Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

A Bound for the Error Covariance of the Recursive Kalman Filter with Markov jump parameters

Eduardo F. Costa and Alessandro Astolfi

Abstract— **In this paper we study the error covariance matrix of the recursive Kalman filter when the parameters of the filter are driven by a Markov chain taking values in a countably infinite set. In this context, the error covariance matrix of the filter depends on the Markov state realizations, and in this sense forms a stochastic process. We show in a rather direct and comprehensive manner that a standard stochastic detectability concept plays the role of a sufficient condition for the mean of the error covariance process of the Kalman filter to be bounded. Illustrative examples are included.**

I. Introduction

The recursive Kalman filter (KF) is one of the most well-known and employed filters for dynamical systems. It is optimal in different senses, linear, recursive, computationally efficient, and allows for off-line implementations, see for instance [1], [21] and references therein. Moreover, it presents structural links with the plant that allows to characterize upper bounds for the error covariance of the estimates, stability and other fundamental aspects of the filter, based on the models of the system and noise. Along this line, it is known that the KF for a linear time varying (LTV) system presents bounded error covariance when the system is uniformly detectable[1] [2]. In general, the more accurate is the information available on the models, the more complete is the characterization of the filter properties. *A priori* exact knowledge of the system matrices for each time instant $k \geq 0$ makes possible to check detectability and thus to infer properties on the error covariance of the recursive KF, while, for instance, an uncertain model description does not allow to check uniform detectability and motivates the use of modifications to achieve robustness [24], [26], [27]

In this paper we investigate an existence condition for bounded error covariance of the KF, assuming that the filter, system and noise matrices are taken from certain sets of matrices, accordingly to a subjacent denumerable Markov chain. More precisely, we consider the recursive KF applied to discrete-time, linear infinite Markov jump parameters (LIMJP) systems. We assume that the jump variable θ (referred to as the jump state) is observed at each time instant k, and takes values in a countably infinite set $\mathscr{X} = \{1, 2, \ldots\}$. Note that this is more general than considering *a priori* exact knowledge of the system matrices, since we only know which matrix is associated with each $k \geq 0$ after observing the jump state; in a sense, an LTV system is an LIMJP system with particular choices for the transition probabilities and initial distribution of the jump state, see Section IV. LIMJP systems comprise a class of systems featuring strong properties that parallel the ones of deterministic linear systems, see for instance [7], [10] and [16]. Different formulations of filters for LIMJP systems without observation of the jump state and/or stationary filters can be found in [3], [12], [17], [18] and [22].

In the scenario when the jump state is accessible, the recursive KF is the filter of choice for many applications, in view of the aforementioned properties. However, existing results for the error covariance of the filter considers deterministic concepts such as uniform detectability, and consequently the stochastic nature of the LIMJP system is not taken into account. This is precisely the gap that this paper fills, by showing that stochastic (S) detectability implies the existence of an upper bound for the error covariance of the KF, as stated in Theorem 1. The method of proof is direct, having as a starting point the idea of comparison with a suboptimal filter, similarly e.g. to [2]. In fact, the optimality of the KF allows for an ordering with the error covariance $\tilde{X}_K$ of a linear estimator with any fixed sequence of gains $K = \{K_1, \ldots\}$ (possibly adapted to the observations). Then, we use the S-detectability hypothesis to set K as a stabilizing gain and, considering the associated conditional covariances $W_{K,i}$ (defined for each jump state i) we show that $\mathscr{E}\{W_{K,i}(k)\} \leq X_i(k)$ where X_i is the conditional second moment of a certain LIMJP system with stationary complete noise excitation. The result follows by showing that $X_i(k)$ is bounded from above, by employing available results on infinite dimensional LIMJP systems. We do not assume ergodicity nor require the existence of limiting probabilities for the Markov chain.

The paper is organised as follows. Section II presents notation, some preliminary results and the S-detectability concept. Section III shows that S-detectability ensures an upper bound for the error covariance. Section IV considers some particular scenarios and existing results, and Section V presents illustrative examples. Finally, Section VI provides some conclusions.

E.F. Costa is with Depto. de Matemática Aplicada e Estatística, Universidade de São Paulo, C.P. 668, 13560-970, São Carlos, SP, Brazil, and currently is an academic visitor at the Electrical and Electronic Engineering Dept., Imperial College London, SW7 2AZ, London, UK. efcosta@icmc.usp.br

A. Astolfi is with the Electrical and Electronic Engineering Dept., Imperial College London, SW7 2AZ, London, UK and with the Dipartimento di Informatica, Sistemi e Produzione, University of Rome "Tor Vergata", 00133 Roma, Italy. a.astolfi@imperial.ac.uk

Research supported in part by FAPESP Grants 06/02004-0 and 06/04210-6, and the EPSRC Research Grant EP/E057438, Nonlinear observation theory with applications to Markov jump systems.

[1]Please refer to Section II for the definition of detectability.

