Proceedings of the 18th World Congress
The International Federation of Automatic Control
Milano (Italy) August 28 - September 2, 2011

Necessary and Sufficient Conditions for Stability of Kalman Filtering with Markovian Packet Losses ⋆

Keyou You *, **Minyue Fu** ** and **Lihua Xie** *

* *School of EEE, Nanyang Technological University, 639798, Singapore. (e-mails: {youk0001,elhxie}@ntu.edu.sg)*
** *Department of Control Science and Engineering, Zhejiang University, China and School of EECS, University of Newcastle, Australia(e-mail: minyue.fu@newcastle.edu.au)*

Abstract: This paper studies the stability of Kalman filtering over a network with random packet losses, which are modeled by a Markov process. Based on the realization of the packet loss process, two stability notions, namely stability in stopping times and stability in sampling times, are introduced to analyze the behavior of the estimation error covariance matrix. For second-order systems, both the stability notions are shown to be equivalent and a necessary and sufficient condition for stability is derived. While for a certain class of higher-order systems, a necessary and sufficient condition is provided to ensure the stability of the estimation error covariance matrix at packet reception times. Their implications and relationships with related results in the literature are discussed.

1. INTRODUCTION

In this paper, we are concerned with the stability problem of Kalman filtering with random packet losses. A motivating example is given by sensor and estimator/controller communicating over a wireless channel for which the quality of the communication link varies over time because of random fading and congestion. This happens in resource limited wireless sensor networks where communications between devices are power constrained and therefore limited in range and reliability.

Kalman filtering is of great importance in networked systems due to various applications ranging from tracking, detection and control. Recently, much attention has been paid to the stability of the Kalman filter with intermittent observations, see the survey paper [Schenato et al., 2007] and references therein. The pioneering work [Sinopoli et al., 2004] analyzes the optimal state estimation problem for discrete-time linear stochastic systems under the assumption that the raw observations are randomly dropped. By modeling the packet loss process as an independent and identically distributed (i.i.d.) Bernoulli process, they prove the existence of a critical packet loss rate above which the mean state estimation error covariance matrix will diverge. However, they are unable to quantify the exact critical dropout rate for general systems except providing its lower and upper bounds, which are attainable for some special cases, e.g., the lower bound is tight if the observation matrix is invertible. A less restrictive condition is given in Plarre and Bullo [2009] where invertibility on the observable subspace is required. Another important work is Mo and Sinopoli [2010] which explicitly characterizes the dropout rate for a wider class of systems, including second-order systems and the so-called non-degenerate higher-order systems. A remarkable discovery in Mo and Sinopoli [2010] is that there are counterexamples of second-order systems for which the lower bound given by Sinopoli et al. [2004] is not tight.

On the other hand, to capture the temporal correlation of network conditions, the packet loss process is modeled by a binary Markov process in Huang and Dey [2007]. This is usually called the Gilbert-Elliott channel model. Under an i.i.d. Bernoulli packet loss, the filter stability may be effectively studied by a modified algebraic Riccati equation. In contrast, this approach is no longer feasible for the Markovian packet loss model, rendering the stability analysis more challenging. An interesting notion of peak covariance stability in the mean sense is introduced in Huang and Dey [2007]. They give sufficient conditions for the peak covariance stability for general vector systems, which is also necessary for the scalar case. A less conservative sufficient condition for some cases is provided by Xie and Xie [2008]. However, their works are restricted to the study of a random Riccati equation and do not exploit effects of system structure on stability. Thus, they fail to offer necessary and sufficient conditions for stability of the estimation error covariance matrices, except for the scalar case.

We continue to study the stability of Kalman filtering with Markovian packet losses in the present work. The system structure is exploited to investigate the effects of Markovian packet loss on stability. Motivated by You and Xie [2010], we first study the stability of the estimation error covariance matrix at packet reception times and introduce the notion of stability in stopping times. It turns out that this problem is equivalent to the stability of Kalman filtering for a stochastic time-varying linear system, whose studies can be traced back to Bougerol [1993]. However, the framework in Bougerol [1993] is quite general and not suitable to quantify the effects of Markovian packet losses on stability. Another stability notion is the usual mean square stability in the literature, which is referred to as stability in sampling times in the present paper.

Our contribution shows that for second-order systems, both stability notions are equivalent and a necessary and sufficient condition for stability of Kalman filter is given in terms of the transition probabilities of the Markov chain and the largest open pole. For higher-order systems satisfying that each eigenvalue

⋆ This work was supported in part by the National Natural Science Foundation of China under grant NSFC 60828006.

of the open-loop matrix has a distinct magnitude and associates with only one Jordan block, a simple necessary and sufficient condition for the stability in stopping times is derived. It should be noted that except for scalar systems, there is no result in the literature on the usual stability for the estimation error covariance matrix under Markovian packet losses [Huang and Dey, 2007].

The rest of the paper is organized as follows. The problem under investigation is precisely formulated in Section 2. A necessary and sufficient condition for the stability of Kalman filtering of second-order systems with Markovian packet losses is provided in Section 3. A corresponding result on higher-order system is presented in Section 4. To improve the readability of the paper, all of the proofs are given in Section 5. Concluding remarks are made in Section 6.

Notation: For a symmetric matrix M, $M \geq 0$ ($M > 0$) means that the matrix is positive semi-definite (definite), and the relation $M_1 \geq M_2$ for symmetric matrices means that $M_1 - M_2 \geq 0$. $\mathbb{N}, \mathbb{R}$ and $\mathbb{C}$ respectively denote the set of nonnegative integers, real numbers and complex numbers. $\mathrm{tr}(M)$ denotes the summation of all the diagonal elements of M.

