Available online at www.sciencedirect.com
ScienceDirect
Automatica 43 (2007) 598–607
automatica
www.elsevier.com/locate/automatica

Stability of Kalman filtering with Markovian packet losses☆

Minyi Huang[a,*,1], Subhrakanti Dey[b]

[a]*Department of Information Engineering, RSISE, The Australian National University, Canberra, ACT, Australia*
[b]*Department of Electrical and Electronic Engineering, The University of Melbourne, VIC, Australia*

Received 11 November 2005; received in revised form 19 March 2006; accepted 19 October 2006

Abstract

We consider Kalman filtering in a network with packet losses, and use a two state Markov chain to describe the normal operating condition of packet delivery and transmission failure. Based on the sojourn time of each visit to the failure or successful packet reception state, we analyze the behavior of the estimation error covariance matrix and introduce the notion of peak covariance, as an estimate of filtering deterioration caused by packet losses, which describes the upper envelope of the sequence of error covariance matrices $\{P_t, t \geqslant 1\}$ for the case of an unstable scalar model. We give sufficient conditions for the stability of the peak covariance process in the general vector case, and obtain a sufficient and necessary condition for the scalar case. Finally, the relationship between two different types of stability notions is discussed.

Keywords: Networked systems; Packet losses; Kalman filtering; Stability

1. Introduction

The problem of state estimation is of great importance in various applications ranging from tracking, detection and control, and in linear stochastic dynamical systems, Kalman filtering (Kailath, Sayed, & Hassibi, 2000; Kalman, 1960) plays an essential role. Recently there has been an increased research attention for filtering in distributed systems where sensor measurements and final signal processing take place in geographically separate locations and the usage of wireless or wireline communication channels is essential for data communication. In contrast to traditional filtering problems, an important feature in these networked systems is that the delivery of measurements to the estimator is not always reliable and losses of data may occur. This leads to estimation schemes which are required to handle missing data.

In this paper, we consider optimal filtering in a linear system with random packet losses. When the observer has full information about the loss of each packet, this leads to a modified filtering structure switching between the conventional Kalman filter when packets are received, and a deterministic predictor when a packet loss occurs.

We focus on the n dimensional linear time-invariant system

$$x_{t+1} = Ax_t + w_t, \quad t \geqslant 0,$$

where the initial state is x_0 at $t = 0$. The sensor measurements are obtained starting from $t \geqslant 1$ in the form

$$y_t^0 = Cx_t + v_t, \quad t \geqslant 1,$$

where $C \in \mathbb{R}^{m\times n}$, and then y_t^0 is transmitted by a channel. Here $\{w_t, t \geqslant 0\}$ and $\{v_t, t \geqslant 1\}$ are two mutually independent sequences of independent and identically distributed (i.i.d.) Gaussian noises with covariance matrices Q and $R > 0$, respectively. The two noise sequences are also independent of x_0, which is a Gaussian random vector with mean $\bar{x}_0 = Ex_0$ and covariance matrix P_{x_0}. The underlying probability space is denoted as $(\Omega, \mathbb{F}, \mathbb{P})$ where $\mathbb{F}$ is the σ-algebra of all events.

We consider a communication channel such that y_t^0 is exactly retrieved or the packet containing y_t^0 is lost due to corrupted data

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor George Yin under the direction of Editor Ian Petersen.

* Corresponding author. Tel.: +61 2 6125 8646; fax: +61 2 6125 8660.
E-mail addresses: minyi.huang@rsise.anu.edu.au (M. Huang), s.dey@ee.unimelb.edu.au (S. Dey).

[1] M. Huang was with Department of Electrical and Electronic Engineering, The University of Melbourne, where he performed his work.

0005-1098/$ - see front matter
doi:10.1016/j.automatica.2006.10.023

or substantial delay. When the packet is successfully received, one obtains the observation

$$y_t = y_t^0$$

and if there is a packet loss, by our convention, the observation obtained by the receiver is

$$y_t \equiv 0.$$

Under this assumption, the underlying communication link may be looked at as an erasure channel at the packet level.

We use $\gamma_t \in \{0, 1\}$ to indicate the arrival (with value 1) or loss (with value 0) of packets. Here γ_t may be interpreted as resulting from the physical operating condition of a network. Specifically, the state 0 for γ_t may correspond to channel error or network congestion which causes a straight packet loss or long delay resulting in packet dropping at the receiver. For facilitating the presentation, 0 and 1 shall be called the failure state and normal state, respectively. To capture the temporal correlation of the channel variation (e.g., in bursty error conditions), γ_t is modelled by a two state Markov chain with the transition matrix

$$\alpha = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix}, \tag{1}$$

where p and q, respectively, are called the failure rate and recovery rate and $p, q > 0$. For instance, $1 - p$ denotes the probability of the channel remaining at the normal state 1 after one step if it starts with state 1. This is usually called the Gilbert–Elliott channel model (Elliott, 1963; Gilbert, 1960). Obviously, a small value (close to 0) for p and a large value (close to 1) for q mean the channel is more reliable.

Based on the history $\mathbb{F}_t = \sigma(y_i, \gamma_i, i \leqslant t)$, which is the σ-algebra generated by the available information up to time t (i.e., all events that can be generated by these random variables), one can write a set of filtering and prediction equations corresponding to the optimal estimate $\hat{x}_t = E[x_t|\mathbb{F}_t]$ and $\hat{x}_{t+1|t} = E[x_{t+1}|\mathbb{F}_t]$, $t \geqslant 0$, respectively, by the same method as in Sinopoli et al. (2004) which dealt with the scenario of i.i.d. packet losses. We use the convention $\mathbb{F}_0 = \{\emptyset, \Omega\}$. The details for the recursion of $\hat{x}_t$ and $\hat{x}_{t+1|t}$ will not be repeated here. In this paper we focus on the estimation error of $\hat{x}_{t+1|t}$ with an associated prediction error covariance matrix

$$P_{t+1|t} \triangleq E(x_{t+1} - \hat{x}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})'.$$

We write $P_{t+1|t} = P_{t+1}$. We use M' to denote the transpose of a vector or matrix M. To characterize the prediction error covariance, one can easily derive the following random Riccati equation

$$P_{t+1} = AP_tA' + Q - \gamma_t AP_tC'(CP_tC' + R)^{-1}CP_tA', \quad t \geqslant 1. \tag{2}$$

The initial condition in (2) is $P_1 = Var(x_1) = AP_{x_0}A' + Q$. Note that γ_t appears as a random coefficient in the recursion.

Under a Bernoulli i.i.d. packet loss modelling, the filtering stability may be effectively studied by a modified algebraic Riccati equation (MARE), which is obtained by replacing γ_t in Eq. (2) by the packet arrival rate λ. Subsequently, the analysis amounts to identifying a critical value λ_c such that stability holds if and only if the arrival rate is greater than λ_c (see Section 4 for additional discussion). This approach is generally termed as being based on the uncertainty threshold principle (Sinopoli et al., 2004). In contrast, when the channel model is given by a Markov chain, such a conversion into a deterministic MARE is no longer feasible, and since the channel is described by several independent parameters, the usual threshold argument is not applicable.

