# Sending the Most Recent Observation is not Optimal in Networked Control: Linear Temporal Coding and Towards the Design of a Control Specific Transport Protocol

C. L. Robinson
Department of Industrial and Enterprise Systems Engineering,
and the Coordinated Science Lab,
University of Illinois at Urbana-Champaign
clrobnsn@uiuc.edu

P. R. Kumar
Department of Electrical and Computer Engineering,
and the Coordinated Science Lab,
University of Illinois at Urbana-Champaign
prkumar@uiuc.edu

***Abstract*— This paper explores diversity of temporal observations in a networked control system or sensor network. We analyze *what* information should be sent between a sensor and a controller (or estimator) in a networked control system (or sensor network) where the two components are separated by an unreliable, bandwidth limited communication. Packets may be dropped at any time. Given a sensor of limited computational and storage capability restricted to transmitting linear combinations of measurements, we consider what should be transmitted following a packet drop - the most recent observation, the previously dropped observation, or a combination of the two?**

**We show that the common practice of sending only the most recent observation is not optimal. We then derive necessary and sufficient conditions for an optimal linear combination of past and present observations. We address a special case where sensor bandwidth (or sampling rate) is higher than communication bandwidth (or throughput), and deal with the case of multiple dropped packets.**

**These results suggest the design of a transport layer specific to networked control which optimizes packet contents contingent on previous packet loss. Alternatively, one could optimize access between contending sensors for a scarce communication medium. The results could be regarded as *network coding across time*.**

**Simulations are used to illustrate the theoretical results.**

## I. Introduction

In networked control systems or sensor networks, communication or computational capacities are limited, and so it is useful to revisit familiar problems involving estimation and control to determine a suitable domain specific architecture.

We consider a networked control system (or sensor network) where communication between the sensor and controller (or state estimator) is subject to unpredictable packet loss, as shown in Figure 1. What should be done when a packet is lost and cannot be retransmitted before new information becomes available? Traditionally, this issue is addressed by the Transport layer. For services such as file transfer, reliability is key, and so packets are simply retransmitted until they are received. This approach is not appropriate in networked control for at least two reasons. First, packets received very late have little value (unlike in a data network) since information in a control system has a *time value*. Traditional networking is especially weak in addressing issues such as latency and timely delivery. Second, communication bandwidth may be small, making packet retransmission infeasible. Is ‘retransmission’ in fact the best response? These questions raise the issue of designing an appropriate *application specific transport layer for networked control systems*. This paper provides results that suggest that a *linear temporal coding* based transport layer may be used for network control. This is attractive because it leads to performance improvement as we show, and allows control system design *including* the transport layer to stay within the linear design paradigm, for which highly effective control design tools have been developed over the decades.

Consider responding to a sensor measurement being dropped en route to the controller. A new observation is taken before retransmission is possible. If the communication network capacity and latency were not an issue, sending *both* measurements represents the most complete information transfer. The optimal solution under this favourable situation, see [16], therefore lower bounds optimal cost. While this approach has theoretical value, excessive bandwidth requirement prevents implementation. Alternatively, using an encoder-decoder pair [6] achieves optimality, but may require significant computational capability.

This leads us to consider sensors with limited communication and computational capacity. What should be transmitted in order to minimize the state estimation error covariance or performance cost? We show that a linear combination of the

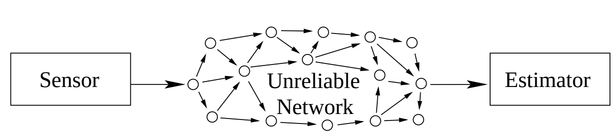


Fig. 1. Information sent from the sensor through the network to the estimation or control logic is subject to unpredictable packet loss.

most recent ‘new’ observation and the dropped ‘old’ observation is strictly better than sending the most recent observation alone. This establishes the benefit of *linear temporal encoding*.

These results also address optimal linear strategies for systems where the sampling frequency (or bandwidth) is much larger than the communication bandwidth (or throughput). If a sensor takes multiple observations before transmitting, our results produce the ‘best’ packet to send. This can lead to designing network scheduling policies and message composition for systems with multiple competing sensors.

Our results give some guidance toward the design of an intelligent transport protocol capable of modifying *packet contents* based on the disposition of previous packets. Such a protocol would manage packets already *in* the transmission queue. This goes beyond active queue management [1], [3], for congestion control, which merely deletes or otherwise deals with packets in buffers, or modulates packet transmission *rate*. Our scheme, specific to networked control, composes packets from those already in the queue. This can be through of as ‘network coding across time’ (as compared to ‘network coding across space’) to enhance performance.

