Adaptive Kalman Filter for Actuator Fault Diagnosis

Qinghua Zhang

Adaptive Kalman Filter for Actuator Fault Diagnosis

Qinghua Zhang *

* *Inria/IFSTTAR, Campus de Beaulieu, 35042 Rennes Cedex, France*
Email: qinghua.zhang@inria.fr

Abstract: An adaptive Kalman filter is proposed in this paper for actuator fault diagnosis in discrete time *stochastic* time varying systems. By modeling actuator faults as parameter changes, fault diagnosis is performed through joint state-parameter estimation in the considered stochastic framework. Under the classical uniform complete observability-controllability conditions and a persistent excitation condition, the exponential stability of the proposed adaptive Kalman filter is rigorously analyzed. The minimum variance property of the combined state and parameter estimation errors is also demonstrated. Numerical examples are presented to illustrate the performance of the proposed algorithm.

Keywords: FDI, adaptive observer, joint state-parameter estimation, LTV/LPV systems.

1. INTRODUCTION

In order to improve the performance and the reliability of industrial systems, and to satisfy safety and environmental requirements, researches and developments in the field of fault detection and isolation (FDI) have been continuously progressing during the last decades (Hwang et al., 2010). Model-based FDI have been mostly studied for linear time invariant (LTI) systems (Gertler, 1998; Chen and Patton, 1999; Isermann, 2005; Ding, 2008), whereas nonlinear systems have been studied to a lesser extent and limited to some particular classes of systems (De Persis and Isidori, 2001; Xu and Zhang, 2004; Berdjag et al., 2006). This paper is focused on actuator fault diagnosis for linear time-varying (LTV) systems, *including the particular case* of linear parameter varying (LPV) systems. The problem of fault diagnosis for a large class of nonlinear systems can be addressed through LTV/LPV reformulation and approximations (Lopes dos Santos et al., 2011; Tóth et al., 2011). It is thus an important advance in FDI by moving from LTI to LTV/LPV systems.

In this paper, actuator faults are modeled as parameter changes, and their diagnosis is achieved through joint estimation of states and parameters of the considered LTV/LPV systems. Usually the problem of joint state-parameter estimation is solved by recursive algorithms known as *adaptive observers*, which are most often studied in *deterministic* frameworks for *continuous* time systems (Marino and Tomei, 1995; Zhang, 2002; Besançon et al., 2006; Farza et al., 2014).

Discrete time systems have been considered in (Guyader and Zhang, 2003; Ţiclea and Besançon, 2016), also in *deterministic* frameworks. In order to take into account *random uncertainties* with a numerically efficient algorithm, this paper considers *stochastic* systems in discrete time, with an *adaptive Kalman filter*, which is structurally inspired by adaptive observers (Zhang, 2002; Guyader and Zhang, 2003), but with well-established stochastic properties.

The main contribution of this paper is an adaptive Kalman filter for discrete time LTV/LPV system joint state-parameter estimation in a *stochastic* framework, *with rigorously proved stability and minimum variance properties.*

Different adaptive Kalman filters have been studied in the literature for state estimation based on inaccurate state-space models. Most of these algorithms address the problem of unknown (or partly known) state noise covariance matrix or output noise covariance matrix (Mehra, 1970; Brown and Rutan, 1985), whereas the case of incorrect state dynamics model is treated as incorrect state covariance matrix. In contrast, *in the present paper*, the new adaptive Kalman filter is designed for actuator fault diagnosis, by jointly estimating states and parameter changes caused by actuator faults.

2. PROBLEM STATEMENT

The discrete time LTV system subject to actuator faults considered in this paper is generally in the form of [1]

$$x(k) = A(k)x(k-1) + B(k)u(k) + \Phi(k)\theta + w(k) \tag{1a}$$

$$y(k) = C(k)x(k) + v(k), \tag{1b}$$

where $k = 0, 1, 2, \dots$ is the discrete time instant index, $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^l$ the input, $y(k) \in \mathbb{R}^m$ the output, $A(k), B(k), C(k)$ are time-varying matrices of appropriate sizes characterizing the nominal state-space model, $w(k) \in \mathbb{R}^n, v(k) \in \mathbb{R}^m$ are mutually independent centered white Gaussian noises of covariance matrices $Q(k) \in \mathbb{R}^{n\times n}$ and $R(k) \in \mathbb{R}^{m\times m}$, and the term $\Phi(k)\theta$ represents actuator faults with a known matrix sequence $\Phi(k) \in \mathbb{R}^{n\times p}$ and a constant (or piecewise constant with rare jumps) parameter vector $\theta \in \mathbb{R}^p$.

