Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Robust adaptive Kalman filter for structural health monitoring

Shenglun Yi*[1] | Tingli Su[2] | ZhenYun Tang[3]

[1]Department of Information Engineering, University of Padova, Padova, Italy
[2]School of Artificial Intelligence, Beijing Technology and Business University, Beijing, China
[3]The Key Laboratory of Urban Security and Disaster Engineering, Ministry of Education, Beijing University of Technology, Beijing, China

Correspondence
*Shenglun Yi, Department of Information Engineering, University of Padova, via Gradenigo 6/B, 35131 Padova, Italy. Email: shenglun@dei.unipd.it

Summary

Health monitoring is critical for the maintenance and risk management of reinforced concrete (RC) structures. In this paper, a robust adaptive Kalman filter is proposed for an interstory drift estimation problem to show the health condition of RC structures in the case that the statistics or internal dynamics describing the signals and measurements are not known precisely. More precisely, we build an adaptive current Jerk model (ACJM) where the model parameters are updated in each time step to presuppose the statistics characterization of the RC dynamic, while the unknown measurement noise covariance is adapted based on a fixed-lag innovation with respect to measurements. Moreover, a robust adaptive Kalman filter is designed for the modeling mismatch in each time increment by solving a minimax game: one "hostile" player tries to select a worst model far from the proposed ACJM with an exponential decay tolerance, while an optimum filter is designed by minimizing the estimation error according to this worst model. Finally, some simulation and experimental results show the effectiveness of the proposed algorithm.

KEYWORDS:
Interstory drift estimation, robust Kalman filtering, adaptive current Jeck model, noise covariance estimation, structural health monitoring

1 | INTRODUCTION

Structural health monitoring (SHM) is an ongoing process involving continuous observation and analysis of engineering structures, e.g., RC structures, over time[1,2]. Recent severe earthquakes, such as the 2019 Peru earthquake and the 2021 Haiti earthquake, have highlighted the application of SHM, particularly for RC structure maintenance and risk management. Over the last two decades, in order to assess the RC structural health and detect its potential structural damage or degradation, various SHM approaches have been proposed, including model-based, data-based, and performance-based approaches[3]. One emerging SHM approach is performance-based structural health monitoring (PSHM)[4], which integrates tools from performance-based earthquake engineering to establish correlations between structural damage states and engineering demand parameters (EDPs). Among the various EDPs, the interstory drift ratio (IDR)[5], defined as the interstory drift (ID) normalized by the corresponding story height, serves as a universal indicator for evaluating structural performance against seismic actions. Here, ID quantifies the relative translational displacement between consecutive floors in multistory buildings. Studies have shown that IDR exhibits a better correlation with observed damage compared to other response quantities, such as peak floor acceleration[6]. However, the raw measurements of ID are extraordinarily noisy and fail to work, which motivates researchers to design more sophisticated estimation approaches to extract useful information from the noisy sensor signal[4,7,8].

Dynamic state estimation (DSE) approaches, such as Kalman filtering (KF) and KF-like filtering, have been extensively used for the estimation problem[9,2]. The estimation task requires a nominal dynamic model to characterize the statistics or internal

dynamics of signals. Several dynamic models, such as constant velocity (CV) model[10], constant acceleration (CA) model[11], Singer model[12], and Jerk model[13], have been proposed over the past 35 years for the typical tracking scenarios where the prior knowledge of such dynamic models is obtained through numerous repeated experiments. Indeed, it is reasonable to apply the tracking dynamic model when estimating the relative displacement, i.e. ID, because the motion of the RC structure in response to the seismic activity also aligns with the inertial law. However, it may lead to poor performances in the seismic test due to practical limitations including the inaccuracy of dynamic models and the complexity of systems. To address such a weakness, adaptive state-space models[14,15] that analyze the dynamic behavior of the system state have been explored to identify the model parameters. Later on, it has been proved that such paradigms can also be used to estimate the RC structural damage through a quasi-static test when the structural damage states do not change significantly over time[16].

In this paper, we focus on estimating the ID in a four-layer RC structure[17] subjected to a structural seismic test. The seismic waves are simulated using El Centro waves[18] from the United States. Unlike the slight structural vibrations observed in the quasi-static test, our seismic test exposes the RC structure to severe vibrations. To capture such severe vibrations, the dynamic characteristics are described using the Jerk model[13], which employs a four-dimensional state space model with state components: displacement, velocity, acceleration, and Jerk. Accordingly, drawing inspiration from the philosophy of adaptive models in References 14,15,16, we use the Yule-Walker algorithm to estimate the unknown parameters in this Jerk model. Moreover, considering that the vibrations of the RC structure should be confined to a range so that it does not exceed its original length, the "current model concept"[19,20] is introduced. In this way, an adaptive current Jerk model (ACJM) is proposed to characterize the dynamic behavior of the RC structure under the seismic testing in the case that the state, i.e. Jerk, in the next instant is around the "current" state, and its corresponding model parameters are adapted in each time increment.

We consider the ACJM to presuppose the statistics characterization of the RC dynamic, however, one is faced with a lack of statistical knowledge of the measurement noise process. To estimate the measurement noise, one possible way has been also investigated in References 14 and 15, however, under the assumption that the measurement noise is zero-mean, white Gaussian, and mutually independent. An alternative approach involves adopting a data-driven approach, wherein the relationship between raw inertial measurements and latent variables is learned. Such data-driven approaches have been widely studied in denoising, tracking, and monitoring applications[21,22]. In this context, a joint estimation technique has been introduced to simultaneously compute model parameters and state vectors. A parameter identification method based on the innovation with respect to measurements is responsible for calibrating the unknown noise parameters, while the primary interest lies in estimating the state trajectory.

In addition, it is worth noting that model uncertainties might arise in the aforementioned approaches since both the model and noise parameters are estimated from measurements[23,24]. Up to this point, some robust versions based on the KF or KF-like algorithms have been developed to solve this problem. These techniques assume a nominal dynamic model, nominal noise processes, and initialization (referred to collectively as the nominal model) have already been estimated from data, but in practice, there might be a mismatch between the actual model and this identified nominal model. Several robust Kalman filters (RKF)[25,26,27] have been proposed recently, taking into account the modeling mismatch. In particular, risk-sensitive filters utilize an exponential quadratic loss function, imposing a more severe penalty for large errors compared to the classical KF with the standard quadratic loss function[28]. Hereafter, manifold robust versions[29,24], which are the continuation of the aforementioned work, have been provided in recent years. For instance, Levy and Nikoukhah[30] presented a robust least-squares estimation method, reformulating risk-sensitive filters into a minimax problem. They further refined their approach into a robust state-space filter[31], which incorporates a separate constraint to counter incremental model perturbations. Then, more extensions of such a paradigm have been presented, for example, the case with the Tau-divergence family[32], the case with degenerate densities[33,23] and the case for the nonlinear system[18]. The aforementioned literatures address the robust estimation problem under the assumption that the actual model is within a ball centered about the nominal model, where the ball is specified by a Kullback-Leibler (KL) tolerance (or other ambiguity sets).

In this paper, we design a robust adaptive Kalman filter (RAKF) for an ID estimation problem under the seismic testing to show the structural health condition of RC in the case that the statistics or internal dynamics describing the signals and measurements are not known. In particular, the contributions of this paper are listed as follows:

1. The dynamic behavior of the RC structure in the seismic test using El Centro waves is reanalyzed in each time step, which results in the prescribed process dynamic model, named ACJM. The model captures the dynamic characteristics of the severe RC structural vibrations while restricting its estimated ID to a valid range.

