

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, $\Pr[P_k \leq M]$, 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 $\Pr[P_k \leq M]$. 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 $\Pr[P_k \leq M]$. We also recover the results in the literature when using $\Pr[P_k \leq M]$ as a metric for scalar systems. Examples are provided to illustrate the theory developed in the paper.

I. INTRODUCTION

In the past decade, networked control systems have gained attention from both the control community and the network and communication community [1]. 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 and maintain and later diagnose in case of malfunctioning, and increase system agility. Although networked control systems have advantages, inserting a network in between the plant and the controller can introduce

Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91106. Email: {shiling, epstein, murray}@cds.caltech.edu.

Tel: (626) 395-2313, Fax: (626) 395-6170.

The work is supported in part by AFOSR grant FA9550-04-1-0169.

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.

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 [2], [3] and further pursued by others (e.g., [4]–[7]). Elia [8] considered the problem of stabilizing a networked control system over fading channels. Sinopoli et al. [9] 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 [10] where stability region of packet arrival rates are provided. Following the spirit of [9], Liu and Goldsmith [11] 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. A scheme based on multi-description coding for packet dropping networks, but limited to the estimation, is considered in [12]. Gupta et al. [13] 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. The readers are referred to [14] 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 [9] and [11], 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 to diverge, and it might be better to ignore such events that occur with 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., $\Pr[P_k \leq M]$. For scalar systems, this is equivalent to considering the cumulative distribution of the random variable P_k .

In [15] 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 [16], 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 $\Pr[P_k \leq M]$, 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 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 [16] that requires A to be unstable and work with arbitrary A .
- 3) We are able to recover the result in [9] for scalar systems, i.e., from the result using $\Pr[P_k \leq M]$ as a metric, we derive the stability result using $\mathbb{E}[P_k \leq M]$ 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 the previously derived lower and upper bounds on $\Pr[P_k \leq M]$ are the same and hence we are able to give an exact expression for $\Pr[P_k \leq M]$.

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 [9] 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 $\Pr[P_k \leq M]$. 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

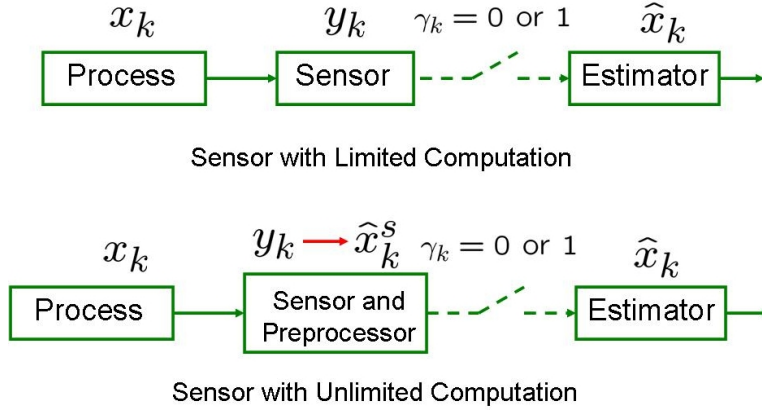


Fig. 1. System Block Diagram

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:

$$x_k = Ax_{k-1} + w_{k-1}, \quad (1)$$

$$y_k = Cx_k + v_k. \quad (2)$$

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 $\mathbb{E}[w_k w_k'] = \delta_{kj} Q \geq 0$, $\mathbb{E}[v_k v_k'] = \delta_{kj} R > 0$, $\mathbb{E}[w_k v_j'] = 0 \forall j, k$, 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

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

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 [17]. 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:

$$\begin{aligned}\hat{x}_k^- &\triangleq \mathbb{E}[x_k | \text{all data packets up to } k-1], \\ \hat{x}_k &\triangleq \mathbb{E}[x_k | \text{all data packets up to } k], \\ P_k^- &\triangleq \mathbb{E}[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)'], \\ P_k &\triangleq \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)'].\end{aligned}$$