1.1. Background and related work

Filtering and estimation constitute an important aspect in sensor network deployment for monitoring, detection or tracking (Chong & Kumar, 2003; Zhang, Moura, & Krogh, 2005; Zhao, Shin, & Reich, 2002), as well as multi-vehicle coordination (Varaiya, 1993), since in reality sensors can only obtain noisy information about a physical activity in its vicinity. And for many linear stochastic models, a useful tool is the standard Kalman filtering theory which has been widely used in various estimation and control scenarios. Recently there is an increased attention for its application in distributed networks while new theoretical questions and implementation issues emerge. In close relation to estimation in lossy sensor networks, there also has been a long history of research on filtering with missing signals at certain points of time, i.e., the output does not necessarily contain the signal in question and it may be only a noise component. Such models were referred to as systems with uncertain observations; see (Hadidi & Schwartz, 1979; Jaffer & Gupta, 1971; Nahi, 1969; Tugnait, 1981). The early work (Nahi, 1969) considered optimal state estimation within the class of linear filters; by modelling the uncertainty via a sequence of i.i.d. binary random variables indicating the signal availability, the author derived a recursion similar to the Kalman filter utilizing the statistics of the unobserved binary uncertainty sequence (Nahi, 1969). The work (Hadidi & Schwartz, 1979) gave conditions for obtaining recursive filtering when the uncertainty sequence is not necessarily i.i.d. Asymptotic stability of the MMSE filter was established in Tugnait (1981) when the loss sequence is i.i.d. with known loss probability; since in this case the estimation covariance is governed by a deterministic equation, one can obtain stability analysis by constructing an equivalent linear system without data losses.

In the more recent research on network models, (Fletcher, Rangan, & Goyal, 2004; Smith & Seiler, 2003) considered state estimation with lossy measurements resulting from time-varying channel conditions. In particular, Smith and Seiler (2003) developed a suboptimal jump linear estimator for complexity reduction in computing the corrector gain using finite loss history where the loss process is modelled by a two state Markov chain. The work (Fletcher et al., 2004) introduced a more general multiple state Markov chain to model the loss and nonloss channel states, and the asymptotic mean square estimation error for suboptimal linear estimators is analyzed and optimized by a linear matrix inequality (LMI) approach. Sinopoli et al. (2004) investigated the filtering stability with i.i.d. packet losses by identifying a threshold condition, and this approach was extended to the two sensor situation in Liu and Goldsmith (2004). In these results, the occurrence of packet losses is known at the estimator and this leads to a random Riccati equation involving the loss indicator sequence $\{\gamma_t, t \geqslant 1\}$, which differs from the deterministic recursion for the covariance function in Tugnait (1981). In networked systems, control problems with packet losses have been examined in Gupta, Spanos, Hassibi, and Murray (2005), Hadjicostis and Touri (2002) and Ling and Lemmon (2003).

1.2. Contributions and organization

In this paper we consider filtering with a Markovian packet loss model which captures the temporal correlation nature of practical channels. Within this channel modelling, on one hand, the averaging technique employed in Sinopoli et al. (2004) is no longer applicable for stability analysis. On the other hand, the filter exhibits an evident bi-modal structure along sample paths while switching between Kalman filtering and open-loop prediction. Traditional stability notions cannot adequately capture the sample path behavior of filtering. Motivated by this situation, we develop a new framework for filtering stability analysis, and our analytic techniques are closely related to the on–off pattern of the channel. In Section 2, for specifying the filter behavior, we introduce the notion of peak covariance. In an unstable scalar model, along a sample path the peal covariance (variance) gives the upper envelope of the actual covariance process. For the general vector model, a sufficient condition is given in Section 3 for peak covariance stability. In Section 4 we examine the stability property for the scalar model, and obtain a sufficient and necessary condition involving the recovery rate of the channel. Section 5 presents some simulation and computational examples, and Section 6 concludes the paper.

1.3. Terminology and notation

For characterizing filter stability and obtaining performance estimates, we need to first introduce some terminologies. For the reader's reference, these terminologies and some preliminary material are described below although they are easily found in textbooks. A stopping time τ (associated with the Markov chain $\gamma_t, t \geqslant 1$) is a measurable map from Ω to the set $\{1, 2, \ldots, \infty\}$ such that for any $k \geqslant 1$, $\{\tau \leqslant k\} \in \sigma(\gamma_t, t \leqslant k)$, which is the σ-algebra generating by $\gamma_{(\cdot)}$ up to time k. In our filtering context, the two sequences of stopping times introduced during the analysis simply describe the random switch time of the filter, or equivalently, the jump time of the Markov chain γ_t. For a real $n \times n$ matrix M, the induced norm is $\|M\| = \sup_{|X|=1} |MX|$ where $|X|$ and $|MX|$ denote the usual Euclidean norm for vectors. For a symmetric matrix M, we indicate it as positive semi-definite by $M \geqslant 0$, and the relation $M_1 \geqslant M_2$ for symmetric matrices means $M_1 - M_2 \geqslant 0$.

2. Evolution of the covariance

In order to simplify the analysis, in the following we assume the initial state for γ_t is $\gamma_1 = 1$. Note that this assumption imposes no essential restriction and the other case with $\gamma_1 = 0$ may be treated in the same manner. Based on Eq. (2), we write two separate equations

$$P_{t+1} = AP_tA' + Q - AP_tC'(CP_tC' + R)^{-1}CP_tA', \quad \gamma_t = 1 \tag{3}$$

$$P_{t+1} = AP_tA' + Q, \quad \gamma_t = 0 \tag{4}$$

depending on the value of γ_t. The covariance process P_t, as a random process, may be regarded as being governed by a bi-modal hybrid system where the evolution of the continuum component is driven by a two state Markov chain. Such a bi-modal structure is especially useful and will be exploited in the stability analysis.

To make the model nontrivial, throughout this paper we make the following assumptions:

(H1) The failure and recovery rate p, q are both in $(0, 1)$.
(H2) The system $[A, C]$ is observable, i.e., the rank of the matrix $[C', A'C', \ldots, (A^{n-1})'C']$ is n.

For the given initial condition $\gamma_1 = 1$, we introduce the following stopping time:

$$\tau_1 = \inf\{t : t > 1, \gamma_t = 0\}.$$

We make the usual convention that the infimum of an empty set is $+\infty$. Thus τ_1 is the first time when a packet loss occurs. Furthermore, we define

$$\beta_1 = \inf\{t : t > \tau_1, \gamma_t = 1\}.$$

It is clear β_1 is the first time the channel recovers from the first failure. The above procedure is repeated to define two sequences

$$\tau_1, \quad \tau_2, \quad \tau_3, \ldots,$$
$$\beta_1, \quad \beta_2, \quad \beta_3, \ldots,$$

which gives

$$\gamma_t = \begin{cases} 0 & \text{if } \tau_i \leqslant t < \beta_i < \infty, \quad i \geqslant 1, \\ 1 & \text{if } \beta_i \leqslant t < \tau_{i+1} < \infty, \quad i \geqslant 1. \end{cases} \tag{5}$$

Obviously the following order relationship holds:

$$1 < \tau_1 < \beta_1 < \cdots < \tau_k < \beta_k < \tau_{k+1} < \cdots, \tag{6}$$

whenever each of the entries is finite on the associated random sample point $\omega \in \Omega$.