2. The measurement noise parameter is updated in each time inclement via a fixed-lag innovation with respect to the measurements.

3. A robust Kalman-like iteration is designed for the modeling mismatch by solving a minimax game in each time increment where the tolerance is tuned in an exponential decay form.

The outline of the paper is as follows. Section 2 shows the adaptive current Jerk model construction including the process model parameters estimation and the measurement noise covariance estimation. In Section 3, we derive the robust adaptive Kalman filter. The simulation study is discussed in Section 4. The experimental results are presented in Section 5. Finally, Section 6 concludes the paper.

2 | ADAPTIVE CURRENT JERK MODEL

We consider a nominal discrete-time state space model of the form

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{1a}$$

$$y_k = C_k x_k + v_k \tag{1b}$$

where $A_k \in \mathbb{R}^{n\times n}$, $B_k \in \mathbb{R}^{n\times p}$, $C_k \in \mathbb{R}^{m\times n}$, $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^p$ is the input vector, and $y_k \in \mathbb{R}^m$ is the observation vector. Accordingly, the process noise and the measurement noise are treated as independent Gaussian noise sequences with time-varying covariances, i.e., $w_k \sim \mathcal{N}(0, Q_k)$ and $v_k \sim \mathcal{N}(0, R_k)$, respectively.

In this paper, measurement data collected from the sensors is the relative translational displacement, i.e. ID, between consecutive floors in multistory buildings shown in Fig. 3(a). Since the RC structure is under seismic vibration, its dynamic characteristics are generally non-uniformly accelerated, but can not be too far away from the current state. Thus, drawn inspiration from the Jerk model[13] and current model[19,20], a current Jerk (CJ) model is proposed to describe the dynamic characteristic of the RC structure. More precisely, the state vector is defined as $x := \left[\mathtt{x}_p, \mathtt{x}_v, \mathtt{x}_a, \mathtt{x}_j\right]^\top$ where $\mathtt{x}_p$, $\mathtt{x}_v$ and $\mathtt{x}_a$ denote the position, velocity and acceleration of the object, respectively, as well as $\mathtt{x}_j$ is the Jerk as shown in Reference 13. Our process model can be written as (1a) with

$$A_k = \begin{bmatrix} 1 & T & \frac{T^2}{2} & a_{1,k} \\ 0 & 1 & T & a_{2,k} \\ 0 & 0 & 1 & a_{3,k} \\ 0 & 0 & 0 & a_{4,k} \end{bmatrix}$$

where

$$\begin{aligned} a_{1,k} &= \frac{2 - 2\alpha_k T + \alpha_k^2 T^2 - 2e^{-\alpha_k T}}{2\alpha_k^3} \\ a_{2,k} &= \frac{e^{-\alpha_k T} - 1 + \alpha_k T}{\alpha_k^2} \\ a_{3,k} &= \frac{1 - e^{-\alpha_k T}}{\alpha_k} \\ a_{4,k} &= e^{-\alpha_k T}, \end{aligned}$$

and T denotes the sampling period; the unknown parameter α_k is the reciprocal of the Jerk time constant in each time increment k of the model. In other words, the higher correlation parameter α_k permits the more rapidly fluctuating Jerk of the object. Next,

$$B_k = \begin{bmatrix} \frac{6}{\alpha_k^3}(\alpha_k^3 T^3 - 3\alpha_k^2 - 6 + 6\alpha_k T + 6e^{-\alpha_k T}) \\ \frac{1}{\alpha_k}\left(-T + \frac{\alpha_k T^2}{2} + \frac{1-e^{-\alpha_k T}}{\alpha_k}\right) \\ T - \frac{1-e^{-\alpha_k T}}{\alpha_k} \\ 1 - e^{-\alpha_k T} \end{bmatrix}.$$

Note that the input vector u_k is defined as a mean value of the Jerk, which is considered as a constant in each sampling period so that what is of concern is only the "current" probability density of Jerk, i.e. $\bar{\mathtt{x}}_{j,k}$, see also in Reference 19 where it first proposed a "current model" concept to describe the statistical distribution of acceleration in the continuous-time state space model. Here, we extend it to the fourth-order discrete-time state space model, so called CJ model. Moreover, the process noise covariance Q_k

$$
\begin{aligned}
Q_k &= 2\alpha_k\sigma_k^2 \begin{bmatrix} q_{11,k} & q_{12,k} & q_{13,k} & q_{14,k} \\ q_{12,k} & q_{22,k} & q_{23,k} & q_{24,k} \\ q_{13,k} & q_{23,k} & q_{33,k} & q_{34,k} \\ q_{14,k} & q_{24,k} & q_{34,k} & q_{44,k} \end{bmatrix} \\
q_{11,k} &= \frac{1}{2\alpha_k^7}\left[\frac{\alpha_k^5T^5}{10} - \frac{\alpha_k^4T^4}{2} + \frac{4\alpha_k^3T^3}{3} - 2\alpha_k^2T^2 + 2\alpha_kT - 3 + 4e^{-\alpha_kT} + 2\alpha_k^2T^2e^{-\alpha_kT} - e^{-2\alpha_kT}\right] \\
q_{12,k} &= \frac{1}{2\alpha_k^6}\left[1 - 2\alpha_kT + 2\alpha_k^2T^2 - \alpha_k^3T^3 + \frac{\alpha_k^4T^4}{4} + e^{-2\alpha_kT} + 2\alpha_kTe^{-\alpha_kT} - 2e^{-\alpha_kT} - \alpha_k^2T^2e^{-\alpha_kT}\right] \\
q_{13,k} &= \frac{1}{2\alpha_k^5}\left[2\alpha_kT - \alpha_k^2T^2 - \frac{\alpha_k^3T^3}{3} - 3 - 2e^{-2\alpha_kT} + 4e^{-\alpha_kT} + \alpha_k^2T^2e^{-\alpha_kT}\right] \\
q_{14,k} &= \frac{1}{2\alpha_k^4}\left[1 + e^{-2\alpha_kT} - 2e^{-\alpha_kT} - 2\alpha_k^2T^2e^{-\alpha_kT}\right] \\
q_{22,k} &= \frac{1}{2\alpha_k^5}\left[1 - e^{-2\alpha_kT} + \frac{2\alpha_k^3T^3}{3} + 2\alpha_kT - 2\alpha_k^2T^2 - 4\alpha_kTe^{-\alpha_kT}\right] \\
q_{23,k} &= \frac{1}{2\alpha_k^4}\left[1 + \alpha_k^2T^2 - 2\alpha_kT + 2\alpha_kTe^{-\alpha_kT} + e^{-2\alpha_kT} - 2e^{-\alpha_kT}\right] \\
q_{24,k} &= \frac{1}{2\alpha_k^3}\left[1 - e^{-2\alpha_kT} - 2\alpha_kTe^{-2\alpha_kT}\right] \\
q_{33,k} &= \frac{1}{2\alpha_k^3}\left[4e^{-\alpha_kT} - e^{-2\alpha_kT} + 2\alpha_kT - 3\right] \\
q_{34,k} &= \frac{1}{2\alpha_k^2}\left[e^{-2\alpha_kT} + 1 - 2\alpha_kT\right] \\
q_{44,k} &= \frac{1}{2\alpha_k}\left[1 - e^{-2\alpha_kT}\right].
\end{aligned} \tag{2}
$$

is shown in the equation (2) where σ_k^2 denotes the unknown variance of the jerk noise. In addition, for the measurement model (1b), we have $C_k = C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$. However, the measurement noise covariance R_k is not known in our practical application. Accordingly, the output y_k denotes the noisy ID measurements.

