Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Estimation with Information Loss: Asymptotic Analysis and Error Bounds

Ling Shi, Michael Epstein, Abhishek Tiwari and Richard M. Murray

Abstract— **In this paper, we consider a discrete time state estimation problem over a packet-based network. In each discrete time step, the measurement is sent to a Kalman filter with some probability that it is received or dropped. Previous pioneering work on Kalman filtering with intermittent observation losses shows that there exists a certain threshold of the packet dropping rate below which the estimator is stable in the expected sense. That work assumes that packets are dropped independently between all time steps. However we give a completely different point of view. On the one hand, it is not required that the packets are dropped independently but just that the information gain π_g, defined to be the limit of the ratio of the number of received packets n during N time steps as N goes to infinity, exists. On the other hand, we show that for any given π_g, as long as $\pi_g > 0$, the estimator is stable almost surely, *i.e.* for any given $\epsilon > 0$, the error covariance matrix P_k is bounded by a finite matrix M, with probability $1-\epsilon$. We also give explicit formula for the relationship between M and ϵ. We consider the case where the observation matrix is invertible.**

I. Introduction

Since the landmark paper by Kalman [1], the Kalman filter has been the subject of extensive research and applications [2], [3], [4], [5]. For example, the Kalman filter has been widely used in autonomous and assisted navigation. In his paper, Kalman showed a recursive solution to the discrete time state estimation problem. State estimation is key to the control community. If the estimate of the state is stable, the corresponding state feedback controller can be independently designed and the overall system can be made stable provided the original system is stabilizable. This separation principle of designing the state estimator and state feedback controller independently is one of the central theorems in modern control [6].

Traditionally the areas of control and communication networks are decoupled from each other as they have almost distinctly different underlying assumptions. For example, control engineers generally assume perfect information within the closed loop control and data processing is done with zero time delay. On the other hand, in communication networks, data packets that carry the information can be dropped, delayed or even reordered due to the network traffic conditions. These different assumptions have for a long time blocked researchers from the two fields from communicating with each other. However, as new applications keep emerging, the two fields are coming closer together. For instance, advances in large scale integration and microelectromechanical system technology have made sensor networks a hot area of research. In sensor networks, the measurement data from different sensors is sent to the controller through a data network where data packets might be dropped if the network has severe traffic.

In recent years, networked control problems have gained much interest. In particular, the state estimation problem over a network has been widely studied. 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 [7], [8] and further pursued by [9], [10], [11], [12]. In [13], Sinopoli et al. discussed how packet loss can affect stable state estimation. They showed there exists a certain threshold of the packet loss rate above which the state estimation diverges in the expected sense, *i.e.* 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. Following the spirit of [13], in [14], Liu and Goldsmith extended their idea to the case where there were multiple sensors and the packets arriving from different sensors were dropped independently. They provided similar bounds on the packet loss rate for a stable estimation, again in the expected sense.

In spite of the great progress that those previous researchers have made, the problems they have studied have certain limitations. For example, in both [13] and [14], they assumed that packets are dropped independently, which is certainly not true in the case where burst packets are dropped or in queuing networks where adjacent packets are not dropped independently. They also use the expected value of the error covariance matrix as the measure of performance. In our opinion this can conceal the fact that events with arbitrarily low probability can make the expected value diverge, and we should ignore such events with extremely low probability. For example, consider the simple unstable scalar system with $a = 2$ in [13]. Let the arrival rate $\lambda = 0.74 < 1 - 1/a^2$. According to [13], $E[P_k]$ is unbounded. This is easily verifiable by considering the event S where no packets are received in k time steps. Then $E[P_k] \geq \Pr[S]E[P_k|S] \geq (0.26^k)4^k P_0 = 1.04^k P_0$. By letting k go to infinity, we see that this event with almost zero probability makes the expected error diverge. We shall study this example in detail later on.

Inspired by the approach in [15], where Abate and et al. gave a stability criterion of a class of stochastic hybrid

Control and Dynamical Systems, California Institute of Technology; Pasadena, CA 91125.
Email:{shiling,epstein,atiwari,murray}@cds.caltech.edu; Tel: (626) 395-2313, Fax: (626) 796-8914.
The authors are all equally contributed to this work

0-7803-9568-9/05/$20.00 ©2005 IEEE

systems, we first define the information gain π_g as

$$\pi_g = \lim_{N\to\infty} \frac{n}{N}, \quad (1)$$

where n is the number of packets received in N time steps. Intuitively, this means that when N is sufficiently large, the total number of packets received in N time steps is $\pi_g N$ with arbitrarily high probability. This certainly includes the case where packet drops occur independently by weak law of large numbers [16]. In that case, π_g simply equals to the packet arrival rate. This also includes other network structures, for example the packets being received or dropped are described by a markov chain [16], which is frequently seen in a queuing network. In this latter case, the steady state probability of receiving the packets coincides with the one we defined above. We then argue that with this new notion of information gain π_g, as long as π_g exists and $\pi_g > 0$, the error covariance matrix P_k is bounded almost surely, *i.e.* for any given $\epsilon > 0$ the error covariance matrix P_k is bounded by a finite matrix M with probability $1 - \epsilon$.