II. Definitions and preliminary results

Let $\mathbb{R}^n$ denote the $n-$th dimensional Euclidean space. Let $\mathbb{D}$ (respectively $\bar{\mathbb{D}}$) be the open (closed) unit disk in the complex plane. Let $\mathscr{R}^{r,s}$ (respectively, $\mathscr{R}^r$) represent the normed linear space formed by all $r \times s$ real matrices (respectively, $r \times r$) and $\mathscr{R}^{r0}$ the closed convex cone $\{U \in \mathscr{R}^r : U = U' \geq 0\}$ where U' denotes the transpose of U. For $V, W \in \mathscr{R}^r$, $\sigma^+(V)$ stands for the largest singular value of V, $\|V\| = \sigma^+(V)$, and $V \geq W$ indicates that $V - W \in \mathscr{R}^{r0}$. Let $\mathscr{Z} = \{1, 2, \ldots\}$ and let $\mathscr{H}^{r,s}$ denote the linear space formed by sequences of matrices $H = \{H_i \in \mathscr{R}^{r,s}; i \in \mathscr{Z}\}$ such that $\sup_{i \in \mathscr{Z}} ||H_i|| < \infty$. For $H \in \mathscr{H}^{r,s}$, we write $\|H\|_\infty = \sup_{i \in \mathscr{Z}} \|H_i\|$ and $\|H\|_1 = \sum_{i \in \mathscr{Z}} ||H_i||$. Let $\mathscr{H}_1^{r,s} \subset \mathscr{H}^{r,s}$ be such that $||H||_1 < \infty$ holds for any $H \in \mathscr{H}_1^{r,s}$. $\mathscr{H}^{r,r}$ is denoted by $\mathscr{H}^r$, and $\mathscr{H}^{r0} \subset \mathscr{H}^r$ is such that $H_i \in \mathscr{R}^{r0}$, $i \in \mathscr{Z}$, holds for any $H \in \mathscr{H}^{r0}$; similarly, $\mathscr{H}_1^r \equiv \mathscr{H}_1^{r,r}$ and $\mathscr{H}_1^{r0}$ is such that $H_i \in \mathscr{R}^{r0}$, $i \in \mathscr{Z}$, holds for any $H \in \mathscr{H}_1^{r0}$. For $U, V \in \mathscr{H}^{r,s}$, $U + V = \{U_i + V_i, i \in \mathscr{Z}\}$, $U \geq V$ indicates that $U_i \geq V_i$ for each $i \in \mathscr{Z}$, and similarly for any operation involving elements of $\mathscr{H}^{r,s}$. For each operator $\mathscr{L} : \mathscr{H}_1^{r0} \to \mathscr{H}_1^{r0}$, $r_\sigma(\mathscr{L})$ represents the spectral radius of $\mathscr{L}$. $\mathscr{E}(\cdot)$ denotes the expected value of a random variable. $I_{\mathscr{A}}(\cdot)$ is the indicator function of a set $\mathscr{A}$ and, with a slight abuse of notation, we write $I_i(\cdot)$ when $\mathscr{A} = \{i\}$.

Consider the LIMJP system defined in a fixed stochastic basis $(\Omega, \mathfrak{F}, (\mathfrak{F}_k), \mathscr{P})$ by

$$\Psi : \begin{cases} x(k+1) = A_{\theta(k)} x(k) + B_{\theta(k)} w(k), \\ y(k) = C_{\theta(k)} x(k) + D_{\theta(k)} v(k), \qquad k \geq 0, \\ x(0) = x_0,\, \theta(0) = \theta_0, \end{cases} \tag{1}$$

where $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^r$ is the output process, $w(k) \in \mathbb{R}^p$ and $v(k) \in \mathbb{R}^q$ form stationary zero-mean independent white noise processes satisfying $\mathscr{E}\{w(k)w(k)'\} = I$, $\mathscr{E}\{v(k)v(k)'\} = I$, x_0 is a (independent) zero-mean random variable satisfying $\mathscr{E}\{x_0 x_0'\} = \Sigma$, and θ is the state of an underlying discrete-time time-stationary Markov chain $\Theta = \{\theta(k); k \geq 0\}$ taking values in $\mathscr{Z}$ and having a transition probability matrix $\mathbb{P} = [p_{ij}]$, $i, j \in \mathscr{Z}$, with initial distribution $\mathscr{P}(\theta_0 = i) = \pi_i(0)$, $i \in \mathscr{Z}$. The state of the system is the pair of variables $(x(k), \theta(k))$, and we say that x is the continuous state and θ is the jump state. $A \in \mathscr{H}^n$, $B \in \mathscr{H}^{n,p}$, $C \in \mathscr{H}^{r,n}$ and $D \in \mathscr{H}^{r,q}$, with $DD' > 0$ (nonsingular measurement noise). In addition, with no loss of generality we consider $C'D = 0$. We assume observation of the output and the jump state, i.e., the available information at the time instant k is $\mathfrak{F}_k = \{y(0), \theta(0), \ldots, y(k), \theta(k)\}$.

For estimating the continuous state $x(k)$, we consider the recursive KF, which provides the estimates $\hat{x}(0) = 0$ and

$$\hat{x}(k+1) = A_{\theta(k)} \hat{x}(k) + L_k [y(k) - C_{\theta(k)} \hat{x}(k)] \tag{2}$$

where the Kalman gain $L_k = A_{\theta(k)} P(k) C'_{\theta(k)} [C_{\theta(k)} P(k) C'_{\theta(k)} + D_{\theta(k)} D'_{\theta(k)}]^{-1}$ is calculated via the Riccati difference equation (RDE)

$$\begin{aligned} P(k+1) = A_{\theta(k)} [P(k) - P(k) C'_{\theta(k)} \times \\ \times (C_{\theta(k)} P(k) C'_{\theta(k)} + D_{\theta(k)} D'_{\theta(k)})^{-1} \times \\ \times C_{\theta(k)} P(k)] A'_{\theta(k)} + B_{\theta(k)} B'_{\theta(k)} \end{aligned} \tag{3}$$