2.1 | Process model parameters estimation

Next, we tend to update the unknown parameters in the process model, i.e., α_k and σ_k^2. Drawing inspiration from the "current concept " in the continuous-time state space model on the basis of a nonzero mean-value and time-correlation model, see Reference 19, we have

$$
\begin{aligned}
\mathrm{x}_{j,t} &= \bar{\mathrm{x}}_j + \delta_t, \\
\dot{\delta}_t &= -\alpha_t\delta_t + \omega_t
\end{aligned} \tag{3}
$$

where δ_t is the maneuvering parameter, and ω_t is a white noise with zero mean and variance $2\alpha\sigma^2$. Here, note that in every single time increment, $\bar{\mathrm{x}}_j$ and α are considered as a constant, however, in the whole time horizon, it should be time-varying indeed. Therefore, in our nominal discrete-time model (1), the input vector u_k as well as the process model A_k, B_k and Q_k are defined as time-varying. Then, via discretizing (3), we have

$$
\delta_k = \beta_{k-1}\delta_{k-1} + \omega_{k-1}
$$

where ω_{k-1} is a discrete-time zero-mean white noise sequence with covariance:

$$
\sigma_{\omega_{k-1}}^2 = \sigma_{k-1}^2(1 - \beta_{k-1}^2), \quad \text{and} \quad \beta_{k-1} = e^{-\alpha_{k-1}T}
$$

with the sampling interval T.

Let $\hat{\mathrm{x}}_{j,k}$ be the estimate of the Jerk at time k, i.e. $\mathrm{x}_{j,k}$, given $Y_k = \{y_0, \cdots y_k\}$. Then, based on the Yule-Walker algorithm, α_k and σ_k^2 in the process model (1a) can be calculated in a recursive way as follows:

$$
\hat{\alpha}_k = -\ln\hat{\beta}_k/T, \quad \text{and} \quad \hat{\sigma}_k^2 = \hat{\sigma}_{\omega,k}^2/(1 - \hat{\beta}_k^2) \tag{4}
$$

where $\hat{\sigma}_{\omega,k}^2$ and $\hat{\beta}_k$ are formulated by

$$
\hat{\beta}_k = r_k/\tilde{r}_k, \quad \text{and} \quad \hat{\sigma}_{\omega,k}^2 = \tilde{r}_k - \hat{\alpha}_k\, r_k.
$$

Here, r_k and $\tilde{r}_k$ are autocorrelation functions of the form:

$$r_k = r_{k-1} + (\hat{\delta}_k \hat{\delta}_{k-1} - r_{k-1})/k,$$
$$\tilde{r}_k = \tilde{r}_{k-1} + (\hat{\delta}_k \hat{\delta}_k - \tilde{r}_{k-1})/k$$

where $\hat{\delta}_k$ is given by

$$\hat{\delta}_k = \hat{\mathrm{x}}_{j,k} - \bar{\hat{\mathrm{x}}}_{j,k} \quad \text{with} \quad \bar{\hat{\mathrm{x}}}_{j,k} = \frac{1}{L+1} \sum_{i=k-L}^{k} \hat{\mathrm{x}}_{j,i}.$$

Note that the mean value of Jerk $\bar{\hat{\mathrm{x}}}_{j,k}$ is considered as the mean in the time horizon $[k-L, k]$ with the lag $L > 0$, which can reduce the computational redundancy. In this way, it is not difficult to update matrices $\hat{A}_k$, $\hat{B}_k$ and $\hat{Q}_k$ based on the adaptive model parameters in view of (4).

Remark 1. Since $k < L + 1$, instead of using (4), we adopt the following strategy to force the Jerk to be within an area:

1. If $\hat{\mathrm{x}}_{j,k} \geq 0$, then
$$\hat{\sigma}_k^2 = (4-\pi)(\mathrm{x}_{j,max} - \hat{\mathrm{x}}_{j,k})^2/\pi;$$
2. If $\hat{\mathrm{x}}_{j,k} < 0$, then
$$\hat{\sigma}_k^2 = (4-\pi)(\hat{\mathrm{x}}_{j,k} - \mathrm{x}_{j,min})^2/\pi,$$

Note that $\hat{\mathrm{x}}_j \in [\mathrm{x}_{j,min}, \mathrm{x}_{j,max}]$ is the bound of the Jerk according to the earthquake magnitude.

2.2 | Measurement noise covariance estimation

Then, we seek to estimate the measurement noise covariance $\hat{R}_k$ at time step k. Let $\hat{x}_k$ and P_k denote the estimate and error covariance matrix of x_k give Y_k. Then, we have the innovation:

$$e_k = y_k - C_k \hat{x}_k.$$

Accordingly, its covariance S_k is.

$$S_k = C_k P_k C_k^\top + R_k. \tag{5}$$

Inspired by the fixed-lag smoother[34], a lag L is introduced. More precisely, for the measurement noise statistics, we tend to estimate the measurement noise covariance R_k using the observations in the $[k-L, k]$ with $L > 1$. Then, the sample covariance of the innovation with respect to lag $L + 1$ is as follows:

$$\bar{S}_{i,k} = \frac{1}{L} \sum_{i=k-L}^{k} (e_i - \bar{e}_k)(e_i - \bar{e}_k)^\top$$

where $\bar{e}_k$ is the mean of the estimation innovation in the time horizon $[k-L, k]$ with

$$\bar{e}_k = \frac{1}{L+1} \sum_{i=k-L}^{k} e_i = \frac{L}{L+1} \bar{e}_{k-1} + \frac{1}{L+1} e_k. \tag{6}$$

In addition, under the assumption that the noise covariance in every time step k is considered as constant, $\bar{S}_{i,k}$ also can be denoted as:

$$\bar{S}_{i,k} = \frac{1}{L+1} \sum_{i=k-L}^{k} C_i P_i C_i^\top + R_k.$$

Thus, it can be shown that an unbiased estimate of R_k is given by Reference 35:

$$\hat{R}_k = \frac{1}{L} \sum_{i=k-L}^{k} (e_i - \bar{e}_k)(e_i - \bar{e}_k)^\top - \frac{1}{L+1} \sum_{i=k-L}^{k} C_i P_i C_i^\top, \tag{7}$$

which can be further divided into two parts:

$$\hat{R}_k = \frac{L}{L+1} \tilde{R} + \Delta R_k \tag{8}$$

where the first one $\tilde{R}$ is with respect to the time horizon $[k-L, k-1]$, i.e.,

$$\tilde{R} = \frac{L+1}{L^2} \sum_{i=k-L}^{k-1} (e_i - \bar{e}_k)(e_i - \bar{e}_k)^\top - \frac{1}{L} \sum_{i=k-L}^{k-1} C_i P_i C_i^\top \tag{9}$$

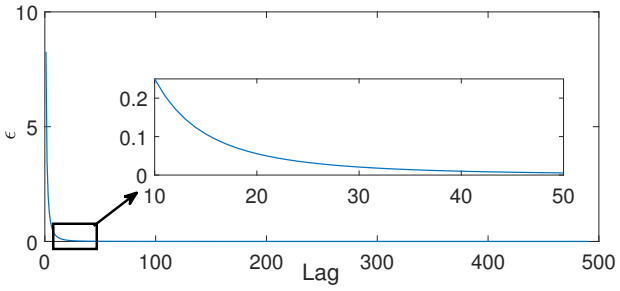


FIGURE 1 Error corresponding to the lag L.