Lemma 1. *Under condition* (H1), *with probability one* (*w.p.*1), *the two sequences* $\{\tau_i, i \geqslant 1\}$ *and* $\{\beta_i, i \geqslant 1\}$ *have finite values for each of their entries.*

Proof. The Markov chain $\{\gamma_t, t \geqslant 1\}$ is ergodic under (H1). It is easy to check that $\mathbb{P}(\tau_1 = \infty) = 0$. Hence, $\tau_1 < \infty$ w.p.1. Since $\gamma_{\tau_1+t}, t \geqslant 1$, is still a Markov process (i.e., having strong Markov property (Freedman, 1983)), we may calculate the condition probability $\mathbb{P}\{\beta_1=\infty|\tau_1<\infty\}=\mathbb{P}\{\gamma_{\tau_1+t}=0, t\geqslant 1|\tau_1<\infty\}=0$. Since $\tau_1<\infty$ w.p.1, we have $\beta_1<\infty$ w.p.1. By induction, we see that for each finite $i\geqslant 1$, $\tau_i<\beta_i<\infty$ w.p.1. Finally, after excluding a null sample set, all $\tau_i, \beta_i, i\geqslant 1$ have finite values along each sample path. □

Lemma 1 forms the basis for the peak covariance notion to be introduced later, which has a strictly increasing subscript after relabelling a subsequence of $\{P_t, t\geqslant 1\}$.

Define

$$\tau_i^* = \tau_i - \beta_{i-1}, \quad i\geqslant 1, \tag{7}$$

$$\beta_i^* = \beta_i - \tau_i, \quad i\geqslant 1, \tag{8}$$

where we adopt the convention $\beta_0=1$. Here τ_i^* and β_i^* denote the sojourn times (i.e., the length of a continuous stay) at the success state 1 and failure state 0, respectively.

Lemma 2. *Under* (H1), *we have*

(i) *The random variables* $\{\tau_i^*, i\geqslant 1\}$ *are i.i.d., and* τ_i^*-1 *is geometrically distributed with* $\mathbb{P}(\tau_i^*-1=k)=(1-p)^k p$, $k\geqslant 0$.
(ii) *The random variables* $\{\beta_i^*, i\geqslant 1\}$ *are i.i.d., and* β_i^*-1 *is geometrically distributed with* $\mathbb{P}(\beta_i^*-1=k)=(1-q)^k q$, $k\geqslant 0$.
(iii) *The random variables* $\{\tau_i^*, \beta_i^*, i\geqslant 1\}$ *are independent of each other.*

Proof. For the proof of (i) and (ii), see Freedman (1983). The property (iii) can be proven by direct computation for any given finite dimensional distribution of $(\tau_i^*, \beta_i^*, i\geqslant 1)$. Indeed, letting $k_i, \bar{k}_i\geqslant 1$, and $k=\sum_{j\leqslant i}k_j+\bar{k}_j$, we have

$$\begin{aligned}
&\mathbb{P}(\tau_1^*=k_1, \beta_1^*=\bar{k}_1, \ldots, \tau_i^*=k_i, \beta_i^*=\bar{k}_i, \tau_{i+1}^*=k_{i+1}) \\
&\quad = \mathbb{P}(\tau_1^*=k_1, \ldots, \tau_i^*=k_i, \beta_i^*=\bar{k}_i, \\
&\qquad \gamma_k=1, \ldots, \gamma_{k+k_{i+1}-1}=1, \gamma_{k+k_{i+1}}=0) \\
&\quad = \mathbb{P}(\tau_1^*=k_1, \ldots, \tau_i^*=k_i, \beta_i^*=\bar{k}_i, \gamma_k=1)(1-p)^{k_{i+1}-1}p \\
&\quad = \mathbb{P}(\tau_1^*=k_1, \ldots, \tau_i^*=k_i, \beta_i^*=\bar{k}_i)(1-p)^{k_{i+1}-1}p,
\end{aligned} \tag{9}$$

where we have used the Markovian property for γ_t to get

$$\begin{aligned}
&\mathbb{P}(\gamma_{k+1}=1, \ldots, \gamma_{k+k_{i+1}-1}=1, \gamma_{k+k_{i+1}}=0|\tau_1^*=k_1, \ldots, \tau_i^* \\
&\quad =k_i, \beta_i^*=\bar{k}_i, \gamma_k=1) \\
&\quad = \mathbb{P}(\gamma_{k+1}=1, \ldots, \gamma_{k+k_{i+1}-1}=1, \gamma_{k+k_{i+1}}=0|\gamma_k=1).
\end{aligned}$$

Repeating the above calculation with (9), it is easy to verify (iii). □

Now we define

$$\beta_k^- = \beta_k - 1. \tag{10}$$

In fact, β_k^- is the last time instant in a period of successive packet losses. In other words, β_k^- is the last time of visit of γ_t, to the failure state 0 since τ_k. The time β_k^- is useful for analyzing the filtering performance in that it provides a basis for estimating to what extent the covariance process may deteriorate resulting from successive packet losses. Immediately from time β_k, a new packet will arrive at the observer, and the state prediction will start to improve. The period $[\tau_i, \beta_i^-]$ and $[\beta_i, \tau_{i+1}-1]$ shall be called the loss cycle and normal cycle, respectively.

Labelling a subsequence of the covariance process P_k by the sequence of times β_k, we denote

$$M_k = P_{\beta_k}. \tag{11}$$

M_k denotes the value of the covariance $P_{\beta_k|\beta_k^-}$ computed by (4) at $t=\beta_k^-$. For an unstable scalar model, starting from τ_k+1, P_t monotonically increases to reach a maximum M_k at time β_k before turning downward; the sequence $\{M_k, k\geqslant 1\}$ gives the *upper envelope* of the covariance sequence. For this reason, we shall call M_k the peak covariance process. In the multi-dimensional (vector) case, P_t does not necessarily change monotonically before or after reaching M_k according to the packet arrival or loss; to facilitate our presentation, however, we shall still refer to M_k as the peak covariance process.

Definition 3. We say the sequence $\{M_k, k\geqslant 1\}$ is stable if $\sup_{k\geqslant 1} E\|M_k\|<\infty$. Accordingly, we say the (filtering) system satisfies peak covariance stability.

3. Sufficient condition for peak covariance stability

Let S^n denote the set of all $n\times n$ nonnegative definite real matrices. Based on Kalman filtering, define the map

$$F(P) = APA' + Q - APC'(CPC'+R)^{-1}CPA', \tag{12}$$

where $P\in S^n$. By completion of squares it is easy to show that for any $P\in S^n$, $F(P)\geqslant F(0_{n\times n})\geqslant Q$ and therefore $F(P)\in S^n$. To analyze the map F, we need to introduce the following definition.

Definition 4. For the observable linear system $[A, C]$, the observability index is the smallest integer I_o such that $[C', A'C', \ldots, (A^{I_o-1})'C']$ has rank n.