The paper is organized as follows. In Section I, we briefly review relevant past work. In Section II, the mathematical model of our problem is given. In Section III, we give the main results in terms of a series of theorems and lemmas that prove the error covariance is bounded almost surely for any nonzero information gain. In Section IV, we give an explicit relationship between the bound and probability of the error covariance staying below the bound. In Section V we compare our metric with that of [13] for a specific example. The paper concludes with a summary of our results and a discussion of the work that lies ahead.

II. Problem set up

Let $\rho(A) = \max |\lambda_i(A)|$ be the spectral radius of a matrix $A \in M_n$, where M_n consists of all n by n matrices. Consider the following discrete-time LTI system

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k. \end{aligned} \quad (2)$$

We assume A is unstable, *i.e.* $\rho(A) > 1$, and C is invertible. We will discuss the general case for C not invertible in the future work. As a result, the pair (C, A) is observable. As usual, $x_k \in \mathbb{R}^n$ is the state vector, $y_k \in \mathbb{R}^m$ is the observation vector, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are Gaussian random vectors with zero mean and covariance matrices $Q \geq 0$ and $R > 0$, respectively. Assume w_s is independent from w_k and v_s is independent from v_k for $s \neq k$, and w_s is also independent from v_k for all s and k.

The problem of interest to us is to get a stable estimate for the state vector x_k at time k given all past measurement $y_0, \ldots, y_k$ sent through a packet based network, where $y_i, i \in \{0, \ldots, k\}$ can be dropped randomly. We make use of a Kalman filter to do the state estimation.

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. Let us also define the following quantities:

$$\begin{aligned} \hat{x}_{k|k} &= E[x_k|\mathbf{y_k}, \gamma_\mathbf{k}] \\ P_{k|k} &= E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)'|\mathbf{y_k}, \gamma_\mathbf{k}] \\ \hat{x}_{k+1|k} &= E[x_{k+1}|\mathbf{y_k}, \gamma_\mathbf{k}] \\ P_{k+1|k} &= E[(x_{k+1} - \hat{x}_{k+1})(x_{k+1} - \hat{x}_{k+1})'|\mathbf{y_k}, \gamma_\mathbf{k}] \\ \hat{y}_{k+1|k} &= E[y_{k+1}|\mathbf{y_k}, \gamma_\mathbf{k}], \end{aligned}$$

where $\mathbf{y_k} = [y_0, y_1, \cdots, y_k]'$ and $\gamma_\mathbf{k} = [\gamma_0, \gamma_1, \cdots, \gamma_k]'$.

Similar to Sinopoli et al. in [13], the Kalman filter update equations are given as:

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} \\ P_{k+1|k} &= AP_{k|k}A' + Q \\ \hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + \gamma_{k+1}K_{k+1}(y_{k+1} - C\hat{x}_{k+1|k}) \\ P_{k+1|k+1} &= P_{k+1|k} - \gamma_{k+1}K_{k+1}CP_{k+1|k}, \end{aligned}$$

where $K_{k+1} = P_{k+1|k}C'(CP_{k+1|k}C' + R)^{-1}$ is the Kalman gain matrix. As we are only interested in the error propagations, *i.e.* how the error covariance matrix evolves, we concentrate on its update equation, which turns out to be the following

$$P_{k+1} = AP_kA' + Q - \gamma_k AP_kC'[CP_kC' + R]^{-1}CP_kA' \quad (3)$$

where we simply write $P_k = P_{k|k-1}$. Intuitively this means that when there no packet arrives, the Kalman filter just performs the time update, and otherwise it performs both the time and measurement updates.

III. Asymptotic Properties of Error Covariance Matrix

As the simple example in the introduction shows, some events with almost zero probability can make the expected value of the error covariance diverge. In practice, those rare events are unlikely to happen and hence should be ignored. Therefore the expected value of the error covariance matrix may not be the best metric to evaluate the Kalman filter performance. By ignoring these low probability events, we hope that the error covariance matrix is stable with arbitrarily high probability. This is precisely captured in the following theorem.

Theorem 1: Let the information gain π_g be defined as in Eqn. (1) and assume it exists. If $\pi_g > 0$, then the error covariance matrix P_k is bounded above *almost surely*, *i.e.* for any given $\epsilon \in (0, 1)$, P_k is bounded by $M(\epsilon) < \infty$ with probability $1 - \epsilon$. Furthermore, $\inf(M)$ depends on the choice of ϵ and the smaller the ϵ, the bigger the $\inf(M)$.

Before we prove the theorem, we briefly compare this with Sinopoli et al. [13]:

1) We do not assume that the packet drops occur independently. This allows more network structures to be considered, for example, a queuing network where the packet drops are described by a Markov chain.

2) Our result shows that such upper bound M always exists, provided $\pi_g > 0$, and only depends on π_g, A, C, R and ϵ. In their paper, the M depends on the existence of the solution of another set of iterative equations.