Section III formulates the problem and provides some intuition. In Section IV we derive necessary and sufficient conditions for the optimal ratio of ‘new’ and ‘old’ observations. In Section V we present the results for an ‘alternate drop’ or ‘oversampling’ system in which more observations are taken than can be transmitted. An optimal ratio is obtained which is identical to the result in Section IV. Finally, Section VI presents a simulation study. Concluding remarks, implications, and future extensions are discussed in Section VII.

## II. Related Work and Underlying Assumptions

The effects of random packet losses have been investigated for Kalman filtering [17], optimal control [10], and stability; see [2], [8] for multiple references. Regulating access to the communication medium [9], [20], reducing communication overhead [23], and controlling systems over bandwidth limited channels [14], [19] have been considered. In general, systems with random unknown packet drops are more difficult to analyze due to non-traditional ‘Witsenhausen’ type information patterns [22].

The results are generally divided based on the transport model used - either ‘TCP’ (which guarantees in-order packet delivery but may incur significant delays), or ‘UDP’ (lower latency but no delivery guarantee). There are intermediate transport layer levels of service such as DCCP, TFRC and SCTP which are potentially beneficial for control. The DCCP protocol [11] is an unreliable transport protocol with end-to-end congestion control. The random delays and in-order packet delivery restrictions of TCP are avoided, while packet delivery notifications are given. The TFRC protocol [7] smooths the abrupt send-rate changes associated with TCP flows by regulating the allowable transmition rate. The SCTP protocol [18] forgoes the byte oriented approach of TCP in favour of a message oriented protocol. This enables simultaneous message delivery despite individual bytes arriving out of order.

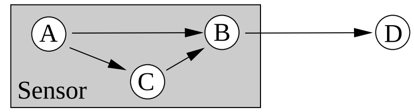


Fig. 2. Potential network coding representation. Node $A$ is the sensor. The unreliable channel is between nodes $B$ and $D$. At time $k$, node $A$ transmits to nodes $B$ and $C$. Transmissions take a single time step. At the next step, node $B$ transmits to $D$, receives a measurements for time $k$ from $C$ and for time $k+1$ from $A$. Node $B$ then chooses what to send to $D$.

Protocols specific for control also exist [13], [15], such as Ethernet (e.g., LonWorks), token based schemes (e.g., PROFIBUS and ControlNET) and Control Area Networks (CAN) (e.g., DeviceNet). The protocols typically provide lower OSI level functionality (link and MAC layers) and the particular implementation provides the higher levels. The Ethernet based schemes provide rapid access to the network and high data rates, but give no guarantees on message delay or delivery. Token ring/bus type architectures bound the maximum delivery delay by the token transfer time. The CAN protocol, used mainly in automotive domains, uses arbitration and priorities to give guarantees to nodes.

Combining packet contents is studied in Network Coding, which under some circumstances can enable a maximal information rate as well as minimal delay [4]. In Linear Network coding, outgoing messages are constructed using *random* linear combinations of incoming messages. Our approach is somewhat similar, as shown in Figure 2. In our work, instead of using a *random* combination, we determine the optimal coefficients and investigate ‘Network Coding *across Time*’.

### *A. Background Assumptions*

We will use a DCCP type protocol with predictable timing properties. Hence, the separation theorem holds, and the optimal control problem is separated into optimal state estimation and certainty equivalent optimal control law design. Thus, we shall use the message disposition knowledge to improve the *state estimate*, which will improve overall system performance for both the control and the state estimation problems.

We will only consider packet loss. Hence, our results are subject to the maximum packet drop rate: $p < 1/\lambda_{max}^2$ described in [10], [17], which is applicable regardless of the information structure [16]. Our results can readily be extended to incorporate delay.

We assume that if a packet is dropped, a new state observation is available before any retransmission is possible. We shall consider limited bandwidth systems which transmit only a *single* observation at any time. We sidestep the issue of infinite resolution and information in a real number since we only consider linear combinations, which are also subject to noise. The question we consider is whether a *linear combination* of the past (*dropped*) observation with the current (*new*) observation in a single ‘combined’ observation can improve the state estimation error covariance at the receiver. Motivation for using linear processing is that it maintains the linear paradigm, and thus the use of well established linear control design principles.

In [6] a Kalman Filter with zero control input is used to create a single value, which is transmitted. A 'smart sensor' is required to execute the Kalman Filter. We shall assume that the sensor capacity is limited such that only one measurement can be stored, and only linear combinations can be computed.

## III. Problem Formulation

For simplicity we will consider a scalar system:

$$\begin{aligned} x_{k+1} &= ax_k + bu_k + w_k, \\ y_k &= cx_k + v_k, \end{aligned} \tag{2}$$

where $x_k$ represents the system state and $y_k$ is the noisy measurement at time $k$. The variances of the jointly Gaussian independent state noise $w_k$, and observation noise $v_k$, are $q_w$ and $r_v$ respectively.

