Kalman Filtering Over a Packet-Dropping Network: A Probabilistic Perspective

Ling Shi, Michael Epstein, and Richard M. Murray

***Abstract*—We consider the problem of state estimation of a discrete time process over a packet-dropping network. Previous work on Kalman filtering with intermittent observations is concerned with the asymptotic behavior of $\mathbb{E}[P_k]$, i.e., the expected value of the error covariance, for a given packet arrival rate. We consider a different performance metric, $\Pr[P_k \le M]$, i.e., the probability that P_k is bounded by a given M. We consider two scenarios in the paper. In the first scenario, when the sensor sends its measurement data to the remote estimator via a packet-dropping network, we derive lower and upper bounds on $\Pr[P_k \le M]$. In the second scenario, when the sensor preprocesses the measurement data and sends its local state estimate to the estimator, we show that the previously derived lower and upper bounds are equal to each other, hence we are able to provide a closed form expression for $\Pr[P_k \le M]$. We also recover the results in the literature when using $\Pr[P_k \le M]$ as a metric for scalar systems. Examples are provided to illustrate the theory developed in the paper.**

***Index Terms*—Kalman filtering, packet-dropping network, random process, state estimation.**

I. Introduction

In the past decade, networked control systems have gained attention from both the control community and the network and communication community. When compared with classical feedback control systems, networked control systems have several advantages. For example, they can reduce the system wiring, make the system easy to operate, maintain and diagnose, and increase system agility. Although networked control systems have advantages, inserting a network in between the plant and the controller can introduce many problems as well. For example, in communication networks, data packets that carry the information can be dropped, delayed or even reordered due to the network traffic conditions. When closing the control loop over such communication networks, the overall system might have poor performance or even become unstable when the aforementioned issues exist. Thus the effect that those issues have on the closed loop system performance must be fully analyzed before networked control systems become commonplace.

Manuscript received August 04, 2008; revised February 05, 2009. First published January 26, 2010; current version published March 10, 2010. Recommended by Associate Editor K. H. Johansson.

L. Shi is with the Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China (e-mail: eesling@ust.hk).

M. Epstein and R. M. Murray are with Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91106 USA (e-mail: epstein@ds.caltech.edu, murray@cds.caltech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2009.2039236

Recently, many researchers have investigated these issues and some significant results were obtained and many are in progress. The problem of state estimation and stabilization of a linear time invariant (LTI) system over a digital communication channel which has a finite bandwidth capacity was introduced by Wong and Brockett [1], [2] and further pursued by others (e.g., [3]–[6]). Elia [7] considered the problem of stabilizing a networked control system over fading channels. Sinopoli *et al.* [8] discussed how packet loss can affect state estimation. They showed there exists a certain threshold of the packet arrival rate below which, $\mathbb{E}[P_k]$, the expected value of the error covariance matrix, becomes unbounded as time goes to infinity. They also provided lower and upper bounds of the threshold value. The authors extended their result from estimation to closed loop control in [9] where stability regions of packet arrival rates are provided. Following the spirit of [8], Liu and Goldsmith [10] extended the idea to the case where there are multiple sensors and the packets arriving from different sensors are dropped independently. They provided similar bounds on the packet arrival rate for a stable estimate, again in the expected sense. Jin *et al.* [11] considered the problem of state estimation over packet-dropping networks using a multi-description (MD) coding scheme. They showed that by using the MD codes the stability region the Kalman filter is increased and the performance is improved. Gupta *et al.* [12] studied the problem of LQG control across packet-dropping networks and showed that it is optimal to let the sensor preprocess the measurement data and sends its local state estimate to the remote estimator over a packet-dropping network. The implicit assumption of their work is that the sensor has unlimited computation capability. In [13], actuation buffers and a receding horizon control strategy is proposed for the LQG control over packet-dropping networks. Huang and Dey [14] considered Kalman filtering over a packet-dropping network where data packet drops are described by a two-state Markov chain. The readers are referred to [15] and references therein for some recent results in the area of networked control systems.

The problem of state estimation of a dynamical system where measurements are sent across a packet-dropping network is also the focus of this work. Despite the great progress of the previous researchers, the problems they have studied have certain limitations. For example, in both [8] and [10], the authors assumed that packets are dropped independently, which is certainly not true in the case where bursts of packets are dropped or in queuing networks where adjacent packets are not dropped independently. They also use $\mathbb{E}[P_k]$ as the measure of performance, which can conceal the fact that events with arbitrarily low probability can cause the expected value diverge, and it might be better to ignore such events that occur with

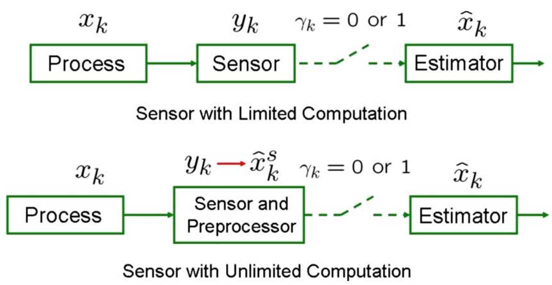


Fig. 1. System block diagram.

arbitrarily low probability. We will provide such an example in Section III-C after necessary definitions are given.

The goal of the present work is to give a different characterization of the estimator performance by considering a probabilistic description of the error covariance, i.e., $\mathbf{Pr}[P_k \leq M]$. For scalar systems, this is equivalent to considering the cumulative distribution of the random variable P_k.

In [16] the present authors first introduced this notion for the same problem setting but under the additional assumption that the measurement matrix, C, is invertible. In [17], the present authors extended the result to the case when C is not invertible. However, extra assumptions are made, e.g., in order to obtain the upper bound of $\mathbf{Pr}[P_k \leq M]$, A is assumed to be purely unstable, i.e., $|\lambda_i(A)| \geq 1$ for all i, where $\lambda_i(A)$ is the i-th eigenvalue of discrete-time system matrix A.

The main contributions of this paper are summarized as follows:

1) Unlike previous work where the *a priori* error covariance is studied, we consider the *a posteriori* error covariance in this paper.
2) We remove the constraint in [17] that requires A to be unstable and work with arbitrary A.
3) We are able to recover the result in [8] for scalar systems, i.e., from the result using $\mathbf{Pr}[P_k \leq M]$ as a metric, we derive the stability result using $\mathbb{E}[P_k]$ as a metric.
4) We study the case when the sensor can preprocess the information and sends its own state estimate to the remote estimator. In this case, we show that the previously derived lower and upper bounds on $\mathbf{Pr}[P_k \leq M]$ are the same and hence we are able to give an exact expression for $\mathbf{Pr}[P_k \leq M]$.

The rest of the paper is organized as follows. In Section II, the mathematical model of the system that we consider is given. In Section III, some frequently used terms are defined, a quick summary of Kalman filter updating equations is provided and some results on $\mathbb{E}[P_k]$ from [8] is reviewed. In Section IV we consider the case when the sensor directly sends its measurement packet to the estimator and we derive lower and upper bounds for $\mathbf{Pr}[P_k \leq M]$. In Section V we consider the case when the sensor preprocesses the measurement and sends its own state estimate to the remote estimator. In Section VI we provide two examples to demonstrate the theory developed. The paper concludes with a summary of our results and a discussion of the work that lies ahead.

II. Problem Setup

We consider the networked control system shown in Fig. 1.

The process dynamics and sensor measurement equation are given as follows:

$$x_k = Ax_{k-1} + w_{k-1}, \tag{1}$$
$$y_k = Cx_k + v_k. \tag{2}$$

In the above equations, $x_k \in \mathbb{R}^n$ is the state vector, $y_k \in \mathbb{R}^m$ is the observation vector, $w_{k-1} \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are zero mean, white, Gaussian random vectors with $\mathbb{E}[w_k w_j'] = \delta_{kj} Q \geq 0$, $\mathbb{E}[v_k v_j'] = \delta_{kj} R > 0$, $\mathbb{E}[w_k v_j'] = 0 \; \forall j, k$, where $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ otherwise. We assume that the pair (A, C) is observable and $(A, \sqrt{Q})$ is controllable.