To prove this theorem, some other propositions and lemmas are needed. For the remainder of the paper, let

$$\begin{aligned} f(X) &= AXA' - AXC'[CXC' + R]^{-1}CXA' \\ g(X) &= f(X) + Q \\ h(X) &= AXA' + Q \end{aligned}$$

where $X = X' \geq 0$, and A, C, Q and R are defined in Section II. Then according to Eqn. 3 the error covariance propagates with g if a packet is received and h if not, *i.e.*

$$P_{k+1} = \begin{cases} h(P_k) & \text{if } \gamma_k = 0 \\ g(P_k) & \text{if } \gamma_k = 1 \end{cases} \tag{4}$$

Proposition 1: Let $\lambda_h(X) = \frac{\mathrm{Tr}(h(X))}{\mathrm{Tr}(X)}$, then

$$\lambda_h(X) \leq 1 + \lambda_n(A'A)$$

for all $X > 0$ such that $\mathrm{Tr}(X) \geq \mathrm{Tr}(Q)$, where $\lambda_n(A'A)$ denotes the largest eigenvalue of $A'A$.

Proof:

$$\begin{aligned} \lambda_h(X) &= \frac{\mathrm{Tr}(AXA')}{\mathrm{Tr}(X)} + \frac{\mathrm{Tr}(Q)}{\mathrm{Tr}(X)} \\ &\leq 1 + \frac{\mathrm{Tr}(AXA')}{\mathrm{Tr}(X)} \\ &= 1 + \frac{\mathrm{Tr}(A'AX)}{\mathrm{Tr}(X)} \\ &= 1 + \frac{\mathrm{Tr}(P'A'APP'XP)}{\mathrm{Tr}(P'XP)} \\ &= 1 + \frac{\mathrm{Tr}(SY)}{\mathrm{Tr}(Y)}, \end{aligned}$$

where $S = P'A'AP$ is diagonal and $Y = P'XP > 0$ and has the same eigenvalues as X. Such P exists and $P' = P^{-1}$, as $A'A$ is real symmetric. Hence,

$$\begin{aligned} \lambda_h(X) &\leq 1 + \frac{\mathrm{Tr}(SY)}{\mathrm{Tr}(Y)} \\ &= 1 + \frac{\sum_{i=1}^n \lambda_i(A'A)Y_{ii}}{\sum_{i=1}^n Y_{ii}} \\ &\leq 1 + \frac{\lambda_n(A'A)\sum_{i=1}^n Y_{ii}}{\sum_{i=1}^n Y_{ii}} \\ &= 1 + \lambda_n(A'A). \end{aligned}$$

Notice that we implicitly used the fact that $Y_{ii} > 0$ for all i, this follows as

$$Y_{ii} = e_i^{'} Y e_i > 0.$$

■

Lemma 1: If $X \geq Y$, then $f(X) \geq f(Y)$ and $h(X) \geq h(Y)$.

Proof: See [13] appendix A. ■

Lemma 2: (Weyl's Theorem) Let A and B be Hermitian, and let the eigenvalues $\lambda_i(A)$, $\lambda_i(B)$, and $\lambda_i(A+B)$ be arranged in increasing order. For each $k = 1, 2, ..., n$ we have

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B) .$$

Proof: See [17], page 181. ■

Lemma 3: If $A > 0$ and $B \geq 0$, then there exists a scalar $t \geq 0$ such that $tA \geq B$.

Proof: Let $t \geq 0$. By Lemma 2,

$$\begin{aligned} \lambda_1(tA - B) &\geq \lambda_1(tA) + \lambda_1(-B) \\ &= t\lambda_1(A) - \lambda_n(B) . \end{aligned}$$

So any $t \geq \frac{\lambda_n(B)}{\lambda_1(A)}$ proves the lemma. ■

Lemma 4: For any scalar $t \geq 0$,

$$f(tC^{-1}RC'^{-1}) \leq AC^{-1}RC'^{-1}A' .$$

Proof:

$$\begin{aligned} f(tC^{-1}RC'^{-1}) &= \frac{t}{t+1}AC^{-1}RC'^{-1}A' \\ &\leq AC^{-1}RC'^{-1}A'. \end{aligned}$$

■

Lemma 5: For all $X \geq 0$,

$$0 \leq f(X) \leq AC^{-1}RC'^{-1}A' .$$

Proof: Clearly $f(X) \geq f(0) = 0$. For any $X \geq 0$, as $C^{-1}RC'^{-1} > 0$, by Lemma 3, there exists $t \geq 0$ such that $tC^{-1}RC'^{-1} \geq X$, and hence by Lemma 1 $f(tC^{-1}RC'^{-1}) \geq f(X)$. Therefore by Lemma 4

$$f(X) \leq f(tC^{-1}RC'^{-1}) \leq AC^{-1}RC'^{-1}A' .$$

■

Proposition 2: Let $\lambda_g(X) = \frac{\mathrm{Tr}(g(X))}{\mathrm{Tr}(X)}$, then

$$\lim_{\mathrm{Tr}(X)\to\infty} \lambda_g(X) = 0 .$$

Proof:

$$\begin{aligned} \lim_{\mathrm{Tr}(X)\to\infty} \lambda_g(X) &= \lim_{\mathrm{Tr}(X)\to\infty} \left\{ \frac{\mathrm{Tr}(f(X))}{\mathrm{Tr}(X)} + \frac{\mathrm{Tr}(Q)}{\mathrm{Tr}(X)} \right\} \\ &= \lim_{\mathrm{Tr}(X)\to\infty} \frac{\mathrm{Tr}(f(X))}{\mathrm{Tr}(X)} \\ &\leq \lim_{\mathrm{Tr}(X)\to\infty} \frac{\mathrm{Tr}(AC^{-1}RC^{'-1}A')}{\mathrm{Tr}(X)} \\ &= 0. \end{aligned}$$

■

Now we are ready to prove Theorem 1.

