Automatica, 2008 (to appear)
http://www.cds.caltech.edu/~murray/papers/estm08-automatica.html

Probabilistic Performance of State Estimation Across a Lossy Network

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

Control and Dynamical Systems, Caltech, Mail Code 107-81, Pasadena, CA 91125
epstein,shiling,atiwari,murray@cds.caltech.edu

Abstract

We consider a discrete time state estimation problem over a packet-based network. In each discrete time step, a measurement packet is sent across a lossy network to an estimator unit consisting of a modified Kalman filter. Using the designed estimator algorithm, the importance of placing a measurement buffer at the sensor that allows transmission of the current and several previous measurements is shown. Previous pioneering work on Kalman filtering with intermittent observation losses is concerned with the asymptotic behavior of the expected value of the error covariance, *i.e.* $\mathbf{E}[P_k] < \infty$ as $k \to \infty$. We consider a different performance metric, namely a probabilistic statement of the error covariance $\mathbf{Pr}[P_k \leq M] \geq 1 - \epsilon$, meaning that with high probability the error covariance is bounded above at any instant in time. Provided the estimator error covariance has an upper bound whenever a measurement packet arrives, we show that for any finite M this statement will hold so long as the probability of receiving a measurement packet is nonzero. We also give an explicit relationship between M and ϵ and provide examples to illustrate the theory.

Key words: Kalman filtering; networked control.

1 Introduction

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 transmission of information within the closed loop and that 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. In the past there was no pressing need to relax these assumptions, however, as new applications emerge 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 an estimator 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 [12,13] and further pursued by [1,6,11,7]. In [10], Sinopoli et al. discussed how packet loss can affect state estimation. They showed there exists a certain threshold of the packet loss rate above which the state estimation error 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 [10], in [5], Liu and Goldsmith 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 loss rate for a stable estimate, again in the expected sense.

The problem of state estimation of a dynamical system where measurements are sent across a lossy 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 [10] and [5], 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. This can conceal the fact that events with arbitrarily low probability can make the expected value diverge, and it might be better to ignore such events that occur with extremely low probability. For example, consider the simple unstable scalar system with $a = 2$ in [10]. Let the arrival rate $\gamma = 0.74 < 1 - 1/a^2$. According to [10], the expected value of the estimation error covariance, $\mathbf{E}[P_k]$, is unbounded. This is easily verifiable by considering the event S where no packets are received in k time steps. Then $\mathbf{E}[P_k] \geq \mathbf{Pr}[S]\mathbf{E}[P_k|S] \geq (0.26^k)4^k P_0 = 1.04^k P_0$. By letting k go to infinity, we see that this event with almost zero probability makes the expected error diverge.

The goal of the present work is to give a more complete characterization of the estimator performance by instead considering a probabilistic description of the error covariance. We show it is bounded above by a given bound with a high probability, *i.e.*

$$\mathbf{Pr}[P_k \leq M] = 1 - \epsilon \ . \tag{1}$$

The importance of this characterization lies in the fact that while the expected value of P_k may diverge due to events with very low probability, in fact the actual value of P_k can be below an acceptable limit for a vast majority of the time. For this expression to hold, it requires an estimator that will have a finite upper bound whenever a measurement packet is received. We will construct such an estimator in this paper. We will also show how to determine the relationship between M and ϵ.

The rest of the paper is organized as follows. In Section 2, the mathematical model of our problem is given. The estimation algorithm that provides an upper bound to the error covariance whenever a measurement packet arrives is described in Section 3. In Section 4, we show that M exists for any given ϵ. In Section 5, we give an explicit relationship between the bound and probability of the error covariance staying below the bound. In Section 6 we compare our metric with that of [10] by means of an example. The paper concludes with a summary of our results and a discussion of the work that lies ahead.

2 Problem Set Up

2.1 Problem Setting

We consider estimating the state of a discrete-time LTI system

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

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 zero mean white Gaussian random vectors with $\mathbf{E}[w_k w_j{}'] = \delta_{kj} Q \geq 0$, $\mathbf{E}[v_k v_j{}'] = \delta_{kj} R > 0$, $\mathbf{E}[w_k v_j{}'] = 0 \ \forall j, k$. Where $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ otherwise. We will assume the pair (A, C) is observable and $(A, Q^{\frac{1}{2}})$ controllable and to make the estimation problem interesting that A is unstable.

We assume the sensor measurements y_k are to be sent across a lossy network, with negligible quantization effects, to the estimator. Thus the estimator will either receive a perfectly communicated measurement packet or none at all. It is assumed the network losses are random events. 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 measurements in a buffer. Therefore each packet sent through the network will 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[4]. 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 loss rates. Figure 1 shows a schematic of the system set up.

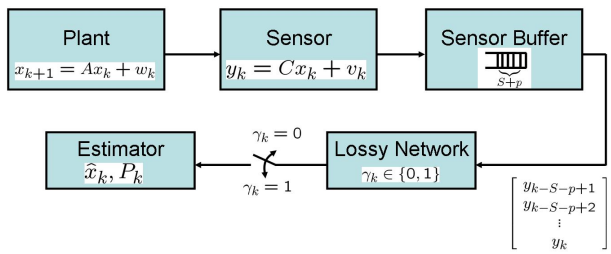

Fig. 1. A schematic diagram of the system set up we are considering. Note the measurement packets sent across the network consist of the previous $S + p$ measurements taken by the sensor.