As mentioned in Section I, we are interested in finding a closed form solution to $\Pr[P_k \leq M]$ given M . In the next few sections, we consider the two scenarios in Fig. 1 and provide results on $\Pr[P_k \leq M]$ 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; $X \in \mathbb{S}_+^n$ where \mathbb{S}_+^n is the set of n by n positive semi-definite matrices; f_i is a function from $\mathbb{S}_+^n \rightarrow \mathbb{S}_+^n$, $i = 1, 2$; Y_i is a random variable where the underlying sample spaces will be clear from its context.

$$\begin{aligned}
\rho(A) &\triangleq \max_i |\lambda_i(A)|, \\
h(X) &\triangleq AXA' + Q, \\
g(X) &\triangleq h(X) - AXC'[CXC' + R]^{-1}CXA', \\
\tilde{g}(X) &\triangleq X - XC'[CXC' + R]^{-1}CX, \\
f_1 \circ f_2(X) &\triangleq f_1(f_2(X)), \\
\Pr[Y_1|Y_2] &\triangleq \Pr[Y_1] \text{ given } Y_2.
\end{aligned}$$

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 Eqn (1) and (2) is a standard Kalman filter, denoted as **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:

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

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

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

Let P^* be the unique positive semi-definite solution¹ to $g(X) = X$, i.e., $P^* = g(P^*)$. Define \bar{P} as $\bar{P} \triangleq \tilde{g}(P^*)$. Then we have

$$\tilde{g} \circ h(\bar{P}) = \tilde{g} \circ h \circ \tilde{g}(P^*) = \tilde{g} \circ g(P^*) = \tilde{g}(P^*) = \bar{P},$$

where we use the fact that $h \circ \tilde{g} = g$. In other words,

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

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

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. [9] 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 **KF** except that

$$\hat{x}_k = \hat{x}_k^- + \gamma_k K_k (y_k - C \hat{x}_k^-) \quad (3)$$

$$P_k = P_k^- - \gamma_k K_k C P_k^- \quad (4)$$

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 [9] that there exists a critical value γ_c such that if $\gamma > \gamma_c$, $\mathbb{E}[P_k]$ converges as $k \rightarrow \infty$ and diverges otherwise. When C^{-1} exists, γ_c is given in exact form as

$$\gamma_c = 1 - \frac{1}{\rho(A)^2}. \quad (5)$$

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 evaluate the performance of the estimator. For example, consider the unstable scalar system with $a = 2$, $q = 1$, $P_0 = 1$ in Eqn (1). Let the packet arrival rate γ be

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

Then from [9] we conclude that

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

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

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

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 $\Pr[P_k \leq M]$ 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 $\Pr[P_k \leq M]$

Similar to [9], the optimal state estimate \hat{x}_k and its error covariance matrix P_k are given by

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

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 4.7 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. ■

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

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

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:

$$k_1(M) \triangleq \min\{t \geq 1 : h^t(\overline{M}) \not\leq M\}, \quad (6)$$

$$k_2(M) \triangleq \min\{t \geq 1 : h^t(\overline{P}) \not\leq M\}. \quad (7)$$

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) $\bar{P} = \tilde{g}(P^*) \leq \bar{M}$ 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

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

which violates the definition of k_2 . Notice that we use the property that h is nondecreasing as well as $h(\bar{P}) \geq \bar{P}$ from Lemma A.1 and A.3 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 $P_0 \geq \bar{P}$. Then for all $k \geq 0$, $P_k \geq \bar{P}$.

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

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

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

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

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

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

The induction step is thus complete. ■

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

$$N_k \triangleq \min\{t \geq 0 : \gamma_{k-t} = 1\}. \quad (8)$$

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 Eqn (1) and (2) with $A = 1.4, C = 1, Q = 0.2, R = 0.5$. It is easy to verify that $\bar{P} = 0.3083, \bar{M} = 0.5$. 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 \bar{P} and \bar{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.