2. PROBLEM FORMULATION

Consider the discrete-time stochastic linear system:

$$\begin{cases} x_{k+1} = Ax_k + w_k; \\ y_k = Cx_k + v_k, \end{cases} \tag{1}$$

where $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}$ are the state vector and the scalar output [1]. $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}$ are white Gaussian noises with zero means and covariance matrices $Q > 0$ and $R > 0$, respectively. The initial state x_0 is a random Gaussian vector of mean $\hat{x}_0$ and covariance matrix $P_0 > 0$. In addition, w_k, v_k and x_0 are mutually independent.

Due to random fading of the channel, packets may be lost while in transit through the network. In the present work, we ignore other effects such as quantization, transmission errors and data delays. The packet loss process is modeled as a time-homogenous binary Markov process $\{\gamma_k\}_{k\geq 0}$, which is more general and realistic than the i.i.d. case studied in Sinopoli et al. [2004] due to possible temporal correlations of network conditions. Furthermore, $\{\gamma_k\}_{k\geq 0}$ does not contain any information of the system output, suggesting that it is independent of the channel input. Let $\gamma_k = 1$ indicate that the packet containing the measurement y_k has been successfully delivered to the state estimation center while $\gamma_k = 0$ corresponds to the dropout of the packet. Moreover, the Markov chain has a transition probability matrix defined by

$$\Pi^+ = (\mathbb{P}\{\gamma_{k+1} = j | \gamma_k = i\})_{i,j\in\mathbb{S}} = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix}, \tag{2}$$

where $\mathbb{S} \triangleq \{0,1\}$ is the state space of the Markov chain. To avoid any trivial case, the failure rate p and recovery rate q of the channel are assumed to be strictly positive and less than 1, i.e., $0 < p, q < 1$, so that the Markov chain $\{\gamma_k\}_{k\geq 0}$ is ergodic. Obviously, a smaller value of p and a larger value of q indicate a more reliable channel.

[1] Due to page limitation, we restrict to the case with scalar outputs in this paper. The case with vector measurements is investigated in the full version of the paper [You et al., 2011].

Denote $(\Omega, \mathcal{F}, \mathbb{P})$ the common probability space for all random variables in the paper and let $\mathcal{F}_k \triangleq \sigma(y_i\gamma_i, \gamma_i, i \leq k) \subset \mathcal{F}$ be an increasing sequence of σ-fields generated by the information received by the estimator up to time k, i.e., all events that are generated by the random variables $\{y_i\gamma_i, \gamma_i, i \leq k\}$. In the sequel, the terminology of almost everywhere (abbreviated as *a.e.*) is always with respect to $\mathbb{P}$. Associated with the Markov chain $\{\gamma_k\}_{k\geq 0}$, the stopping time sequence $\{t_k\}_{k\geq 0}$ is introduced to denote the time at which a packet is successfully delivered. Without loss of generality, let $\gamma_0 = 1$. Then, $t_0 = 0$ and $t_k, k \geq 1$ is precisely defined by

$$\begin{aligned} t_1 &= \inf\{k : k \geq 1, \gamma_k = 1\}, \\ t_2 &= \inf\{k : k > t_1, \gamma_k = 1\}, \\ &\vdots \quad \vdots \\ t_k &= \inf\{k : k > t_{k-1}, \gamma_k = 1\}. \end{aligned} \tag{3}$$

By the ergodic property of the Markov chain $\{\gamma_k\}_{k\geq 0}$, t_k is finite *a.e.* for any fixed k [Meyn et al., 1996]. Thus, the integer valued sojourn times $\{\tau_k\}_{k>0}$ to denote the time duration between two successive packet received times are well-defined *a.e.*, where

$$\tau_k \triangleq t_k - t_{k-1} > 0. \tag{4}$$

We recall the following interesting result.

Lemma 1. [Xie and Xie, 2009] The sojourn times $\{\tau_k\}_{k>0}$ are independent and identically distributed. Furthermore, the distribution of τ_1 is explicitly expressed as

$$\mathbb{P}\{\tau_1 = i\} = \begin{cases} 1-p, & i = 1; \\ pq(1-q)^{i-2}, & i > 1. \end{cases} \tag{5}$$

As in You and Xie [2010], we shall exploit this fact to establish our results. To this purpose, denote the state estimate and prediction corresponding to the minimum mean square estimate respectively by $\hat{x}_{k|k} = \mathbb{E}[x_k|\mathcal{F}_k]$ and $\hat{x}_{k+1|k} = \mathbb{E}[x_{k+1}|\mathcal{F}_k]$. The associated error covariance matrices are defined by $P_{k|k} = \mathbb{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^H|\mathcal{F}_k]$ and $P_{k+1|k} = \mathbb{E}[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^H|\mathcal{F}_k]$, where A^H is the conjugate transpose of A. By Sinopoli et al. [2004], it is known that Kalman filter is still optimal. To be exact, the following recursions are in force:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \gamma_k K_k(y_k - C\hat{x}_{k|k-1}); \tag{6}$$

$$P_{k|k} = P_{k|k-1} - \gamma_k K_k C P_{k|k-1}, \tag{7}$$

where $K_k = P_{k|k-1}C^H(CP_{k|k-1}C^H + R)^{-1}$. In addition, the time update equations continue to hold: $\hat{x}_{k+1|k} = A\hat{x}_{k|k}, P_{k+1|k} = AP_{k|k}A^H + Q$ and $\hat{x}_{0|-1} = \bar{x}_0, P_{0|-1} = P_0$. Let $P_{k+1} = P_{k+1|k}$ and $M_k = P_{t_k+1}$, the following two types of stability notion will be investigated.