A typical example of actuator faults represented by the term $\Phi(k)\theta$ is actuator gain losses. When affected by such faults, the nominal control term $B(k)u(k)$ becomes

$$B(k)(I_l - \mathrm{diag}(\theta))u(k) = B(k)u(k) - B(k)\mathrm{diag}(u(k))\theta$$

[1] There exists a "forward" variant form of the state-space model, typically with $x(k+1) = A(k)x(k) + B(k)u(k) + w(k)$. While this difference is important for control problems, it is not essential for estimation problems, like the one considered in this paper. The form chosen in this paper corresponds to the convention that data are collected at $k = 1, 2, 3, \dots$ and the initial state refers to $x(0)$.

where I_l is the $l \times l$ identity matrix, the diagonal matrix $\mathrm{diag}(\theta)$ contains gain loss coefficients within the interval $[0, 1]$, and $\Phi(k) \in \mathbb{R}^{n\times l}$ $(p{=}l)$ is, in this particular case,

$$\Phi(k) = -B(k)\mathrm{diag}(u(k)). \tag{2}$$

A straightforward solution for the joint estimation of $x(k)$ and θ is to apply the Kalman filter to the augmented system

$$\begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} A(k) & \Phi(k) \\ 0 & I_p \end{bmatrix} \begin{bmatrix} x(k-1) \\ \theta(k-1) \end{bmatrix} + \begin{bmatrix} B(k) \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} w(k) \\ 0 \end{bmatrix}$$

$$y(k) = [C(k) \;\; 0] \begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} + v(k).$$

However, to ensure the stability of the Kalman filter, this *augmented* system should be uniformly completely observable and uniformly completely controllable regarding the state noise (Kalman, 1963; Jazwinski, 1970). Notice that, even in the case of time *invariant* matrices A and C, the augmented system is time varying because of $\Phi(k)$, which is typically time varying. The uniform complete observability of an LTV system is defined as the uniform positive definiteness of its observability Gramian (Kalman, 1963; Jazwinski, 1970). In practice, it is not natural to directly assume properties (observability and controllability) of the *augmented system* (a more exaggerated way would be *directly assuming* the stability of its Kalman filter, or anything else that should be proved!). Moreover, the augmented system is clearly *not* uniformly completely controllable regarding the state noise $w(k)$, since the augmented states $\theta(k)$ are not controlled at all by $w(k)$.

In contrast, *in the present paper*, the classical uniform complete observability and uniform complete controllability are assumed for the *original* system (1), in terms of the Gramian matrices defined for the $[A(k), C(k)]$ pair and the $[A(k), Q^{\frac{1}{2}}(k)]$ pair. These conditions, together with a persistent excitation condition (see Assumption 3 formulated later), ensure the stability of the adaptive Kalman filter presented in this paper.

3. THE ADAPTIVE KALMAN FILTER

In the adaptive Kalman filter, the state estimate $\hat{x}(k|k) \in \mathbb{R}^n$ and the parameter estimate $\hat{\theta}(k) \in \mathbb{R}^p$ are recursively updated at every time instant k. This algorithm involves also a few other recursively updated auxiliary variables: $P(k|k) \in \mathbb{R}^{n\times n}, \Upsilon(k) \in \mathbb{R}^{n\times p}, S(k) \in \mathbb{R}^{p\times p}$ and a forgetting factor $\lambda \in (0, 1)$.

At the initial time instant $k = 0$, the initial state $x(0)$ is assumed to be a Gaussian random vector

$$x(0) \sim \mathcal{N}(x_0, P_0). \tag{3}$$

Let $\theta_0 \in \mathbb{R}^p$ be the initial guess of θ, $\lambda \in (0, 1)$ be a chosen forgetting factor, and ω be a chosen positive value for initializing $S(k)$, then the adaptive Kalman filter consists of the initialization step and the recursion steps described below. Each part of this algorithm separated by horizontal lines will be commented after the algorithm description.

Initialization

$$P(0|0) = P_0 \qquad \Upsilon(0) = 0 \qquad S(0) = \omega I_p \tag{4a}$$
$$\hat{\theta}(0) = \theta_0 \qquad \hat{x}(0|0) = x_0 \tag{4b}$$

Recursions for $k = 1, 2, 3, \ldots$

$$P(k|k-1) = A(k)P(k-1|k-1)A^T(k) + Q(k) \tag{5a}$$
$$\Sigma(k) = C(k)P(k|k-1)C^T(k) + R(k) \tag{5b}$$
$$K(k) = P(k|k-1)C^T(k)\Sigma^{-1}(k) \tag{5c}$$
$$P(k|k) = [I_n - K(k)C(k)]P(k|k-1) \tag{5d}$$

$$\Upsilon(k) = [I_n - K(k)C(k)]A(k)\Upsilon(k-1) + [I_n - K(k)C(k)]\Phi(k) \tag{5e}$$
$$\Omega(k) = C(k)A(k)\Upsilon(k-1) + C(k)\Phi(k) \tag{5f}$$
$$\Lambda(k) = [\lambda\Sigma(k) + \Omega(k)S(k-1)\Omega^T(k)]^{-1} \tag{5g}$$
$$\Gamma(k) = S(k-1)\Omega^T(k)\Lambda(k) \tag{5h}$$
$$S(k) = \frac{1}{\lambda}S(k-1) - \frac{1}{\lambda}S(k-1)\Omega^T(k)\Lambda(k)\Omega(k)S(k-1) \tag{5i}$$

$$\tilde{y}(k) = y(k) - C(k)\big[A(k)\hat{x}(k-1|k-1) + B(k)u(k) + \Phi(k)\hat{\theta}(k-1)\big] \tag{5j}$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \Gamma(k)\tilde{y}(k) \tag{5k}$$
$$\hat{x}(k|k) = A(k)\hat{x}(k-1|k-1) + B(k)u(k) + \Phi(k)\hat{\theta}(k-1) + K(k)\tilde{y}(k) + \Upsilon(k)[\hat{\theta}(k) - \hat{\theta}(k-1)]. \tag{5l}$$

