Novel adaptive Bayesian scheme for real-time simultaneous anomaly detection and system identification

Dukang Huang[1, 2], Ke Huang[1*], Lei Xiao[1], Yafei Ma[1], Ka-Veng Yuen[3, 4], Lei Wang[1]

[1]School of Civil Engineering, Changsha University of Science and Technology, Changsha, China
[2]Department of Structural, Building & Geotechnical Engineering, Politecnico di Torino, Turin, Italy
[3]State Key Laboratory of Internet of Things for Smart City and Department of Civil and Environmental Engineering, University of Macau, Macao, China
[4]Guangdong-Hong Kong-Macau Joint Laboratory for Smart Cities, University of Macau, Macao, China

ABSTRACT

This study presents a novel approach for real-time anomaly detection and system identification. The approach eliminates the need for fixed threshold settings in anomaly detection and provides an efficient solution for simultaneous recognition of multiple anomaly types and identification of time-varying systems. Statistical models for random and gross errors are introduced to represent typical measurement anomalies, and Bernoulli random vectors are used for anomaly detection. Once potential anomalies are recognized, they are either excluded or compensated in further real-time system identification through detect-to-reject and detect-to-fix procedures. An adaptive Bayesian scheme updates both the Bernoulli and model parameters, allowing for real-time simultaneous anomaly detection and system identification. The approach is verified through numerical simulation and laboratory experiment. Moreover, it is implemented in a full-scale monitoring system. The proposed method effectively detects multiple anomaly types and achieves reliable identification results for time-varying systems.

Keywords

System identification; anomaly detection; real-time identification; Bernoulli distribution; Bayes' theorem

1 Introduction

System identification is fundamental to assessing structural behavior as it facilitates the estimation of key metrics and essential parameters that characterize the dynamic properties of a structure. Various system identification methods have been developed, including filter-based approaches [1-4], statistical and machine learning techniques [5-7], and signal-based methods

[8-13]. As a critical application of structural health monitoring (SHM), system identification provides essential insights into structural integrity and performance by detecting variations in material properties and geometric configurations through response measurements. SHM extends beyond system identification to encompass broader research areas, such as optimal sensor placement [14-16], damage detection [17-20], and structural reliability and risk assessment [21-23].

Although numerous system identification methods have been investigated, most rely on high-quality observations. Abnormal measurements are inevitable in practical applications due to harsh environments, sensor aging or failure, and electromagnetic interference [24]. In SHM, data anomalies are typically classified into two main categories: random errors and gross errors [25]. Random errors (shown in Fig. 1a) are unpredictable variations arising from uncontrollable factors in the measurement process. These errors lead to fluctuations in measured values around the actual value and are typically caused by environmental factors and instrument limitations. When a sufficient number of measurements are taken, random errors generally follow a Gaussian distribution centered around the mean. In contrast, gross errors (shown in Fig. 1b) are significant deviations in measured data caused by human mistakes, instrument failures, or procedural flaws, leading to highly inaccurate results. Bias and drift errors are common examples of gross errors. In particular, bias refers to a constant, time-invariant offset, whereas drift describes a dynamic deviation that progressively accumulates over time. Unlike random errors, which are statistical in nature, gross errors often stem from misreading instruments, recording incorrect values, or using faulty equipment. Identification results may be erroneous if measurement anomalies are not properly addressed during monitoring, potentially resulting in false alarms or overlooked issues [26, 27]. Hence, detecting and mitigating these anomalies is desired before assessing structural health.

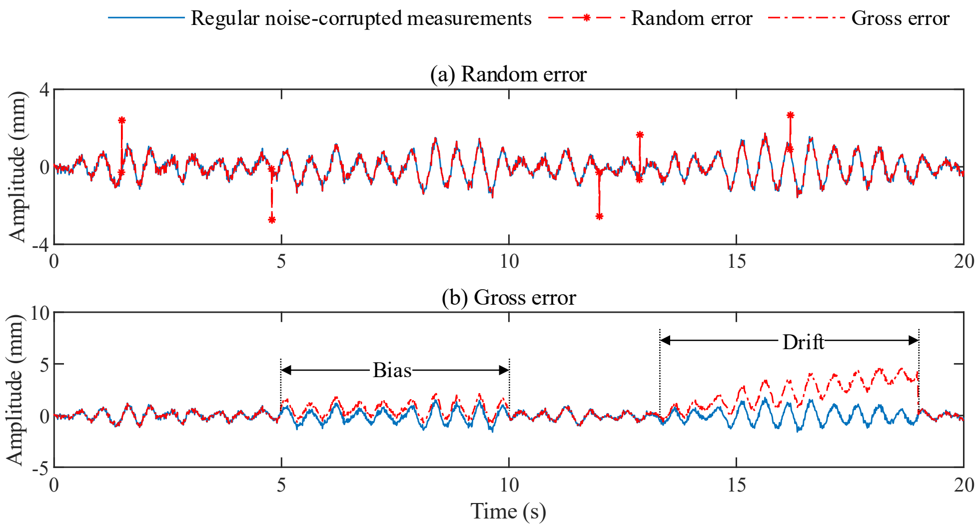

Fig. 1. Illustration of typical data anomalies encountered in SHM: (a) random error and (b) gross error.

Anomaly detection aims to refine and enhance measurements to ensure reliable and robust system identification. Generally, anomaly detection methods for SHM are divided into five categories: statistical methods [28-30], distance-based methods [31, 32], clustering methods [33, 34], artificial intelligence (AI)-based methods [35, 36], and online methods [37-39]. Statistical methods establish a probabilistic model based on pristine data and use statistical hypothesis tests to detect anomalies by assessing the consistency between the current data distribution and the pristine distribution. However, most of these methods assume that the probabilistic model remains unchanged throughout the detection process, rendering them unsuitable for time-varying systems. Distance-based methods detect anomalies by calculating distances between data points from a global perspective. A data anomaly alert is activated when the distance of the currently measured data exceeds predefined thresholds. Clustering methods group sampled data points into clusters using clustering algorithms. Data points that do not belong to regular clusters are considered anomalies. However, these methods require extra effort to extract suitable data features that maximize cluster separation. Otherwise, they may fail to detect anomalies when anomaly clusters closely resemble regular clusters. AI-based methods use machine learning algorithms to extract features from raw measurements to identify anomalies. However, AI-based methods, especially deep learning neural networks, require extensive training data to achieve satisfactory generalization. The effectiveness of previously discussed techniques is heavily reliant on the proper configuration of anomaly detection thresholds which can be determined using objective techniques. For instance, kernel density estimation [40] can be employed with a chosen significance level α. The detection threshold is set based on a $100 \times (1-\alpha)\%$ confidence criterion. Alternatively, thresholds may be optimized through cross-validation using offline historical data. However, such thresholds are typically static. In other words, they are fixed for a particular operational regime. In SHM, environmental and operational conditions can evolve significantly over time, potentially invalidating fixed thresholds for online use.

Several online anomaly detection methods have been established, including Bayesian online changepoint detection (BOCD) [37], interacting multiple models (IMM) [38], and switching Kalman filter (SKF) [39]. BOCD computes the posterior distribution over run lengths at each time point without requiring a fixed threshold. However, in practice, a decision rule is still needed to trigger an alarm. For example, a changepoint may be declared if the probability of run length zero given the data exceeds a predetermined threshold. This threshold effectively reintroduces a cutoff, which can be subjective or tuned offline, thereby limiting the adaptability of the method. IMM and SKF detect and isolate anomalies by selecting among predefined candidate models. However, when applied to systems with a large number of sensing channels, the number of possible combinations of model hypotheses may suffer from a combinatorial explosion, thereby increasing complexity and reducing scalability.

Recently, Yuen and Mu [41] proposed a novel concept for detecting outliers in monitoring data, namely the outlier probability. They suggested that observations with an outlier probability of 0.5 or higher should be considered anomalies and excluded from the system identification process. This method is non-subjective as it uses the probability threshold of 0.5 without any prior information. However, it can only detect a single type of anomaly, i.e., random errors. The method presented in [41] is ineffective in detecting initial gross errors, such as bias or drift.

Despite extensive research efforts on anomaly detection methods, several key challenges persist. While many existing anomaly detection methods allow threshold estimation from offline data, they typically depend on manual tuning and practitioner-defined configurations for different application scenarios, thereby limiting their applicability to online settings. Moreover, many methods are limited to identifying only a single type of anomaly. Furthermore, most investigations focus solely on anomalies caused by either measurement errors or structural damage. They overlook the scenarios where measurement anomalies and structural damage occur simultaneously. Additionally, extra effort is required to determine whether an anomaly is caused by structural damage, which degrades the efficiency of system identification.

To resolve the above problems, we propose an adaptive Bayesian scheme for real-time simultaneous anomaly detection and system identification. Statistical models are introduced to characterize typical types of measurement anomalies. This modeling enables the detection of multi-type anomaly by using Bernoulli random variables. Upon detecting potential anomalies, they are either excluded or compensated in subsequent real-time system identification using detect-to-reject and detect-to-fix approaches. The proposed method eliminates the requirements for subjective and manual threshold adjustments in anomaly detection. Furthermore, it facilitates real-time simultaneous anomaly detection and system identification, enhancing both robustness and efficiency.

The remainder of this paper is structured as follows. Section 2 presents the mathematical formulation for real-time system identification. Section 3 proposes a novel approach for real-time simultaneous anomaly detection and system identification. First, statistical models for random and gross errors are developed to characterize measurement anomalies, and Bernoulli random vectors are used to indicate the presence of both types of anomalies. Then, the Bernoulli and model parameters are updated concurrently using an adaptive Bayesian scheme. The proposed method is validated in Section 4 through two case studies: a numerical bridge and an experimental shear building model. In Section 5, the proposed approach is implemented in a full-scale monitoring system. Finally, the closing section presents the conclusions.

2 Mathematical formulation of the problem

For a structural system with N_d degrees of freedom (DOFs), the dynamics are governed by:

$$\mathbf{M}[\boldsymbol{\theta}_M(t)]\ddot{\mathbf{x}}(t) + \mathbf{C}[\boldsymbol{\theta}_C(t)]\dot{\mathbf{x}}(t) + \mathbf{K}[\boldsymbol{\theta}_K(t)]\mathbf{x}(t) = \mathbf{T}\mathbf{f}(t) \tag{1}$$

where $\mathbf{M} \in \mathbb{R}^{N_d \times N_d}$, $\mathbf{C} \in \mathbb{R}^{N_d \times N_d}$, and $\mathbf{K} \in \mathbb{R}^{N_d \times N_d}$ represent the mass, damping, and stiffness matrices of the system, respectively. These matrices are characterized by model parameters $\boldsymbol{\theta}(t) \equiv [\boldsymbol{\theta}_M(t)^{\mathrm{T}}, \boldsymbol{\theta}_C(t)^{\mathrm{T}}, \boldsymbol{\theta}_K(t)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{N_\theta}$, which may change over time, and N_θ represents the number of model parameters; $\mathbf{x}(t) \in \mathbb{R}^{N_d}$ denotes the displacement vector at time t; $\mathbf{f}(t) \in \mathbb{R}^{N_f}$ represents the external excitations acting on the structure at time t; N_f denotes the total number of excitations, and the excitations are characterized as zero-mean Gaussian process; $\mathbf{T} \in \mathbb{R}^{N_d \times N_f}$ denotes the location matrix of the external excitations.

An augmented state vector $\mathbf{y}(t) = [\mathbf{x}(t)^{\mathrm{T}}, \dot{\mathbf{x}}(t)^{\mathrm{T}}, \boldsymbol{\theta}(t)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2N_d + N_\theta}$ is defined to incorporate the displacement $\mathbf{x}$, velocity $\dot{\mathbf{x}}$, and parameters $\boldsymbol{\theta}$. Eq. (1) can be reformulated as the nonlinear state-space equation:

$$\begin{aligned}\dot{\mathbf{y}}(t) &= \begin{bmatrix} \dot{\mathbf{x}}(t) \\ -\mathbf{M}(\boldsymbol{\theta}_M)^{-1}[\mathbf{K}(\boldsymbol{\theta}_K)\mathbf{x}(t) + \mathbf{C}(\boldsymbol{\theta}_C)\dot{\mathbf{x}}(t) - \mathbf{T}\mathbf{f}(t)] \\ -\eta \mathbf{1}_{N_\theta \times 1} \end{bmatrix} \\ &\equiv \boldsymbol{\mathcal{G}}[\mathbf{y}(t), \mathbf{f}(t); \boldsymbol{\theta}(t)]\end{aligned} \tag{2}$$

where η represents a small positive scalar to characterize the variation of the parameter vector $\boldsymbol{\theta}$. The linearized form of Eq. (2) is obtained by applying the Taylor expansion:

$$\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{f}(t) + \tilde{\mathbf{h}} \tag{3}$$

where $\mathbf{A} = \partial\boldsymbol{\mathcal{G}}[\mathbf{y}(t), \mathbf{f}(t); \boldsymbol{\theta}(t)]/\partial\mathbf{y}(t)|_{\mathbf{y}(t)=\hat{\mathbf{y}}}$, and $\hat{\mathbf{y}}$ represents an arbitrary state; $\mathbf{B} = \partial\boldsymbol{\mathcal{G}}[\mathbf{y}(t), \mathbf{f}(t); \boldsymbol{\theta}(t)]/\partial\mathbf{f}(t)|_{\mathbf{y}(t)=\hat{\mathbf{y}}}$; $\tilde{\mathbf{h}} = \boldsymbol{\mathcal{G}}[\mathbf{y}(t), \mathbf{f}(t); \boldsymbol{\theta}(t)]|_{\mathbf{y}(t)=\hat{\mathbf{y}}} - \mathbf{A}\hat{\mathbf{y}}$. Given that dynamic responses are often sampled discretely over time, Eq. (3) can be converted to a discrete-time form:

$$\mathbf{y}_k = \mathbf{A}_{k-1}\mathbf{y}_{k-1} + \mathbf{B}_{k-1}\mathbf{f}_{k-1} + \tilde{\mathbf{h}}_{k-1} \tag{4}$$

where $\mathbf{y}_k \equiv \mathbf{y}(k\Delta t)$; k denotes the discrete-time sampling step, and Δt represents the fixed time interval between two consecutive sampling steps; $\mathbf{A}_{k-1} = \exp(\mathbf{A}\Delta t)$, $\mathbf{B}_{k-1} = \mathbf{A}^{-1}(\mathbf{A}_{k-1} - \mathbf{I}_{2N_d+N_\theta})\mathbf{B}$, $\mathbf{f}_{k-1} \equiv \mathbf{f}[(k-1)\Delta t]$, $\tilde{\mathbf{h}}_{k-1} = \mathbf{A}^{-1}(\mathbf{A}_{k-1} - \mathbf{I}_{2N_d+N_\theta})\tilde{\mathbf{h}}$; $\mathbf{I}_a$ denotes an identity matrix of size $a \times a$.