Depending on its computational capability, the sensor can either send y_k or preprocess y_k and send $\hat{x}_k^s$ to the remote estimator, where $\hat{x}_k^s$ is defined at the sensor as

$$\hat{x}_k^s \triangleq \mathbb{E}[x_k | y_1, \cdots, y_k].$$

The two cases correspond to the two scenarios in Fig. 1, i.e., sensor with limited or unlimited computation.

We assume that the data packets from the sensor (either y_k or $\hat{x}_k^s$) are to be sent across a packet-dropping network, with negligible quantization effects, to the estimator. Thus the estimator will either receive a perfectly communicated data packet or none at all. Let γ_k be the random variable indicating whether a packet is dropped at time k or not, i.e., $\gamma_k = 0$ if a packet is dropped and $\gamma_k = 1$ otherwise.

In addition, we assume the sensor has the ability to store some previous measurements in a buffer when needed. Therefore each packet sent through the network could contain a finite number of the previous measurements. In packet based networks the transmitted packet usually contains a fixed amount space for data, therefore if less than this amount is needed to be transmitted, the packet is padded to meet the required length [18]. We assume all the data from the buffered measurements can fit into a single packet and therefore the additional measurements do not increase the bandwidth required nor the packet-dropping rate (we require this when C is not full rank).

We define the following state quantities at the remote state estimator:

$$\hat{x}_k \triangleq \mathbb{E}[x_k | \text{all data packets up to } k],$$
$$P_k \triangleq \mathbb{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)' | \text{all data packets up to } k\right].$$

As mentioned in Section I, we are interested in finding a closed form solution to $\mathbf{Pr}[P_k \leq M]$ given M. In the next few sections, we consider the two scenarios in Fig. 1 and provide results on $\mathbf{Pr}[P_k \leq M]$ for each of them.

III. Preliminaries

A. *Definitions*

The following terms that are frequently used in subsequent sections are defined in this section. It is assumed that (A, C, Q, R) are the same as they appear in Section II. $\mathbb{S}_+^n$ is the set of n by n positive semidefinite matrices. When $X \in \mathbb{S}_+^n$, we simply write $X \geq 0$; when X is positive definite, we write $X > 0$. We define the function $h : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$h(X) \triangleq AXA' + Q. \tag{3}$$

As we shall see shortly, applying h to the previous error covariance matrix corresponds to the time update of the standard Kalman filter. Similarly, we define the function $g : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$g(X) \triangleq AXA' + Q - AXC'[CXC' + R]^{-1}CXA' \tag{4}$$

and the function $\tilde{g} : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$\tilde{g}(X) \triangleq X - XC'[CXC' + R]^{-1}CX. \tag{5}$$

Then g and $\tilde{g}$ correspond to the measurement update for the a priori and a posteriori error covariance matrices respectively in the standard Kalman filter. It is easy to see that

$$g = h \circ \tilde{g}. \tag{6}$$

We denote $\rho(A)$ as spectral radius of A, i.e., $\rho(A) = \max_i |\lambda_i(A)|$. We say A is stable if $\rho(A) < 1$, and A is unstable if A is not stable. For functions $f, f_1, f_2 : \mathbb{S}_+^n \to \mathbb{S}_+^n$, $f_1 \circ f_2$ is defined as

$$f_1 \circ f_2(X) \triangleq f_1(f_2(X)) \tag{7}$$

and f^t is defined as

$$f^t(X) \triangleq \underbrace{f \circ f \circ \cdots \circ f}_{t \text{ times}}(X). \tag{8}$$

For a random variable X, we write its expectation value as $\mathbb{E}[X]$ and its conditional probability given another random variable Y as $\mathbf{Pr}[X|Y]$.

B. Kalman Filtering Preliminaries

If the network between the sensor and the estimator is perfect, i.e., no packet is dropped, then it is well known that the optimal linear estimator for the system described by (1) and (2) is a standard Kalman filter, denoted as $\mathbf{KF}$. We write $(\hat{x}_k, P_k)$ in compact form as

$$(\hat{x}_k, P_k) = \mathbf{KF}(\hat{x}_{k-1}, P_{k-1}, y_k)$$

which represents the follow set of equations:

$$\begin{cases} \hat{x}_k^- = A\hat{x}_{k-1}, \\ P_k^- = AP_{k-1}A' + Q, \\ K_k = P_k^- C' \left[CP_k^- C' + R_k\right]^{-1}, \\ \hat{x}_k = A\hat{x}_{k-1} + K_k(y_k - CA\hat{x}_{k-1}), \\ P_k = (I - K_k C)P_k^-. \end{cases}$$

With some manipulation, P_k^- and P_k can be shown to satisfy

$$P_k^- = g\left(P_{k-1}^-\right), \quad P_k = \tilde{g} \circ h(P_{k-1}).$$

Let P^* be the unique positive semi-definite solution[1] to $g(X) = X$, i.e., $P^* = g(P^*)$. Define $\overline{P}$ as $\overline{P} \triangleq \tilde{g}(P^*)$. Then we have

$$\tilde{g} \circ h(\overline{P}) = \tilde{g} \circ h \circ \tilde{g}(P^*) = \tilde{g} \circ g(P^*) = \tilde{g}(P^*) = \overline{P}.$$

In other words

$$P^* = \lim_{k\to\infty} P_k^-, \quad \overline{P} = \lim_{k\to\infty} P_k.$$

C. Kalman Filtering With Intermittent Observations

Consider the case when the sensor sends the measurement data to the estimator without processing it. Sinopoli *et al.* [8] showed that the Kalman filter is still the optimal linear estimator in this setting. There is a slight change to the standard Kalman filter in that only the time update is performed when a measurement packet is dropped. When a measurement is received, both the time and measurement update steps are performed. The filtering equations are thus the same as $\mathbf{KF}$ except that

$$\hat{x}_k = \hat{x}_k^- + \gamma_k K_k \left(y_k - C\hat{x}_k^-\right) \tag{9}$$
$$P_k = P_k^- - \gamma_k K_k C P_k^-. \tag{10}$$

Unlike the standard Kalman filtering setting where P_k is a deterministic quantity (given an initial value P_0), the randomness of the data packet drops makes it a random variable as well.

When γ_k's are independent and identically distributed Bernoulli random variables with mean γ, it was shown in [8] that there exists a critical value γ_c such that if $\gamma > \gamma_c$, $\mathbb{E}[P_k]$ converges as $k \to \infty$ and diverges otherwise. When C^{-1} exists, γ_c is given in exact form as

$$\gamma_c = 1 - \frac{1}{\rho(A)^2}. \tag{11}$$

Using $\mathbb{E}[P_k]$ as a metric, however, may conceal the fact that events with arbitrarily small probability can make the expected value diverge, and it might be better to ignore such events when evaluating the performance of the estimator. For example, consider the unstable scalar system with $a = 2$, $q = 1$, $P_0 = 1$ in (1). Let the packet arrival rate γ be

$$\gamma = 0.74 < \gamma_c = 1 - \frac{1}{a^2} = 0.75.$$

Then from [8] we conclude that

$$\lim_{k\to\infty} \mathbb{E}[P_k] = \infty.$$

This is easily verifiable by considering the event σ that no packets are received in all k time steps. Then

$$\mathbb{E}[P_k] \geq \mathbb{E}[P_k|\sigma]\mathbf{Pr}[\sigma] > (0.26^k)4^k P_0 = 1.04^k P_0 = 1.04^k.$$

By letting k go to infinity, we see that $\mathbb{E}[P_k]$ diverges. Thus σ alone can make $\mathbb{E}[P_k]$ diverge, and the probability that σ occurs approaches zero when k goes to infinity. This partially motivates us to consider $\mathbf{Pr}[P_k \leq M]$ as a metric to evaluate the performance of the estimator subject to packet drops.

