Optimal Redundant Transmission for State Estimation with Packet Drops

Girish N. Nair Alexandre Mesquita Joao P. Hespanha

Abstract: In networked systems that suffer from data packet drops, transmitting multiple, redundant packets during each sampling interval can improve estimation performance, but at the expense of a higher communication rate. In recent work, this idea was developed to introduce the notion of a dynamic redundant transmission policy, with Markov decision theory used to find the optimal policy numerically. The purpose of this paper is to present an alternative approach to this problem. By relaxing the integer requirement on the number of packets transmitted, it becomes possible to explicitly find a real-time recursion for the optimal transmission function. The theoretical properties of this recursion are then analysed to propose a simple numerical procedure for finding the minimum over all real-valued transmission functions, by searching over a one-dimensional parameter space. This then yields an implementable, suboptimal policy when discretised to integer values. These results are supported by numerical studies in MATLAB.

Keywords: Packet drops, state estimation, limited communication.

1. INTRODUCTION

In networked sensor and actuator systems, it is known that communication imperfections can have a significant impact on real-time estimation and control performance (see, e.g. Antsaklis and Baillieul [2007] and the papers therein). Examples of such imperfections include channel noise, finite bit rates, limited communication bandwidth, dropped packets, etc. For such systems, designing the communications protocols with reference to the specific system objectives can improve performance significantly.

In this paper, we are concerned with the problem of linear state estimation over a communication channel that largely operates in a dichotomous fashion, i.e. either carrying a transmitted data packet perfectly with no or constant delay, or dropping it entirely. Two instances include data links that have low noise and bit error rates but that lose packets at congested intermediate nodes, and digital channels with error detection coding, whereby the receiver can determine if packets have been corrupted by bit errors but cannot correct them. Furthermore, we suppose that the number of bits in each packet is sufficiently large that the effect of quantisation errors can be safely ignored. Thus, we may assume that each packet carries a real-valued scalar or vector.

Most previous work on this topic assumes that only one packet can be sent over the channel from the sensor to the estimator during each sampling interval. If the channel drops packets in an independent and identically distributed (iid) way and there is sufficient computational ability at the transmitter side, then the mean-square-optimal strategy is to pre-filter the system measurements by using a Kalman filter at the transmitter side Gupta et al. [2007]. One the other hand, if there is no computation possible at the transmitter and only the raw measurement can be transmitted, then in Sinopoli et al. [2004] the optimal filter at the receiver side is derived, and furthermore there is a critical drop probability above which mean-square bounded estimation errors are impossible.

The recent paper Mesquita et al. [2009], explored the possibility of transmitting multiple packets per sampling interval, e.g. by use of an orthogonal scheme such as time- or frequency-division multiplexing (TDM/FDM). It is easy to see that by transmitting multiple redundant packets at each time instant, the probability that at least one packet arrives at the receiver increases, and consequently the estimation error improves at the expense of a higher communication load. With the further assumption that the receiver sends acknowledgements back to the transmitter, this led to the notion of an optimal *redundant transmission policy*, whereby the number of packets transmitted at each time instant is chosen dynamically based on available information, so as to minimise a suitable cost that captures both the average estimation error and the average communication rate. For situations where only moderate computational power was available at the receiver, the methods of Markov decision process (MDP) theory, in particular value iteration, were used to numerically determine optimal "simple" policies, whereby the number of packets transmitted at each time depends in a stationary way on just the current number of past consecutive transmission failures.

The purpose of this paper is to explore an alternative approach to finding "simple" policies, without using the tools of MDP theory. By relaxing the integer requirement on the number of packets transmitted, it becomes possible to explicitly find a real-time recursion for the optimal transmission function. We analyse the properties of this recursion to propose a simple numerical procedure for finding the minimum over all real-valued transmission functions, by searching over a one-dimensional parameter $\alpha \in \mathbb{R}$. This then yields an implementable, suboptimal policy when discretised to integer values. These results are supported by numerical studies in MATLAB.