with initial condition $P(0) = \Sigma \in \mathscr{R}^{n0}$. Note that the information $\mathfrak{F}_k$ allows to calculate $P(k)$ and L_k at the time instant k (off-line implementations in which the gains L are calculated *a priori* cannot be considered); moreover, given a realization of $\theta(0), \ldots, \theta(k)$, the above filter coincides with the standard KF for LTV systems. Among the interesting features of the KF, the fact that it is a linear minimal mean square estimator is of central importance, and it is formalized in what follows in a suitable form. For a sequence of gains $K \in \mathscr{H}^{n,r}$, consider the state estimate $\hat{x}_K(0) = 0$ and

$$\hat{x}_K(k+1) = A_{\theta(k)} \hat{x}_K(k) + K_k [y(k) - C_{\theta(k)} \hat{x}_K(k)]. \tag{4}$$

We assume that K is possibly $\mathfrak{F}_k$-adapted, in the form $K_k = K_{\theta(k)}$. The estimation error is obtained from (1) and (4), $\tilde{x}(0) = x(0) - \hat{x}_K(0) = x(0)$ and

$$\begin{aligned} \tilde{x}_K(k+1) = (A_{\theta(k)} - K_k C_{\theta(k)}) \tilde{x}_K(k) \\ + B_{\theta(k)} w(k) - K_k D_{\theta(k)} v_k. \end{aligned} \tag{5}$$

Consider $\tilde{X}_K(k) \in \mathscr{R}^{n0}$ defined recursively as

$$\begin{aligned} \tilde{X}_K(k+1) &= (A_{\theta(k)} - K_k C_{\theta(k)}) \tilde{X}_K(k) (A_{\theta(k)} - K_k C_{\theta(k)})' \\ &\quad + K_k D_{\theta(k)} D'_{\theta(k)} K'_k + B_{\theta(k)} B'_{\theta(k)}, \\ \tilde{X}_K(0) &= \Sigma, \qquad k \geq t \geq 0. \end{aligned} \tag{6}$$

The connection with the error covariance is as follows, see e.g. [25].

Proposition 1. *Consider* $\tilde{X}_K$ *defined in* (6). $\mathscr{E}\{\tilde{x}(k)\tilde{x}(k)' | \mathfrak{F}_k\} = \tilde{X}_K(k)$.

It is simple to check from (3) and (6) that $P(k) = \tilde{X}_L(k)$, $k \geq 0$. The minimal mean square property of the KF can be formalized as follows.

Proposition 2. $P(k) = \tilde{X}_L(k) \leq \inf_{K \in \mathscr{H}^{n,r}} \tilde{X}_K(k)$, $k \geq 0$.

Note from (5) and Proposition 1 that the error covariance matrix $\tilde{X}_K(k)$ is a function of the random variables $\theta(0), \ldots, \theta(k)$, in such a manner that $\tilde{X}_K(k)$ (and in particular $P(k) = \tilde{X}_L(k)$) forms a stochastic process. We refer to the process as the error covariance process of the KF. We are seeking for a condition for $\mathscr{E}\{P(k)\} = \mathscr{E}\{\mathscr{E}\{\tilde{x}(k)\tilde{x}(k) | \mathfrak{F}_k\}\} \leq \bar{P}$ for some $\bar{P} \in \mathscr{R}^{n0}$. We say that a KF satisfying this relation is a *mean bounded error covariance* KF.

Regarding the system Ψ, certain linear operators have been introduced in [14]. These play an important role in characterizing stability and other aspects of LIMJP systems, see for instance [10], [16], [9], [13]. Following the notation of [7], we consider $\mathscr{L}_V : \mathscr{H}_1^r \to \mathscr{H}_1^n$, defined for $V \in \mathscr{H}_\infty^{n,r}$ and $H \in \mathscr{H}_1^{r0}$ by

$$\mathscr{L}_{V,i}(H) = \sum_{j \in \mathscr{Z}} p_{ji} V_j H_j V'_j, \quad i \in \mathscr{Z}. \tag{7}$$

It is shown in [11] that $\mathscr{L}$ in (7) is well-defined, and is a positive linear operator. We denote $\mathscr{L}^0(H) = H$, and for $k \geq 1$, we can define $\mathscr{L}^k(H)$ recursively by $\mathscr{L}^k(H) = \mathscr{L}(\mathscr{L}^{k-1}(H))$. The following result is simple to check by inspection of (7).

Proposition 3. *If $U,H \in \mathscr{H}_1^{n0}$ are such that $U \geq H$, then $\mathscr{L}_V(U) \geq \mathscr{L}_V(H)$, for each $V \in \mathscr{H}_\infty^{n,r}$.*

Notions of stochastic stability and detectability can be traced back to [19], [20]. We consider the next formulation.

Definition 1. *We say that $(A,\mathbb{P})$ is stochastically stable (S-stable) if, for each $V \in \mathscr{H}_1^{n0}$, there exists $\bar{V} \in \mathscr{R}^{n0}$ such that $\sum_{k=0}^{\infty} \mathscr{L}_A^k(V) \leq \bar{V}$. We say that $(A,C,\mathbb{P})$ is S-detectable if there exists $K \in \mathscr{H}^{n,r}$ such that $(A+KC,P)$ is S-stable.*

A sharp condition for S-stability in terms of the spectral radius of the operator $\mathscr{L}$ was obtained in [11], see also [10], [16] for similar results.