We shall use a quadratic cost criterion:

$$J = E\sum_{k=0}^{N}\{x_k' Q_x x_k + u_k' R_u u_k\}, \tag{3}$$

where $Q_x \geq 0$ and $R_u > 0$. Reconditioning with respect to measurements one can write:

$$J = E\sum_{k=0}^{N-1}\left(\hat{x}_k' Q_x \hat{x}_k + u_k' R_u u_k + Tr(Q_x P_k)\right), \tag{4}$$

where $\hat{x}_k$ is the state estimate and $P_k$ is the conditional covariance of the state estimation error. As discussed above, our aim is to improve the state estimate, i.e., minimize $P_k$.

### A. Accounting for Dropped Packets

To incorporate an unreliable network into (2), we assume an independent Bernoulli erasure channel with packet acknowledgments. A single measurement is stored at the sensor, which has taken another measurement. Hence, at time $k$ the sensor has measurements for times $k$ and $k-1$. With this in mind we formulate a new concatenated system as:

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ x_k \end{bmatrix} &= \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_k + \begin{bmatrix} w_k \\ 0 \end{bmatrix}, \\ y_k &= cx_k + v_k, \end{aligned} \tag{5}$$

$$\begin{aligned} z_k &= \begin{bmatrix} \alpha & \beta \end{bmatrix}\begin{bmatrix} y_k \\ y_{k-1} \end{bmatrix}, \\ &= \begin{bmatrix} \alpha & \beta \end{bmatrix}\begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} \alpha & \beta \end{bmatrix}\begin{bmatrix} v_k \\ v_{k-1} \end{bmatrix}, \end{aligned}$$

$$Q_w = \begin{bmatrix} q_w & 0 \\ 0 & 0 \end{bmatrix}, \tag{6}$$

$$R_v = (\alpha^2 + \beta^2) r_v. \tag{7}$$

Here, $y_k$ is the measurement, while $z_k$ is the information sent at time $k$. Without loss of generality we assume $c = 1$. We use $C = (\alpha, \beta)$ to trade off the composition of the transmitted observation and define it as the *linear temporal code*. If there was *never* a packet loss, there is no benefit in retransmitting an observation. Hence we would set $\alpha = 1$ and $\beta = 0$. Note that scaling $(\alpha, \beta)$ to $(\theta\alpha, \theta\beta)$ makes not difference for $\theta \neq 0$. Hence we need only consider $(\alpha, \beta)$ on the unit circle $\alpha^2 + \beta^2 = 1$. The state noise covariance in this formulation is represented by $Q_w$, as shown in (6), and the observation noise is given in (7) as $R_v$. We have used the fact that $E[v_k v_{k-1}'] = 0$ and $E[v_k^2] = E[v_{k-1}^2]$ since they are i.i.d. with zero mean.

For ease of reference, we now present the Kalman filter equations [21]. The time update equations are:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k, \tag{8}$$

$$P_{k+1|k} = AP_{k|k}A' + Q_w. \tag{9}$$

The measurement update equations are:

$$K_{k+1} = P_{k+1|k}C'\left(CP_{k+1|k}C' + R_v\right)^{-1}, \tag{10}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left(z_k - c\hat{x}_{k+1|k}\right), \tag{11}$$

$$P_{k+1|k+1} = \left(I - K_{k+1}C\right)P_{k+1|k}, \tag{12}$$

where $\hat{x}_{k+1|k}$ represents the state estimate at time $k+1$ given state observations up to time $k$. When combined, the Kalman filter represents the discrete Riccati equation:

$$P_{k+1} = A\left[P_k - P_k C'[CP_kC' + R_v]^{-1}CP_k\right]A' + Q_w, \tag{13}$$

where by convention $P_{k+1} = P_{k+1|k}$. We partition the covariance matrix before a packet drop as:

$$P_{k|k} = \begin{bmatrix} p_{k|k}^{1,1} & p_{k|k}^{1,2} \\ p_{k|k}^{1,2} & p_{k|k}^{2,2} \end{bmatrix}. \tag{14}$$

The optimal state estimate, when no observation is received by the estimator, is obtained by doing an open loop time prediction of the Kalman filter using (9), which yields:

$$\begin{aligned} P_{k+1|k} &= AP_{k|k}A' + Q_w \\ &= \begin{bmatrix} a^2 p_{k|k}^{1,1} + q_w & ap_{k|k}^{1,1} \\ ap_{k|k}^{1,1} & p_{k|k}^{1,1} \end{bmatrix}. \end{aligned} \tag{15}$$