Definition 1. We say the mean estimation error covariance matrices are stable in sampling times if $\sup_{k\in\mathbb{N}} \mathbb{E}[P_k] < \infty$ while it is stable in stopping times if $\sup_{k\in\mathbb{N}} \mathbb{E}[M_k] < \infty$ for any non-negative P_0. Here the expectation is taken w.r.t. packet loss process $\{\gamma_k\}_{k\geq 0}$.

Note that the above two types of stability notion have different meanings. $\mathbb{E}[P_k]$ represents the mean prediction error covariance at time k whereas $\mathbb{E}[M_k]$ denotes the mean prediction error covariance at the stopping times of t_k. To some extent, the former is time-driven while the latter is event-driven. Our objective of this paper is to derive a necessary and sufficient condition for the two stability notions and discuss their relationship.

In consideration of Theorems 3 and 8 of Mo and Sinopoli [2010], there is no loss of generality to assume that:

A1: P_0, Q, R are all identity matrices with compatible dimensions.
A2: All the eigenvalues of A lie outside the unit circle.
A3: (C, A) is observable.

3. SECOND-ORDER SYSTEMS

At first, we look at second-order systems and make the following assumption:

A4: $A = \text{diag}(\lambda_1, \lambda_2)$ and $\lambda_2 = \lambda_1 \exp(\frac{2\pi r}{d}\mathbb{I}), \mathbb{I}^2 = -1$, where $d > r \geq 1$ and $r, d \in \mathbb{N}$ are irreducible.

The case that A does not contain complex eigenvalues is a special case in Section 4. Note that (C, A^d) is not observable. This essentially indicates that the measurements received at times $kd, k > 1$ do not help to reduce the estimation error, which will become clear in the sequel. Thus, it is intuitive that a smaller d may require a stronger condition to ensure stability of the mean estimation error covariance matrices. This observation is rigourously confirmed in the rest of this section.

3.1 Stability in stopping times

Theorem 2. Consider the system (1) satisfying A1-A4) and the packet loss process of the output follows a Markov chain with transition probability matrix (2), a necessary and sufficient condition for $\sup_{k\in\mathbb{N}} \mathbb{E}[M_k] < \infty$ is that $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$.

3.2 Stability in sampling times

Theorem 3. Consider the system (1) satisfying A1-A4) and the packet loss process of the output follows a Markov chain with transition probability matrix (2), a necessary and sufficient condition for $\sup_{k\in\mathbb{N}} \mathbb{E}[P_k] < \infty$ is that $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$.

By Theorem 2 and 3, we immediately obtain that the notions of stability in stopping times and stability in sampling times are equivalent.

Remark 4. Since $d \geq 2$, it can be verified that the function $(1+\frac{pq}{(1-q)^2})(1-q)^d$ is decreasing w.r.t. $q \in (0, 1)$ but increasing w.r.t. $p \in (0, 1)$. Thus, for a smaller p and a larger q, which corresponds to a more reliable network, we can allow a more unstable system to guarantee stability of the estimation error covariance matrices. This is consistent with our intuition.

Remark 5. The inequality conditions in Theorems 2 and 3 imply that $|\lambda_1|^2(1-q) < 1$ for all $d \geq 1$. If the conjugate complex eigenvalues satisfy that $\lambda_2 = \lambda_1 \exp(\varphi\mathbb{I})$, where $\frac{\varphi}{2\pi}$ is an irrational number, corresponding to the case $d \to \infty$ in A4), the necessary and sufficient conditions for both the types of stability simply reduce to $|\lambda_1|^2(1-q) < 1$. Under this situation, the failure rate p becomes immaterial.

In Huang and Dey [2007], the authors establish the equivalence of the usual stability (stability in sampling times) and the so-called peak covariance stability for scalar systems. But for vector systems, they give a conservative sufficient condition for the peak covariance stability and do not consider the usual stability.

Remark 6. If the packet loss process is an i.i.d. process, corresponding to $q = 1 - p$, $(1 + \frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$ is reduced to that $q > 1 - |\lambda_1|^{-\frac{2d}{d-1}}$, which recovers the result in Mo and Sinopoli [2010].

4. HIGHER-ORDER SYSTEMS

Under the i.i.d. packet loss model, an implicit characterization of a necessary and sufficient condition for stability of general vector linear systems is known to be extremely challenging [Mo and Sinopoli, 2010, Sinopoli et al., 2004, Plarre and Bullo, 2009]. Fortunately, for higher-order systems such that A^{-1} is *cyclic*, we can give a necessary and sufficient condition for the stability in stopping times.

A5: $A^{-1} = \text{diag}(J_1, \ldots, J_m)$, where $J_i = \lambda_i^{-1} I_i + N_i \in \mathbb{R}^{n_i \times n_i}$ and $|\lambda_i| > |\lambda_{i+1}|$. I_i is an identity matrix with a compatible dimension and the (u, v)-th element of N_i is 1 if $v = u + 1$ and 0, otherwise.

Now, we are in the position to deliver our main result with regards to higher-order systems.