2.2 Kalman Filtering Across a Lossy Network

Sinopoli et al. [10] showed that the Kalman filter is still the optimal estimator in this setting. There is a slight change to the standard Kalman filter in that only the time update is performed when the measurement packets are dropped. When a measurement is received the time and measurement update steps are performed. The filtering equations become

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} \tag{3}$$
$$P_{k+1|k} = AP_{k|k}A' + Q \tag{4}$$
$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \gamma_{k+1}K_{k+1}(y_{k+1} - C\hat{x}_{k+1|k}) \tag{5}$$
$$P_{k+1|k+1} = P_{k+1|k} - \gamma_{k+1}K_{k+1}CP_{k+1|k}, \tag{6}$$

where $'$ is the transpose operator, $\gamma_{k+1} \in \{0,1\}$ indicates if the measurement y_{k+1} was received and $K_{k+1} = P_{k+1|k}C'(CP_{k+1|k}C' + R)^{-1}$ is the Kalman gain matrix. Note Eqn. (3) -(4) are the Kalman Filter Time Update equations and Eqn. (5) - (6) are the Measurement Update equations and $P_{k+1|k}$ and $P_{k+1|k+1}$ are the a priori and a posteriori error covariances respectively.

Unlike the standard Kalman filtering setting where the error covariance matrix is a deterministic quantity (given an initial value), the randomness of the lossy network makes it a random variable as well. Nonetheless, its update equation can still be characterized as

$$P_{k+1} = AP_kA' + Q - \gamma_k AP_kC'[CP_kC' + R]^{-1}CP_kA' \tag{7}$$

where we simply write $P_k = P_{k|k-1}$. Given the system parameters A, C, Q, R, then for any positive semidefinite matrix $X \geq 0$ define the following functions

$$h(X) = AXA' + Q \tag{8}$$
$$g(X) = AXA' + Q - AXC'(CXC' + R)^{-1}CXA' \ . \tag{9}$$

From[10] we have that $X \geq Y \geq 0 \Rightarrow g(X) \geq g(Y)$ and $h(X) \geq h(Y)$. We will adopt the notation that $g^m(X)$ and $h^m(X)$ mean to apply the g and h functions m times starting from X with $g^0(X) = h^0(X) = X$. Note that Eqn. (9) is the discrete Algebraic Riccati Equation. We will denote the solution to this equation by $\overline{P} = g(\overline{P})$, which is also the steady state covariance if all measurements are received (*i.e.* $\lim_{k\to\infty} P_k = \overline{P}$ if $\gamma_k = 1 \ \forall k$ for any $P_0 \geq 0$).

For the case where γ_k is an independent and identically distributed random variable with mean γ, Sinopoli et al. [10] showed that there exists a critical value which determines the stability of the expected value of the estimation error covariance $\mathbf{E}[P_k]$ as $k \to \infty$. As mentioned in Section 1, we are interested in a different metric to evaluate the estimator performance,

$$\mathbf{Pr}[P_k \leq M] = 1 - \epsilon \ . \tag{10}$$

In [9] the present authors first introduced this notion for this same problem setting but under the additional assumption that the measurement matrix, C, is invertible. With C invertible the error covariance has an upper bound whenever the Kalman filter time and measurement updates are applied, *i.e.* whenever a measurement packet arrives. This was the key feature that allows the expression in Eqn. (10) to be evaluated.

With C not invertible, then given a single measurement update step no such upper bound on the Kalman filter error covariance can be determined. We seek an estimator algorithm that will provide an upper bound whenever a measurement packet arrives. As shown below, a suboptimal estimator can be constructed that uses a series of previous measurements but has a fixed upper bound. This estimator can be run in parallel with the Kalman filter that uses a single measurement update, switching to the suboptimal estimator when the error covariance of the Kalman filter is above this bound.

2.3 Observer Based Estimator

The observer based estimator described in this section will provide a state estimate by inverting out the known dynamics from a finite sequence of past measurements. Define

$$\mathcal{O}(r) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix} \tag{11}$$

for any positive integer $r \geq 1$. Next define S to be the smallest integer such that the matrix is rank n, *i.e.*

$$S = \min\{r \geq 1 : \text{rank}\,(\mathcal{O}(r)) = n\} \ . \tag{12}$$

Since (A, C) is observable, S is guaranteed to exist and $S \leq n$. Thus by concatenating the previous S consecutive measurements the augmented observation vector $\mathcal{O}(S)$ is full rank and hence has a pseudo-inverse. Now denote the pseudo-inverse by

$$\mathcal{C} = \left(\mathcal{O}(S)'\mathcal{O}(S)\right)^{-1}\mathcal{O}(S)' \ . \tag{13}$$

Then at time k, given the sequence of measurements $y_{k-S+1}, y_{k-S+2}, \ldots, y_k$, we can construct an estimate of the state according to

$$\overline{x}_k = A^{S-1}\mathcal{C}\left[\, y_{k-S+1}' \ y_{k-S+2}' \ \cdots \ y_k' \,\right]' \ , \tag{14}$$

and define the estimation error as $\overline{e}_k = x_k - \overline{x}_k$.

