\mathbf{v}})^{\mathrm{T}} + \mathbf{A}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{\nu}\mathbf{r}_k^{\mathrm{T}} \tag{A.10}$$

$$\frac{\partial}{\partial \boldsymbol{\nu}}\mathcal{L}(\mathbf{L},\boldsymbol{\nu}) = \mathbf{A}_{\mathrm{a}}\left(\hat{\mathbf{x}}_{k|k-1}+\mathbf{L}\mathbf{r}_k\right)-\mathbf{b}_{\mathrm{a}} \tag{A.11}$$

The optimality requires that both partial derivatives are zero. Assume $\mathbf{A}_{\mathrm{a}} \in \mathbb{R}^{n_{\mathrm{ac}}\times n_{\mathbf{x}}}$ is a full row-rank matrix with $\mathrm{rank}(\mathbf{A}_{\mathrm{a}}) = n_{\mathrm{ac}} \leq n_{\mathbf{x}}$. We can express the Kalman gain $\mathbf{L}$ as a function of the Lagrange multiplier $\boldsymbol{\nu}$. First, the partial derivative $\frac{\partial}{\partial \mathbf{L}}\mathcal{L}(\mathbf{L},\boldsymbol{\nu})$ in Eq. (A.10) is set as zero. Solving the equation $\frac{\partial}{\partial \mathbf{L}}\mathcal{L}(\mathbf{L},\boldsymbol{\nu}) = \mathbf{0}$ for $\mathbf{L}$ provides:

$$\mathbf{L} = \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1} - \frac{1}{2}\mathbf{A}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{\nu}\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1} \tag{A.12}$$

Substituting Kalman gain expression Eq. (A.12) into the partial derivative Eq. (A.11) and solving the equation $\frac{\partial}{\partial \boldsymbol{\nu}}\mathcal{L}(\mathbf{L},\boldsymbol{\nu}) = \mathbf{0}$ for $\boldsymbol{\nu}$ provides:

$$\boldsymbol{\nu} = 2(\mathbf{A}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}}^{\mathrm{T}})^{-1}\left(\mathbf{A}_{\mathrm{a}}\left(\hat{\mathbf{x}}_{k|k-1}+\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k\right)-\mathbf{b}_{\mathrm{a}}\right)\left(\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k\right)^{-1} \tag{A.13}$$

Denote the regular Kalman gain and the *a posterior* state estimate of the unconstrained EKF as:

$$\tilde{\mathbf{L}}_k = \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}(\mathbf{H}_k^{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1} \tag{A.14}$$

$$\tilde{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1}+\tilde{\mathbf{L}}_k\mathbf{r}_k \tag{A.15}$$

The Lagrange multiplier $\boldsymbol{\nu}$ in Eq. (A.13) is simplified as:

$$\boldsymbol{\nu} = 2(\mathbf{A}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}}^{\mathrm{T}})^{-1}\left(\mathbf{A}_{\mathrm{a}}\tilde{\mathbf{x}}_{k|k}-\mathbf{b}_{\mathrm{a}}\right)\left(\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k\right)^{-1} \tag{A.16}$$

Finally, substituting the simplified $\boldsymbol{\nu}$ into Eq. (A.12), the Kalman gain of CEKF can be rewritten as:

$$\mathbf{L}_k = \tilde{\mathbf{L}}_k - \mathbf{A}_{\mathrm{a}}^{\mathrm{T}}(\mathbf{A}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}}^{\mathrm{T}})^{-1}\left(\mathbf{A}_{\mathrm{a}}\tilde{\mathbf{x}}_{k|k-1} - \mathbf{b}_{\mathrm{a}}\right)\left(\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1}\mathbf{r}_k\right)^{-1}\mathbf{r}_k^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{y}_k}^{-1} \tag{A.17}$$