The measurement $\mathbf{z}_k$, affected by observation noise, is written in the form below:

$$\mathbf{z}_k = \mathbf{o}(\mathbf{y}_k) + \mathbf{n}_k \tag{5}$$

where $\mathbf{z}_k \equiv \left[\mathrm{z}_{1,k}, \mathrm{z}_{2,k}, \dots, \mathrm{z}_{N_z,k}\right]^{\mathrm{T}} \in \mathbb{R}^{N_z}$, and $\mathrm{z}_{l,k}, l = 1,2, \dots, N_z$ is the measurement of the lth observed DOF; $\mathbf{o}(\cdot)$ denotes the observation equation; $\mathbf{n}_k$ is the observation noise and is modeled as a zero-mean Gaussian process with covariance matrix $\mathbf{\Sigma}^n \in \mathbb{R}^{N_z \times N_z}$; N_z indicates the total number of observed DOFs.

Various system identification approaches [1-4] can be employed to identify the time-varying system in Eq. (1) based on Eqs. (4) and (5). However, measurements are inevitably corrupted with anomalies due to data transmission errors, sensor malfunctions, and environmental disturbances [26, 27]. The performance of system identification relies heavily on the quality of measurements [42]. Therefore, detecting abnormal measurements before estimation is desired. Conventional methods for detecting abnormal measurements typically rely on subjective thresholds, and they are limited to detecting a single type of anomaly [32, 41]. Moreover, most existing investigations focus solely on detecting anomalies caused by either measurement errors or structural damage. This narrow focus not only leads to inefficient system identification but also overlooks the scenarios where measurement anomalies and structural damage occur simultaneously. Such concurrent anomalies can interact in complex ways, compounding errors and producing highly misleading identification results. Addressing these combined anomalies is critical to ensure the robustness and reliability of real-time system identification. In the next section, an adaptive Bayesian scheme for real-time simultaneous anomaly detection and system identification is introduced. By characterizing typical types of measurement anomalies using a Bernoulli probabilistic model, multi-type abnormal data can be detected, and the time-varying systems can be identified simultaneously in a real-time manner.

3 Proposed Bayesian framework for anomaly detection and system identification

This section presents a Bayesian framework for real-time simultaneous anomaly detection and system identification. First, a Bernoulli probabilistic model is introduced to characterize random and gross errors. Then, an adaptive Bayesian scheme is used to detect measurement anomalies and estimate model parameters simultaneously.

3.1 Probabilistic modeling of measurement anomalies

Recall that data anomalies are predominantly classified into two types: random errors and gross errors [25]. To consider both errors, the anomaly-corrupted measurement $\tilde{\mathbf{z}}_k \in \mathbb{R}^{N_z}$ is formulated as:

$$\tilde{\mathbf{z}}_k = \mathbf{o}(\mathbf{y}_k) + \mathbf{n}_k + \mathbf{\Lambda}(\boldsymbol{\varsigma}_k^1)\mathbf{r}_k + \mathbf{\Lambda}(\boldsymbol{\varsigma}_k^2)\mathbf{g}_k \tag{6}$$

where $\mathbf{\Lambda}(\boldsymbol{a})$ indicates a square diagonal matrix with the elements of $\boldsymbol{a}$ on the main diagonal; $\mathbf{r}_k \in \mathbb{R}^{N_z}$ is the random error vector, and $\mathbf{r}$ is characterized as a Gaussian process with zero mean and covariance matrix $\mathbf{\Sigma}^r \in \mathbb{R}^{N_z \times N_z}$; $\mathbf{g}_k \in \mathbb{R}^{N_z}$ is the gross error vector, and $\mathbf{g}$ is characterized as a Gaussian process with mean $\boldsymbol{\mu}_k$ and covariance matrix $\mathbf{\Sigma}^\mu \in \mathbb{R}^{N_z \times N_z}$. To capture the effect of gross errors at the kth time step, $\boldsymbol{\mu}_k$ is modeled as a first-order Markovian process [43]:

$$\boldsymbol{\mu}_k = \mathbf{\Lambda}(\boldsymbol{\varsigma}_{k-1}^0 + \boldsymbol{\varsigma}_{k-1}^1)\boldsymbol{\mu}_k^0 + \mathbf{\Lambda}(\boldsymbol{\varsigma}_{k-1}^2)(\boldsymbol{\mu}_{k-1} + \boldsymbol{\varepsilon}_{k-1}) \tag{7}$$

where $\boldsymbol{\mu}_k^0 \equiv \left[\mu_{1,k}^0, \mu_{2,k}^0, \dots, \mu_{N_z,k}^0\right]^{\mathrm{T}} \in \mathbb{R}^{N_z}$, and $\mu_{l,k}^0, l = 1,2,\dots,N_z$, is sampled from the uniform distribution over $[c_l, d_l]$. This model captures the possibility of gross errors ranging from c_l to d_l in the lth observed DOF; $\boldsymbol{\varepsilon}_{k-1}$ indicates the potential variation of gross errors, and $\boldsymbol{\varepsilon}$ is characterized as a Gaussian process with zero mean and covariance matrix $\mathbf{\Sigma}^\varepsilon \in \mathbb{R}^{N_z \times N_z}$; $\boldsymbol{\varsigma}_{k-1}^w \in \mathbb{R}^{N_z}, w = 0,1,2$, is given as follows:

$$\boldsymbol{\varsigma}_{k-1}^w = \left[\delta(\ell_{1,k-1} - w), \delta(\ell_{2,k-1} - w), \dots, \delta(\ell_{N_z,k-1} - w)\right]^{\mathrm{T}}, \quad w = 0,1,2 \tag{8}$$

where $\delta(\cdot)$ represents the Kronecker delta function and $\delta(\ell_{l,k-1} - w)$ is defined by:

$$\delta(\ell_{l,k-1} - w) = \begin{cases} 1, & \ell_{l,k-1} = w \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where $\ell_{l,k-1}, l = 1,2,\dots,N_z$, is considered as a nuisance parameter and is designed to determine the presence of anomalies at the $(k-1)$th time step for the lth measured DOF. A vector $\boldsymbol{\ell}_{k-1} \equiv \left[\ell_{1,k-1}, \ell_{2,k-1}, \dots, \ell_{N_z,k-1}\right]^{\mathrm{T}} \in \mathbb{R}^{N_z}$ is defined to aggregate all the anomaly indexes; $\boldsymbol{\varsigma}_{k-1}^{w}, w = 0,1,2$, in Eq. (8) is considered as a Bernoulli random vector and represents an indicator vector to model whether a specific location of the measurement vector is free from abnormalities. Specifically, $\boldsymbol{\ell}_{k-1} = [0,0,0,\dots,0]^{\mathrm{T}}$ results in $\boldsymbol{\varsigma}_{k-1}^{0} = [1,1,1,\dots,1]^{\mathrm{T}}$, $\boldsymbol{\varsigma}_{k-1}^{1} = [0,0,0,\dots,0]^{\mathrm{T}}$ and $\boldsymbol{\varsigma}_{k-1}^{2} = [0,0,0,\dots,0]^{\mathrm{T}}$ and this indicates that there is no anomaly for all measurement channels according to Eq. (6). For example, $\boldsymbol{\ell}_{k-1} = [0,0,0,\dots,1]^{\mathrm{T}}$ implies $\boldsymbol{\varsigma}_{k-1}^{0} = [1,1,1,\dots,0]^{\mathrm{T}}, \boldsymbol{\varsigma}_{k-1}^{1} = [0,0,0,\dots,1]^{\mathrm{T}}$ and $\boldsymbol{\varsigma}_{k-1}^{2} = [0,0,0,\dots,0]^{\mathrm{T}}$ and this represents a random error for the last measured DOF. $\boldsymbol{\ell}_{k-1} = [2,0,0,\dots,0]^{\mathrm{T}}$ implies $\boldsymbol{\varsigma}_{k-1}^{0} = [0,1,1,\dots,1]^{\mathrm{T}}$, $\boldsymbol{\varsigma}_{k-1}^{1} = [0,0,0,\dots,0]^{\mathrm{T}}$ and $\boldsymbol{\varsigma}_{k-1}^{2} = [1,0,0,\dots,0]^{\mathrm{T}}$ and this refers to a gross error for the first measured DOF. It should be clarified that $\boldsymbol{\ell}_{k}$ is modeled as a Markovian process.

3.2 Bayesian inference for joint anomaly detection and system identification

Bayesian inference enables the joint estimation of system states and uncertain parameters. States and uncertain parameters are often incorporated into an augmented state vector, which is updated through Bayesian inference to achieve joint estimation. However, a major challenge in robust joint identification arises from the large number of unknown parameters. As the number of unknowns increases, ensuring parameter identifiability becomes more challenging. Whether a parameter is identifiable depends on the structure of the model, and the number of effectively observed DOFs [44, 45]. While a large number of unknown parameters does not necessarily imply unidentifiability, increasing the number of unknowns relative to a limited set of observations can reduce the information available for parameter estimation, thereby leading to unidentifiability. In particular, when the number of unknown parameters exceeds the number of effectively independent constraints imposed by the measurements, the estimation problem becomes ill-posed or ill-conditioned [46].

To enhance the identifiability of measurement anomalies and model parameters while simplifying the proposed algorithm, a two-stage scheme is implemented at each time step for anomaly detection and time-varying system identification based on an adaptive Bayesian algorithm. In the first stage, Bernoulli random vectors are updated using raw recorded data, and potential anomalies are detected based on the updated vectors. The detected random anomalies are excluded through a detect-to-reject strategy, while the detected gross anomalies are compensated using a detect-to-fix strategy. In the second stage, real-time system identification is performed by updating the model parameter vector using the refined data provided from the first stage.

3.2.1 Anomaly detection using Bernoulli probabilities

A vector is defined as $\mathbf{e}_k \equiv \left[\boldsymbol{\mathcal{X}}_k^{\mathrm{T}}, \boldsymbol{\mu}_k^{\mathrm{T}}, \boldsymbol{\ell}_k^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{2N_d+2N_z}$ to include the state vector $\boldsymbol{\mathcal{X}} = [\mathbf{x}^{\mathrm{T}}, \dot{\mathbf{x}}^{\mathrm{T}}]^{\mathrm{T}}$, the vector $\boldsymbol{\mu}$ for characterizing gross errors, and the anomaly indicator vector $\boldsymbol{\ell}$. Given the set of observed data $\widetilde{D}_{k-1} = \{\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \dots, \tilde{\mathbf{z}}_{k-1}\}$, the posterior probability density function (PDF) of $\mathbf{e}_k$ is obtained by:

$$p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}, \tilde{\mathbf{z}}_k\right) = c_e p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right) p\left(\tilde{\mathbf{z}}_k \middle| \widetilde{D}_{k-1}, \mathbf{e}_k\right) \tag{10}$$

where c_e is a normalizing factor; $p\left(\tilde{\mathbf{z}}_k \middle| \widetilde{D}_{k-1}, \mathbf{e}_k\right)$ is the likelihood function; $p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right)$ is the prior PDF, and it is defined as follows:

$$p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right) = \int p(\mathbf{e}_k | \mathbf{e}_{k-1})\, p\left(\mathbf{e}_{k-1} \middle| \widetilde{D}_{k-1}\right) d\mathbf{e}_{k-1} \tag{11}$$

Due to the challenges in solving high-dimensional integrals of Eqs. (10) and (11), a Monte Carlo simulation strategy is used to approximate the posterior PDF by generating N particles $\left\{\mathbf{e}_{k|k-1}^{(1)}, \mathbf{e}_{k|k-1}^{(2)}, \dots, \mathbf{e}_{k|k-1}^{(N)}\right\}$ from $p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right)$, which is a properly selected importance distribution. It should be noted that the subscript $|k-1$ represents the conditional expectation given the set of all past measurements up to the $(k-1)$th time step. As a result, $p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}, \tilde{\mathbf{z}}_k\right)$ can be empirically obtained by:

$$p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}, \tilde{\mathbf{z}}_k\right) \approx \sum_{j=1}^{N} \omega_{k|k-1}^{(j)} \delta\left(\mathbf{e}_k - \mathbf{e}_{k|k-1}^{(j)}\right) \tag{12}$$

where $\omega_{k|k-1}^{(j)}, j = 1,2,\dots,N$ is the weight of the jth particle, and it can be calculated as:

$$\omega_{k|k-1}^{(j)} \propto p\left(\mathbf{e}_{k|k-1}^{(j)} \middle| \widetilde{D}_{k-1}, \tilde{\mathbf{z}}_k\right) / p\left(\mathbf{e}_{k|k-1}^{(j)} \middle| \widetilde{D}_{k-1}\right) = p(\tilde{\mathbf{z}}_k \left| \mathbf{e}_{k|k-1}^{(j)}\right.), \quad j = 1,2, \ldots, N \tag{13}$$

where $\propto$ is the proportionality symbol; $p(\tilde{\mathbf{z}}_k \left| \mathbf{e}_{k|k-1}^{(j)}\right.)$ can be determined by:

$$p\left(\tilde{\mathbf{z}}_k \middle| \mathbf{e}_{k|k-1}^{(j)}\right) \propto (2\pi)^{-N_z/2} |\mathbf{\Sigma}|^{-1/2} \exp\left(-\boldsymbol{\delta}^{\mathrm{T}} \mathbf{\Sigma}^{-1} \boldsymbol{\delta}/2\right), \quad j = 1,2, \ldots, N \tag{14}$$

where $\mathbf{\Sigma}$ and $\boldsymbol{\delta}$ are given as follows:

$$\mathbf{\Sigma} = \mathbf{\Sigma}^n + \mathbf{\Lambda}\left(\boldsymbol{\varsigma}_{k|k-1}^{1,(j)}\right) \mathbf{\Sigma}^r + \mathbf{\Lambda}\left(\boldsymbol{\varsigma}_{k|k-1}^{2,(j)}\right) \mathbf{\Sigma}^\mu \tag{15}$$

$$\boldsymbol{\delta} = \tilde{\mathbf{z}}_k - \mathbf{o}\left(\mathcal{X}_{k|k-1}^{(j)}, \boldsymbol{\theta}_{k-1|k-1}\right) - \mathbf{\Lambda}(\boldsymbol{\varsigma}_{k|k-1}^{2,(j)}) \boldsymbol{\mu}_{k|k-1}^{(j)} \tag{16}$$

These weights $\left\{\omega_{k|k-1}^{(1)}, \omega_{k|k-1}^{(2)}, \ldots, \omega_{k|k-1}^{(N)}\right\}$ are normalized to sum to unity as:

$$\widetilde{\omega}_{k|k-1}^{(j)} = p(\tilde{\mathbf{z}}_k \left| \mathbf{e}_{k|k-1}^{(j)}\right.) / \sum_{j=1}^{N} p(\tilde{\mathbf{z}}_k \left| \mathbf{e}_{k|k-1}^{(j)}\right.), \quad j = 1,2, \ldots, N \tag{17}$$

where $\widetilde{\omega}_{k|k-1}^{(j)}$ denotes the normalized form of the weight $\omega_{k|k-1}^{(j)}$. In Eq. (16), when calculating the weights of particles $\left\{\mathbf{e}_{k|k-1}^{(1)}, \mathbf{e}_{k|k-1}^{(2)}, \ldots, \mathbf{e}_{k|k-1}^{(N)}\right\}$, the parameter vector is considered known and retains the value it held at the previous time step.

