Journal of Sound and Vibration 325 (2009) 1023–1046

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Extended Kalman filtering for the detection of damage in linear mechanical structures

X. Liu, P.J. Escamilla-Ambrosio, N.A.J. Lieven *

Department of Aerospace Engineering, University of Bristol, Queens Building, University Walk, Bristol BS8 1TR, UK

ARTICLE INFO

Article history:
Received 25 March 2008
Received in revised form
31 March 2009
Accepted 1 April 2009
Handling Editor: C.L. Morfey
Available online 9 May 2009

ABSTRACT

This paper addresses the problem of assessing the location and extent of damage in a vibrating structure by means of vibration measurements. Frequency domain identification methods (e.g. finite element model updating) have been widely used in this area while time domain methods such as the extended Kalman filter (EKF) method, are more sparsely represented. The difficulty of applying EKF in mechanical system damage identification and localisation lies in: the high computational cost, the dependence of estimation results on the initial estimation error covariance matrix $\mathbf{P}_{ex}(0)$, the initial value of parameters to be estimated, and on the statistics of measurement noise $\mathbf{R}_{ex}$ and process noise $\mathbf{Q}_{ex}$. To resolve these problems in the EKF, a multiple model adaptive estimator consisting of a bank of EKF in modal domain was designed, each filter in the bank is based on different $\mathbf{P}_{ex}(0)$. The algorithm was iterated by using the weighted global iteration method. A fuzzy logic model was incorporated in each filter to estimate the variance of the measurement noise $\mathbf{R}_{ex}$. The application of the method is illustrated by simulated and real examples.

1. Introduction

Damage in structures is defined as '…changes to the material and/or geometric properties of these systems, including changes to the boundary conditions and system connectivity, which adversely affect the current or future performance of these systems' [1]. For safety reasons and because of the economic benefits that can result, the interest in the ability to detect, locate and quantify structural damage is pervasive throughout the engineering communities.

There exist many ways to classify vibration-based structural damage identification methods and the damage classification system presented by Rytter is adopted here, which defines four levels of damage assessment [2]:

(1) level 1 (detection): determination that damage is present in the structure;
(2) level 2 (localisation): level 1 plus determination of the geometric location of the damage;
(3) level 3 (quantification): level 2 plus quantification of the severity of the damage; and
(4) level 4 (prediction): level 3 plus prediction of the remaining service life of the structure.

The level 1 and level 2 damage assessments are both based on the comparison of features before and after damage. The comparison-based techniques are 'forward problems' and relatively simple. On the contrary, level 3 damage assessment

* Corresponding author.
E-mail address: Nick.Lieven@bristol.ac.uk (N.A.J. Lieven).

0022-460X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2009.04.005

Nomenclature

$\mathbf{A}$	continuous system matrix
$\mathbf{A}_f$	the sensitivity matrix of nonlinear function to state vector
$\mathbf{A}_F$	the discrete fundamental matrix at time t_k
$\mathbf{B}$	input matrix of dynamic system
$\mathbf{B}_e$	input matrix of state space model
$\mathbf{C}$	damping matrix
$\mathbf{f}(\mathbf{z}(t),\mathbf{u}(t),t)$	nonlinear function of continuous state vector
$\mathbf{f}_r(\mathbf{z}_r(t),\mathbf{u}(t),t)$	nonlinear function of reduced continuous state vector
$\mathbf{G}$	input matrix in measurement equation
$\mathbf{H}$	measurement matrix of state space model
$\mathbf{H}_0$	measurement matrix of dynamic system
$\mathbf{H}_{\mathrm{ex}}$	measurement matrix of extended Kalman filter
J^i	the cost function of ith Kalman filter in the bank
$\mathbf{K}$	stiffness matrix
$\mathbf{K}_g$	Kalman gain
$\mathbf{L}$	process noise matrix for dynamic system
$\mathbf{M}$	mass matrix
$\mathbf{p}$	state vector in the modal coordinate system defined by $\mathbf{q}\triangleq\mathbf{\Phi p}$
$\mathbf{p}_r$	the first rth elements of $\mathbf{P}$
Pr^i	the probability of the ith extended Kalman filter in the bank
P_R	the range that the bank of EKF could cover
$\mathbf{P}_{\mathrm{ex}}(k)$	the error covariance matrix associated with the updated estimate $\hat{\mathbf{z}}(k)$
$\mathbf{P}_{\mathrm{ex}}(k+1/k)$	error covariance matrix associated with the prior estimate $\hat{\mathbf{z}}(k+1/k)$
$\mathbf{P}_{\mathrm{ex}}^{\mathrm{opt}}(0)$	optimal initial error covariance matrix in the bank of Kalman filters
$\mathbf{P}_{xx}$	error covariance matrix related with $\hat{\mathbf{x}}$
$\mathbf{P}_{xx}^{\mathrm{mod}}$	error covariance matrix related with $\hat{\mathbf{x}}$ in modal coordinates
$\mathbf{P}_{\alpha\alpha}$	error covariance matrix related with $\hat{\boldsymbol{\alpha}}$
$\mathbf{q}(t)$	the displacement vector (m)
$\dot{\mathbf{q}}(t)$	the velocity vector (m/s)
$\ddot{\mathbf{q}}(t)$	the acceleration vector (m/s^2)
$\mathbf{Q}_s$	continuous covariance matrix of process noise $\mathbf{w}(t)$
$\mathbf{Q}_{\mathrm{ex}}$	discrete covariance matrix of process noise for extended Kalman filter
$\mathbf{Q}_{\mathrm{exs}}$	continuous covariance matrix of process noise for extended Kalman filter
$\mathbf{R}(k)$	variance matrix of measurement noise $\mathbf{v}(k)$
$\mathbf{R}_{\mathrm{ex}}$	variance matrix of measurement noise for extended Kalman filter, $\mathbf{R}_{\mathrm{ex}}=\mathbf{R}$
$\mathbf{R}(k)_{(i,i)}$	the ith diagonal element of $\mathbf{R}(k)$
$\mathbf{S}(k)$	covariance matrix of residual vector $\mathbf{v}(k)$
$\mathbf{u}$	force (N)
$\mathbf{v}(k)$	measurement noise
$\mathbf{w}(t)$	continuous process noise vector
$\mathbf{w}_{\boldsymbol{\alpha}}(k)$	process noise on the parameter vector equation
w_f	weighting factor which balance the TAAU and TANIS criteria.
$\mathbf{w}_{\mathrm{exs}}(t)$	the continuous process noise for EKF
W	weighting factor in weighted global iteration method
$\mathbf{x}$	state vector
$\mathbf{z}(t)$	the state vector in extend Kalman filter
$\mathbf{z}(k+1/k)$	the prediction of state vector at t_{k+1} given estimate $\mathbf{z}(k)$
$\mathbf{z}_r(t)$	the state vector in the reduced modal extended Kalman filter
$\mathbf{Z}_{\boldsymbol{\alpha}}(k)$	covariance matrix of $\mathbf{w}_{\boldsymbol{\alpha}}(k)$
$\boldsymbol{\alpha}$	parameter vector
$\hat{\boldsymbol{\alpha}}_{\mathrm{opt}}$	the optimal parameter vector estimated by the whole bank of EKFs
γ_i	the ith modal damping defined by $\gamma_i=2\zeta_i\omega_i$
$\mathbf{\Gamma}$	modal damping matrix defined by $\mathbf{\Gamma}\triangleq\mathbf{\Phi}^{\mathbf{T}}\mathbf{C}\mathbf{\Phi}$
$\mathbf{\Gamma}_r$	the reduced damping matrix containing the first rth modal dampings
$\boldsymbol{\delta}\mathbf{P}^i$	the adjustment of the previous optimal $\mathbf{P}_{\mathrm{ex}}^{\mathrm{opt}}(0)$ for the ith EKF
$\hat{\varepsilon}_v$	time-averaged normalised innovation squared
ζ_i	the damping ratio of the ith mode.
λ_i	the square of ith eigenfrequency of the undamped system
$\mathbf{\Lambda}$	eigenvalue matrix
$\mathbf{\Lambda}_r$	the reduced eigenvalue matrix containing the first rth eigenvalues
$\boldsymbol{\nu}(k)$	residual vector at time step t_k
$\mathbf{\Phi}$	mass normalised modal matrix
$\mathbf{\Phi}_r$	reduced modal matrix containing the first rth column of $\mathbf{\Phi}$
$\hat{\rho}(l)$	time-average autocorrelation of residuals
$\varphi(\hat{\mathbf{z}}(k),\mathbf{u}(k),t_k,t_{k+1})$	propagation function from $\hat{\mathbf{z}}(k)$ to $\hat{\mathbf{z}}(k+1)$
ω_i	the ith eigenfrequency of the undamped system
DoM	degree of mismatch
EKF	extended Kalman filter
EK–WGI	weighted global iteration method for extended Kalman filter
FIS	fuzzy inference system
FL–AKF	fuzzy logic-based adaptive Kalman filter
KF	Kalman filter
MMAE	multiple model adaptive estimator
MOKF	modal Kalman filter
TAAU	time-average autocorrelation of residuals
TANIS	time-averaged normalised innovation squared

(i.e. damage quantification) is a significant challenge because it is an 'inverse problem' [3]. In techniques for level 3 damage assessment, a correlated analytical finite element (FE) model is generally needed. Therefore, level 3 damage assessment is closely connected with FE model updating.

Many FE model-based identification methods have been developed. In terms of the equations they adopt, these methods can be loosely classified as frequency domain methods and time domain methods [4]. Compared with the frequency domain methods, which use modal parameters to update physical parameters in FE model, time domain methods directly update FE model parameters from input/output (I/O) time series. Therefore, they generally reveal more information for the purpose of damage identification, albeit at a higher computational cost. These methods include instrumental variables (IV) method [5], restoring force method [6], and extended Kalman filter (EKF) method [7]. Among these methods, the EKF method has been used by many researchers [8], and has been shown to behave well with relatively large I/O noise.

However, the difficulties of applying EKF hinder its application in structural damage localisation. These difficulties include: the high computational cost, the dependence of estimation results on the initial estimation error covariance matrix $\mathbf{P}_{ex}(0)$, on the initial value of parameters to be estimated, and on the statistics of measurement noise $\mathbf{R}_{ex}$ and process noise $\mathbf{Q}_{ex}$.

Accordingly, this paper strives to resolve these problems in the EKF. The proposed method is a modified EKF, where all the previous problems have been handled by means of combining different techniques. This includes a multiple model adaptive estimator (MMAE) consisting of a bank of EKF designed in modal domain (MOKF), each filter in the bank is based on a different $\mathbf{P}_{ex}(0)$. The algorithm iterates using the weighted global iteration (WGI) method. Finally, a fuzzy logic (FL) model was incorporated in each filter to estimate the variance of the measurement noise $\mathbf{R}_{ex}$.

The structure of the paper is organised as follows. In Section 2, the traditional EKF is introduced. In Section 3, the modified EKF is proposed. The application of the proposed method is illustrated by simulated and real examples in Section 4. Finally, conclusions are given in Section 5.

2. Mathematical model of the system and the extended Kalman filter

2.1. Equation of motion

Within the framework of finite element modelling, the equations of motion of a linear dynamical system can be written as

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{B}\mathbf{u}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ denote structural mass, damping and stiffness matrices, respectively; $\mathbf{u}(t)$ is the external force vector, $\mathbf{B}$ is the input matrix, $\mathbf{q}(t)$, $\dot{\mathbf{q}}(t)$ and $\ddot{\mathbf{q}}(t)$ are, respectively, the displacement, velocity and the acceleration vectors.

The corresponding state space formulation is

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_e\mathbf{u}(t) \tag{2a}$$

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) \tag{2b}$$

where $\mathbf{A}$, $\mathbf{B}_e$, $\mathbf{x}(t)$ take the well-known form

$$\mathbf{A} = \begin{bmatrix} [\mathbf{0}] & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, \quad \mathbf{B}_e = \begin{bmatrix} [\mathbf{0}] \\ \mathbf{M}^{-1}\mathbf{B} \end{bmatrix}, \quad \mathbf{x}(t) = \begin{bmatrix} \mathbf{q}(t) \\ \dot{\mathbf{q}}(t) \end{bmatrix} \tag{3}$$

The matrices of the measurement equation $\mathbf{H}$ and $\mathbf{G}$ depend on the nature of the measured quantities (displacements, velocities, accelerations, etc.).

(1) Displacements:

$$\text{if } \mathbf{y}(t) = \mathbf{H}_0\mathbf{q}(t) \tag{4a}$$

$$\text{then } \mathbf{H} = [\mathbf{H}_0 \quad [\mathbf{0}]], \quad \mathbf{G} = [\mathbf{0}] \tag{4b}$$

(2) Velocities:

$$\text{if } \mathbf{y}(t) = \mathbf{H}_0\dot{\mathbf{q}}(t) \tag{5a}$$

$$\text{then } \mathbf{H} = [[\mathbf{0}] \quad \mathbf{H}_0], \quad \mathbf{G} = [\mathbf{0}] \tag{5b}$$

(3) Accelerations:

$$\text{if } \mathbf{y}(t) = \mathbf{H}_0\ddot{\mathbf{q}}(t) \tag{6a}$$

$$\text{then } \mathbf{H} = [-\mathbf{H}_0\mathbf{M}^{-1}\mathbf{K} \quad -\mathbf{H}_0\mathbf{M}^{-1}\mathbf{C}], \quad \mathbf{G} = [\mathbf{H}_0\mathbf{M}^{-1}\mathbf{B}] \tag{6b}$$

where the matrix $\mathbf{H}_0$ is composed by zeros and ones and determines which quantities have been measured.

2.2. Extended Kalman filter for mechanical structure

The EKF approach applies the standard Kalman filter (for linear systems) to nonlinear systems with additive white noise by continually updating a linearisation around the previous state estimate, starting with an initial guess. In other words, only a linear Taylor approximation of the system function at the previous state estimate and the observation function at the corresponding predicted position are considered. This approach gives a simple and efficient algorithm to handle a nonlinear model. Furthermore, the EKF can be applied to the simultaneous identification of states and parameters. Assuming in Eq. (1) that only $\mathbf{K}$ and $\mathbf{C}$ contain the parameters $\boldsymbol{\alpha}$ to be estimated, then the state vector $\mathbf{x}$ can be extended to be

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\alpha} \end{bmatrix} \tag{7a}$$

where $\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_a]^{\mathrm{T}}$ is the parameter vector.