Lemma 1 *The a posteriori covariance,* $\overline{P}_{k|k} = \mathbf{E}[\overline{e}_k\overline{e}_k{}']$ *is*

$$\overline{P}_{k|k} = \widetilde{Q} + A^{S-1}\mathcal{C}\widetilde{R}\mathcal{C}'A^{S-1\prime} - A^{S-1}\mathcal{C}\widetilde{T} - \widetilde{T}'\mathcal{C}'A^{S-1\prime} \tag{15}$$

where

$$\widetilde{R} = diag\left(\widetilde{R}_S\right) + U_S + U_S{}' \tag{16}$$
$$\widetilde{Q} = \widetilde{Q}_S \tag{17}$$
$$\widetilde{T} = \left[0_n^m, \widetilde{T}_2, \widetilde{T}_3, \cdots, \widetilde{T}_S\right] \ , \tag{18}$$

with

$$\widetilde{R}_S = \left[R, R + C\widetilde{Q}_2 C', R + C\widetilde{Q}_3 C', \cdots, R + C\widetilde{Q}_S C'\right]$$

$$\widetilde{Q}_i = \sum_{j=0}^{i-2} A^j Q A^{j'} \ , \ \textit{for } i = 2, \ldots, S$$

$$U_S = \begin{cases} \left[0_m^{S\cdot m'}, u_2', u_3', \cdots, u_{S-1}', 0_m^{S\cdot m'}\right]' & , \ \textit{if } S \geq 2 \\ 0_m^{S\cdot m} & , \ \textit{if } S = 1 \end{cases}$$

$$u_i = \left[0_m^{i\cdot m}, C\widetilde{Q}_i A' C', C\widetilde{Q}_i A^{2'} C', \cdots, C\widetilde{Q}_i A^{S-i'} C'\right]$$

$$\widetilde{T}_i = \sum_{j=S-i}^{S-2} A^j Q A^{j-S+i'} C' \ , \ \textit{for } i = 2, \ldots, S \ .$$

The term 0_i^j is used to represent a matrix with i rows and j columns whose elements are all identically zero.

Proof: Note that for $j \geq k - S + 1$ we can write

$$y_j = C\left(A^{j-k+S-1} x_{k-S+1} + \sum_{i=0}^{j-k+S-2} A^i w_{j-i-1}\right) + v_j \ .$$

The term in the parenthesis is x_j, and note we have separated the expression in terms of dependence on the state at time $k - S + 1$ and the noise sequence from $k - S + 1$ to $j - 1$. We can then write the estimator in Eqn. (14) in terms of x_{k-S+1} by using these expressions for the measurement signals.

$$\begin{aligned} \overline{x}_k = & \\ & A^{S-1}\mathcal{C} \begin{bmatrix} Cx_{k-S+1} + v_{k-S+1} \\ C(Ax_{k-S+1} + w_{k-S+1}) + v_{k-S+2} \\ \vdots \\ C\left(A^{S-1}x_{k-S+1} + \sum_{i=0}^{j-k+S-2} A^i w_{j-i-1}\right) + v_k \end{bmatrix} \\ & = A^{S-1}\ \mathcal{C}\ \mathcal{O}(S)\ x_{k-S+1} \\ & \quad + A^{S-1}\mathcal{C} \begin{bmatrix} v_{k-S+1} \\ Cw_{k-S+1} + v_{k-S+2} \\ \vdots \\ C\left(\sum_{i=0}^{j-k+S-2} A^i w_{j-i-1}\right) + v_k \end{bmatrix} \end{aligned}$$

Since $x_k = A^{S-1}x_{k-S+1} + \sum_{i=0}^{S-2} A^i w_{k-i-1}$, the estimation error for this estimator can then be easily seen to be

$$\overline{e}_k = \sum_{i=0}^{S-2} A^i w_{k-i-1} - A^{S-1}\mathcal{C} \begin{bmatrix} v_{k-S+1} \\ Cw_{k-S+1} + v_{k-S+2} \\ \vdots \\ C\left(\sum_{i=0}^{S-2} A^i w_{k-i-1}\right) + v_k \end{bmatrix} .$$

The error covariance $\overline{P}_{k|k} = \mathbf{E}[\overline{e}_k \overline{e}_k{}']$ is then found by making use of the standard assumptions on the covariances of the process and sensor noise terms, resulting in Eqn. (15). □

Remark 2 *Note that $\overline{P}_{k|k}$ in Eqn. (15) independent of $\overline{P}_{k-1|k-1}$. In fact, it is a fixed quantity which depends only on A, C, Q, R and S. So whenever S consecutive measurements are available the error covariance using this observer based estimator has an upper bound.*

Remark 3 *To assure that this upper bound exists whenever a measurement packet is received simply requires the sensor transmit the previous S measurements at each time step.*

Since Eqn. (7) gives the a priori covariance update for the Kalman filter, we will likewise be concerned with the a priori covariance of this estimator. Denote the a priori covariance by $\bar{S} = \overline{P}_{k+1|k} = \mathbf{E}[\overline{e}_{k+1}\overline{e}_{k+1}{}'|y_{k-S+1}, \cdots, y_k]$, then we have

$$\overline{S} = A\overline{P}_{k|k}A' + Q \ . \tag{19}$$

3 Estimator Algorithm

We saw in the previous section that the error covariance of the observer based estimator is a fixed value. Thus if the previous S measurements are transmitted to the estimator, it guarantees that using the observer based estimator yields $P_k \leq \overline{S}$ at each time step the measurement packet is received. As to be expected, inverting the dynamics to obtain the state estimate could cause $\overline{S}$ to be quite large. If we include an additional p measurements in the buffer, we can run the observer based algorithm using the first S measurements and then a Kalman filter using the additional p measurements to decrease the covariance upper bound after a measurement packet is received.