⋆ The 1st author is at Dept. Elect. & Electr. Eng., Uni. Melbourne, VIC 3010, Australia (email: gnair@unimelb.edu.au). The 2nd and 3rd authors are at Dept. Elect. & Comp. Eng., Uni. California, Santa Barbara, CA 93106-9560, USA. The 1st author acknowledges the partial support of the Australian Research Council under grant DP0985397.

Preprint submitted to 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems. Received May 3, 2010.

2. FORMULATION

Consider a fully observed, stochastic, linear time-invariant (LTI) system

$$X_{t+1} = AX_t + W_t, \quad \forall t \in \mathbb{Z}_{\geq 0}, \tag{1}$$

with state $X_t \in \mathbb{R}^n$, noise $W_t \in \mathbb{R}^n$ and dynamical matrix $A \in \mathbb{R}^{n\times n}$. Assume that A has at least one eigenvalue of magnitude ≥ 1 and that the noise process W_0^∞ is independent of X_0, has zero mean, and is uncorrelated i.e.

$$\mathrm{E}\{W_s W_t'\} = \begin{cases} 0 & \text{when } s \neq t \\ \Sigma_W & \text{when } s = t \end{cases}, \quad \forall s,t \in \mathbb{Z}_{\geq 0}, \tag{2}$$

where Σ_W is a constant $n \times n$ covariance matrix.

Suppose that at each time t the state is coded into one or more packets for transmission over a communication channel to an estimator located elsewhere. For simplicity suppose each packet is dropped or transmitted with constant probabilities, independently of every other packet and of the initial state and dynamical noise in (1). That is, any packet fails to arrive at the estimator, with probability $p \in (0,1)$, or else arrives uncorrupted before the next time instant, with probability $1-p$. Further suppose that the number of bits in each packet is large enough that the effect of quantisation errors is negligible compared to that of packet loss.

Assume the transmitter has only modest computational capabilities. A simple strategy it can adopt to counteract packet losses is to transmit N_t identical packets at time t, each carrying the value of X_t. The transmission at time t is a success if at least one of the copies arrives at the receiver, in which case at time $t+1$ the controller knows the value of X_t exactly. Plainly, the probability of transmission failure is p^{N_t}.[1] Similar to *transmission control protocol*, just before time $t+1$ the receiver then sends a binary-valued acknowledgement $\Phi_t = 0$ or 1 back to the transmitter, to indicate whether or not the transmission at time t failed or succeeded, respectively. For simplicity, assume that this acknowledgement is received before time $t+1$.

In principle, the transmitter has available the entire past record of transmission successes and failures, and could use this information, together with the state history, to determine how many packets N_t. However, suppose that, due to the computational restrictions mentioned above, the transmitter can store only the number $L_t \in \mathbb{Z}_{\geq 0}$ of consecutive failures since the last successful transmission prior to time t and then chooses N_t to be a static function of this number, i.e. $N_t \equiv \nu(L_t)$. The process Φ_0^∞ then evolves according to

$$\mathrm{P}\left\{\Phi_t = 0 | \Phi_0^{t-1}, X_0, W_0^\infty, V_0^\infty\right\} = \mathrm{P}\{\Phi_t = 0 | L_t\} = p^{\nu(L_t)} \tag{3}$$

$$\mathrm{P}\left\{\Phi_t = 1 | \Phi_0^{t-1}, X_0, W_0^\infty, V_0^\infty\right\} = \mathrm{P}\{\Phi_t = 1 | L_t\} = 1 - p^{\nu(L_t)} \tag{4}$$

Note that Φ_t is dependent on past values, via L_t.