In this manner, the estimation problem now becomes nonlinear even if the original system dynamics are linear due to the nonlinear coupling of the states $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t)$ with $\boldsymbol{\alpha}$. If the parameters are constant ($\dot{\boldsymbol{\alpha}} = [0]$), then the system becomes

$$\dot{\mathbf{z}}(t) = \begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{q}}(t) \\ \ddot{\mathbf{q}}(t) \\ \dot{\boldsymbol{\alpha}} \end{bmatrix} \tag{8a}$$

$$= \begin{bmatrix} \dot{\mathbf{q}}(t) \\ -\mathbf{M}^{-1}\mathbf{K}\mathbf{q}(t) - \mathbf{M}^{-1}\mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{M}^{-1}\mathbf{B}\mathbf{u}(t) \\ [\mathbf{0}] \end{bmatrix} \tag{8b}$$

$$\triangleq \mathbf{f}(\mathbf{z}(t), \mathbf{u}(t), t) \tag{8c}$$

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) \tag{9a}$$

$$\triangleq \mathbf{h}(\mathbf{z}(t), \mathbf{u}(t), t) \tag{9b}$$

By discretising and linearising Eq. (8c) at time t_k ($k = 1, 2, \ldots$) using first-order Taylor series expansion, the discrete-time state space model is obtained:

$$\mathbf{z}(k+1) = \mathbf{A}_F\mathbf{z}(k) + \bar{\mathbf{U}} \tag{10a}$$

where $\mathbf{A}_F$ is the fundamental matrix at time t_k which takes the form

$$\mathbf{A}_F = \mathrm{e}^{\mathbf{A}_f \Delta t} = \mathbf{I} + \mathbf{A}_f\Delta t + (\mathbf{A}_f\Delta t)^2/2! + \cdots (\Delta t = t_{k+1} - t_k) \tag{10b}$$

$$\bar{\mathbf{U}} = \int_{t_k}^{t_{k+1}} \mathrm{e}^{\mathbf{A}_f(t_{k+1}-\tau)}(\mathbf{f}(\mathbf{z}(k), \mathbf{u}(k)) - \mathbf{A}_f\mathbf{z}(k))\,\mathrm{d}\tau \tag{10c}$$

and

$$\mathbf{A}_f = \frac{\partial \mathbf{f}}{\partial \mathbf{z}}\Big/ \mathbf{z} = \mathbf{z}(k) \tag{10d}$$

$$= \begin{bmatrix} [\mathbf{0}] & \mathbf{I} & [\mathbf{0}] & \cdots & [\mathbf{0}] \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} & -\mathbf{M}^{-1}\dfrac{\partial \mathbf{K}}{\partial \alpha_1}\mathbf{q}(k) - \mathbf{M}^{-1}\dfrac{\partial \mathbf{C}}{\partial \alpha_1}\dot{\mathbf{q}}(k) & \cdots & -\mathbf{M}^{-1}\dfrac{\partial \mathbf{K}}{\partial \alpha_a}\mathbf{q}(k) - \mathbf{M}^{-1}\dfrac{\partial \mathbf{C}}{\partial \alpha_a}\dot{\mathbf{q}}(k) \\ [\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}] & \cdots & [\mathbf{0}] \end{bmatrix} \tag{10e}$$

The measurement Eq. (9b) can be linearised and discretised in a similar way to obtain

$$\mathbf{y}(k) = \mathbf{H}_{\mathrm{ex}}\mathbf{z}(k) + \bar{\mathbf{V}} \tag{11a}$$

where $\mathbf{H}_{\mathrm{ex}}$ is called the measurement matrix and has the form

$$\mathbf{H}_{\mathrm{ex}} = \frac{\partial \mathbf{h}}{\partial \mathbf{z}}\Big/ \mathbf{z} = \mathbf{z}(k) \tag{11b}$$

$$= [[\mathbf{H}_0, \quad [\mathbf{0}], \quad [\mathbf{0}]] \quad (\text{if } \mathbf{y}(t) = \mathbf{H}_0\mathbf{q}(t)) \tag{11c}$$

$$= [[\mathbf{0}], \quad \mathbf{H}_0, \quad [\mathbf{0}]] \quad (\text{if } \mathbf{y}(t) = \mathbf{H}_0\dot{\mathbf{q}}(t)) \tag{11d}$$

$$= \left[-\mathbf{H}_0\mathbf{M}^{-1}\mathbf{K}, \quad -\mathbf{H}_0\mathbf{M}^{-1}\mathbf{C}, \quad -\mathbf{H}_0\mathbf{M}^{-1}\frac{\partial \mathbf{K}}{\partial \boldsymbol{\alpha}}\mathbf{q}(k) - \mathbf{H}_0\mathbf{M}^{-1}\frac{\partial \mathbf{C}}{\partial \boldsymbol{\alpha}}\dot{\mathbf{q}}(k) \right] \quad (\text{if } \mathbf{y}(t) = \mathbf{H}_0\ddot{\mathbf{q}}(t)) \tag{11e}$$

and

$$\bar{\mathbf{V}} = \mathbf{h}(\mathbf{z}(k), \mathbf{u}(k), t_k) - \frac{\partial \mathbf{h}}{\partial \mathbf{z}}\mathbf{z}(k) \tag{11f}$$

The preceding first-order Taylor expansions for the fundamental matrix $\mathbf{A}_F$ and measurement matrix $\mathbf{H}_{ex}$ only have to be used in the computation of the Kalman gains. The old estimates that have to be propagated forward do not have to be determined with the fundamental matrix but instead can be propagated directly by integrating the actual nonlinear differential equations (i.e. Eq. (8c)) forward at each sampling interval.

$$\mathbf{z}(k+1) = \varphi(\mathbf{z}(k), \mathbf{u}(k), t_k, t_{k+1}) \tag{12}$$

where $\varphi(\mathbf{z}(k), \mathbf{u}(k), t_k, t_{k+1})$ is realised by applying numerical integration methods to Eq. (8c).

The discrete process noise covariance matrix $\mathbf{Q}_{ex}$, the measurement noise covariance matrix $\mathbf{R}_{ex}$, and the initial estimation error covariance matrix $\mathbf{P}_{ex}(0)$ are still required to start the EKF algorithm. It will be shown how the continuous process noise $\mathbf{w}(t)$ and the discrete measurement noise $\mathbf{v}(t)$ evolve to $\mathbf{Q}_{ex}$ and $\mathbf{R}_{ex}$, respectively, and how $\mathbf{P}_{ex}(0)$ is determined.

Reconsider Eq. (1) and add process noise $\mathbf{w}(t)$ with constant covariance matrix $\mathbf{Q}_s \triangleq E[\mathbf{w}(t)\mathbf{w}^{\mathrm{T}}(t)]$, then Eq. (1) becomes

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{B}\mathbf{u}(t) + \mathbf{L}\mathbf{w}(t) \tag{13}$$

where $\mathbf{L}$ is the process noise matrix. Thus the corresponding Eq. (8c) becomes

$$\dot{\mathbf{z}}(t) = \mathbf{f}(\mathbf{z}(t), \mathbf{u}(t), t) + \mathbf{w}_{exs}(t) \tag{14a}$$

where $\mathbf{w}_{exs}(t)$ is the continuous process noise for EKF

$$\mathbf{w}_{exs}(t) = \begin{bmatrix} [\mathbf{0}] \\ \mathbf{M}^{-1}\mathbf{L}\mathbf{w}(t) \\ [\mathbf{0}] \end{bmatrix} \tag{14b}$$

Directly calculating the covariance matrix of the process noise for the EKF in continuous form (denoted as $\mathbf{Q}_{exs}$) gives

$$\mathbf{Q}_{exs} = E\left[\begin{bmatrix} [\mathbf{0}] \\ \mathbf{M}^{-1}\mathbf{L}\mathbf{w}(t) \\ [\mathbf{0}] \end{bmatrix} [[\mathbf{0}] \quad (\mathbf{M}^{-1}\mathbf{L}\mathbf{w}(t))^{\mathrm{T}} \quad [\mathbf{0}]] \right] \tag{15a}$$

$$= \begin{bmatrix} [\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}] \\ [\mathbf{0}] & \mathbf{M}^{-1}\mathbf{L}\mathbf{Q}_s\mathbf{L}^{\mathrm{T}}(\mathbf{M}^{-1})^{\mathrm{T}} & [\mathbf{0}] \\ [\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}] \end{bmatrix} \tag{15b}$$

where $\mathbf{Q}_s = E[\mathbf{w}(t)\mathbf{w}^{\mathrm{T}}(t)]$, is the continuous covariance matrix of the process noise $\mathbf{w}(t)$.

The discrete process noise covariance matrix thus becomes

$$\mathbf{Q}_{ex}(k) = \int_0^{\Delta t} \mathbf{e}^{\mathbf{A}_f\tau}\mathbf{Q}_{exs}\mathbf{e}^{(\mathbf{A}_f)^{\mathrm{T}}\tau}\,\mathrm{d}\tau \tag{16}$$

Moreover, when the EKF is applied in parameter estimation, the parameter vector $\boldsymbol{\alpha}$ is treated as a random constant vector such as

$$\boldsymbol{\alpha}(k+1) = \boldsymbol{\alpha}(k) + \mathbf{w}_{\boldsymbol{\alpha}}(k) \tag{17}$$

where $\mathbf{w}_{\boldsymbol{\alpha}}(k)$ is any zero-mean Gaussian white noise sequence uncorrelated with the process noise and with pre-assigned positive definite variance $\mathrm{Var}(\mathbf{w}_{\boldsymbol{\alpha}}(k)) = \mathbf{Z}_{\boldsymbol{\alpha}}(k)$. Thus, the discrete covariance matrix of the process noise for extended Kalman filtering is modified from Eq. (16) to be

$$\mathbf{Q}_{ex}(k) = \int_0^{\Delta t} \mathbf{e}^{\mathbf{A}_f\tau}\mathbf{Q}_{exs}\mathbf{e}^{(\mathbf{A}_f)^{\mathrm{T}}\tau}\,\mathrm{d}\tau + \begin{bmatrix} [\mathbf{0}] & [\mathbf{0}] \\ [\mathbf{0}] & \mathbf{Z}_{\boldsymbol{\alpha}}(k) \end{bmatrix} \tag{18}$$

For the measurement noise, since the actual measurements are always sampled in discrete time steps, instead of starting from continuous measurement noise $\mathbf{v}(t)$ and discretise it to $\mathbf{v}(k)$, directly starting from the discrete measurement noise $\mathbf{v}(k)$, with constant covariance matrix $\mathbf{R} \triangleq E[\mathbf{v}(k)\mathbf{v}^{\mathrm{T}}(k)]$, on the measurement, Eq. (9a) becomes

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{G}\mathbf{u}(t) + \mathbf{v}(k) \tag{19a}$$

$$= \mathbf{h}(\mathbf{z}(k), \mathbf{u}(k), k) + \mathbf{v}(k) \tag{19b}$$

From the above equation, the discrete measurement noise covariance matrix for the EKF is

$$\mathbf{R}_{\mathrm{ex}} = \mathbf{R} \tag{20}$$

In the case of the covariance matrix $\mathbf{P}_{\mathrm{ex}}(0)$, which represents the initial errors in the state estimates of the EKF, it takes the form

$$\mathbf{P}_{\mathrm{ex}}(0) = \begin{bmatrix} \mathbf{P}_{xx} & [\mathbf{0}] \\ [\mathbf{0}] & \mathbf{P}_{\alpha\alpha} \end{bmatrix} \tag{21}$$

where $\mathbf{P}_{xx}$ is the covariance matrix related with the estimate of the dynamic state variable (i.e. $\hat{\mathbf{x}}$), $\mathbf{P}_{\alpha\alpha}$ is the part corresponding to the parameters (i.e. $\hat{\boldsymbol{\alpha}}$).

From this premise, all the information to start the EKF algorithm has been obtained.

The EKF equations can be grouped into the three stages:

Propagation:

$$\hat{\mathbf{z}}(k+1/k) = \varphi(\hat{\mathbf{z}}(k), \mathbf{u}(k), t_k, t_{k+1}) \tag{22a}$$

$$\mathbf{P}_{\mathrm{ex}}(k+1/k) = \mathbf{A}_F \mathbf{P}_{\mathrm{ex}}(k) \mathbf{A}_F^{\mathrm{T}} + \mathbf{Q}_{\mathrm{ex}}(k) \tag{22b}$$

Measurement:

$$\hat{\mathbf{y}}(k+1/k) = \mathbf{h}(\hat{\mathbf{z}}(k+1/k), \mathbf{u}(k+1), t_{k+1}) \tag{22c}$$

$$\boldsymbol{\nu}(k+1) = \mathbf{y}(k+1) - \hat{\mathbf{y}}(k+1) \tag{22d}$$

$$\mathbf{S}(k+1) = \mathbf{H}_{\mathrm{ex}} \mathbf{P}_{\mathrm{ex}}(k+1/k) \mathbf{H}_{\mathrm{ex}}^{\mathrm{T}} + \mathbf{R}_{\mathrm{ex}} \tag{22e}$$

Correction:

$$\mathbf{K}_g(k+1) = \mathbf{P}_{\mathrm{ex}}(k+1/k) \mathbf{H}_{\mathrm{ex}}^{\mathrm{T}} \mathbf{S}^{-1}(k+1) \tag{22f}$$

$$\hat{\mathbf{z}}(k+1) = \hat{\mathbf{z}}(k+1/k) + \mathbf{K}_g(k+1) \boldsymbol{\nu}(k+1) \tag{22g}$$

$$\mathbf{P}_{\mathrm{ex}}(k+1) = \mathbf{P}_{\mathrm{ex}}(k+1/k) - \mathbf{K}_g(k+1) \mathbf{S}(k+1) \mathbf{K}_g^{\mathrm{T}}(k+1) \tag{22h}$$

where $\boldsymbol{\nu}(k+1)$ is called the measurement innovation, or the residual. The residual reflects the discrepancy between the predicted measurement and the actual measurement. $\mathbf{S}(k+1)$, defined as $\mathbf{S}(k+1) \triangleq \mathrm{Cov}[\boldsymbol{\nu}(k+1)]$ is the covariance matrix of the residual $\boldsymbol{\nu}(k+1)$.