The idea is to include a total of $S + p$ measurements in the buffer. The first S measurements, $\{y_{k-S-p+1}, y_{k-S-p+1}, \ldots, y_{k-p}\}$, are used to construct the estimate $\overline{x}_{k-p}$ according to Eqn. (14). The remaining measurements, $\{y_{k-p+1}, \ldots, y_k\}$, are then used in

running the Kalman filter time and measurement updates (Eqn. (3) - Eqn. (6) with $\gamma_{k+1} = 1$ since all the necessary measurements are included in the measurement packet) which are initialized with $\hat{x}_{k-p} = \overline{x}_{k-p}$ and $P_{k-p} = \overline{S}$. After running a total of p Kalman filter time and measurement updates we will have an estimate $\hat{x}_k$ whose error covariance will be

$$\overline{M} = g^p(\overline{S}) \ . \tag{20}$$

This will provide a smaller (than $\overline{S}$) upper bound on the error covariance that will hold whenever a measurement packet is received, *i.e.*

$$P_{k+1} \leq \overline{M}, \text{ if } \gamma_k = 1 \ . \tag{21}$$

Note that since the covariance of the observer based estimator is a fixed quantity, $\overline{S}$, the subsequent p Kalman gains can be computed off-line and stored in advance. We will call the estimator just described an observer based estimator with Kalman filter extension.

The estimator algorithm consists of running both the modified Kalman filter algorithm and the observer based estimator with the Kalman filter extension (as described above) along with some logic to choose the estimate with the lower covariance. When no measurement packet is received the estimator algorithm simply performs the time update steps according to Eqn. (3) - (4) using the previous estimate and covariance. If a measurement packet is received the measurement update steps of the Kalman filter are run Eqn. (5) - (6) using only the most recent measurement (which is from the current time-step y_k) from the measurement packet. The computed covariance is checked against $\overline{M}$. If the Kalman filter covariance is not less than this $\overline{M}$, then the observer based estimator with Kalman filter extension is run and the estimate and covariance are set to these computed values. This estimation algorithm will assure an upper bound on the error covariance $\overline{M}$ always exists whenever a packet is received. The algorithm is described in Table 1, that it consists of two distinct estimators: (i) the Kalman filter using only the measurement from the current time-step and (ii) the observer based estimator with Kalman filter extension using the sequence of previous measurements.

Remark 4 *It would be possible to use all the measurements in the packet with the Kalman filter by storing the previous estimate and covariance from time-step $k-S-p$ and then recomputing the Kalman filter time and measurement update steps from time-step $k-S-p+1$ to k once a packet arrives. This would make use of any lost information that eventually arrives at the filter, and as shown in [2] the resulting covariance after applying S time and measurement updates has an upper bound that is equivalent to the upper bound of the observer based estimator under the extra condition that $\mathcal{O}(S)$ is square. The disadvantage of this approach is that the S Kalman gains would need to be computed every time-step that a packet arrives (actually $S+p$ if we then use the extra p measurements to further reduce the covariance), this would be computationally burdensome. As noted above, the observer based estimator with Kalman filter extension does not suffer from this computational burden since the gains can be computed off-line. Furthermore, recomputing the Kalman filter estimate with the older measurements in the packet is only necessary if those measurements were never received. It will not improve the estimate if the measurement update for that corresponding time-step was already computed, i.e. if it is not providing new information, so it would not unnecessary to constantly utilize the old measurements to recompute the Kalman filter estimate.*

Remark 5 *From the description of the algorithm and the remark above, one can see that when the estimation algorithm is implemented, the Kalman filter that uses the current measurement only will be selected during a sequence of packet receives. When a packet is received after a long enough string of drops, however, the algorithm will utilize the older measurements that had not yet been received by choosing the observer based estimator with Kalman filter extension, ultimately switching back to the Kalman filter once a sequence of receives begins again.*

4 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 estimator 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 6 *Assume the packet arrival sequences are i.i.d. Let π_g be the expected value of the packet arrival rate. If $\pi_g > 0$, then for any $0 < \epsilon < 1$, there exists $M(\epsilon) < \infty$ such that the error covariance matrix P_k is bounded by M with probability $1-\epsilon$.*

Though we assume here that the packet drops occur independently, it is shown later when we determine the relationship between M and ϵ, the condition can be relaxed to include the case where the packet drops are described by an underlying markov chain.

The theorem also suggests that for a given error tolerance $M > 0$, we can find $\min(\pi_g)$ such that the error covariance matrix P_k is bounded by M with any given

Table 1
Algorithm for estimation scheme.

0) Given A, C, Q, R ;
- Determine $\overline{S}$ and $\overline{P}$;
- Choose the number of additional measurements, p, to buffer and transmit so that $\overline{M} = g^p(\overline{S})$ is as close to $\overline{P}$ as desired and so that $P_k \leq \overline{M}$ will hold whenever a packet is received ;
- Initialize $\hat{x}_0$ and P_0 ;

1) Wait for packet at time k ;
- Kalman Filter Time Update ;
- If packet received at time k ;
  - Kalman Filter Measurement Update ;
  - If $P_k \not\leq \overline{M}$;
    - Compute $\overline{x}_{k-p}$ using Eqn. (15) ;
    - Set $\hat{x}_{k-p} \leftarrow \overline{x}_{k-p}$ and $P_{k-p} \leftarrow \overline{S}$;
    - Loop $j = 1$ to p ;
      - Kalman Filter Time and Measurement Updates using measurement y_{k-p+j} ;
    - EndLoop ;
  - EndIf ;
  - $k \leftarrow k + 1$;
- EndIf ;
- Goto 1 ;

specified probability .

Before we prove the theorem, we introduce the following proposition.

Proposition 7 *Define* $\lambda_h(X) = \frac{\mathrm{Tr}(h(X))}{\mathrm{Tr}(X)}$ *. Then,*

$$\lambda_h(X) \leq 1 + \lambda_n(A'A) \triangleq \bar{\lambda}_n$$

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

$$Y_{ii} = e_i{}'Ye_i > 0 \,. \qquad \square$$

Now we are ready to prove Theorem 6.