while the other part ΔR_k is only corresponding to the time instant k

$$\Delta R_k = \frac{1}{L}(e_k - \bar{e}_k)(e_k - \bar{e}_k)^\top - \frac{1}{L+1}(C_k P_k C_k^\top). \tag{10}$$

In addition, according to (7), the measurement noise covariance at the time instant $k-1$ can be computed as:

$$\hat{R}_{k-1} = \frac{1}{L-1}\sum_{i=k-L}^{k-1}(e_i - \bar{e}_{k-1})(e_i - \bar{e}_{k-1})^\top - \frac{1}{L}\sum_{i=k-L}^{k-1} C_i P_i C_i^\top. \tag{11}$$

Here, it is easy to see that comparing (9) with (11), they would be identical if and only if $\lim_{L\to\infty} \epsilon \to 0$ where the error ϵ is defined as $\epsilon = (L+1)/L^2 - 1/(L-1)$. Clearly, as shown in Fig. 1, the error ϵ is negligible when $L \geq 50$. Thus, we can update the estimated measurement noise covariance in a recursive way:

$$\hat{R}_k = \frac{L}{L+1}\hat{R}_{k-1} + \Delta R_k. \tag{12}$$

where ΔR_k is shown in (10).

Remark 2. It is worth noting that if $k < L+1$, then in the equation (6), $\bar{e}_k$ should be redefined as $\bar{e}_k = y_k - C_k\hat{x}_k$. Also, the computations of $\hat{R}_k$ and ΔR_k in (8) and (10) are needed to be restructured into:

$$\hat{R}_k = k \times \hat{R}_{k-1}/(k-1) + \Delta R_k, \text{where } \Delta R_k = \|e_k - \bar{e}_k\|^2/(k-1) - C_k P_k C_k^\top/k.$$

3 | ROBUST ADAPTIVE KALMAN FILTER

In last section, we derived the nominal model, which is estimated from measurements in each time increment. Accordingly, we define $z_k := [\, x_{k+1}^\top \; y_k^\top \,]^\top$. Then, let $\phi(z_k|x_k)$ denote the transition probability density function of z_k given x_k with respect to the nominal model. Thus, the nominal conditional density that can characterize the nominal model derived by Section 2, is $\phi_k(z_k|x_k) \sim \mathcal{N}\left(m_{z_k}, K_{z_k}\right)$ with

$$m_{z_k} = \begin{bmatrix} \hat{A}_k x_k + \hat{B}_k u_k \\ C_k x_k \end{bmatrix}, \quad K_{z_k} = \begin{bmatrix} \hat{Q}_k & 0 \\ 0 & \hat{R}_k \end{bmatrix}.$$

In addition, the joint nominal probability density of the nominal model (1) over a finite time horizon $k \in [0, N]$ is as follows:

$$p\left(X_{N+1}, \mathrm{Y}_N\right) = \tilde{p}_0\left(x_0\right)\prod_{k=0}^{N}\phi_k\left(z_k|x_k\right) \tag{13}$$

where

$$X_{N+1}^\top = \left[\, x_0^\top \; \ldots \; x_k^\top \; \ldots \; x_{N+1}^\top \,\right] \;, \; Y_N^\top = \left[\, y_0^\top \; \ldots \; y_k^\top \; \ldots \; y_N^\top \,\right],$$

with $\tilde{p}_0(x_0) \sim \mathcal{N}(\hat{x}_0, V_0)$. However, it might be poor performance due to the emergence of model uncertainties when the estimated model or noise parameters, especially in the initial phase, are not exactly accurate. Thus, we assume that the probability density of the actual model is different from the one in the equation (13), which is given by

$$\tilde{p}\left(X_{N+1}, Y_N\right) = \tilde{p}_0\left(x_0\right)\prod_{k=0}^{N}\tilde{\phi}_k\left(z_k|x_k\right) \tag{14}$$

where $\tilde{\phi}_k(z_k|x_k)$ is defined as the actual transition probability density function of z_k given x_k according to the actual model. Then, in view of the KL-divergence philosophy[23], the modeling mismatch between the p and $\tilde{p}$ can be measured as:

$$D(\tilde{p}, p) := \iint \tilde{p}\left(X_{N+1}, Y_N\right) \ln \frac{\tilde{p}\left(X_{N+1}, Y_N\right)}{p\left(X_{N+1}, Y_N\right)} dY_N dX_{N+1}.$$

Accordingly, we have

$$D(\tilde{p}, p) = \sum_{k=0}^{N} D\left(\tilde{\phi}_k, \phi_k\right) \tag{15}$$

where

$$D(\tilde{\phi}_k, \phi_k) = \iint \tilde{\phi}_k\left(z_k|x_k\right) \tilde{p}_k\left(x_k\right) \ln\left(\frac{\tilde{\phi}_k\left(z_k|x_k\right)}{\phi_k\left(z_k|x_k\right)}\right) dz_k dx_k,$$

and $\tilde{p}_k(x_k)$ is the actual marginal density of x_k. In what follows, we assume that the actual model is within a ball centered about our nominal Gaussian model (1), where the ball is specified by a KL tolerance, hereafter called c_k. Thus, we assume that the actual conditional density $\tilde{\phi}_k$ given Y_{k-1} is included in an ambiguity set shown as:

$$\mathcal{B}_k := \left\{ \tilde{\phi}_k \text{ s.t. } \tilde{\mathbb{E}}\left[\ln\left(\frac{\tilde{\phi}_k\left(z_k|x_k\right)}{\phi_k\left(z_k|x_k\right)}\right)\middle| Y_{k-1}\right] \le c_k \right\}$$

where

$$\tilde{\mathbb{E}}\left[\ln\left(\frac{\tilde{\phi}_k\left(z_k|x_k\right)}{\phi_k\left(z_k|x_k\right)}\right)\middle| Y_{k-1}\right] := \iint \tilde{\phi}_k\left(z_k|x_k\right) \tilde{p}_k\left(x_k|Y_{k-1}\right) \ln\left(\frac{\tilde{\phi}_k\left(z_k|x_k\right)}{\phi_k\left(z_k|x_k\right)}\right) dz_k dx_k.$$

Note that $c_k > 0$ is the budget of modeling mismatch allowed at each sampling time instant k.

Remark 3. In Section 2, the nominal model is estimated from the data, this could introduce the larger mismodeling budget in the beginning because of the lack of the prior knowledge. For such a reason, the value of the tolerance c_k is designed as a decadent exponential function with a decay rate of 0.1 starting from the initial value c_0 and a constant c_1: $c_k = c_0 e^{-0.1k} + c_1$.