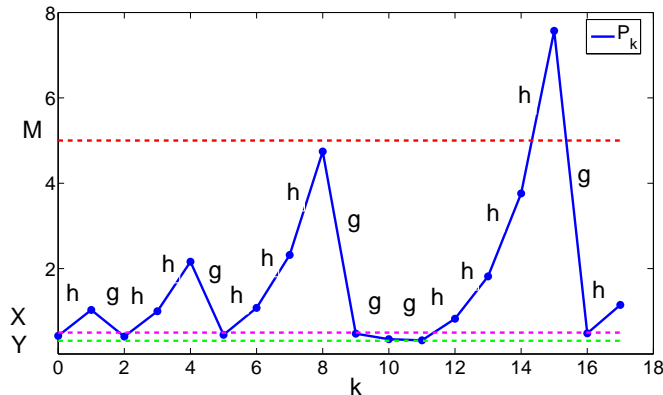


Fig. 2. A Packet Arrival Sequence

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

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

$$1 - \Pr[N_k \geq k_1] \leq \Pr[P_k \leq M] \leq 1 - \Pr[N_k \geq k_2]. \quad (9)$$

Proof: We divide the proof into two parts.

1) Let us first prove

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

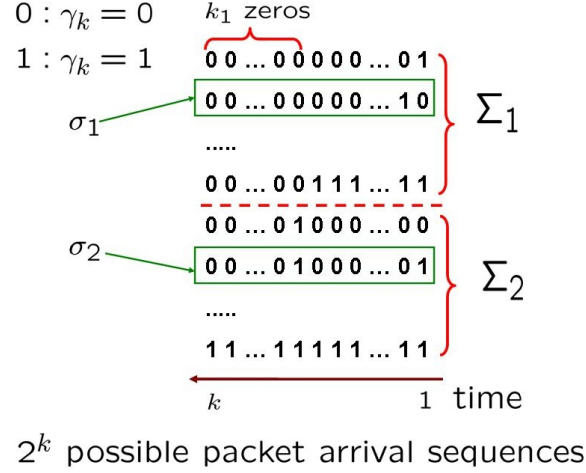
As $\gamma_k = 1$ or 0 , there are in total 2^k possible realizations of γ_1 to γ_k as seen from Fig. 3.

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, $P_{k-k_1+1} \leq \bar{M}$. Therefore we have

$$P_k(\sigma_2) \leq h^{k_1-1}(P_{k-k_1+1}) \leq h^{k_1-1}(\bar{M}) \leq M,$$

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,

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

Fig. 3. $N_k \geq k_1$

Therefore we have

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

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

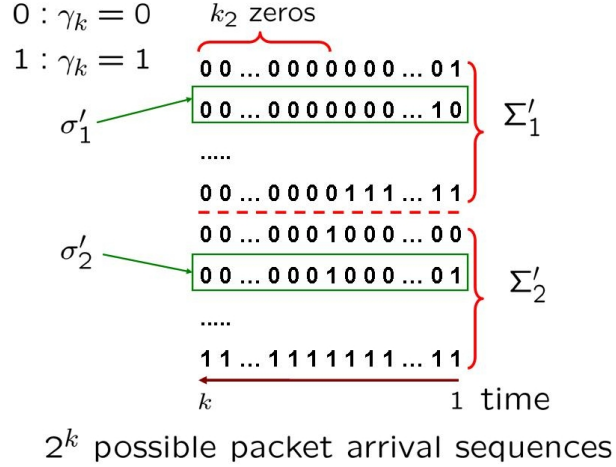
$$\Pr[P_k \leq M] \leq 1 - \Pr[N_k \geq k_2].$$

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,

$$P_k(\sigma'_1) = h^{s(\sigma'_1)}(P_{k-s(\sigma'_1)}) \geq h^{s(\sigma'_1)}(\bar{P}),$$

Fig. 4. $N_k \geq k_2$

where the inequality is from Lemma 4.4. Therefore we conclude $P_k(\sigma'_1) \not\leq M$. Otherwise $h^{s(\sigma'_1)}(\bar{P}) \leq P_k(\sigma'_1) \leq M$, which violates the definition of k_2 . In other words, $\Pr[P_k \leq M | \sigma'_1] = 0$. Therefore we have