Proposition 4. $(A,\mathbb{P})$ *is S-stable if and only if $r_\sigma(\mathscr{L}_A) < 1$.*

S-stability has the following interpretation, regarding the process formed by the continuous state [11].

Proposition 5. *Consider the system Ψ with $B=0$. $(A,\mathbb{P})$ is S-stable if and only if $\mathscr{E}\{\sum_{k=0}^{\infty} \|x(k)\|^2\} < \infty$.*

For $U \in \mathscr{H}_1^{n,0}$, consider the sequence $X(k) \in \mathscr{H}_1^{n,0}$ defined as $X(0) = X_0 \in \mathscr{H}_1^{n,0}$ and

$$X(k+1) = \mathscr{L}_A(X(k)) + U, \; k \geq 0. \tag{8}$$

A condition for existence of an upper bound for $X(k)$ is presented next. If we additionally assume that the Markov chain is ergodic, the result follows as a particular case of the results in [10], which ensure convergence of $X(k)$ to a limiting X; however, since the Markov chain has no limiting distribution in general, $X(k)$ does not converge and those results cannot be used here.

Lemma 1. *Assume that $(A,\mathbb{P})$ is S-stable. Then, for each $X_0 \in \mathscr{H}_1^{n,0}$ and $U \in \mathscr{H}_1^{n,0}$ there exists $\bar{X} \in \mathscr{H}_1^{n,0}$ such that $X(k) \leq \bar{X}$, $k \geq 0$.*

Proof: Set $\varepsilon > 0$ such that $(1+\varepsilon) r_\sigma(\mathscr{L}_A) < 1$, and $\bar{A} = (1+\varepsilon)^{1/2} A$. From (7) we have that $\mathscr{L}_{\bar{A}} = (1+\varepsilon)\mathscr{L}_A < 1$ and $r_\sigma(\mathscr{L}_{\bar{A}}) = (1+\varepsilon) r_\sigma(\mathscr{L}_A) < 1$. Note from (8) that $X(k+1) = \mathscr{L}_A(X(k)) + U = (1+\varepsilon)^{-1}\mathscr{L}_{\bar{A}}(X(k)) + U$, with the solution

$$X(k) = (1+\varepsilon)^{-k}\mathscr{L}_{\bar{A}}^k(X_0) + \sum_{\ell=0}^{N-1}(1+\varepsilon)^{-\ell}\mathscr{L}_{\bar{A}}^{\ell}(U).$$

Note that $\mathscr{L}_A^k(V) \leq \sum_{k=0}^{\infty} \mathscr{L}_A^k(V)$ and, since $(\bar{A},\mathbb{P})$ is S-stable, we have from Definition 1 that $\mathscr{L}_{\bar{A}}^k(V) \leq \sum_{k=0}^{\infty} \mathscr{L}_{\bar{A}}^k(V) \leq \bar{V}$. A similar evaluation holds for $\mathscr{L}_{\bar{A}}^k(U)$. Hence,

$$\begin{aligned} X(k) &\leq (1+\varepsilon)^{-k}\bar{X} + \sum_{\ell=0}^{N-1}(1+\varepsilon)^{-\ell}\bar{U} \\ &\leq \bar{X} + (1-(1+\varepsilon)^{-1})^{-1}\bar{U}. \end{aligned}$$

Lemma 2. *There exists $U \in \mathscr{H}_1^{n,0}$ such that $\sum_{j\in\mathscr{Z}} p_{ji}\pi_j(k)\big(K_j D_j D_j' K_j' + B_j B_j'\big) \leq U$.*

Proof: Note that

$$\begin{aligned} \sum_{j\in\mathscr{Z}} & p_{ji}\pi_j(k)\big(K_j D_j D_j' K_j' + B_j B_j'\big) \\ &\leq \sum_{j\in\mathscr{Z}} \pi_j(k)\big(K_j D_j D_j' K_j' + B_j B_j'\big) \\ &\leq (\|K\|_\infty^2 \|D\|_\infty^2 + \|B\|_\infty^2) I \sum_{j\in\mathscr{Z}} \pi_j(k) \\ &\leq (\|K\|_\infty^2 \|D\|_\infty^2 + \|B\|_\infty^2) I. \end{aligned}$$

III. S-DETECTABILITY ENSURES AN UPPER BOUND FOR THE AVERAGE ERROR COVARIANCE OF THE KF

Theorem 1. *Assume $(A,C,\mathbb{P})$ is S-detectable. Then, there exists $\bar{P} \in \mathscr{R}^{n0}$ such that $\mathscr{E}\{P(k)\} \leq \bar{P}$.*