The recursion starts with the initial values $\hat{\mathbf{z}}(0)$ and the initial covariance matrix $\mathbf{P}_{\mathrm{ex}}(0)$, which represents the initial estimation error. $\mathbf{Q}_{\mathrm{ex}}$ reflects the model error and unknown inputs, while $\mathbf{R}_{\mathrm{ex}}$ describes the measurement noise.

The above traditional EKF method, when applied to the actual system identification, encountered many practical problems:

(1) The order of the EKF based on FE model is $2n + a$ (n: number of DOFs, a: number of unknown parameters) thus it becomes very high even for a very simple structure. The computational effort of the Kalman filter (KF) or EKF is on the order of n^3. An EKF with high order suffers from long computational time and increased computational error, making highly difficult for the filter to get the expected structural parameters.

(2) A large $\mathbf{P}_{\mathrm{ex}}(0)$ can cause the divergence of the EKF. Since normally it is difficult to obtain the variances of the initial parameters, $\mathbf{P}_{\mathrm{ex}}(0)$ is often chosen as a diagonal matrix with very large values. In this case, divergence of the filter can occur, especially for the increasingly complex EKF models.

(3) The initial values of the parameters to be estimated are also important for the EKF algorithm. Since the EKF is derived using a linear model, a local minimum may be reached if the initial parameters are far from the actual ones. Normally, the more accurate the initial values, the better results will be expected from the EKF.

(4) It is difficult to obtain accurate $\mathbf{Q}_{\mathrm{ex}}$ and $\mathbf{R}_{\mathrm{ex}}$. The EKF assumes complete *a-priori* knowledge of the process and measurement noise statistics, $\mathbf{Q}_{\mathrm{ex}}$ and $\mathbf{R}_{\mathrm{ex}}$, respectively. Whilst they are often assumed to be constant matrices, normally it is difficult to give accurate values to these two matrices. On the other hand, the statistics pair ($\mathbf{Q}_{\mathrm{ex}}$ and $\mathbf{R}_{\mathrm{ex}}$) play an important role in the EKF: they determine the bandwidth of the filter and thus affects the convergence and stability of the parameter estimation of the EKF. A large $\mathbf{Q}_{\mathrm{ex}}$ or small $\mathbf{R}_{\mathrm{ex}}$ means a wide bandwidth. In this condition, the filter can follow the state well provided good quantity and quality measurements. But the filter must pay the price for it: by ignoring the model, the parameters (the very thing which is expected from the EKF) become insensitive to the model thus they remain around their original values and reluctant to change. On the other hand, a small $\mathbf{Q}_{\mathrm{ex}}$ or large $\mathbf{R}_{\mathrm{ex}}$ represents a small bandwidth of the filter, which makes the parameters sensitive to the residuals. Although convergence speed is increased in this condition, in the presence of model structure error, the filter may not follow the state thus the estimation of parameters becomes difficult as well because of the coupling relationship between state

and parameters. Based on the discussion above, the pair $(\mathbf{Q}_{ex}, \mathbf{R}_{ex})$ is important for parameter estimation in EKF wh ile accurate values are normally difficult to get.

3. Modified extended Kalman filter

To address the four problems above, the modifications to the EKF are used to make it more suitable for structural identification. These modifications, each addressing one problem, will be presented in the following four sections.

3.1. Modal extended Kalman filter and system order reduction

Modal transformation provides an efficient way to reduce the order of the system and therefore has been widely adopted in extended Kalman filters [8,9].

Instead of using actual displacement, velocity, and acceleration to construct the mechanical system (Eq. (1)), a new variable vector is defined by modal transformation

$$\mathbf{q} \triangleq \mathbf{\Phi p} \tag{23}$$

where $\mathbf{\Phi} \triangleq [\boldsymbol{\varphi}_1 \quad \boldsymbol{\varphi}_2 \quad \cdots \quad \boldsymbol{\varphi}_n]$ is the modal matrix constructed by the mass normalised eigenvectors which satisfy

$$\mathbf{\Phi}^{\mathrm{T}}\mathbf{M}\mathbf{\Phi} = \mathbf{I} \tag{24a}$$

$$\mathbf{\Phi}^{\mathrm{T}}\mathbf{K}\mathbf{\Phi} = \mathbf{\Lambda} \tag{24b}$$

with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1 \ \ \lambda_2 \ \ \cdots \lambda_n)$, where $\lambda_i = \omega_i^2$ is the ith eigenfrequency of the undamped system.

With the help of Eq. (23), Eq. (1) is changed to

$$\ddot{\mathbf{p}} + \mathbf{\Lambda p} + \mathbf{\Gamma}\dot{\mathbf{p}} = \mathbf{\Phi}^{\mathbf{T}}\mathbf{B}\mathbf{u} \tag{25}$$

where $\mathbf{\Gamma} \triangleq \mathbf{\Phi}^{\mathbf{T}}\mathbf{C}\mathbf{\Phi}$. Assume modal damping, thus $\mathbf{\Gamma}$ takes the form of $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1 \ \ \gamma_2 \ \ \cdots \gamma_n)$, where $\gamma_i = 2\zeta_i\omega_i$ is the ith modal damping, and ζ_i is the damping ratio of the ith mode.

The corresponding state space formulation has the same form with Eqs. (2a) and (2b), for simplicity recapitulate here

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_e\mathbf{u}(t) \tag{26a}$$

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) \tag{26b}$$

However, $\mathbf{A}$,$\mathbf{B}_e$ and $\mathbf{x}$ take the modal form

$$\mathbf{A} = \begin{bmatrix} [\mathbf{0}] & \mathbf{I} \\ -\mathbf{\Lambda} & -\mathbf{\Gamma} \end{bmatrix}, \quad \mathbf{B}_e = \begin{bmatrix} [\mathbf{0}] \\ \mathbf{\Phi}^{\mathrm{T}}\mathbf{B} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbf{p}(t) \\ \dot{\mathbf{p}}(t) \end{bmatrix} \tag{27}$$

The corresponding EKF state space formulation becomes

$$\dot{\mathbf{z}}(t) = \begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{p}}(t) \\ \ddot{\mathbf{p}}(t) \\ \dot{\boldsymbol{\alpha}} \end{bmatrix} \tag{28a}$$

$$= \begin{bmatrix} \dot{\mathbf{p}}(t) \\ -\mathbf{\Lambda p}(t) - \mathbf{\Gamma}\dot{\mathbf{p}}(t) + \mathbf{\Phi}^{\mathrm{T}}\mathbf{B}\mathbf{u}(t) \\ [\mathbf{0}] \end{bmatrix} \tag{28b}$$

$$\triangleq \mathbf{f}(\mathbf{z}(t), \mathbf{u}(t), t) \tag{28c}$$

$$\mathbf{y}(t) = \mathbf{H}\mathbf{z}(t) + \mathbf{G}\mathbf{u}(t) \tag{29a}$$

$$\triangleq \mathbf{h}(\mathbf{z}(t), \mathbf{u}(t), t) \tag{29b}$$

The fundamental matrix of the EKF is obtained by discretising and linearising Eq. (28c) at time t_k:

$$\mathbf{A}_F = \mathrm{e}^{\mathbf{A}_f\Delta t} = \mathbf{I} + \mathbf{A}_f\Delta t + (\mathbf{A}_f\Delta t)^2/2! + \ldots (\Delta t = t_{k+1} - t_k) \tag{30a}$$

where

$$\mathbf{A}_f = \frac{\partial \mathbf{f}}{\partial \mathbf{z}} \Big/ \mathbf{z} = \mathbf{z}(k) = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \frac{\partial \mathbf{f}}{\partial \dot{\mathbf{p}}} \frac{\partial \mathbf{f}}{\partial \boldsymbol{\alpha}}\right] \tag{30b}$$

$$= \begin{bmatrix} [\mathbf{0}] & \mathbf{I} & [\mathbf{0}] & \cdots & [\mathbf{0}] \\ -\boldsymbol{\Lambda} & -\boldsymbol{\Gamma} & -\dfrac{\partial \boldsymbol{\Lambda}}{\partial \alpha_1}\mathbf{p} - \dfrac{\partial \boldsymbol{\Gamma}}{\partial \alpha_1}\dot{\mathbf{p}} + \left(\dfrac{\partial \boldsymbol{\Phi}}{\partial \alpha_1}\right)^{\mathrm{T}} \mathbf{Bu} & \cdots & -\dfrac{\partial \boldsymbol{\Lambda}}{\partial \alpha_a}\mathbf{p} - \dfrac{\partial \boldsymbol{\Gamma}}{\partial \alpha_a}\dot{\mathbf{p}} + \left(\dfrac{\partial \boldsymbol{\Phi}}{\partial \alpha_a}\right)^{\mathrm{T}} \mathbf{Bu} \\ [\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}] & \cdots & [\mathbf{0}] \end{bmatrix} \tag{30c}$$

where

1. $\partial\boldsymbol{\Lambda}/\partial\alpha_j$ is the sensitivity of the eigenvalue matrix to the parameter α_j:

$$\frac{\partial \boldsymbol{\Lambda}}{\partial \alpha_j} = \operatorname{diag}\left(\frac{\partial \lambda_1}{\partial \alpha_j} \quad \frac{\partial \lambda_2}{\partial \alpha_j} \quad \cdots \quad \frac{\partial \lambda_n}{\partial \alpha_j}\right) \tag{31a}$$

The ith diagonal entry is the rate of the ith system eigenvalue to the jth updating parameter

$$\frac{\partial \lambda_i}{\partial \alpha_j} = \boldsymbol{\Phi}_i^{\mathrm{T}} \left(\frac{\partial \mathbf{K}}{\partial \alpha_j} - \lambda_i \frac{\partial \mathbf{M}}{\partial \alpha_j}\right) \boldsymbol{\Phi}_i \tag{31b}$$

2. $\partial\boldsymbol{\Phi}/\partial\alpha_j$ is the sensitivity of the eigenvector matrix to the parameter α_j:

$$\frac{\partial \boldsymbol{\Phi}}{\partial \alpha_j} = \left[\frac{\partial \boldsymbol{\varphi}_1}{\partial \alpha_j} \quad \frac{\partial \boldsymbol{\varphi}_2}{\partial \alpha_j} \quad \cdots \quad \frac{\partial \boldsymbol{\varphi}_n}{\partial \alpha_j}\right] \tag{32a}$$

The ith term in the matrix is the rate of the ith system eigenvector to the jth updating parameter, which was determined by [10]:

$$\frac{\partial \boldsymbol{\varphi}_i}{\partial \alpha_j} = \boldsymbol{\Phi}\mathbf{C}_{ij} = [\boldsymbol{\varphi}_1 \ \ \boldsymbol{\varphi}_2 \ \ \cdots \ \ \boldsymbol{\varphi}_n] \begin{bmatrix} \mathbf{C}_{ij}(1) \\ \mathbf{C}_{ij}(2) \\ \vdots \\ \mathbf{C}_{ij}(n) \end{bmatrix} \tag{32b}$$

where the kth coefficient $\mathbf{C}_{ij}(k)$ is

$$\mathbf{C}_{ij}(k) = \frac{-\boldsymbol{\varphi}_k^{\mathrm{T}} \left(\dfrac{\partial \mathbf{K}}{\partial \alpha_j} - \lambda_i \dfrac{\partial \mathbf{M}}{\partial \alpha_j}\right) \boldsymbol{\varphi}_i}{\lambda_k - \lambda_i} \quad (k \neq i) \tag{32c}$$

$$\mathbf{C}_{ij}(k) = -\frac{1}{2}\boldsymbol{\phi}_k^{\mathrm{T}} \frac{\partial \mathbf{M}}{\partial \alpha_j} \boldsymbol{\varphi}_i \quad (k = i) \tag{32d}$$

3. $\partial\boldsymbol{\Gamma}/\partial\alpha_j$ is the sensitivity of the damping matrix to the parameter α_j, for a uniform damping, $\boldsymbol{\Gamma} = 2\zeta \cdot \operatorname{diag}(\omega_1 \ \ \omega_2 \ \ \cdots \ \ \omega_n) = 2\zeta\sqrt{\boldsymbol{\Lambda}}$, thus

$$\frac{\partial \boldsymbol{\Gamma}}{\partial \alpha_j} = 2\frac{\partial \zeta}{\partial \alpha_j}\sqrt{\boldsymbol{\Lambda}} + \zeta \boldsymbol{\Lambda}^{-1/2} \frac{\partial \boldsymbol{\Lambda}}{\partial \alpha_j} \tag{33a}$$

Normally, when ζ is one of the parameters to be estimated, then

$$\frac{\partial \boldsymbol{\Gamma}}{\partial \alpha_j} = 2\sqrt{\boldsymbol{\Lambda}} \tag{33b}$$

For other parameters except ζ

$$\frac{\partial \boldsymbol{\Gamma}}{\partial \alpha_j} = \zeta \boldsymbol{\Lambda}^{-1/2} \frac{\partial \boldsymbol{\Lambda}}{\partial \alpha_j} \tag{33c}$$

Likewise, the measurement matrix $\mathbf{H}_{\mathrm{ex}}$ in the EKF is

$$\mathbf{H}_{\mathrm{ex}} = \frac{\partial \mathbf{h}}{\partial \mathbf{z}} \Big/ \mathbf{z} = \mathbf{z}(k) \tag{34a}$$

$$= \left[\mathbf{H}_0\boldsymbol{\Phi}, \quad [\mathbf{0}], \quad \mathbf{H}_0 \frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\alpha}}\mathbf{p}(k)\right] \quad (\text{if } \mathbf{y}(t) = \mathbf{H}_0\mathbf{q}(t)) \tag{34b}$$

$$= \left[[\mathbf{0}], \quad \mathbf{H}_0\boldsymbol{\Phi}, \quad \mathbf{H}_0 \frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\alpha}}\dot{\mathbf{p}}(k)\right] \quad (\text{if } \mathbf{y}(t) = \mathbf{H}_0\dot{\mathbf{q}}(t)) \tag{34c}$$

$\mathbf{Q}_{\mathrm{ex}}(k)$ and $\mathbf{R}_{\mathrm{ex}}$ are redefined as

$$\mathbf{Q}_{\mathrm{ex}}(k)=\int_0^{\Delta t}\mathbf{e}^{\mathbf{A}_f\tau}\mathbf{Q}_{\mathrm{exs}}\mathbf{e}^{(\mathbf{A}_f)^{\mathrm{T}}\tau}\,\mathrm{d}\tau+\begin{bmatrix}[\mathbf{0}] & [\mathbf{0}]\\ [\mathbf{0}] & \mathbf{S}(k)\end{bmatrix} \tag{35a}$$

where

$$\mathbf{Q}_{\mathrm{exs}}=\begin{bmatrix}[\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}]\\ [\mathbf{0}] & \mathbf{\Phi}^{\mathrm{T}}\mathbf{L}\mathbf{Q}_s\mathbf{L}^{\mathrm{T}}\mathbf{\Phi} & [\mathbf{0}]\\ [\mathbf{0}] & [\mathbf{0}] & [\mathbf{0}]\end{bmatrix} \tag{35b}$$

and

$$\mathbf{R}_{\mathrm{ex}}=\mathbf{R} \tag{36}$$

$\mathbf{P}_{\mathrm{ex}}(0)$ has the same form as defined in Eq. (21). However, since the dynamic part of state vector (i.e. $\mathbf{x}$) has changed to be in modal coordinates, the current $\mathbf{P}_{xx}$ (denoted as $\mathbf{P}_{xx}^{\mathrm{mod}}$) has the following relationship with the previous one (i.e. $\mathbf{P}_{xx}$ in normal coordinate, denoted as $\mathbf{P}_{xx}^{\mathrm{nor}}$)

$$\mathbf{P}_{xx}^{\mathrm{nor}}=\mathbf{\Phi}\mathbf{P}_{xx}^{\mathrm{mod}}\mathbf{\Phi}^{\mathrm{T}} \tag{37}$$

The modal EKF has more flexibility and thus can be used in model reduction.