Theorem 7. Consider the system (1) satisfying A1-A3) and A5) and the packet loss process of the output follows a Markov chain with transition probability matrix (2), a necessary and sufficient condition for $\sup_{k\in\mathbb{N}} \mathbb{E}[M_k] < \infty$ is that $|\lambda_1|^2(1-q) < 1$.

Proof. Due to page limitation, the proof is given in the full version of this paper [You et al., 2011].

5. PROOFS

Lemma 8. [Solo, 1991] For any $A \in \mathbb{R}^{n\times n}$ and $\epsilon > 0$, it holds that

$$A^k \leq M\eta^k, \forall k \geq 0, \tag{8}$$

where $M = \sqrt{n}(1 + \frac{2}{\epsilon})^{n-1}$ and $\eta = \rho(A) + \epsilon A$.

If A is invertible, define $\phi(k, i) = A^{t_i - t_k}$ if $k > i$ and $\phi(k, i) = I$ if $k \leq i$. Let

$$\Lambda_k = \sum_{j=0}^{k} \phi^H(k, j) C^H C \phi(k, j) + \phi^H(k, 0)\phi(k, 0), \tag{9}$$

$$\Xi_k = \sum_{j=0}^{k} \phi^H(j, 0) C^H C \phi(j, 0) + \phi^H(k, 0)\phi(k, 0), \tag{10}$$

$$\Xi = \sum_{j=0}^{\infty} \phi^H(j, 0) C^H C \phi(j, 0). \tag{11}$$

Lemma 9. Under A1-A3), there exist strictly positive constant numbers α and β such that $\forall k \in \mathbb{N}$,

$$\alpha A \Lambda_k^{-1} A^H \leq M_k \leq \beta A \Lambda_k^{-1} A^H. \tag{12}$$

Proof. By revising Lemma 2 in Mo and Sinopoli [2010] and $\gamma_j = 0, \forall j \notin \{t_k, k \in \mathbb{N}\}$, the proof can be established.

Lemma 10. Under A1-A3), there exist strictly positive constant numbers $\tilde{\alpha}$ and $\tilde{\beta}$ such that

$$\tilde{\alpha}A\mathbb{E}[\Xi^{-1}]A^H \le \sup_{k\in\mathbb{N}}\mathbb{E}[M_k] \le \tilde{\beta}A\mathbb{E}[\Xi^{-1}]A^H. \quad (13)$$

Proof. By Lemma 1, it is clear that the following random vectors are with an identical distribution, e.g., $(\tau_k, \tau_k + \tau_{k-1}, \ldots, \tau_k + \ldots + \tau_1) \stackrel{d}{=} (\tau_1, \tau_1 + \tau_2, \ldots, \tau_1 + \ldots + \tau_k)$, where $\stackrel{d}{=}$ means equal in distribution on its both sides. Thus, it yields that $\mathbb{E}[\Lambda_k^{-1}] = \mathbb{E}[\Xi_k^{-1}]$. Jointly with Lemma 9, it follows that

$$\mathbb{E}[M_k] \le \beta A\mathbb{E}[\Xi_k^{-1}]A^H. \quad (14)$$

Select $\epsilon < \frac{1-|\lambda_1|^{-1}}{A^{-1}}$ and $\eta = |\lambda_1|^{-1} + \epsilon A^{-1} < 1$, it follows from Lemma 8 that for any $k \in \mathbb{N}$,

$$\sum_{j=k+1}^{\infty} \phi^H(j,k)C^HC\phi(j,k) \le \|C\|^2 \sum_{j=k+1}^{\infty} \|A^{t_j - t_k}\|^2 I$$
$$\le M\|C\|^2 \sum_{j=k+1}^{\infty} \eta^{2(t_k - t_j)} I \le \frac{M\|C\|^2}{1-\eta^2} I \triangleq \beta_0 I, \quad (15)$$

where the last inequality is due to that $\tau_k \ge 1, \forall k \in \mathbb{N}$. Let $\beta_1 = \min(1, \beta_0^{-1})$ and $\tilde{\beta} = \beta\beta_1$, we further obtain that

$$\Xi_k \ge \sum_{j=0}^{k} \phi^H(j,0)C^HC\phi(j,0)$$
$$+\beta_0^{-1}\phi^H(k,0)(\sum_{j=k+1}^{\infty} \phi^H(j,k)C^HC\phi(j,k))\phi(k,0) \ge \beta_1\Xi,$$

where the second inequality is due to (15). Then, the right hand side of the inequality trivially follows from (14).

Similar to (14), the left hand side of (13) can be shown by using Fatou's Lemma [Ash and Doléans-Dade, 2000].