IV. Sensors With Limited Computation

A. Lower and Upper Bounds of $\mathbf{Pr}[P_k \leq M]$

Similar to [8], the optimal state estimate $\hat{x}_k$ and its error covariance matrix P_k are given by

$$(\hat{x}_k, P_k) = \begin{cases} (A\hat{x}_{k-1}, h(P_{k-1})), & \text{if } \gamma_k = 0, \\ \mathbf{KF}(\hat{x}_{k-1}, P_{k-1}, y_k), & \text{if } \gamma_k = 1. \end{cases}$$

[1]Since (A, C) is assumed to be observable and $(A, \sqrt{Q})$ controllable, from standard Kalman filtering analysis, P^* exists.

As a result

$$P_k = \begin{cases} h(P_{k-1}), & \text{if } \gamma_k = 0,\\ \tilde{g}\circ h(P_{k-1}), & \text{if } \gamma_k = 1.\end{cases}$$

Assume C is full rank, and without loss of generality, assume C^{-1} exists. We show in Remark 4.8 that the main result developed in Theorem 4.6 extends naturally to the general case.

Define $\overline{M} \triangleq C^{-1}RC^{-1'}$. Then we have the following result that shows the relationship between P_k and $\overline{M}$.

Lemma 4.1: For any $k \geq 1$, if $\gamma_k = 1$, then $P_k \leq \overline{M}$.

Proof: As $\gamma_k = 1$, we have $P_k = \tilde{g}\circ h(P_{k-1}) \leq \overline{M}$, where the inequality is due to Lemma A.2 in Appendix A and the fact that $h(P_{k-1}) \geq 0$. ■

Remark 4.2: We can also interpret Lemma 4.1 as follows. One way to obtain an estimate $\tilde{x}_k$ when $\gamma_k = 1$ is simply by inverting the measurement, i.e., $\tilde{x}_k = C^{-1}y_k$. Therefore

$$\tilde{e}_k = C^{-1}v_k \text{ and } \tilde{P}_k = \mathbb{E}\left[\tilde{e}_k\tilde{e}_k'\right] = C^{-1}RC^{-1'} = \overline{M}.$$

Since Kalman filter is optimal among the set of all linear filters, we must have $P_k \leq \tilde{P}_k = \overline{M}$.

For $M \geq \overline{M}$, define $k_1(M)$ and $k_2(M)$ as follows:

$$k_1(M) \triangleq \min\left\{t \geq 1 : h^t(\overline{M}) \nleq M\right\}, \tag{12}$$

$$k_2(M) \triangleq \min\left\{t \geq 1 : h^t(\overline{P}) \nleq M\right\}. \tag{13}$$

Notice that $h^t(P_k)$ corresponds to the error covariance evolution when there are t consecutive packet drops from time k. Therefore if the current error covariance is $\overline{M}$ (or $\overline{P}$), then k_1 (or k_2) will be the minimum number of consecutive packet drops such that the error covariance will grow and exceed the given M. We sometimes write $k_i(M)$ as k_i, $i = 1, 2$ for simplicity for the rest of the paper. The following lemma shows the relationship between $\overline{P}$ and $\overline{M}$ as well as k_1 and k_2.

Lemma 4.3: (1) $\overline{P} \leq \overline{M}$; (2) $k_1 \leq k_2$ whenever either k_i is finite, $i = 1, 2$.

Proof: (1) $\overline{P} = \tilde{g}(P^*) \leq \overline{M}$ where the inequality is from Lemma A.2 in Appendix A. (2) Without loss of generality, we assume k_2 is finite. If k_1 is finite, and $k_1 > k_2$, then according to their definitions, we must have

$$M \geq h^{k_1-1}(\overline{M}) \geq h^{k_1-1}(\overline{P}) \geq h^{k_2}(\overline{P})$$

which violates the definition of k_2. Notice that we use the property that h is nondecreasing as well as $h(\overline{P}) \geq \overline{P}$ from Lemma A.1 and A.2 in Section A in the Appendix. Similarly we can show that k_1 cannot be infinite. Therefore we must have $k_1 \leq k_2$. ■

Lemma 4.4: Assume $P_0 \geq \overline{P}$. Then for all $k \geq 0$, $P_k \geq \overline{P}$.

Proof: We prove this by induction. Assume $P_k \geq \overline{P}$ for some $k \geq 0$. This clearly holds when $k = 0$. Let us consider P_{k+1}.

If $\gamma_{k+1} = 1$, then

$$P_{k+1} = \tilde{g}\circ h(P_k) \geq \tilde{g}\circ h(\overline{P}) = \overline{P}$$

where the inequality is due to Lemma A.1 in Appendix A.

If $\gamma_{k+1} = 0$, then

$$P_{k+1} = h(P_k) \geq h(\overline{P}) \geq \overline{P}.$$

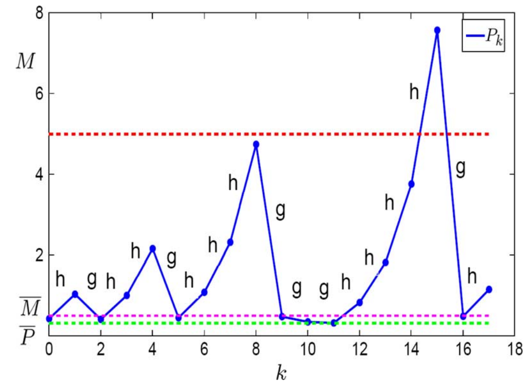


Fig. 2. Packet arrival sequence.

The induction step is thus complete. ■

Define N_k as the number of consecutive packet drops at time k, i.e.,

$$N_k \triangleq \min\{t \geq 0 : \gamma_{k-t} = 1\}. \tag{14}$$

Notice that N_k is also a random variable which depends on the underlying packet arrival sequence. Let us consider a scalar example to illustrate k_1, k_2 and N_k.

Example 4.5: Consider (1) and (2) with $A = 1.4$, $C = 1$, $Q = 0.2$, $R = 0.5$. It is easy to verify that $\overline{P} = 0.3083$, $\overline{M} = 0.5$. For $M = 5$, it is calculated that $k_1(M) = 3$, $k_2(M) = 4$. We plot P_k for one possible packet arrival sequence in Fig. 2 with h in the figure indicating that the data packet is lost at that time and g indicating that the data packet arrives at the estimator. Notice that whenever the estimator receives a packet, P_k is seen to be between $\overline{P}$ and $\overline{M}$, no matter how large P_{k-1} is. For this particular example, we have $N_1 = 1$, $N_2 = 0$, $N_3 = 1$, $N_4 = 2$, etc.

With the definitions of k_1, k_2 and N_k, we have the following theorem that provides lower and upper bounds on $\mathbf{Pr}[P_k \leq M]$.

Theorem 4.6: Assume $\overline{P} \leq P_0 \leq \overline{M}$. For any $M \geq \overline{M}$, we have

$$1 - \mathbf{Pr}[N_k \geq k_1] \leq \mathbf{Pr}[P_k \leq M] \leq 1 - \mathbf{Pr}[N_k \geq k_2]. \tag{15}$$

Proof: We divide the proof into two parts.

1) Let us first prove

$$1 - \mathbf{Pr}[N_k \geq k_1] \leq \mathbf{Pr}[P_k \leq M].$$

As $\gamma_k = 1$ or 0, there are in total 2^k possible realizations of γ_1 to γ_k as seen from Fig. 3.

