# Kalman Filtering with Markovian Packet Losses and Stability Criteria

**Conference Paper** *in* Proceedings of the IEEE Conference on Decision and Control · January 2007

DOI: 10.1109/CDC.2006.376710 · Source: IEEE Xplore



2 authors:

Minyi Huang
Carleton University


Subhrakanti Dey
University of Melbourne

Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006



# Kalman Filtering with Markovian Packet Losses and Stability Criteria

Minyi Huang and Subhrakanti Dey

***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. We analyze the behavior of the estimation error covariance matrix and introduce the notion of peak covariance, which describes the upper envelope of the sequence of error covariance matrices $\{P_t, t \geq 1\}$ 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; for the scalar case we obtain a sufficient and necessary condition, and derive upper and lower bounds for the tail distribution of the peak variance. For practically verifying the stability condition, we further introduce a suboptimal estimator and develop a numerical procedure to generate tighter estimate for the constants involved in the stability criterion.**

## I. 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 [12], [11] 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.

In this paper, we consider the optimal filtering of a linear system with random packet losses. We focus on the $n$ dimensional linear time-invariant system

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

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

$$y_t^0 = Cx_t + v_t \qquad t \geq 1,$$

where $C \in \mathbb{R}^{m\times n}$, and then $y_t^0$ is transmitted by a channel. Here $\{w_t, t \geq 1\}$ and $\{v_t, t \geq 1\}$ are two mutually independent sequences of 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 $(\Omega, \mathbb{F}, P)$ where $\mathbb{F}$ is the $\sigma$-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 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 $\gamma_t$ may be interpreted as resulting from the physical operating condition of a network and is assumed to be known at the filter. Specifically, the state 0 for $\gamma_t$ 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), $\gamma_t$ 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 transition if it starts with state 1. This is usually called the Gilbert-Elliott channel model [5], [2]. Obviously, a small value (close to 0) for $p$ and large value (close to 1) for $q$ mean the channel is more reliable.

Based on the history $\mathbb{F}_t = \sigma(y_i, \gamma_i, i \leq t)$, which is the $\sigma$-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 $\hat{x}_{t+1|t} = E[x_{t+1}|\mathbb{F}_t]$, $t \geq 0$, respectively, by the same method as in [16] which dealt with i.i.d. packet losses. 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

Work partially supported by Australian Research Council (ARC).

M. Huang is with Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra 0200, ACT, Australia. Email: minyi.huang@rsise.anu.edu.au

S. Dey is with Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria 3010, Australia. Email: s.dey@ee.unimelb.edu.au

1-4244-0171-2/06/$20.00 ©2006 IEEE.

error of $\hat{x}_{t+1|t}$ with an associated 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 also write $P_{t+1|t} = P_{t+1}$. To characterize the filtering covariance conditioned on the past history, one can easily derive the following random Riccati equation

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

where $M'$ denotes the transpose of a vector or matrix $M$. The initial condition in (2) is $P_1 = Var(x_1) = AP_{x_0}A' + Q$. Note that $\gamma_t$ appears as a random coefficient in the recursion.

Under a Bernoulli i.i.d. packet loss modelling, the filtering stability can be efficiently studied by a modified algebraic Riccati equation (MARE), which is obtained by replacing $\gamma_t$ in equation (2) by the arrival rate $\lambda$. Subsequently, the analysis amounts to identifying a critical value $\lambda_c$ (as a threshold) such that stability holds if and only if the arrival rate is greater than $\lambda_c$ (see Section IV for additional discussion) [16]. In contrast, when the channel model is given by the Markov chain $\gamma_t$, 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.

### A. Background and Related Work

Nowadays, filtering and estimation constitute an important aspect in sensor network deployment for monitoring, detection or tracking [1], [21], [20], as well as multi-vehicle coordination [19], 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 [15], [10], [7], [18], where a typical method for stability analysis is to construct a deterministic recursion utilizing the statistics of the uncertainty sequence indicating the availability of signals.