The PDF of $\mathbf{e}_k$ conditioned on $\widetilde{D}_{k-1}$ is obtained as:

$$p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right) = \left[\sum_{j=1}^{N} p\left(\mathbf{e}_k \middle| \mathbf{e}_{k-1|k-1}^{(j)}\right)\right] / N \tag{18}$$

where $\left\{\mathbf{e}_{k-1|k-1}^{(1)}, \mathbf{e}_{k-1|k-1}^{(2)}, \ldots, \mathbf{e}_{k-1|k-1}^{(N)}\right\}$ are particles drawn from $p(\mathbf{e}_{k-1} | \widetilde{D}_{k-1})$, and $p(\mathbf{e}_k \left| \mathbf{e}_{k-1|k-1}^{(j)}\right.)$ is given by:

$$p\left(\mathbf{e}_k \left|\mathbf{e}_{k-1|k-1}^{(j)}\right.\right) = p\left(\boldsymbol{\mathcal{X}}_k \left|\boldsymbol{\mathcal{X}}_{k-1|k-1}^{(j)}\right.\right) \times p\left(\boldsymbol{\mu}_k \left|\boldsymbol{\mu}_{k-1|k-1}^{(j)}, \boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right) \times p_m\left(\boldsymbol{\ell}_k \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right) \tag{19}$$

where $p\left(\boldsymbol{\mathcal{X}}_k \left|\boldsymbol{\mathcal{X}}_{k-1|k-1}^{(j)}\right.\right)$ is the PDF describing the underlying dynamical system, given by:

$$p(\boldsymbol{\mathcal{X}}_k \left|\boldsymbol{\mathcal{X}}_{k-1|k-1}^{(j)}\right.) = (2\pi)^{-N_d}\left|\boldsymbol{\Sigma}^f\right|^{-1/2}\exp\left[-(\boldsymbol{\mathcal{X}}_k - \bar{\boldsymbol{\mathcal{X}}})^{\mathrm{T}}{\boldsymbol{\Sigma}^f}^{-1}(\boldsymbol{\mathcal{X}}_k - \bar{\boldsymbol{\mathcal{X}}})/2\right] \tag{20}$$

where $\bar{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{G}}\left(\boldsymbol{\mathcal{X}}_{k-1|k-1}^{(j)}, \boldsymbol{\theta}_{k-1|k-1}\right)$; $\boldsymbol{\Sigma}^f$ denotes the covariance matrix of the unrecorded external excitations; $p\left(\boldsymbol{\mu}_k \left|\boldsymbol{\mu}_{k-1|k-1}^{(j)}, \boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right)$ is the PDF representing the evolution of gross error and it is given by:

$$p\left(\boldsymbol{\mu}_k \left|\boldsymbol{\mu}_{k-1|k-1}^{(j)}, \boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right) = \prod_{l=1}^{N_z}\left[\delta\left(\ell_{l,k-1|k-1}^{(j)}\right) + \delta\left(\ell_{l,k-1|k-1}^{(j)} - 1\right)\right]/(d_l - c_l) + \delta\left(\ell_{l,k-1|k-1}^{(j)} - 2\right)(2\pi)^{-1/2}{\sigma_{l,k}^{\varepsilon}}^{-1}\exp\left[-\left(\mu_{l,k} - \mu_{l,k-1|k-1}^{(j)}\right)^2/\left(\sqrt{2}\sigma_{l,k}^{\varepsilon}\right)^2\right] \tag{21}$$

where $\left(\sigma_{l,k}^{\varepsilon}\right)^2$ is lth variance of the vector $\boldsymbol{\varepsilon}_{k-1}$; $p_m\left(\boldsymbol{\ell}_k \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right)$ is the conditional probability mass function for the discrete abnormality indicator $\boldsymbol{\ell}_k$ and it is given as follows:

$$\begin{aligned} p_m\left(\boldsymbol{\ell}_k \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right) &= \prod_{l=1}^{N_z} p_m\left(\ell_{l,k} \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right) \\ &= \prod_{l=1}^{N_z}\sum_w p_m\left(\ell_{l,k} = w \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right)\delta\left(\ell_{l,k} - w\right) \\ &= \prod_{l=1}^{N_z}\sum_w \delta\left(\ell_{l,k} - w\right)/3, \quad w = 0,1,2 \end{aligned} \tag{22}$$

where $\delta\left(\ell_{l,k} - w\right)$ is defined in Eq. (9) and the conditional probability mass function $p_m\left(\ell_{l,k} = w \left|\boldsymbol{\ell}_{k-1|k-1}^{(j)}\right.\right)$ is taken as $1/3$ indicating no prior preference regarding the occurrence of measurement anomalies. It is noted that if prior knowledge about the types and statistical characteristics of anomalies is obtainable, it can be incorporated into Eq. (22) to enhance prediction.

Finally, the mean value $\mathbf{e}_{k|k}$ for $p\left(\mathbf{e}_k \middle| \widetilde{D}_k\right)$ is given as:

$$\mathbf{e}_{k|k} = (\sum_{j=1}^{N} \widetilde{\omega}_{k|k-1}^{(j)} \mathbf{e}_{k|k-1}^{(j)})/N \tag{23}$$

Eqs. (12) to (23) provide the fundamental propagation formulas necessary for an adaptive Bayesian algorithm to detect anomalies in a real-time manner. In addition, a resampling step is performed to prevent particle degeneracy [47].

Once potential anomalies are detected, they are either excluded or compensated for efficient and robust real-time system identification using detect-to-reject and detect-to-fix strategies.

3.2.2 Real-time system identification via unscented Kalman filter

After potential anomalies are detected, they are discarded or compensated to construct anomaly-free measurements, enabling efficient real-time system identification. An unscented Kalman filter (UKF) [48] is applied to real-time system identification. N particles $\left\{\mathbf{y}_{k-1|k-1}^{(1)}, \mathbf{y}_{k-1|k-1}^{(2)}, \dots, \mathbf{y}_{k-1|k-1}^{(N)}\right\}$ are generated from $p(\mathbf{y}_{k-1}|D_{k-1})$. For each particle $\mathbf{y}_{k-1|k-1}^{(j)}, j = 1,2, \dots, N$, $2N_y + 1$ sigma points are generated using unscented transformation:

$$\mathbf{y}_{k-1|k-1}^{(j),(i)} = \mathbf{y}_{k-1|k-1}^{(j)}, \quad i = 0 \tag{24}$$

$$\mathbf{y}_{k-1|k-1}^{(j),(i)} = \mathbf{y}_{k-1|k-1}^{(j)} + \left\{\sqrt{\left(N_y + \lambda\right)\mathbf{P}_{y,k-1|k-1}^{(j)}}\right\}_i, \quad i = 1,2, \dots, N_y \tag{25}$$

$$\mathbf{y}_{k-1|k-1}^{(j),(i)} = \mathbf{y}_{k-1|k-1}^{(j)} - \left\{\sqrt{\left(N_y + \lambda\right)\mathbf{P}_{y,k-1|k-1}^{(j)}}\right\}_{i-N_y}, \quad i = N_y + 1, N_y + 2, \dots, 2N_y \tag{26}$$

where $\lambda = \phi^2\left(N_y + \gamma\right) - N_y$ is a scaling parameter; $\phi \in [10^{-4}, 1]$, $N_y = 2N_d + N_\theta$, and $\gamma = 0$; $\mathbf{P}_{y,k-1|k-1}^{(j)}$ indicates the covariance matrix of $\mathbf{y}_{k-1|k-1}^{(j)}$; $\left\{\sqrt{\left(N_y + \lambda\right)\mathbf{P}_{y,k-1|k-1}^{(j)}}\right\}_i$ denotes the ith column of the matrix square root of $(N_y + \lambda)\mathbf{P}^{(j)}_{y,k-1|k-1}$. Then, each sigma point is propagated by using Eq. (4):

$$\mathbf{y}^{(j),(i)}_{k|k-1} = \mathbf{A}^{(j),(i)}_{k-1}\mathbf{y}^{(j),(i)}_{k-1|k-1} + \tilde{\mathbf{h}}^{(j),(i)}_{k-1} \tag{27}$$

The augmented state vector $\mathbf{y}^{(j)}_{k|k-1}$ and its covariance matrix are given as:

$$\mathbf{y}^{(j)}_{k|k-1} = \sum_{i=0}^{2N_y} \omega^{(i)}_m \mathbf{y}^{(j),(i)}_{k|k-1} \tag{28}$$

$$\mathbf{P}^{(j)}_{y,k|k-1} = \sum_{i=0}^{2N_y} \omega^{(i)}_c \left(\mathbf{y}^{(j),(i)}_{k|k-1} - \mathbf{y}^{(j)}_{k|k-1}\right)\left(\mathbf{y}^{(j),(i)}_{k|k-1} - \mathbf{y}^{(j)}_{k|k-1}\right)^{\mathrm{T}} + \mathbf{B}_{k-1}\mathbf{\Sigma}^{f}\mathbf{B}^{\mathrm{T}}_{k-1} \tag{29}$$

where $\omega^{(i)}_m$ represents the weight associated with the ith sigma point; $\omega^{(i)}_c$ corresponds to the weight for the covariance matrix of the ith sigma point; $\omega^{(i)}_m$ and $\omega^{(i)}_c$ are given by:

$$\omega^{(0)}_m = \lambda/(N_y + \lambda) \tag{30}$$

$$\omega^{(0)}_c = \lambda/(N_y + \lambda) + 1 - \phi^2 + \beta \tag{31}$$

$$\omega^{(i)}_m = \omega^{(i)}_c = 1/2(N_y + \lambda), \quad i = 1,2,\ldots,2N_y \tag{32}$$

where β is a scalar controlling the weighting of the zeroth sigma point for calculating the covariance and is taken as 2 in this study according to [48]. On the other hand, each sigma point is propagated by using Eq. (5):

$$\mathbf{z}^{(j),(i)}_{k|k-1} = \mathbf{o}\left(\mathbf{y}^{(j),(i)}_{k|k-1}\right) \tag{33}$$

The expected observation vector $\mathbf{z}^{(j)}_{k|k-1}$ and the corresponding covariance matrix are given below:

$$\mathbf{z}_{k|k-1}^{(j)} = \sum_{i=0}^{2N_y} \omega_m^{(i)} \mathbf{z}_{k|k-1}^{(j),(i)} \tag{34}$$

$$\mathbf{P}_{z,k|k-1}^{(j)} = \sum_{i=0}^{2N_y} \omega_c^{(i)} \left(\mathbf{z}_{k|k-1}^{(j),(i)} - \mathbf{z}_{k|k-1}^{(j)}\right)\left(\mathbf{z}_{k|k-1}^{(j),(i)} - \mathbf{z}_{k|k-1}^{(j)}\right)^{\mathrm{T}} + \mathbf{\Sigma}^n \tag{35}$$

The augmented state vector $\mathbf{y}_{k|k}^{(j)}$ and its covariance matrix $\mathbf{P}_{y,k|k}^{(j)}$ are updated as follows:

$$\mathbf{y}_{k|k}^{(j)} = \mathbf{y}_{k|k-1}^{(j)} + \mathbf{G}_k^{(j)}\left(\mathbf{z}_k - \mathbf{z}_{k|k-1}^{(j)}\right) \tag{36}$$

$$\mathbf{P}_{y,k|k}^{(j)} = \xi\left[\mathbf{P}_{y,k|k-1}^{(j)} - \mathbf{G}_k^{(j)}\mathbf{P}_{z,k|k-1}^{(j)}{\mathbf{G}_k^{(j)}}^{\mathrm{T}}\right] \tag{37}$$

where ξ is a fading factor [49]; $\mathbf{G}_k^{(j)}$ represents the Kalman gain:

$$\mathbf{G}_k^{(j)} = \mathbf{P}_{yz,k|k-1}^{(j)}\left(\mathbf{P}_{z,k|k-1}^{(j)}\right)^{-1} \tag{38}$$

where $\mathbf{P}_{yz,k|k-1}^{(j)}$ denotes the covariance matrix between $\mathbf{y}_{k|k-1}^{(j)}$ and $\mathbf{z}_{k|k-1}^{(j)}$:

$$\mathbf{P}_{yz,k|k-1}^{(j)} = \sum_{i=0}^{2N_y} \omega_c^{(i)} \left(\mathbf{y}_{k|k-1}^{(j),(i)} - \mathbf{y}_{k|k-1}^{(j)}\right)\left(\mathbf{z}_{k|k-1}^{(j),(i)} - \mathbf{z}_{k|k-1}^{(j)}\right)^{\mathrm{T}} \tag{39}$$

As a result, the proposal distribution $\mathcal{N}\left(\mathbf{y}_{k|k}^{(j)}, \mathbf{P}_{y,k|k}^{(j)}\right), j = 1,2,\ldots,N$ can be obtained to sample N particles $\left\{\mathbf{y}_{k|k-1}^{*(1)}, \mathbf{y}_{k|k-1}^{*(2)}, \ldots, \mathbf{y}_{k|k-1}^{*(N)}\right\}$. Finally, the updated augmented state vector and the corresponding covariance matrix are given as:

$$\mathbf{y}_{k|k} = \left(\sum_{j=1}^{N} \tilde{\tau}_{k|k-1}^{(j)} \mathbf{y}_{k|k-1}^{*(j)}\right)/N \tag{40}$$

$$\mathbf{P}_{y,k|k} = \sum_{j=1}^{N} \tilde{\tau}_{k|k-1}^{(j)} \left(\mathbf{y}_{k|k-1}^{*(j)} - \mathbf{y}_{k|k}\right)\left(\mathbf{y}_{k|k-1}^{*(j)} - \mathbf{y}_{k|k}\right)^{\mathrm{T}} \tag{41}$$

where $\tilde{\tau}_{k|k-1}^{(j)}$ is the normalized weight of the jth particle:

$$\tilde{\tau}_{k|k-1}^{(j)} = \tau_{k|k-1}^{(j)} / \sum_{j=1}^{N} \tau_{k|k-1}^{(j)} = p\left(\mathbf{z}_k \middle| \mathbf{y}_{k|k-1}^{*(j)}\right) / \sum_{j=1}^{N} p\left(\mathbf{z}_k \middle| \mathbf{y}_{k|k-1}^{*(j)}\right) \tag{42}$$

where $p\left(\mathbf{z}_k \middle| \mathbf{y}_{k|k-1}^{*(j)}\right)$ is the likelihood given by:

$$p\left(\mathbf{z}_k \middle| \mathbf{y}_{k|k-1}^{*(j)}\right) \propto (2\pi)^{-N_z/2}|\boldsymbol{\Sigma}^n|^{-1/2}\exp\left\{-\left[\mathbf{z}_k - \mathbf{o}\left(\mathbf{y}_{k|k-1}^{*(j)}\right)\right]^{\mathrm{T}} \boldsymbol{\Sigma}^{n^{-1}} \left[\mathbf{z}_k - \mathbf{o}\left(\mathbf{y}_{k|k-1}^{*(j)}\right)\right]/2\right\} \tag{43}$$

The resampling step is used to avoid particle degeneracy [47]. Fig. 2 presents the probabilistic graphical model of the proposed approach for real-time simultaneous anomaly detection and system identification.