$$\begin{aligned}
\Pr[P_k \leq M] &= \sum_{\sigma \in \Sigma'_1 \cup \Sigma'_2} \Pr[P_k \leq M | \sigma] \Pr(\sigma) \\
&= \sum_{\sigma'_1 \in \Sigma'_1} \Pr[P_k \leq M | \sigma'_1] \Pr(\sigma'_1) + \sum_{\sigma'_2 \in \Sigma'_2} \Pr[P_k \leq M | \sigma'_2] \Pr(\sigma'_2) \\
&= \sum_{\sigma'_2 \in \Sigma'_2} \Pr[P_k \leq M | \sigma'_2] \Pr(\sigma'_2) \\
&\leq \sum_{\sigma'_2 \in \Sigma'_2} \Pr(\sigma'_2) = \Pr(\Sigma'_2) = 1 - \Pr(\Sigma'_1) = 1 - \Pr[N_k \geq k_2],
\end{aligned}$$

where the inequality is from the fact that $\Pr[P_k \leq M | \sigma'_2] \leq 1$ for any $\sigma'_2 \in \Sigma'_2$. \blacksquare

Remark 4.7: 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

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

is full rank. In this section, we consider the special case when $r = 2$, and in particular, we

assume $\begin{bmatrix} C \\ CA \end{bmatrix}^{-1}$ 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 \hat{x}_k

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

The corresponding error covariance can be calculated as

$$P_k = AM_1A' + Q,$$

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

$$\bar{P} \leq P_k = AM_1A' + Q \triangleq \bar{M} \text{ if } \gamma_k = 1.$$

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

B. Computing $\Pr[N_k \geq k_i]$

Theorem 4.6 provides a lower and an upper bound for $\Pr[P_k \leq M]$. Both bounds involve the term $\Pr[N_k \geq k_i]$. In this section, we show how we can compute $\Pr[N_k \geq k_i]$ 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 $\Pr[N_k \geq k_i]$, 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 Bernoulli random variables with mean γ , e.g., [9], [11].
- 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., [18], [19].

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

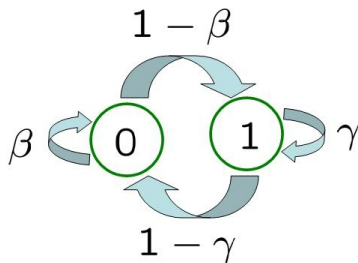


Fig. 5. Gilbert-Elliott Model

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

$$\Pr[N_k \geq k_i] = \Pr[\gamma_k = 0, \dots, \gamma_{k-k_i+1} = 0] = (1 - \gamma)^{k_i}. \quad (10)$$

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

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

Let z_k be defined as

$$z_k = \begin{bmatrix} z_k^1 \\ z_k^2 \end{bmatrix} \triangleq \begin{bmatrix} \Pr[\gamma_k = 0] \\ \Pr[\gamma_k = 1] \end{bmatrix}.$$

Then

$$z_k^1 = \Pr[\gamma_k = 0] = \sum_{i=0}^1 \Pr[\gamma_k = 0 | \gamma_{k-1} = i] \Pr[\gamma_{k-1} = i] = \beta z_{k-1}^1 + (1 - \gamma) z_{k-1}^2.$$

Similarly, we have

$$z_k^2 = (1 - \beta) z_{k-1}^1 + \gamma z_{k-1}^2.$$

Therefore $z_k = T' z_{k-1}$. Assume z_0 is given, then z_k satisfies the following equation

$$z_k = (T')^k z_0, k \geq 1.$$

Furthermore, for k sufficiently large, $z_k \approx \pi'$, i.e.,

$$z_k^1 \approx \pi_0, \quad z_k^2 \approx \pi_1.$$

Therefore we have

$$\begin{aligned} \Pr[N_k \geq k_i] &= \Pr[\gamma_k = 0, \dots, \gamma_{k-k_i+1} = 0] \\ &= \sum_{i=0}^1 \alpha_i \Pr[\gamma_{k-k_i} = i] \\ &= \beta^{k_i} z_{k-k_i}^1 + (1 - \gamma) \beta^{k_i-1} z_{k-k_i}^2 \\ &= \pi_0 \beta^{k_i} + \pi_1 (1 - \gamma) \beta^{k_i-1} \\ &= \frac{1 - \gamma}{\beta(2 - \gamma - \beta)} \beta^{k_i}, \end{aligned} \tag{11}$$

where

$$\alpha_i = \Pr[\gamma_k = 0, \dots, \gamma_{k-k_i+1} = 0 | \gamma_{k-k_i} = i].$$

Remark 4.8: If $\beta = 1 - \gamma$, i.e., we have the i.i.d model, Eqn (11) then becomes:

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

which is exactly the same as in Eqn (10).