Under the observability assumption (H2), the integer I_o specified in Definition 4 obviously exists. For a deterministic system, I_o specifies the minimum number of observations which are required in order to reconstruct the initial condition of an observable system.

We define

$$S_0^n = \{P : 0\leqslant P\leqslant A\tilde{P}A' + Q, \ \text{for some } \tilde{P}\geqslant 0\}, \tag{13}$$

which is a convex subset of S^n.

Lemma 5. *Letting F be the map defined by* (12), *there exists a constant $K>0$ such that*

(i) *for any* $\bar{P}\in S_0^n$, $F^k(\bar{P})\leqslant KI$ *for all* $k\geqslant I_o$;
(ii) *for any* $\bar{P}\in S^n$, $F^{k+1}(\bar{P})\leqslant KI$ *for all* $k\geqslant I_o$,

where I is the $n\times n$ identity matrix.

Proof. In the proof we use the same set of notation as in Section 1 to the case without packet losses, i.e., $\gamma_t \equiv 1$.

(i) We begin by considering the case $\bar{P} = A\tilde{P}A' + Q$ with $\tilde{P} \geqslant 0$. By running a standard Kalman filter, we may interpret $F^k(\bar{P})$ as the optimal prediction error covariance matrix for x_{k+1} resulting from k measurements $(y_1, \ldots, y_k)$ with the initial covariance $P_1 = P_{1|0} = \bar{P}$ (after the covariance of x_0 is set as $\tilde{P}$). By use of $(y_1, y_2, \ldots, y_{I_o})$, we can construct a suboptimal estimator satisfying the bound condition, which further ensures that $F^k(\bar{P})$, $k \geqslant I_o$, is bounded by a fixed constant independent of $\bar{P} \in S_0^n$. In fact, we have

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{I_o} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{I_o-1} \end{pmatrix} x_1 + \mathbb{W}_{I_o-1} \triangleq Jx_1 + \mathbb{W}_{I_o-1},$$

where the random vector $\mathbb{W}_{I_o-1}$ depends only on the noise $(w_1, \ldots, w_{I_o-1}, v_1, \ldots, v_{I_o})$. Denote $\mathbb{Y}_{I_o} = (y_1, \ldots, y_{I_o})'$. For state x_1, we construct the estimate

$$\check{x}_1 = (J'J)^{-1}J'\mathbb{Y}_{I_o} \tag{14}$$

and the suboptimal estimator for x_{I_o+1} given $\mathbb{Y}_{I_o}$, as

$$\check{x}_{I_o+1} = A^{I_o}(J'J)^{-1}J'\mathbb{Y}_{I_o}. \tag{15}$$

It is easy to show that $E|x_{I_o+1} - \check{x}_{I_o+1}|^2 \leqslant L$ for a constant L independent of x_0 and x_1. Hence, by optimality of Kalman filtering, it follows that $P_{I_o+1} = F^{I_o}(P_1) \leqslant KI$ with a fixed constant K regardless of the value of $\bar{P} = P_1 = E(x_1 - x_{1|0})(x_1 - x_{1|0})'$.

Now we consider the general case $\bar{P} \leqslant A\tilde{P}A' + Q$ for some $\tilde{P} \geqslant 0$. Combing the monotone property (Anderson & Moore, 1979) of the map F (i.e., $F(P_1) \leqslant F(P_2)$ if $0 \leqslant P_1 \leqslant P_2$) with the result proved above, we see $F^{I_o}(\bar{P}) \leqslant KI$. Subsequently, for $k > I_o$, we have $F^k(\bar{P}) = F^{I_o}(F^{k-I_o}(\bar{P})) \leqslant KI$ since $F^{k-I_o}(\bar{P}) \in S_0^n$, and this completes the proof of (i).

(ii) This part follows from (i) and the fact $F^{k+1}(\bar{P}) = F^k(F(\bar{P}))$ where $0 \leqslant F(\bar{P}) \leqslant A\bar{P}A' + Q$. □

Remark. The observability condition may be relaxed to detectability, and one can identify an associated index I_o such that Lemma 5 holds. Then the subsequent analysis in this paper can be extended to the detectable model in a straightforward manner.

We introduce some positive constants. For $1 \leqslant i \leqslant (I_o - 1) \vee 1$, let $d_i^{(0)}$ and $d_i^{(1)}$ satisfy the following inequality:

$$\|F^i(P)\| \leqslant d_i^{(1)}\|P\| + d_i^{(0)}, \quad \forall P \in S_0^n, \tag{16}$$

where we use $\|\cdot\|$ to denote the induced norm for matrices. By the fact $F(P) \leqslant APA' + Q$, it is clear the pair $(d_i^{(0)}, d_i^{(1)})$ always exists. For the case $I_o = 1$, we may take $d_1^{(1)} = 0$ and $d_1^{(0)} > 0$ by Lemma 5.

Theorem 6. *The peak covariance process is stable if the following two conditions hold*:

(i) $|\lambda_A|^2(1-q) < 1$,
(ii) $pqd_1^{(1)}[1 + \sum_{i=1}^{I_o-1} d_i^{(1)}(1-p)^i]\sum_{j=1}^{\infty}\|A^j\|^2(1-q)^{j-1} < 1$,

where λ_A is an eigenvalue of the largest magnitude for matrix A.

The proof is given in Appendix. Here we give a brief discussion on condition (ii). Note that under condition (i), the infinite series in condition (ii) converges. Now let the pair A and q be fixed such that (i) holds. Then it is easy to check that for the given pair (A, q), if p is sufficiently small, condition (ii) is always satisfied.

Condition (i) may be regarded as specifying the minimum requirement for the recovery rate. Thus, based on the theorem, we may roughly state an intuitive fact: any given q fulfilling the minimum recovery rate requirement may be combined with a sufficiently small failure rate p such that the peak covariance stability holds.

Corollary 7. *If C is invertible, condition* (ii) *in Theorem* 6 *vanishes and the peak covariance stability holds under condition* (i).

Proof. When C is invertible, we have $I_o = 1$ and hence $\|F(P)\| \leqslant K$ for $P \in S_0^n$, by Lemma 5. This means $d_1^{(1)} = 0$ in (16), and therefore condition (ii) in Theorem 6 vanishes. □

4. Stability of the scalar model

For the scalar case, condition (ii) in Theorem 6 vanishes since in this case $d_1^{(1)} = 0$. The reason is that for the scalar Riccati equation, once there is an arrival of one packet at t, the covariance P_{t+1} becomes bounded by a fixed constant, no matter what value it has at the previous step. Furthermore, assuming $Q > 0$ to avoid triviality, we can show that condition (i) in Theorem 6 is also necessary. This leads to a sufficient and necessary condition in the following theorem. It is of interest to note that this condition only depends on the recovery rate of the Markov chain $\{\gamma_t, t \geqslant 1\}$.

For the scalar case, we set the coefficients A and C in the dynamics to their lower case form, i.e., we take $A = a$ and $C = c \neq 0$. The term covariance is also appropriately replaced by variance.