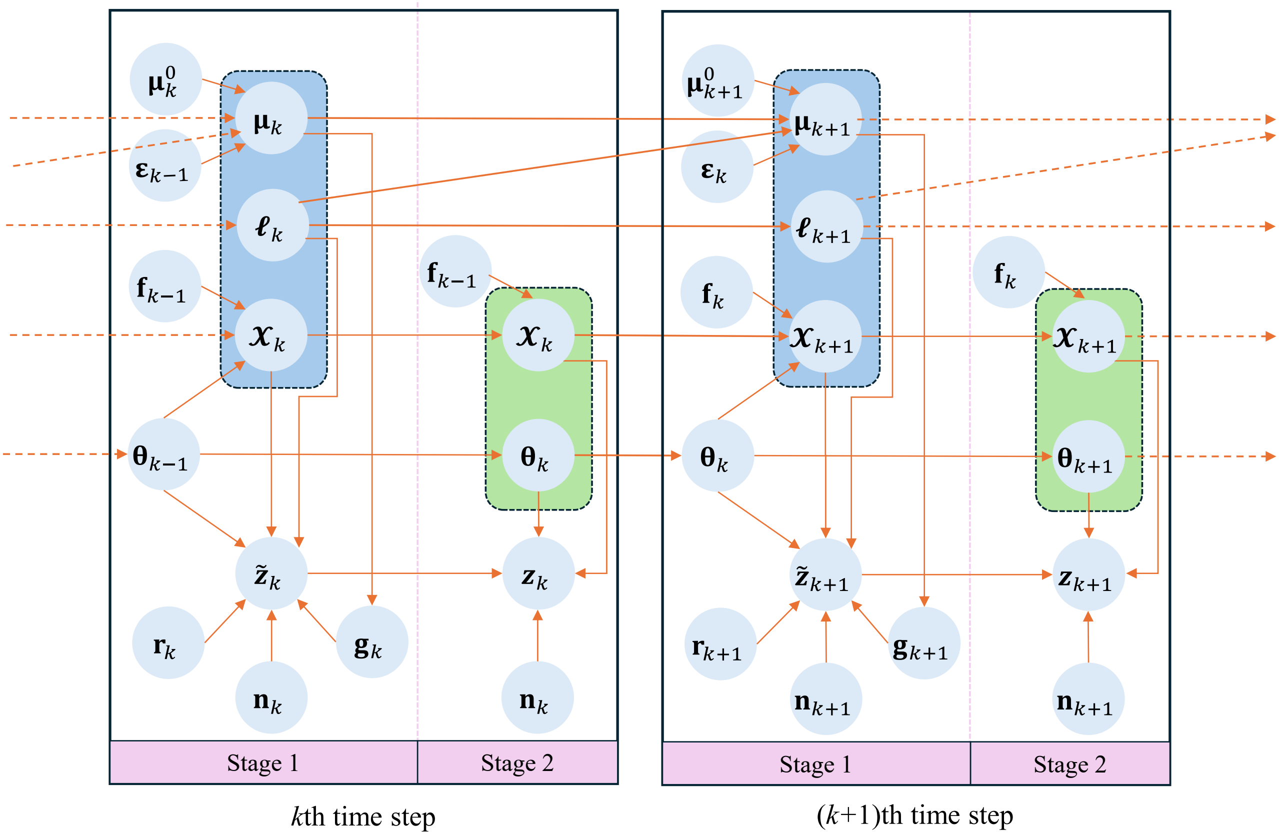

Fig. 2. Probabilistic graphical model visualizing the interdependencies among states, model parameters, anomaly indicators, and measurement errors.

The procedure of the adaptive Bayesian scheme for real-time simultaneous anomaly detection and system identification is outlined as follows:

For $k = 0$,

(1) Start with initial PDF $p(\mathbf{e}_0)$ and $p(\mathbf{y}_0)$. Draw particles $\left\{\mathbf{e}_0^{(1)}, \mathbf{e}_0^{(2)}, \dots, \mathbf{e}_0^{(N)}\right\}$ from $p(\mathbf{e}_0)$

and particles $\left\{\mathbf{y}_0^{(1)}, \mathbf{y}_0^{(2)}, \dots, \mathbf{y}_0^{(N)}\right\}$ from $p(\mathbf{y}_0)$.

For $k = 1,2, \dots$

Stage 1: Anomaly detection

(2) Generate N particles $\left\{\mathbf{e}_{k|k-1}^{(1)}, \mathbf{e}_{k|k-1}^{(2)}, \dots, \mathbf{e}_{k|k-1}^{(N)}\right\}$ from $p\left(\mathbf{e}_k \middle| \widetilde{D}_{k-1}\right)$.

(3) Calculate the particle weights $\left\{\widetilde{\omega}_{k|k-1}^{(1)}, \widetilde{\omega}_{k|k-1}^{(2)}, \dots, \widetilde{\omega}_{k|k-1}^{(N)}\right\}$ using Eq. (17).

(4) Update $\mathbf{e}_{k|k}$ using Eq. (23). If a random error is detected, it is excluded; if a gross error is detected, it is compensated based on the updating results. Consequently, the refined measurement vector $\mathbf{z}_k$ can be obtained at the kth time step.

(5) Resample $\left\{\mathbf{e}_{k|k-1}^{(1)}, \mathbf{e}_{k|k-1}^{(2)}, \dots, \mathbf{e}_{k|k-1}^{(N)}\right\}$ based on $\left\{\widetilde{\omega}_{k|k-1}^{(1)}, \widetilde{\omega}_{k|k-1}^{(2)}, \dots, \widetilde{\omega}_{k|k-1}^{(N)}\right\}$.

Stage 2: System identification

(6) Generate $2N_y + 1$ sigma points for each particle $\mathbf{y}_{k-1|k-1}^{(j)}, j = 1,2, \dots, N$.

(7) Calculate $\mathbf{y}_{k|k-1}^{(j)}$ and $\mathbf{P}_{y,k|k-1}^{(j)}, j = 1,2, \dots, N$, using Eqs. (28) and (29), respectively.

(8) Update $\mathbf{y}_{k|k}^{(j)}$ and $\mathbf{P}_{y,k|k}^{(j)}, j = 1,2, \dots, N$, to obtain the proposal distribution $\mathcal{N}\left(\mathbf{y}_{k|k}^{(j)}, \mathbf{P}_{y,k|k}^{(j)}\right)$ using Eqs. (36) and (37), respectively.

(9) Generate N particles $\left\{\mathbf{y}_{k|k-1}^{*(1)}, \mathbf{y}_{k|k-1}^{*(2)}, \dots, \mathbf{y}_{k|k-1}^{*(N)}\right\}$ from $\mathcal{N}\left(\mathbf{y}_{k|k}^{(j)}, \mathbf{P}_{y,k|k}^{(j)}\right), j = 1,2, \dots, N$.

(10) Calculate the particle weight $\left\{\tilde{\tau}_{k|k-1}^{(1)}, \tilde{\tau}_{k|k-1}^{(2)}, \dots, \tilde{\tau}_{k|k-1}^{(N)}\right\}$ using Eq. (42).

(11) Update $\mathbf{y}_{k|k}$ and $\mathbf{P}_{y,k|k}$ using Eqs. (40) and (41), respectively.

(12) Resample $\left\{\mathbf{y}_{k|k-1}^{(1)}, \mathbf{y}_{k|k-1}^{(2)}, \dots, \mathbf{y}_{k|k-1}^{(N)}\right\}$ based on $\left\{\tilde{\tau}_{k|k-1}^{(1)}, \tilde{\tau}_{k|k-1}^{(2)}, \dots, \tilde{\tau}_{k|k-1}^{(N)}\right\}$.

(13) Repeat the procedure from step (2).

Fig. 3 presents the flowchart of the proposed Bayesian scheme for real-time simultaneous anomaly detection and system identification.

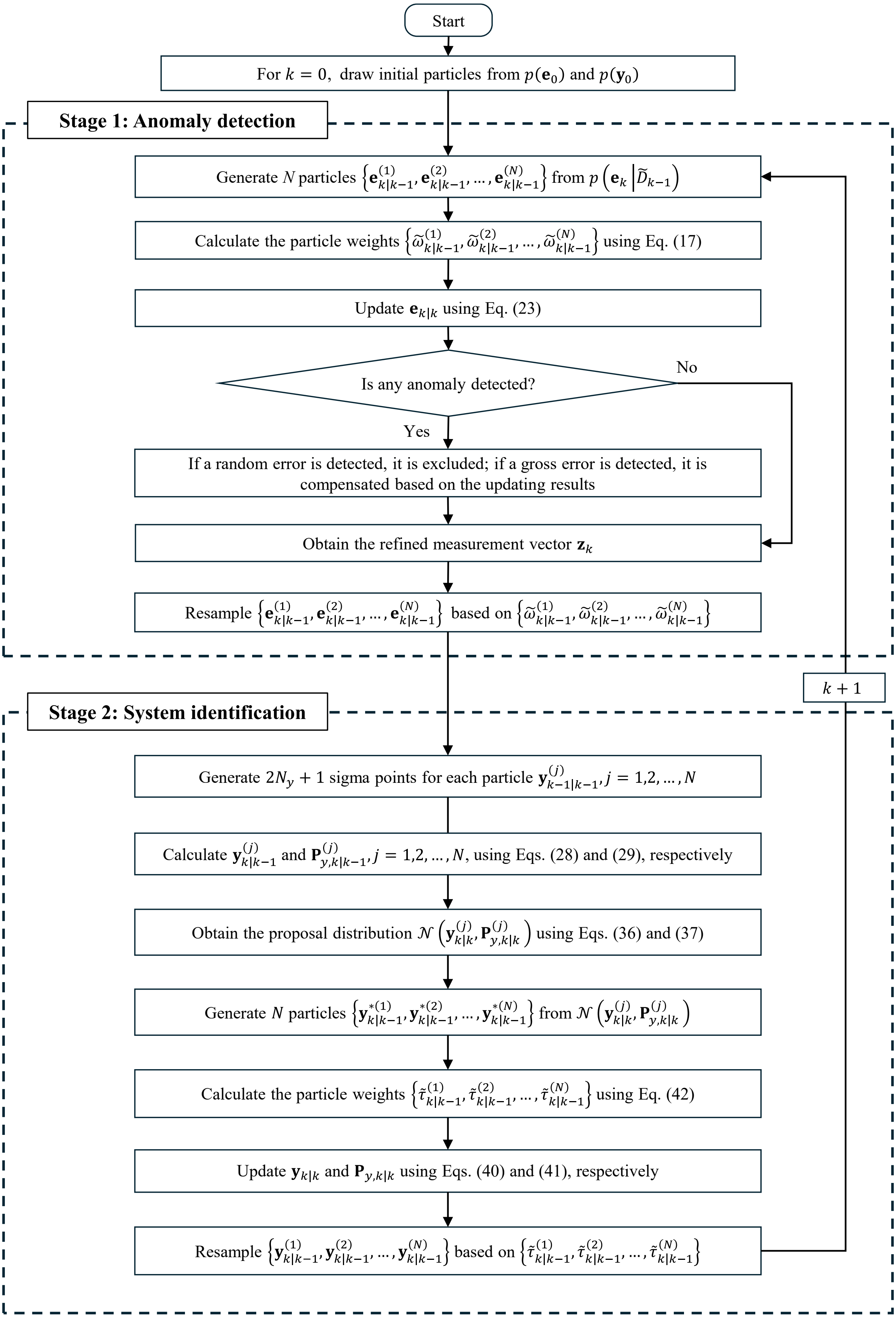

Fig. 3. Flowchart outlining the proposed approach consisting of two primary stages: Stage 1: Anomaly detection for raw measurements; Stage 2: System identification using refined measurements.

4 Numerical and experimental validation

4.1 Numerical case study: bridge

A bridge with three spans (45+90+45 m) and two piers is considered to validate the proposed approach. The bridge model consists of 19 nodes and 18 elements (shown in Fig. 4). The girder consists of 12 elements of equal length, and each pier is composed of 3 elements of equal length. Table 1 provides detailed information on the girder and piers. The first natural frequency of the bridge is 1.536 Hz. The structural damping is characterized by the Rayleigh model with coefficients $\alpha_1 = 0.2977\ \mathrm{s}^{-1}$ and $\alpha_2 = 0.0011\ \mathrm{s}$. Consequently, the first two modal damping ratios of the bridge are 2%. Six stiffness parameters are utilized to characterize the stiffness matrix $\mathbf{K} = \sum_{q=1}^{6} \theta_K^{(q)} \widehat{\mathbf{K}}^{(q)}$, where $\theta_K^{(q)}$ and $\widehat{\mathbf{K}}^{(q)}$ represent the stiffness parameter and the nominal stiffness matrix of the qth substructure, respectively. The girder is assigned four stiffness parameters, each representing one of the four equal-length segments, and each pier is assigned one stiffness parameter.

Table 1
Material and geometric properties of the girder and piers.