Up to this point, in view of the robust philosophy[30,31], a minimax game is designed for this modeling mismatch where one player, say nature, tends to find the worst model in the ball $\mathcal{B}_k$, while the other one conspires to select the optimum filter minimizing the estimation error within the worst model:

$$\hat{x}_k = \underset{f_k \in \mathcal{F}_k}{\operatorname{argmin}} \max_{\tilde{\phi}_k \in \mathcal{B}_k} J_k(\tilde{\phi}_k, f_k) \tag{16}$$

where

$$\begin{aligned} J_k(\tilde{\phi}_k, f_k) =& \frac{1}{2}\tilde{\mathbb{E}}\left[\| \left(x_k - f_k\left(y_k\right)\right) \|^2 | Y_{k-1}\right] \\ =& \frac{1}{2}\iint \| \left(x_k - f_k\left(y_k\right)\right) \|^2 \tilde{\phi}_k\left(z_k|x_k\right) \tilde{p}_k\left(x_k|Y_{k-1}\right) dz_k dx_k, \end{aligned}$$

$\mathcal{F}_k$ denotes the class of estimators with finite second-order moments with respect to all the densities $\tilde{\phi}_k\left(z_k|x_k\right) \tilde{p}_k(x_k|Y_{k-1})$ such that $\tilde{\phi}_k \in \mathcal{B}_k$. Here, the optimization variables are the actual model $\tilde{\phi}_k$ and the robust filter f_k. Moreover, it is not difficult to see that the actual probability density function (PDF) $\tilde{\phi}_k$ must be under the following constraint:

$$I_k(\tilde{\phi}_k) \triangleq \iint \tilde{\phi}_k\left(z_k|x_k\right) \tilde{p}_k\left(x_k|Y_{k-1}\right) dz_k dx_k = 1. \tag{17}$$

Lemma 1. For a fixed estimator $f_k \in \mathcal{F}_k$ with the given nominal PDF, the actual PDF $\tilde{\phi}_k\left(z_k|x_k\right) \in \mathcal{B}_k$ that maximizes the objective function

$$J_k(\tilde{\phi}_k, f_k) = \tilde{\mathbb{E}}\left[\left\|x_k - f_k(y_k)\right\|^2 | Y_{k-1}\right]$$

under the constraint $D(\tilde{\phi}_k, \phi_k) \le c_k$ is given by

$$\tilde{\phi}_k = \frac{1}{M_k\left(\lambda_k\right)} \exp\left(\frac{1}{2\lambda_k} \left\|x_k - f_k(y_k)\right\|^2\right) \phi_k. \tag{18}$$

Moreover, $M_k(\lambda_k)$ is the normalizing constant such that (17) holds. Finally, for $c_k > 0$ sufficiently small, there exists a unique $\lambda_k > 0$ such that $D(\tilde{\phi}_k^0, \phi_k) = c_k$.

Proof. The proof is similar to the Lemma 1 in Reference 31. □

Algorithm 1 Robust adaptive Kalman filter (RAKF)

Require: C, V_0, $\hat{R}_0$, $\hat{\alpha}_0$, $\hat{\sigma}_0^2$, r_{L+1}, $\tilde{r}_{L+1}$, $\mathtt{x}_{j,min}$, $\mathtt{x}_{j,max}$, $\hat{x}_0$, c_0, c_1, $y_0 \dots y_N$.
Ensure: $\hat{x}_{k+1}, k = 0 \dots N$.
1: Find $\hat{A}_0$, $\hat{B}_0$, $\hat{Q}_0$, w.r.t. $\hat{\alpha}_0$, $\hat{\sigma}_0^2$.
2: Find $\bar{\hat{\mathtt{x}}}_{j,0}$ w.r.t. $\hat{x}_0$.
3: **for** $k = 0 : L$ **do**
4: $G_k = \hat{A}_k V_k C^\top (C V_k C^\top + \hat{R}_k)^{-1}$;
5: $\hat{x}_{k+1} = \hat{A}_k \hat{x}_k + \hat{B}_k \bar{\hat{\mathtt{x}}}_{j,k} + G_k \left(y_k - C\hat{x}_k\right)$;
6: $P_{k+1} = \hat{A}_k V_k \hat{A}_k^\top - G_k \left(C V_k C^\top + \hat{R}_k^\top\right) G_k^\top + \hat{Q}_k$;
7: $c_k = c_0 e^{-0.1k} + c_1$;
8: Find λ_k^{-1} s.t. $\gamma(P_{k+1}, \lambda_k) = c_k$;
9: $V_{k+1} = (P_{k+1}^{-1} - \lambda_k^{-1} I)^{-1}$;
10: $e_{k+1} = y_{k+1} - C\hat{x}_{k+1}$;
11: $\bar{e}_{k+1} = \sum_{i=1}^{k+1} e_i/(k+1)$;
12: $\Delta R_{k+1} = \|e_{k+1} - \bar{e}_{k+1}\|^2/k - C V_{k+1} C^\top/(k+1)$;
13: $\hat{R}_{k+1} = (k+1)\hat{R}_k/k + \Delta R_{k+1}$;
14: $\hat{\alpha}_{k+1} = \hat{\alpha}_0$;
15: Find $\hat{\mathtt{x}}_{j,k+1}$ w.r.t. $\hat{x}_{k+1}$;
16: **if** $\hat{\mathtt{x}}_{j,k+1} \geq 0$ **then**
17: $\hat{\sigma}_{k+1}^2 = (4-\pi)(\mathtt{x}_{j,max} - \hat{\mathtt{x}}_{j,k+1})^2/\pi$;
18: **else**
19: $\hat{\sigma}_{k+1}^2 = (4-\pi)(\hat{\mathtt{x}}_{j,k+1} - \mathtt{x}_{j,min})^2/\pi$;
20: **end if**
21: Find $\hat{A}_{k+1}$, $\hat{B}_{k+1}$, $\hat{Q}_{k+1}$, w.r.t. $\hat{\alpha}_{k+1}$, $\hat{\sigma}_{k+1}^2$.
22: **end for**
23: **for** $k = L+1 : N$ **do**
24: Run Step 4 - Step 10;
25: $\bar{e}_{k+1} = (L+1)\bar{e}_k/L + e_{k+1}/(L+1)$;
26: $\Delta R_{k+1} = \|e_{k+1} - \bar{e}_{k+1}\|^2/L - C V_{k+1} C^\top/(L+1)$;
27: $\hat{R}_{k+1} = (L+1)\hat{R}_k/L + \Delta R_{k+1}$;
28: Find $\hat{\mathtt{x}}_{j,k+1}$ w.r.t. $\hat{x}_{k+1}$;
29: $\bar{\hat{\mathtt{x}}}_{j,k+1} = \sum_{i=k-L+1}^{k+1} \hat{\mathtt{x}}_{j,i}/(L+1)$;
30: $\hat{\delta}_{k+1} = \hat{\mathtt{x}}_{j,k+1} - \bar{\hat{\mathtt{x}}}_{j,k+1}$;
31: $r_{k+1} = r_k + (\hat{\delta}_{k+1}\hat{\delta}_k - r_k)/(k+1)$;
32: $\tilde{r}_{k+1} = \tilde{r}_k + (\hat{\delta}_{k+1}\hat{\delta}_{k+1} - \tilde{r}_k)/(k+1)$;
33: $\hat{\alpha}_{k+1} = (\ln \tilde{r}_{k+1} - \ln r_{k+1})/T$;
34: $\hat{\sigma}_{k+1}^2 = (\tilde{r}_{k+1} - \hat{\alpha}_{k+1}\, r_{k+1})/(1 - (r_{k+1}/\tilde{r}_{k+1})^2)$;
35: Find $\hat{A}_{k+1}$, $\hat{B}_{k+1}$, $\hat{Q}_{k+1}$, w.r.t. $\hat{\alpha}_{k+1}$, $\hat{\sigma}_{k+1}^2$.
36: **end for**

At this point, it is not difficult to see that the worst density solution to the minimax problem (16) depends on the estimator $\hat{x}_k$, while the optimal estimator depends on the worst density. Hence, in order to break this deadlock, we assume that the actual *a priori* probability density $\tilde{p}_t(x_k|Y_{k-1}) \sim \mathcal{N}(\hat{x}_k, V_k)$ with $V_k > 0$. Since $\tilde{p}_k\left(z_k|Y_{k-1}\right)$ enjoys the aforementioned properties and $V_t > 0$, in view of the nominal model shown in Section 2 and the results shown in Theorem 1 of Reference 30 , it follows that the optimal robust estimator solution to (16) is given by Algorithm 1. Note that to find the optimal solution of the dynamic game (16), we do not compute numerically any integral. Indeed, the dynamic game (16) admits the solution through Lagrange multipliers theory which can provide a similar structure of the standard Kalman filter, however, the filtering performance is computed in a different way, see Step 9 in Algorithm 1. Then, $\gamma(P_k, \theta_k)$ in Step 8 is given by (19):

$$\gamma(P_k, \lambda_k) := \frac{1}{2}\left[\ln\det(I - \lambda_k^{-1} P_k) + \operatorname{tr}\left((I - \lambda_k^{-1} P_k)^{-1} - I\right)\right] = c_k \tag{19}$$

and $\theta_k = \lambda_k^{-1}$. Furthermore, the inclusion of V_t in the process necessitates the computation of θ_t at Step 8, which is crucial for the success of our RAKF in handling model uncertainties. It should be noted that determining the risk sensitivity parameter θ_t can not be achieved through a straightforward closed-form solution and requires the Bisection method.