Proof of Theorem 6: Without loss of generality assume at time k the packet is not received, $\gamma_k = 0$, otherwise $P_k \leq \overline{M} \triangleq M(\epsilon)$ for any ϵ. Define $\pi_h = 1 - \pi_g$ and let $k' = \max\{s : s \leq k, \gamma_s = 1\}$. Then $k - k' = N$ with probability $\pi_g \pi_h^N$. Further define $M_0 = \mathrm{Tr}(P_0)$, $M_1 = \mathrm{Tr}(\overline{M})$ and $\alpha_N = \bar{\lambda}_h^N$. We discuss two cases for a given ϵ.

(1) $0 < \epsilon \leq \pi_g$
Solve the following equation for N

$$\pi_g \pi_h^N = \epsilon$$

to get

$$N = \left\lceil \frac{\log \epsilon - \log \pi_g}{\log \pi_h} \right\rceil,$$

where $\lceil x \rceil$ denotes the smallest integer that is bigger or equal to x. Assume first that $k \geq N$ so that $k' \geq 0$. Since $\gamma_{k'} = 1$, $P_{k'} \leq \overline{M}$. Therefore

$$P_k \leq \alpha_N M_1 I \triangleq M(\epsilon)$$

with probability $1-\epsilon$, where I is the identity matrix of appropriate dimension.

Now consider the case $k < N$, it is easy to see

$$P_k \leq \alpha_N M_0 I \triangleq M(\epsilon)$$

with probability at least $1-\epsilon$.

(2) $\pi_g < \epsilon \leq 1$.
Assume first $k \geq 2$. Let $N = 1$ so that $k' = k-1$, *i.e.* , the previous packet is received and $P_{k-1} \leq \overline{M}$. Then

$$P_k \leq \alpha_1 M_1 I \triangleq M(\epsilon)$$

with probability at least $1-\epsilon$. When $k = 1$,

$$P_k \leq \alpha_1 M_0 I \triangleq M(\epsilon)$$

with probability at least $1-\epsilon$. □

5 Determining the M-ϵ Relationship

It is apparent that $M(\epsilon)$ given in the Theorem above is very conservative and we seek a tighter bound for the expression

$$\mathbf{Pr}[P_k \leq M] = 1 - \epsilon \ . \tag{22}$$

We begin by finding an upper bound on ϵ given M. Recall the bound on the error covariance after a packet is received is given by $\overline{M}$ as in Eqn. (21). Then define $\epsilon_i(k)$ as the probability that at least the previous i consecutive packets are dropped at time k, *i.e.*

$$\epsilon_i(k) = \Pr[N_k \geq i] \ , \tag{23}$$

with N_k the number of consecutive packets dropped at time k. Note that $N_k = (1-\gamma_k)(1+N_{k-1})$. Clearly $\epsilon_i \geq \epsilon_j$ for $i \leq j$. Next define

$$k_{\min} \triangleq \min\{k \in Z^+ : h^k(\overline{M}) \not\leq M\} \tag{24}$$

Lemma 8 *For $0 \leq M < \infty$, the quantity $k_{\min}$ will always exist.*

Proof: To prove the existence of $k_{\min}$ note 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 $k_{\min}$ such that $h^{k_{\min}}(\overline{M}) \not\leq tI$ and t can be chosen such that $tI \geq M$. This means $\lambda_n(h^{k_{\min}}(\overline{M})) > t$ and $\lambda_n(M) < t$, where λ_n is the maximum eigenvalue. Then using Weyl's Theorem [3] we see $\lambda_n(h^{k_{\min}}(\overline{M}) - M) \geq \lambda_n(h^{k_{\min}}(\overline{M})) - \lambda_n(M) > 0$ which implies $h^{k_{\min}}(\overline{M}) \not\leq M$. □

Theorem 9 *For unstable A assume the initial error covariance matrix P_0 is given by $P_0 \leq \overline{M}$. Given a matrix bound $M \geq \overline{M}$ then we have the following lower bound*

$$\mathbf{Pr}[P_k \leq M] = 1 - \epsilon \geq 1 - \epsilon_{k_{\min}}(k) \ . \tag{25}$$

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 $P_0 \leq \overline{M}$, then assuming the next k packets are dropped we have $P_k = h^k(P_0)$ and it is clear that $P_0 \leq \overline{M} \Rightarrow h^k(P_0) \leq h^k(\overline{M})$ so

$$P_k \leq h^k(\overline{M}) \ .$$

So the necessary condition that $P_k \not\leq M$ is

$$h^k(\overline{M}) \not\leq M \ ,$$

but by definition $h^k(\overline{M}) \leq M \ \forall k < k_{\min}$. Thus for $P_k \not\leq M$ it is necessary to drop at least the previous $k_{\min}$ consecutive packets.