Property	Unit	Girder	Piers
Mass density	kg/m^3	2700	2650
Modulus of elasticity	GPa	34.5	32.0
Cross-sectional area	m^2	9.8	17.5
Moment of inertia	m^4	10.2	54.5
Element length	m	15	6

The bridge is subjected to horizontal and vertical ground excitations modeled as Gaussian white noise with a spectral intensity of $1.5 \times 10^{-9}\ \mathrm{m}^2 \cdot \mathrm{s}^{-3}$. The monitoring duration is 100 s, using a 100 Hz sampling rate. Measurements were collected using accelerometers to record the vertical acceleration responses at the 2nd, 3rd, 6th, 7th, 9th, 10th, 11th, and 12th nodes, as well as the horizontal acceleration responses at the 14th, 15th, 17th, and 18th nodes. This sensor configuration measures dynamic responses at 12 DOFs ($N_z = 12$). The noise level of the regular noise-corrupted measurements is taken as 10% root mean square (RMS) of the corresponding noise-free response quantities. During the monitoring period, two instances of damage occurred. Specifically, the values of $\theta_K^{(2)}$ and $\theta_K^{(6)}$ decrease by 5% and 10% at 30 s and 60 s, respectively. The initial stiffness parameter vector is assumed to have a 5% uniformly distributed error relative to the actual values.

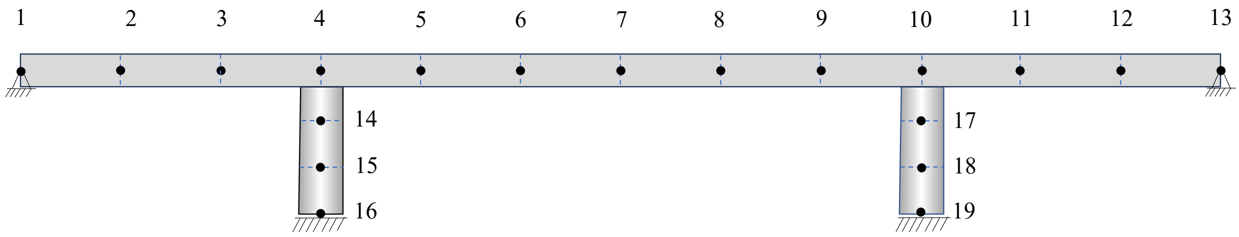

Fig. 4. Finite element model of the bridge with numbered nodes (black dots).

Both random and gross errors are considered in this example. Random errors are simulated as isolated outliers, with a consistent occurrence rate of 0.02 applied to the measurements at the 2nd, 7th, and 9th nodes. These errors follow a uniform distribution covering the range from the minimum to the maximum values of the corresponding noise-free response quantities. In addition, random errors appear between 20 s and 40 s at the 2nd and 7th nodes, and between 40 s and 50 s at the 9th node. Gross errors are modeled as bias and drift. Bias errors occur between 25 s and 45 s at the 6th node, and between 60 s and 70 s at the 9th node. The bias levels are set as 10% and 20% of the RMS values of the corresponding noise-free response quantities for the measurements of the 6th and 9th nodes, respectively. Drift errors appear between 50 s and 70 s at the 7th node, and between 70 s and 90 s at the 9th node. The drift growth factors, representing the proportional increase in drift error over time, are set to 9.0×10^{-7} and 5.0×10^{-7} for the 7th and 9th nodes, respectively. It should be emphasized that, while this study simulates random errors using isolated outliers and gross errors using biases and drifts for illustration, the proposed method is versatile and applicable to more complex scenarios. The details of measurement anomalies remain unknown during the entire estimation process and are utilized solely for simulation purposes. In this study, N is taken as 100. Fig. 5 presents the time-history noisy acceleration measurements at the 2nd, 6th, 7th, and 9th nodes, corrupted by random and gross errors.

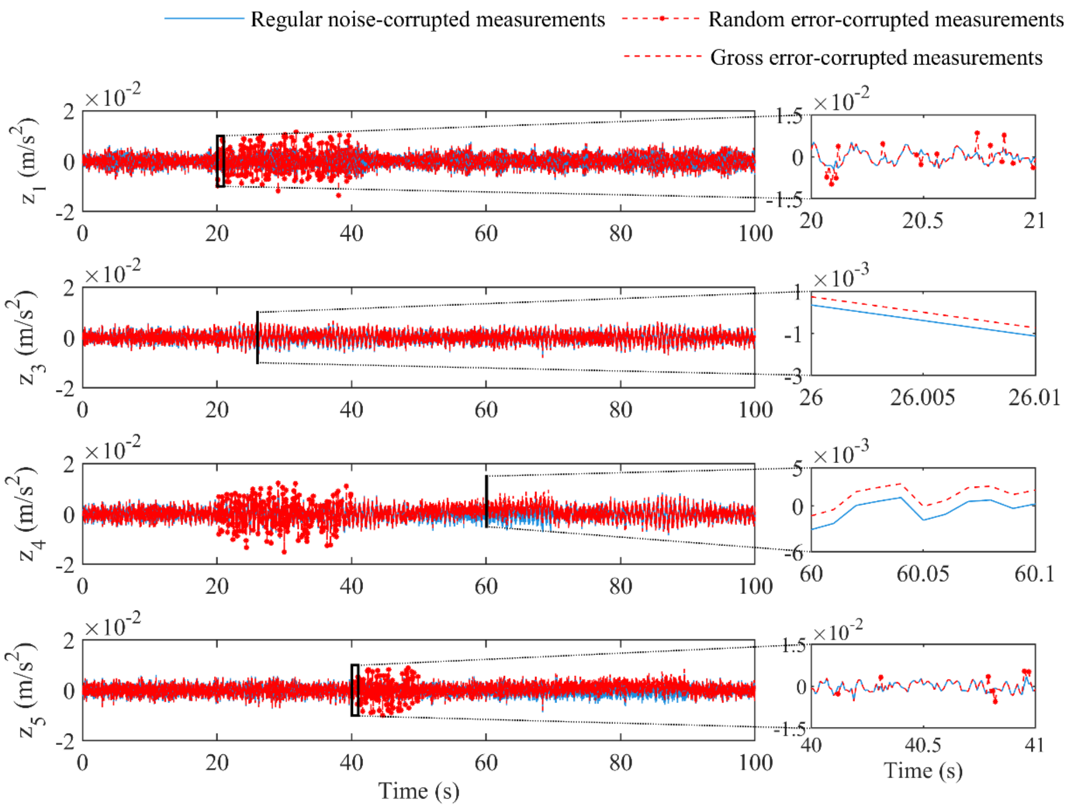

Fig. 5. Time-history noisy acceleration measurements at the 2nd, 6th, 7th, and 9th nodes, corrupted by random and gross errors (left column); zoomed-in views highlighting detailed anomaly-corrupted regions (right column).

The results of anomaly detection are first presented. The estimated values of $\boldsymbol{\mu}_k$ characterizing the mean value of gross errors are shown in Fig. 6. It can be observed that data points are closely distributed around the actual gross error values and remain uniformly distributed in the absence of gross errors. These results demonstrate the effectiveness of the proposed method in tracking gross errors. Fig. 7 shows the estimated values of $\boldsymbol{\ell}_k$ indicating the presence of anomalies. The results reveal that data points cluster around 1 in the presence of random errors, around 2 in the presence of gross errors, and around 0 when no measurement anomalies occur. Therefore, the estimated anomaly indicators can effectively identify both the presence and types of measurement anomalies. These findings confirm that the proposed method accurately detects multiple types of measurement anomalies.

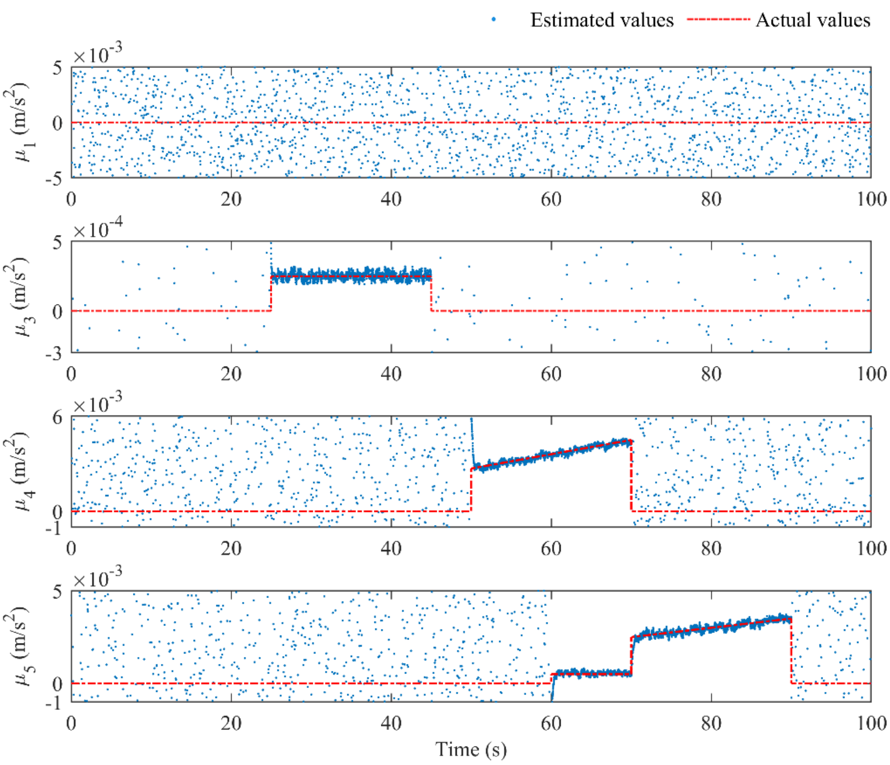

Fig. 6. Time-history estimates of $\boldsymbol{\mu}_k$ characterizing the mean value of gross errors.

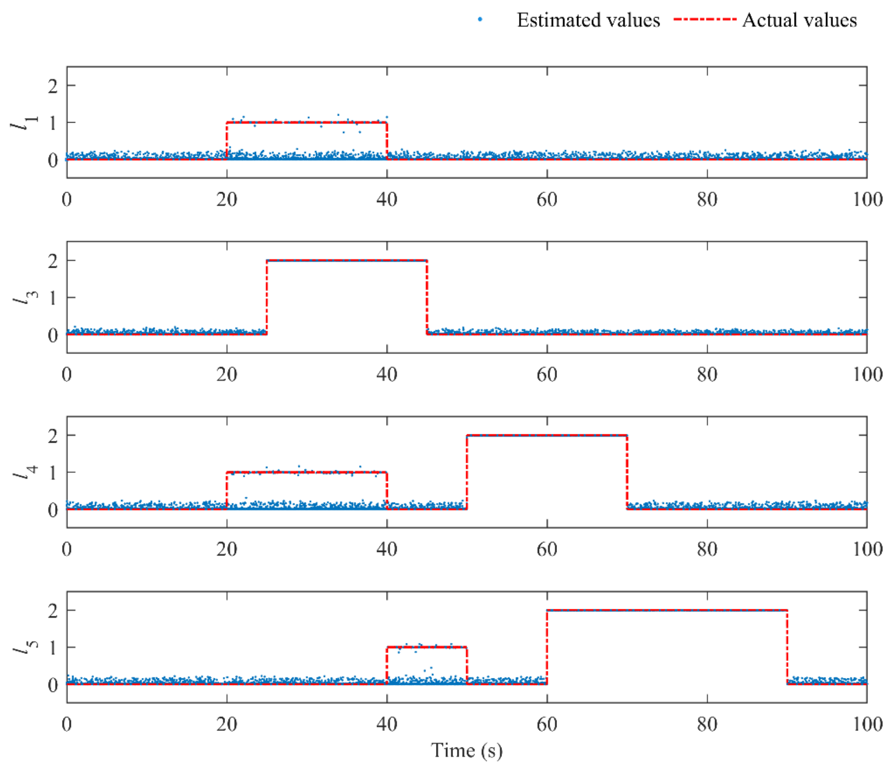

Fig. 7. Time-history estimates of $\boldsymbol{\ell}_k$ indicating the presence and types of measurement anomalies.

Two metrics, i.e., the false alarm rate and the miss detection rate, are employed to assess the effectiveness of anomaly detection. Table 2 lists the calculated false alarm rates and miss detection rates for the 12 measured DOFs. The four columns in the table represent the measurement channel, the corresponding location, and the two metrics. The results indicate that both metrics approach zero, demonstrating the excellent performance of the proposed approach in anomaly detection.

Table 2
Calculated false alarm rates and miss detection rates for assessing the effectiveness of the proposed method.

Channel	Location	False alarm rate	Miss detection rate
1	2nd	0.03	0.04
2	3rd	0.02	-
3	6th	0.00	0.00
4	7th	0.00	0.00
5	9th	0.00	0.00
6	10th	0.00	-
7	11th	0.00	-
8	12th	0.02	-
9	14th	0.01	-
10	15th	0.00	-
11	17th	0.02	-
12	18th	0.00	-

Two additional metrics, namely precision and recall, are used to evaluate detection performance for anomalies of small magnitude and specific types. Precision is defined as the number of true positives divided by the sum of true positives and false positives, while recall is defined as the number of true positives divided by the sum of true positives and false negatives. These metrics characterize the ability of the proposed method in detecting random and gross errors and the propensity for false alarms. Table 3 lists the calculated precision and recall for the four anomaly-corrupted channels. Specifically, the 1st and 2nd columns indicate the measurement channel and its corresponding location, respectively; the 3rd and 4th columns present the precision and recall for random errors, respectively; and the 5th and 6th columns list the precision and recall for gross errors, respectively. The results indicate that the proposed method achieves high precision and recall, suggesting a low tendency towards false alarms and missed detections in recognizing random errors. For gross errors, the method attains perfect performance, as both precision and recall are equal to one. This superior performance is attributed to the nature of gross errors which typically cause sustained deviations from actual measurement values, making them easier to detect than random errors.

Table 3
Calculated precision and recall for assessing the effectiveness of the proposed method.

Channel	Location	Random error		Gross error	
		Precision	Recall	Precision	Recall
1	2nd	0.97	0.96	-	-
3	6th	-	-	1.0	1.0
4	7th	0.96	0.98	1.0	1.0
5	9th	0.95	0.96	1.0	1.0

Fig. 8 presents the estimated stiffness parameters using the proposed method. As shown in Fig. 8, the stiffness parameters are effectively tracked, with the estimates falling within the 99.7% credible intervals. In contrast, Fig. 9 presents the stiffness parameter estimates using the traditional Bayesian inference [50]. The estimates exhibit significant fluctuations and divergence. This underscores the detrimental impact of anomalies on system identification and highlights the necessity of anomaly detection and removal. To quantitatively assess the effectiveness of system identification, the root mean square errors (RMSEs) between the estimated and actual states are calculated. Table 4 compares the RMSE values by using both methods. The notable reductions in prediction errors confirm the superior performance of the proposed approach.