Proof of Theorem 1:

Recall that the error covariance matrix update equation is

$$P_{k+1} = h(P_k),$$

when the packet at time k is dropped, and

$$P_{k+1} = g(P_k),$$

when the packet at time k is received. The functions $h(X)$ and $g(X)$ are defined above. Let $\lambda_h = 1 + \lambda_n(A'A)$, *i.e.* λ_h is the upper bound of the ratio between the trace of $h(X)$ and X obtained in Proposition 1. Let $\pi_g > 0$ be the information gain. Let $\pi_h = 1 - \pi_g$ denote the information loss coefficient. Solve the following equation

$$\lambda_g^{\pi_g} \lambda_h^{\pi_h} = K < 1 \tag{5}$$

involving variable λ_g for some $K \in (0, 1)$. Let the solution to this to be δ, *i.e.*

$$\delta = \exp\left(\frac{\ln K - \pi_h \ln \lambda_h}{\pi_g}\right).$$

From Proposition 2, as the limit of $\lambda_g(X)$ is zero when $\mathrm{Tr}(X)$ tends to infinity, then there exist $M_1 < \infty$ such that $\mathrm{Tr}(X) > M_1$ implies $\lambda_g(X) < \delta$. In fact M_1 can be simply chosen as

$$M_1 = \frac{1}{\delta}\mathrm{Tr}(AC^{-1}RC^{'-1}A').$$

Let $N > 0$ be such that the number of packets received during the N time steps equals to $\pi_g N$ almost surely, *i.e.* for any given $\epsilon > 0$, that is true with probability $1 - \epsilon$. As a result, the number of packets dropped equals to $\pi_h N$ almost surely. Let $M = \lambda_h^N M_1 = M(\epsilon)$. We argue that $\mathrm{Tr}(P_k)$ will be bounded by M for $k > k_0$ almost surely. Clearly by choosing a smaller ϵ, we need to pick up a larger N, hence M is also increased.

Without loss of generality, let us suppose that at time k_0, $\mathrm{Tr}(P_{k_0}) > M_1$. Notice that during the time steps from $k = k_0$ to k_N, there exist at least a k_n where $0 \le n \le N$ such that $\mathrm{Tr}(P_{k_n}) \le M_1$. Otherwise, if $\mathrm{Tr}(P_{k_n}) > M_1$ for all $0 \le n \le N$, then $\lambda_g(P_{k_n}) \le \delta$ for all $0 \le k \le N$, and hence the following inequalities are obtained

$$\begin{aligned}\mathrm{Tr}(P_{k_N}) &\le \delta^{\pi_g N}\lambda_h^{\pi_h N}\mathrm{Tr}(P_{k_0}) \\ &= [\delta^{\pi_g}\lambda_h^{\pi_h}]^N \mathrm{Tr}(P_{k_0}) \\ &\le K^N \mathrm{Tr}(P_{k_0}),\end{aligned}$$

which tends to zero and hence produces a contradiction. Now let $k_0 = \inf\{k : k > k_n, \mathrm{Tr}(P_k) > M_1\}$ and repeat the same argument again. Notice that during the N time steps, $\mathrm{Tr}(X)$ shall be at most $\lambda_h^N M_1$ which corresponds to the event that no packets are received during the N time steps which occurs with almost zero probability.

Lastly, notice that for $X > 0$, if $\mathrm{Tr}(X) \le \alpha$, then $X \le \alpha I$ where I is the identity matrix. This is because $\lambda_i(\alpha I - X) = \alpha - \lambda_i(X)$, which is positive. In other words, if the trace of a positive definite matrix is bounded, the matrix itself is also bounded. This completes the proof. ∎

Remark 1: The upper bound M in the theorem is obviously a conservative bound as we see in the proof that we have considered an event with almost zero probability. As we show in Section IV, M is usually a much smaller value than presented in the theorem. The key point, however, is that by choosing M to be sufficiently large yet finite, we are almost sure that the error covariance matrix is bounded above by M.

IV. Determining the M-ϵ Relationship

We wish to derive the relationship between the M and ϵ given in Theorem 1. That is we seek to determine the probability, $1-\epsilon$, that the error covariance P_k will be below the bound M at any time. More formally $\Pr[P_k \le M] = 1 - \epsilon$. We present such a relationship below. We consider unstable A and invertible C.

Denote the solution to the algebraic Riccati equation as $\bar{P}$ and the upper bound on $g(X)$ as $\bar{M}$, namely

$$g(\bar{P}) = \bar{P} \tag{6}$$

$$g(X) \le \bar{M},\ \forall X \ge 0 \tag{7}$$

where from Lemma 5 we know $\bar{M} = AC^{-1}RC'^{-1}A' + Q$ and $\bar{P} \le \bar{M}$. Then $\epsilon_{k_i}(k)$ is defined to be the probability that at least the previous k_i consecutive packets are dropped at time k, *i.e.*

$$\epsilon_{k_i}(k) = \Pr[N_k \ge k_i]\ , \tag{8}$$

with N_k the number of consecutive packets dropped at time k. Clearly $\epsilon_{k_i} \ge \epsilon_{k_j}$ for $k_i \le k_j$. Next define the following quantities

$$k_1 \triangleq \min\{k \in Z^+ : h^k(\bar{M}) \nleq M\} \tag{9}$$

$$k_2 \triangleq \min\{k \in Z^+ : h^k(\bar{P}) \ge M\}\ , \tag{10}$$

where $h^k(X)$ means the operator h is applied k times to X.