For example, if only the first r modes are excited, then $(p_{r+1}\ \ p_{r+2}\ \ \cdots\ \ p_n)$, $(\dot{p}_{r+1}\ \ \dot{p}_{r+2}\ \ \cdots\ \ \dot{p}_n)$ and $(\ddot{p}_{r+1}\ \ \ddot{p}_{r+2}\ \ \cdots\ \ \ddot{p}_n)$ can be neglected, without losing any generality, here assuming displacement measurements, it results:

$$\mathbf{y}=\mathbf{H}_0\mathbf{\Phi}\mathbf{p}=\mathbf{H}_0[\boldsymbol{\varphi}_1\ \ \cdots\ \ \boldsymbol{\varphi}_r\ \ \cdots\boldsymbol{\varphi}_n]\begin{bmatrix}p_1\\ \vdots\\ p_r\\ \vdots\\ p_n\end{bmatrix} \tag{38a}$$

$$=\mathbf{H}_0[\boldsymbol{\varphi}_1\ \ \cdots\ \ \boldsymbol{\varphi}_r\begin{bmatrix}p_1\\ \vdots\\ p_r\end{bmatrix} \tag{38b}$$

The original Eq. (28) and (29) can be written as reduced EKF state space formulations:

$$\dot{\mathbf{z}}_r(t)=\begin{bmatrix}\dot{\mathbf{x}}_r(t)\\ \dot{\boldsymbol{\alpha}}\end{bmatrix}=\begin{bmatrix}\dot{\mathbf{p}}_r(t)\\ \ddot{\mathbf{p}}_r(t)\\ \dot{\boldsymbol{\alpha}}\end{bmatrix} \tag{39a}$$

$$=\begin{bmatrix}\dot{\mathbf{p}}_r(t)\\ -\mathbf{\Lambda}_r\mathbf{p}_r(t)-\mathbf{\Gamma}_r\dot{\mathbf{p}}_r(t)+\mathbf{\Phi}_r^{\mathrm{T}}\mathbf{B}\mathbf{u}(t)\\ [\mathbf{0}]\end{bmatrix} \tag{39b}$$

$$\triangleq\mathbf{f}_r(\mathbf{z}_r(t),\mathbf{u}(t),t) \tag{39c}$$

$$\mathbf{y}(t)=\mathbf{H}_{\mathrm{exr}}\mathbf{z}_r(t) \tag{40}$$

where

$$\mathbf{p}_r\triangleq[p_1\ \ p_2\ \ \cdots\ \ p_r]^{\mathrm{T}} \tag{41a}$$

$$\mathbf{\Lambda}_r=\mathrm{diag}(\omega_1^2\ \ \omega_2^2\ \ \cdots\ \ \omega_r^2) \tag{41b}$$

$$\mathbf{\Gamma}_r=\mathrm{diag}(\gamma_1\ \ \gamma_2\ \ \cdots\ \ \gamma_r) \tag{41c}$$

$$\mathbf{\Phi}_r\triangleq[\boldsymbol{\varphi}_1\ \ \boldsymbol{\varphi}_2\ \ \cdots\ \ \boldsymbol{\varphi}_r] \tag{41d}$$

$$\mathbf{H}_{\mathrm{exr}}=[\mathbf{H}_0\mathbf{\Phi}_r\ \ [\mathbf{0}]\ \ [\mathbf{0}]] \tag{41e}$$

The fundamental matrix, the measurement matrix and related covariance matrices are determined in the similar way.

Now the order of the system has decreased from $2n+a$ (n: number of degree of freedom of the system, a: number of parameters to be estimated) to $2r+a$ (r: modes included in the model). Thus the computation time is highly decreased.

3.2. Multiple model EKF based on different initial estimation covariance $\mathbf{P}_{ex}(0)$

The idea of multiple model EKF came from the 'multiple model adaptive estimator' [11]. This scheme consists of a bank of parallel Kalman filters, each with a different model, and a hypothesis testing algorithm as shown in Fig. 1.

The internal models in the Kalman filters can be represented by discrete values in a parameter vector ($\boldsymbol{\alpha}^i, i = 1, 2 \ldots N$). The Kalman filters are provided with a measurement vector ($\mathbf{y}$) and the input vector ($\mathbf{u}$), and produce a state estimate ($\hat{\mathbf{x}}^i$) and a residual ($\boldsymbol{\nu}^i$). Here the superscript i corresponds to the ith EKF in the bank. The hypothesis testing algorithm uses the residuals to compute conditional probabilities (Pr^i) of the various hypotheses that are modelled in the Kalman filters.

At each recursive step the adaptive filter carries out three tasks:

1. each filter in the bank of filters computes its own estimate, which is based on its own model;
2. the system computes the posterior probabilities for each of the hypotheses; and
3. the scheme forms the adaptive optimal estimate of state variable vector as a weighted sum of the estimates produced by each of the individual Kalman filters.

As measurements evolve with time, the adaptive scheme learns which of the filters is the correct one, and its weight factor approaches unity while the others converge to zero. The filter who 'wins' is the correct model.

Although MMAE finds many applications, for example, in flight control failure [12,13] etc., some drawbacks must be overcome before it can be applied to mechanical structure damage detection.

1. It is assumed that the model which could describe the actual system is contained in the bank of Kalman filters. Although this is possible when the system has been well-defined and damage always occurs in a few certain locations, it is difficult to include the true model for a mechanical structure where damage could happen (even simultaneously) in many places. In this case, the number of Kalman filters in the bank is unmanageably large.
2. The hypothesis testing algorithm determines that only one filter will be finally selected as the 'correct' one. Thus the information from other filters in the bank is totally discarded. From the optimal point of view, although not as good as the optimal filter, the other filters still contain information (more or less) that would be useful for the estimation result.

Based on the discussion above, the MMAE method has been revised to be applied to the damage detection of mechanical structures. Fritzen proposed a scheme of MMAE to identify the cracks in a structure [9]. In his method, each of the filters represents a model with one special location of crack. The depth of the crack is an additional free parameter which also has to be identified. The proposed method, in essence, helps to improve the EKF estimation results by decreasing the number of parameters to be identified (in this case, only one parameter for each EKF). However, when two cracks happen simultaneously, the number of EKF in the bank must be increased.

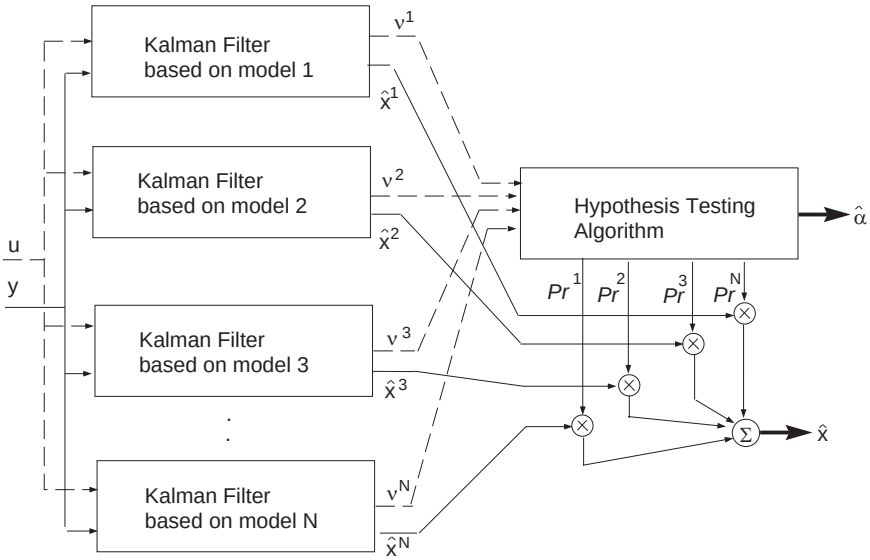

Fig. 1. Traditional MMAE method.

As was mentioned before, $\mathbf{P}_{ex}(0)$ is the covariance matrix representing initial errors of the state estimate vector $\hat{\mathbf{z}}(t)$

$$\left(\hat{\mathbf{z}}(t) = \begin{bmatrix} \hat{\mathbf{x}}(t) \\ \hat{\boldsymbol{\alpha}}(t) \end{bmatrix}\right).$$

$\mathbf{P}_{ex}(0)$, especially the lower part corresponding to the parameters (i.e. $\mathbf{P}_{\alpha\alpha}(0)$), plays a crucial role in the parameter estimation. A large $\mathbf{P}_{ex}(0)$ normally accompanies higher speed of estimation convergence, while at the price of lower stability. A small $\mathbf{P}_{ex}(0)$ increases the stability of convergence while convergence becomes very slow. Since $\mathbf{P}_{ex}(0)$ is important for the estimation while difficult to obtain, the $\mathbf{P}_{ex}(0)$-based EKF MMAE scheme takes the form as shown in Fig. 2.

Each filter in the bank is a modal-based EKF, but starts from different $\mathbf{P}_{ex}(0)$. The estimate result (i.e. $\hat{\boldsymbol{\alpha}}^i$) from each EKF is used to construct the following Kalman filter which only estimates $\mathbf{x}(t)$. The residuals of each KF is used to evaluate the model of KF itself, (i.e. the validity of $\hat{\boldsymbol{\alpha}}^i$). Thus, the information of all the Kalman filters in the bank was incorporated based on some criteria, and the optimal $\hat{\boldsymbol{\alpha}}^{opt}$ is given as the final estimation result.

The hypotheses testing algorithm in the scheme is based on the statistical tests for filter consistency, which is described by the following criteria [14]:

1. the state errors should be acceptable as zero mean and have magnitude commensurate with the state covariance as yielded by the filter;
2. the residuals should also have the same property; and
3. the residuals should be acceptable as white.

The last two criteria are the only ones that can be tested in real data applications, and can be further quantified by the following two criteria:

1. under the hypothesis that the filter is consistent, the time-averaged normalised innovation squared (TANIS)

$$\hat{\varepsilon}_v = \frac{1}{K}\sum_{k=1}^{K} \mathbf{v}^{\mathrm{T}}(k)\mathbf{S}(k)\mathbf{v}(k) \tag{42}$$

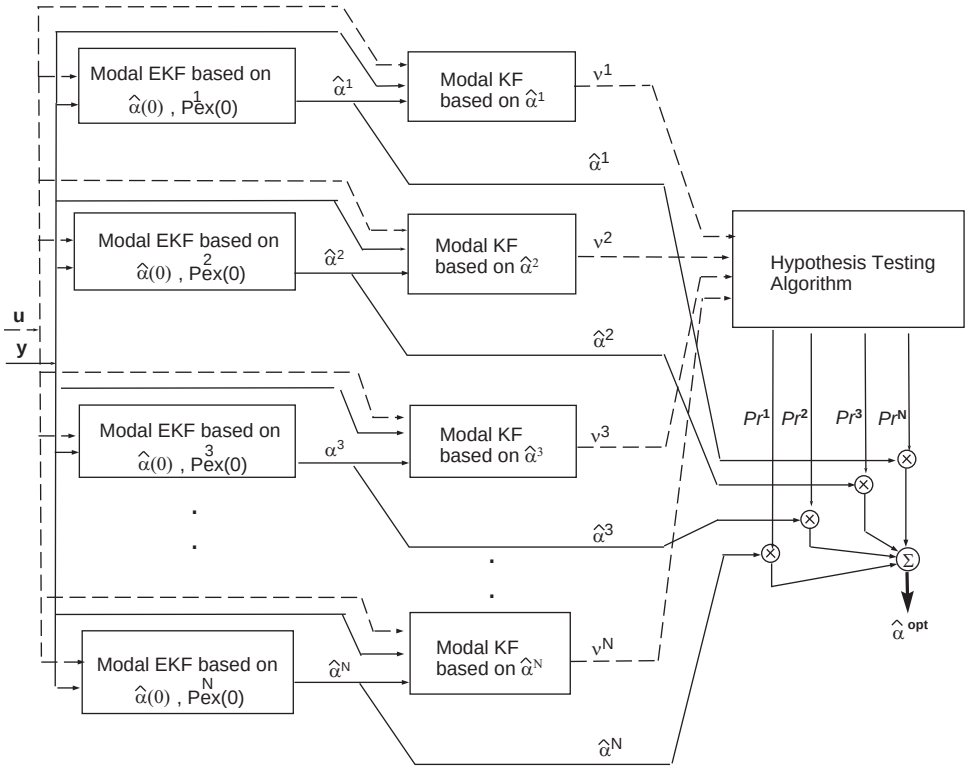

Fig. 2. Scheme of a modified multiple model MOKF.

has a chi-square distribution with n_z DOF, where $\mathbf{v}(k)$ is the residual of the Kalman filter, $\mathbf{S}(k)$ is the corresponding covariance matrix, n_z is the dimension of the measurement.