Proof: Assume that $K \in \mathscr{H}^{n,r}$ is an $\mathfrak{F}_k$-adapted sequence, with $K_k = K_{\theta(k)}$, such that $(A-KC,\mathbb{P})$ is S-stable or, equivalently, $r_\sigma(\mathscr{L}_{A-KC}) < 1$, see Proposition 4. From Proposition 2, for each realization of $\theta(0),\ldots,\theta(k)$ we have that $P(k) = \tilde{X}_L(k) \leq \tilde{X}_K(k)$, $k \geq 0$. Employing basic properties of expected value (see e.g. [4]) we write

$$\mathscr{E}\{P(k)\} \leq \mathscr{E}\{\tilde{X}_K(k)\}. \tag{9}$$

Associated with the random process $\tilde{X}_K(k)$, consider the quantity $W_K(k) \in \mathscr{H}^{n,0}$, $k \geq 0$, defined by $W_{K,i}(k) = \tilde{X}_K(k) I_i(\theta(k))$, $i \in \mathscr{Z}$. From (9) we obtain

$$\begin{aligned} \|\mathscr{E}\{P(k)\}\| &\leq \|\mathscr{E}\{\tilde{X}_K(k)\}\| = \|\sum_{j\in\mathscr{Z}} \mathscr{E}\{\tilde{X}_K(k) I_j(\theta)\}\| \\ &= \|\sum_{j\in\mathscr{Z}} \mathscr{E}\{W_{K,j}(k)\}\|. \end{aligned} \tag{10}$$

Now, consider $X(k)$ defined as in (8) with $X_{0,i} = \mathscr{E}\{W_{K,i}(0)|\mathfrak{F}_0\} = \Sigma\pi_i(0)$, $U \in \mathscr{H}_1^{n,0}$ as in Lemma 2 and A replaced by $A-KC$. We show inductively that

$$\mathscr{E}\{W_K(k)\} \leq X(k). \tag{11}$$

For $k=0$ we have that

$$\begin{aligned} \mathscr{E}\{W_{K,i}(0)\} &= \mathscr{E}\{\tilde{X}_K(0) I_i(\theta(0))\} = \Sigma\mathscr{E}\{I_i(\theta(0)=i)\} \\ &= \Sigma\pi_i(0) = X(0). \end{aligned}$$

Denoting $\Phi_i = K_i D_i D_i' K_i' + B_i B_i'$, assuming $\mathscr{E}\{W_{K,i}(k)\} \leq X_i(k)$, and employing (6) and Lemma 2, we evaluate

$$\begin{aligned} \mathscr{E}\{W_{K,i}(k+1)\} &= \mathscr{E}\{\tilde{X}_K(k+1) I_i(\theta(k+1))\} \\ &= \mathscr{E}\{\big((A_{\theta(k)} - K_{\theta(k)} C_{\theta(k)})\tilde{X}_K(k) \times \\ &\qquad \times (A_{\theta(k)} - K_{\theta(k)} C_{\theta(k)})' + \Phi_{\theta(k)}\big) I_i(\theta(k+1))\} \\ &= \sum_{j\in\mathscr{Z}} \big(\mathscr{E}\{\big((A_{\theta(k)} - K_{\theta(k)} C_{\theta(k)})\tilde{X}_K(k) \times \\ &\qquad \times (A_{\theta(k)} - K_{\theta(k)} C_{\theta(k)})' + \Phi_{\theta(k)}\big) I_i(\theta(k+1) I_j(\theta(k))\} \\ &= \sum_{j\in\mathscr{Z}} p_{ji}(A_j - K_j C_j)\mathscr{E}\{W_{K,j}(k)\}(A_j - K_j C_j)' \\ &\quad + \sum_{j\in\mathscr{Z}} p_{ji}\pi_j(k)\Phi_j \\ &\leq \sum_{j\in\mathscr{Z}} p_{ji}(A_j - K_j C_j) X_j(k) (A_j - K_j C_j)' + U_i \\ &\leq \mathscr{L}_{A-KC,i}(X(k)) + U_i = X_i(k+1), \quad i \in \mathscr{Z}, \end{aligned}$$

which completes the induction. Since $r_\sigma(\mathscr{L}_{A-KC}) < 1$, we have from Lemma 1 that there exists $\bar{X} \in \mathscr{H}_1^{n,0}$ such that $X(k) \le \bar{X}$, hence (10) and (11) yield

$$\begin{aligned}\|\mathscr{E}\{P(k)\}\| &\le \sum_{j\in\mathscr{Z}} \|\mathscr{E}\{W_{K,j}(k)\}\| \\ &\le \sum_{j\in\mathscr{Z}} \|X_j(k)\| = \|X(k)\|_1 \le \|\bar{X}\|_1.\end{aligned}$$

IV. Particular scenarios and existing results and conditions

One interesting feature of infinite dimensional LIMJP systems is that they comprise LTV systems in the sense that, if we set $p_{i,i+1} = 1$ and assume initial distribution $\pi_1(0) = 1$ and $\pi_i(0) = 0$, $i \ne 1$, then it is simple to check that $\theta(k) = k+1$, $k \ge 0$, almost surely, implying that $A_{\theta(k)} = A_{k+1}$ almost surely and similarly for other system matrices. Taking this setup into account, Theorem 1 reads as follows: if there exists a sequence of gains $K \in \mathscr{H}^{n,r}$ such that state trajectories of the system $x(k+1) = (A_k - K_k C_k)x(k)$ converges exponentially to zero, then $P(k) = \mathscr{E}\{P(k)\} \le \bar{P}$. This retrieves a result for LTV systems, see e.g. [2].

Now, consider finite dimensional LIMJP systems with $\mathscr{Z} = \{1,\dots,Z\}$. To the best of our knowledge there is no previous result ensuring bounded error covariance for the KF that takes into account the stochastic characteristics of Ψ. A question that arises in this scenario is whether the concept of weak detectability presented in [6] (see the Appendix for a definition), which is weaker than S-detectability, ensures mean bounded error covariance. Weak detectability implies that finite cost controls are stabilizing, thus playing an important role in control of LIMJP systems. The next example establishes that the answer is negative. It is a remarkable feature of LIMJP systems that the most adequate concepts for filtering and control, respectively S-detectability and weak detectability, are not equivalent.