C. Computing k_i

In the previous section, we calculate the term $\Pr[N_k \geq k_i]$ in Eqn (10) and (11) 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 $h(\overline{M}) \leq M$ (or $h(\overline{P}) \leq M$) is satisfied. If the answer is yes, we check whether $h^2(\overline{M}) \leq M$ (or $h^2(\overline{P}) \leq M$) 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

$$\Pr[N_k \geq k_i] \approx 0.$$

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

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

we only need to set

$$k_i = U_i \triangleq \lceil \frac{\log \epsilon}{\log(1 - \gamma)} \rceil. \quad (12)$$

When using the Gilbert-Elliott model, in order that

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

we only need to set

$$k_i = U_i \triangleq \lceil \frac{\log \epsilon - \log c_3}{\log \beta} \rceil. \quad (13)$$

where $c_3 = \frac{1 - \gamma}{\beta(2 - \gamma - \beta)}$.

When $\gamma = 0.5$, $\epsilon = 10^{-20}$, Eqn (12) returns $k_i = 67$. Using the same ϵ and let $\beta = 0.5$, $\gamma = 0.8$, Eqn (13) 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 $\Phi_1 = \bar{M}$ and $\Phi_2 = \bar{P}$.

Algorithm 1 INCREMENTAL 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

$$\bar{P} \leq h(\bar{P}) \leq h^2(\bar{P}) \leq \dots,$$

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

$$k_2 = \min\{t \geq 1 : h^t(\bar{P}) \not\leq M\} = \infty,$$

and as a result,

$$\Pr[N_k \geq k_2] = 0.$$

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  $h^{k_2}(\bar{P}) \leq M$  and  $k_2 \leq U_2$  do
     $t := t + 1$ 
     $k_2 := 2^t$ 
end while
if  $t > 0$  then
     $l := 2^{t-1}$ 
     $u := 2^t$ 
     $m := \lceil \frac{l+u}{2} \rceil$ 
    while  $l < u$  do
        if  $h^m(\bar{P}) \leq M$  then
             $l := m$ 
             $m := \lceil \frac{l+u}{2} \rceil$ 
        else
             $u := m$ 
             $m := \lceil \frac{l+u}{2} \rceil$ 
        end if
    end while
end if

```

D. $\mathbb{E}[P_k]$ as a Metric

In this section, we show that we are able to recover the results in [9] using $\Pr[P_k \leq M]$ as a metric for scalar systems. Let us consider Eqn (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

$$\mathbb{E}[P_k] = \int_0^\infty (1 - \Pr[P_k \leq M])dM = \int_0^{\bar{M}} (1 - \Pr[P_k \leq M])dM + \int_{\bar{M}}^\infty (1 - \Pr[P_k \leq M])dM.$$

Using the fact

$$0 \leq \Pr[P_k \leq M] \leq 1,$$

we have

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

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

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

Since in [9], i.i.d packet drop model is used, from Eqn (10), we have

$$\Pr[N_k \geq k_1] = (1 - \gamma)^{k_1},$$

$$\Pr[N_k \geq k_2] = (1 - \gamma)^{k_2}.$$

Therefore we obtain

$$\int_{\bar{M}}^\infty (1 - \gamma)^{k_2(M)} dM \leq \mathbb{E}[P_k] \leq \int_{\bar{M}}^\infty (1 - \gamma)^{k_1(M)} dM + \bar{M}. \quad (14)$$

Recall that $k_1(M) = \min\{t \geq 1 : h^t(\bar{M}) \not\leq M\}$ and

$$h^t(\bar{M}) = a^{2t}\bar{M} + q(1 + a^2 + \dots + a^{2t-2}) = (\bar{M} + \frac{q}{a^2 - 1})a^{2t} - \frac{q}{a^2 - 1} = c_1 a^{2t} - c_2,$$