Recursions (5a)-(5d) compute the covariance matrix $P(k|k) \in \mathbb{R}^{n\times n}$ of the state estimate, the innovation covariance matrix $\Sigma(k) \in \mathbb{R}^{m\times m}$ and the state estimation gain matrix $K(k) \in \mathbb{R}^{n\times m}$. These formulas are identical to those of the classical Kalman filter. Inspired by the recursive least square (RLS) estimator with an exponential forgetting factor, recursions (5e)-(5i) compute the parameter estimate gain matrix $\Gamma(k) \in \mathbb{R}^{p\times m}$ through the auxiliary variables $\Upsilon(k) \in \mathbb{R}^{n\times p}, \Omega(k) \in \mathbb{R}^{m\times p}, S(k) \in \mathbb{R}^{p\times p}$. Equation (5j) computes the innovation $\tilde{y}(k) \in \mathbb{R}^m$. Finally, recursions (5k)-(5l) compute the parameter estimate and the state estimate.

Part of equation (5l), namely,

$$\hat{x}(k|k) \sim A(k)\hat{x}(k-1|k-1) + B(k)u(k) + K(k)\tilde{y}(k)$$

can be easily recognized as part of the classical Kalman filter, with the traditional *prediction step* and *update step* combined into a single step. The term $\Phi(k)\hat{\theta}(k-1)$ corresponds to the actuator fault term $\Phi(k)\theta$ in (1a), with θ replaced by its estimate $\hat{\theta}(k-1)$. The extra term $\Upsilon(k)[\hat{\theta}(k)-\hat{\theta}(k-1)]$ is for the purpose of compensating the error caused by $\hat{\theta}(k-1) \neq \theta$. This term is essential for the analysis of the properties of the adaptive Kalman filter in the following sections. It has also been introduced in the deterministic adaptive observer in (Guyader and Zhang, 2003), and its continuous time counterpart in (Zhang, 2002).

4. STABILITY OF THE ADAPTIVE KALMAN FILTER

Assumption 1. The matrices $A(k), B(k), C(k), \Phi(k), Q(k), R(k)$ and the input $u(k)$ are *upper bounded*, $Q(k)$ is symmetric positive semidefinite, and $R(k)$ is symmetric positive-definite with a strictly positive lower bound, for all $k \geq 0$. □

Assumption 2. The $[A(k), C(k)]$ pair is uniformly completely observable, and the $[A(k), Q^{\frac{1}{2}}(k)]$ pair is uniformly completely controllable, in the sense of the uniform positive definiteness of the corresponding Gramian matrices. (Kalman, 1963; Jazwinski, 1970). □

Assumption 3. The signals contained in the matrix $\Phi(k)$ are persistently exciting in the sense that there exit an integer $h > 0$ and a real constant $\alpha > 0$ such that, for all integer $k \geq 0$, the matrix sequence $\Omega(k)$ driven by $\Phi(k)$ through the linear system (5e)-(5f), satisfies

$$\sum_{s=0}^{h-1} \Omega^T(k+s)\Sigma^{-1}(k+s)\Omega(k+s) \geq \alpha I_p. \qquad (6)$$

□

Proposition 1. The matrices $P(k|k), \Upsilon(k), \Sigma(k), K(k), \Omega(k)$ all have a finite upper bound, and $\Sigma(k)$ has a strictly positive lower bound. □

Proof. The proof based on classical results is quite straightforward. The recursive computations (5a)-(5d) for $\Sigma(k), K(k), P(k|k)$ are identical to the corresponding part in the classical Kalman filter, hence like in the classical Kalman filter theory (Kalman, 1963; Jazwinski, 1970), the boundedness of $P(k|k)$ is ensured by the complete uniform observability of the $[A(k), C(k)]$ pair and the complete uniform controllability of the $[A(k), Q^{\frac{1}{2}}(k)]$ pair stated in Assumption 2. As simple corollaries of this result and Assumption 1, the matrices $\Sigma(k)$ and $K(k)$ are also bounded.

Equation (5b) implies $\Sigma(k) > R(k)$. The covariance matrix $R(k)$ is assumed to have a strictly positive lower bound, that is also a lower bound of $\Sigma(k)$.

Again under the complete uniform observability and the complete uniform controllability conditions, the homogenous part of the Kalman filter, corresponding to the homogenous system

$$\zeta(k) = [I_n - K(k)C(k)]A(k)\zeta(k-1), \qquad (7)$$

is exponentially stable (Kalman, 1963; Jazwinski, 1970). This result implies that $\Upsilon(k)$ driven by bounded $\Phi(k)$ through (5e) is bounded. It is then follows from (5f) that $\Omega(k)$ is also bounded. □

Notice that $S(k)$ was missing in Proposition 1. It is the object of the following proposition.