1. the whiteness test statistic for residuals l steps apart from a single run can be written as time-average autocorrelation (TAAU)

$$\hat{\rho}(l) = \frac{\sum_{k=1}^{K} \mathbf{v}^{\mathrm{T}}(k)\mathbf{v}(k+l)}{\sqrt{\sum_{k=1}^{K} \mathbf{v}^{\mathrm{T}}(k)\mathbf{v}(k) \sum_{k=1}^{K} \mathbf{v}^{\mathrm{T}}(k+l)\mathbf{v}(k+l)}} \tag{43}$$

This statistics is, for large enough K, in view of the central limit theorem, normally distributed. Furthermore, it has zero mean and covariance $1/K$.

Here, the first-order statistic properties (i.e. mean) of random variable $\hat{\varepsilon}_v$ and $\hat{\rho}(l)$ are tested. Since $\hat{\varepsilon}_v$ is a chi-square distribution with n_z DOF, the mean of $\hat{\varepsilon}_v$ is n_z; $\hat{\rho}(l)$ is a normally distributed variable with zero mean, thus the mean of $\hat{\rho}(l)$ is zero.

Based on the discussion above, the probability of the ith Kalman filter is defined by

$$Pr^i = \frac{1}{J^i} \Big/ \sum_{j=1}^{N} \frac{1}{J^j} \tag{44a}$$

where N is the number of filters in the bank, J^i is the cost function of the ith Kalman filter with the form

$$J^i = (\hat{\varepsilon}_v^i - n_z)^2 + w_f(\hat{\rho}^i(l))^2 \tag{44b}$$

where $\hat{\varepsilon}_v^i$, $\hat{\rho}^i$ are the statistic properties from the ith Kalman filter determined according to Eq. (42) and Eq. (43), w_f is a weighting factor which balances the two criteria.

Subsequently, the optimal parameters estimated by the whole bank of EKF are determined by

$$\hat{\boldsymbol{\alpha}}_{\mathrm{opt}} = \sum_{i=1}^{N} Pr^i \hat{\boldsymbol{\alpha}}^i \tag{45}$$

3.3. Weighted global iterative method in MOKF

The weighted global iterative method for the extended Kalman filter (EK-WGI) was first proposed by Hoshiya and Saito [15,16], which comprises the following steps [17] and is shown in Fig. 3:

1. Supply the *a-priori* estimates $\hat{\mathbf{z}}(0)$, $\mathbf{P}_{\mathrm{ex}}(0)$.
2. Use the extended Kalman filter and find at the end of n time steps, the updated state vector, $\hat{\mathbf{z}}(n/n)$, $\mathbf{P}_{\mathrm{ex}}(n/n)$.
3. Weight the final updated covariance from step 2 and use them as new *a-priori* estimates for iteration. For instance, after i iterations, the i+1 iteration will start with the new *a-priori* estimates as follows:

$$\hat{\mathbf{z}}^{i+1}(0) = \begin{bmatrix} \hat{\mathbf{x}}^0(0) \\ \hat{\boldsymbol{\alpha}}^i(n/n) \end{bmatrix} \tag{46a}$$

$$\mathbf{P}_{\mathrm{ex}}^{i+1}(0) = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}}^0(0) & \mathbf{P}_{\mathrm{x}\alpha}(0) = [\mathbf{0}] \\ \mathbf{P}_{\alpha\mathrm{x}}(0) = [\mathbf{0}] & W \cdot \mathbf{P}_{\alpha\alpha}^i(n/n) \end{bmatrix} \tag{46b}$$

where the superscripts 0, i and i+1 corresponds to the iteration number. In other words, only the state and error covariance associated with parameters to be identified are updated.

4. Repeat until the parameters can no longer be improved or a local minimum identification error has been achieved.

The method described above can also be used in different sets of data in a consecutive way, which is shown in Fig. 4.

The EK-WGI method, in essence, is based on the following two facts:

1. The physical parameters to be estimated by the EKF are stationary. The parameters are normally improved after one iteration, thus using the parameters from the previous iteration to start a new one will hopefully further improve the estimation result even using the same set of I/O data.
2. The weight factor W is used to increase $\mathbf{P}_{\mathrm{ex}}(0)$ thus to accelerate the convergence at the beginning of the new iteration. There are so far no guidelines on the choice of W except that it should be greater than 1. In their application, Hoshiya and Saito used $W = 100$.

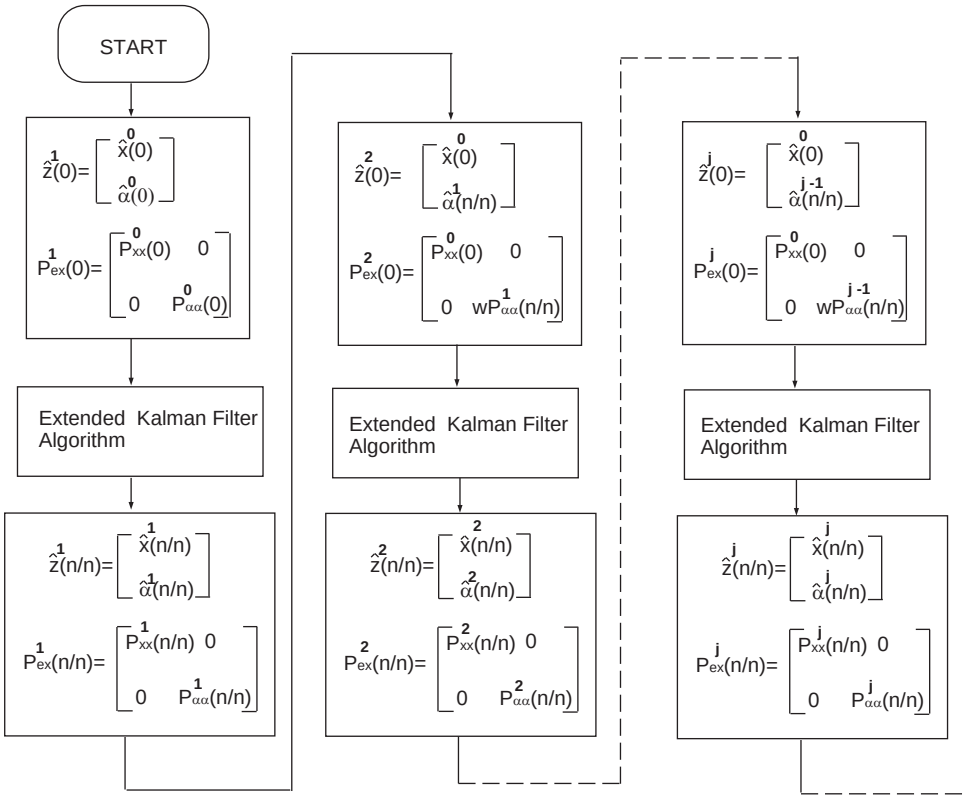

Fig. 3. Local j time iteration on first set of data.

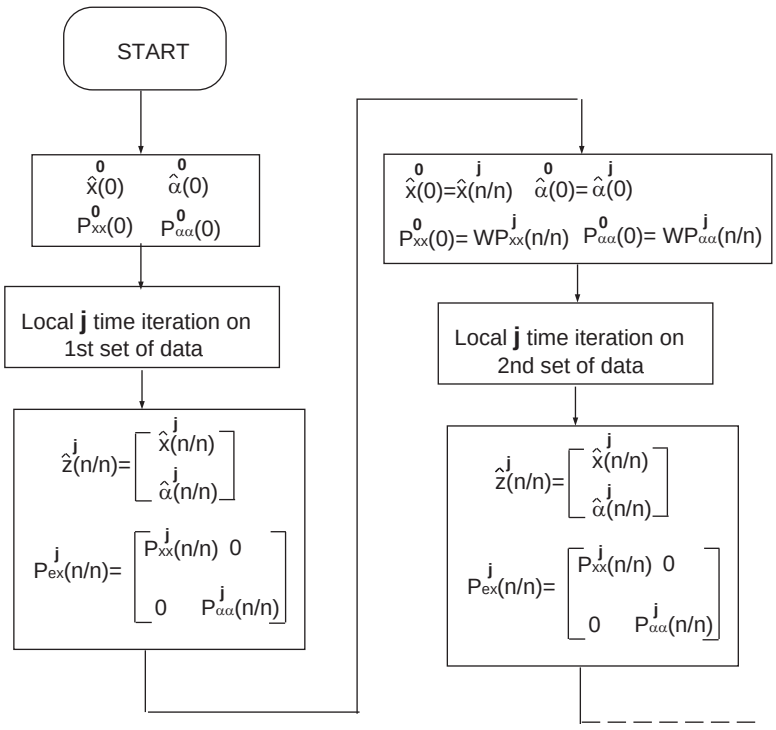

Fig. 4. Weighted global method using different sets of data.

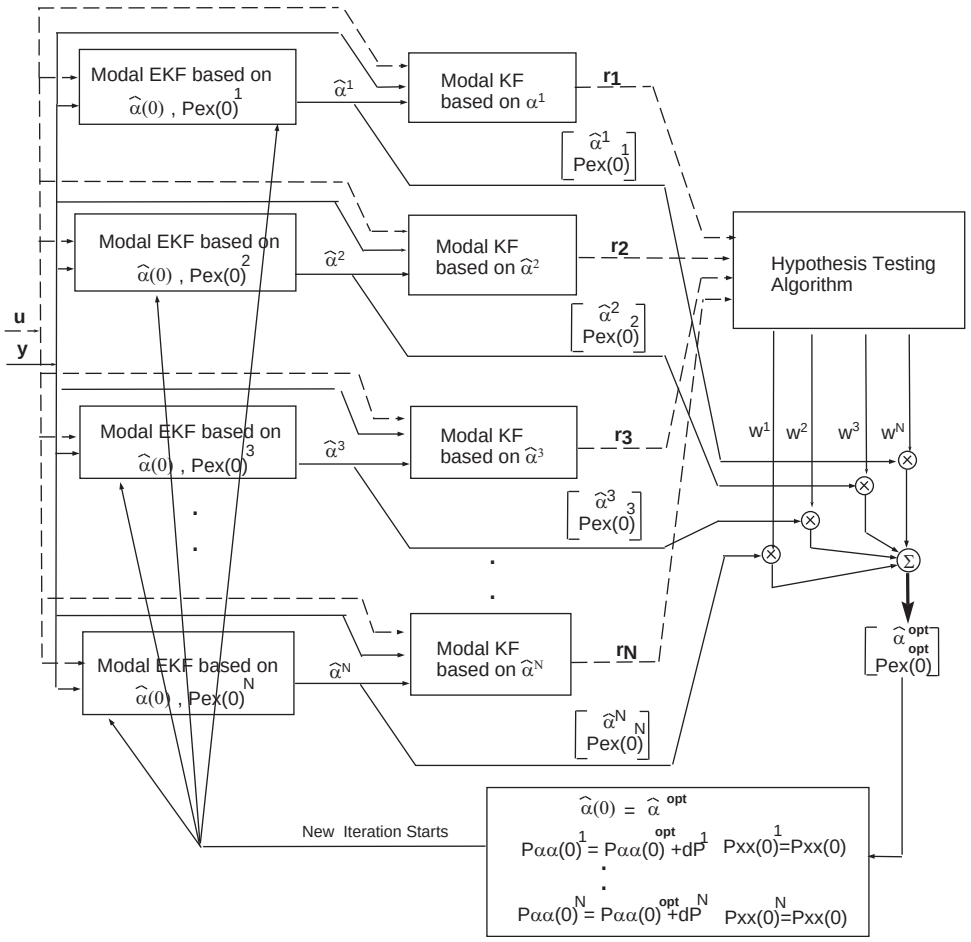

Fig. 5. Revised MMAE to incorporate EK-WGI.

The multiple model MOKF scheme described in Section 3.2 is revised to incorporate EK-WGI method, as shown in Fig. 5.

Here the superscript i corresponds to the ith MOKF in the bank. Not only the parameters $\hat{\boldsymbol{\alpha}}^i$ estimated from each filter in the bank, but also the initial error covariance matrix $\mathbf{P}^i_{\mathrm{ex}}(0)$ are incorporated at the end of one iteration. The optimal $\hat{\boldsymbol{\alpha}}^{\mathrm{opt}}$ and $\mathbf{P}^{\mathrm{opt}}_{\mathrm{ex}}(0)$ are determined by

$$\hat{\boldsymbol{\alpha}}^{\mathrm{opt}} = \sum_{i=1}^{N} Pr^i \hat{\alpha}^i \tag{47a}$$

$$\mathbf{P}^{\mathrm{opt}}_{\mathrm{ex}}(0) = \sum_{i=1}^{N} Pr^i \mathbf{P}^i_{\mathrm{ex}}(0) \tag{47b}$$

Before the new iteration starts, $\hat{\boldsymbol{\alpha}}^i(0)$ and $\mathbf{P}^i_{\mathrm{ex}}(0)$ are assigned to each filters in the bank according to

$$\hat{\boldsymbol{\alpha}}^i(0) = \hat{\boldsymbol{\alpha}}^{\mathrm{opt}} \tag{48a}$$

$$\mathbf{P}^i_{\mathrm{ex}}(0) = \begin{bmatrix} \mathbf{P}_{xx}(0) & [\mathbf{0}] \\ [\mathbf{0}] & \mathbf{P}^{\mathrm{opt}}_{\alpha\alpha}(0) + \boldsymbol{\delta}\mathbf{P}^i \end{bmatrix} \tag{48b}$$

where $\delta\mathbf{P}^i$ $(i = 1,\ldots,N)$ normally takes the form

$$\delta\mathbf{P}^i = \frac{2}{N-1}\left(i - \frac{N+1}{2}\right) P_R \tag{48c}$$

where P_R is the range that the bank of EKF can cover.

For example, for the first filter in the bank ($i=1$), $\delta P^i=-P_R$; for the last filter in the bank ($i=N$), $\delta P^i=P_R$; for the filter in the middle and assume an odd N, ($i=(N+1)/2$), $\delta P^i=[0]$). The others δP^i evenly separated between $-P_R$ and P_R. It is obvious that, according to Eq. (48c), in each new iteration, the estimate error covariance $\mathbf{P}_{\mathrm{ex}}(0)$ of all the filters in the bank will be evenly separated with centre of optimal $\mathbf{P}_{\alpha\alpha}^{\mathrm{opt}}(0)$ from pervious iteration and the variation of P_R.