Let Σ_1 denote those packet arrival sequences of γ_1 to γ_k such that $N_k \geq k_1$. Similarly let Σ_2 denote those packet arrival sequences such that $N_k < k_1$. Let $P_k(\sigma_i)$ be the error covariance at time k when the underlying packet arrival sequence is σ_i, where $\sigma_i \in \Sigma_i$, $i = 1, 2$. Consider a particular $\sigma_2 \in \Sigma_2$. As $\gamma_{k-k_1+1} = 1$, from Lemma 4.1, $P_{k-k_1+1} \leq \overline{M}$. Therefore we have

$$P_k(\sigma_2) \leq h^{k_1-1}\left(P_{k-k_1+1}\right) \leq h^{k_1-1}(\overline{M}) \leq M$$

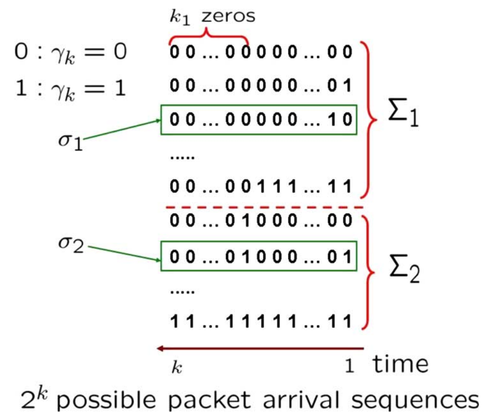


Fig. 3. $N_k \geq k_1$.

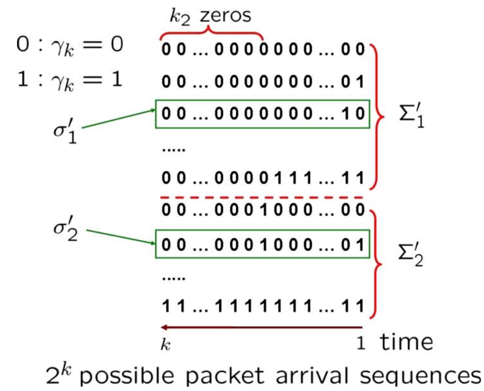


Fig. 4. $N_k \geq k_2$.

where the first and second inequalities are from Lemma A.1 in Appendix A and the last inequality is from the definition of k_1. In other words

$$\mathbf{Pr}[P_k \leq M|\sigma_2] = 1.$$

Therefore we have

$$\begin{aligned}\mathbf{Pr}[P_k \leq M] &= \sum_{\sigma\in\Sigma_1\cup\Sigma_2} \mathbf{Pr}[P_k \leq M|\sigma]\mathbf{Pr}(\sigma)\\ &= \sum_{\sigma_1\in\Sigma_1} \mathbf{Pr}[P_k \leq M|\sigma_1]\mathbf{Pr}(\sigma_1)\\ &\quad + \sum_{\sigma_2\in\Sigma_2} \mathbf{Pr}[P_k \leq M|\sigma_2]\mathbf{Pr}(\sigma_2)\\ &\geq \sum_{\sigma_2\in\Sigma_2} \mathbf{Pr}[P_k \leq M|\sigma_2]\mathbf{Pr}(\sigma_2)\\ &= \sum_{\sigma_2\in\Sigma_2} \mathbf{Pr}(\sigma_2) = \mathbf{Pr}(\Sigma_2)\\ &= 1 - \mathbf{Pr}(\Sigma_1) = 1 - \mathbf{Pr}[N_k \geq k_1]\end{aligned}$$

where the first equality is from the total probability theorem, the second equality holds as Σ_1 and Σ_2 are disjoint, the third inequality holds as the first sum is non-negative, the rest equalities are easy to see.

2) We now prove

$$\mathbf{Pr}[P_k \leq M] \leq 1 - \mathbf{Pr}[N_k \geq k_2].$$

Let Σ_1' denote those packet arrival sequences of γ_1 to γ_k such that $N_k \geq k_2$, and Σ_2' denote those packet arrival sequences such that $N_k < k_2$ (Fig. 4). Consider $\sigma_1' \in \Sigma_1'$. Let

$$s(\sigma_1') = \min\{t \geq 1 : \gamma_{k-t} = 1|\sigma_1'\}.$$

As $\sigma_1' \in \Sigma_1'$, we must have $s \geq k_2$. Consequently

$$P_k(\sigma_1') = h^{s(\sigma_1')}\left(P_{k-s(\sigma_1')}\right) \geq h^{s(\sigma_1')}(\overline{P})$$

where the inequality is from Lemma 4.4. Therefore we conclude $P_k(\sigma_1') \not\leq M$. Otherwise $h^{s(\sigma_1')}(\overline{P}) \leq P_k(\sigma_1') \leq M$, which violates the definition of k_2. In other words

$$\mathbf{Pr}[P_k \leq M|\sigma_1'] = 0.$$

Therefore we have

$$\begin{aligned}\mathbf{Pr}[P_k \leq M] &= \sum_{\sigma\in\Sigma_1'\cup\Sigma_2'} \mathbf{Pr}[P_k \leq M|\sigma]\mathbf{Pr}(\sigma)\\ &= \sum_{\sigma_1'\in\Sigma_1'} \mathbf{Pr}[P_k \leq M|\sigma_1']\mathbf{Pr}(\sigma_1')\\ &\quad + \sum_{\sigma_2'\in\Sigma_2'} \mathbf{Pr}[P_k \leq M|\sigma_2']\mathbf{Pr}(\sigma_2')\\ &= \sum_{\sigma_2'\in\Sigma_2'} \mathbf{Pr}[P_k \leq M|\sigma_2']\mathbf{Pr}(\sigma_2')\\ &\leq \sum_{\sigma_2'\in\Sigma_2'} \mathbf{Pr}(\sigma_2') = \mathbf{Pr}(\Sigma_2')\\ &= 1 - \mathbf{Pr}(\Sigma_1') = 1 - \mathbf{Pr}[N_k \geq k_2]\end{aligned}$$

where the inequality is from the fact that $\mathbf{Pr}[P_k \leq M|\sigma_2'] \leq 1$ for any $\sigma_2' \in \Sigma_2'$. ■

Remark 4.7: We assume in the theorem that $P_0 \geq \overline{P}$. This is without loss of generality as $\overline{P}$ is the steady-state error covariance of the Kalman filter. Furthermore, for any $P_0 \not\leq M$, as long as $\gamma_t = 1$, from Lemma 4.1, we have $P_t \leq \overline{M}$. Therefore the theorem applies to any $k \geq t$. Notice that $\mathbf{Pr}[t = \infty] = 0$, i.e., $\mathbf{Pr}[t < \infty] = 1$.

Remark 4.8: We point out in this remark that the result in Theorem 4.6 extends naturally to the case when C is not full rank. Since (A, C) is observable, there exists r $(2 \leq r \leq n)$ such that

$$\begin{bmatrix} C \\ CA \\ \cdots \\ CA^{r-1} \end{bmatrix}$$

is full rank. In this section, we consider the special case when $r = 2$, and in particular, we assume $\left[\begin{smallmatrix} C \\ CA \end{smallmatrix}\right]^{-1}$ exists. The idea readily extends cases where $r > 2$. Unlike the case when C^{-1} exists, and y_k is sent across the network, here we assume that the previous measurement y_{k-1} is sent along with y_k. This only requires that the sensor has a buffer that stores y_{k-1}. Then if $\gamma_k = 1$, both y_k and y_{k-1} are received. Thus we can use the following linear estimator to generate the state estimate:

$$\tilde{x}_k = A\begin{bmatrix} CA \\ C \end{bmatrix}^{-1}\begin{bmatrix} y_k \\ y_{k-1} \end{bmatrix}.$$

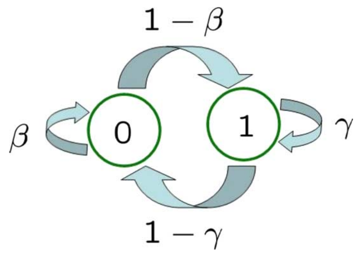


Fig. 5. Gilbert–Elliott model.