Proposition 2. The matrix $S(k)$ recursively computed through (5i) has a finite upper bound and a strictly positive lower bound for all $k \geq 0$. □

Proof. In order to show the upper and lower bounds of $S(k)$, let us first study another recursively defined matrix sequence

$$M(0) = S^{-1}(0) = \omega^{-1} I_p > 0 \qquad (8)$$

$$M(k) = \lambda M(k-1) + \Omega^T(k)\Sigma^{-1}(k)\Omega(k). \qquad (9)$$

It will be shown later that $S(k) = M^{-1}(k)$, but for the moment this relationship is not used.

According to Proposition 1, $\Omega(k)$ is upper bounded and $\Sigma(k)$ has a strictly positive lower bound, hence $\Omega^T(k)\Sigma^{-1}(k)\Omega(k)$ is upper bounded. Then $M(k)$ recursively generated from (9) with $\lambda \in (0,1)$ is also upper bounded. The lower bound of $M(k)$ is investigated in the following.

Repeating the recursion of $M(k)$ in (9) yields

$$M(k) = \lambda^k M(0) + \sum_{j=0}^{k-1} \lambda^j \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j)$$

For $0 \leq k \leq h$, $M(k) \geq \lambda^h M(0)$.

For $k > h$, let $[k/h]$ denote the largest integer smaller than or equal to k/h, and break the sum $\sum_{j=0}^{k-1}$ into sub-sums of h terms, then

$$\begin{aligned} M(k) &\geq \sum_{i=1}^{[k/h]} \sum_{j=(i-1)h+0}^{(i-1)h+(h-1)} \lambda^j \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j) \\ &\geq \sum_{i=1}^{[k/h]} \lambda^{ih-1} \sum_{j=(i-1)h+0}^{(i-1)h+(h-1)} \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j) \\ &\geq \sum_{i=1}^{[k/h]} \lambda^{ih-1} \alpha I_p \end{aligned}$$

where the last inequality is based on Assumption 3 (persistent excitation). The geometric sequence sum in this last result can be explicitly computed, so that

$$M(k) \geq \frac{\alpha\lambda^{h-1}(1-\lambda^{h[k/h]})}{1-\lambda^h} I_p \geq \alpha\lambda^{h-1} I_p.$$

Therefore, $M(k)$ has a strictly positive lower bound for either $0 \leq k \leq h$ or $k > h$.

Now take the matrix inverse of both sides of equation (9), and apply the matrix inversion formula

$$(A+V^TBV)^{-1} = A^{-1} - A^{-1}V^T(B^{-1}+VA^{-1}V^T)^{-1}VA^{-1}$$

with $A = \lambda M(k-1)$, $B = \Sigma^{-1}(k)$ and $V = \Omega(k)$, then

$$\begin{aligned} M^{-1}(k) =& \frac{1}{\lambda} M^{-1}(k-1) - \frac{1}{\lambda} M^{-1}(k-1)\Omega^T(k)\big[\lambda\Sigma(k) \\ &+ \Omega(k)M^{-1}(k-1)\Omega^T(k)\big]^{-1}\Omega(k)M^{-1}(k-1) \end{aligned}$$

This recursion in $M^{-1}(k)$ coincides exactly with that of $S(k)$ as formulated in (5i). Moreover, as defined in (8), $M(0) = S^{-1}(0)$, hence $M^{-1}(k) = S(k)$ for *all* $k = 0, 1, 2, \ldots$. It then follows from the already proved upper and lower bounds of $M(k)$ that $S(k)$ has also a finite upper bound and a strictly positive lower bound. □

Define the state and parameter estimation errors

$$\tilde{x}(k|k) \triangleq x(k) - \hat{x}(k|k) \qquad (10)$$

$$\tilde{\theta}(k) \triangleq \theta - \hat{\theta}(k) \qquad (11)$$

The main result of this section is stated in the following proposition.

Proposition 3. The mathematical expectations $\mathrm{E}\tilde{x}(k|k)$ and $\mathrm{E}\tilde{\theta}(k)$ tend to zero exponentially when $k \to \infty$. □

In other words, the state and parameter estimates of the adaptive Kalman filter converge respectively to the true state $x(x)$ and to the true parameter θ *in mean.* It also means that the deterministic part of the error dynamic system, ignoring the random noise terms, is exponentially stable.

Proof. It is straightforward to compute from (1), (5j) and (5l) that

$$\begin{aligned}\tilde{x}(k|k) &= A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1) + w(k) \\ &\quad - K(k)\tilde{y}(k) - \Upsilon(k)[\hat{\theta}(k) - \hat{\theta}(k-1)] \\ &= [I_n - K(k)C(k)][A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] \\ &\quad + \Upsilon(k)[\tilde{\theta}(k) - \tilde{\theta}(k-1)] \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k),\end{aligned}$$

and from (5j) and (5k) that

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) - \Gamma(k)\tilde{y}(k) \quad (12)$$

$$= \tilde{\theta}(k-1) - \Gamma(k)C(k)[A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] - \Gamma(k)C(k)w(k) - \Gamma(k)v(k). \quad (13)$$

Like in (Zhang, 2002), define

$$\xi(k) \triangleq \tilde{x}(k|k) - \Upsilon(k)\tilde{\theta}(k). \quad (14)$$

Simple substitutions lead to

$$\begin{aligned}\xi(k) &= [I_n - K(k)C(k)][A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] \\ &\quad + \Upsilon(k)[\tilde{\theta}(k) - \tilde{\theta}(k-1)] \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k) \\ &\quad - \Upsilon(k)\tilde{\theta}(k).\end{aligned}$$