Lemma 6: For unstable A, we have $h(\bar{P}) > \bar{P}$ and if A is purely unstable, then $h(X) \ge X$ for all $X \ge 0$

Proof: To prove the first statement, using Eqn. 6 we write

$$\begin{aligned}\bar{P} &= g(\bar{P}) \\ &= A\bar{P}A' + Q - A\bar{P}C'(C\bar{P}C' + R)^{-1}C\bar{P}A' \\ &< A\bar{P}A' + Q \\ &= h(\bar{P})\ .\end{aligned}$$

To prove the second statement note that since A is strictly unstable all of the eigenvalues lie outside the unit circle, so it has no zero eigenvalues and A^{-1} exists. Moreover the eigenvalues of A^{-1} all lie inside the unit circle and are nonzero. Then using the discrete Lyapunov equation for $X \ge 0$ we can write $A^{-1}XA'^{-1} \le X$ since A^{-1} is stable. Then multiply on the left by A and on the right by A' to get $X \le AXA'$ from which it is obvious $h(X) = AXA' + Q \ge X$ since $Q \ge 0$. ∎

Lemma 7: For finite M, the quantity k_1 will always exist while k_2 is guaranteed to exist if A is purely unstable.

Proof: To prove the existence of k_1 note that Lemma 4 in [18] says that for any $X > 0$, $\lim_{k\to\infty} \mathrm{Tr}(h^k(X)) = \infty$ if A is unstable. Thus for any scalar $t > 0$ there exists a k_1 such that $h^{k_1}(\bar{M}) \nleq tI$ and t can be chosen such that $tI \ge M$. This means $\lambda_n(h^{k_1}(\bar{M})) > t$ and $\lambda_n(M) < t$, where λ_n is the maximum eigenvalue. Then using Lemma 2 we see $\lambda_n(h^{k_1}(\bar{M}) - M) \geq \lambda_n(h^{k_1}(\bar{M})) - \lambda_n(M) > 0$ which implies $h^{k_1}(\bar{M}) \nleq M$.

If A is purely unstable then $\lim_{k\to\infty} \lambda_1(h^k(X)) = \infty$. Thus we can again pick any finite scalar $t > 0$ such that $tI \geq M$ and find a k_2 such that $h^{k_2}(\bar{P}) \geq tI \geq M$. ■

Lemma 8: With the definitions above, if they both exist then $k_1 \leq k_2$.

Proof: This can easily be shown by contradiction. Assume $k_1 > k_2$. We know $\bar{P} \leq \bar{M}$, and since $X \geq Y \Rightarrow h(X) \geq h(Y)$ it is easy to see $h^{k_2}(\bar{P}) \leq h^{k_2}(\bar{M}) \leq M$. From the definition of k_2, however, we see $h^{k_2}(\bar{P}) \geq M$ which is a contradiction of the previous inequality. Hence it must be true that $k_1 \leq k_2$. ■

Theorem 2: For unstable A and invertible C, assume the initial error covariance matrix P_0 is given by $\bar{P} \leq P_0 \leq \bar{M}$. Given a matrix bound $M \geq \bar{M}$ then ϵ in the expression $\Pr[P_k \leq M] = 1 - \epsilon$ is bounded by

$$\epsilon \leq \epsilon_{k_1}(k) \; . \tag{11}$$

That is the probability only depends on the number of consecutive packets dropped at the current time and is independent of the packet drop/receive sequence prior to the previous received packet.

Proof: Since $\bar{P} \leq P_0 \leq \bar{M}$, then assuming the next k packets are dropped we have $P_k = h^k(P_0)$ and it is clear from Lemma 1 that

$$h^k(\bar{P}) \leq P_k \leq h^k(\bar{M}) \; .$$

So the necessary condition that $P_k \nleq M$ is

$$h^k(\bar{M}) \nleq M \; ,$$

but this will only hold for $k \geq k_1$. Thus for $P_k \nleq M$ it is necessary to drop at least the previous k_1 consecutive packets.

Now assume a packet is not received until time $m > k_1$, that is $\gamma_k = 0$ for $k = 0, \cdots, m-1$ and $\gamma_m = 1$, then $P_{m+1} = g(P_m) \leq \bar{M}$ from Eqn. (7). It is also true that $P_{m+1} \geq \bar{P}$ since, from the concavity of g, $g(X) \geq \bar{P} \; \forall X \geq \bar{P}$. Thus for a packet received at time m, we have

$$\bar{P} \leq P_{m+1} \leq \bar{M} \; . \tag{12}$$

Regardless of how large m is, *i.e.* how long between packet receives, and how large the error covariance gets, Eqn. (12) holds. Hence the analysis above can always be repeated with P_{m+1} replacing P_0, and the probability $P_k \nleq M$ depends only on the number of consecutive packets dropped and is independent of what happens prior to the last packet received. ■