Example 1. *Consider the system Ψ with*

$$A_1 = \begin{bmatrix} 0.9 & 0 \\ 1 & 1.5 \end{bmatrix} A_2 = \begin{bmatrix} 0.9 & 0.01 \\ 0.01 & 0.9 \end{bmatrix},$$
$$B_1 = B_2 = \Sigma = I, C_1 = C_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}, D_1 = D_2 = 1,$$
$$\mathbb{P} = \begin{bmatrix} 0.95 & 0.05 \\ 1 & 0 \end{bmatrix} \text{ and } \pi(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Employing the S-detectability test of Proposition 6 (in the Appendix), one can check that $(A,C,\mathbb{P})$ is not S-detectable and is weakly detectable. Figure 1 presents the estimated $\mathscr{E}\{P(k)\}$, obtained via Monte Carlo simulation based on $10{,}000$ realizations of the Markov chain, clearly suggesting that is is not bounded, in spite of the fact that $(A,C,\mathbb{P})$ is weakly detectable.

Finally, linear periodic systems are of course a particular case of the above ones. In this context, S-detectability, weak detectability, uniform detectability and exponential detectability are equivalent. Moreover, detectability is a sufficient condition for the KF to present bounded error covariance, which follows immediately from available results on Riccati difference equations [23].

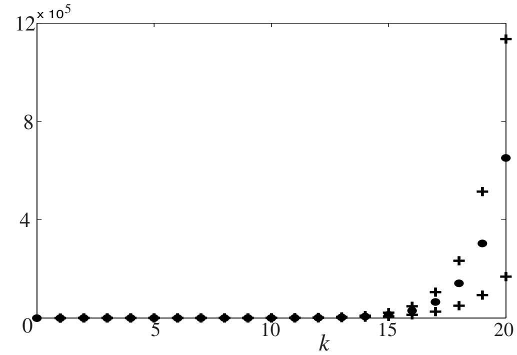

Fig. 1. Covariance error for the KF in Example 1. $\overline{\|P(k)\|}$ (marked by $\bullet$) and $\overline{\|P(k)\|} \pm \sigma$ $(+)$ where σ is the estimated standard deviation of $\|P(k)\|$.

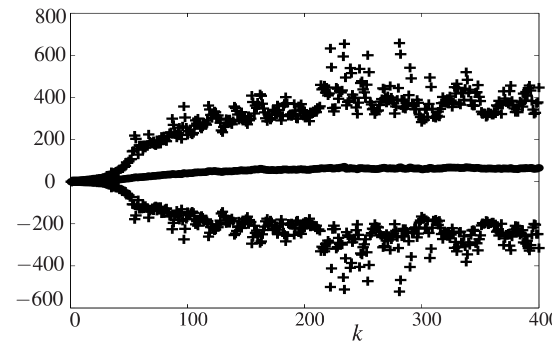

Fig. 2. Covariance error for the KF in Example 2. $\overline{\|P(k)\|}$ (marked by $\bullet$) and $\overline{\|P(k)\|} \pm \sigma$ $(+)$ where σ is the estimated standard deviation of $\|P(k)\|$.

V. Illustrative examples

Example 2. *Consider the system Ψ of Example 1 with*

$$\mathbb{P} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

The main qualitative difference from the system of Example 1 is that, now, the average occupation time in the jump state $i = 2$ increases, and A_2 is a stable matrix, leading to S-detectability of $(A,C,\mathbb{P})$ (checked using Proposition 6 in the Appendix). Figure 2 presents the estimated $\mathscr{E}\{P(k)\}$, obtained via Monte Carlo simulation based on $1{,}000$ realizations, clearly suggesting that it is bounded, as expected from Theorem 1.

Example 3. *Consider the system Ψ of Example 1 with $B_i = \Sigma_i = 0$, $i = 1,2$. Clearly, $P(k) = 0$, $k \ge 0$, for each realization of θ. On the other hand, recall from Example 1 that the system is not S-detectable, hence we conclude that S-detectability is not a necessary condition for bounded error covariance. It is interesting to mention that, in cases when the noise does not provide a complete excitation for the system, as in this example, a bounded error covariance does not ensures that the KF is stable. Figure 3 illustrates the behaviour of the estimation error $\tilde{x}(k)$ (obtained via Monte Carlo simulation based on $10{,}000$ realizations) when we consider a non-modelled error $x(0) = 1 \times 10^{-4} \ne \hat{x}(0) = 0$.*

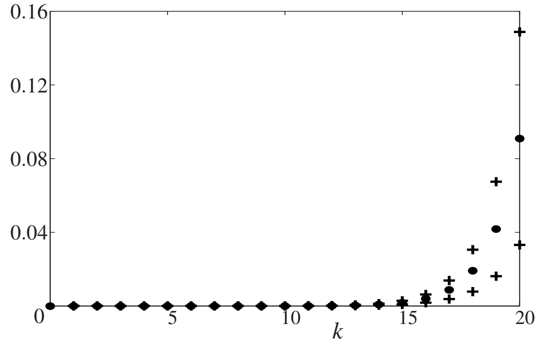

Fig. 3. Estimation error for the KF in Example 3. $\overline{\|\tilde{x}(k)\|^2}$ (marked by •) and $\overline{\|\tilde{x}(k)\|^2} \pm \sigma$ (+) where σ is the estimated standard deviation of $\|\tilde{x}(k)\|^2$.