Remark 4. It is worth noting that the proposed robust adaptive Kalman filter allows for exploiting the recursive expressions with a lack of statistical knowledge of the dynamic and noise processes. The same philosophy can be adopted not only for our interstory drift estimation problem but also for the adaptive tracking scenarios.

4 | SIMULATION STUDY

We analyze the impact of the model parameters on the estimated results based on a maneuvering target tracking problem. In particular, analysis of the model parameters such like c_k, $\hat{\alpha}_k$ and $\hat{\sigma}_k^2$ has been studied References 12,13 and 24. Thus, we only focus on the measurement noise covariance $\hat{R}_k$. More precisely, we generated a random two-dimensional trajectory in a two-dimensional interval $[-10, 10] \times [-10, 10]$, see the red line in Fig. 2, to simulate the maneuvering target movement. Next, a white noise sequence with variance $\bar{R}_k = \bar{r}_k \cdot I_2$ where $\bar{r}_k \sim \mathbb{U}\{1.5, 2.5\}$ has been added to the reference trajectory. Thus, we obtained a measurement trajectory as well, see the blue line.

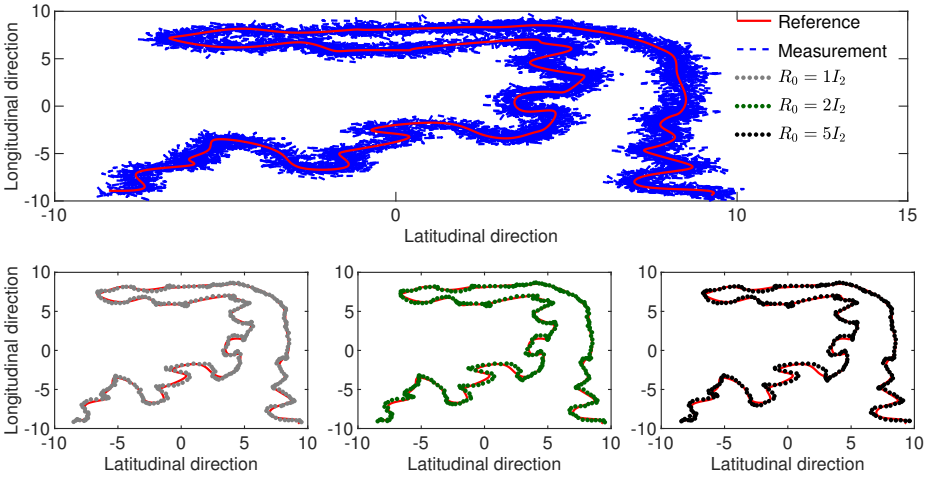


FIGURE 2 Reference trajectory, measurement trajectory and estimated trajectories with $R_0 = 1 \cdot I_2$, $R_0 = 2 \cdot I_2$ and $R_0 = 5 \cdot I_2$.

Here, we select $T = 0.01$, and then the initialization of the model parameters are: $c_0 = 1$, $\hat{\alpha}_0 = 0.05$, $\hat{\sigma}_0^2 = 10$, $r_{L+1} = 0$, $\tilde{r}_{L+1} = 0$, $\mathtt{x}_{j,min} = -1$, $\mathtt{x}_{j,max} = 1$, and $c_1 = 0.15$. Subgraphs in Fig. 2 show estimated trajectories with different initial measurement noise covariances: $R_0 = 1 \cdot I_2$, $R_0 = 2 \cdot I_2$ and $R_0 = 5 \cdot I_2$. Clearly, the results are not affected by the initial value. Besides, we also design the following two methods: robust Signer adaptive Kalman filter (RSAKF) and robust Jerk adaptive Kalman filter (RJAKF). In other words, the only difference between RAKF and RJAKF is that the measurement noise covariance in RAKF is adaptive by using our method 2.2, while RSAKF is the third-order paradigm of RJAKF. Table 1 shows RAKF has the better performance than RSAKF and RJAKF. More importantly, the results of the proposed RAKF are not affected by the initial value of the noise covariance even it was assumed time varying.

5 | EXPERIMENT

Structural seismic test, based on the earthquake simulation shaking table, can evaluate the seismic performance and understand the seismic response of the RC structure[36]. In our test, we generated seismic waves that simulate El Centro waves[18] in the United States to study the level of the structural damage of a four-story RC structure. Fig. 3 shows our experiment platform

TABLE 1 RMSE for RAKF, RSAKF and RJAKF

	RAKF		RSAKF		RJAKF	
	Latitudinal	Longitudinal	Latitudinal	Longitudinal	Latitudinal	Longitudinal
R_0=5	0.214	0.192	2.016	2.246	0.821	0.982
R_0=2	0.180	0.172	0.676	0.644	0.363	0.233
R_0=1	0.164	0.162	1.987	1.433	0.540	0.699

where the yellow rope on the top was used to provide backup support without the direct contact, the red part is our four-layer RC structure with the story height $h = 0.53m$, and in the bottom, it is the earthquake simulation shaking table which can provide the simulated seismic signal. Moreover, it is worth noticing that the structural damage usually occurs on the first floor, as shown in Fig. 3(b). This means, in general case, we only need to focus on the ID of the first floor. Accordingly, before the observable structural damage occurs, we obtained two trajectories over a finite time interval $N = [0, 10000]$ of the ID of the first floor: one is from a low-cost sensor with the massive measurement noise shown in the dark green line in Fig. 4, the other is from a high precision sensor that is impossible to assemble on a large scale, shown in the red line in Fig. 4, which is treated as the reference signal. Clearly, our measurement data is extraordinarily noisy and the dynamic characteristic of the reference displacement is almost covered by noise. In addition, we notice that the reference data is also slightly disturbed. Indeed, this disturbance, which could be from the measuring error of our high-precision sensor or the vibration of the RC structure, is acceptable because of its negligible size. Accordingly, we seek to match our estimated trajectory with the reference trajectory as closely as possible.

(a) Before the test. (b) After the test.

FIGURE 3 Red four-layer RC frame, yellow rope providing backup support but no direct contact and earthquake simulation shaking table.