In the more recent research on network models, the work [17] and [3] considered state estimation with lossy measurements resulting from time-varying channel conditions. In particular, the authors in [17] 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 [3] introduced a more general multiple state Markov chain to model the loss and non-loss channel states, and the asymptotic mean square estimation error for suboptimal linear estimators is analyzed and optimized by a linear matrix inequality (LMI) approach. The MARE based analysis in [16] for i.i.d. packet losses was extended to the two sensor situation in [14]. 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. Control problems with packet losses have been examined in [8], [13], [6].

### B. Contributions and Organization

In this paper we consider a Markovian packet loss model which captures the temporal correlation nature of practical channels, and we develop new analytic techniques for filtering stability analysis. In Section II, we introduce the notion of peak covariance. The general sufficient condition in Section III was initially obtained in our earlier work [9]. In Section IV we examine the stability property for the scalar model, and present a tail distribution analysis for the peak variance. To find practically more verifiable sufficient conditions than in [9], in Section V we introduce an appropriately parameterized suboptimal estimator, and optimize the parameter in the suboptimal estimator to produce a tighter selection of the constants in the stability criteria. Section VI presents some simulation and computational examples.

## II. Evolution of the Covariance

In order to simplify the analysis, in the following we assume the initial state for $\gamma_t$ 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 equation (2), we write two separate equations

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

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

depending on the value of $\gamma_t$. 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 $[C', A'C', \cdots, (A^{n-1})'C']$ is $n$. □

For the reader's convenience, we introduce the basic definition of stopping times although it is easily found in textbooks (see, e.g., [4]). A stopping time $\tau$ (associated with the Markov chain $\gamma_t$, $t \geq 1$) is a measurable map from $\Omega$ to the set $\{1, 2, \cdots, \infty\}$ such that $\{\tau \leq k\}$ depends on $\gamma$ 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 $\gamma_t$.

Given the initial condition $\gamma_1 = 1$, we introduce the following stopping time: $\tau_1 = \inf\{t > 1, \ \gamma_t = 0\}$. We make the usual convention that the infimum of an empty set is $+\infty$. Thus $\tau_1$ 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 $\beta_1$ is the first time the channel recovers from the first failure. The above procedure is repeated to define two sequences

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

which gives the value of $\gamma_t$ at the switch times:

$$\gamma_t = \begin{cases} 0, & \text{if } t = \tau_i < \infty, \\ 1, & \text{if } t = \beta_i < \infty. \end{cases} \tag{5}$$

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 sample point $\omega \in \Omega$.

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

Lemma 1 forms the basis for the peak covariance notion to be introduced later.

Define

$$\tau_i^* = \tau_i - \beta_{i-1}, \qquad i \geq 1$$
$$\beta_i^* = \beta_i - \tau_i, \qquad i \geq 1$$

where we adopt the convention $\beta_0 = 1$. Here $\tau_i^*$ and $\beta_i^*$ 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 $\{\tau_i^*, i \geq 1\}$ are i.i.d., and $\tau_i^* - 1$ is geometrically distributed with $P(\tau_i^* - 1 = k) = (1-p)^k p$, $k \geq 0$.

(ii) the random variables $\{\beta_i^*, i \geq 1\}$ are i.i.d., and $\beta_i^* - 1$ is geometrically distributed with $P(\beta_i^* - 1 = k) = (1-q)^k q$, $k \geq 0$.

(iii) The two sequences of random variables $\{\tau_i^*, i \geq 1\}$ and $\{\beta_i^*, i \geq 1\}$ are independent of each other. □

Now we define

$$\beta_k^- = \beta_k - 1. \tag{7}$$

In fact, $\beta_k^-$ is the last time of visit of $\gamma_t$, to the failure state 0 since $\tau_k$. The time $\beta_k^-$ 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 $\beta_k$, a new packet will arrive at the observer, and the state prediction for the next step 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 $\beta_k$, we denote

$$M_k = P_{\beta_k}. \tag{8}$$

$M_k$ denotes the prediction error covariance $P_{\beta_k|\beta_k - 1}$ computed by (4) at $t = \beta_k^-$. For an unstable scalar model, starting from $\tau_k + 1$, $P_t$ monotonically increases to reach a maximum $M_k = P_{\beta_k}$ at time $\beta_k$ before turning downward; the sequence $\{M_k, k \geq 1\}$ 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 $\{M_k, k \geq 1\}$ is stable if $\sup_{k\geq 1} E\|M_k\| < \infty$. Accordingly, we say the (filtering) system satisfies peak covariance stability. □

## III. 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

$$F(P) = APA' + Q - APC'(CPC' + R)^{-1}CPA', \tag{9}$$

where $P \in S^n$. It is easy to show $F(P) \in S^n$. To analyze the map $F$, we introduce the following definition.