At the other end of the channel, the estimator stores an estimate $\hat{X}_t \in \mathbb{R}^n$ of the plant state at time t and updates its estimate as

$$\hat{X}_{t+1} = \begin{cases} A\hat{X}_t & \text{if } \Phi_t = 0 \\ AX_t & \text{if } \Phi_t = 1 \end{cases}, \quad \forall t \in \mathbb{Z}_{\geq 0}. \tag{5}$$

We do not address here the question of whether or not the optimal estimator has the structure above. Instead, our focus is solely on how to design the transmission redundancy function ν, given this structure. In doing so, two competing objectives must be balanced. The first is the estimation error, as defined by the average quadratic cost

$$J_e(\nu) := \limsup_{k\to\infty} \frac{1}{k} \sum_{t=0}^{k-1} \mathrm{E}\{E_t' Q E_t\}, \tag{6}$$

where $E_t := \hat{X}_t - X_t$ and Q is a given positive definite matrix, while the second is the average communication rate

$$R(\nu) := \limsup_{k\to\infty} \frac{1}{k} \sum_{t=0}^{k-1} \mathrm{E}\{N_t\} \quad \text{(packets/sample)}. \tag{7}$$

It is clear that these criteria are in conflict, since increasing N_t reduces the chances of transmission failure and therefore estimation errors, but directly increases the communication rate. In order to achieve a satisfactory trade-off, we seek to minimise the linear combination

$$J(\nu) = J_e(\nu) + \lambda R(\nu), \tag{8}$$

where $\lambda > 0$ is a given weight.

This optimisation problem was first posed in Mesquita et al. [2009], where it was formulated in terms of a Markov decision process (MDP). By placing a bound on N_t and using value iteration, a numerical procedure for finding the optimal stationary redundant transmission policy was proposed and studied.

In this paper, we return to the same formulation but with a different point of view. By relaxing the integer constraints on N_t and using alternative techniques, we explicitly derive the form of the optimal real-valued policy. Rounding this up then yields a simple, near-optimal integer-valued policy.

3. COST IN TERMS OF ν

The analysis of this problem hinges on the dynamics of the past consecutive failure process L_0^∞. As described in Mesquita et al. [2009], this is a Markov chain with dynamics that are independent of the noise process W_0^∞ and the initial system state X_0. There are two possible transitions from each possible state of the Markov chain: if transmission were to succeed at time t then the number of past consecutive failures would be reset to 0; otherwise it would increment by 1. That is, $\forall l \in \mathbb{Z}_{\geq 0}$,

$$\begin{aligned} \mathrm{P}\{L_{t+1} = l+1 | L_t = l, L_0^t, X_0, W_0^\infty\} &= \mathrm{P}\{L_{t+1} = l+1 | L_t = l\} \\ &= p^{\nu(l)}, \end{aligned} \tag{9}$$

$$\begin{aligned} \mathrm{P}\{L_{t+1} = 0 | L_t = l, L_0^{t-1}, X_0, W_0^\infty\} &= \mathrm{P}\{L_{t+1} = 0 | L_t = l\} \\ &= 1 - p^{\nu(l)}, \end{aligned} \tag{10}$$

with $\mathrm{P}\{L_0 = 0\} = 1$. It can be shown that this Markov chain is irreducible and aperiodic. Thus by] there is a stationary probability mass function (pmf) μ s.t. $\mathrm{P}\{L_t = l\} \to \mu(l)$, $\forall l \in \mathbb{Z}_{\geq 0}$. From the transition probabilities (9)–(10), it is straightforward to see that μ is uniquely defined by

$$\mu(l) = \mu(l-1) p^{\nu(l-1)} = \mu(0) p^{\sum_{j=0}^{l-1} \nu(j)}, \quad \forall l \in \mathbb{Z}_{\geq 0}, \tag{11}$$

where

$$\mu(0) = \frac{1}{\sum_{l\geq 0} p^{\sum_{j=0}^{l-1} \nu(j)}} \tag{12}$$

since the pmf must sum to 1. Note that for the stationary distribution to exist, we require $\mu(0) > 0$, i.e.

$$\sum_{l\geq 0} p^{\sum_{j=0}^{l-1} \nu(j)} < \infty. \tag{13}$$

The stationary pmf is thus directly determined by the transmission redundancy function ν. The next step is to use this

[1] We assume that p is not dependent on N_t.