In this last result, according to (14), replace $\tilde{x}(k-1|k-1)$ with

$$\tilde{x}(k-1|k-1) = \xi(k-1) + \Upsilon(k-1)\tilde{\theta}(k-1), \quad (15)$$

then

$$\begin{aligned}\xi(k) &= [I_n - K(k)C(k)]A(k)\xi(k-1) \\ &\quad + \big\{[I_n - K(k)C(k)]A(k)\Upsilon(k-1) \\ &\quad + [I_n - K(k)C(k)]\Phi(k) - \Upsilon(k)\big\}\tilde{\theta}(k-1)] \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k).\end{aligned}$$

The content of the curly braces $\{\cdots\}$ is zero, because $\Upsilon(k)$ satisfies (5e). Then

$$\begin{aligned}\xi(k) &= [I_n - K(k)C(k)]A(k)\xi(k-1) \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k). \quad (16)\end{aligned}$$

The noises $w(k)$ and $v(k)$ are assumed to have zero mean values (centered noises), then

$$\mathrm{E}\xi(k) = [I_n - K(k)C(k)]A(k)\mathrm{E}\xi(k-1).$$

Like (7), this recurrent equation is exponentially stable. Starting from

$$\mathrm{E}\xi(0) = \mathrm{E}\tilde{x}(0) - \Upsilon(0)\mathrm{E}\tilde{\theta}(0) = 0,$$

it is recursively shown that $\mathrm{E}\xi(k) = 0$ for all $k \geq 0$.

In (13) replace $\tilde{x}(k-1|k-1)$ with (15), then

$$\begin{aligned}\tilde{\theta}(k) &= \big[I_p - \Gamma(k)C(k)A(k)\Upsilon(k-1) \\ &\quad - \Gamma(k)C(k)\Phi(k)\big]\tilde{\theta}(k-1) - \Gamma(k)C(k)A(k)\xi(k-1) \\ &\quad - \Gamma(k)C(k)w(k) - \Gamma(k)v(k) \\ &= [I_p - \Gamma(k)\Omega(k)]\tilde{\theta}(k-1) - \Gamma(k)e(k) \quad (17)\end{aligned}$$

where $\Omega(k)$ is as defined in (5f), and

$$e(k) \triangleq C(k)A(k)\xi(k-1) + C(k)w(k) + v(k). \quad (18)$$

It was already shown $\mathrm{E}\xi(k) = 0$ for all $k \geq 0$, therefore $\mathrm{E}e(k) = 0$.

Take the mathematical expectation at both sides of (17) and denote

$$\bar{\tilde{\theta}}(k) \triangleq \mathrm{E}\tilde{\theta}(k), \quad (19)$$

then

$$\bar{\tilde{\theta}}(k) = [I_p - \Gamma(k)\Omega(k)]\,\bar{\tilde{\theta}}(k-1). \quad (20)$$

Before analyzing the convergence of $\bar{\tilde{\theta}}(k)$ governed by (20), let us combine the two equations (5h) and (5i) into

$$S(k) = \frac{1}{\lambda}[I_p - \Gamma(k)\Omega(k)]S(k-1). \quad (21)$$

Accordingly, $M(k) = S^{-1}(k)$ satisfies [2]

$$M(k) = \lambda M(k-1)[I_p - \Gamma(k)\Omega(k)]^{-1}. \quad (22)$$

Let us define the Lyapunov function candidate

$$V(k) \triangleq \left(\bar{\tilde{\theta}}(k)\right)^T M(k)\bar{\tilde{\theta}}(k), \quad (23)$$

it then follows from (20) and (22) that

$$\begin{aligned}V(k) &= \left(\bar{\tilde{\theta}}(k-1)\right)^T [I_p - \Gamma(k)\Omega(k)]^T \lambda M(k-1)\bar{\tilde{\theta}}(k-1) \\ &= \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T M(k-1)\bar{\tilde{\theta}}(k-1) \\ &\quad - \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T \Omega^T(k)\Gamma^T(k)M(k-1)\bar{\tilde{\theta}}(k-1). \quad (24)\end{aligned}$$

By recalling (5h), (5g) and $M(k-1) = S^{-1}(k-1)$, it yields

$$\begin{aligned}\Xi(k) &\triangleq \Omega^T(k)\Gamma^T(k)M(k-1) \quad (25)\\ &= \Omega^T(k)\Lambda(k)\Omega(k) \quad (26)\\ &= \Omega^T(k)\left[\lambda\Sigma(k) + \Omega(k)S(k-1)\Omega^T(k)\right]^{-1}\Omega(k), \quad (27)\end{aligned}$$

which is a symmetric positive semidefinite matrix. Then (24) becomes

$$\begin{aligned}V(k) &= \lambda V(k-1) - \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T \Xi(k)\bar{\tilde{\theta}}(k-1) \\ &\leq \lambda V(k-1), \quad (28)\end{aligned}$$

implying that $V(k) = \left(\bar{\tilde{\theta}}(k)\right)^T M(k)\bar{\tilde{\theta}}(k)$ converges to zero exponentially. It is already shown that $M(k)$ is lower and upper bounded, with a strictly positive lower bound, $\mathrm{E}\tilde{\theta}(k) = \bar{\tilde{\theta}}(k)$ then converges to zero exponentially.