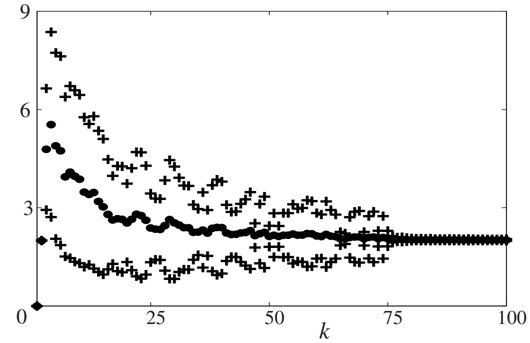

Fig. 4. Covariance error for the KF in Example 4. $\overline{\|P(k)\|}$ (marked by •) and $\overline{\|P(k)\|} \pm \sigma$ (+) where σ is the estimated standard deviation of $\|P(k)\|$.

Conditions for stability of the KF can be found in [25], [5] for LTV systems. For LIMJP systems, some case studies were presented in [8] suggesting that a notion of weak stabilizability ensures stochastic stability, assuming mean bounded error covariance*; hence, this paper represents a first step for the study of stochastic stability of the KF.*

Example 4. *Consider the system Ψ with*

$$A_i = \frac{1}{i}\begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix} \text{ and } B_i = C_i = D_i = \Sigma = I,\ i \in \mathscr{Z}.$$

Let $\mathbb{P} = [p_{ij}]$ be given by $p_{11} = p$, $p_{12} = 1-p$, $p_{i,i-1} = p$ and $p_{i,i+1} = 1-p$, $i \geq 2$. θ forms the so-called random walk process, and it is known [4, Section 6.2] that $\pi(k)$ diverges when $p < 0.5$. Let $p = 0.4$. It is difficult, in general, to check S-detectability of an infinite dimensional LIMJP system. However, in this particular setup, the matrix A_i has a simple relation with the matrix A_{i+1}, yielding that $\|A_i v\|^2 \leq \|A_{i+1} v\|^2$ for each $v \in \mathbb{R}^2$, and we can test S-detectability via a simplified version of the system Ψ. Indeed, consider the "truncated" version $\bar{\Psi}$ of system Ψ with $\bar{\mathscr{Z}} = \{1, \ldots, z\}$, and $p_{z,z-1} = p$, $p_{zz} = 1-p$, and p_{ij} as above for each $1 \leq i < z$. Let $\bar{x}$ be the corresponding continuous state. The average occupation time of the jump states $1 \leq i \leq z$ increases, and it can be checked from the above property of A that $\mathscr{E}\{\bar{x}(k)\} \leq \mathscr{E}\{x(k)\}$, $k \geq 0$. For system $\bar{\Psi}$ with $z = 10$, the test for S-detectability in Proposition 6 (in the Appendix) yields that $\bar{\Phi}$ is S-detectable, hence Proposition 5 provides that $\mathscr{E}\{\sum_{k=0}^{\infty} \|\bar{x}(k)\|^2\} < \infty$; the evaluation above allows to write $\mathscr{E}\{\sum_{k=0}^{\infty} \|x(k)\|^2\} < \mathscr{E}\{\sum_{k=0}^{\infty} \|\bar{x}(k)\|^2\} < \infty$ and from Proposition 5 we have that $(A, C, \mathbb{P})$ is S-detectable. The $\mathscr{E}\{P(k)\}$ estimated via Monte Carlo simulation based on $10,000$ realizations is illustrated in Figure 4.

VI. CONCLUDING REMARKS

The error covariance of the standard recursive KF is studied in the situation when the system matrices A_i, B_i, C_i and D_i evolves accordingly to a Markov chain. We take into account the stochastic nature of the problem ($P(k)$ forms a stochastic process) to show that S-detectability of $(A, C, \mathbb{P})$ is a sufficient condition for the existence of an upper bound in the form $\mathscr{E}\{P(k)\} \leq \bar{P}$ for some matrix $P = P' \geq 0$. The result retrieves available sufficient conditions for bounded error covariance of the KF in particular scenarios of LTV and periodic systems. Examples illustrate the obtained result and clarify that S-detectability is not necessary for mean bounded error covariance, whereas weak detectability (an existing notion that is weaker than S-detectability) is not a sufficient condition.

APPENDIX

The following testable conditions for S-detectability and weak detectability are adapted from [6] and [15].

Proposition 6. *Consider the system Ψ in (1) with $\mathscr{Z} = \{1, \ldots, z\}$. $(A, C, \mathbb{P})$ is S-detectable if and only if the LMI*

$$X_i > 0,$$
$$\begin{bmatrix} X_i & \left(A_i G_i + (C_i' C_i) M_i\right)' \\ A_i G_i + (C_i' C_i) M_i & G_i G_i' - \mathscr{E}_i(X) \end{bmatrix} < 0, \quad 1 \leq i \leq z. \tag{12}$$