The corresponding error covariance can be calculated as

$$\tilde{P}_k = AM_1A' + Q - A\begin{bmatrix} CA \\ C \end{bmatrix}^{-1}\begin{bmatrix} CQ \\ 0 \end{bmatrix} - \begin{bmatrix} CQ \\ 0 \end{bmatrix}'\begin{bmatrix} CA \\ C \end{bmatrix}^{-1'}A'$$

where

$$M_1 = \begin{bmatrix} CA \\ C \end{bmatrix}^{-1}\begin{bmatrix} CQC' + R & 0 \\ 0 & R \end{bmatrix}\begin{bmatrix} CA \\ C \end{bmatrix}^{-1'}.$$

Since Kalman filter is optimal among the set of all linear estimators, we conclude that if $\gamma_k = 1$, then

$$\begin{aligned}\overline{P} &\le \tilde{P}_k \\ &= AM_1A' + Q - A\begin{bmatrix} CA \\ C \end{bmatrix}^{-1}\begin{bmatrix} CQ \\ 0 \end{bmatrix}\begin{bmatrix} CQ \\ 0 \end{bmatrix}'\begin{bmatrix} CA \\ C \end{bmatrix}^{-1'}A' \\ &\triangleq \overline{M}.\end{aligned}$$

Therefore we obtain the same results as in Section IV with the new $\overline{M}$.

B. *Computing* $\mathbf{Pr}[N_k \ge k_i]$

Theorem 4.6 provides a lower and an upper bound for $\mathbf{Pr}[P_k \le M]$. Both bounds involve the term $\mathbf{Pr}[N_k \ge k_i]$. In this section, we show how we can compute $\mathbf{Pr}[N_k \ge k_i]$ given a packet arrival and drop model.

Let k_1 and k_2 be given (see next section for their computation and approximation). In order to compute $\mathbf{Pr}[N_k \ge k_i]$, we need to have a model that describes packet arrival and drop behaviors. The most commonly used models in literature are

1) I.I.D model: i.e., γ_k's are independent and identically distributed (I.I.D) Bernoulli random variables with mean γ, e.g., [8], [10].
2) Gilbert-Elliott model: i.e., a two state Markov chain is used to describe the transition from γ_k to γ_{k+1}, e.g., [19], [20].

We give closed form solutions to both models in this section.

1) I.I.D Model: If γ_k's are i.i.d Bernoulli random variables with rate γ, then

$$\mathbf{Pr}[N_k \ge k_i] = \mathbf{Pr}\left[\gamma_k = 0, \cdots, \gamma_{k-k_i+1} = 0\right] = (1-\gamma)^{k_i}. \tag{16}$$

2) Gilbert-Elliott Model: Now consider a two-state (0 or 1) Markov chain that represents packet drops and arrivals (Fig. 5). Let T denote the state transition probability matrix, i.e.,

$$T = \begin{bmatrix} \beta & 1-\beta \\ 1-\gamma & \gamma \end{bmatrix}.$$

Let $\pi = [\pi_0\ \pi_1]$ be the steady-state distribution of the Markov chain, i.e., $\pi = \pi T$. Then π can be computed as

$$\pi = \left[\frac{1-\gamma}{2-\gamma-\beta}\ \frac{1-\beta}{2-\gamma-\beta}\right].$$

Assume the Markov chain starts from the steady-state, then with some manipulation, $\mathbf{Pr}[N_k \ge k_i]$ can be shown to be

$$\mathbf{Pr}[N_k \ge k_i] = \frac{1-\gamma}{\beta(2-\gamma-\beta)}\beta^{k_i} \tag{17}$$

Remark 4.9: If $\beta = 1-\gamma$, i.e., we have the i.i.d model, (17) then becomes:

$$\mathbf{Pr}[N_k \ge k_i] = \frac{1-\gamma}{\beta(2-\gamma-\beta)}\beta^{k_i} = \beta^{k_i} = (1-\gamma)^{k_i}$$

which is exactly the same as in (16).

C. *Computing* k_i

In the previous section, we calculate the term $\mathbf{Pr}[N_k \ge k_i]$ in (16) and (17) for two different models that describe packet arrival and drop behavior. In this section, we show how we can compute k_i.

In general, k_i can be computed from their definitions, i.e., we check whether $h(\overline{M}) \le M$ (or $h(\overline{P}) \le M$) is satisfied. If the answer is yes, we check whether $h^2(\overline{M}) \le M$ (or $h^2(\overline{P}) \le M$) is satisfied, and so on and so forth until k_i is found. However, k_i could be unbounded when A is stable and M is sufficiently large. Even when A is unstable, depending on the parameters, k_i could be very large. Therefore computing k_i from their definitions may be time consuming or even result in infinite computations. The good news is that from previous section, we see that using either the i.i.d or the Gilbert-Elliott model, when k_i is sufficiently large

$$\mathbf{Pr}[N_k \ge k_i] \approx 0.$$

For example, when using the i.i.d model, in order that

$$(1-\gamma)^{k_i} \le \epsilon$$

we only need to set

$$k_i = U_i \triangleq \left\lceil \frac{\log \epsilon}{\log(1-\gamma)} \right\rceil. \tag{18}$$

When using the Gilbert-Elliott model, in order that

$$\frac{1-\gamma}{\beta(2-\gamma-\beta)}\beta^{k_i} \le \epsilon$$

we only need to set

$$k_i = U_i \triangleq \left\lceil \frac{\log \epsilon - \log c_3}{\log \beta} \right\rceil \tag{19}$$

where $c_3 = (1-\gamma)/\beta(2-\gamma-\beta)$.

When $\gamma = 0.5$, $\epsilon = 10^{-20}$, (18) returns $k_i = 67$. Using the same ϵ and let $\beta = 0.5$, $\gamma = 0.8$, (19) returns $k_i = 66$. Hence we can use 67 or 66 to approximate the true k_i. Thus we propose Algorithm 1 to compute (and approximate) k_i, where in the algorithm $\Phi_1 = \overline{M}$ and $\Phi_2 = \overline{P}$.

Algorithm 1 INCREMENT SEARCH ALGORITHM

$k_i := 1$

while $h^{k_i}(\Phi_i) \leq M$ and $k_i \leq U_i$ **do**

  $k_i := k_i + 1$

end while

There are a few cases where we can make Algorithm 1 run faster or get better approximation, depending on whether A is stable or not. We discuss those cases below.

1) When A is Stable: It is well known that the Lyapunov equation $AXA' + Q = X$ for A being stable and $Q \geq 0$ has a unique solution M^*. Since

$$\overline{P} \leq h(\overline{P}) \leq h^2(\overline{P}) \leq \cdots$$

we immediately obtain $h^t(\overline{P}) \leq M^*$ for all $t \geq 1$. Thus if $M \geq M^*$

$$k_2 = \min\left\{t \geq 1 : h^t(\overline{P}) \nleq M\right\} = \infty$$

and as a result

$$\mathbf{Pr}[N_k \geq k_2] = 0.$$

2) When A is Unstable: In this case we can find k_2 efficiently via Algorithm 2. The efficiency and the correctness of the algorithm is easily seen.

Algorithm 2 BINARY SEARCH ALGORITHM

$t := 0$

$k_2 := 2^t$

while $h^{k_2}(\overline{P}) \leq M$ and $k_2 \leq U_2$ **do**

  $t := t + 1$

  $k_2 := 2^t$

end while

if $t > 0$ **then**

  $l := 2^{t-1}$

  $u := 2^t$

  $m := \lceil (l + u)/2 \rceil$

  while $l < u$ **do**

    if $h^m(\overline{P}) \leq M$ **then**

      $l := m$

      $m := \lceil (l + u)/2 \rceil$

    else

      $u := m$

      $m := \lceil (l + u)/2 \rceil$

    end if

  end while

end if

D. *Finding $\mathbb{E}[P_k]$: Scalar Case*

In this section, we show that we are able to recover the results in [8] using $\mathbf{Pr}[P_k \leq M]$ as a metric for scalar systems. Let us consider (1) and (2) with

$$A = a > 1,\ Q = q > 0,\ C = c > 0,\ R = r > 0.$$

Notice that in the scalar case, the assumption that (a, c) is observable and $(a, \sqrt{q})$ is controllable holds trivially.