In summary, the proposed method accurately detects multiple types of measurement anomalies while simultaneously enabling real-time system identification.

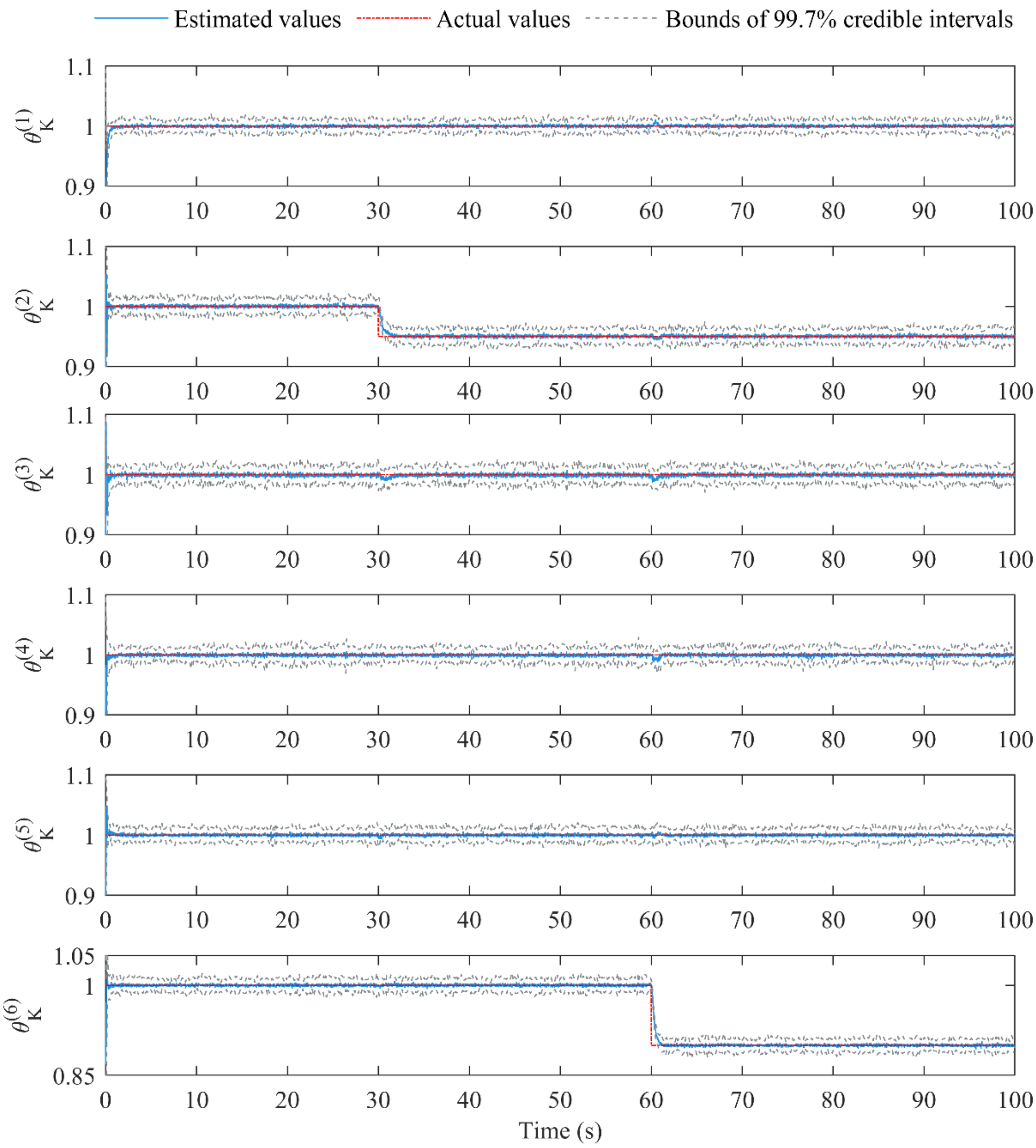

Fig. 8. Time-history estimates of the stiffness parameters obtained using the proposed method.

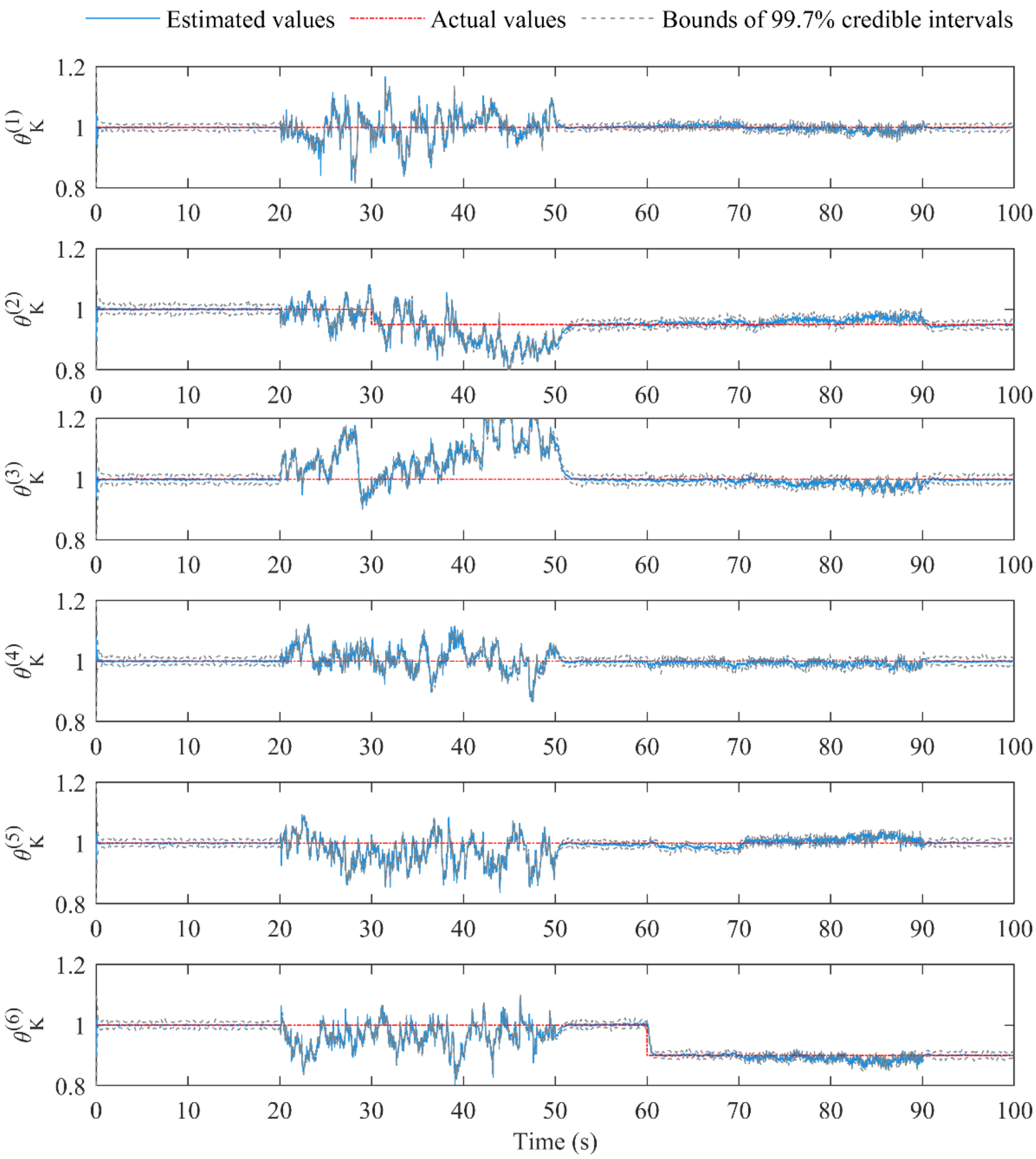

Fig. 9. Time-history estimates of the stiffness parameters obtained using the traditional method in [50].

Table 4
RMSEs of the estimated states using the proposed method and the traditional method in [50].

Estimated state	Location	Traditional method	Proposed method	RMSE reduction (%)
x_{v}	4th node	5.46×10^{-6}	1.08×10^{-6}	80.2
$\dot{x}_{\mathrm{h}}$		1.90×10^{-5}	5.18×10^{-6}	72.8
x_{v}	5th node	4.77×10^{-6}	1.06×10^{-6}	77.7
$\dot{x}_{\mathrm{h}}$		8.70×10^{-5}	6.95×10^{-6}	92.0
x_{v}	8th node	7.26×10^{-6}	1.77×10^{-6}	75.6
$\dot{x}_{\mathrm{h}}$		4.57×10^{-5}	2.49×10^{-6}	94.6

x_{v} and $\dot{x}_{\mathrm{h}}$ denote the vertical displacement and horizontal velocity, respectively.

4.2 Experimental case study: shear building

The proposed method is assessed in this section using a laboratory-scale aluminum structure introduced in [51]. The structure is modeled as a one-story shear building, with a plan dimension of 200 mm × 150 mm and a story height of 130 mm. Further details on the experimental setup are available in [51]. The mass parameter θ_M characterizes the mass matrix: $\mathbf{M} = \theta_M \mathbf{M}_0$, where $\mathbf{M}_0$ denotes the nominal mass matrix. Acceleration responses were measured using a triaxial piezoelectric accelerometer mounted at the center of the roof. The monitoring duration is 100 s with a sampling rate of 256 Hz. To consider time-varying properties, an experimental setup involving roof mass changes is implemented. Specifically, the initial roof mass is 1791 g, which is reduced to 1691 g at $t = 50$ s by removing a 100 g weight. As a result, the mass parameter θ_M decreases by 6% at $t = 50$ s. In addition, a 5% uniformly distributed error is introduced to the initial mass parameter. Acceleration responses along the x- and y- axes are analyzed representing the weak and strong directions of the structure, respectively. Fig. 10 shows the shear building model and the experimental setup.

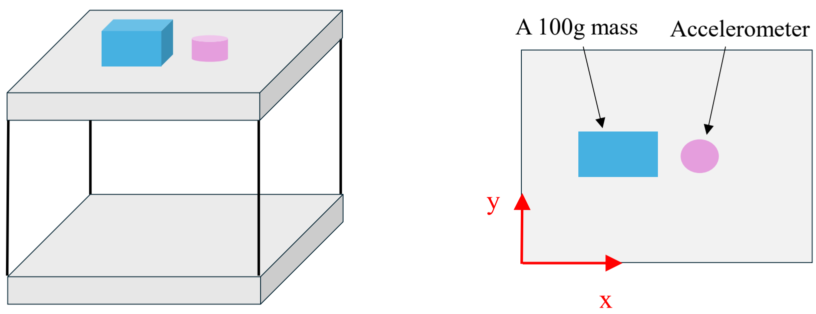

Fig. 10. Shear building model and experimental setup: side view (left), top view (right).

Fig. 11 presents the estimated values of $\boldsymbol{\mu}_k$ corresponding to the x- and y-direction channels. The estimated values remain uniformly distributed throughout the monitoring period, except near $t = 50$ s at which the sudden decrease of the roof mass occurred. This coincides well with the actual situation that no gross error is present in the measured values. Fig. 12 presents the estimated values of $\boldsymbol{\ell}_k$ for the two observed DOFs. As shown in Fig. 12, most data points cluster around 0 for both channels, while a small number of points cluster around 1 in the x-direction, indicating a few detected random errors for the raw measurements in the x-direction. These random error points may result from significant noise or transient structural disturbances. In addition, the data points are clustered around 2 at approximately $t = 50$ s, especially in the x-direction channel. These points, identified as gross errors, are attributed to transients induced by the removal of the weight. Table 5 displays the anomaly detection rates for the two testing channels. The first two columns present the measurement channel and its corresponding direction; the third and fourth columns present the anomaly detection rates for random and gross errors, respectively. The results show that only a small proportion of measurements in the x-direction channel are identified as random or gross errors, while no anomalies are detected in the y-direction channel.

Table 5
Calculated anomaly detection rates for assessing the effectiveness of the proposed method.

Channel	Direction	Anomaly detection rate	
		Random error	Gross error
1	x	0.01	0.02
2	y	0.00	0.00

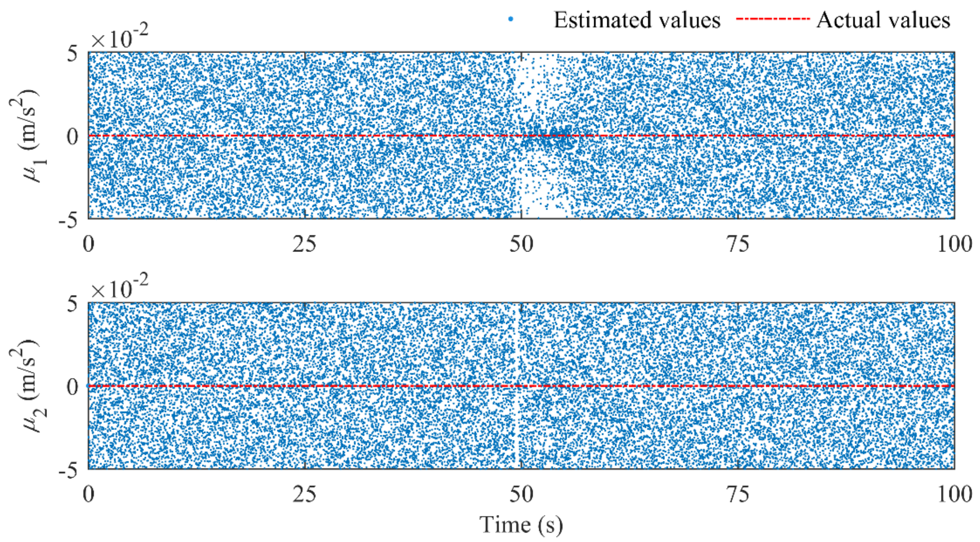

Fig. 11. Time-history estimates of $\boldsymbol{\mu}_k$ characterizing the mean value of gross errors.