4.1. The sufficient and necessary condition for stability

Theorem 8. *For the scalar model with $Q > 0$, the peak variance process is stable if and only if*

$$a^2(1-q) < 1,$$

where q is the recovery rate.

Proof. Under the observability condition, condition (ii) in Theorem 6 vanishes for the scalar model, and sufficiency of condition (i) in Theorem 6 follows easily. Now we show that condition (i) in Theorem 6 is necessary. Below it suffices to establish necessity for the case $|a|>1$; it is evident that

$$P_{\tau_k} < P_{\tau_k+1} < P_{\tau_k+2} < \cdots < P_{\beta_k},$$

$$P_{\beta_k} > P_{\beta_k+1} > P_{\beta_k+2} > \cdots > P_{\tau_{k+1}}.$$

By use of the recursion with packet arrivals, it can be checked that

$$Q \leqslant P_{\tau_k} \leqslant Q + a^2 R c^{-2}.$$

This implies that for each loss cycle, the variance will evolve from an initial condition P_{τ_k} which is both lower and upper bounded by two strictly positive numbers.

On the other hand, we see that $\Delta_k = \beta_k^* - 1 = \beta_k - \tau_k - 1$ has a geometric distribution with parameter $1-q$. The magnitude of P_{β_k} depends on the difference $\beta_k - \tau_k$, where τ_k is the initial time. Using the lower bound for P_{τ_k}, it follows that

$$\begin{aligned} EP_{\beta_k} \geqslant\, & (a^2 Q + Q)\mathbb{P}(\Delta_k = 0) \\ & + (a^4 Q + a^2 Q + Q)\mathbb{P}(\Delta_k = 1) \\ & + \cdots + (a^{2i} Q + a^{2i-2} Q + \cdots + Q) \\ & \times \mathbb{P}(\Delta_k = i-1) + \cdots, \end{aligned} \tag{17}$$

where we have

$$\mathbb{P}(\Delta_k = l) = q(1-q)^l$$

for $l \geqslant 0$. We can verify that the series on the right-hand side of (17) converges if and only if $a^2(1-q)<1$, This completes the proof. □

Moreover, we have the following stability results in higher order moments.

Corollary 9. *For $r \geqslant 1$, we have $\sup_k E|P_{\beta_k}|^r < \infty$ if and only if $|a|^{2r}(1-q) < 1$.*

Proof. This corollary can be proven by following the same method as in the proof of Theorem 8. □

It is clearly seen from Corollary 9 that for obtaining higher order stability results, we need to put a more stringent condition on the recovery rate for an unstable system ($|a|>1$).

In the following we establish a stability result for the usual variance process P_t. To simplify the estimates, we only analyze the symmetric case with $p=q$, in which the distribution of the random variable $\tau_k - 2k + 1$ is the convolution of $2k-1$ i.i.d. geometric distributions, and this substantially simplifies the calculations. For the general case with $p \neq q$, the calculation is much more involved.

Theorem 10. *For the scalar model with $p=q$, if $a^2(1-q)<1$, then the variance process has the usual stability property, i.e., $\sup_{t\geqslant 1} EP_t < \infty$.*

Proof. For any given $t>1$, we write

$$\begin{aligned} P_t &= \sum_{i=1}^{\infty} P_t 1_{(\tau_i < t \leqslant \beta_i)} + \sum_{i=0}^{\infty} P_t 1_{(\beta_i < t \leqslant \tau_{i+1})} \\ &\leqslant \sum_{i=1}^{\infty} P_t 1_{(\tau_i < t \leqslant \beta_i)} + \sum_{i=0}^{\infty} (Q + a^2 R c^{-2}) 1_{(\beta_i < t \leqslant \tau_{i+1})} \\ &\leqslant \sum_{i=1}^{\infty} P_t 1_{(\tau_i < t \leqslant \beta_i)} + (Q + a^2 R c^{-2}). \end{aligned}$$

Let $\theta = a^2(1-q) < 1$. The case with $\theta = 0$ is trivial. In the following estimate we only consider the case $\theta > 0$. We may pick a large but fixed $D>0$ such that

$$\begin{aligned} \zeta_k &\triangleq EP_t 1_{(\tau_k < t \leqslant \beta_k)} \\ &\leqslant DE1_{(\tau_k < t \leqslant \beta_k)} \\ &\quad + DE[a^{2(t-\tau_k)}(1-q)^{t-\tau_k}] 1_{(\tau_k < t)}. \end{aligned}$$

In order to show $\sup_t EP_t < \infty$, now it suffices to show $\sup_t \sum_{k=1}^{\infty} E\theta^{t-\tau_k} 1_{(\tau_k \leqslant t)} < \infty$.

Recall that the distribution of $\tau_k - (2k-1)$ is the convolution of $(2k-1)$ i.i.d. geometric distributions with the same parameter $\delta = 1-q$, and by use of the probability generating function $(1-\delta)^{2k-1}/(1-\delta z)^{2k-1} = \sum_{i=0}^{\infty} \mathbb{P}\{\tau_k - (2k-1) = i\} z^i$, $|z| < \delta^{-1}$, of $\tau_k - (2k-1)$, it can be checked that

$$\begin{aligned} &\mathbb{P}(\tau_k - (2k-1) = i) \\ &\quad = (1-\delta)^{2k-1} \frac{(2k-1)\cdots(2k-1+i-1)}{i!} \delta^i, \end{aligned}$$

where $i \geqslant 0$.

By use of the distribution of τ_k, we have

$$\begin{aligned} &\sum_{k=1}^{\infty} E\theta^{t-\tau_k} 1_{(\tau_k \leqslant t)} \\ &\quad = \sum_{k\geqslant 1, 0\leqslant i \leqslant t-(2k-1)} \theta^{t-i-(2k-1)} \mathbb{P}(\tau_k - (2k-1) = i) \\ &\quad = \sum_{k\geqslant 1, 0\leqslant i \leqslant t-(2k-1)} \theta^{t-i-(2k-1)} (1-\delta)^{2k-1} \delta^i \\ &\qquad \times \frac{(2k-1)\cdots(2k-1+i-1)}{i!} \\ &\quad \leqslant \sum_{k\geqslant 1, 0\leqslant i \leqslant t-k} \theta^{t-i-k} (1-\delta)^k \delta^i \frac{k\cdots(k+i-1)}{i!} \\ &\quad = \theta^{-1}(1-\delta) \sum_{k\geqslant 0, 0\leqslant i\leqslant t-k} \theta^{t-i-k}(1-\delta)^k \delta^i \frac{(k+i)!}{k! \times i!} \\ &\quad = \theta^{-1}(1-\delta) \sum_{l=0}^{t} \sum_{k+i=l} \theta^{t-i-k}(1-\delta)^k \delta^i \frac{(k+i)!}{k! \times i!} \\ &\quad = \theta^{-1}(1-\delta) \sum_{l=0}^{t} \theta^{t-l} < \frac{1-\delta}{\theta(1-\theta)} = \frac{q}{\theta(1-\theta)}, \end{aligned}$$

which completes the proof. □

4.2. The relation between different stability notions

For showing the relationship between the peak covariance stability with other existing stability results in the literature, we specialize to the scalar model with i.i.d. packet losses. In this case, the transition matrix of the channel given by (1) reduces to $\begin{bmatrix} 1-q & q \\ 1-q & q \end{bmatrix}$, with an associated packet loss probability $p = 1 - q$. It is shown in Sinopoli et al. (2004, Theorem 2 & Section IV) that for the scalar model with i.i.d. packet losses, $\sup_{t \geqslant 1} E|P_t| < \infty$ (we term this as the usual stability of P_t) if and only if the packet arrival rate λ satisfies

$$\lambda > \lambda_c = 1 - 1/a^2. \tag{18}$$

By translating into our notation for channel parameters, λ is equal to q, so that (18) is equivalent to

$$q > 1 - 1/a^2. \tag{19}$$

Recalling Theorem 8, (19) is also a necessary and sufficient condition for the peak variance stability for the special case of i.i.d. packet losses. Then we may immediately claim the following relationship.