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

Define $S_0^n = \{P : 0 \leq P \leq A\tilde{P}A' + Q, \text{ for some } \tilde{P} \geq 0\}$, which is a convex subset of $S^n$.

*Lemma 5:* Letting $F$ be defined by (9), there exists a constant $K > 0$ such that (i) for any $\bar{P} \in S_0^n$, $F^k(\bar{P}) \leq KI$ for all $k \geq I_0$; (ii) for any $\bar{P} \in S^n$, $F^{k+1}(\bar{P}) \leq KI$ for all $k \geq I_0$, where $I$ is the $n \times n$ identity matrix. □

The strategy to prove the lemma is to run an auxiliary Kalman filer; see [9] for details.

*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 a few constants. For $1 \leq i \leq (I_o - 1) \vee 1$, let $C_i^{(0)}$ and $C_i^{(1)}$ satisfy the following inequality

$$\|F^i(P)\| \leq C_i^{(1)}\|P\| + C_i^{(0)}, \qquad \forall\ P \in S_0^n, \tag{10}$$

where $\|\cdot\|$ denotes the induced norm for matrices. By the fact $F(P) \leq APA' + Q$, it is clear the above pair $(C_i^{(0)}, C_i^{(1)})$ always exists. For the case $I_0 = 1$, we may take $C_1^{(1)} = 0$.

*Theorem 6:* [9] The peak covariance process is stable if the following two conditions hold:

(i) $|\lambda_A|^2(1-q) < 1$,

(ii) $pqC_1^{(1)}\left[1 + \sum_{i=1}^{I_o - 1} C_i^{(1)}(1-p)^i\right] \sum_{j=1}^{\infty} \|A^j\|^2(1-q)^{j-1} < 1$,

where $\lambda_A$ is an eigenvalue of $A$ with the largest absolute value. □

We give a brief discussion on condition (ii). Notice 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.

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

## IV. Stability of the Scalar Model

For the scalar case, condition (ii) in Theorem 6 vanishes since in this case $C_1^{(1)} = 0$. The reason is that for the scalar Riccati equation, once there is an arrival of one packet at $t$, $P_{t+1}$ becomes bounded by a fixed constant, regardless of the value of $P_t$. Furthermore we can show that condition (i) in Theorem 6 is also necessary. This leads to a sufficient and necessary condition. Note that this condition only depends on the recovery rate of the Markov chain $\{\gamma_t, t \geq 1\}$.

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

### A. The Sufficient and Necessary Condition for Stability

*Theorem 8:* Letting $r \geq 1$, we have $\sup_k E|P_{\beta_k}|^r < \infty$ if and only if $|a|^{2r}(1-q) < 1$. □

It is clearly seen from Theorem 8 that, with a given $|a| > 1$, for obtaining higher order stability results, we need to put a more stringent condition on the recovery rate $q$. By taking $r = 1$ in Theorem 8, we conclude that in the scalar model a sufficient and necessary condition for peak covariance stability is $a^2(1-q) < 1$.

In the following we establish the stability on the standard 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 9:* For the scalar model with $p = q$, if $a^2(1 - q) < 1$, then the variance process has the usual stability property, i.e., $\sup_{t\geq 1} EP_t < \infty$. □

### B. The Relation between Different Stability Notions

For illuminating the relationship between our 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 [16] (Theorem 2 and Sec. IV) that for the scalar model with i.i.d. packet losses, $\sup_{t\geq 1} E|P_t| < \infty$ (we term this as the usual stability of $P_t$) if and only if the packet arrival rate $\lambda > \lambda_c = 1 - 1/a^2$, or equivalently,

$$q > 1 - 1/a^2. \tag{11}$$

Recalling Theorem 8, (11) is also a necessary and sufficient condition for the peak variance stability for the special case of i.i.d. packet losses. Then we can immediately claim the following relationship.