in the variables $X \in \mathscr{H}^{n0}$, $G \in \mathscr{H}^n$ and $M \in \mathscr{H}^n$ is feasible. Consider $O(k) \in \mathscr{H}^{n0}$ defined recursively by $O_i(0) = C_i C_i'$ and $O_i(k) = \sum_{j=1}^{z} p_{ij} A_i' O_j(k-1) A_i$, $i \in \mathscr{Z}$. $(A, C, \mathbb{P})$ is weakly detectable if and only if $(A, O(0) + \cdots + O(n^2 z), \mathbb{P})$ is S-detectable.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, London, first edition, 1979.
[2] B. D. O. Anderson and J. B. Moore. Detectability and stabilizability of discrete-time linear systems. *SIAM Journal of Control and Optimization*, 19(1):20–32, 1981.
[3] P. E. Caines and J. Zhang. On the adaptive control of jump parameter systems via nonlinear filtering. *SIAM J. Control and Optimization*, 33(6):1758–1777, 1995.
[4] E. Çinlar. *Introduction to stochastic processes*. Prentice Hall, New York, 1975.
[5] E. F. Costa and A. Astolfi. A necessary and sufficient condition for exponential divergence of the recursive Kalman filter for linear time-varying systems. *Submitted to the SIAM Journal on Control and Optimization*.
[6] E. F. Costa and J. B. R. do Val. Weak detectability and the linear-quadratic control problem of discrete-time Markov jump linear systems. *International Journal of Control, Special Issue on Switched and Polytopic Linear Systems*, 75(16/17):1282–1292, 2002.

[7] E. F. Costa, J. B. R. do Val, and M. D. Fragoso. A new approach to detectability of discrete-time infinite Markov jump linear systems. *SIAM Journal on Control and Optimization*, 43(6):2132–2156, 2005.

[8] E. F. Costa, A. L. P. Manfrim, and J. B. R. do Val. Weak controllability and weak stabilizability concepts for linear systems with Markov jump parameters. In *Proc. ACC'06 American Control Conference*, Minneapolis,USA, 2006.

[9] O. L. V. Costa, J. B. R. do Val, and J. C. Geromel. Continuous-time state-feedback H_2-control of Markovian jump linear systems via convex analysis. *Automatica*, 35:259–268, 1999.

[10] O. L. V. Costa, M. D. Fragoso, and R. P. Marques. *Discrete-Time Markovian Jump Linear Systems*. Springer-Verlag, New York, 2005.

[11] O. L. V. Costa and M.D. Fragoso. Discrete-time LQ-optimal control problems for infinite Markov jump parameter systems. *IEEE Transactions on Automatic Control*, AC-40:2076–2088, 1995.

[12] O. L. V. Costa and S. G. Jimenez. Stationary filter for linear minimum mean square error estimator of discrete-time Markovian jump systems. *International Journal of Control*, 75:712–727, 2002.

[13] O. L. V. Costa and R.P. Marques. Robust H_2-control for discrete-time Markovian jump linear systems. *International Journal of Control*, 73(1):11–21, 2000.

[14] O. L.V. Costa and M. D. Fragoso. Stability results for discrete-time linear systems with Markovian jumping parameters. *Journal of Mathematical Analysis and Applications*, 179:154–178, 1993.

[15] J. B. R. do Val, J. C. Geromel, and A. P. C. Gonçalves. The H_2-control for jump linear systems: Cluster observations of the Markov state. *Automatica*, 38(2):343–349, 2002.

[16] M. D. Fragoso and O. L. V. Costa. A unified approach for stochastic and mean square stability of continuous-time linear systems with Markovian jumping parameters and additive disturbances. *SIAM Journal on Control and Optimization*, 44(4):1165–1191, 2005.

[17] M. D. Fragoso, O. L. V. Costa, J. Baczynski, and N. C. S. Rocha. Optimal linear mean square filter for continuous-time jump linear systems. *IEEE Transactions on Automatic Control*, 50(9):1364–1369, 2005.

[18] M. D. Fragoso and N. C. S. Rocha. Stationary filter for continuous-time Markovian jump linear systems. *SIAM Journal on Control and Optimization*, 44(3):801–815, 2005.

[19] Y. Ji and H. J. Chizeck. Controllability, observability and discrete-time Markovian jump linear quadratic control. *International Journal of Control*, 48(2):481–498, 1988.

[20] Y. Ji and H. J. Chizeck. Jump linear quadratic Gaussian control: Steady-state solution and testable conditions. *Control Theory and Advanced Technology*, 6(3):289–319, 1990.

[21] P. R. Kumar and P. Varaiya. *Stochastic Systems: estimation, identification and adaptive control*. Prentice-Hall, Inc., New Jersey, 1986.

[22] B. M. Miller and W. J. Runggaldier. Kalman filtering for linear systems with coefficients driven by a hidden Markov jump process. *Systems & Control Letters*, 31:93–102, 1997.

[23] P. Park and T. Kailath. Convergence of the RDE solution to the ARE strong solution. *IEEE Transactions on Automatic Control*, 42(4):573–578, 1997.

[24] I. R. Petersen and A. V. Savkin. *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*. Birkhauser, Boston, 1999.

[25] S. Sangsuk-Iam and T. E. Bullock. Analysis of discrete-time Kalman filtering under incorrect noise covariances. *IEEE Transactions on Automatic Control*, 35(12):1304–1309, 1990.

[26] A. H. Sayed. A framework for state-space estimation with uncertain models. *IEEE Transactions on Automatic Control*, 46(7):998–1013, 2001.

[27] F. Wang and V. Balakrishnan. Robust Kalman filters for linear time-varying systems with stochastic parametric uncertainties. *IEEE Transactions on Signal Processing*, 50(4):803–813, 2002.