where

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

therefore for any $t \geq 1$, $k_1(M) = t$, if $c_1 a^{2t-2} - c_2 \leq M < c_1 a^{2t} - c_2$. From Eqn (14), we have

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

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

$$\gamma > 1 - \frac{1}{a^2}. \quad (15)$$

Similarly from Eqn (14), we have

$$\begin{aligned} \mathbb{E}[P_k] &\geq \int_0^\infty (1 - \gamma)^{k_2(M)} dM \\ &= \sum_{t=1}^\infty \int_{c'_1 a^{2t-2} - c_2}^{c'_1 a^{2t} - c_2} (1 - \gamma)^t dM \\ &= \sum_{t=1}^\infty (c'_1 a^{2t} - c_1 a^{2t-2}) (1 - \gamma)^t \\ &= \sum_{t=1}^\infty c'_1 \left(1 - \frac{1}{a^2}\right) (a^2 - \gamma a^2)^t, \end{aligned}$$

where $c'_1 = \bar{P} + \frac{q}{a^2-1}$. Hence $\mathbb{E}[P_k]$ diverges if $a^2 - \gamma a^2 \geq 1$, i.e.,

$$\gamma \leq 1 - \frac{1}{a^2}. \quad (16)$$

From Eqn (15) and (16), we conclude that

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

which is exactly the same as Eqn (5) for scalar systems. Furthermore if we assume

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

then we have

$$\frac{c'_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2} \leq \mathbb{E}[P_k] \leq \frac{c_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2} + \bar{M}.$$

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

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

If we assume k is sufficiently large, e.g., when the Kalman filter enters steady state at the sensor side and $P_k^s = \bar{P}$, we can write P_k as

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

In Section IV, we have defined k_2 and N_k in Eqn (7) and (8) respectively. With these two numbers, we have the following result that gives the exact form of $\Pr[P_k \leq M]$.

Theorem 5.1: Assume k is sufficiently large such that $P_k^s = \bar{P}$. Then for any $M \geq \bar{P}$, we have

$$\Pr[P_k \leq M] = 1 - \Pr[N_k \geq k_2]. \quad (17)$$

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

We show an alternative way to prove the theorem. Let σ'_i and $\Sigma'_i, i = 1, 2$ be defined in the same way as in the proof of Theorem 4.6 (see Fig. 4). Clearly for any $\sigma'_2 \in \Sigma'_2$,

$$P_k(\sigma'_2) \leq h^{k_2-1}(\bar{P}) \leq M.$$

The first inequality is from the fact that $\gamma_{k-k_2+1} = 1$ and hence $P_{k-k_2+1} = \bar{P}$. The second inequality is from the definition of k_2 . In other words,

$$\Pr[P_k \leq M | \sigma'_2] = 1.$$

Similar to the proof of Theorem 4.6, for $\sigma'_1 \in \Sigma'_1$, let us define

$$s = s(\sigma'_1) \triangleq \min\{t \geq 1 : \gamma_{k-t} = 1 | \sigma'_1\}.$$

As $\sigma'_1 \in \Sigma'_1$, $s \geq k_2$. Therefore $P_k(\sigma'_1) = h^s(\bar{P}) \not\leq M$. In other words, $\Pr[P_k \leq M | \sigma'_1] = 0$.

Therefore

$$\begin{aligned} \Pr[P_k \leq M] &= \sum_{\sigma' \in \Sigma'_1 \cup \Sigma'_2} \Pr[P_k \leq M | \sigma'] \Pr(\sigma') \\ &= \sum_{\sigma'_1 \in \Sigma'_1} \Pr[P_k \leq M | \sigma'_1] \Pr(\sigma'_1) + \sum_{\sigma'_2 \in \Sigma'_2} \Pr[P_k \leq M | \sigma'_2] \Pr(\sigma'_2) \\ &= \sum_{\sigma'_2 \in \Sigma'_2} \Pr[P_k \leq M | \sigma'_2] \Pr(\sigma'_2) \\ &= \sum_{\sigma'_2 \in \Sigma'_2} \Pr(\sigma'_2) = \Pr(\Sigma'_2) = 1 - \Pr(\Sigma'_1) = 1 - \Pr[N_k \geq k_2]. \end{aligned}$$