to express the average cost (8) in terms of ν. Observe that as $t \to \infty$,

$$\mathrm{E}\{N_t\} \equiv \sum_{l\geq 0} \nu(l)\mathrm{P}\{L_t = l\} \to \sum_{l\geq 0} \nu(l)\mu(l). \quad (14)$$

That is, the summand in (7) approaches a limiting value with time, forcing the long-term average to take the same value, i.e.

$$R(\nu) = \sum_{l\geq 0} \nu(l)\mu(l) = \frac{\sum_{l\geq 0} \nu(l) p^{\sum_{j=0}^{l-1}\nu(j)}}{\sum_{l\geq 0} p^{\sum_{j=0}^{l-1}\nu(j)}}. \quad (15)$$

Next, observe from (1) and (5) that the estimation error dynamics are given by

$$E_t = \begin{cases} AE_{t-1} - W_{t-1} & \text{if } \Phi_{t-1} = 0 \\ -W_{t-1} \text{if } \Phi_{t-1} = 1 \end{cases}. \quad (16)$$

As the last successful transmission was, by definition, at time $t - L_t - 1$,

$$E_t = -\sum_{i=0}^{L_t} A^i W_{t-1-i}.$$

$$\Rightarrow \mathrm{E}\{E_t E_t' | L_t\} = \sum_{i=0}^{L_t} A^i \mathrm{E}\{W_{t-1-i} W_{t-1-i}'\} (A^i)' \equiv \sum_{i=0}^{L_t} A^i \Sigma_W (A^i)',$$

since W_0^∞ is an uncorrelated process and is also independent of L_0^∞. Thus as $t \to \infty$,

$$\begin{aligned} \mathrm{E}\{E_t E_t'\} &= \sum_{l\geq 0} \mathrm{P}\{L_t = l\} \sum_{i=0}^{l} A^i \Sigma_W (A^i)' \\ &\to \sum_{l\geq 0} \mu(l) \sum_{i=0}^{l} A^i \Sigma_W (A^i)'. \\ \Rightarrow \mathrm{E}\{E_t' Q E_t\} &= \mathrm{tr}\left(Q\mathrm{E}\{E_t E_t'\}\right) \\ &\to \sum_{l\geq 0} \mu(l) g(l) \equiv \frac{\sum_{l\geq 0} g(l) p^{\sum_{j=0}^{l-1}\nu(j)}}{\sum_{l\geq 0} p^{\sum_{j=0}^{l-1}\nu(j)}}, \quad (17) \end{aligned}$$

where

$$g(l) := \sum_{i=0}^{l} \mathrm{tr}\left(QA^i \Sigma_W (A^i)'\right), \quad \forall l \in \mathbb{Z}_{\geq 0}. \quad (18)$$

As the LHS of (17) is the summand in the long-term average cost (6), the latter must equal the limiting value of the former. Substituting this and (15) into (8) then yields

$$J(\nu) = \frac{\sum_{l\geq 0} (g(l) + \lambda\nu(l))\, p^{\sum_{j=0}^{l-1}\nu(j)}}{\sum_{l\geq 0} p^{\sum_{j=0}^{l-1}\nu(j)}} \equiv \frac{N(\nu)}{D(\nu)}, \quad (19)$$

which expresses the cost directly in terms of ν.

Observe that for this functional to be well-defined, we require the denominator $D(\nu) \equiv \mu(0)^{-1} < \infty$, which is equivalent to the existence of the stationary distribution μ (11)–(12). For convenience, define the sequence space

$$\mathbf{V} := \left\{\nu \in \mathbb{R}_{\geq 0}^\infty : N(\nu), D(\nu) < \infty\right\}. \quad (20)$$

The following result is straightforward to establish:

Lemma 1. The cost J (19) and its numerator and denominator are continuous on $\mathbf{V}$ with respect to the metric

$$d_{\mathbf{V}}(\nu, \upsilon) := \sum_{l\geq 0} |\nu(l) - \upsilon(l)|, \quad \forall \nu, \upsilon \in \mathbf{V}.$$

Furthermore, J possesses one or more global minima over $\mathbf{V}$.

Proof: Omitted.