One remaining issue which should be noticed is the iteration number. By iterating, it is being implicitly assumed that the measured data are trustworthy. In the case when the measurement data contain system error or large noise, the identified parameters will inevitably deviate from their true values and moreover, divergence may occur after several iterations. Therefore, it would be suggested by the authors that filter consistency should be checked in each iteration and once the KF in current iteration is found to be less consistent than the previous one (according to Eqs. (42) and (43)), the whole process should be stopped and the identification result from the previous iteration is adopted.

So far, some solutions to the first two questions proposed at the end of Section 2.2 are provided, the last one will be answered in the next section.

3.4. Fuzzy logic-based adaptive Kalman Filter in mechanical structures

In the context of the standard EKF algorithm the measurement noise covariance matrix $\mathbf{R}(k)$ represents the accuracy of the measurement instrument (note: since $\mathbf{R}_{\mathrm{ex}}=\mathbf{R}$, $\mathbf{R}$ is used in this section for simplicity). Thus, the enlargement of the covariance matrix $\mathbf{R}(k)$ for measured data means that less trust is put on this measured data and more faith is put on the prediction. Assuming that the process noise covariance matrix $\mathbf{Q}_{\mathrm{ex}}(k)$ is completely known, in this section an algorithm employing the principles of fuzzy logic is derived to adaptively adjust the matrix $\mathbf{R}(k)$. In the Kalman filter literature, there are reported some approaches to do this task, e.g. [18]. But in this work, the approach described in this section is adopted. For simplicity, it is assumed that the measurement noise covariance matrix $\mathbf{R}(k)$ is restricted to be a diagonal matrix whose elements represent the variances of the individual components of the measurement noise vector $\mathbf{v}(k)$, this is

$$\mathbf{R}(k)=\mathrm{diag}(\mathbf{R}(k)_{(1,1)}\ \ \mathbf{R}(k)_{(2,2)}\ \ \cdots\ \ \mathbf{R}(k)_{(m,m)}) \tag{49}$$

subject to $\mathbf{R}(k)_{(i,i)}>0$ ($i=1,\ldots,m$). Therefore, the tuning of $\mathbf{R}(k)$ reduces to tuning its m diagonal elements (recall that m is the dimension of the measurement vector). In the fuzzy logic-based adaptive Kalman filter (FL-AKF) this is achieved as follows. For an optimal filter, the residual sequence $\mathbf{v}(k)$ is a combination of independent Gaussian random variables. As a result, the residual sequence also is a Gaussian white random sequence with zero mean and theoretic covariance from Eq. (50):

$$\mathbf{S}(k+1)=\mathbf{H}_{\mathrm{ex}}\mathbf{P}_{\mathrm{ex}}(k+1/k)\mathbf{H}_{\mathrm{ex}}^{\mathrm{T}}+\mathbf{R}(k) \tag{50}$$

Therefore, if it is noted that the theoretical residual covariance $\mathbf{S}(k)$ has discrepancies with its actual value $\hat{\mathbf{C}}_r(k)$ which is defined by

$$\hat{\mathbf{C}}_r(k)=\frac{1}{WS}\sum_{i=i_0}^{k}\mathbf{v}(i)\mathbf{v}(i)^{\mathrm{T}} \tag{51}$$

where $i_0=k-WS+1$ is the first sample inside a sliding window whose size WS is chosen empirically to give some statistical smoothing, then a fuzzy inference system (FIS) derives adjustments for $\mathbf{R}(k)$ based on the knowledge of the size of this discrepancy. The objective of these adjustments is to correct this mismatch as much as possible and, in this way, maintain the consistency between the theoretical and actual residual statistics.

In order to detect and monitor the size of the discrepancy between $\mathbf{S}(k)$ and its actual value $\hat{\mathbf{C}}_r(k)$, a new variable called the degree of mismatch (referred to as $\mathbf{DoM}(k)$) is defined as:

$$\mathbf{DoM}(k)=\mathbf{S}(k)-\hat{\mathbf{C}}_r(k) \tag{52}$$

In qualitative terms, the logic of the adaptation algorithm implemented by a FIS can be described as follows. From Eq. (50) it is clear that the value of $\mathbf{S}(k)$ depends directly on the value of $\mathbf{R}(k)$. This means that an increment in $\mathbf{R}(k)$ will increment $\mathbf{S}(k)$ and vice versa. Thus, $\mathbf{R}(k)$ can be used to vary $\mathbf{S}(k)$ in accordance with the value of $\mathbf{DoM}(k)$ in order to reduce the discrepancies between $\mathbf{S}(k)$ and $\hat{\mathbf{C}}_r(k)$. Note that all matrices $\mathbf{S}(k)$, $\hat{\mathbf{C}}_r(k)$, $\mathbf{R}(k)$ and $\mathbf{DoM}(k)$ have the same dimension $m\times m$. Therefore, the diagonal elements of $\mathbf{R}(k)$ can be adapted in accordance with the diagonal elements of $\mathbf{DoM}(k)$. This is, if the actual residual covariance value $\hat{\mathbf{C}}_r(k)_{(i,i)}$ is observed, whose value is within the range predicted by theory $\mathbf{S}(k)_{(i,i)}$, then $\mathbf{DoM}(k)_{(i,i)}$ will be near to zero indicating that both covariance values match almost perfectly, and then no changes to $\mathbf{R}(k)_{(i,i)}$ are needed. If the actual residual covariance $\hat{\mathbf{C}}_r(k)_{(i,i)}$ is less than its predicted theoretic value $\mathbf{S}(k)_{(i,i)}$, then $\mathbf{DoM}(k)_{(i,i)}$ will be greater than zero, and then $\mathbf{R}(k)_{(i,i)}$ needs to be decreased to compensate this positive mismatch. Conversely, if the actual residual covariance $\hat{\mathbf{C}}_r(k)_{(i,i)}$ is greater than its predicted theoretic value $\mathbf{S}(k)_{(i,i)}$, then $\mathbf{DoM}(k)_{(i,i)}$ will be less than zero, and then $\mathbf{R}(k)_{(i,i)}$ needs to be increased to compensate this negative mismatch. The previous adjusting mechanism lends itself very well to being dealt with using a fuzzy logic approach based on rules of the kind 'IF x is A_j THEN y is B_j', where x and y are linguistic variables in the universes of discourse U and V; A_j and B_j are linguistic values (fuzzy sets) of the linguistic variables x and y, characterised by the membership functions $\mu_{A_j}(x)$ and $\mu_{B_j}(y)$, respectively.

Hence, a single-input-single-output (SISO) FIS can be implemented to sequentially generate the tuning or correction factors for the diagonal elements of $\mathbf{R}(k)$. First define $\mathbf{DoM}(k)_{(i,i)}$ as the FIS linguistic input variable and $\Delta R(k)$, the adjusting factor for $\mathbf{R}(k)_{(i,i)}$, as the FIS linguistic output variable. Next, define three fuzzy sets for $\mathbf{DoM}(k)_{(i,i)}$: $N =$ Negative, $ZE =$ Zero, and $P =$ Positive, describing the degree of mismatch; and three fuzzy sets for $\Delta R(k)$: $I =$ increase, $M =$ maintain, and $D =$ decrease, describing the action or correction to be taken. Different membership functions may be considered to define the fuzzy sets of the FIS input and output linguistic variables. However, from simulation studies of different processes it was found that triangular membership functions are sufficient for this task. Therefore, the fuzzy sets for both linguistic variables are defined as shown in Fig. 6. There, the parameters a and b can be selected in accordance with the knowledge available about the system at hand. For example, if there is knowledge about the maximum possible value that the measurement noise variances can take, then this value can be used to determine the parameter a. In the problems at hand, from experimentation it was observed that good results were obtained by defining $a = 1e-4$ (maximum possible measurement noise covariance value). The parameter b, which defines the maximum size of adjustment, can be selected as a percentage of a, for example 5% empirically demonstrated to produce smooth adjustments. If nothing is known about the measurement noise statistics, then the parameters a and b can be initially guessed and further adjusted based on simulation results.

Therefore, only three fuzzy rules are required to complete the FIS rule base:

- Rule 1. If $\mathbf{DoM}(k)_{(i,i)} = P$, then $\Delta R(k) = D$.
- Rule 2. If $\mathbf{DoM}(k)_{(i,i)} = ZE$, then $\Delta R(k) = M$.
- Rule 3. If $\mathbf{DoM}(k)_{(i,i)} = N$, then $\Delta R(k) = I$.

Hence, using the compositional rule of inference sum-prod and the centre of area (COA) defuzzification method, the adjusting factor for the diagonal elements of $\mathbf{R}(k)$, are calculated by

$$\Delta R(k) = \frac{\mu_P(\mathbf{DoM}(k)_{(i,i)})\mu_D + \mu_{ZE}(\mathbf{DoM}(k)_{(i,i)})\mu_M + \mu_N(\mathbf{DoM}(k)_{(i,i)})\mu_I}{\mu_P(\mathbf{DoM}(k)_{(i,i)}) + \mu_{ZE}(\mathbf{DoM}(k)_{(i,i)}) + \mu_N(\mathbf{DoM}(k)_{(i,i)})} \tag{53}$$

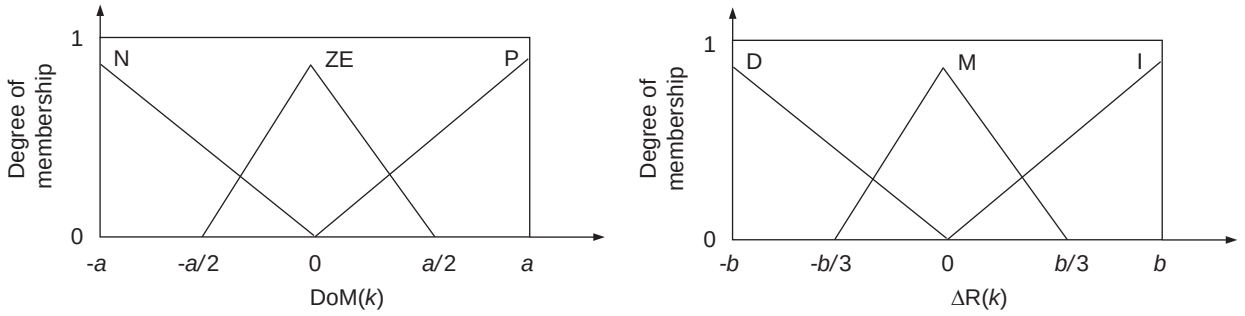

Fig. 6. Membership functions for DOM and ΔR.

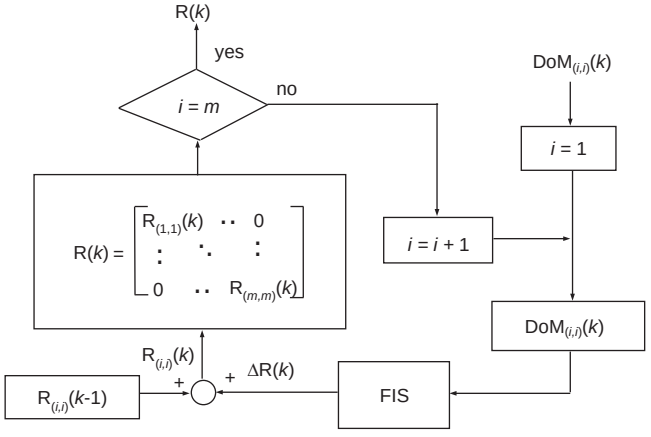

Fig. 7. Algorithm to sequentially generate the correction factors for R.

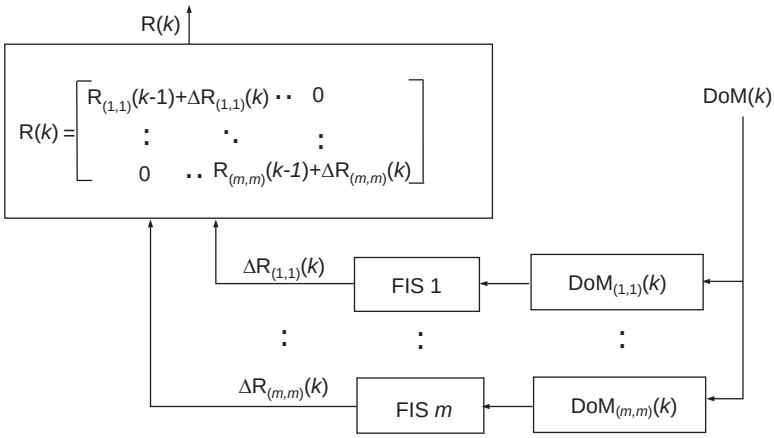

Fig. 8. Parallel SISO FISs implementation.

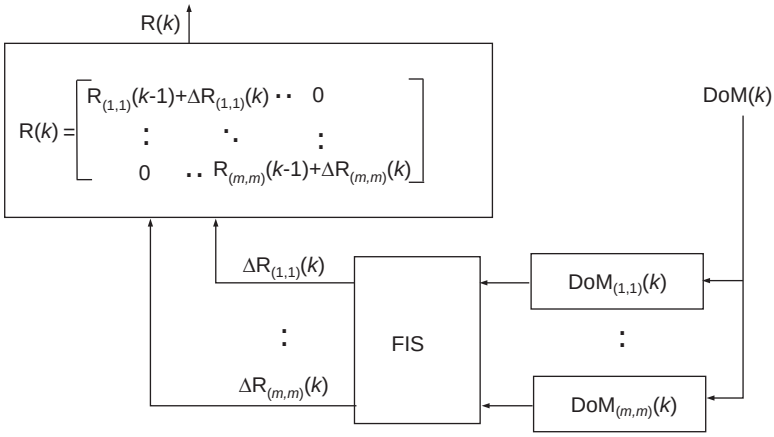

Fig. 9. MIMO FIS implementation.