Corollary 1: If A is purely unstable the a lower bound on ϵ is given by

$$\epsilon_{k_2}(k) \leq \epsilon \; . \tag{13}$$

Proof: Following the proof of Theorem 2, assume the first k packets are dropped so $P_k = h^k(P_0)$. A sufficient condition for $P_k \nleq M$ is then

$$h^k(\bar{P}) \geq M \; ,$$

which will only hold for $k \geq k_2$. Thus dropping the previous k_2 consecutive packets guarantees $P_k \nleq M$. Now assume a packet is not received until time $m > k_2$, then we know $P_m = h^m(P_0) \geq h^m(\bar{P}) \geq M$ and $\bar{P} \leq P_{m+1} \leq \bar{M}$ so the analysis is repeated with P_{m+1} replacing P_0 as before. ■

The following example can help visualize the concepts of the theorem.

Example 1: Consider the scalar system $A = 1.3$, $C = 1$, $Q = 0.5$ and $R = 1$. For this system we have $\bar{P} = 1.519$ and $\bar{M} = 2.19$. Picking $M = 6.25$ it is easy to show $k_1 = 2$ and $k_2 = 3$. Thus there exists an $\bar{P} < X^* \leq M$ such that all for $\bar{P} \leq X < X^*$ it requires 3 consecutive packets to be dropped before the error covariance is greater than M, while for the region $X^* \leq X \leq \bar{M}$ it only requires 2 consecutive packets be dropped. In fact it can be easily shown that $X^* = 1.7174$.

Figure 1 shows the evolution of the error covariance for a particular sequence of packet acceptance/rejection. The sequence used is $hhhhggghhghhhgh(P_0)$. As can be seen, it requires at least 2 consecutive packets be dropped for the error covariance to rise above the bound.

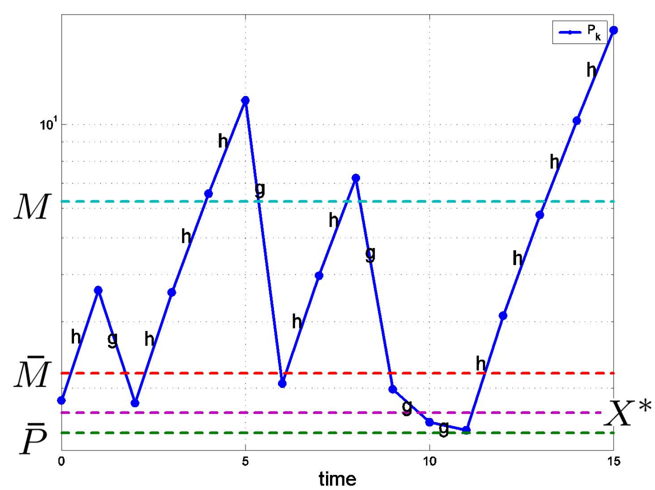

Fig. 1. Error covariance (log scale) for Example 1. For this system it will take at least 2 and no more than 3 consecutive dropped packets for $P_k \nleq M$.

Remark 2: Theorem 2 provides a different metric than the previous study [13] for assessing the performance of estimation across a network. Here it is the probability of bounding the error covariance that is important rather than the expected value of the error covariance. The method in this paper allows systems to be estimated that according to [13] could not be estimated. In addition, Theorem 2 is applicable to any packet dropping network while [13] is only for i.i.d networks.

Remark 3: With the definition of $\epsilon_{k_i}(k)$ as in Eqn. (8) it is easy to see $\epsilon_{k_i}(k) = 0$, $\forall k < k_i$. Which leads to $\Pr[P_k \leq M] = 1$, $\forall k < k_1$. So we only consider time greater than k_1.

We seek a method for calculating $\epsilon_{k_i}(k)$. Figure 2 shows all possible packet sequences at time k for a packet dropping network. From this it is clear to see that $\epsilon_{k_i}(k)$ will be the sum of the probabilities of each of the instances with at least the previous k_i packets dropped occurring.

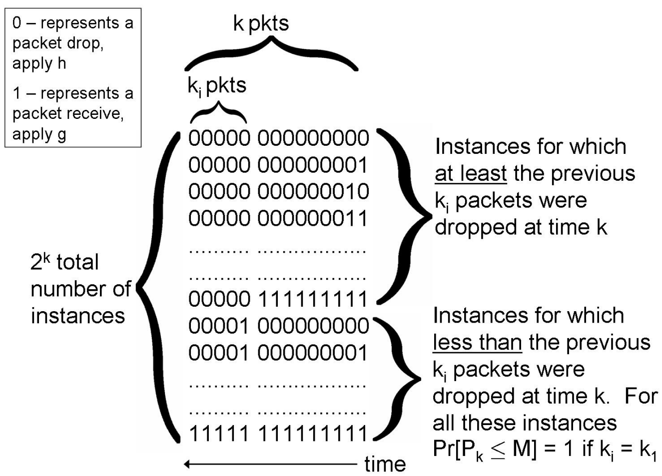

Fig. 2. A binary representation of the possible packet sequences (*i.e.* drop/ receive) at time k. A 0 signifies a packet was dropped and h was applied to the error covariance. A 1 signifies the packet was received and g was applied.