Corollary 11. *For the scalar model with i.i.d. packet losses, the peak variance stability is equivalent to the usual stability* (*i.e.*, $\sup_{t \geqslant 1} E|P_t| < \infty$).

For the scalar model with i.i.d. packet losses, it is of interest to note that the peak variance stability is seemingly stronger than the usual stability as the former characterizes a certain boundedness property along the upper envelope of the variance trajectories, but actually it is not, as stated in Corollary 11.

For the vector case when P_t is a matrix, the relation between the two stability notions as discussed above is much more complicated as the stability condition is not just reduced to the inequality (19). In addition, in the general vector case there is no obvious method for calculating the critical arrival rate λ_c, except one can show its existence.

5. Numerical examples

We first consider a scalar system with parameters $[A, C] \triangleq [a, c] = [1.4, 1]$, $Q = R = 1$ and $P_1 = P_{1|0} = 1$.

For this model, in order to guarantee stability, the minimum recovery rate is $q_c = 1 - 1/a^2 = 0.489796$. Fig. 1 shows a typical sample path with the parameter $q = 0.6 > q_c$, which ensures stability of the peak variance process. The horizontal axis in the figure is the discrete time. Along that sample path, we have $\tau_1 = 3$, $\beta_1 = 6$, $\tau_2 = 23$, $\beta_2 = 25$, etc. In Fig. 1-top, the curve displays the change of the variance along that sample path, and Fig. 1-bottom, shows the associated channel state jumping between 0 and 1. A high peak value for the variance is observed near $t = 60$, and this is due to the multiple successive packet losses.

Fig. 2 shows a sample path with $q = 0.32 < q_c$. Since in this case the recovery rate is low, the variance process has more chances to reach a high level.

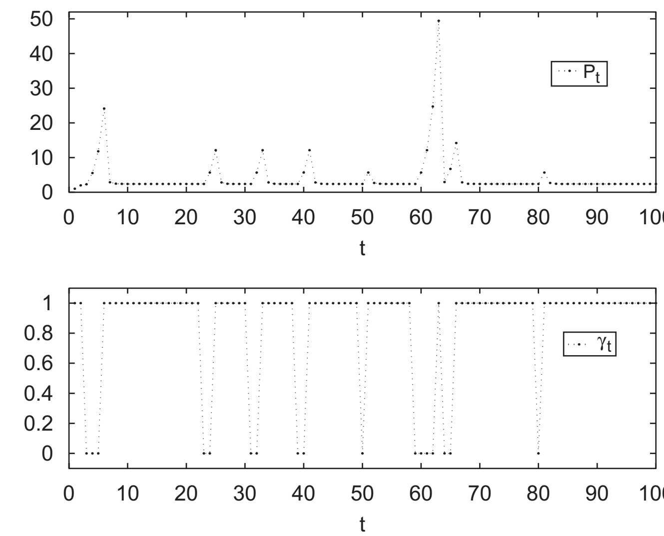

Fig. 1. The variance P_t and channel state γ_t, $q = 0.6$.

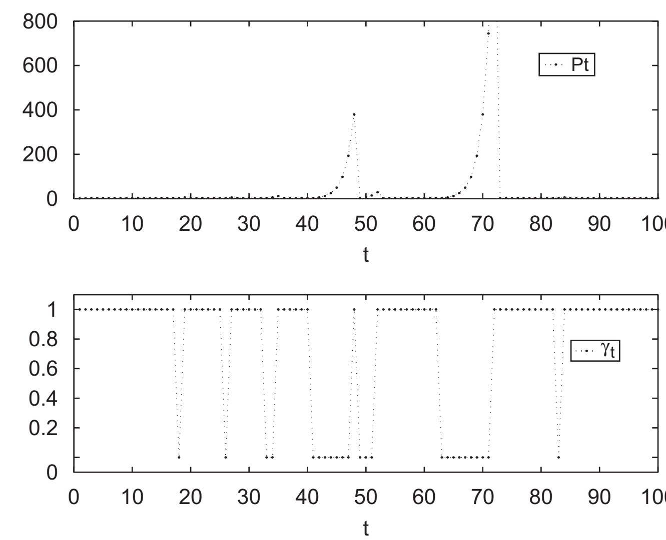

Fig. 2. The variance P_t and channel state γ_t, $q = 0.32$.

We continue to examine a vector example. Let the system be specified by

$$A = \begin{bmatrix} 1.3 & 0.3 \\ 0 & 1.2 \end{bmatrix}, \quad C = [1, 1].$$

The covariance of w_t is $Q = I \in \mathbb{R}^{2\times 2}$, and the variance of v_t is $R = 1$. We have $\|F(P)\| \leqslant \|AA'\| \cdot \|P\|$. It is easily checked that the observability index is $I_o = 2$ and we may take

$$d_1^{(1)} = 2.00813, \tag{20}$$

since AA' has two eigenvalues $\lambda_1 = 1.211879$ and $\lambda_2 = 2.008121$. By condition (i) in Theorem 6, the recovery rate must satisfy $q > 1 - |\lambda_A|^{-2} = 0.408285$. From now on we take $q = 0.65$.

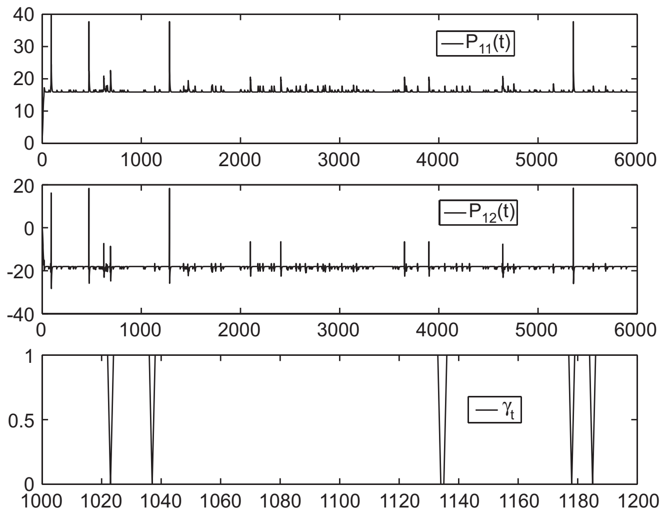

Fig. 3. $P_{11}(t)$, $P_{12}(t)$ and channel state γ_t, $q=0.65$ for the vector case.