*Corollary 10:* For the scalar model with i.i.d. packet losses, the peak variance stability is equivalent to the usual stability (i.e., $\sup_{t\geq 1} E|P_t| < \infty$). □

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 10.

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 (11).

### C. Tail Distribution of the Peak Variance

Now we examine the tail distribution of the peak variance when it is stable, i.e., $a^2(1-q) < 1$. We restrict to the case $|a| > 1$ and $Q > 0$. For $M > 0$, define the tail distribution of $P_{\beta_k}$, $k \geq 1$, as $P_{tail}(M) = P\{P_{\beta_k} \geq M\}$.

It is easy to show that $Q \leq P_{\tau_k} \leq a^2R/c^2 + Q \triangleq \bar{P}$. Below we restrict to $M \geq \bar{P} \vee (a^2Q)$. It can be verified that $\{P_{\beta_k} \geq M\} \supset \{\beta_k^* - 1 \geq \frac{\ln(M/Q)}{\ln a^2} - 1\}$. Denote by $\lceil x \rceil$ the smallest integer no less than $x$. Hence

$$P\{P_{\beta_k} \geq M\} \geq (1-q)^{\lceil \frac{\ln(M/Q)}{\ln a^2} - 1 \rceil} \geq (1-q)^{\frac{\ln(M/Q)}{\ln a^2}}.$$

Denote $\kappa = \frac{\ln(1-q)^{-1}}{\ln a^2} > 1$, where $|a| > 1$. We have $P_{tail}(M) \geq Q^\kappa M^{-\kappa}$.

In a similar manner, we have

$$\begin{aligned}\{P_{\beta_k} \geq M\} &= \{F^{\beta_k^*}(P_{\tau_k}) \geq M\} \\ &\subset \{\beta_k^* - 1 \geq \frac{\ln(M/\bar{P})}{\ln a^2} + \frac{\ln(a^2-1)}{\ln a^2} - 2\}.\end{aligned}$$

Hence $P\{P_{\beta_k} \geq M\} \leq (1-q)^{\frac{\ln(M/\bar{P})}{\ln a^2} + \frac{\ln(a^2-1)}{\ln a^2} - 2}$, which gives $P_{tail}(M) \leq \zeta_0 \bar{P}^\kappa M^{-\kappa}$, where $\zeta_0 = (1-q)^{\frac{\ln(a^2-1)}{\ln a^2} - 2}$. For $M \geq \bar{P} \vee (a^2Q)$, it follows that

$$Q^\kappa M^{-\kappa} \leq P_{tail}(M) \leq \zeta_0 \bar{P}^\kappa M^{-\kappa}$$

which gives the lower and upper bound estimates. It is seen that the decaying rate of the tail distribution depends on the ratio $\kappa$ which in turn is related to the stability margin associated with the condition $a^2(1-q) < 1$ and $|a| > 1$.

## V. Stability Check by Linear Suboptimal Estimators

Theorem 6 gives a criterion for checking the stability of the peak covariance process. In particular, condition (ii) depends on some constants related to the operator $F$ defined by (10). In general, it is difficult to explicitly compute these constants, and a rough selection may lead to very conservative conditions for the pair $(p, q)$. In this section, we combine the analytic technique in Section III with a numerical procedure to give a practically useful selection of these constants. The basic idea is as follows. First we construct a suboptimal estimator and introduce its prediction error covariance $P_k^s$, where $s$ indicates it is yielded by the suboptimal estimator. Next, it is easy to establish a dominance relationship between $P_k$ and $P_k^s$, i.e., $P_k \le P_k^s$, Subsequently, if the system dynamics together with the channel statistics ensure stability of $P_{\beta_k}^s$, then the stability of $P_{\beta_k}$ naturally follows. Below we restrict to linear suboptimal estimators where the filtering corrector gain $L$ for the case of a packet arrival is a constant matrix to be selected.