Finally, the adjustment of $\mathbf{R}(k)$ is carried out by cyclically applying

$$\mathbf{R}(k)_{(i,i)} = \mathbf{R}(k-1)_{(i,i)} + \Delta R(k) \tag{54}$$

A graphical representation of this adjusting process is shown in Fig. 7. The algorithm starts by assigning an initial value to $\mathbf{R}(0)_{(i,i)}$. Usually, an initial estimate of the measurement noise variances can be determined by observing raw measurements in a stationary situation. However, if this is not possible, then $\mathbf{R}(0)_{(i,i)}$ can be assigned a value in function of the process noise covariance, e.g. $\mathbf{R}(0)_{(i,i)} = q_w$, where q_w is the maximum process noise covariance component in $\mathbf{Q}_{\mathrm{ex}}(k)$.

As an alternative to a sequential implementation, the adaptation algorithm can be implemented in two additional ways. In the first alternative, m parallel SISO FISs can be considered in order to adapt simultaneously all the diagonal elements of $\mathbf{R}(k)$, as is graphically represented in Fig. 8. In the second alternative, a multiple-input-multiple-output (MIMO) FIS with $3 \times m$ rules in the rule base can be used to adjust at once all the diagonal elements of $\mathbf{R}(k)$. This last alternative is graphically represented in Fig. 9. The use of any of these three ways of implementation: sequential, parallel or MIMO, depends on the computational resources and the problem at hand.

Thus, finally, by incorporating the fuzzy logic-based block into the EKF model, the final algorithm denoted as MMAE-WGI-FL MOKF method is shown in Fig. 10.

4. Applications of the proposed MMAE-WGI-FL MOKF method

4.1. The MMAE-WGI-FL MOKF method applied to simulated data

As an illustration, Fig. 11 shows a typical beam member, AB, with 6 elements. The structural member is discretised into a number of finite elements. The number of elements depends on the required resolution of damage identification. It is

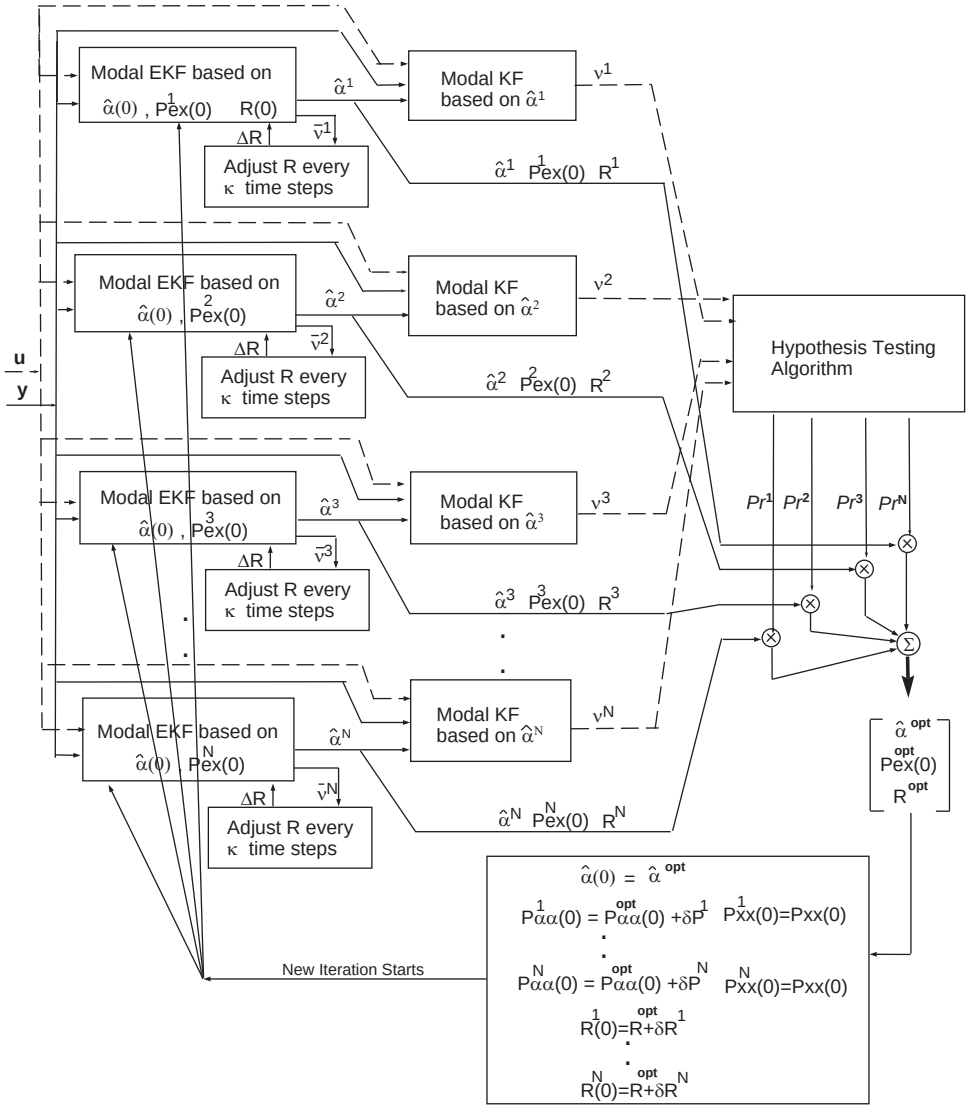

Fig. 10. MMAE-WGI-FL MOKF scheme for mechanical structure estimation.

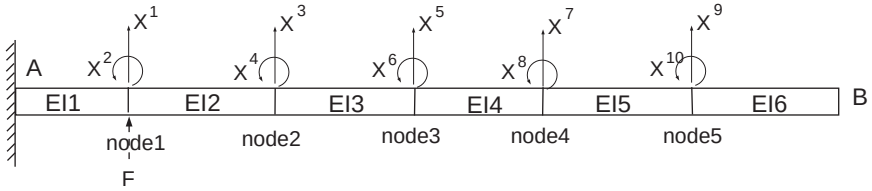

Fig. 11. A cantilever beam AB for damage detection.

assumed that damage would result in stiffness reduction (the change in *EI* value). Hence, if the beam stiffness can be identified before and after damage, then it may be possible to deduce the damage location and even the severity of damage.

Assuming the cantilever beam shown in Fig. 11 has 6 beam elements and imposed uniform modal damping. The parameters to be estimated are the ratio of the damaged stiffness value to the undamaged value for all 6 elements, plus one damping ratio ζ:

$$\boldsymbol{\alpha} = \begin{bmatrix} \dfrac{(EI)_1^D}{(EI)_1^U} & \cdots & \dfrac{(EI)_6^D}{(EI)_6^U} & \zeta \end{bmatrix}^{\mathrm{T}} \tag{55a}$$

$$\triangleq[\alpha_1 \quad \cdots \quad \alpha_6 \quad \zeta]^{\mathrm{T}} \tag{55b}$$

where superscript U denotes undamaged (or original) condition and superscript D denotes damaged (or changed) condition. Since damage is assumed to result in corresponding change in stiffness parameters, $\alpha_i\,(i=1,\ldots,6)$ is below 1 if damage takes place. The two extreme values are 1 denoting no damage at all and 0 denoting complete damage.

In a numerical simulation study, the dynamic response is simulated. Since in reality, it is difficult to obtain full measurements for all the DOFs in the finite element model, particularly for rotational DOFs. Here only translational displacements from the five nodes are simulated.

To achieve wideband response for better identification, band-limited white Gaussian random noise is used as the input signal. The reason for this kind of input lies in the two facts below:

1. The source that drives the parameter estimation in the EKF is the fundamental matrix $\mathbf{A}_F$. From Eq. (30), it can be seen that for the identification of parameters, it is desirable to have a large eigen-sensitivity to parameters to be estimated. Thus, theoretically, the more modes are excited, the better the parameter estimation will be. Thus, white Gaussian random noise is preferred.
2. However, considering the reduced MOKF proposed in Section 3.1, it is based on the assumption that only the first r modes are excited. Therefore, instead of using a random input and a low pass filter on the output (which violate the linear relationship between input and output), a band-limited white noise is adopted as the input to excite the system. By doing this and assuming the linearity of the system, the spectrum of the output will be confined within the range of certain modes. From a computational point of view, the fewer modes included, the less computation load. However, as was mentioned just now, from identification point of view, the more the modes are included, the better. Thus a balance needs to be found between identification results and computation load.

The clean signal of the dynamic response (output) of the beam is computed numerically using the Trapezoidal Rule for numerical integration. To account for the effects of I/O noise, both input and output clean signals are artificially contaminated with zero-mean Gaussian white noise.

Local damage of 30%, 10% and 20% are assumed to take place at element 1, 2, 3, respectively, and the corresponding integrity index is 0.7, 0.9, 0.8. The input and output are sampled at 10 kHz for 5000 points per set. The related parameters for the system are presented in Table 1.

The modified EKF method adopted in this study includes 3 extended Kalman filters. The initial parameters set for the 3 filters in the bank are presented in Table 2. The effectiveness of proposed method is presented by Tables 3, 4, 5 and Fig. 12. It's clear that the residuals are getting better with iterations going on, and the estimation results converged to the actual ones. The problems proposed in Section 2.2 have been solved, respectively:

1. By using modal equation instead of normal dynamic equation, the computational effort is decreased from the order of $(2n+a)^3$ to $(2r+a)^3$ (n: the number of DOFs, a: the number of unknown parameters, r: the number of modes included).

Table 1
Parameters used in simulation.

Parameters	Value
Length (mm)	900
Sectional area (mm × mm)	25.45 × 6.47
Density (kg/m^3)	2700
Young's modulus (GP)	70
Integrity index $\alpha1$ $\alpha2$ $\alpha3$	0.7 0.9 0.8
Damping ratio ζ	5e−3
Noise on measurement $R_{i,i}(0)$ $(i=1\ldots6)$	1e−8

Table 2
Initial parameters for the three filters.

Initial para	First filter	Second filter	Third filter
$\alpha_1(0)$	1	1	1
$\alpha_2(0)$	1	1	1
$\alpha_3(0)$	1	1	1
$\zeta(0)$	1e−3	1e−3	1e−3
Initial $Pex_{i,i}(0)$ $(i=7\ldots10)$	1e−4	3e−4	5e−4
Initial $R_{i,i}(0)$ $(i=1\ldots6)$	5e−6	5e−6	5e−6

Table 3
Identification results after the first iteration.

Estimated Para	Exact value	First filter	Second filter	Third filter
α_1	0.7	0.7200	0.6944	0.7079
α_2	0.9	0.9617	0.9417	0.8669
α_3	0.8	0.7501	0.7188	0.7515
ζ	0.005	0.0554	0.0082	0.0054
TANIS		1.9397	0.7619	0.5756
TAAU		−0.2728	−0.1712	−0.0720
Probability		0.1913	0.3064	0.5024
Optimal para				
α_1			0.7061	
α_2			0.9080	
α_3			0.7412	
ζ			0.01578	
Optimal $Pex_{i,i}$ $(i = 7\ldots10)$			6.746691e−4	
Optimal $R_{i,i}$ $(i = 1\ldots3)$		5.506191e−008	5.588902e−008	5.580108e−008
Optimal $R_{i,i}$ $(i = 4\ldots6)$		5.490538e−008	5.424910e−008	5.753701e−008

Table 4
Identification results after the second iteration.

Estimated para	Exact value	First filter	Second filter	Third filter
α_1	0.7	0.7053	0.6892	0.7091
α_2	0.9	0.8816	0.9511	0.8634
α_3	0.8	0.8002	0.8038	0.7982
ζ	0.005	0.0036	0.0044	0.0039
TANIS		5.7982	5.9127	5.8565
TAAU		−0.0127	−0.0295	−0.0145
Probability		0.4740	0.1091	0.4168
Optimal para				
α_1			0.7051	
α_2			0.8816	
α_3			0.7997	
ζ			0.00381	
Optimal $Pex_{i,i}$ $(i = 7\ldots10)$			6.669537e−004	
Optimal $R_{i,i}$ $(i = 1\ldots3)$		1.045916e−008	1.030602e−008	1.031147e−008
Optimal $R_{i,i}$ $(i = 4\ldots6)$		9.828832e−009	9.983579e−009	1.046926e−008

Table 5
Identification results after the third iteration.

Estimated para	Exact value	First filter	Second filter	Third filter
α_1	0.7	0.6965	0.7055	0.6922
α_2	0.9	0.9125	0.8778	0.9345
α_3	0.8	0.8023	0.7973	0.8036
ζ	0.005	0.0047	0.0050	0.0047
TANIS		6.1197	6.1604	6.1973
TAAU		−0.0129	−0.0132	−0.0217
Probability		0.4436	0.3993	0.1572
Optimal para				
α_1			0.6994	
α_2			0.9021	
α_3			0.8005	
ζ			0.00483	
Optimal $Pex_{i,i}$ $(i = 7\ldots10)$			6.287487e−004	
Optimal $R_{i,i}$ $(i = 1\ldots3)$	1.019385e−008	1.000552e−008	1.015951e−008	
Optimal $R_{i,i}$ $(i = 4\ldots6)$	9.592652e−009	9.815434e−009	1.016976e−008	

2. By using multiple EKF with different $\mathbf{P}_{ex}(0)$, the estimation process becomes more robust and information of all 3 filters is incorporated.
3. The estimation results are improved with each iteration thus demonstrating the effect of WGI method.
4. The fuzzy logic model can estimate $\mathbf{R}$ with a very rough initial estimation, and still converge to the actual value. The correct $\mathbf{R}$ helps the estimation of parameters.