Corollary 2: For $k > k_i$ any packet dropping network that is either i.i.d or reaches a steady state (for example a Markov network), $\epsilon_{k_i}(k) = \epsilon_{k_i}$.

The above corollary says the probability of dropping at least the previous k_i packets is the same for all time. To calculate ϵ_{k_i} we can make use of the Markov chain model in Figure 3.

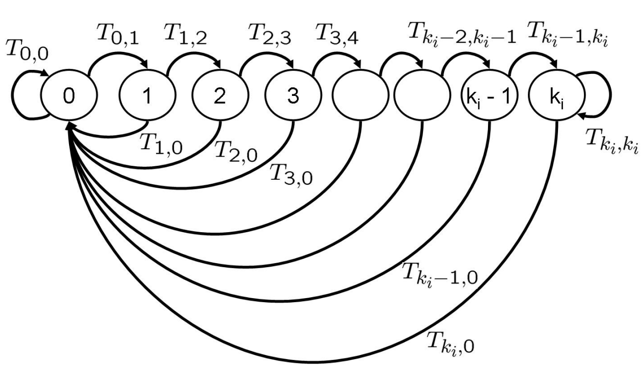

Fig. 3. Markov chain used to determine ϵ_{k_i}. The states of the Markov chain represent the number of consecutive packets dropped at the current time, the final state represents k_i or more consecutive packets dropped. The transition probability from state i to state j is given by $T_{i,j}$.

The states of the Markov chain represent the number of consecutive packets dropped at the current time, the final state represents k_i or more consecutive packets dropped. The transition probability from state i to state j is given by $T_{i,j}$. It is clear $\epsilon_{k_i} = \pi_{k_i}$, the steady state probability of the Markov chain being in state k_i. This is easily determined to be given by

$$\pi_{k_i} = \frac{D}{D + T_{k_i,0} + T_{k_i,0} \sum_{l=1}^{k_i-1} \prod_{j=0}^{l-1} T_{j,j+1}} \tag{14}$$

with

$$D = 1 - T_{0,0} - \sum_{l=1}^{k_i-1} T_{l,0} \prod_{j=0}^{l-1} T_{j,j+1} \; .$$

Note that π_{k_i} decreases as k_i increases.

The $T_{i,j}$ are determined based on the type of network. For example, an i.i.d network with packet arrival rate λ and drop rate $1-\lambda$ has $T_{j,0} = \lambda \; \forall j \geq 0$, $T_{j,j+1} = 1-\lambda \; \forall j \geq 0$, and $T_{k_i,k_i} = 1-\lambda$. This leads to $\pi_{k_i} = (1-\lambda)^{k_i}$. A first order Markov network with transition probabilities T_{hh}, T_{hg}, T_{gh}, and T_{gg} leads to $\pi_{k_i} = \frac{1-T_{gg}}{2-T_{hh}-T_{gg}}(T_{hh})^{k_i-1}$. The probability π_{k_i} for any arbitrary order Markov network can be determined in this manner.

Theorem 2 and Corollary 1 provide bounds on ϵ for a given M and the network properties, *i.e.* π_{k_1} and π_{k_2}. It is also possible to determine bounds on M and π_{k_i}.

Corollary 3: With the same assumptions as Theorem 2 and given the transition probabilities $T_{i,j}$ of the Markov model in Figure 3 and an upper bound ϵ_{max} it is possible to determine a suitable M such that $\Pr[P_k \leq M] \geq 1 - \epsilon_{max}$. To do so, define

$$k_M \triangleq \min\{k \in Z^+ : \pi_k \leq \epsilon_{max}\} \; , \tag{15}$$

with π_k given in Eqn. (14). Then the tightest such bound is

$$M = h^{k_M}(\bar{M}) \; . \tag{16}$$

Corollary 4: Likewise, given M and an upper bound ϵ_{max} it is possible to determine limits on the transition probabilities $T_{i,j}$ of the Markov model in Figure 3 such that $\Pr[P_k \leq M] \leq 1 - \epsilon_{max}$. With k_1 as defined in Eqn. (9), it is easy to see that we require

$$\pi_{k_1} \leq \epsilon_{max} \; . \tag{17}$$

For the i.i.d network this reduces to $\lambda \geq 1 - \epsilon_{max}{}^{\frac{1}{k_1}}$.

Remark 4: Theorem 2 is based on the assumption that $\bar{P} \leq P_0 \leq \bar{M}$. This assumption can be relaxed by slightly modifying the definition of ϵ. For arbitrary P_0 the definition of ϵ becomes $\Pr[P_k \leq M] = 1 - \epsilon \; \forall k \geq k^*$, where k^* is defined to be the first time instance such that $\bar{P} \leq P_{k^*} \leq \bar{M}$. That is for arbitrary P_0 we simply consider the probability for time after the first time instance the error covariance is between $\bar{P}$ and $\bar{M}$. If $P_0 \geq \bar{M}$ this will correspond to the time instance the first packet is received. If $P_0 \leq \bar{P}$ it is not as easy to determine the time, however the upper bound ϵ_{max} is still valid for all time.

V. Examples and Simulations

In this section we will compare our metric to that in [13] using a scalar example.