When an observation packet is dropped, a measurement update (12) is skipped, and another time update is performed:

$$P_{k+2|k} = \begin{bmatrix} q_w + a^2(q_w + a^2 p_{k|k}^{1,1}) & a(q_w + a^2 p_{k|k}^{1,1}) \\ a(q_w + a^2 p_{k|k}^{1,1}) & q_w + a^2 p_{k|k}^{1,1} \end{bmatrix}. \tag{16}$$

If an observation then arrives, it will be associated with an $\alpha$, $\beta$, and a $R_v$. Performing a measurement update using (10) and (12) yields $P_{k+2|k+2}$. The $p_{k+2|k+2}^{1,1}$ term is given by (1) at the bottom of the page. This term will incur a cost in (3).

$$p_{k+2|k+2}^{1,1} = q_w + a^2\left(a^2 p_k^{1,1} + q_w\right) - \frac{\left(q_w + \left(a^2 p_k^{1,1} + q_w\right)\left(a^2\alpha + a\beta\right)\right)\left(\left(a^2 p_k^{1,1} + q_w\right)\left(a\beta + a^2\alpha^2\right) + \alpha q_w\right)}{\left(a^2 p_k^{1,1} + q_w\right)(\alpha a + \beta)^2 + \alpha^2 q_w + r_v(\alpha^2 + \beta^2)} \tag{1}$$

We will investigate the optimal choice of $\alpha$ and $\beta$ so as to minimize this term.

For insight, making the substitutions $a = 2$, $r_v = 1$ and $q_w = 1$ into (1) and assuming $p_{k|k}^{1,1} = 1$, yields:

$$p_{k+2|k+2}^{1,1} = \frac{21\alpha^2 + 26\beta^2}{22\alpha^2 + 20\alpha\beta + 6\beta^2}, \tag{17}$$

which can be minimized to a value of $\frac{13}{16}$ with $(\alpha, \beta) = \left(\frac{13}{5}, 1\right)$. Using only the most recent measurement, $(\alpha, \beta) = (1, 0)$, yields a covariance of $\frac{21}{22}$. A plot of the cost is shown in Figure 3. We examine necessary and sufficient conditions for a minimum in Section IV.

## IV. Theoretical Results

The following theorems give necessary and sufficient conditions for a global optimal minimizing $(\alpha, \beta)$. First, we restate a well known result [12] for convenience.

*Lemma 4.1:* Monotonicity of Riccati equation iterations. Let $P_k$ and $\bar{P}_k$ represent the $k^{th}$ iteration of (13) with initial conditions $P_0$ and $\bar{P}_0$ respectively. Let $P_\infty$ represent the steady state solution to (13). If $P_0 \geq \bar{P}_0 \geq P_\infty$ then $P_k \geq \bar{P}_k \geq P_\infty$.

Lemma 4.2 establishes that there is no trade-off between present and future. I.e., optimizing a single Kalman filter update at time $k$ is optimal for *all* subsequent time too, regardless of the future policy:

*Lemma 4.2:* Minimizing $P_{k|k}^{1,1}$ is an optimal policy for all $P_j$ for all $j > k$.

*Proof:* Consider $P$ after a Kalman time update. The only term in $P_{k+1|k}$ from $P_{k|k}$ is $p_{k|k}^{1,1}$, as illustrated in (15). If the optimal $(\alpha^*, \beta^*)$ was used to generate $p_{k|k}^{1,1\,*}$, then $P_{k+1|k}$ generated using $p_{k|k}^{1,1\,*}$ in (9) is also minimal.

Consider a second system where a non-optimal choice of $(\tilde{\alpha}, \tilde{\beta})$ was used to compute $\tilde{p}_{k|k}^{1,1}$, which is used to find $\tilde{P}_{k+1|k}$. Since $\tilde{p}_{k|k}^{1,1} > p_{k|k}^{1,1\,*}$, we have $\tilde{P}_{k+1|k} > P_{k+1|k}^*$. Any subsequent choice of $(\alpha, \beta)$ by the second system can also be used by the first system. Hence by Lemma 4.1, $\tilde{P}_j > P_j^*$. ■

We show later that forming the optimal estimate for $p_{k+2|k+2}^{1,1}$ does *not* yield the optimal estimate for $p_{k+2|k+2}^{2,2}$. Consider first a *single* packet drop with no consecutive drops:

*Theorem 4.3:* Following a single packet drop, the optimal linear combination of past and present measurements is:

$$\frac{\alpha^*}{\beta^*} = a + \frac{q_w}{a r_v} + \frac{q_w}{a(a^2 p_k^{1,1} + q_w)}. \tag{18}$$