5.1 Proof of Theorem 2

Proof. Define the integer valued set $\mathcal{S}_d = \{kd|\ \forall k \in \mathbb{N}\}$ and $\theta = \sum_{j\in\mathcal{S}_d} \mathbb{P}\{\tau_1 = j\}$. Let $E_k, k \ge 1$ be a sequence of events defined as follows: $E_1 = \{\tau_1 \notin \mathcal{S}_d\}, E_k \triangleq \{\tau_1 \in \mathcal{S}_d, \ldots, \tau_{k-1} \in \mathcal{S}_d, \tau_k \notin \mathcal{S}_d\}, \forall k \ge 2$. By Lemma 1, it is obvious that $\mathbb{P}(E_k) = \theta^{k-1}(1-\theta)$ and $E_i \bigcap E_j = \emptyset, \forall i \ne j$. Let $F_k = \bigcup_{j=1}^k E_j$ and $F = \bigcup_{j=1}^\infty E_j$, it follows that F_k asymptotically increases to F and $\mathbb{P}(F) = \mathbb{P}(\bigcup_{j=1}^\infty E_j) = \sum_{j=1}^\infty \mathbb{P}(E_j) = 1$. Define the indicator function $1_{F_k}(w)$ which is one if $w \in F_k$, otherwise 0. It is clear that $1_{F_k} = \sum_{j=1}^k 1_{E_j}$ asymptotically increases to 1_F. Since $\mathbb{P}(F) = 1$, then $1_F = 1$ *a.e.*. By the monotone convergence theorem [Ash and Doléans-Dade, 2000], it follows that $\mathbb{E}[\Xi^{-1}] = \mathbb{E}[\Xi^{-1}1_F] = \mathbb{E}[\Xi^{-1}(\lim_{k\to\infty} 1_{F_k})] = \lim_{k\to\infty}\sum_{j=1}^k \mathbb{E}[\Xi^{-1}1_{E_j}]$.

Sufficiency: It is clear that

$$\mathbb{E}[\Xi^{-1}1_{E_j}] \le \mathbb{E}[(\sum_{i=j-1}^{j} \phi^H(i,0)C^HC\phi(i,0))^{-1}1_{E_j}].$$

In addition, we obtain that

$$\sum_{i=j-1}^{j} \phi^H(i,0)C^HC\phi(i,0) = \phi^H(j-1,0)\begin{bmatrix} c_1 & \\ & c_2 \end{bmatrix}$$
$$\times \begin{bmatrix} 1+\lambda_1^{-2\tau_j} & 1+\lambda_1^{-\tau_j}\lambda_2^{-\tau_j} \\ 1+\lambda_1^{-\tau_j}\lambda_2^{-\tau_j} & 1+\lambda_2^{-2\tau_j} \end{bmatrix} \begin{bmatrix} c_1 & \\ & c_2 \end{bmatrix} \phi(j-1,0). (16)$$

Define $\Sigma_j = \begin{bmatrix} 1+\lambda_1^{-2\tau_j} & 1+\lambda_1^{-\tau_j}\lambda_2^{-\tau_j} \\ 1+\lambda_1^{-\tau_j}\lambda_2^{-\tau_j} & 1+\lambda_2^{-2\tau_j} \end{bmatrix}$, it follows that if $\tau_j \notin \mathcal{S}_d$, we have $\Sigma_j^{-1} \le \frac{4}{\lambda_1^{-2\tau_j}+\lambda_2^{-2\tau_j}-2\lambda_1^{-\tau_j}\lambda_2^{-\tau_j}} I \le \frac{2|\lambda_1|^{2\tau_j}}{1-\cos(\frac{2\pi}{d})} I$. Thus, let $c = \max(c_1^{-2}, c_2^{-2})$, it follows from (16) that if $\tau_j \notin \mathcal{S}_d$, $(\sum_{i=j-1}^j \phi^H(i,0)C^HC\phi(i,0))^{-1} \le \frac{2c|\lambda_1|^{2t_k}}{1-\cos(\frac{2\pi}{d})} I$. Combining the above, we obtain that $\mathbb{E}[\Xi^{-1}] \le \frac{2cI}{1-\cos(\frac{2\pi}{d})} \lim_{k\to\infty}\sum_{j=1}^k \mathbb{E}[|\lambda_1|^{2t_j}1_{E_j}]$. In addition, the following statements are true: $\lim_{k\to\infty}\sum_{j=1}^k \mathbb{E}[|\lambda_1|^{2t_j}1_{E_j}] = \lim_{k\to\infty}\sum_{j=1}^k \mathbb{E}[(\prod_{i=1}^{j-1}|\lambda_1|^{2\tau_i}1_{\{\tau_i\in\mathcal{S}_d\}})|\lambda_1|^{2\tau_j}1_{\{\tau_j\notin\mathcal{S}_d\}}] \le \lim_{k\to\infty}\mathbb{E}[|\lambda_1|^{2\tau_1}]\sum_{j=1}^k(\mathbb{E}[|\lambda_1|^{2\tau_1}1_{\{\tau_1\in\mathcal{S}_d\}}])^{j-1}$, which is finite if $\mathbb{E}[|\lambda_1|^{2\tau_1}] < \infty$ and $\mathbb{E}[|\lambda_1|^{2\tau_1}1_{\{\tau_1\in\mathcal{S}_d\}}] < 1$. By some algebraic manipulations, it is clear that $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$ is equivalent to that $|\lambda_1|^2(1-q) < 1$ and $\frac{pq}{(1-q)^2}\frac{(|\lambda_1|^2(1-q))^d}{1-(|\lambda_1|^2(1-q))^d} < 1$. Together with Lemma 1, it implies that $\mathbb{E}[|\lambda_1|^{2\tau_1}] < \infty$ and

$$\mathbb{E}[|\lambda_1|^{2\tau_1}1_{\{\tau_1\in\mathcal{S}_d\}}] = \frac{pq}{(1-q)^2}\frac{(|\lambda_1|^2(1-q))^d}{1-(|\lambda_1|^2(1-q))^d} < 1.$$

Necessity: Denote $\Xi'_k = \sum_{j=0}^k \phi^H(j,0)C^HC\phi(j,0)$. By (15), we have that

$$\Xi = \Xi'_{j-1} + \phi^H(j,0)(C^HC + \sum_{i=j+1}^{\infty} \phi^H(i,j)C^HC\phi(i,j))$$
$$\times\phi(j,0) \le \Xi'_{j-1} + \phi^H(j,0)(C^HC + \beta_0 I)\phi(j,0). \quad (17)$$