From Lemma A.4 in Appendix A, we can write $\mathbb{E}[P_k]$ as

$$\begin{aligned}\mathbb{E}[P_k] &= \int_0^\infty (1 - \mathbf{Pr}[P_k \leq M])\, dM \\ &= \int_0^{\overline{M}} (1 - \mathbf{Pr}[P_k \leq M])\, dM \\ &\quad + \int_{\overline{M}}^\infty (1 - \mathbf{Pr}[P_k \leq M])\, dM.\end{aligned}$$

Using the fact

$$0 \leq \mathbf{Pr}[P_k \leq M] \leq 1$$

we have

$$\mathbb{E}[P_k] \geq \int_{\overline{M}}^\infty (1 - \mathbf{Pr}[P_k \leq M])\, dM$$

and

$$\mathbb{E}[P_k] \leq \overline{M} + \int_{\overline{M}}^\infty (1 - \mathbf{Pr}[P_k \leq M])\, dM.$$

From Theorem 4.6, we know that when $M \geq \overline{M}$

$$1 - \mathbf{Pr}[N_k \geq k_1] \leq \mathbf{Pr}[P_k \leq M] \leq 1 - \mathbf{Pr}[N_k \geq k_2].$$

Since in [8], i.i.d packet drop model is used, from (16), we have

$$\begin{aligned}\mathbf{Pr}[N_k \geq k_1] &= (1-\gamma)^{k_1},\\ \mathbf{Pr}[N_k \geq k_2] &= (1-\gamma)^{k_2}.\end{aligned}$$

Therefore we obtain

$$\mathbb{E}[P_k] \leq \int_{\overline{M}}^\infty (1-\gamma)^{k_1(M)} dM + \overline{M} \tag{20}$$

and

$$\mathbb{E}[P_k] \geq \int_{\overline{M}}^\infty (1-\gamma)^{k_2(M)} dM. \tag{21}$$

Recall that $k_1(M) = \min\{t \geq 1 : h^t(\overline{M}) \nleq M\}$ and

$$\begin{aligned} h^t(\overline{M}) &= a^{2t}\overline{M} + q(1 + a^2 + \cdots + a^{2t-2}) \\ &= \left(\overline{M} + \frac{q}{a^2-1}\right)a^{2t} - \frac{q}{a^2-1} = c_1 a^{2t} - c_2 \end{aligned}$$

where

$$c_1 = \overline{M} + \frac{q}{a^2-1},\ c_2 = \frac{q}{a^2-1}.$$

Therefore for any $t \geq 1$, if

$$c_1 a^{2t-2} - c_2 \leq M < c_1 a^{2t} - c_2$$

then $k_1(M) = t$. From (20), we have

$$\begin{aligned} \mathbb{E}[P_k] &\leq \overline{M} + \int_{\overline{M}}^{\infty} (1-\gamma)^{k_1(M)} dM \\ &= \overline{M} + \sum_{t=1}^{\infty} \int_{c_1 a^{2t-2} - c_2}^{c_1 a^{2t} - c_2} (1-\gamma)^t dM \\ &= \overline{M} + \sum_{t=1}^{\infty} (c_1 a^{2t} - c_1 a^{2t-2})(1-\gamma)^t \\ &= \overline{M} + \sum_{t=1}^{\infty} c_1 \left(1 - \frac{1}{a^2}\right)(a^2 - \gamma a^2)^t. \end{aligned}$$

Clearly $\mathbb{E}[P_k]$ converges if $a^2 - \gamma a^2 < 1$, i.e.,

$$\gamma > 1 - \frac{1}{a^2}. \tag{22}$$

Similarly from (21), we have

$$\begin{aligned} \mathbb{E}[P_k] &\geq \int_{\overline{M}}^{\infty} (1-\gamma)^{k_2(M)} dM \\ &= \sum_{t=1}^{\infty} \int_{c_1' a^{2t-2} - c_2}^{c_1' a^{2t} - c_2} (1-\gamma)^t dM \\ &= \sum_{t=1}^{\infty} \left(c_1' a^{2t} - c_1 a^{2t-2}\right)(1-\gamma)^t \\ &= \sum_{t=1}^{\infty} c_1' \left(1 - \frac{1}{a^2}\right)(a^2 - \gamma a^2)^t \end{aligned}$$

where $c_1' = \overline{P} + (q/(a^2-1))$. Hence $\mathbb{E}[P_k]$ diverges if $a^2 - \gamma a^2 \geq 1$, i.e.,

$$\gamma \leq 1 - \frac{1}{a^2}. \tag{23}$$

From (22) and (23), we conclude that

$$\lambda_c = 1 - \frac{1}{a^2}$$

which is exactly the same as (11) for scalar systems. Furthermore if we assume

$$\gamma > 1 - \frac{1}{a^2}$$

then we have

$$\frac{c_1'(a^2-1)(1-\gamma)}{1 - a^2 + \gamma a^2} \leq \mathbb{E}[P_k] \leq \frac{c_1(a^2-1)(1-\gamma)}{1 - a^2 + \gamma a^2} + \overline{M}.$$

V. Sensors With Unlimited Computation

We now consider the second scenario in Fig. 1, i.e., when the sensor has unlimited computation capability, and it can preprocesses y_k and send $\hat{x}_k^s$ to the remote estimator. At the estimator side, it is clear that the optimal state estimate and error covariance evolve as

$$(\hat{x}_k, P_k) = \begin{cases} (A\hat{x}_{k-1}, h(P_{k-1})), & \text{if } \gamma_k = 0, \\ (\hat{x}_k^s, P_k^s), & \text{if } \gamma_k = 1. \end{cases}$$

For any $P_0^s \geq 0, P_k^s$ converges to $\overline{P}$ exponentially fast. Therefore without loss of generality, we assume the Kalman filter enters steady-state at the sensor side and hence $P_k^s = \overline{P}$, then we can write P_k as

$$P_k = \begin{cases} h(P_{k-1}), & \text{if } \gamma_k = 0, \\ \overline{P}, & \text{if } \gamma_k = 1. \end{cases}$$

In Section IV, we have defined k_2 and N_k in (13) and (14) respectively. With these two numbers, we have the following result that gives the exact form of $\mathbf{Pr}[P_k \leq M]$.