*Proof:* Computing the necessary condition for a stationary point $\left(\frac{dp_{k+2|k+2}^{1,1}}{d\alpha} = \frac{dp_{k+2|k+2}^{1,1}}{d\beta} = 0\right)$ yields:

$$\frac{-1}{\beta}\frac{dp_{k+2|k+2}^{1,1}}{d\alpha} = \frac{1}{\alpha}\frac{dp_{k+2|k+2}^{1,1}}{d\beta} = -\frac{2\Theta\Gamma}{\Phi^2} = 0, \tag{19}$$

where

$$\begin{aligned}
\Theta &= \left(a^4\,p_k^{1,1} + q_w + a^2\,q_w\right)\,\alpha + a\,\left(a^2\,p_k^{1,1} + q_w\right)\beta,\\
\Gamma &= \left(a^2 p_k^{1,1} + q_w\right)\left(a r_v \alpha - a^2 r_v \beta - q_w \beta\right) - q_w r_v \beta,\\
\Phi &= \left(a^2 p_k^{1,1} + q_w\right)(a\alpha + \beta)^2 + (q_w + r_v)\,\alpha^2 + r_v \beta^2.
\end{aligned}$$

Hence, a necessary condition is that either $\Theta$ or $\Gamma$ are zero, since $\Phi$ is bounded. Now consider positive semi-definiteness of the Hessian $\mathcal{H}$, which requires all principal minors have non-negative determinant. Substituting $\Gamma = 0$ into the first term in $\mathcal{H}$ yields:

$$\frac{2\,\Psi\beta^4\,(\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4)^2\left(a^4\,p_k^{1,1} + q_w + a^2\,q_w\right)^6}{a^2\left(a^2\,p_k^{1,1} + q_w\right)^2 r_v^2 \beta^6 (\Delta_1 + \Delta_2 + \Pi_1 + \Pi_2)^3} \geq 0, \tag{20}$$

where

$$\begin{aligned}
\Psi &= \left(a^4\,p_k^{1,1}\,r + (q_w + r_v)^2 + a^2\,\left(q_w\,r_v + p_k^{1,1}\,(q_w + r_v)\right)\right),\\
\Delta_1 &= a^8\,{p_k^{1,1}}^2\,r_v + {q_w}^2\,(q_w + r_v)\,,\\
\Delta_2 &= a^6\,p_k^{1,1}\,\left(2\,q_w\,r_v + p_k^{1,1}\,(q_w + r_v)\right),\\
\Delta_3 &= a^4\,q_w\,\left(q_w\,r_v + 2\,p_k^{1,1}\,(q_w + 2\,r_v)\right),\\
\Delta_4 &= a^2\,{q_w}^2\,\left(p_k^{1,1} + q_w + 3\,r_v\right),\\
\Pi_1 &= a^4\,q_w\,\left(q_w\,r_v + 2\,p_k^{1,1}\,(q_w + 2\,r_v)\right),\\
\Pi_2 &= a^2\,{q_w}^2\,\left(p_k^{1,1} + q_w + 3\,r_v\right).
\end{aligned}$$

This expression is positive definite for all non-degenerate cases (e.g., $a = r_v = q_w = 0$). The condition $\Theta = 0$ yields a *negative* definite solution, and need not be considered further. The determinant of $\mathcal{H}$ is:

$$det(\mathcal{H}) = \frac{-4\,\Theta^2\,\Gamma^2}{\Phi^4}, \tag{21}$$

which is clearly positive semi-definite for $\Gamma = 0$. Hence, $\Gamma = 0$ is a minimum. This is infact a global minimum because we can restrict our attention to the circle $\alpha^2 + \beta^2 = 1$. See Figure 3 for a graphical representation. ■

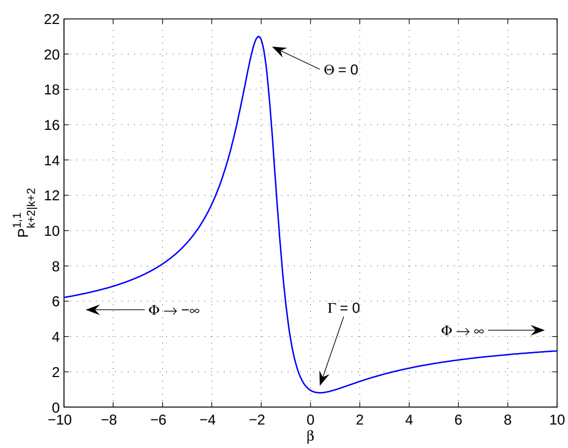


Fig. 3. The magnitude of the error covariance $p_{k+2|k+2}^{1,1}$ as a function of $\beta$ with $\alpha = 1$, $a = 2$, $p_{k|k}^{1,1} = 1$, $r_v = 1$ and $q_w = 1$. Note that $\lim_{\Phi \to \pm\infty} P_{k+2|k+2}^{1,1} = q_w + a^2\left(a^2 p_k^{1,1} + q_w\right)$, as can be seen in (1).