Then, let $\beta_2^{-1} = \max(\frac{1}{1-|\lambda_1|^{-2}}, 1, \beta_0)$, it follows that

$$\Xi^{-1}1_{E_j} \ge (\Xi'_{j-1} + \phi^H(j,0)(C^HC+\beta_0 I)\phi(j,0))^{-1}1_{E_j}$$
$$= (\sum_{i=0}^{j-1}|\lambda_1|^{-2t_i}C^HC + \phi^H(j,0)(C^HC+\beta_0 I)\phi(j,0))^{-1}1_{E_j}$$
$$\ge (\frac{1}{1-|\lambda_1|^{-2}}C^HC + \phi^H(j,0)(C^HC+\beta_0 I)\phi(j,0))^{-1}1_{E_j}$$
$$\ge \beta_1(C^HC + \phi^H(j,0)(C^HC+I)\phi(j,0))^{-1}1_{E_j}. \quad (18)$$

In view of Lemma 10, $\sup_{k\in\mathbb{N}}\mathbb{E}[M_k] < \infty$ is equivalent to $\mathbb{E}[\Xi^{-1}] < \infty$. This implies that

$$\lim_{k\to\infty}\sum_{j=1}^k \mathbb{E}[(C^HC+\phi^H(j,0)(C^HC+I)\phi(j,0))^{-1}1_{E_j}] < \infty. \quad (19)$$

By some manipulations, one can verify that there exists a positive constant $\beta_3 > 0$ such that $\mathrm{tr}(C^HC+\phi^H(j,0)(C^HC+I)\phi(j,0))^{-1}1_{E_j}) \ge \beta_3|\lambda_1|^{2t_j}1_{E_j}$. In light of (19), we obtain that

$$\lim_{k\to\infty}\sum_{j=1}^{k}\mathbb{E}[|\lambda_1|^{2t_j}1_{E_j}]$$

$$= \mathbb{E}[|\lambda_1|^2 1_{\{\tau_1\notin\mathcal{S}_d\}}]\lim_{k\to\infty}\sum_{j=1}^{k}(\mathbb{E}[|\lambda_1|^2 1_{\{\tau_1\in\mathcal{S}_d\}}])^{j-1} < \infty,$$

which is equivalent to $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$ by Lemma 1.

5.2 Proof of Theorem 3

Proof. Given an arbitrary $j \in \mathbb{N}$, there exists an $i \in \mathbb{N}$ such that $id \le j \le (i+1)d$. Then, $P_{j+1} \le |\lambda_1|^{2(j-id)}P_{id+1} + \sum_{i=0}^{j-id-1}|\lambda_1|^{2i}I \le |\lambda_1|^{2d}P_{id+1} + \frac{1-|\lambda_1|^{2d}}{1-|\lambda_1|^2}I$. This implies that $\sup_{i\in\mathbb{N}}\mathbb{E}[P_{id+1}] \le \sup_{j\in\mathbb{N}}\mathbb{E}[P_j] \le |\lambda_1|^{2d}\sup_{i\in\mathbb{N}}\mathbb{E}[P_{id+1}] + \frac{1-|\lambda_1|^{2d}}{1-|\lambda_1|^2}I$. Thus, it follows that $\sup_{k\in\mathbb{N}}\mathbb{E}[P_k] < \infty$ if and only if $\sup_{k\in\mathbb{N}}\mathbb{E}[P_{kd+1}] < \infty$.

Sufficiency: Consider a large $k \in \mathbb{N}$ and define
$k_0 = \inf\{j \le k | jd \in \{t_k, k \in \mathbb{N}\} \text{ and } \forall t_i \in (jd, kd], \tau_i \in \mathcal{S}_d\}$.
With regard to the value of k_0, two cases are separately discussed.

C1: If $k_0 = 0$, let $i_0 = 0$. By A1), we obtain that $P_{k_0 d+1} \le AP_0A^H + Q = AA^H + I$.
C2: If $k_0 > 0$, there must exist an $i_0 > 0$ such that $k_0 d = t_{i_0}$ and $\tau_{i_0} < d$. In view of Lemma 9, it follows that

$$\begin{aligned}P_{k_0 d+1} &\le \beta A(\sum_{j=i_0-1}^{i_0}\phi^H(i_0,j)C^HC\phi(i_0,j))^{-1}A^H\\ &= \beta A(C^HC + (A^{-\tau_{i_0}})^HC^HCA^{-\tau_{i_0}})^{-1}A^H\\ &= \beta A\begin{bmatrix}c_1^{-1} & \\ & c_2^{-1}\end{bmatrix}\Sigma_{i_0}^{-1}\begin{bmatrix}c_1^{-1} & \\ & c_2^{-1}\end{bmatrix}A^H\\ &\le \frac{2c\beta|\lambda_1|^{2d}}{1-\cos(\frac{2\pi}{d})}AA^H. \end{aligned}\tag{20}$$