Theorem 5.1: Assume the Kalman filter enters steady-state at the sensor side and hence $P_k^s = \overline{P}$. Then for any $M \geq \overline{P}$, we have

$$\mathbf{Pr}[P_k \leq M] = 1 - \mathbf{Pr}[N_k \geq k_2]. \tag{24}$$

Proof: A simple way to prove this runs as follows. In this case, when $\gamma_k = 1$, $P_k = \overline{P}$, hence if we let $\overline{M} = \overline{P}$, we immediately obtain $P_k \leq \overline{M}$. As a result, $k_1 = k_2$ and Theorem 5.1 follows directly from Theorem 4.6. ■

Computing $\mathbf{Pr}[N_k \geq k_2]$ and k_i follows exactly the same way as in Sections IV-B and IV-C. Since we get a strict equality for $\mathbf{Pr}[P_k \leq M]$, the special cases we considered in Section IV-B have simpler forms. For example, when A is stable and when $M \geq M^*$, we have shown that $\mathbf{Pr}[N_k \geq k_2] = 0$, therefore we obtain $\mathbf{Pr}[P_k \leq M] = 1$. Assume

$$\gamma > 1 - \frac{1}{a^2}$$

then similar to Section IV-D, we have

$$\frac{c_1'(a^2-1)(1-\gamma)}{1 - a^2 + \gamma a^2} \leq \mathbb{E}[P_k] \leq \frac{c_1'(a^2-1)(1-\gamma)}{1 - a^2 + \gamma a^2} + \overline{P}$$

where $c_1' = \overline{P} + (q/(a^2-1))$. If

$$\frac{c_1'(a^2-1)(1-\gamma)}{1 - a^2 + \gamma a^2} \gg \overline{P}$$

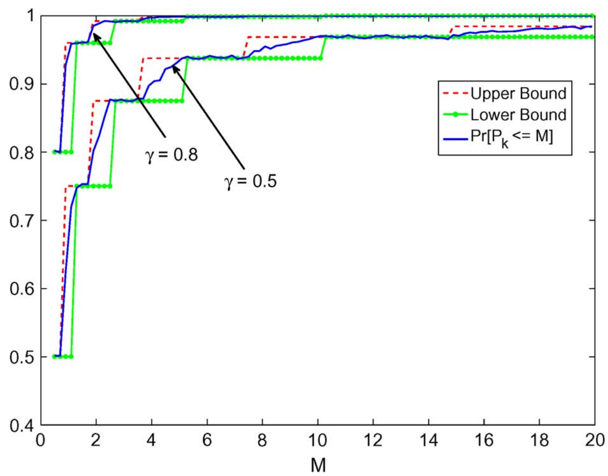


Fig. 6. Sensor with Limited Computation: $\mathbf{Pr}[P_k \leq M]$ and its lower and upper bounds.

then

$$\mathbb{E}[P_k] \approx \frac{c_1'(a^2-1)(1-\gamma)}{1-a^2+\gamma a^2}.$$

As pointed out in [12], it is always better to let the sensor (if it is able to) preprocess the measurement before send it out. This is also seen from the fact that the upper bound of $\mathbf{Pr}[P_k \leq M]$ in (15) is achieved when the sensor has unlimited computation (i.e., (24)).

Remark 5.2: An interesting thing to notice is that P_k can only take values over a countable set $\{h^m(\overline{P}) : m = 0, 1, 2, \ldots\}$. Take the I.I.D packet-drop model for example, we have $\mathbf{Pr}[P_k = h^m(\overline{P})] = \gamma(1-\gamma)^m$, and in particular, $\mathbf{Pr}[P_k \neq h^m(\overline{P})] = 0$. From this, we can easily compute every quantity associated with P_k such as $\mathbb{E}[P_k]$, $\mathbf{Pr}[P_k \leq M]$, etc. This also explains the stair-like values of $\mathbf{Pr}[P_k \leq M]$ (e.g., Figs. 7 and 9) and its associated lower and upper bounds (e.g., Fig. 6) in the example session. For the first scenario, i.e., the sensor has limited computation, to the best of our knowledge, it is still an open problem whether the steady-state distribution of P_k exists.

VI. Example

A. Scalar System With I.I.D Packet Arrivals

Consider (1) and (2) with

$$A = 1.4,\ C = 1,\ Q = 0.2, R = 0.5.$$

The packet arrivals are assumed to be I.I.D and we run a Monte Carlo simulation for both scenarios considered in Sections IV and V, respectively. In the simulation, we use the empirical probability of the event $P_k \leq M$.

For scenario one, i.e., when the sensor has limited computation and only the measurement packet is sent across the network, we plot the value of $\mathbf{Pr}[P_k \leq M]$ and its lower and upper bounds for two different values of γ. As we can see from Fig. 6 that the lower and upper bounds that we have derived in (15) provide tight approximation of $\mathbf{Pr}[P_k \leq M]$. We also notice that $\mathbf{Pr}[P_k \leq M]$ increases with larger γ which leads to better estimator performance.

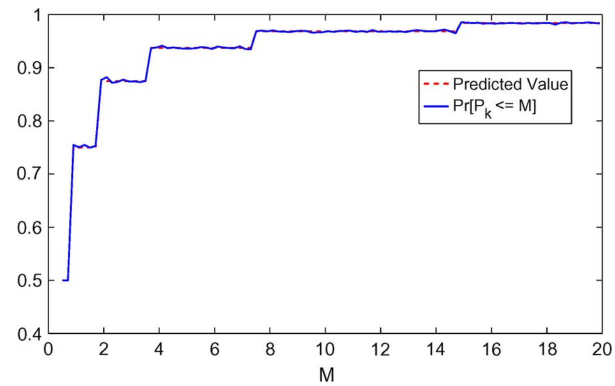


Fig. 7. Sensor with Unlimited Computation: $\mathbf{Pr}[P_k \leq M]$.

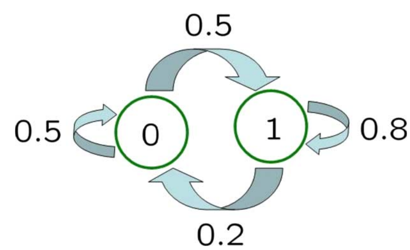


Fig. 8. Packet arrival and drop model.

For scenario two, i.e., when the local estimate is sent across the network, we can also see from Fig. 7 that the predicted value of $\mathbf{Pr}[P_k \leq M]$ given by (24) agrees well with the true value of $\mathbf{Pr}[P_k \leq M]$.

B. Vector System With Markov Packet Arrivals

Consider a vehicle moving in a two dimensional space according to the standard constant acceleration model, which assumes that the vehicle has zero acceleration except for a small perturbation. The example was considered in [21]. The state of the vehicle consists of its x and y positions as well as velocities. Assume a sensor measures the positions of the vehicle and sends the measurements to a remote estimator over a packet-dropping network. The system parameters are given according to (1) and (2) as follows:

$$A = \begin{bmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The process and measurement noise covariances are $Q = \mathbf{diag}(0.01, 0.01, 0.01, 0.01)$ and $R = \mathbf{diag}(0.001, 0.001)$. The packet arrival and drop is modeled as a Markov chain with state transition probabilities shown in Fig. 8.

Similar to the scalar example, when the local estimate is sent across the network, the predicted values of $\mathbf{Pr}[P_k \leq M]$ from (24) matches well with the actual value as seen from Fig. 9. As C is not invertible in this case, the sensor stores its previous measurement y_{k-1} and sends it along with y_k at time k. Unlike the scalar system example, where the actual value of $\mathbf{Pr}[P_k \leq M]$ lies midway between the lower and upper bounds given by (15),

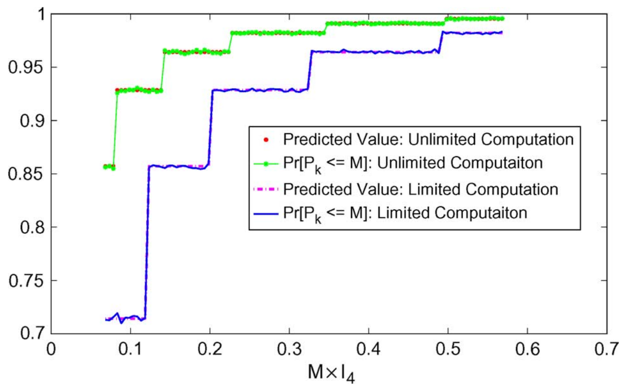


Fig. 9. Vector system example.

the actual value here approaches the lower bound. This happens as whenever a data packet is received, the error covariance is reset to $\overline{M}$ using the estimation scheme in Remark 4.8. In both cases, $\mathbf{Pr}[N_k \geq k_i]$ is obtained from (17).