4. RELAXATION AND DIRECT OPTIMISATION

The number $N_t = \nu(L_t)$ of packets transmitted at time t is a positive integer and the MDP-based optimisation method studied in Mesquita et al. [2009] accomodates this constraint quite easily. However, the explicit cost representation (19) reveals a smooth dependence on $\nu(0), \nu(1), \ldots$ if the integer requirement is relaxed. We exploit this observation to propose an alternative approach, based on using calculus to directly minimise (19) over nonnegative, *real*-valued $\nu(0), \nu(1), \ldots$. We show that the global minimum in the enlarged solution space $\mathbf{V}$ is, somewhat surprisingly, uniquely defined by a simple, finite-dimensional recursion.

Next, take a partial derivative of (19) with respect to $\nu(l)$ to yield

$$\begin{aligned} \frac{\partial J(\nu)}{\partial \nu(l)} &= \frac{1}{D(\nu)} \frac{\partial N(\nu)}{\partial \nu(l)} - \frac{N(\nu)}{D(\nu)^2} \frac{\partial D(\nu)}{\partial \nu(l)} \\ &\equiv \frac{1}{D(\nu)} \left(\frac{\partial N(\nu)}{\partial \nu(l)} - J(\nu) \frac{\partial D(\nu)}{\partial \nu(l)} \right). \quad (21) \end{aligned}$$

At a global minimum ν_* we must thus have

$$\left.\frac{\partial N(\nu)}{\partial \nu(l)}\right|_{\nu=\nu_*} \begin{cases} = J(\nu) \left.\frac{\partial D(\nu)}{\partial \nu(l)}\right|_{\nu=\nu_*} & \text{if } \nu_*(l) > 0 \\ \geq J(\nu) \frac{\partial D(\nu)}{\partial \nu(l)}|_{\nu=\nu_*} & \text{if } \nu_*(l) = 0 \end{cases}. \quad (22)$$

For any $\alpha \in \mathbb{R}$ let us now define sequences $\nu_\alpha \in \mathbb{R}_{\geq 0}^\infty$ and $\kappa_\alpha \in \mathbb{R}^\infty$ by coupled recursions s.t. $\forall l \in \mathbb{Z}_{\geq 0}$,

$$\nu_\alpha(l) := \left[\frac{\alpha - g(l)}{\lambda} - \frac{1}{r} + \kappa_\alpha(l-1)\right]^+, \quad (23)$$

$$\kappa_\alpha(l) := \begin{cases} \exp(r\nu_\alpha(l))/r & \text{when } \nu_\alpha(l) > 0 \\ (\alpha - g(l))/\lambda + \kappa_\alpha(l-1) & \text{when } \nu_\alpha(l) = 0 \end{cases}, \quad (24)$$

where $[\cdot]^+ := \max\{0, \cdot\}$, $\kappa_\alpha(-1) := 0$, g is given by (18), and

$$r := -\ln p > 0. \quad (25)$$

We have our first main result:

Theorem 2. The global minimiser $\nu_* \in \mathbf{V}$ (20) of the cost $J(\nu)$ (19) is uniquely given by the sequence ν_α (23)–(24), setting parameter $\alpha = \min_{\nu\in\mathbf{V}} J(\nu)$.

Proof: Let $J_* = J(\nu_*)$ and $\eta = r/\lambda$. After substituting the expressions for the numerator $N(\nu)$ and denominator $D(\nu)$ of (19) into (22), and performing some lengthy manipulations, we obtain

$$r\nu_*(l) + \eta(g(l) - J_*) + 1 \begin{cases} = \frac{\sum_{j=0}^{l-1} (\eta(J_* - g(j)) - r\nu_*(j))\, p^{\sum_{i=0}^{j-1}\nu_*(i)}}{p^{\sum_{i=0}^{l-1}\nu_*(i)}} & \text{if } \nu_*(l) > 0 \\ \geq \frac{\sum_{j=0}^{l-1} (\eta(J_* - g(j)) - r\nu_*(j))\, p^{\sum_{i=0}^{j-1}\nu_*(i)}}{p^{\sum_{i=0}^{l-1}\nu_*(i)}} & \text{if } \nu_*(l) = 0 \end{cases},$$