■

Computing $\Pr[N_k \geq k_2]$ and k_i follows exactly the same way as in Section IV-B and IV-C. Since we get a strict equality for $\Pr[P_k \leq M]$, 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 $\Pr[N_k \geq k_2] = 0$, therefore we obtain

$$\Pr[P_k \leq M] = 1.$$

Assume

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

then similar to Section IV-D, we have

$$\frac{c'_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2} \leq \mathbb{E}[P_k] \leq \frac{c'_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2} + \bar{P},$$

where $c'_1 = \bar{P} + \frac{q}{a^2 - 1}$. If

$$\frac{c'_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2} \gg \bar{P},$$

then

$$\mathbb{E}[P_k] \approx \frac{c'_1(a^2 - 1)(1 - \gamma)}{1 - a^2 + \gamma a^2}.$$

VI. EXAMPLE

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

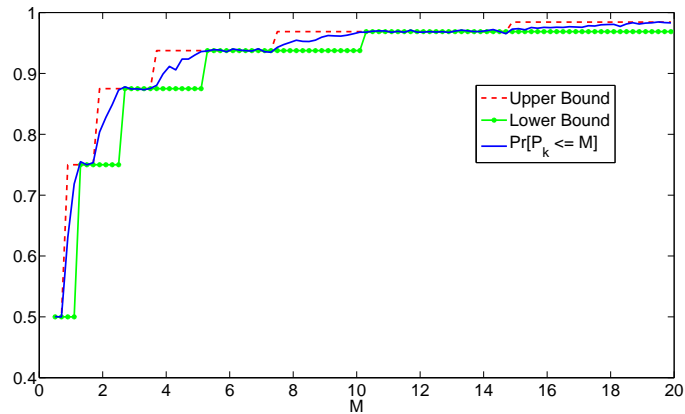
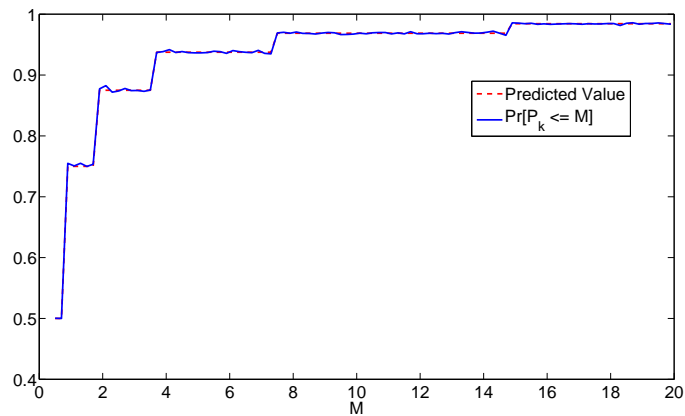
Consider Eqn (1) and (2) with

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

The packet arrivals are assumed to be I.I.D with arrival rate $\gamma = 0.5$. We run a Monte Carlo simulation for both scenarios considered in Section IV and V respectively. During the simulation, we use the relative frequency to measure $\Pr[P_k \leq M]$.

For scenario one, i.e., when measurement packet is sent across the network, we can see from Fig. 6 that the lower and upper bounds that we have derived in Eqn (9) provide tight approximation of $\Pr[P_k \leq M]$.

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 $\Pr[P_k \leq M]$ given by Eqn (17) agrees well with the true value of $\Pr[P_k \leq M]$.

Fig. 6. y_k is sentFig. 7. \hat{x}_k^s is sent

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 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 Eqn (1)

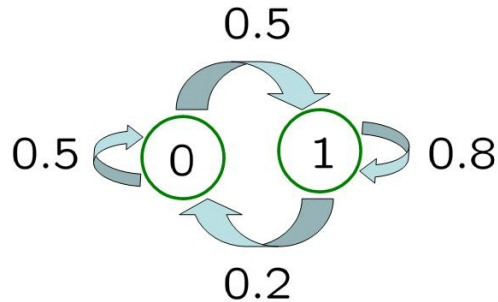


Fig. 8. Packet Arrival and Drop Model

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 $\Pr[P_k \leq M]$ from Eqn (17) 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 $\Pr[P_k \leq M]$ lies in between the lower and upper bounds given by Eqn (9), the actual value here approaches the lower bound. This happens as whenever a data packet is received, the error covariance is reset to \bar{M} using the estimation scheme in Remark 4.7. In both cases, $\Pr[N_k \geq k_i]$ is obtained from Eqn (11).

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., $\Pr[P_k \leq M]$. The advantage of the new metric is easy to

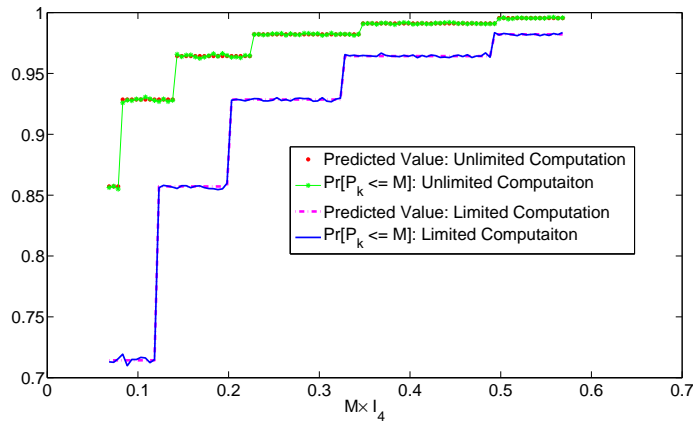


Fig. 9. Vector System Example

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 $\Pr[P_k \leq M]$. 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 $\Pr[P_k \leq M]$ in an exact form. We also recover the result for scalar systems in [9].

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 4.7; 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 [9]. 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

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