The following theorem addresses the more general case of *multiple* successive packet drops.

*Theorem 4.4:* Following $D$ successive dropped observations, the optimal linear combination of the last and current measurements is:

$$\frac{\alpha^*}{\beta^*} = a + \frac{q_w}{a r_v} + \frac{q_w}{a(a^{2D} p_k^{1,1} + \sum_{i=1}^{D} a^{2(i-1)} q_w)}. \quad (22)$$

*Proof:* Follows directly from the previous proof since $a^{2D} p_k^{1,1} + \sum_{i=1}^{D} a^{2(i-1)} q_w$ is the state estimation error covariance projected forward $D$ drops and simply replaces the single drop state error covariance in (15). ■

### A. Interpretation of Results

Theorem 4.3 provides some interesting insight and results:

- $r_v$ Small values of $r_v$ represent near perfect observations, and in the limit leads to $\frac{\alpha}{\beta} \to \infty$, indicating that only the most recent observation should be sent. For large noise, weighting of the missed observation is lower bounded by $\frac{\alpha}{\beta} \geq a + \frac{q_w}{a(a^2 p_k^{1,1} + q_w)}$.
- $a$ As the system becomes kinematic ($a \to 0$), any present control action or system noise can arbitrarily change the state. Thus, there is no correlation between states at observation instants, and no advantage in communicating past information. Hence, $\frac{\alpha}{\beta} \to \infty$ as $a \to 0$.
- $q_w$ The limit $q_w \to 0$ represents noiseless state prediction. It is intriguing that in this case $\frac{\alpha}{\beta} \to a$, for which we still have no intuitive explanation.
- $p_k^{1,1}$ The point of interest here is that the ratio depends on the estimation covariance before the packet drop, but is bounded for all $p_k^{1,1}$ by $a + \frac{q_w}{a r_v} \leq \frac{\alpha}{\beta} \leq a + \frac{q_w}{a r_v} + \frac{1}{a}$.

## V. Oversampling

We now study the case where measurements are delivered reliably, but are taken at double the rate they are transmitted. Starting with the formulation in (5):

$$\begin{aligned}
\begin{bmatrix} x_{k+2} \\ x_{k+1} \end{bmatrix} &= \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{k+1} \\ x_k \end{bmatrix} + u_{k+1} + \begin{bmatrix} w_{k+1} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} a^2 & 0 \\ a & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_k \\ u_{k+1} \end{bmatrix} \\
&\quad + \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} w_k \\ w_{k+1} \end{bmatrix}, \\
z_k &= \alpha y_k + \beta y_{k-1} = \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} y_k \\ y_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} v_k \\ v_{k-1} \end{bmatrix}.
\end{aligned}$$

This represents the system evolution between transmitted observations. Since measurements are delivered regularly, consider only even $k$ by setting $k = 2n$, and define:

$$\begin{aligned}
X_n &:= \begin{bmatrix} x_{2n} \\ x_{2n-1} \end{bmatrix}, \\
\bar{W}_n &:= \begin{bmatrix} w_{2n} \\ w_{2n+1} \end{bmatrix} \rightsquigarrow N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} q_w & 0 \\ 0 & q_w \end{bmatrix} \right), \\
\bar{V}_n &:= \begin{bmatrix} v_{2n} \\ v_{2n-1} \end{bmatrix} \rightsquigarrow N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} r_v & 0 \\ 0 & r_v \end{bmatrix} \right), \\
Z_n &:= z_{2n}.
\end{aligned}$$

This can be used to specify the system:

$$\begin{aligned}
X_{n+1} &= \begin{bmatrix} a^2 & 0 \\ a & 0 \end{bmatrix} X_n + \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix} \bar{W}_n, \\
Q_w &= cov\left( \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix} \bar{W}_n \right) = \begin{bmatrix} a^2+1 & a \\ a & 1 \end{bmatrix} q_w, \\
R_v &= cov\left( \begin{bmatrix} \alpha & \beta \end{bmatrix} v_n \right) = \left(\alpha^2 + \beta^2\right) r_v, \\
Z_n &= \begin{bmatrix} \alpha & \beta \end{bmatrix} X_n + \begin{bmatrix} \alpha & \beta \end{bmatrix} \bar{V}_n.
\end{aligned}$$

We will now study convergence of the discrete Riccati equation for state estimation of this system. For existence and uniqueness we require stabilizability of $(A, Q_w)$ and detectability of $(A, C)$, where $C = [\alpha, \beta]$. This can be easily shown for $(a, q_w)$ stabilizable, and $(a, \alpha)$ or $(a, \beta)$ detectable.