which can be rearranged to yield

$$\nu_*(l) = \Bigg[\frac{J_* - g(l)}{\lambda} - \frac{1}{r} + p^{-\sum_{i=0}^{l-1}\nu_*(i)} \sum_{j=0}^{l-1}\left(\frac{J_*-g(j)}{\lambda} - \nu_*(j)\right) p^{\sum_{i=0}^{j-1}\nu_*(i)}\Bigg]^+ \quad (26)$$

$$\equiv \left[\frac{J_*-g(l)}{\lambda} - \frac{1}{r} + \kappa_*(l-1)\right]^+, \quad (27)$$

where

$$\kappa_*(l) := p^{-\sum_{i=0}^{l}\nu_*(i)} \sum_{j=0}^{l}\left(\frac{J_*-g(j)}{\lambda} - \nu_*(j)\right) p^{\sum_{i=0}^{j-1}\nu_*(i)} \quad (28)$$

$$\equiv p^{-\nu_*(l)}\left(\kappa_*(l-1) + (J_* - g(l)/\lambda - \nu_*(l)\right)$$

$$\overset{(27)}{=} \begin{cases} p^{-\nu_*(l)}/r & \text{if } \nu_*(l) > 0 \\ \kappa_*(l-1) + (\alpha - g(l))/\lambda & \text{if } \nu_*(l) = 0 \end{cases}, \quad (29)$$

with $\kappa_*(-1) := 0$. Observe that (27) and (29) are identical to (23) and (24) with $\alpha = J_*$. Furthermore, since these recursions uniquely define $\nu_*(0), \nu_*(1), \ldots$, there can be at most one minimiser that achieves the global minimum cost J_*. □

Thus we can compute the global minimiser $\nu_* \in \mathbf{V}$ via a relatively simple, finite-dimensional recursion. In addition, we have reduced the potentially infinite-dimensional search for ν_* into the tuning of a single real parameter α. The real numbers $\nu_*(0), \nu_*(1), \ldots$ must of course be rounded to integer values for implementation, yielding a possibly suboptimal transmission redundancy policy. Nonetheless the potential computational savings afforded by the explicit recursion (23)–(24) make it a viable alternative to MDP-based numerical optimisation. In addition, when memory at the transmitter is limited, we can compute this transmission policy on the fly, instead of having to store a large, pre-computed look-up table.

However, two important questions remain. Firstly, is there a way to find J_* that is simpler and more structured than generating an infinite sequences ν_α for different candidate values of α and evaluating the corresponding infinite sums in (19)? Secondly, for an arbitrary value of α, what is the value of $J(\nu_\alpha)$?

The answers to both questions will be seen to be related. Before presenting them, $\forall \alpha > 0$ let

$$l_\alpha := \min\{l \in \mathbb{Z}_{\geq 0} : g(l) > \alpha\}, \quad (30)$$

and $\forall n \in \mathbb{Z}_{\geq 0}$ and $l \geq l_\alpha$ define

$$\chi_{\alpha,n}(l) := \ln\left(\gamma_\alpha(l) + \ln\left(\gamma_\alpha(l+1) + \ln\left(\cdots + \ln\left(\gamma_\alpha(l+n)\right)\right)\right)\right) \quad (31)$$

$$\gamma_\alpha(l) := 1 + (g(l+1) - \alpha)r/\lambda, \quad (32)$$

with g the increasing, positive function of (18). Observe that when $l \geq l_\alpha$, $\gamma_\alpha(l) > 1$ and is increasing and consequently $\gamma_{\alpha,n}(l)$ is positive and increasing. It is straighforward to show that l_α is right-continuous w.r.t. α and that $\gamma_\alpha(l)$ and $\chi_{\alpha,n}(l)$ are continuous w.r.t. α. The following result will also be useful.