By defining $\Delta = (1+\frac{2c\beta|\lambda_1|^{2d}}{1-\cos(\frac{2\pi}{d})})AA^H + I$, it yields that

$$P_{k_0 d+1} \le \Delta, \forall k_0 \in \{0,\ldots,k\}. \tag{21}$$

Consider the integer compositions of $k-k_0$. Here a composition of $k-k_0$ is a way of writing $k-k_0$ as the sum of a sequence of strictly positive integers, i.e., write $k-k_0 = \delta_1+\ldots+\delta_l$, where $1 \le l \le k-k_0-1$ and $\delta_i \ge 1, \forall i \in \{1,\ldots,l\}$. By fixing the composition size l, it is clear that there are $\binom{k-k_0-1}{l-1}$ possible integer compositions, where $\binom{n}{k}$ is the binomial coefficient. In addition, for each composition of size l, it follows from Lemma 1 that the probability of $\{\tau_{i_0+1} = \delta_1 d,\ldots,\tau_{i_0+l-1} = \delta_{l-1}d, \tau_{i_0+l} \ge \delta_l d\}$ is computed by

$$\begin{aligned}&\mathbb{P}\{\tau_{i_0+1} = \delta_1 d,\ldots,\tau_{i_0+l-1} = \delta_{l-1}d,\tau_{i_0+l} \ge \delta_l d\}\\ &= \prod_{j=1}^{l-1}\mathbb{P}\{\tau_j = \delta_j d\}\mathbb{P}\{\tau_l \ge \delta_l d\}\\ &= (\prod_{j=1}^{l-1}\frac{pq}{(1-q)^2}(1-q)^{\delta_j d})(\frac{pq}{(1-q)^2}(1-q)^{\delta_l d}\sum_{j=0}^{\infty}(1-q)^j)\\ &= \frac{1}{q}(\frac{pq}{(1-q)^2})^l(1-q)^{(k-k_0)d}.\end{aligned}\tag{22}$$

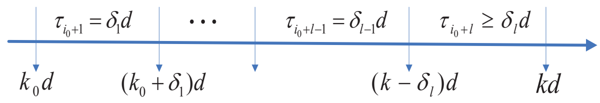

Fig. 1. A composition

For the integer $k-k_0$, there are $\binom{k-k_0-1}{l-1}$ possible compositions of size l.

The total probability for all the possible compositions with size l is denoted by $\mathbb{P}(l|k_0)$ and computed as follows:

$$\mathbb{P}(l|k_0) = \frac{1}{q}\binom{k-k_0-1}{l-1}(\frac{pq}{(1-q)^2})^l(1-q)^{(k-k_0)d}.$$

Let $s_k = \sum_{k_0=0}^{k-1}\sum_{l=1}^{k-k_0-1}|\lambda_1|^{2(k-k_0)d}\mathbb{P}(l|k_0)$, we obtain that

$$\begin{aligned}s_k &= \frac{1}{q}\sum_{j=1}^{k}\sum_{l=1}^{j-1}\binom{j-1}{l-1}(\frac{pq}{(1-q)^2})^l(|\lambda_1|^2(1-q))^{jd}\\ &= \frac{p}{(1-q)^2}\sum_{j=1}^{k}(|\lambda_1|^2(1-q))^{jd}\sum_{l=0}^{j-1}\binom{j-1}{l}(\frac{pq}{(1-q)^2})^l\\ &= \frac{p}{(1-q)^2}\sum_{j=1}^{k}(|\lambda_1|^2(1-q))^{jd}(1+\frac{pq}{(1-q)^2})^{j-1}.\end{aligned}\tag{23}$$

Now, we are ready to compute $\mathbb{E}[P_{kd+1}]$ as follows:

$$\begin{aligned}&\mathbb{E}[P_{kd+1}] = \mathbb{E}[\mathbb{E}[P_{kd+1}|k_0]] \le \sum_{k_0=0}^{k}\mathbb{E}[P_{kd+1}|k_0]\\ &= \sum_{k_0=0}^{k-1}\sum_{l=1}^{k-k_0-1}\mathbb{E}[P_{kd+1}|k_0,l]\mathbb{P}(l|k_0) \le \sum_{k_0=0}^{k-1}\sum_{l=1}^{k-k_0-1}\\ &\mathbb{E}\left[|\lambda_1|^{2(k-k_0)d}P_{k_0 d+1} + \frac{|\lambda_1|^{2(k-k_0)d}-1}{|\lambda_1|^2-1}I\right]\mathbb{P}(l|k_0)\\ &\le s_k\Delta + \frac{s_k}{|\lambda_1|^2-1}I,\end{aligned}\tag{24}$$

where the first inequality is due to that $P_{k+1} \le |\lambda_1|^2P_k + I, \forall k \in \mathbb{N}$ under A1). Together with (23), we immediately concludes that $\sup_{k\in\mathbb{N}}\mathbb{E}[P_{kd+1}] < \infty$ if $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$.

Necessity: By (2), it is clear that the binary state Markov chain is irreducible and aperiodic. Thus, it has an unique stationary distribution. In particular,

$$\mathbb{P}\{\gamma_\infty = i\} = \lim_{k\to\infty}\mathbb{P}\{\gamma_k = i\} = \frac{p^{1-i}q^i}{p+q}, \forall i \in \mathbb{S}. \tag{25}$$

Now, consider a special case that the Markov chain starts at this stationary distribution, i.e., $\mathbb{P}\{\gamma_0 = i\} = \frac{p^{1-i}q^i}{p+q}, \forall i \in \mathbb{S}$. It follows that the distribution of $\gamma_k, \forall k \in \mathbb{N}$ is the same distribution as that of γ_0. Under this situation, it can be verified that

$$\Pi^- = (\mathbb{P}\{\gamma_k = j|\gamma_{k+1} = i\})_{i,j\in\mathbb{S}} = \begin{bmatrix}1-q & q\\ p & 1-p\end{bmatrix}. \tag{26}$$