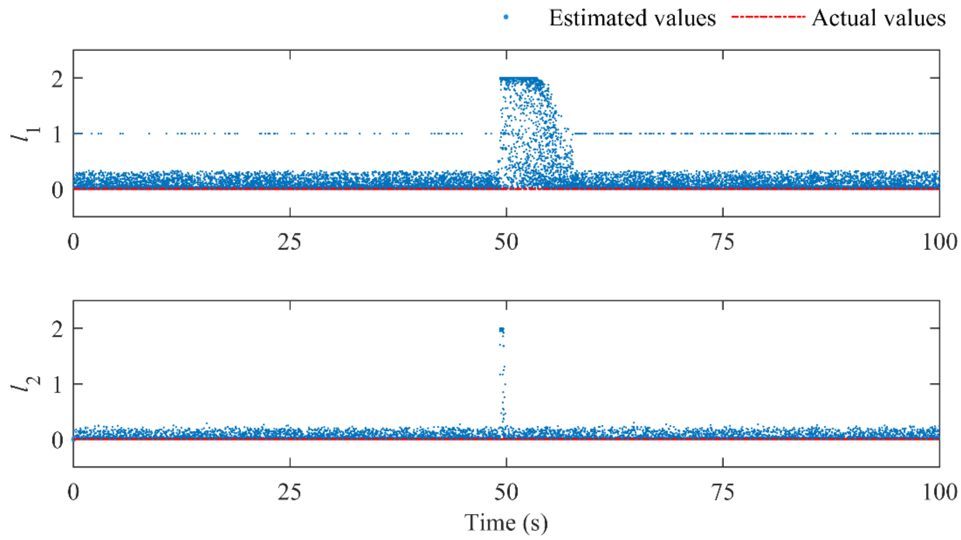

Fig. 12. Time-history estimates of $\boldsymbol{\ell}_k$ indicating the presence and types of measurement anomalies.

Fig. 13 (a) and (b) present the estimated mass parameter using the proposed method and the traditional approach in [50], respectively. The results demonstrate that the proposed method effectively tracks the mass parameter, with estimates remaining within the bounds of 99.7% credible intervals. In contrast, the traditional method provides highly fluctuated results when tracking the sudden decrease in mass, with inadequate credible intervals to capture posterior uncertainty. Additionally, Table 6 shows the RMSEs of the estimated mass parameter using both methods. The notably lower RMSE value obtained using the proposed method underscores its superior effectiveness compared to the traditional approach. Although no obvious measurement anomalies are observed in the raw data, the proposed method provides reliable estimation results for time-varying systems by identifying, removing, and compensating for substantial errors in the observed data. This process enhances the overall reliability and precision of system identification.

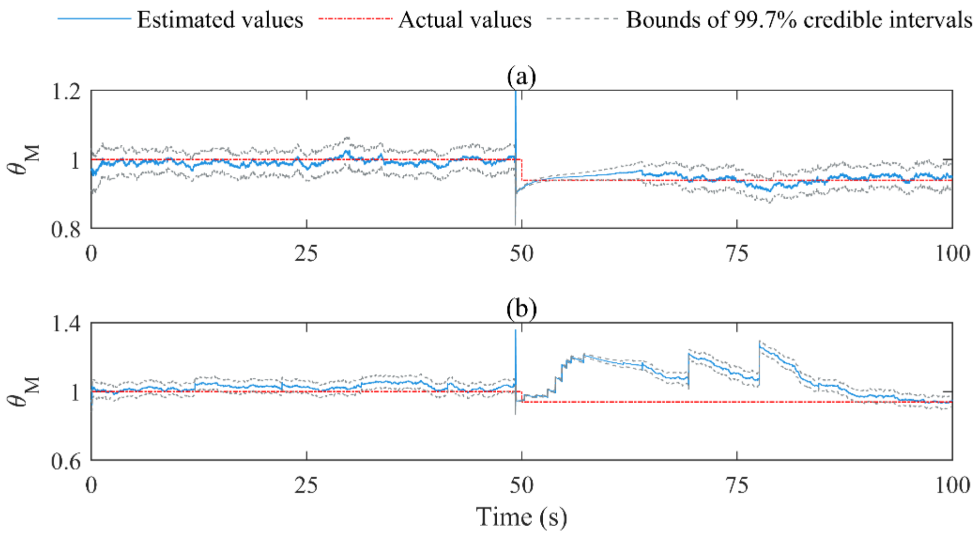

Fig. 13. Time-history estimates of structural mass parameters obtained using (a) the proposed method, and (b) the traditional method in [50].

Table 6
RMSEs of the estimated mass parameter using the proposed method and the traditional method in [50].

Estimated parameter	Traditional method	Proposed method	RMSE reduction (%)
θ_M	0.11	0.02	81.8

In summary, this experimental case study demonstrates the effectiveness of the proposed approach in real-time system identification by accurately detecting and addressing measurement anomalies.

5 Implementation in a full-scale monitoring system

This section utilizes field measurements from the Canton tower to evaluate the practicality of the proposed approach. Located in Guangzhou, China, the Canton tower is a 610 m, supertall structure, consisting of a 454 m main tower and a 156 m antenna mast. A comprehensive SHM system has been deployed to guarantee its stability and safety [52, 53]. This system utilizes uniaxial accelerometers to capture dynamic responses. These accelerometers were installed at different levels and oriented along both the long and short axes of the inner structure. Acceleration data from sixteen sensors, recorded between 18:00 and 19:00 on January 19, 2010, are analyzed in this study. The total monitoring duration is 400 s, with a 50 Hz sampling rate. The validation in Section 5 is based on actual field measurements.

The Canton tower is characterized as a three-dimensional cantilever beam using a reduced-order model comprising 185 DOFs. Further details about the Canton tower model can be found in [53]. The stiffness matrix is characterized using two parameters: $\mathbf{K} = \sum_{q=1}^{2} \theta_K^{(q)} \widehat{\mathbf{K}}^{(q)}$. Specifically, the model elements are divided into two clusters: the main tower with stiffness parameter $\theta_K^{(1)}$, and the antenna with stiffness parameter $\theta_K^{(2)}$. Both stiffness parameters were assigned reference values of 1, based on the model updating results in [54]. The initial stiffness parameters are taken with a 5% uniformly distributed error.

Fig. 14 shows the estimated values of $\boldsymbol{\mu}_k$ corresponding to the 11th and 12th measurement channels. The results show that the estimates remain uniformly distributed throughout the monitoring period. Fig. 15 presents the estimated values of $\boldsymbol{\ell}_k$ for the two observed DOFs. Only a small number of data points cluster around 1, while most data points cluster around 0. These results indicate that a few random errors were detected for the raw measurements. No gross errors were detected throughout the monitoring period. Table 7 presents the anomaly detection rates obtained using the proposed method. The first three columns present the sensor number, observation channel, and accelerometer orientation, respectively. The fourth and fifth columns present the anomaly detection rates for random errors and gross errors, respectively. Only a small proportion of measurements in each channel are detected as random errors. These anomalies may result from large measurement noise or structural disturbances.

Fig. 16 shows the stiffness parameter estimates using the proposed method. The estimates remain within the 99.7% credible intervals, indicating high accuracy. In contrast, Fig. 17 presents the stiffness parameter estimates using the traditional method in [50] based on raw measurements and they exhibit significantly greater fluctuations due to measurement anomalies. Moreover, Table 8 compares the RMSE values of the estimates using both methods. It is clear that the proposed method yields lower RMSE values compared to the traditional one, highlighting its superior performance. Again, although no obvious anomalies were observed in the raw measurements, the proposed method successfully identified and removed raw observations with substantial errors. This ensures reliable and accurate system identification, further validating the robustness of the proposed approach.

This application validates the proposed approach for real-time system identification in practical engineering scenarios. By effectively detecting and removing measurement anomalies, the approach ensures reliable and accurate identification results.

Table 7
Calculated anomaly detection rates for assessing the effectiveness of the proposed method.

Number	Channel	Orientation	Anomaly detection rate	
			Random error	Gross error
1	1st	y	0.01	0.00
2	2nd	x	0.01	0.00
3	3rd	y	0.03	0.00
4	4th	x	0.02	0.00
5	5th	y	0.01	0.00
6	6th	x	0.02	0.00
7	8th	y	0.04	0.00
8	9th	x	0.02	0.00
9	11th	y	0.03	0.00
10	12th	x	0.02	0.00
11	13th	y	0.03	0.00
12	14th	x	0.04	0.00
13	15th	y	0.01	0.00
14	16th	x	0.04	0.00
15	18th	y	0.02	0.00
16	19th	x	0.02	0.00

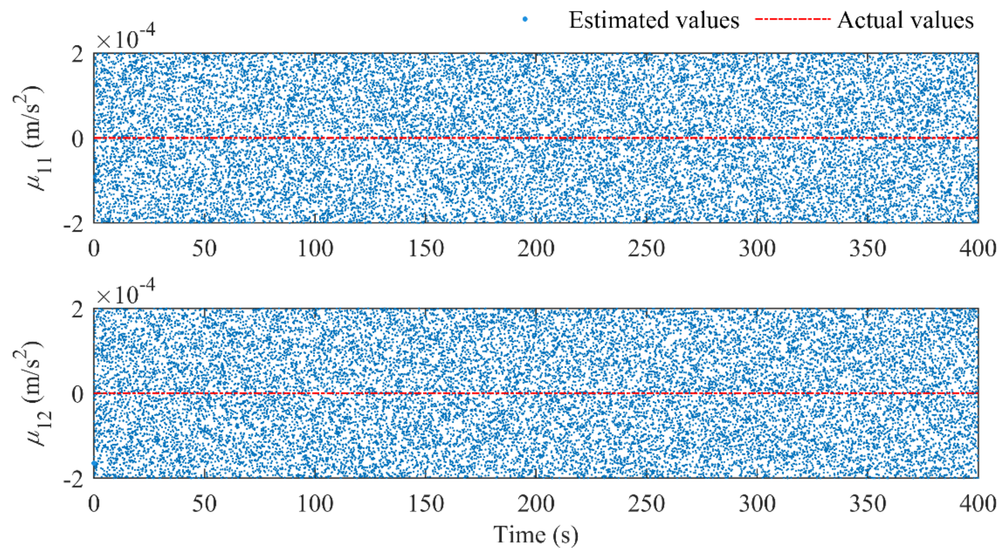

Fig. 14. Time-history estimates of $\boldsymbol{\mu}_k$ characterizing the mean value of gross errors.

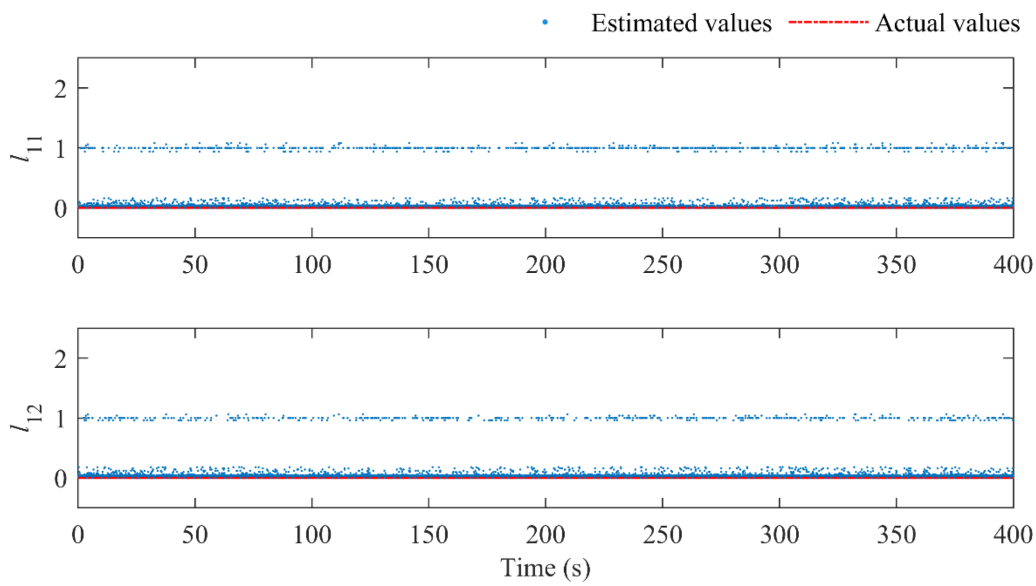

Fig. 15. Time-history estimates of $\boldsymbol{\ell}_k$ indicating the presence and types of measurement anomalies.

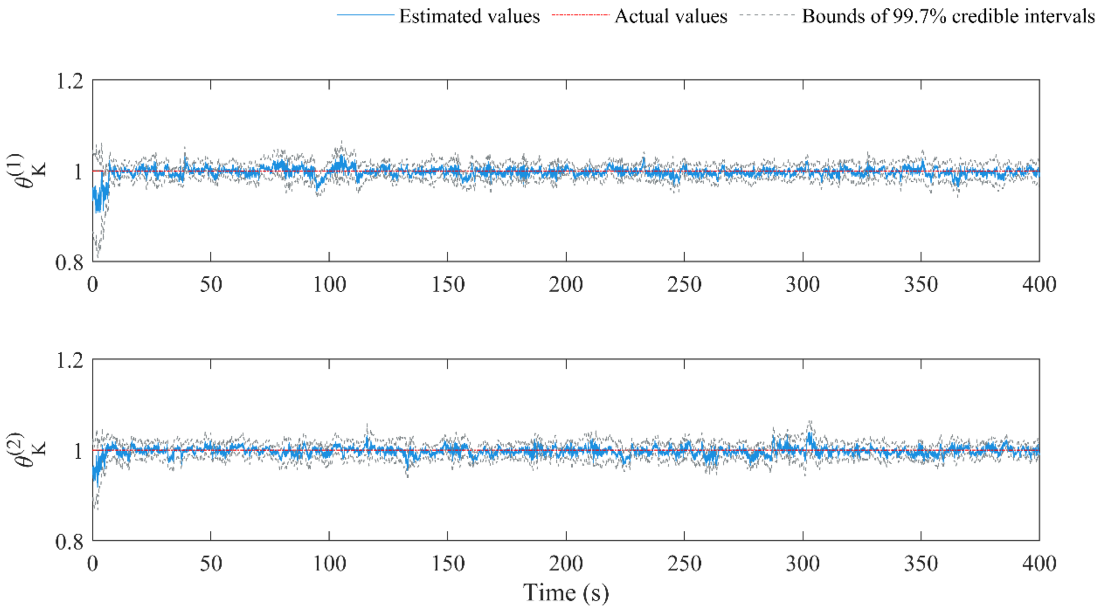

Fig. 16. Time-history estimates of the stiffness parameters obtained using the proposed method.