In what follows, we consider the following three filters over the time horizon N: the proposed RAKF in Algorithm 1, the robust Kalman filter[23] based on the Singer model (RSKF)[12] and the robust Kalman filter[23] based on the Jerk model (RJKF)[13]. The model parameters are selected to be the same as those in the simulation. Note that the initial conditions of the predicted state vector and the covariance matrix for the initial state are given by $\hat{x}_0 = [0\ 0\ 0\ 0]^\top$ and $V_0 = 10 \cdot I_4$, respectively. Here, for RSKF, they are $\hat{x}_0 = [0\ 0\ 0]^\top$ and $V_0 = 10 \cdot I_3$, respectively. Finally, $\hat{R}_0 = 10$.

Fig. 5 shows the online estimated trajectories obtained by RAKF (blue line), RSKF (gray line) and RJKF (orange line) compared with the reference trajectory (red line). Clearly, RAKF outperforms RSKF and RJKF. Moreover, the first subgraph in Fig. 5 shows the convergence speed of three filters. As excepted, both RAKF and RJKF are convergent rapidly with the help of our robust approach. Interestingly, RSKF converges almost in an instant. This is because the vibration of the RC structure is relatively weak when the seismic excitation first arrived. At this time, the jerk has increased the modeling mismatch instead,

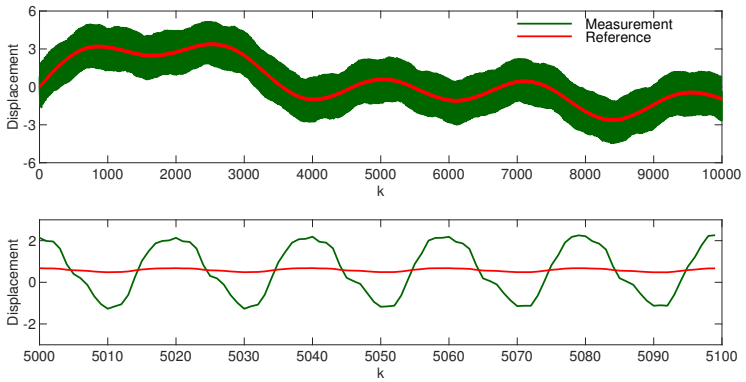


FIGURE 4 Reference trajectory and measurement trajectory of ID.

which affects the estimation accuracy. However, as time goes on, it is easy to see in the next two subgraphs, due to the update of model parameters, only RAKF can accurately match the reference trajectory. Moreover, we quantify the performance of three

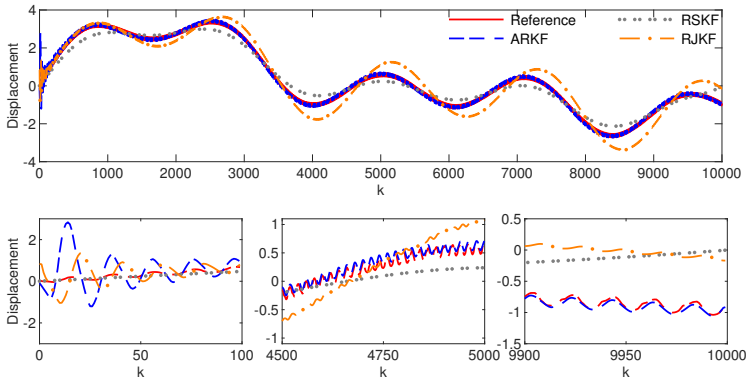


FIGURE 5 Reference trajectory and estimated trajectories obtained by RAKF, RSKF and RJKF.

filters via mean, standard deviation (Std) and root-mean-squared error (RMSE). More precisely, the RMSE is given by:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\tilde{\mathrm{x}}_{p,k} - \hat{\mathrm{x}}_{p,k})^2}$$

where $\tilde{\mathrm{x}}_{p,k}$ represents the reference ID and $\hat{\mathrm{x}}_{p,k}$ is the estimated ID extracted from $\hat{x}_k$. Then, values of mean, Std and RMSE for aforementioned three filters are displayed in Table 2. As expected, compared to RSKF and RJKF, RAKF can track the reference trajectory more accurately and process measurement noise better.

TABLE 2 Performance for RAKF, RSKF and RJKF

	Mean	Std	RMSE
RAKF	0.0514	0.0026	0.0968
RSKF	0.2739	0.4071	0.3206
RJKF	0.5046	0.1034	0.4239

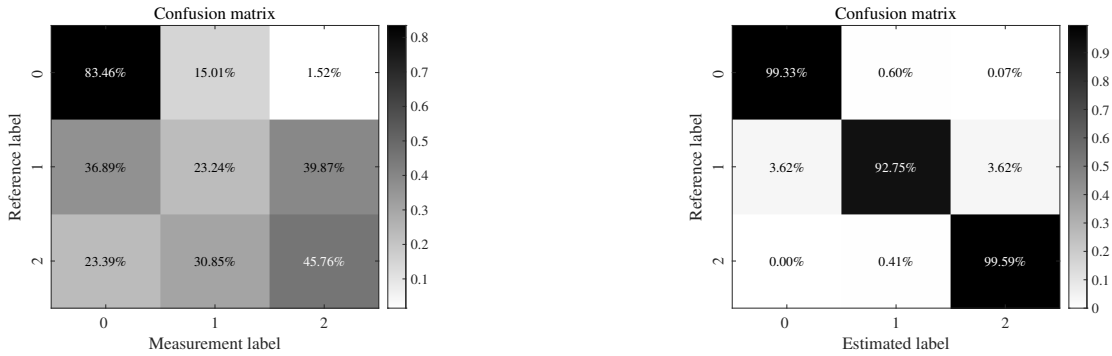
Then, it is not difficult to show the IDR, namely, θ_k, which is the ratio of the ID to the story height in every time step[37], defined as:

$$\hat{\theta}_k = \hat{\mathtt{x}}_{p,k}/(100h), \quad \tilde{\theta}_k = \tilde{\mathtt{x}}_{p,k}/(100h), \quad \check{\theta}_k = \check{\mathtt{x}}_{p,k}/(100h)$$

where $\hat{\theta}_k$, $\tilde{\theta}_k$ and θ_k are the estimated IDR with respect to $\hat{\mathtt{x}}_{p,k}$ by means of RAKF, the reference IDR with respect to $\tilde{\mathtt{x}}_{p,k}$ and the measurement IDR with respect to the measurement ID, i.e., $\check{\mathtt{x}}_{p,k}$, respectively. In addition, based on "General Rule for Performance-based Seismic Design of Buildings" with the No. CECS 160-2004 in China, we have the Table 3 showing the IDR corresponding to different building status. According to Table 3, we classified our obtained IDR into three labels, i.e., 0, 1 2. Fig. 6 shows the corresponding confusion matrices: due to our proposed algorithm, the accuracy of the structural health monitoring is obviously improved.

TABLE 3 IDR corresponding to different building status

Building status	Criteria	Label
No damage	$\theta_k < 1/550$	0
Slight damage	$1/550 < \theta_k < 1/250$	1
Moderate damage	$1/250 < \theta_k < 1/125$	2



(a) Confusion matrix between $\tilde{\theta}_k$ and $\check{\theta}_k$.

(b) Confusion matrix between $\tilde{\theta}_k$ and $\hat{\theta}_k$.

FIGURE 6 Results of the structural health monitoring.