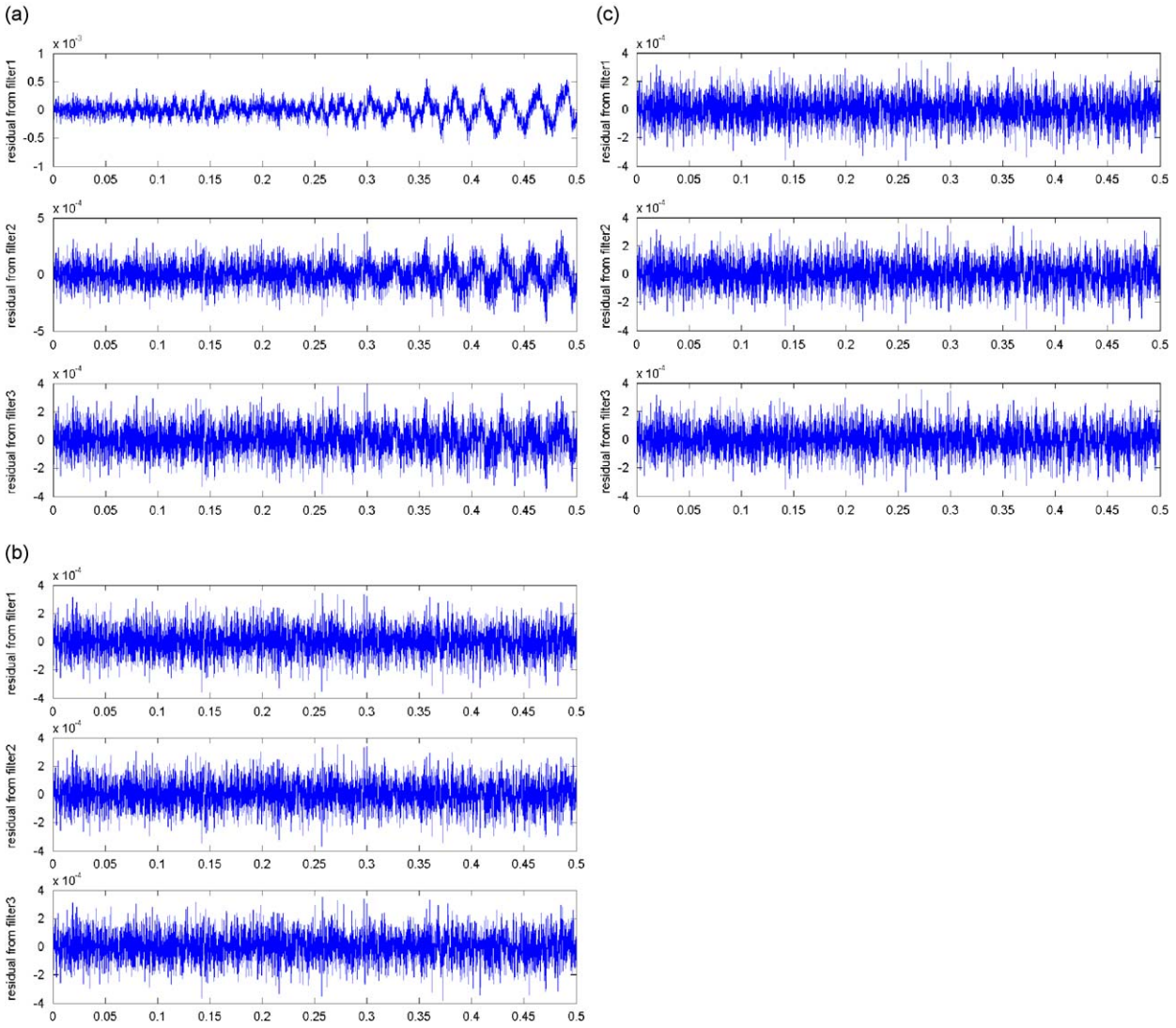

Fig. 12. Iterations: (a) results from the first iteration 1, (b) results from the second iteration and (c) results from the third iteration.

4.2. The MMAE-WGI-FL MOKF method applied to experimental data

In this research, an experimental study involving shaker-excited vibration tests of an aluminium cantilever beam was carried out in the laboratory as shown in Fig. 13. Zero-mean band-limited Gaussian white noise was used as the input signal to the shaker. A force transducer (23 g) was screwed on the bottom surface of the beam. Six accelerometers (7 g each) were screwed to the top surface along the centreline at selected positions that corresponded to the nodes in the finite element model described in Section 4.1.

The signals from the force transducer and accelerometers were fed into a DSpace acquisition and control system [19]. Experimental data were acquired at 10 kHz for 0.5 s. Instead of integrating acceleration to derive the velocities and displacements, acceleration signals were directly used as the measurement data. The reason for doing that is

1. Integration and double integration accompany loss of formation and the results drift significantly from the actual values if accelerations are noisy or have spurious mean.
2. In order to avoid the spurious mean level and drift due to the integration, normally a band pass filter is employed. However, after this process, white noise in accelerations after integration becomes non-white in the velocities and displacements. This would violate one basic requirement of the Kalman filter that noises in the observations be white.
3. The big difference between the spectrum of the displacement measurement and acceleration measurement is the percentage of higher modes spectrum. The higher modes are clearly displayed in the accelerations while not in the

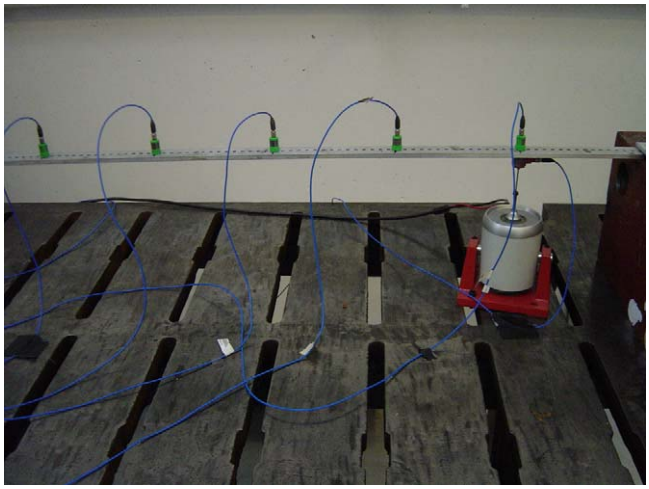

Fig. 13. Experiment setup.

Table 6
Identification results for the first experiment.

Initial para		First filter	Second filter	Third filter
Initial parameters for the three filters in experiment 1				
$\alpha_1(0)$		1	1	1
$\alpha_2(0)$		1	1	1
$\alpha_3(0)$		1	1	1
Initial $Pex_{i,i}(0)$ $(i = 7 \ldots 10)$		1e−4	3e−4	5e−4
Initial $R_{i,i}(0)$ $(i = 1 \ldots 6)$		1	1	1
Estimated para	Exact value	First filter	Second filter	Third filter
Identification results after the first iteration				
α_1	1.0	0.9496	1.0110	0.9241
α_2	4.3	3.5096	3.8776	4.0099
α_3	1.0	1.7348	1.6657	1.5909
Probability		0.3065	0.3129	0.3805
Optimal para				
α_1			0.9591	
α_2			3.8151	
α_3			1.6584	
Estimated para	Exact value	First filter	Second filter	Third filter
Identification results after the second iteration				
α_1	1.0	0.9359	0.9402	0.8742
α_2	4.3	4.2249	4.3098	*4.6051*
α_3	1.0	1.5326	1.5141	1.4183
Probability		0.3230	0.3310	0.3460
Optimal para				
α_1			0.9160	
α_2			4.3845	
α_3			1.4869	

displacements. Those higher modes are normally more sensitive to the local damage thus more preferable for the identification purpose.

The finite element model used to describe the cantilever beam includes 12 beam elements. Since the weight of force transducer and accelerometers can not be neglected compared with the beam itself, each sensor is modelled as a lumped mass in the finite element model. In addition, since many holes are drilled for the sensors setup, the original density has been re-measured and the EI values of the undamaged beam were first identified.

Damage of the cantilevered beam was created in a different way in the experimental study: by adding a lumped mass (23 g) under some certain locations (under one of six accelerometers). The aim was to demonstrate whether the proposed identification strategy can detect the change in the lumped mass and its location.

Table 7
Identification results for the second experiment.

Initial para		First filter	Second filter	Third filter
Initial parameters for the three filters in experiment 2				
$\alpha_1(0)$		1	1	1
$\alpha_2(0)$		1	1	1
$\alpha_3(0)$		1	1	1
Initial $Pex_{i,i}(0)$ $(i=7\ldots10)$		1e−5	3e−5	5e−5
Initial $R_{i,i}(0)$ $(i=1\ldots6)$		1	1	1
Estimated Para	Exact value	First filter	Second filter	Third filter
Identification results after the first iteration				
α_1	1.0	1.1253	1.0935	1.6478
α_2	1.0	1.3462	2.4136	0.9952
α_3	4.3	1.6449	1.8348	1.4087
Probability		0.3324	0.4442	0.2235
Optimal para				
α_1			1.2279	
α_2			1.2977	
α_3			1.6765	
Estimated para	Exact value	First filter	Second filter	Third filter
Identification results after the second iteration				
α_1	1.0	1.3657	1.0043	1.0333
α_2	1.0	1.0962	1.2100	1.3106
α_3	4.3	1.9947	2.2372	1.9753
Probability		0.2730	0.4474	0.2797
Optimal para				
α_1			1.1111	
α_2			1.2071	
α_3			2.0978	

In the first experiment, the lumped mass was put under the second sensor (node2). Here, 'damage' is assumed only happened under the node1, node2, and node3. Thus the parameters to be estimated are α_1, α_2, α_3 (the ratio of the changed mass value to the original mass value under the node1, node2, node3, respectively) and damping ratio ζ. Thus the true parameter values are: $\alpha_1=1$, $\alpha_2=4.3$, $\alpha_3=1$.

The initial parameters for the filters and estimation results for the first experiment are presented in Table 6. For comparison, the sensitivity based frequency model updating method [4] is also used for the same data, but the whole time series data is used (10 s). The frequency domain updating results are $\alpha_1=0.9923$, $\alpha_2=2.9140$, $\alpha_3=0.9001$.

In the second experiment, the lumped mass is put under the 6th sensor and 'damage' is assumed only happened under the node4, node5, and node6. α_1, α_2, α_3 become the ratio of the changed mass value to the original mass value under the node4, node5, node6, respectively. Thus the true values are $\alpha_1=1$, $\alpha_2=1$, $\alpha_3=4.3$.

The initial parameters for the filters and estimation results for the second experiment are presented in Table 7. Similarly, for comparison, the frequency domain updating results are $\alpha_1=1.4236$, $\alpha_2=1.4013$, $\alpha_3=2.5496$.

Note that the estimated parameters are not strictly equivalent to the theoretical values. This is mainly caused by the inaccuracy of the FE model of the cantilever beam. FE model including 12 beam elements is not perfect enough to represent the actual beam. Moreover, the boundary condition of the beam can be more accurately described by adding a coil spring with its stiffness to be estimated. However, it can be seen from the identification results that the current FE model (12 beam elements with one fixed boundary condition) is good enough to for damage identification purpose (at least for damage detection and localisation), and at the same time, the computational cost which is related to the complexity of the model is kept as low as possible.

5. Concluding remarks

The problem of damage detection and localisation has been treated by the proposed MMAE-WGI-FL MOKF method in the time domain. The application to an experimental and simulated cantilever beam shows that it is possible to detect, locate and quantify the damage by means of the proposed method.

It should still be mentioned that, the success of the damage identification by the EKF or its derivatives is strongly dependent on the quality of the original model. The advantage of EKF methods, when compared with the methods in frequency domain (like FE model updating by natural frequencies), lies in the higher sensitivity to the damage and its

information efficiency (in the examples above, using only 0.5 s period of I/O time series would provide a good result). On the other hand, unfortunately, the EKF methods are also sensitive to the modelling errors when compared with frequency domain estimation methods. Thus, a method where the sensitivity due to modelling errors is low while sensitivity with respect to the damage parameters is high is the ultimate goal for model identification. The proposed MMAE-WGI-FL MOKF, to some extent, is actually trying to provide a solution in this direction.

References

[1] H. Sohn, C.R. Farrar, F.M. Hemez, D.D. Shunk, S.W. Stinemates, B.R. Nadler, J.J. Czarnecki, A review of structural health monitoring literature form 1996–2001, Report LA-13976-MS, Los Alamos National Laboratory, 2004.
[2] A. Rytter, Vibration Based Inspection of Civil Engineering Structures, PhD Thesis, Department of Building Technology and Structural Engineering, Aalborg University, 1993.
[3] C.R. Farrar, S.W. Doebling, An overview of modal-based damage identification methods, *Proceedings of DAMAS Conference*, Sheffield, UK, June 1997.
[4] M. Friswell, J.E. Mottershead, *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, The Netherlands, 1995.
[5] P.C. Young, An instrumental variables method for real-time identification of a noisy process, *Automatica* 6 (1970) 271–288.
[6] S.F. Masri, T.K. Caughey, A nonparametric identification technique or nonlinear dynamic problems, *Journal of Applied Mechanics* 46 (1979) 433–447.
[7] A.H. Jazwinski, *Stochastic Process and Filtering Theory*, Academic Press, New York, 1970.
[8] C.-P. Fritzen, S. Zhu, Updating of finite element models by means of measured information, *Computers and Structures* 40 (1991) 475–486.
[9] C.-P. Fritzen, S. Seibold, D. Buchen, Application of filter techniques for damage detection in linear and nonlinear mechanical structures, *Proceedings of the 13th International Modal Analysis Conference*, 1995, pp. 1874–1881.
[10] K.F. Alvin, Efficient computation of eigenvector sensitivities for structural dynamics via conjugate gradients, *Proceedings of the 16th International Modal Analysis Conference*, 1998, pp. 652–659.
[11] D.T. Magill, Optimal adaptive estimation of sampled stochastic processes, *IEEE Transactions on Automatic Control* AC-10 (4) (1965) 434–439.
[12] P.D. Hanlon, P.S. Maybeck, Multiple-model adaptive estimation using a residual correlation Kalman filter bank, *IEEE Transactions on Aerospace and Electronic Systems* 36 (2) (2000) 393–406.
[13] P.S. Maybeck, P.D. Hanlon, Performance enhancements of a multiple model adaptive estimator, *IEEE Transactions on Aerospace and Electronic Systems* 31 (4) (1995) 1240–1254.
[14] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*, Wiley-Interscience, USA, 2001.
[15] M. Hoshiya, A. Sutoh, Kalman filter-finite element method in identification, *Journal of Engineering Mechanics* 119 (2) (1993) 197–210.
[16] M. Hoshiya, E. Saito, Structural identification by extended Kalman filter, *Journal of Engineering Mechanics (ASCE)* 110 (12) (1984) 1757–1771.
[17] J.-S. Lin, Y. Zhang, Nonlinear structural identification using extended Kalman filter, *Computers and Structures* 52 (4) (1994) 757–764.
[18] J.Z. Sasiadek, Q. Wang, Low cost automation using INS/GPS data fusion for accurate positioning, *Robotica* 21 (2003) 255–260.
[19] N. Quijano, K. Passino, A tutorial introduction to control systems development and implementation with dSPACE, 〈http://www.ece.osu.edu/_passino/dSPACEtutorial.doc.pdf〉.