Now assume a packet is not received until time $m > k_{\min}$, that is $\gamma_k = 0$ for $k = 0, \cdots, m-1$ and $\gamma_m = 1$, then $P_{m+1} \leq \overline{M}$ from Eqn. (21). Thus for a packet received at time m, we have

$$P_{m+1} \leq \overline{M} \ . \tag{26}$$

Regardless of how large m is, *i.e.* how long between packet receives, and how large the error covariance gets, Eqn. (26) holds. Hence the analysis above can always be repeated with P_{m+1} replacing P_0, and the probability $P_k \not\leq M$ depends only on the number of consecutive packets dropped and is independent of what happens prior to the last packet received. □

Now we will also establish an upper bound on $1-\epsilon$ that is valid under certain conditions. Recall $\overline{P}$ is the solution to the Riccati equation, $g(\overline{P}) = \overline{P}$. The extra condition we will require to establish a lower bound on ϵ is that the relation

$$\overline{P} < \overline{M} \tag{27}$$

holds. Now define

$$k_{\max} \triangleq \min\{k \in Z^+ : h^k(\overline{P}) > M\} \ , \tag{28}$$

Lemma 10 *It is always true that $h(\overline{P}) \geq \overline{P}$ which implies $h^{k+1}(\overline{P}) \geq h^k(\overline{P})$.*

Proof: Since $\overline{P}$ is the solution to the DARE we can write

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

With $h(\overline{P}) \geq \overline{P}$ if we apply h to both sides k times we get $h^{k+1}(\overline{P}) \geq h^k(\overline{P})$. □

Lemma 11 *If A is purely unstable, $k_{\max}$ is guaranteed to exist.*

Proof: If A is purely unstable then $\lim_{k\to\infty} \lambda_{\min}(h^k(X)) = \infty$. Thus we can again pick any finite scalar $t > 0$ such that $tI > M$ and find a $k_{\max}$ such that $h^{k_{\max}}(\overline{P}) \geq tI > M$. □

Lemma 12 *With the definitions above, if $k_{\min}$ and $k_{\max}$ both exist then $k_{\min} \leq k_{\max}$.*

Proof: This can easily be shown by contradiction. Assume $k_{\min} > k_{\max}$. By assumption $\overline{P} < \overline{M}$ implying $h^{k_{\max}}(\overline{P}) < h^{k_{\max}}(\overline{M})$ and if $k_{\min} > k_{\max}$ then $h^{k_{\max}}(\overline{M}) \leq M$. From the definition of $k_{\max}$, however, we see $h^{k_{\max}}(\overline{P}) > M$ which is a contradiction of the previous inequality. Hence it must be true that $k_{\min} \leq k_{\max}$. □

Corollary 13 *If A is purely unstable and assuming $\overline{P} \leq P_0 \leq \overline{M}$, then we have the upper bound*

$$\mathbf{Pr}[P_k \leq M] = 1 - \epsilon \leq 1 - \epsilon_{k_{\max}}(k) \ . \qquad (29)$$

Proof: Following the proof of Theorem 9, assume the first k packets are dropped so $P_k = h^k(P_0)$. A sufficient condition for $P_k \nleq M$ is then

$$h^k(\overline{P}) > M \ ,$$

since $P_k = h^k(P_0) \geq h^k(\overline{P})$. By definition $h^k(\overline{P}) > M$ will first hold when $k = k_{\max}$. Then since $h^{k+1}(\overline{P}) \geq h^k(\overline{P})$, it will also hold for $k > k_{\max}$. Thus dropping at least the previous $k_{\max}$ consecutive packets guarantees $P_k \nleq M$. Now assume a packet is not received until time $m > k_{\max}$, then we know $P_m = h^m(P_0) \geq h^m(\overline{P}) > M$ and $\overline{P} \leq P_{m+1} \leq \overline{M}$ 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 14 *Consider the scalar system $A = 1.3$, $C = 1$, $Q = 0.5$ and $R = 1$. For this system we have $\overline{P} = 1.519$ and with $S = 1$ and picking $p = 0$ we get $\overline{M} = 2.19$. Picking $M = 6.25$ it is easy to show $k_{\min} = 2$ and $k_{\max} = 3$. Thus there exists an $\overline{P} < X^* \leq M$ such that all for $\overline{P} \leq X < X^*$ it requires 3 consecutive packets to be dropped before the error covariance is greater than M, while for the region $X^* \leq X \leq \overline{M}$ it only requires 2 consecutive packets be dropped. In fact it can be easily shown that $X^* = 1.7174$.*

Figure 2 shows the evolution of the error covariance for a particular sequence of packet drops. 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.

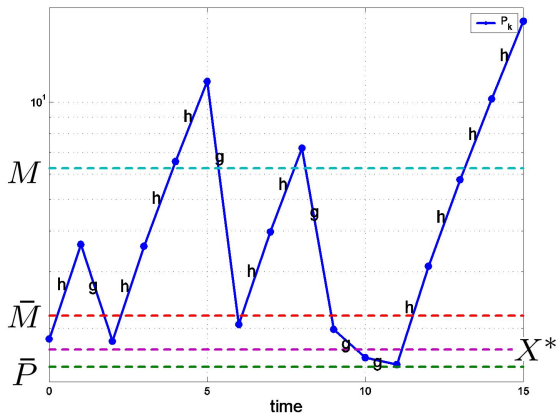

Fig. 2. Error covariance (*log* scale) for Example 14. For this system it will take at least 2 and no more than 3 consecutive dropped packets for $P_k \nleq M$.

Remark 15 *With the definition of $\epsilon_i(k)$ in Eqn. (23) it is easy to see $\epsilon_i(k) = 0$, $\forall k < i$. Which leads to $Pr[P_k \leq M] = 1$, $\forall k < k_{\min}$. In addition, for the estimator algorithm to be implemented and Eqn. (21) to hold requires at least $S + p$ measurements. Therefore we only consider time greater than* $\max\{k_{\min}, S + p\}$.

For these results to be useful we need to calculate $\epsilon_i(k)$. Figure 3 shows all possible packet sequences at time k for a packet dropping network. From this it is clear to see that $\epsilon_i(k)$ will be the sum of the probabilities of each of the instances with at least the previous i packets dropped occurring.