By numerical calculation, we have $\sum_{j=1}^{\infty}\|A^j\|^2(1-q)^{j-1}\approx 6.433363$. Then if $p<0.04$, condition (ii) holds. Fig. 3 shows a sample path for this model with parameters $p=0.03$ and $q=0.65$; $P_{11}(t)$ and $P_{12}(t)$ are two entries in the 2×2 covariance matrix P_t, and the channel state is displayed between $t=1000$ and $t=1200$. For the associated channel with $(p,q)=(0.03,0.65)$, the stationary distribution of the failure state is $\mathbb{P}(\gamma_t=0)=0.044118$. Thus the long term packet loss rate is about 4.41%.

Unlike the scalar case, we only have a sufficient condition for stability of the filter, and condition (ii) in Theorem 6 specifying the region for (p,q) may be conservative. However, this criterion is still of usefulness since it covers some practical models with packet loss rate as high as several percents.

6. Conclusion and further discussion

In this paper we consider the problem of optimal linear filtering with packet losses modelled by a two-state Markov chain. The behavior of the error covariance process is examined under a certain stability notion and a sufficient condition is derived for ensuring stability. In the scalar case we obtain a sufficient and necessary condition for stability, and the relationship between different types of stability notions is also illustrated. For future work, it is of interest to develop filtering performance analysis by relating the filtering covariance process to the channel statistics.

Moreover, it is potentially useful to develop LMI analysis within our framework. For any suboptimal estimator with its associated covariance sequence $\{\tilde{P}_k, k\geqslant 1\}$, it is easy to show the relation $P_{\beta_k}\leqslant \tilde{P}_{\beta_k}$, for all $k\geqslant 1$. This property may provide an alternative approach for determining the region of (p,q) for which one may construct a suboptimal estimator such that the resulting $\{\tilde{P}_{\beta_k}, k\geqslant 1\}$ is stable, which in turn implies peak covariance stability. This will be addressed in depth in future research.

Acknowledgments

This work was partially supported by the Australian Research Council (ARC).

Appendix: Proof of Theorem 6

To analyze stability, we calculate the expectation of $\|P_{\beta_{k+1}+1}\|$ conditioned on $P_{\beta_k+1}=P\geqslant 0$. We have the relation:

$$
\begin{aligned}
&E[\|P_{\beta_{k+1}+1}\|\,|P_{\beta_k+1}=P]\\
&\quad=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}E[\|P_{\beta_{k+1}+1}\|\\
&\qquad\times 1_{\{\tau_{k+1}-\beta_k=i,\beta_{k+1}-\tau_{k+1}=j\}}|P_{\beta_k+1}=P]\\
&\quad=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\|F[A^jF^{i-1}(P)(A')^j\\
&\qquad+A^{j-1}Q(A')^{j-1}+\cdots+AQA'+Q]\|\\
&\qquad\times(1-p)^{i-1}p(1-q)^{j-1}q\\
&\quad=\Gamma(P).
\end{aligned}
\tag{21}
$$

Since $P_{\beta_k+1}=F(\beta_k)\in S_0^n$, we have the estimate:

$$
\begin{aligned}
\Gamma(P)&\leqslant\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}d_1^{(1)}\|A^jF^{i-1}(P)(A')^j+A^{j-1}Q(A')^{j-1}\\
&\qquad+\cdots+AQA'+Q\|\times(1-p)^{i-1}p(1-q)^{j-1}q+d_1^{(0)}\\
&\leqslant\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}d_1^{(1)}\|A^{j-1}Q(A')^{j-1}+\cdots+AQA'+Q\|\\
&\qquad\times(1-p)^{i-1}p(1-q)^{j-1}q+\sum_{j=1}^{\infty}\sum_{i=I_o+1}^{\infty}\\
&\qquad d_1^{(1)}\|A^jF^{i-1}(P)(A')^j\|\times(1-p)^{i-1}p(1-q)^{j-1}q\\
&\qquad+\sum_{j=1}^{\infty}\sum_{i=1}^{I_o}d_1^{(1)}\|A^jF^{i-1}(P)(A')^j\|\\
&\qquad\times(1-p)^{i-1}p(1-q)^{j-1}q\\
&\qquad+d_1^{(0)}\triangleq\Gamma_1+\Gamma_2+\Gamma_3+d_1^{(0)}.
\end{aligned}
\tag{22}
$$

We have

$$
\begin{aligned}
\Gamma_1&=\sum_{j=1}^{\infty}d_1^{(1)}\left\|\sum_{k=0}^{j-1}A^kQA'^k\right\|(1-q)^{j-1}q\\
&\leqslant\sum_{j=1}^{\infty}d_1^{(1)}\sum_{k=0}^{j-1}\|A^k\|^2\cdot\|Q\|(1-q)^{j-1}q\\
&=\sum_{k=0}^{\infty}\sum_{j=k+1}^{\infty}d_1^{(1)}\|A^k\|^2\cdot\|Q\|(1-q)^{j-1}q\\
&=d_1^{(1)}\|Q\|\sum_{k=0}^{\infty}\|A^k\|^2(1-q)^k<\infty,
\end{aligned}
\tag{23}
$$

where the series converges since $|\lambda_A|^2(1-q)<1$, and

$$\Gamma_2 \leqslant K d_1^{(1)} \sum_{j=1}^{\infty} \|A^j\|^2 (1-q)^{j-1} q \times \sum_{i=I_o+1}^{\infty} (1-p)^{i-1} p$$
$$= K d_1^{(1)} q (1-p)^{I_o} \sum_{j=1}^{\infty} \|A^j\|^2 (1-q)^{j-1} < \infty, \tag{24}$$

where the constant K is determined in Lemma 5. And finally we have

$$\Gamma_3 \leqslant \sum_{j=1}^{\infty} d_1^{(1)} \|A^j\|^2 p (1-q)^{j-1} q \times \left[\|P\| + \sum_{i=1}^{I_o-1} (d_i^{(1)}\|P\| + d_i^{(0)})(1-p)^i \right]$$
$$= \left\{ \left[1 + \sum_{i=1}^{I_o-1} d_i^{(1)} (1-p)^i \right] \|P\| + \sum_{i=1}^{I_o-1} d_i^{(0)} (1-p)^i \right\} \times pq d_1^{(1)} \sum_{j=1}^{\infty} \|A^j\|^2 (1-q)^{j-1}. \tag{25}$$

Combining the estimates in (22)–(25), we have

$$\Gamma(P) = E[\|P_{\beta_{k+1}+1}\| \; |P_{\beta_k+1} = P] \leqslant \delta \|P\| + C_0,$$

where the constant $C_0 > 0$ is independent of k and

$$\delta = pq d_1^{(1)} \left[1 + \sum_{i=1}^{I_o-1} d_i^{(1)} (1-p)^i \right] \sum_{j=1}^{\infty} \|A^j\|^2 (1-q)^{j-1} < 1$$

and this further implies

$$\Gamma(P_{\beta_k+1}) = E[\|P_{\beta_{k+1}+1}\| \; |P_{\beta_k+1}] \leqslant \delta \|P_{\beta_k+1}\| + C_0$$

which leads to

$$E\|P_{\beta_{k+1}+1}\| \leqslant \delta E \|P_{\beta_k+1}\| + C_0. \tag{26}$$

It immediately follows that $\limsup_k E\|P_{\beta_k+1}\| < \infty$.