VII. Conclusion

In this paper, we study the problem of state estimation of a discrete time process over a packet-dropping network based on a modified Kalman filter. We consider a probabilistic metric on the error covariance matrix, i.e., $\mathbf{Pr}[P_k \leq M]$. The advantage of the new metric is easy to see compared with the most widely used performance metric in literature, e.g., $\mathbb{E}[P_k]$, as the new metric completely characterizes the behavior of P_k.

When the sensor has limited computation capability, we derive lower and upper bounds for $\mathbf{Pr}[P_k \leq M]$. Both bounds depend on the underlying model that describes packet arrival and drop behavior of the communication network between the sensor and the estimator. When the sensor has unlimited computation capability, we are able to compute $\mathbf{Pr}[P_k \leq M]$ in an exact form. We also recover the result for scalar systems in [8].

There are many interesting directions for continuing this work, which include: finding better estimation scheme that outperforms the simple linear estimation scheme presented in Remark 4.8; finding better bounds of $\mathbf{Pr}[P_k \leq M]$ for the first scenario that the sensor has limited computation; extending the results in Section IV-D to general vector systems; studying closed loop system performance from a probabilistic angle; looking at distributed and cooperative control problems over packet-dropping networks; and experimentally evaluating the theory developed in the paper.

Appendix

A. Supporting Lemmas

Lemma A.1: For any $0 \leq X \leq Y$

$$\begin{aligned} h(X) &\leq h(Y), \\ g(X) &\leq g(Y), \\ \tilde{g}(X) &\leq \tilde{g}(Y), \\ \tilde{g}(X) &\leq X, \\ h \circ \tilde{g}(X) &= g(X), \\ g(X) &\leq h(X). \end{aligned}$$

Proof: $h(X) \leq h(Y)$ holds as $h(X)$ is affine in X. Proof for $g(X) \leq g(Y)$ can be found in Lemma 1-c in [8]. As $\tilde{g}$ is a special form of g by setting $A = I$ and $Q = 0$, we immediately obtain $\tilde{g}(X) \leq \tilde{g}(Y)$. Next we have

$$\tilde{g}(X) = X - XC'[CXC' + R]^{-1}CX \leq X$$

and

$$\begin{aligned} h \circ \tilde{g}(X) &= h\left(X - XC'[CXC' + R]^{-1}CX\right) \\ &= A\left(X - XC'[CXC' + R]^{-1}CX\right)A' + Q \\ &= g(X). \end{aligned}$$

Finally we have

$$g(X) = h(X) - AXC'[CXC' + R]^{-1}CXA' \leq h(X).$$

■

Lemma A.2: For any $X \geq 0$, $\tilde{g}(X) \leq \overline{M}$.

Proof: For any $t > 0$, we have

$$\begin{aligned} \tilde{g}(t\overline{M}) &= t\overline{M} - t^2\overline{M}C'[tC\overline{M}C' + R]^{-1}C\overline{M} \\ &= t\overline{M} - \frac{t^2}{t+1}\overline{M}C'R^{-1}C\overline{M} \\ &= \frac{t}{t+1}\overline{M} \leq \overline{M}. \end{aligned}$$

For all $X \geq 0$, since $\overline{M} > 0$, it is clear that there exists $t_1 > 0$ such that $t_1\overline{M} > X$. Therefore

$$\tilde{g}(X) \leq \tilde{g}(t_1\overline{M}) \leq \overline{M}$$

by using the fact that $\tilde{g}(X) \leq \tilde{g}(Y)$ if $0 \leq X \leq Y$. ■

Lemma A.3: $\overline{P} \leq h(\overline{P})$.

Proof:

$$h(\overline{P}) = h \circ \tilde{g}(P^*) = g(P^*) = P^* \geq \tilde{g}(P^*) = \overline{P}$$

where the first and the last equality are from the definition of $\overline{P}$, the third equality is from the definition of P^*. The rest equality and inequality are from Lemma A.1. ■

Lemma A.4 : Let X be a continuous random variable defined on $[0, \infty)$ and let $F(x) = \mathbf{Pr}[X \leq x]$. Then

$$\mathbb{E}[X] = \int_0^\infty [1 - F(x)]\, dx.$$

Proof: See Lemma (4) in [22], page 93. ■

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions which helped improve the paper.

References

[1] W. S. Wong and R. W. Brockett, “Systems with finite communication bandwidth-part i: State estimation problems,” *IEEE Trans. Autom. Control*, vol. 42, no. 9, pp. 1294–1299, Sep. 1997.

[2] W. S. Wong and R. W. Brockett, “Systems with finite communication bandwidth-part ii: Stabilization with limited information feedback,” *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

[3] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1279–1289, Jul. 2000.

[4] G. N. Nair and R. J. Evans, "Communication-limited stabilization of linear systems," in *Proc. 39th Conf. Decision Control*, Dec. 2000, vol. 1, pp. 1005–1010.

[5] S. C. Tatikonda, "Control Under Communication Constraints," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2000.

[6] I. R. Petersen and A. V. Savkinc, "Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel," in *Proc. 40th Conf. Decision Control*, Dec. 2001, vol. 1, pp. 304–309.

[7] N. Elia, "Remote stabilization over fading channels," *Syst. Control Lett.*, vol. 54, no. 3, pp. 237–249, Mar. 2005.

[8] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[9] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[10] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. 43rd IEEE Conf. Decision Control*, Dec. 2004, vol. 4, pp. 4180–4186.

[11] Z. Jin, V. Gupta, and R. M. Murray, "State estimation over packet dropping networks using multiple description coding," *Automatica*, vol. 42, no. 9, pp. 1441–1452, 2006.

[12] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "On lqg control across a stochastic packet-dropping link," in *Proc. Amer. Control Conf.*, 2005, pp. 360–365.

[13] V. Gupta, B. Sinopoli, S. Adlakha, and A. Goldsmith, "Receding horizon networked control," in *Proc. Allerton Conf. Commun., Control Comput.*, Monticello, IL, Sep. 2006, [CD ROM].

[14] M. Huang and S. Dey, "Stability of kalman filtering with markovian packet losses," *Automatica*, vol. 43, pp. 598–607, 2007.

[15] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[16] L. Shi, M. Epstein, A. Tiwari, and R. Murray, "Estimation with information loss: Asymptotic analysis and error bounds," in *Proc. IEEE Conf. Decision Control*, Dec. 2005, pp. 1215–1221.

[17] M. Epstein, L. Shi, A. Tiwari, and R. M. Murray, "Probabilistic performance of state estimation across a lossy network," *Automatica*, vol. 44, pp. 3046–3053, 2008.

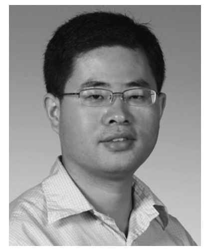
[18] F. Lian, J. Moyne, and D. Tillbury, "Peformance evaluation of control networks," *IEEE Control Syst. Mag.*, vol. 21, no. 1, pp. 66–83, Jan. 2001.

[19] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sep. 1960.

[20] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sep. 1963.

[21] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[22] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. London, U.K.: Oxford Univ. Press, 2001.



Ling Shi received the B.S. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology, Hong Kong, China, in 2002 and the Ph.D. degree in control and dynamical systems from The California Institute of Technology, Pasadena, in 2008.

He is currently an Assistant Professor in the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks, and distributed estimation and control.

Michael Epstein received the B.S. degree (with highest honors) from the University of California at Los Angeles in 2002 and the M.S. and Ph.D. degrees from the California Institute of Technology, Pasadena, in 2003 and 2007, respectively, all in mechanical engineering.

His general research interests are in all topics related to control and dynamical systems.

Richard M. Murray received the B.S. degree in electrical engineering from the California Institute of Technology, Pasadena, in 1985 and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, in 1988 and 1991, respectively.

He is currently the Thomas E. and Doris Everhart Professor of Control and Dynamical Systems and the Director for Information Science and Technology at the California Institute of Technology. His research is in the application of feedback and control to mechanical, information, and biological systems.