We construct the following suboptimal estimator:

$$\hat{x}_{k|k}^s = \hat{x}_{k|k-1}^s + 1_{\{\gamma_k=1\}} L(y_k - C\hat{x}_{k|k-1}^s)$$
$$\hat{x}_{k|k-1}^s = A\hat{x}_{k-1|k-1}^s.$$

The prediction error covariance $P_{t+1}^s \triangleq P_{t+1|t}^s$, $t \ge 0$, is described by

$$P_{t+1}^s = \{A(I-LC)P_t^s(I-LC)'A' + ALRL'A'\}1_{\{\gamma_t=1\}} + AP_t^sA'1_{\{\gamma_t=0\}} + Q.$$

For the case $\gamma_t = 1$, we get

$$P_{t+1}^s = A(I-LC)P_t^s(I-LC)'A' + ALRL'A' + Q. \quad (12)$$

Based on (12) we introduce the operator

$$F_L(P) = A(I-LC)P(I-LC)'A' + ALRL'A' + Q.$$

For the Kalman filter and the suboptimal estimator, we have the following dominance relationship .

*Theorem 11:* Let the common initial condition $P_{x_0} \ge 0$ be given for the Kalman filter and the suboptimal estimator. With probability one, we have

$$P_k^s \ge P_k, \quad \forall k \ge 1. \quad (13)$$

*Proof.* By using Kalman filtering without packet losses, we can establish the relation $F(P) \le F_L(P)$ for all $P \ge 0$. We assume $P = Var(x_1)$; then for any matrix $H$, $H\hat{x}_{2|1}$ is the minimum covariance estimate of $Hx_2$. Thus,

$$H(x_2 - \hat{x}_{2|1})(x_2 - \hat{x}_{2|1})'H' \le H(x_2 - \hat{x}_{2|1}^s)(x_2 - \hat{x}_{2|1}^s)'H'$$

which means $HF(P)H' = HP_{2|1}H' \le HP_{2|1}^sH' = HF_L(P)H'$, and therefore $F(P) \le F_L(P)$. Next, the Kalman filtering covariance sequence $P_k(P_1)$ as a function of $P_1$ is monotone, i.e., $P_k(P_1) \le P_k(\bar{P}_1)$ if $0 \le P_1 \le \bar{P}_1$. It is easy to check that $\Psi(P)$, standing for $F_L(P)$ or $APA'+Q$, also has the same monotone property. Finally, by applying these monotone properties along the sequence $\gamma_t$, we see that (13) holds with probability one. □

*Corollary 12:* If $\{P_{\beta_k}^s, k \ge 1\}$ is stable, then the peak covariance of $P_k$ is also stable. □

For the operator $F_L$, in analogy to Section III, we also define the constants as follows. For $1 \le i \le (I_o - 1) \vee 1$, let $C_i^{L,(0)}$ and $C_i^{L,(1)}$ satisfy the following inequality

$$\|F_L^i(P)\| \le C_i^{L,(1)}\|P\| + C_i^{L,(0)}, \qquad \forall\ P \in S_0^n. \quad (14)$$

It is obvious the above pair $(C_i^{L,(0)}, C_i^{L,(1)})$ always exists. The following stability result can be proved by the same method as in proving Theorem 6.

*Corollary 13:* The covariance process $\{P_{\beta_k}^s, k \ge 1\}$ is stable if the following two conditions hold:

(i) $|\lambda_A|^2(1-q) < 1$,
(ii) $pqC_1^{L,(1)}\left[1 + \sum_{i=1}^{I_o-1} C_i^{L,(1)}(1-p)^i\right]\sum_{j=1}^{\infty}\|A^j\|^2(1-q)^{j-1} < 1$,

where $\lambda_A$ is an eigenvalue of $A$ with the largest absolute value. □

For a practical application to the stability Theorem 6 for the Kalman filter, we may choose a suitable $L$ to reduce the magnitude of $C_i^{L,(1)}$. Then by the fact $F(P) \le F_L(P)$, we may set $C_i^{(1)} = C_i^{L,(1)}$, $1 \le i \le I_0 - 1$.

## VI. Numerical Examples

### A. Simulations for the Peak Covariance

We first consider a scalar system with parameters $[A, C] \triangleq [a, c] = [1.4, 1]$, $Q = R = 1$ and $P_0 = 1$.

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.