Consider a single iteration of (13). Note that if $P_{n+1|n} \to P$ as $n \to \infty$, then $P_{n|n} \to \tilde{P}$ where $\tilde{P}$ is related to $P_{n+1|n}$ through (12). Computing $p_{n+1|n+1}^{1,1}$ and taking the derivative with respect to $\beta$ and $\alpha$ yields precisely expression (19) in the proof of Theorem 4.3. Thus, the minimizing ratio in this case is exactly the same:

$$\frac{\alpha^*}{\beta^*} = a + \frac{q_w}{a r_v} + \frac{q_w}{a(a^2 p_n^{1,1} + q_w)}. \quad (23)$$

### A. Intermediate State Estimation

If state estimation at intermediate times when the observation is dropped/not sent is of importance (i.e., $p_{k+1|k+2}^{1,1}$ or equivalently $p_{k+2|k+2}^{2,2}$), we can find the optimal $(\alpha, \beta)$ ratio by similar arguments as in Theorem 4.3. The optimal ratio is:

$$\frac{\alpha}{\beta} = \frac{a r_v}{q_w + r_v}. \quad (24)$$

Hence, the choice of $C$ effects both state estimation and estimate smoothing since the optimizer for each is different. This is shown graphically in Figure 4.

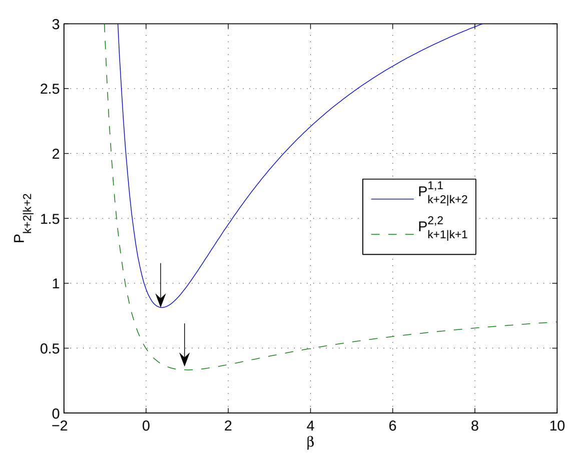


Fig. 4. The graph illustrates the minimum points of the error covariance $p_{k+2|k+2}^{1,1}$ and $p_{k+2|k+2}^{2,2}$ as a function of $\beta$, with $\alpha = 1$, $a = 2$, $p_{k|k}^{1,1} = 1$ and $q_w = 1$. The arrows indicate the minimum point. As expected, more weighting is given to the 'missed' observation (larger $\beta$) at time $k+1$ when the state covariance at that time is to be minimized.

These results only consider sampling at double the transmission rate. This Corollary deals with multiple observations:

*Corollary 5.1:* If observations are transmitted every $D$ samples, the optimal linear combination of the current and previous measurement is given by

$$\frac{\alpha}{\beta} = a + \frac{q_w}{ar_v} + \frac{q_w}{a(a^{2D}p_k^{1,1} + \sum_{i=1}^{D} a^{2(i-1)}q_w)}.$$

*Proof:* Follows directly from Theorem 4.4. ■

## VI. Simulations

We illustrate our results with a simulation study using system parameters $a = 2$, $r_v = 1$, $q_w = 1$ and $\alpha = 1$. The estimation error cost is computed as $\frac{1}{N}\sum_{k=0}^{N} p_k^{1,1}$ with $N = 10^6$. The results are shown in Figure 5, where for a baseline comparison we have used the lower bound from [16].

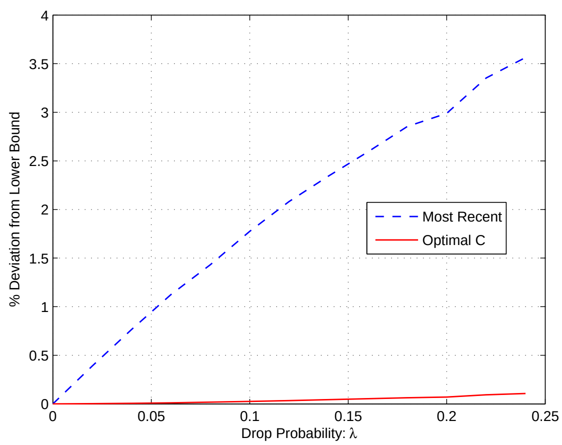


Fig. 5. Simulation results showing the percentage deviation in cost from the lower bound described in [16]. $\lambda$ is the packet loss probability.

## VII. Implications and Conclusions

We have demonstrated that sending the most recent observation, in systems where observations are occasionally lost, is not optimal. We have derived conditions for the existence of a linear combination of past and present measurements which minimizes the state estimation error covariance. This highlights the utility of *linear temporal coding*.