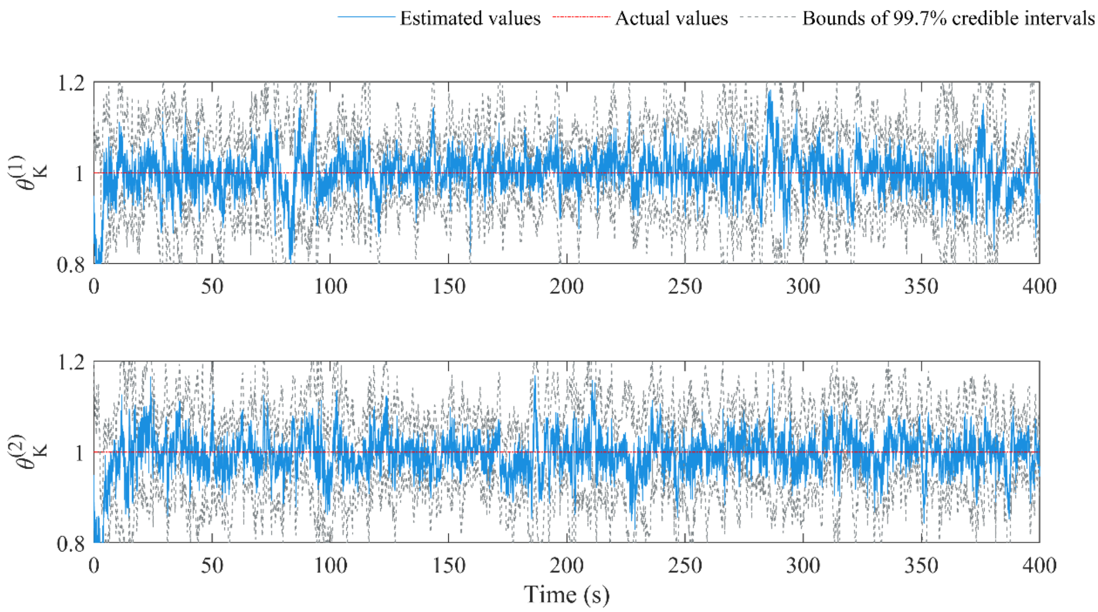

Fig. 17. Time-history estimates of the stiffness parameters obtained using the traditional method in [50].

Table 8
RMSEs of the estimated stiffness parameters using the proposed method and the traditional method in [50].

Estimated parameter	Traditional method	Proposed method	RMSE reduction (%)
$\theta_K^{(1)}$	0.0585	0.0146	75.0
$\theta_K^{(2)}$	0.0467	0.0121	74.1

6 Conclusion

This study proposes a novel approach for real-time anomaly detection and system identification. The approach tackles the challenge of performing real-time system identification under multi-type abnormal measurements. The detection of typical measurement anomalies is realized by updating Bernoulli random vectors that indicate the presence of anomalies. The detected anomalies are excluded or compensated in the subsequent system identification process. Notably, the proposed approach requires no prior information on anomaly characteristics. An adaptive Bayesian scheme is employed to recursively update both the Bernoulli and model parameters, enabling real-time simultaneous anomaly detection and system identification. The proposed approach is validated using two time-varying structural systems covering a numerical bridge and a laboratory shear building. The results demonstrate reliable and consistent performance in identifying both anomalies and system parameters in a real-time manner. The proposed method is applied to field measurements from the Canton tower monitoring system, demonstrating promising results and indicating its potential practicality. While these findings are encouraging, further studies involving diverse structural types and complex operational environments are necessary to comprehensively assess the general applicability of the proposed method in real-world SHM scenarios.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is financially supported by the National Science Fund for Distinguished Young Scholars (Grant No. 52425802), the National Natural Science Foundation of China (Grant No. 52008037), the Natural Science Foundation of Hunan Province, China (Grant No. 2024JJ5026), the Science and Technology Development Fund, Macau SAR (Grant 001/SKL/2024), University of Macau (CPG2025-00025-FST), and Guangdong-Hong Kong-Macau Joint Laboratory Program (Project no. 2020B1212030009). These generous supports are gratefully acknowledged.

References

[1] R.E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng. 82 (1960) 35–45.

[2] R. Astroza, H. Ebrahimian, J.P. Conte, Performance comparison of Kalman-based filters for nonlinear structural finite element model updating, J. Sound Vib. 438 (2019) 520–542.

[3] K. Huang, K.V. Yuen, Bayesian real-time system identification: from centralized to distributed approach, Springer, Singapore, 2023.

[4] G. Mzyk, Identification of time-varying non-linear systems with adaptive bootstrap-based tracking, Mech. Syst. Signal Proc. 223 (2025) 111896.

[5] D. Hajializadeh, Deep learning-based indirect bridge damage identification system, Struct. Health Monit. 22 (2) (2022) 897–912.

[6] S. Mammeri, B Barros, B. Conde-Carnero, B. Riveiro, From traditional damage detection methods to physics-informed machine learning in bridges: A review, Eng. Struct. 330 (2025) 119862.

[7] M.M. Azad, S. Kim, Y.B. Cheon, H.S. Kim, Intelligent structural health monitoring of composite structures using machine learning, deep learning, and transfer learning: A review, Adv. Compos. Mater. 33 (2) (2023) 162–188.

[8] K. Sepahvand, S. Marburg, Identification of composite uncertain material parameters from experimental modal data, Probabilist. Eng. Mech. 37 (2014) 148–153.

[9] K.A. Hoppe, M.G.T. Kronthaler, K. Sepahvand, S. Marburg, Identification of a cantilever beam's spatially uncertain stiffness, Sci. Rep. 13 (2023) 1169.

[10] C.A. Gaviria, L.A. Montejo, Optimal wavelet parameters for system identification of civil engineering structures, Earthq. Spectra 34 (1) (2018) 197–216.

[11] M. Maeder, S. Chandra, S. Marburg, Identification of temperature dependent material properties in composite plates utilizing experimental vibration data. In: H. Altenbach, M. Beitelschmidt, M. Kästner, K. Naumenko, T. Wallmersperger (Eds.), Material Modeling and Structural Mechanics, Advanced Structured Materials, Springer, Cham, 2022, pp. 115–134.

[12] C.W. Zhang, A.A. Mousavi, S.F. Masri, G. Gholipour, K. Yan, X.L. Li, Vibration feature extraction using signal processing techniques for structural health monitoring: A review, Mech. Syst. Signal Proc. 177 (2022) 109175.

[13] K. Huang, K.V. Yuen, Y.F. Ma, L. Wang, Online operational modal decomposition, Mech. Syst. Signal Proc. 214 (2024) 111376.

[14] M. Civera, M.L. Pecorelli, R. Ceravolo, C. Surance, L.Z. Fragonara, A multi-objective genetic algorithm strategy for robust optimal sensor placement, Comput.-Aided Civil Infrastruct. Eng. 36 (9) (2021) 12646.

[15] L. Eichner, R. Schneider, M. Baeßler, Optimal vibration sensor placement for jacket support structures of offshore wind turbines based on value of information analysis, Ocean Eng. 288 (2023) 115407.

[16] S. Raorane, T. Ercan, C. Papadimitriou, P. Packo, T. Uhl, Bayesian optimal sensor placement for acoustic

emission source localization with clusters of sensors in isotropic plates, Mech. Syst. Signal Proc. 214 (2024) 111342.
[17] D.S. Aulakh, S. Bhalla, Piezo sensor based multiple damage detection under output only structural identification using strain modal flexibility, Mech. Syst. Signal Proc. 194 (2023) 110272.
[18] B. Ahmed, S. Mangalathu, J.S. Jeon, Unveiling out-of-distribution data for reliable structural damage assessment in earthquake emergency situations, Automat. Constr. 156 (2023) 105142.
[19] F. Brighenti, M.F. Bado, F. Romeo, D. Zonta, Forecasting bridge damage within a predictive structural reliability-based DSS, Automat. Constr. 168 (2024) 105740.
[20] M. Kim, J. Song, C.W. Kim, Near-real-time damage identification under vehicle loads using dynamic graph neural network based on proper orthogonal decomposition, Mech. Syst. Signal Proc. 224 (2025) 112175.
[21] L. Hofer, K. Toska, M.A. Zanini, Impact of epistemic and aleatory uncertainties on the seismic reliability assessment of existing structures, Structures 57 (2023) 105235
[22] A. Shehadeh, O. Alshboul, E. Saleh, Enhancing safety and reliability in multistory construction: a multi-state system assessment of shoring/reshoring operations using interval-valued belief functions, Reliab. Eng. Syst. Safe. 252 (2024) 110458.
[23] M.A. Aladsani, H.V. Burton, Reliability-based quantification of the benefits of machine learning predictive models in seismic structural design and performance assessment, Struct. Saf. 115 (2025) 102596.
[24] A. Rovera, A. Tancau, N. Boetti, M.D.L. Dalla Vedova, P. Maggiore, D. Janner, Fiber optic sensors for harsh and high radiation environments in aerospace applications, Sensors 23 (5) (2023) 2512.
[25] R. Foorthuis, On the nature and types of anomalies: A review of deviations in data, Int. J. Data Sci. Anal. 12 (2021) 297–331.
[26] Y. Fu, C. Peng, F. Gomez, Y. Narazaki, B.F. Spencer, Sensor fault management techniques for wireless smart sensor networks in structural health monitoring, Struct. Control. Health Monit. 26 (7) (2019) e2362.
[27] P. Cheema, M.M. Alamdari, G. Vio, L. Azizi, S. Luo, On the use of matrix profiles and optimal transport theory for multivariate time series anomaly detection within structural health monitoring, Mech. Syst. Signal Proc. 204 (2023) 110797.
[28] G. Park, A.C. Rutherford, H. Sohn, C.R. Farrar, An outlier analysis framework for impedance-based structural health monitoring, J. Sound Vib. 286 (1–2) (2005) 229–250.
[29] S. Baek, D.Y. Kim, Abrupt variance and discernibility analyses of multi-sensor signals for fault pattern extraction, Comput. Ind. Eng. 128 (2019) 999–1007.
[30] Z. Uddin, A. Qamar, F. Alam, ICA based sensors fault diagnosis: an audio separation application, Wireless Pers. Commun. 118 (4) (2021) 3369–3384.
[31] M. Gul, F.N. Catbas, Statistical pattern recognition for structural health monitoring using time series modeling: theory and experimental verifications, Mech. Syst. Signal Proc. 23 (7) (2009) 2192–2204.
[32] N. Dervilis, E.J. Cross, R.J. Barthorpe, K. Worden, Robust methods of inclusive outlier analysis for structural health monitoring, J. Sound Vib. 333 (20) (2014) 5181–5195.
[33] Y. Yang, S. Nagarajaiah, Blind denoising of structural vibration responses with outliers via principal component pursuit, Struct. Control. Health Monit. 21 (6) (2014) 962–978.
[34] F.A. Mazarbhuiya, M. Shenify, A mixed clustering approach for real-time anomaly detection, Appl. Sci-Basel 13 (7) (2023) 4151.

[35] H. Sarmadi, A. Entezami, F. Magalhaes, Unsupervised data normalization for continuous dynamic monitoring by an innovative hybrid feature weighting-selection algorithm and natural nearest neighbor searching, Struct. Health Monit. 22 (6) (2023) 4005–4026.

[36] S. Lee, J. Lee, M. Kim, S. Lee, Y.J. Lee, Deep learning-based anomaly detection in acceleration data of long-span cable-stayed bridges, Smart Struct. Syst. 33 (2) (2024) 93–103.

[37] X. Jiang, B. Stephen, S. McArthur, A sequential Bayesian approach to online power quality anomaly segmentation, IEEE T. Ind. Inform. 17(4) (2021): 2675-2685.

[38] V. Judalet, S. Glaser, D. Gruyer, S. Mammar, Fault detection and isolation via the interacting multiple model approach applied to drive-by-wire vehicles, Sensors 18(7) (2018): 2332.

[39] L.H. Nguyen, J.A. Goulet, Anomaly detection with the switching Kalman filter for structural health monitoring, Struct. Control Health Monit. 25(4) (2018): e2136.

[40] H.B. Huang, T.H. Yi, H.N. Li, Sensor fault diagnosis for structural health monitoring based on statistical hypothesis test and missing variable approach, J. Aerospace Eng. 30(2) (2017): B4015003.

[41] K.V. Yuen, H.Q. Mu, A novel probabilistic method for robust parametric identification and outlier detection, Probabilist. Eng. Mech. 30 (2012) 48–59.

[42] J.D.A. Santos, G.A. Barreto, An outlies-robust kernel RLS algorithm for nonlinear system identification, Nonlinear Dynam. 90 (2017) 1707–1726.

[43] A.C. Hussain, T. Muhammad, U. Momin, A robust Bayesian approach for online filtering in the presence of contaminated observations, IEEE Trans. Instrum. Meas. 70, (2021) 1–15.

[44] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, J. Eng. Mech. 124(4) (1998): 455-461.

[45] K.V. Yuen, Bayesian Methods for Structural Dynamics and Civil Engineering, John Wiley & Sons, 2010.

[46] K.V. Yuen, L. Dong, Real-time system identification using hierarchical interhealing model classes, Struct. Control. Health Monit. 27 (2020) e2628.

[47] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proc. F 140 (1993) 107–113.

[48] E.A. Wan, R. Van Der Merwe, The unscented Kalman filter for nonlinear estimation, in: Proc. IEEE 2000 Adapt. Syst. Signal Proc. Commun. Control Symp. (Cat. No.00EX373), IEEE, Lake Louise, AB, Canada, 2000.

[49] H.W. Sorenson, J.E. Sacks, Recursive fading memory filtering, Inform. Sciences 3 (2) (1971) 101–119.

[50] G. Capellari, S. Eftekhar Azam, S. Mariani, Damage detection in flexible plates through reduced-order modeling and hybrid particle-Kalman filtering, Sensors 16 (1) (2016) 2.

[51] Y.P. Yang, Z. Zhu, S.K. Au, Tracking time-varying properties using quasi time-invariant models with Bayesian dynamic programming, Mech. Syst. Signal Proc. 223 (2025) 111546.

[52] Y.Q. Ni, Y. Xia, W.Y. Liao, J.M. Ko, Technology innovation in developing the structural health monitoring system for Guangzhou New TV Tower, Struct. Control Health Monit. 16 (1) (2009) 73–98.

[53] Y. Xia, Y.Q. Ni, J.M. Ko, H.B. Chen, Development of a structural health monitoring benchmark problem for high rise slender structures, Adv. Sci. Technol. 56 (2008) 489–94.

[54] S.C. Kuok, K.V. Yuen, Structural health monitoring of Canton tower using Bayesian framework, Smart Struct. Syst. 10(4-5) (2012): 375-391.