Corollary 16 *For $k > i$ any packet dropping network that is either i.i.d. or reaches a steady state (for example a Markov network), $\epsilon_i(k) = \epsilon_i$ is independent of k.*

The above corollary says the probability of dropping at least the previous i packets is the same for all time. To calculate $\epsilon_{k_{\min}}$ (or $\epsilon_{k_{\max}}$) we can make use of the Markov chain model in Figure 4.

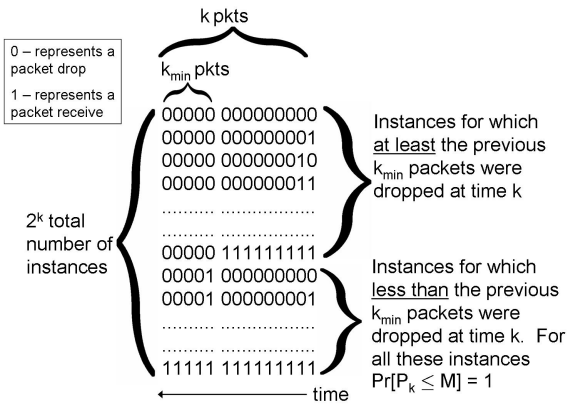

Fig. 3. A binary representation of the possible packet sequences (*i.e.* drop/receive) at time k. A 0 signifies a packet was dropped and 1 signifies the packet was received.

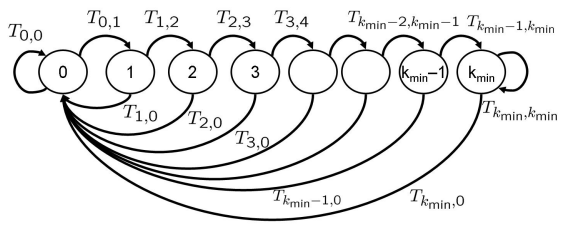

Fig. 4. The states of the Markov chain represent the number of consecutive packets dropped at the current time, the final state represents $k_{\min}$ or more consecutive packets dropped. The transition probability from state i to state j is given by $T_{i,j}$. The same figure can be made for $k_{\max}$.

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

$$\pi_{k_{\min}} = \frac{D}{D + T_{k_{\min},0} + T_{k_{\min},0} \sum_{l=1}^{k_{\min}-1} \prod_{j=0}^{l-1} T_{j,j+1}} \tag{30}$$

with

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

Note that $\pi_{k_{\min}}$ decreases as $k_{\min}$ increases. The same formula holds for $k_{\max}$ by replacing $k_{\min} \leftarrow k_{\max}$.

The $T_{i,j}$ are determined based on the type of network. For example, an i.i.d. network with packet arrival rate γ and drop rate $1-\gamma$ has $T_{j,0} = \gamma \ \forall j \geq 0$, $T_{j,j+1} = 1-\gamma \ \forall j \geq 0$, and $T_{k_{\min},k_{\min}} = 1-\gamma$. This leads to $\pi_{k_{\min}} = (1-\gamma)^{k_{\min}}$. A first order Markov network with transition probabilities T_{hh}, T_{hg}, T_{gh}, and T_{gg} leads to $\pi_{k_{\min}} = \frac{1-T_{gg}}{2-T_{hh}-T_{gg}} (T_{hh})^{k_{\min}-1}$. The probability $\pi_{k_{\min}}$ for any arbitrary order Markov network can be determined in this manner. All the equations above can be used to calculate $\epsilon_{k_{\max}}$ as well by simply replacing $k_{\min} \leftarrow k_{\max}$

Theorem 9 and Corollary 13 provide bounds on ϵ for a given M and the network properties, *i.e.* $\pi_{k_{\min}}$ and $\pi_{k_{\max}}$. It is also possible to determine bounds on M and $\pi_{k_{\min}}$.

Corollary 17 *With the same assumptions as Theorem 9 and given the transition probabilities $T_{i,j}$ of the Markov model in Figure 4 and a lower bound $1-\epsilon_{k_{\min}}$ it is possible to determine a suitable M such that $Pr[P_k \leq M] \geq 1-\epsilon_{k_{\min}}$. To do so, define*

$$k_M \triangleq \min \{k \in Z^+ : \pi_k \leq \epsilon_{k_{\min}}\} \ , \tag{31}$$

with π_k given in Eqn. (30). Then the tightest such bound is

$$M = h^{k_M}(\overline{M}) \ . \tag{32}$$

Corollary 18 *Likewise, given M and a lower bound $1-\epsilon_{k_{\min}}$ it is possible to determine limits on the transition probabilities $T_{i,j}$ of the Markov model in Figure 4 such that $Pr[P_k \leq M] \geq 1-\epsilon_{k_{\min}}$. With $k_{\min}$ as defined in Eqn. (24), it is easy to see that we require*

$$\pi_{k_{\min}} \leq \epsilon_{max} \ . \tag{33}$$

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

6 Simulation Example

Consider the linearized pendubot system in [8] with

$$A = \begin{bmatrix} 1.001 & 0.005 & 0.000 & 0.000 \\ 0.35 & 1.001 & -0.135 & 0.000 \\ -0.001 & 0.000 & 1.001 & 0.005 \\ -0.375 & -0.001 & 0.590 & 1.001 \end{bmatrix}, \ B = \begin{bmatrix} 0.001 \\ 0.540 \\ -0.002 \\ -1.066 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ R = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix},$$

$$Q = qq', \ q' = \begin{bmatrix} 0.003 & 1.000 & -0.005 & -2.150 \end{bmatrix}$$

and an i.i.d. network with packet acceptance rate $\gamma = 0.75$. For this system we have $S = 2$, meaning $[C, A'C']'$ is full rank and we need to transmit at least 2 measurements at each time step. Using the analysis presented in this paper, we can predict with what probability the error will remain below certain bounds. The value of the trace of $\overline{M}$ as a function of the number of additional measurements to buffer p is shown in Figure 5. This

curve can be used as a guide to pick the amount of data measurements to buffer. Note the error covariance if all measurements were received, $\overline{P}$, has a trace of 16.27. For the simulations presented below we used $p = 7$ (so we transmit a total of $S+p = 9$ measurements) which gives $Tr(\overline{M}) = 16.99$.