We continue to examine a vector example specified by

$$A = \begin{bmatrix} 1.3 & 0.3 \\ 0 & 1.2 \end{bmatrix}, \qquad C = [1, 1]. \quad (15)$$

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 $\|F(P)\| \le \|AA'\| \cdot \|P\|$. It is easily checked that the observability index $I_o = 2$ and we may take

$$C_1^{(1)} = 2.00813, \quad (16)$$

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 $q = 0.65$. 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 $2\times 2$ 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 $P(\gamma_t = 0) = 0.044118$. Thus the long term packet loss rate is about $4.41\%$.

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

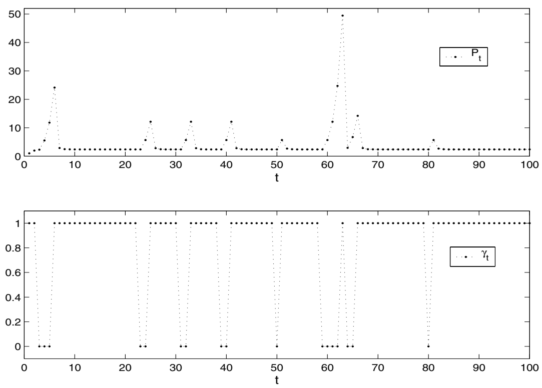


Fig. 1. The variance $P_t$ and channel state $\gamma_t$, $q = 0.6$.

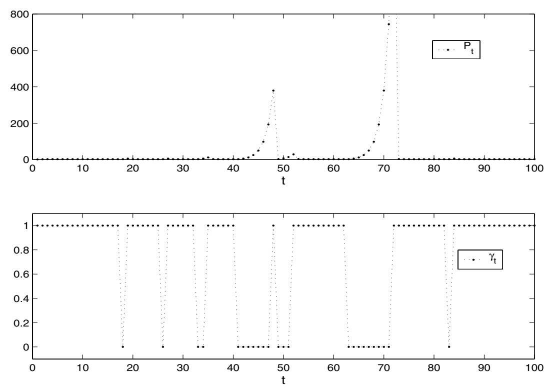


Fig. 2. The variance $P_t$ and channel state $\gamma_t$, $q = 0.32$.

## *B. Suboptimal Estimator Aided Stability Check*

For the two dimensional example given by (15), with the aid of the suboptimal estimator, we may take $C_1^{(1)} = \|(A - ALC)(A - ALC)'\|$ by recalling $F(P) \leq F_L(P)$ and the structure of $F_L(P)$. Now the task is to find $L$ such that $\|A_L\| \stackrel{\triangle}{=} \|(A - ALC)(A - ALC)'\|$ is minimized. Here we use a numerical method to search for an ideal value for $L = [l_1, l_2]^T$. We compute the value $\|A_L\|$ for $(l_1, l_2)$ on a grid $[-0.6, 2] \times [-0.7, 2]$ with a step size 0.05 on both edges, which gives a total of $53 \times 55$ points. On the grid, the minimum of 1.22 for $\|A_L\|$ is attained by $(l_1, l_2) = (0.5, 0.5)$. Now we may take $C_1^{(1)} = 1.22$ with a significant improvement from the rough estimate in (16).

On the other hand, the steady-state solution to the Riccati equation $P = F(P)$ is $P = \begin{pmatrix} 15.8733 & -18.0630 \\ -18.0630 & 24.9392 \end{pmatrix}$. Accordingly, the steady-state corrector gain for the state prediction equation in a standard Kalman filter is $L_0 = PC'(CPC+R)^{-1} = (-0.3851, 1.2092)^T$. It can be checked that $\|(A-AL_0C)(A-AL_0C)'\| = 4.4106$ $(> 1.22)$, which only gives a very poor bound for $C_1^{(1)}$. This suggests a careful search of $L$ is important.

## REFERENCES


[1] C.-Y. Chong and S. P. Kumar. Sensor networks: evolution, opportunities, and challenges, *Proc. IEEE*, vol. 91, pp. 1247-1256, Aug. 2003.

[2] E. O. Elliott. Estimation of error rates for codes on burst-noise channels. *Bell Syst. Tech. J.*, vol. 42, pp. 1977-1997, Sept. 1963.