For any measurable function $f : \mathbb{R}^{k+1} \to \mathbb{R}^{n\times n}$, it is easy to establish the following result:

$$
\begin{aligned}
&\mathbb{E}[f(\gamma_k,\ldots,\gamma_0)] \\
&= \sum_{i_j\in\mathbb{S},0\le j\le k} f(i_k,\ldots,i_0)\mathbb{P}\{\gamma_k=i_k,\ldots,\gamma_0=i_0\} \\
&= \sum_{i_j\in\mathbb{S},0\le j\le k} f(i_k,\ldots,i_0)\mathbb{P}\{\gamma_0=i_0\} \\
&\qquad\times\prod_{j=0}^{k-1}\mathbb{P}\{\gamma_{j+1}=i_{j+1}|\gamma_j=i_j\} \qquad (27)\\
&= \sum_{i_j\in\mathbb{S},0\le j\le k} f(i_k,\ldots,i_0)\mathbb{P}\{\gamma_k=i_0\} \\
&\qquad\times\prod_{j=0}^{k-1}\mathbb{P}\{\gamma_j=i_{j+1}|\gamma_{j+1}=i_j\} \qquad (28)\\
&= \mathbb{E}[f(\gamma_0,\ldots,\gamma_k)] = \mathbb{E}[f(\gamma_1,\ldots,\gamma_{k+1})], \qquad (29)
\end{aligned}
$$

where (27) follows from (2) while (28) is due to (2), (26) and that the distribution of γ_k is the same as that of γ_0. The last equality is due to the strict stationarity of the Markov chain. By Lemma 3 of Mo and Sinopoli [2010], there exists a positive constant β_4 such that

$$
\begin{aligned}
&P_{k+1} \\
&\ge \beta_4\Big(\sum_{i=1}^{k+1}\gamma_{k+1-i}(A^{-i})^H C^H C A^{-i} + (A^{-k-1})^H A^{-k-1}\Big)^{-1}.
\end{aligned}
$$

Together with (29), we have that

$$
\begin{aligned}
&\mathbb{E}[P_{k+1}] \\
&\ge \beta_4\mathbb{E}\Big(\sum_{i=1}^{k+1}\gamma_i(A^{-i})^H C^H C A^{-i} + (A^{-k-1})^H A^{-k-1}\Big)^{-1} \\
&\ge \beta_4\mathbb{E}\Big(\sum_{i=1}^{\infty}\gamma_i(A^{-i})^H C^H C A^{-i} + (A^{-k-1})^H A^{-k-1}\Big)^{-1}. \qquad (30)
\end{aligned}
$$

Note that under A2) and A4), the term in (30) is decreasing w.r.t. k, which, jointly with monotone convergence theorem [Ash and Doléans-Dade, 2000], implies that $\sup_{k\in\mathbb{N}}\mathbb{E}[P_{k+1}] \ge \beta_4\mathbb{E}[(\sum_{i=1}^{\infty}\gamma_i(A^{-i})^H C^H C A^{-i})^{-1}] = \beta_4\mathbb{E}[\Xi^{-1}]$, where the last equality follows from the definition of Ξ in (11). Finally, by Lemma 10 and the proof of necessity in Theorem 2, we get that $(1+\frac{pq}{(1-q)^2})(|\lambda_1|^2(1-q))^d < 1$.

6. CONCLUSION

We have examined the stability of Kalman filtering with Markovian packet losses. Two stability notions have been introduced to analyze estimation error covariances of the Kalman filtering. For second-order systems, the two stability notions have been shown to be equivalent and a necessary and sufficient condition is obtained for ensuring stability. Under a certain class of higher-order systems, a necessary and sufficient condition has been derived to guarantee the stability of the estimation error covariance matrix in stopping times. All of the results can recover the related results in the existing literature. Our future works are to characterize the relationship of the two stability notions and find the corresponding stability conditions for general vector systems.

REFERENCES

R. Ash and C. Doléans-Dade. *Probability and Measure Theory*. Academic Press, 2000.

P. Bougerol. Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, 31:942–959, 1993.

M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica*, 43(4):598–607, 2007.

S. Meyn, R. Tweedie, and J. Hibey. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1996.

Y. Mo and B. Sinopoli. Towards finding the critical value for Kalman filtering with intermittent observations. *http://arxiv.org/abs/1005.2442*, 2010.

K. Plarre and F. Bullo. On Kalman filtering for detectable systems with intermittent observations. *IEEE Transactions on Automatic Control*, 54(2):386–390, 2009.

L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163–187, 2007.

B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9): 1453–1464, 2004.

V. Solo. One step ahead adaptive controller with slowly time-varying parameters. Technical report, Dept. ECE, John Hopkins University, Baltimore, 1991.

L. Xie and L. Xie. Stability of a random Riccati equation with Markovian binary switching. *IEEE Transactions on Automatic Control*, 53(7):1759–1764, 2008.

L. Xie and L. Xie. Stability analysis of networked sampled-data linear systems with markovian packet losses. *IEEE Transactions on Automatic Control*, 54(6):1368–1374, 2009.

K. You and L. Xie. Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses. *IEEE Transactions on Automatic Control(in press)*, 2010.

K. You, M. Fu, and L. Xie. Mean square stability for Kalman filtering with markovian packet losses. *Automatica(provisionally accepted)*, January 2011.