Finally, it follows from (14) that

$$\mathrm{E}\tilde{x}(k|k) = \mathrm{E}\xi(k) + \Upsilon(k)\mathrm{E}\tilde{\theta}(k) = \Upsilon(k)\mathrm{E}\tilde{\theta}(k), \quad (29)$$

it is then concluded that the mathematical expectations $\mathrm{E}\tilde{x}(k|k)$ and $\mathrm{E}\tilde{\theta}(k)$ tend to zero exponentially when $k \to \infty$. □

5. MINIMUM COVARIANCE OF COMBINED ESTIMATION ERRORS

The following result is a generalization of the minimum variance property of the classical Kalman filter.

Proposition 4. In the adaptive Kalman filter (5), relax the Kalman gain $K(k)$ computed through the recurrent equations (5a)-(5d) to *any* matrix sequence $L(k) \in \mathbb{R}^{n\times m}$. Consider the combined state and parameter estimation error $\xi(k) = \tilde{x}(k|k) - \Upsilon(k)\tilde{\theta}(k)$. Its covariance matrix depending on the gain sequence $L(k)$ and denoted by $\mathrm{cov}[\xi(k)|L]$ reaches its minimum when $L(k) = K(k)$, in the sense of the positive definiteness of the difference matrix:

$$\mathrm{cov}[\xi(k)|L] - \mathrm{cov}[\xi(k)|K] \geq 0 \quad (30)$$

for any $L(k) \in \mathbb{R}^{n\times m}$. □

This result means that $v^T\mathrm{cov}[\xi(k)|L]v \geq v^T\mathrm{cov}[\xi(k)|K]v$ for *any* vector $v \in \mathbb{R}^n$. In fact, the term $\Upsilon(k)\tilde{\theta}(k)$

[2] According to Proposition 2, $S(k)$ is positive definite for all $k \geq 0$, hence (21) implies that the matrix $[I_p - \Gamma(k)\Omega(k)]$ is invertible.

is the part of the state estimation error $\tilde{x}(k|k)$ due to the parameter estimation error $\tilde{\theta}(k)$ (this part would be zero if the parameter estimate $\hat{\theta}(k)$ was replaced by the true parameter value θ, and then the adaptive Kalman filter (5) would be reduced to the standard Kalman filter). Therefore, the meaning of this proposition is that the remaining part of the state estimation error reaches its minimum variance if the Kalman gain $K(k)$ is used.

Proof.

Following (16) where $K(k)$ is replaced by $L(k)$, and noticing that $\xi(k-1), w(k)$ and $v(k)$ are pairwise independent, compute the covariance matrix of $\xi(k)$:

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= \mathrm{E}[\xi(k)\xi^T(k)] \\ &= [I_n - L(k)C(k)]A(k)\mathrm{cov}[\xi(k-1)|L] \\ &\quad \cdot A^T(k)[I_n - L(k)C(k)]^T \\ &\quad + [I_n - L(k)C(k)]Q(k)[I_n - L(k)C(k)]^T \\ &\quad + L(k)R(k)L^T(k) \end{aligned} \tag{31}$$

For shorter notations, let us denote

$$\Pi(k) \triangleq A(k)\mathrm{cov}[\xi(k-1)|L]A^T(k) + Q(k), \tag{32}$$

which is independent of $L(k)$ (of course, $\Pi(k)$ depends on $L(k-1)$). Then

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= [I_n - L(k)C(k)]\Pi(k)[I_n - L(k)C(k)]^T \\ &\quad + L(k)R(k)L^T(k).\end{aligned} \tag{33}$$

Define also

$$H(k) \triangleq C(k)\Pi(k)C^T(k) + R(k). \tag{34}$$

Because $C(k)\Pi(k)C^T(k) \geq 0$ and $R(k) > 0$ ($R(k)$ is a positive definite matrix, see Assumption 1), $H(k)$ is also *positive definite*, and thus invertible.

Rearrange (33) as

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= [L(k) - \Pi(k)C^T(k)H^{-1}(k)]H(k) \\ &\quad \cdot [L(k) - \Pi(k)C^T(k)H^{-1}(k)]^T + \Pi(k) \\ &\quad - \Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k).\end{aligned} \tag{35}$$

The equivalence between (33) and (35) can be shown by first developing (35) and then by incorporating (34).

The matrix $H(k)$ is *positive definite*, hence the first term in (35) (a symmetric matrix product) is positive semidefinite, that is,

$$\begin{aligned}&[L(k) - \Pi(k)C^T(k)H^{-1}(k)]H(k) \\ &\quad \cdot [L(k) - \Pi(k)C^T(k)H^{-1}(k)]^T \geq 0.\end{aligned}$$

When $L(k)$ is chosen such that this inequality becomes equality, *i.e.*, the first term of (35) is zero, $\mathrm{cov}[\xi(k)]$ reaches its minimum, since the other terms in (35) are independent of $L(k)$. This optimal choice of $L(k)$, denoted by $L_*(k)$, is

$$L_*(k) \triangleq \Pi(k)C^T(k)H^{-1}(k). \tag{36}$$

It remains to show that $L_*(k)$ is identical to the Kalman gain $K(k)$ in order to complete the proof.