[3] A. K. Fletcher, S. Rangan, and V. K. Goyal. Estimation from lossy sensor data: jump linear modeling and Kalman filtering. *Proc. the 3rd ACM/IEEE International Symposium on Information Processing in Sensor Networks*, Berkeley, CA, pp. 251-258, Apr. 2004.

[4] D. Freedman. *Markov Chains*, Springer-Verlag, New York, 1983.

[5] E. N. Gilbert. Capacity of burst-noise channels. *Bell Syst. Tech. J.*, vol. 39, pp. 1253-1265, Sept., 1960.

[6] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray. Optimal LQG control across packet-dropping links. *Proc. American Control Conference*, Portland, OR, pp. 360-365, June 2005.

[7] M. T. Hadidi and S. C. Schwartz. Linear recursive state estimation under uncertain observations. *IEEE Trans. Automat. Control*, vol. 24, no. 6, pp. 944-948, Dec. 1979.

[8] C. N. Hadjicostis and R. Touri. Feedback control utilizing packet dropping network links. *Proc. IEEE Conf. Decision and Control*, Las Vegas, NV, pp. 1205-1210, Dec. 2002.

[9] M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica* (accepted), 2006.

[10] A. G. Jaffer and S. C. Gupta. Recursive Bayesian estimation with uncertain observation. *IEEE Trans. Inform. Theory*, vol. 17, pp. 614-616, Sept. 1971.

[11] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*, Prentice-Hall, Upper Saddle River, NJ, 2000.

[12] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME – J. Basic Eng., Ser. D.*, vol. 82, pp. 35-45, Mar. 1960.

[13] Q. Ling and M. Lemmon. Soft real-time scheduling of networked control systems with dropouts governed by a Markov chain. *Proc. American Control Conf.*, Denver, CO, pp. 4845-4550, June, 2003.

[14] X. Liu and A. J. Goldsmith. Kalman filtering with partial observation losses. *Proc. IEEE Conf. Decision Contr.*, Bahamas, pp. 1413-1418, Dec. 2004.

[15] N. E. Nahi. Optimal recursive estimation with uncertain observation. *IEEE Trans. Inform. Theory*, vol. 15, pp. 457-462, July 1969.

[16] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. Kalman filtering with intermittent observations, *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1453-1464, Sept. 2004.

[17] S. C. Smith and P. Seiler. Estimation with lossy measurements: jump estimators for jump systems. *IEEE Trans. Automat. Contr.*, vol. 48, pp. 2163-2171, Dec. 2003.

[18] J. K. Tugnait. Asymptotic stability of the MMSE linear filter for systems with uncertain observations. *IEEE Trans. Inform. Theory*, vol. 27, no. 2, pp. 247-250, Mar 1981.

[19] P. Varaiya. Smart cars on smart roads: Problems of control. *IEEE Trans. Automat. Control*, vol. 38, pp. 195-207, Feb. 1993.

[20] H. Zhang, J. M. F. Moura and B. Krogh. Estimation in sensor networks: A graph approach. *Proc. 4th Internat. Symp. Inform. Process. in Sensor Networks (IPSN)*, Los Angeles, pp. 203-209, April, 2005

[21] F. Zhao, J. Shin and J. Reich. Information-driven dynamic sensor collaboration, *IEEE Sig. Process. Mag.*, vol. 19, pp. 61-72, Mar. 2002.


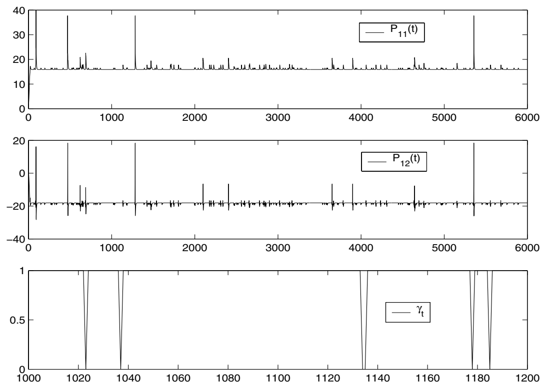


Fig. 3. $P_{11}(t), P_{12}(t)$ and channel state $\gamma_t$, $q = 0.65$ for the vector case.