By a similar technique as in (21), we estimate $E\|P_{\beta_{k+1}}\|$ starting with P_{β_k+1}. First, we have

$$E[\|P_{\beta_{k+1}}\| | P_{\beta_k+1}, \beta_k]$$
$$= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} E[\|P_{\beta_{k+1}}\| 1_{(\tau_{k+1}-\beta_k=i, \beta_{k+1}-\tau_{k+1}=j)} | P_{\beta_k+1}, \beta_k]$$
$$= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \|A^j F^{i-1}(P_{\beta_k+1})(A')^j + A^{j-1} Q (A')^{j-1} + \cdots + AQA' + Q\| (1-p)^{i-1} p (1-q)^{j-1} q$$
$$= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \|A^j F^{i-1}(P_{\beta_k+1})(A')^j\| \times (1-p)^{i-1} p (1-q)^{j-1} q + O(1) \tag{27}$$
$$= \sum_{i=1}^{\infty} \|F^{i-1}(P_{\beta_k+1})\| (1-p)^{i-1} p + O(1) \tag{28}$$
$$= \sum_{i=1}^{I_o} \|F^{i-1}(P_{\beta_k+1})\| (1-p)^{i-1} p + O(1) \tag{29}$$
$$\leqslant L_1 \|P_{\beta_k+1}\| + L_2, \tag{30}$$

for constant $L_1, L_2 > 0$. In the above we obtained (27)–(28) by condition (i), (29) by Lemma 5 since $P_{\beta_k+1} \in S_0^n$, and (30) by (16). Then it readily follows that $\sup_{k \geqslant 1} E\|P_{\beta_{k+1}}\| < \infty$ and we obtain the stability of the peak covariance process $\{P_{\beta_k}, k \geqslant 1\}$. □

References

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. Englewood Cliffs, NJ: Prentice-Hall.

Chong, C.-Y., & Kumar, S. P. (2003). Sensor networks: Evolution, opportunities, and challenges. *Proceedings of IEEE*, *91*, 1247–1256.

Elliott, E. O. (1963). Estimation of error rates for codes on burst-noise channels. *Bell System Technical Journal*, *42*, 1977–1997.

Fletcher, A. K., Rangan, S., & Goyal, V. K. (2004). Estimation from lossy sensor data: Jump linear modeling and Kalman filtering. In *Proceedings of the third ACM/IEEE international symposium on information processing in sensor networks* (pp. 251–258). Berkeley, CA.

Freedman, D. (1983). *Markov chains*. New York: Springer.

Gilbert, E. N. (1960). Capacity of burst-noise channels. *Bell System Technical Journal*, *39*, 1253–1265.

Gupta, V., Spanos, D., Hassibi, B., & Murray, R. M. (2005). Optimal LQG control across packet-dropping links. In *Proceedings of American control conference* (pp. 360–365). Portland, OR.

Hadidi, M. T., & Schwartz, S. C. (1979). Linear recursive state estimation under uncertain observations. *IEEE Transactions on Automatic Control*, *24*, 944–948.

Hadjicostis, C. N., & Touri, R. (2002). Feedback control utilizing packet dropping network links. In *Proceedings of IEEE conference on decision and control* (pp. 1205–1210). Las Vegas, NV.

Jaffer, A. G., & Gupta, S. C. (1971). Recursive Bayesian estimation with uncertain observation. *IEEE Transactions on Information Theory*, *17*, 614–616.

Kailath, T., Sayed, A. H., & Hassibi, B. (2000). *Linear estimation*. Upper Saddle River, NJ: Prentice-Hall.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME—Journal of Basic Engineering Series D*, *82*, 35–45.

Ling, Q., & Lemmon, M. (2003). *Soft real-time scheduling of networked control systems with dropouts governed by a Markov chain* (pp. 4845–4550). Denver, CO.

Liu, X., & Goldsmith, A. J. (2004). Kalman filtering with partial observation losses. In *Proceedings of IEEE conference on decision control* (pp. 1413–1418). Paradise Island, Bahamas.

Nahi, N. E. (1969). Optimal recursive estimation with uncertain observation. *IEEE Transactions on Information Theory*, *15*, 457–462.

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

Smith, S. C., & Seiler, P. (2003). Estimation with lossy measurements: Jump estimators for jump systems. *IEEE Transactions on Automatic Control*, *48*, 2163–2171.

Tugnait, J. K. (1981). Asymptotic stability of the MMSE linear filter for systems with uncertain observations. *IEEE Transactions on Information Theory*, *27*, 247–250.

Varaiya, P. (1993). Smart cars on smart roads: Problems of control. *IEEE Transactions on Automatic Control*, *38*, 195–207.

Zhang, H., Moura, J. M. F., & Krogh, B. (2005). Estimation in sensor networks: A graph approach. In *Proceedings of the fourth international symposium on information processing in sensor networks* (*IPSN*) (pp. 203–209). Los Angeles.

Zhao, F., Shin, J., & Reich, J. (2002). Information-driven dynamic sensor collaboration. *IEEE Signal Processing Magazine*, *19*, 61–72.

Minyi Huang received the B.Sc. degree from Shandong University in 1995, the M.Sc. degree from the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China, in 1998, and the Ph.D. degree from the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada, in July 2003, all in the area of systems and control.
From February 2004 to March 2006, he was a Research Fellow in the Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria, Australia. Since April 2006, he has been a Research Fellow with the Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra, ACT, Australia. His research interests include stochastic control and its applications, game theory, power control in wireless networks, and distributed signal processing in sensor networks.

Subhrakanti Dey was born in Calcutta, India in 1968. He obtained his Bachelor of Technology and Master's of Technology degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively. He obtained his Doctor of Philosophy degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering in the Australian National University (ANU), Canberra, Australia, in 1996.
During 1995–2000, he was a Postdoctoral Research Fellow in the Department of Systems Engineering, ANU, and a Postdoctoral Research Associate at the Institute for Systems Research, University of Maryland, College Park, USA. He is currently appointed as an Associate Professor with the Department of Electrical and Electronic Engineering in the University of Melbourne where he has been since 2000. His current research interests include signal processing for telecommunications, wireless communications and networks, performance analysis of communication networks, stochastic and adaptive estimation and control, and statistical and adaptive signal processing. He is currently an academic staff member of the ARC Special Research Center for Ultra-Broadband Information Networks in the Department of Electrical and Electronic Engineering, University of Melbourne. He has held numerous research grants, authored and co-authored over 60 internationally refereed papers and been on several major international conference program committees. Dr. Dey is an Associate Editor for the international journal *Systems and Control Letters* and *IEEE Transactions on Automatic Control*.