Rewrite (33) while incorporating (34):

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= \Pi(k) - L(k)C(k)\Pi(k) - \Pi(k)C^T(k)L^T(k) \\ &\quad + L(k)H(k)L^T(k).\end{aligned} \tag{37}$$

In the particular case $L(k) = L_*(k) = \Pi(k)C^T(k)H^{-1}(k)$,

$$\begin{aligned}\mathrm{cov}[\xi(k)|L_*] &= \Pi(k) - 2\Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k) \\ &\quad + \Pi(k)C^T(k)H^{-1}(k)H(k)[\Pi(k)C^T(k)H^{-1}(k)]^T. \\ &= \Pi(k) - \Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k) \qquad (38) \\ &= [I_n - L_*(k)C(k)]\Pi(k) \qquad (39)\end{aligned}$$

Assemble (32), (34), (36) and (39) together, for $L = L_*$,

$$\Pi(k) = A(k)\mathrm{cov}[\xi(k-1)|L_*]A^T(k) + Q(k) \tag{40a}$$

$$H(k) = C(k)\Pi(k)C^T(k) + R(k) \tag{40b}$$

$$L_*(k) = \Pi(k)C^T(k)H^{-1}(k) \tag{40c}$$

$$\mathrm{cov}[\xi(k)|L_*] = [I_n - L_*(k)C(k)]\Pi(k). \tag{40d}$$

These equations allow recursive computation of $\mathrm{cov}[\xi(k)|L_*]$ and the other involved matrices. It turns out that these recursive computations are *exactly the same* as those in (5a)-(5d), with $\Pi(k), H(k), L_*(k)$ and $\mathrm{cov}[\xi(k)|L_*]$ corresponding respectively to $P(k|k-1), \Sigma(k), K(k)$ and $P(k|k)$.

It remains to show that these two recursive computations have the same initial condition, *i.e.* $\mathrm{cov}[\xi(0)|L_*] = P(0|0)$, in order to conclude $L_*(k) = K(k)$ for all $k \geq 0$.

Because $\Upsilon(0) = 0$ as specified in (4a), and according to the definition of $\xi(k)$ in (14),

$$\xi(0) = \tilde{x}(0|0) - \Upsilon(0)\tilde{\theta}(0) = \tilde{x}(0|0), \tag{41}$$

hence

$$\mathrm{cov}[\xi(0)|L_*] = \mathrm{cov}[\tilde{x}(0|0)] = P(0|0). \tag{42}$$

Therefore, $L_*(k) = K(k)$ for all $k \geq 0$.

It is then established that the covariance matrix $\mathrm{cov}[\xi(k)|L]$ reaches its minimum when $L(k) = K(k)$. □

6. NUMERICAL EXEMPLE

Consider a piecewise constant system randomly switching within 4 third order ($n = 3$) state-space models with one input ($l = 1$) and 2 outputs ($m = 2$). Each of the 4 state-space models is randomly generated with the Matlab code (requiring the System Identification Toolbox):

```
zreal = rand(1,1)*1.2-0.6;    pmodul = rand(1,1)*0.1+0.4;
pphase = rand(1,1)*2*pi;      preal = rand(1,1)-0.5;
ssk = idss(zpk(zreal,[pmodul.*exp(1i*pphase)  ...
      pmodul.*exp(-1i*pphase) preal],rand(1,1)+0.5,1));
```

In the simulation for $k = 0, 1, 2, \ldots, 1000$, the the actual model at each instant k randomly switches among the 4 randomly generated state-space models, which are kept unchanged. The random switching sequence among the 4 state-space models is plotted in Figure 1. The input $u(k)$ is randomly generated with a Gaussian distribution and the standard deviation equal to 2. The noise covariance matrices are chosen as $Q(k) = 0.1I_3$ and $R(k) = 0.05I_2$ for $k \geq 0$. The matrix $\Phi(k)$ is as in (2) so that θ represents actuator gain loss. During the numerical simulation running from $k = 0$ to $k = 1000$, a gain loss of 50% at the time instant $k = 500$ is simulated, corresponding to a jump of θ (a scalar parameter) from 0 to 0.5. The result of parameter estimation by the adaptive Kalman filter is presented in Fig. 2.

The above results are based on a single numerical simulation trial. In order to statistically evaluate the performance of the proposed method, 1000 simulated trials are performed, each corresponding to a different set of 4 state-space models and to a different random realization of the input and noises. At each time instant k, the histogram of the parameter estimation error based on the 1000

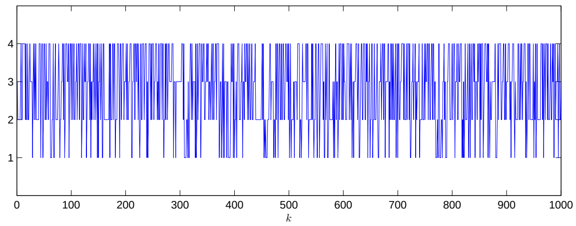

Fig. 1. Random switching index among 4 state-space models.