Showing the benefit of sending combined measurements suggests several future research directions. One is regarding a transmission protocol for networked control systems, where the contents of packets waiting to be sent are modified depending on the previous transmission status. This is similar to "in-network information processing" in sensor networks [5], except that now it is at the transport rather than the network layer. Sensors could be allocated static amounts of communication bandwidth and modify packet *contents* to meet the requirement, as compared to TCP type protocols that regulate the packet transmission rate. Other extensions include examining the multi-dimensional case, examining the case where delivery status is unknown, as well as finding the optimal combination of all dropped packets.

## Acknowledgment

This material is based upon work partially supported by NSF under Contract Nos. ECCS-070160, CNS 05-19535 and CCR-0325716, AFOSR under Contract No. F49620-02-1-0217, DARPA/AFOSR under Contract No. F49620-02-1-0325, DARPA under Contact No. N00014-0-1-1-0576, and Oakridge-Battelle under Contract DOE BATT 4000044522.

## References

[1] R. Braden, D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, and L. Zhang. Recommendations on queue management and congestion avoidance in the internet. RFC 2309, April 1998.

[2] M. Chow and Y. Tipsuwan. Network-based control systems: A tutorial. In *The $27^{th}$ Annual Conference of the IEEE Industrial Electronics Society.*, pages 1593–1602, 2001.

[3] S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *ACM/IEEE Transactions on Networking*, 1(4), August 1993.

[4] C. Fragouli, J.-Y. L. Boudec, and J. Widmer. Network coding: an instant primer. *SIGCOMM Comput. Commun. Rev.*, 36(1):63–68, 2006.

[5] A. Giridhar and P. R. Kumar. Towards a theory of in-network computation in wireless sensor networks. *IEEE Communications Magazine*, 44(4):98–107, April 2006.

[6] V. Gupta, B. Hassibi, and R. M. Murray. Optimal LQG control across packet-dropping links. *Systems & Control Letters*, June 2006.

[7] M. Handley, S. Floyd, J. Pahdye, and J. Widmer. TCP friendly rate control (TFRC): Protocol specification. Technical report, ftp://ftp.isi.edu/in-notes/rfc3448.txt, 2003.

[8] J. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE: Special Issue: Technology of Networked Control Systems*, 95(1), Jan. 2007.

[9] D. Hristu-Versakelis. Feeback control systems as users of a shared netowrk: Communication sequences that guarantee stability. In *Proc. of the IEEE Conf. on Decision and Control.*, 2001.

[10] O. C. Imer, S. Yuksel, and T. Basar. Optimal control of dynamical systems over unreliable communication links. In *NOLCOS*, 2004.

[11] E. Kohler, M. Handley, and S. Floyd. Designing DCCP: Congestion control without reliability. In *Proceedings of ACM SIGCOMM*, 2006.

[12] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*, volume 1. Englewood Cliffs, 1986.

[13] F. Lian, J. Moyne, and D. Tilbury. Performance evaluation of control networks: Ethernet, ControlNet, and DeviceNet. *IEEE Control Systems Magazine*, 21(1):66–83, 2001.

[14] G. N. Nair and R. J. Evans. Exponential stabilisability of finitedimensional linear systems with limited data rates. *Automatica*, 39(4), April 2003.

[15] R. S. Raji. Smart networks for control. *IEEE Spectrum*, 31(6):49–55, 1994.

[16] C. L. Robinson and P. R. Kumar. Control over networks of unreliable links: Location of controllers, control laws and bounds on performance. In *Proceedings of Control over Communication Channels (ConCom) - To Appear*, Cyprus, April 2007.

[17] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. Kalman filtering with intermittent observations. In *IEEE Transactions on Automatic Control*, volume 49, 2004.

[18] R. Stewart and Q. Xie. *Stream Control Transmission Protocol (SCTP): A Reference*. Addison-Wesley, 1 edition, 2001.

[19] A. Tatikonda, S. Sahai and S. Mitter. Control of lqg systems under communication constraints. In *Proc. of the IEEE Conf. on Decision and Control.*, pages 1165–1170, Tampa, Fl., December 1998.

[20] G. C. Walsh and H. Ye. Scheduling of networked control systems. *IEEE Control Systems Magazine*, 2001.

[21] G. Welch and G. Bishop. An introduction to the kalman filter. http://www.cs.unc.edu/∼welch/media/pdf/kalman_intro.pdf, April 2004.

[22] H. S. Witsenhausen. A counter example in stochastic optimal control. *Siam J. Control*, 6:131–147, 1968.

[23] J. K. Yook, D. M. Tilbury, and N. R. Soparkar. Trading computation for bandwidth: Reducing communication in distributed control systems using state estimators. *IEEE Transactions on Control Systems Technology*, 10(4):503–518, July 2002.