6 | CONCLUSION

In this paper, we have proposed an adaptive robust Kalman filter for an interstory drift estimation problem to show the structural health condition of RC. Since the statistics or internal dynamics describing the signals and measurements are not known, we build an adaptive current Jerk model where the model parameters are updated in each time step to presuppose the statistics characterization of the RC dynamic, while the unknown measurement noise covariance is adapted based on a fixed-lag innovation with respect to measurements. Then, considering the possible modeling mismatch between the identified nominal model and the actual one, a robust adaptive Kalman filter is designed with a exponential decay tolerance in each time increment by solving a minimax game: one "hostile" player tries to select an "actual" model far from the proposed ACJM, while an optimum filter is designed by minimizing the estimation error according to this actual model. Finally, in our experimental test, we obtained the measured ID on the first floor of a four-story RC structure under a seismic wave that simulates El Centro waves in the United States. The experimental results show that compared with two other filters, RAKF offers the better estimation performance. Moreover, we have also presented a simulation study showing the satisfactory adaptability of the proposed algorithm in the absence of prior knowledge of the measurement noise process.

References

1. Deraemaeker A, Worden K. A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mechanical Systems and Signal Processing* 2018; 105: 1–15.

2. Smarra F, Tjen J, D'Innocenzo A. Learning methods for structural damage detection via entropy-based sensors selection. *International Journal of Robust and Nonlinear Control* 2022; 32(10): 6035–6067.

3. Barthorpe RJ. *On model-and data-based approaches to structural health monitoring*. PhD thesis. University of Sheffield, 2010.

4. Hou S, Zeng C, Zhang H, Ou J. Monitoring interstory drift in buildings under seismic loading using MEMS inclinometers. *Construction and Building Materials* 2018; 185: 453–467.

5. Building Seismic Safety Council (BSSC), Prestandard and commentary for the seismic rehabilitation of buildings, FEMA-356. *Washington, D.C.* 2000.

6. Algan BB. *Drift and damage considerations in earthquake-resistant design of reinforced concrete buildings*. University of Illinois at Urbana-Champaign . 1982.

7. Chang CY, Huang CW. Non-contact measurement of inter-story drift in three-layer RC structure under seismic vibration using digital image correlation. *Mechanical Systems and Signal Processing* 2020; 136: 106500.

8. Sarkar S, Roy A, Kumar S, Das B. Seismic Intensity estimation using Multilayer Perceptron for Onsite Earthquake Early Warning. *IEEE Sensors Journal* 2021.

9. Stojanovic V, He S, Zhang B. State and parameter joint estimation of linear stochastic systems in presence of faults and non-Gaussian noises. *International Journal of Robust and Nonlinear Control* 2020; 30(16): 6683–6700.

10. Schöller C, Aravantinos V, Lay F, Knoll A. What the constant velocity model can teach us about pedestrian motion prediction. *IEEE Robotics and Automation Letters* 2020; 5(2): 1696–1703.

11. Reece S, Roberts S. The near constant acceleration Gaussian process kernel for tracking. *IEEE Signal Processing Letters* 2010; 17(8): 707–710.

12. Singer RA. Estimating optimal tracking filter performance for manned maneuvering targets. *IEEE Transactions on Aerospace and Electronic Systems* 1970(4): 473–483.

13. Mehrotra K, Mahapatra PR. A jerk model for tracking highly maneuvering targets. *IEEE Transactions on Aerospace and Electronic Systems* 1997; 33(4): 1094–1105.

14. Bian X, Li XR, Chen H, Gan D, Qiu J. Joint estimation of state and parameter with synchrophasors—Part I: State tracking. *IEEE Transactions on Power Systems* 2011; 26(3): 1196–1208.

15. Bian X, Li XR, Chen H, Gan D, Qiu J. Joint estimation of state and parameter with synchrophasors—Part II: Parameter tracking. *IEEE Transactions on Power Systems* 2011; 26(3): 1209–1220.

16. Yi SL, Jin XB, Su TL, et al. Online denoising based on the second-order adaptive statistics model. *Sensors* 2017; 17(7): 1668.

17. Hann CE, Singh-Levett I, Deam BL, Mander JB, Chase JG. Real-time system identification of a nonlinear four-story steel frame structure—Application to structural health monitoring. *IEEE Sensors Journal* 2009; 9(11): 1339–1346.

18. Longhini A, Perbellini M, Gottardi S, Yi S, Liu H, Zorzi M. Learning the tuned liquid damper dynamics by means of a robust EKF. In: 2021 American Control Conference (ACC). IEEE. ; 2021: 60–65.

19. Kumar K, Zhou H. A'current'statistical model and adaptive algorithm for estimating maneuvering targets. *Journal of guidance, control, and dynamics* 1984; 7(5): 596–602.

20. Zhang H, Xie J, Ge J, Lu W, Zong B. Adaptive Strong Tracking Square-Root Cubature Kalman Filter for Maneuvering Aircraft Tracking. *IEEE Access* 2018; 6: 10052-10061.

21. Yi S, Zorzi M. Robust fixed-lag smoothing under model perturbations. *Journal of the Franklin Institute* 2023; 360(1): 458–483.

22. Salim M, Mahdavi Z, Kharrati H. An H_∞ approach to data-driven fault estimation, and isolation for Hammerstein-Wiener systems. *International Journal of Robust and Nonlinear Control* 2022; 32(13): 7348–7362.

23. Yi S, Zorzi M. Robust kalman filtering under model uncertainty: The case of degenerate densities. *IEEE Transactions on Automatic Control* 2021; 67(7): 3458–3471.

24. Zorzi M. On the robustness of the Bayes and Wiener estimators under model uncertainty. *Automatica* 2017; 83: 133–140.

25. Wang Y, Puig V, Cembrano G. Robust fault estimation based on zonotopic Kalman filter for discrete-time descriptor systems. *International Journal of Robust and Nonlinear Control* 2018; 28(16): 5071–5086.

26. Chunshan Y, Zili D. Robust time-varying Kalman estimators for systems with packet dropouts and uncertain-variance multiplicative and linearly correlated additive white noises. *International Journal of Adaptive Control and Signal Processing* 2018; 32(1): 147–169.

27. Zenere A, Zorzi M. On the coupling of model predictive control and robust Kalman filtering. *IET Control Theory & Applications* 2018; 12(13): 1873–1881.

28. Whittle P. *Risk-sensitive optimal control*. 2. Wiley . 1990.

29. Zorzi M. Distributed Kalman filtering under model uncertainty. *IEEE Transactions on Control of Network Systems* 2019; 7(2): 990–1001.

30. Levy BC, Nikoukhah R. Robust least-squares estimation with a relative entropy constraint. *IEEE Transactions on Information Theory* 2004; 50(1): 89-104.

31. Levy BC, Nikoukhah R. Robust State Space Filtering Under Incremental Model Perturbations Subject to a Relative Entropy Tolerance. *IEEE Transactions on Automatic Control* 2013; 58(3): 682-695.

32. Zorzi M. Robust Kalman Filtering Under Model Perturbations. *IEEE Transactions on Automatic Control* 2017; 62(6): 2902-2907.

33. Yi S, Zorzi M. Low-rank Kalman filtering under model uncertainty. In: Proc. 59th IEEE Conf. Decision and Control (CDC). ; 2020: 2930-2935.

34. Kailath T, Sayed AH, Hassibi B. *Linear estimation*. Prentice Hall . 2000.

35. Myers K, Tapley B. Adaptive sequential estimation with unknown noise statistics. *IEEE Transactions on Automatic Control* 1976; 21(4): 520–523.

36. Ji X, Fenves GL, Kajiwara K, Nakashima M. Seismic damage detection of a full-scale shaking table test structure. *Journal of Structural Engineering* 2011; 137(1): 14–21.

37. Hou H, Zhang S, Qu B, et al. Low-rise steel moment-resisting frames with novel re-centering beam-to-column connections: connection modelling and system performance evaluation. *Journal of Constructional Steel Research* 2021; 180.