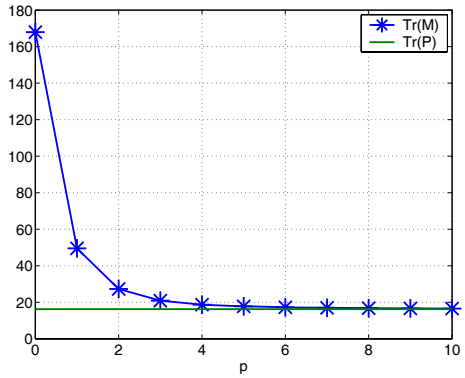

Fig. 5. The trace of $\overline{M}$ bound vs. p.

Figure 6 shows the M - ϵ relationship for this system. A total of 10,000 simulations were run with a random initial error covariance in the range $\overline{P} \leq P_0 \leq \overline{M}$ chosen for each simulation. The simulations were run for 500 time steps and the $1-\epsilon$ 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 $1-\epsilon$ are given by $1-\epsilon_{k_{\min}}$ and $1-\epsilon_{k_{\max}}$ which exhibit sharp jumps, *i.e.* the staircase, as $k_{\min}$ and $k_{\max}$ change integer values.

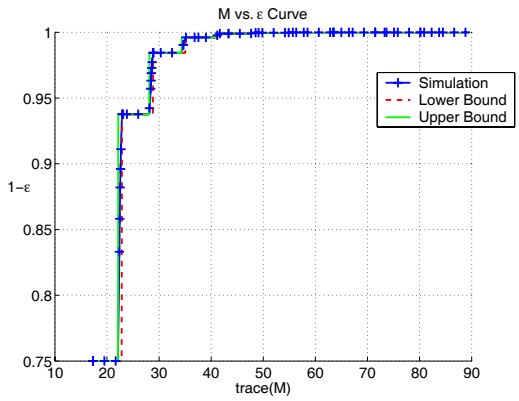

Fig. 6. M bound vs. ϵ. The solid (blue) line is the simulated $1-\epsilon$ and the dashed (red and green) lines are the predicted $1-\epsilon_{\max}$ and $1-\epsilon_{\min}$.

7 Conclusions and Future Work

We analyzed the problem of state estimation where measurement packets are sent across a lossy network. We designed an estimator algorithm that is guaranteed to have an upper bound on the estimation error covariance whenever a measurement packet is received that relies on transmitting the current and several previous sensor measurements.

We showed with this upper bound that as long as the expected value of receiving packets is not identically zero, then for any given $0 < \epsilon < 1$ there exists an $M(\epsilon) < \infty$ such that the error covariance matrix P_k is bounded by M with probability $1-\epsilon$. This analysis is independent of the probability distribution of packet drops.

Next, we give explicit relations for upper and lower bounds on the probability $1-\epsilon_{\min} \leq \mathbf{Pr}[P_k \leq M] \leq 1-\epsilon_{\max}$. We observe that $P_k \not\leq M$ only if a large enough consecutive burst of packets are dropped before time k. The size of the required burst is dependent on M.

References

[1] R.W. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Trans. Automatic Control*, 45(7):1279–1289, July 2000.

[2] M. Epstein. *Managing Information in Networked and Multi-Agent Control Systems*. PhD thesis, California Institute of Technology, 2007.

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

[4] F.L. Lian, J. Moyne, and D. Tillbury. Peformance evaluation of control networks. *IEEE Control Systems Magazine*, 21(1):66–83, 2001.

[5] X. Liu and A. Goldsmith. Kalman filtering with partial observation losses. In *IEEE Conf. on Decision and Control*, pages 4180–4186, Dec 2004.

[6] G.N. Nair and R.J. Evans. Communication-limited stabilization of linear systems. In *IEEE Conf. on Decision and Control*, pages 1005–1010, Dec 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 *IEEE Conf. on Decision and Control*, pages 304–309, Dec 2001.

[8] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Foundations of control and esimation over lossy networks. *Proceedings of the IEEE, Special Issue on Networked Control Systems*, 95(1):163–187, January 2007.

[9] L. Shi, M. Epstein, A. Tiwari, and R.M. Murray. Estimation with information loss: Asymptotic analysis and error bounds. In *IEEE Conf. on Decision and Control*, pages 1215–1221, Dec 2005.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. on Automatic Control*, 49(9):1453–1464, 2004.

[11] S.C. Tatikonda. *Control Under Communication Constraints*. PhD thesis, Massachusetts Institute of Technology, 2000.

[12] W.S. Wong and R.W. Brockett. Systems with finite communication bandwidth-part i: State estimation problems. *IEEE Trans. Automatic Control*, 42(9):1294–1299, Sept 1997.

[13] W.S. Wong and R.W. Brockett. Systems with finite communication bandwidth-part ii: Stabilization with limited information feedback. *IEEE Trans. Automatic Control*, 44(5):1049–1053, May 1999.