Example 2: Consider the scalar system given with $A = 2$, $C = 1$, $Q = 1$ and $R = 1$ and an i.i.d. network with packet acceptance rate λ. According to [13] the expected value of the error covariance will diverge for any $\lambda < \frac{3}{4}$. Assume we are given $\lambda = 0.74$, then according to the metric used in [13] this system cannot be estimated, as the expected value of the error goes unbounded. Using the analysis presented in this paper, however, we can predict with what probability the error will remain below certain bounds. Note that for this system $\bar{P} = 4.236$ and $\bar{M} = 5$.

Figure 4 shows the M - ϵ relationship for this system. The bound M was chosen to vary from $\bar{M}$ to 10,000 and the corresponding ϵ was determined. A total of 100 simulations were run for each value of M, and a random initial error covariance in the range $\bar{P} \le P_0 \le \bar{M}$ was chosen for each simulation. The simulations were run for 10,000 time steps and the ϵ calculated from the simulations corresponds to the average over all simulations of the percent of time the error covariance was larger than the M bound. The staircase like plot can be explained by the fact the probability bounds for ϵ are given by ϵ_{k_1} and ϵ_{k_2} which exhibit sharp jumps, *i.e.* the staircase, as k_1 and k_2 change integer values.

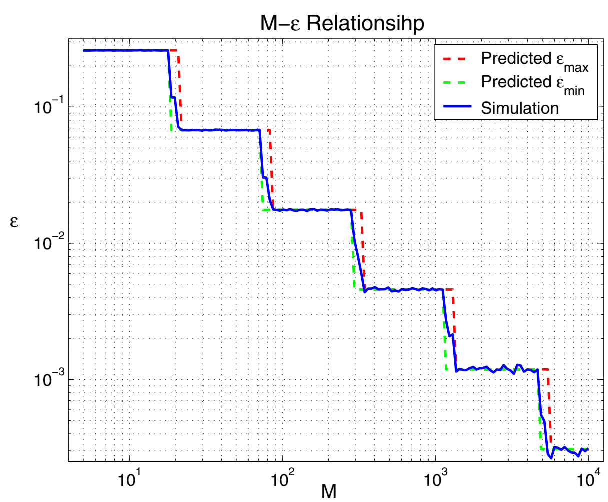

Fig. 4. M bound vs. ϵ for Example 2. The solid (blue) line is the simulated ϵ and the dashed (red and green) lines are the predicted ϵ_{max} and ϵ_{min}.

VI. Conclusions and Future Work

The contributions of this papers can be divided into two major parts. Firstly it gives results for the asymptotic behavior of the error covariance in a discrete estimation problem over a packet based network. In the asymptotic sense an upper bound to the error covariance can always be found as long as the information gain is not exactly equal to zero. This analysis is independent of the probability distribution of packet drops.

Secondly, for a given finite positive definite matrix M we give relations for upper and lower bounds on the probability $1 - \epsilon = \Pr[P_k \le M]$, where P_k is the error covariance at time k. We observe that $P_k \not\le M$ only if a large enough burst of packets are dropped before time k. The size of this burst is only dependent on M and not on the particular time instant k. This observation is not surprising as in most network applications, bursts of packet losses are responsible for failure.

In this paper we have assumed that the observation matrix is invertible. This assumption provides us with an upper matrix bound on the error covariance matrix recursion. Future work will focus on deriving similar results for a non-invertible observation matrix, but with an assumption that the pair (A, C) is detectable. The asymptotic analysis presented in this paper and the expected value based analysis of previous work, need to be compared for closed loop estimation and control problems. We are also looking at distributed and cooperative control problems over packet based networks.

References

[1] R. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASMEłJournal of Basic Engineering*, pp. 35–45, March 1960.

[2] M. S. Grewal and A. P. Andrews, *Kalman Filtering Theory and Practice*. Prentice Hall, 1993.

[3] P. S. Maybeck, *Stochastic Models, Estimation, and Control*. Academic Press, Inc, 1979, vol. 1.

[4] T.Kailath, A.Sayed, and B.Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[5] B.Anderson and J.Moore, *Optimal Filtering*. Prentice Hall, 1990.

[6] F. Callier and C. Desoer, *Linear System Theory*. Springer Verlag, 1991.

[7] W.S.Wong and R.W.Brochett, "Systems with finite communication bandwidth-part i: State estimation problems," *IEEE Trans. Automat. Contr.*, vol. 42, Sept 1997.

[8] ——, "Systems with finite communication bandwidth-part ii: Stabilization with limited information feedback," *IEEE Trans. Automat. Contr.*, vol. 44, May 1999.

[9] R.W.Brockett and D.Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Automat. Contr.*, vol. 45, July 2000.

[10] 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.

[11] S. C. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Massachusetts Institute of Technology, 2000.

[12] 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.

[13] 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.

[14] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses." IEEE Control and Decision, 2004.

[15] A.Abate, L.Shi, S.Simic, and S.Sastry, "A stability criterion on stochastic hybrid systems." Mathematical theory of Networks and Systems Conference, Leuwen, BG, 2004.

[16] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed. Addison-Wesley Pub Co, 1993.

[17] R. A.Horn and C. R.Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[18] M. Epstein, L. Shi, A. Tiwari, and R. Murray, "Estimation of linear stochastic systems over a queueing network." International Conference on Sensor Networks, 2005.