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

$$\tilde{g}(t\overline{M}) = \frac{t}{t+1}\overline{M} \leq \overline{M}.$$

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

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

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

Proof:

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

where the first and the last equality are from the definition of \bar{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 $F(x) = \Pr[X \leq x]$. Then

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

Proof: See Lemma (4) in [20], page 93. ■

REFERENCES

- [1] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "Networked control systems: Analysis and design," *Proc. of IEEE, Special Issue on Networked Control Systems*, 2007.
- [2] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth-part i: State estimation problems," *IEEE Trans. Automat. Contr.*, vol. 42, Sept 1997.
- [3] —, "Systems with finite communication bandwidth-part ii: Stabilization with limited information feedback," *IEEE Trans. Automat. Contr.*, vol. 44, May 1999.
- [4] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Automat. Contr.*, vol. 45, July 2000.
- [5] G. N. Nair and R. J. Evans, "Communication-limited stabilization of linear systems," in *Proceedings of the 39th Conf. on Decision and Contr.*, vol. 1, Dec 2000, pp. 1005–1010.
- [6] S. C. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Massachusetts Institute of Technology, 2000.
- [7] I. R. Petersen and A. V. Savkin, "Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel," in *Proceedings of the 40th Conf. on Decision and Contr.*, vol. 1, Dec 2001, pp. 304–309.
- [8] N. Elia, "Remote stabilization over fading channels," *Syst. and Contr. Lett.*, vol. 54(3), pp. 237–249, March 2005.
- [9] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.
- [10] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks." *Proceedings of the IEEE*, 2007.
- [11] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses." *IEEE Control and Decision*, 2004.
- [12] 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.
- [13] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "On lqg control across a stochastic packet-dropping link." *American Control Conference*, 2005.

- [14] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems." Proceedings of the IEEE, 2007.
- [15] L. Shi, M. Epstein, A. Tiwari, and R. Murray, "Estimation with information loss: Asymptotic analysis and error bounds," in *IEEE Conf. on Decision and Control*, Dec 2005, pp. 1215–1221.
- [16] M. Epstein, L. Shi, A. Tiwari, and R. M. Murray, "Probabilistic performance of state estimation across a lossy network," *Automatica*, 2008, To Appear.
- [17] F. Lian, J. Moyne, and D. Tillbury, "Performance evaluation of control networks," *IEEE Control Systems Magazine*, vol. 21, no. 1, pp. 66–83, 2001.
- [18] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, September 1960.
- [19] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, September 1963.
- [20] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. Oxford University Press, 2001.