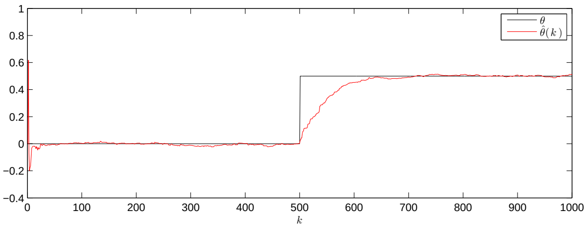

Fig. 2. The simulated "true" parameter θ and parameter estimate $\hat{\theta}(k)$ by the adaptive Kalman filter.

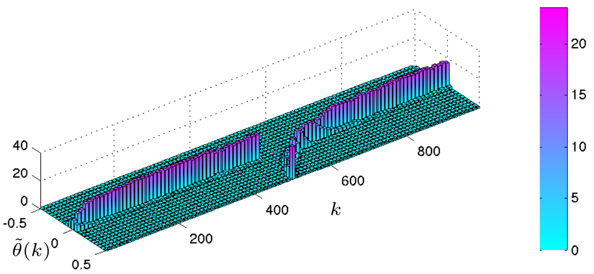

Fig. 3. Histogram per instant k of the parameter estimation error of the adaptive Kalman filter.

simulated trials is generated, and all the histograms are depicted as a 3D illustration in Fig, 3. The histograms are normalized so that they are similar to probability density functions.

7. CONCLUSION

Unlike classical adaptive Kalman filters, which have been designed for state estimation in case of uncertainties about noise covariances, the adaptive Kalman filter proposed in this paper is for the purpose actuator fault diagnosis, through joint state-parameter estimation. The stability and minimum variance properties of the adaptive Kalman filter have been rigorously analyzed. Through LTV/LPV reformulation and approximations, this method for actuator fault diagnosis is also applicable to a large class of nonlinear systems.

REFERENCES

Berdjag, D., Christophe, C., Cocquempot, V., and Jiang, B. (2006). Nonlinear model decomposition for robust fault detection and isolation using algebraic tools. *Int. Jal Innovative Computing, Information and Control*, 2(6), 1337–1354.

Besançon, G., De Leon-Morales, J., and Huerta-Guevara, O. (2006). On adaptive observers for state affine systems. *International Journal of Control*, 79(6), 581–591.

Brown, S.D. and Rutan, S.C. (1985). Adaptive Kalman filtering. *Jounal of Research of the National Bureau of Standards*, 90(6), 403–407.

Chen, J. and Patton, R. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer, Boston.

De Persis, C. and Isidori, A. (2001). A geometric approach to nonlinear fault detection and isolation. *IEEE Trans. Automatic Control*, 46(6), 853–865.

Ding, S. (2008). *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer, Berlin.

Farza, M., Bouraoui, I., Menard, T., Ben Abdennour, R., and M'Saad, M. (2014). Adaptive observers for a class of uniformly observable systems with nonlinear parametrization and sampled outputs. *Automatica*, 50(11), 2951–2960.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, New York.

Guyader, A. and Zhang, Q. (2003). Adaptive observer for discrete time linear time varying systems. In *13th IFAC/IFORS Symposium on System Identification (SYSID)*. Rotterdam.

Hwang, I., Kim, S., Kim, Y., and Seah, C. (2010). A survey of fault detection, isolation, and reconfiguration methods. *IEEE Trans. Control Systems Technology*, 18(3), 636–653. doi:10.1109/TCST.2009.2026285.

Isermann, R. (2005). *Fault Diagnosis Systems: An Introduction From Fault Detection To Fault Tolerance*. Springer, Berlin.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*, volume 64 of *Mathematics in Science and Engineering*. Academic Press, New York.

Kalman, R.E. (1963). New methods in Wiener filtering theory. In J.L. Bogdanoff and F. Kozin (eds.), *Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability*. John Wiley & Sons, New York.

Lopes dos Santos, P., Azevedo-Perdicoúlis, T.P., Ramos, J., Martins de Carvalho, J., Jank, G., and Milhinhos, J. (2011). An LPV modeling and identification approach to leakage detection in high pressure natural gas transportation networks. *IEEE Trans. Control Systems Technology*, 19(1), 77–92.

Marino, R. and Tomei, P. (1995). *Nonlinear control design*. Information and system sciences. Prentice Hall, London, New York.

Mehra, R.K. (1970). On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, 15(2), 175–184.

Ţiclea, A. and Besançon, G. (2016). Adaptive observer design for discrete time LTV systems. *International Journal of Control*, 89(12), 2385–2395.

Tóth, R., Willems, J., Heuberger, P., and Van den Hof, P. (2011). The behavioral approach to linear parameter-varying systems. *IEEE Trans. Automatic Control*, 56(11), 2499–2514.

Xu, A. and Zhang, Q. (2004). Nonlinear system fault diagnosis based on adaptive estimation. *Automatica*, 40(7), 1181–1193.

Zhang, Q. (2002). Adaptive observer for multiple-input-multiple-output (MIMO) linear time varying systems. *IEEE Trans. on Automatic Control*, 47(3), 525–529.