Lemma 3. For any $\alpha > 0$ and $l \geq l_\alpha$ (30), the limit

$$\lim_{n\to\infty} \chi_{\alpha,n}(l) =: \chi_\alpha(l) > 0 \quad (33)$$

exists and is continuous with respect to α.

Proof: Omitted. Briefly, it relies on induction, the inequality $\ln(1+x) \leq x$ and the fact that $g(l)$ grows at most exponentially. □

We have the following result.

Theorem 4. For any $\alpha > 0$, the following statements concerning the sequence ν_α (23)–(24), the cost $J(\nu_\alpha)$ (19) and its denominator $D(\nu_\alpha)$ are all equivalent.

(1) $\nu_\alpha \in \mathbf{V}$, the sequence space defined in (20).
(2) $\nu_\alpha(l) \geq \chi_\alpha(l)/r$, $\forall l \geq l_\alpha$, where l_α, χ_α are given by (30)–(33).
(3) $\nu_\alpha(l) \geq \chi_\alpha(l)/r$ when $l = l_\alpha$.
(4) $D(\nu_\alpha) < \infty$ and $J(\nu_\alpha) = \alpha$.

Proof: We show $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 2 \Rightarrow 4 \Rightarrow 1$. For convenience let $\eta := r/\lambda$.

$1 \Rightarrow 2$: We first show that $\nu_\alpha(l) > 0$, $\forall l \geq l_\alpha$. Suppose in contradiction that $\nu_\alpha(l) = 0$ for some $l \geq l_\alpha$. Then $r\kappa_\alpha(l-1) \leq 1 + \eta(g(l) - \alpha)$ by (23) and $r\kappa_\alpha(l) = r\kappa_\alpha(l-1) - \eta(g(l) - \alpha) \leq r\kappa_\alpha(l-1)$ by (24). Consequently,

$$\begin{aligned} 0 \leq r\nu_\alpha(l+1) &\equiv [r\kappa_\alpha(l) - 1 - \eta(g(l+1) - \alpha)]^+ \\ &\leq [r\kappa_\alpha(l-1) - 1 - \eta(g(l+1) - \alpha)]^+ \\ &\leq [r\kappa_\alpha(l-1) - 1 - \eta(g(l) - \alpha)]^+ \equiv r\nu_\alpha(l) = 0, \end{aligned}$$

since g is increasing. By the same token $g(l+1) > \alpha$ as well and thus by upward induction $\nu_\alpha(i) = 0, \forall i \geq l$. However, this would make $D(\nu_\alpha) = \infty$, contradicting the hypothesis of statement 1.

So $\nu_\alpha(l) > 0$ $\forall l \geq l_\alpha$, implying that

$$r\nu_\alpha(l+1) \overset{(23)}{=} r\kappa_\alpha(l) - \gamma_\alpha(l) \overset{(24)}{=} e^{r\nu_\alpha(l)} - \gamma_\alpha(l) \geq 0, \forall l \geq l_\alpha \quad (34)$$

$$\Rightarrow r\nu_\alpha(l) \geq \ln\gamma_\alpha(l). \quad (35)$$

Replacing l with $l+1$ in (35) and then using (34) yields the tighter bound

$$r\nu_\alpha(l) \geq \ln\left(\gamma_\alpha(l) + \ln\gamma_\alpha(l+1)\right), \ \ \forall l \geq l_\alpha.$$

Replacing l with $l+1$ in this bound and then using (34) yields the even tighter bound

$$r\nu_\alpha(l) \geq \ln\left(\gamma_\alpha(l) + \ln\left(\gamma_\alpha(l+1) + \ln\gamma_\alpha(l+2)\right)\right), \ \ \forall l \geq l_\alpha.$$

Continuing in this way indefinitely thus yields statement (2).

$2 \Rightarrow 3$: Trivial.

$3 \Rightarrow 2$: Suppose $r\nu_\alpha(l) \geq \chi_\alpha(l)$ for some $l \geq l_\alpha$, noting that by hyothesis this is true when $l = l_\alpha$. As $\chi_\alpha(l) > 0$, (24) implies that $r\kappa_\alpha(l) = e^{r\nu_\alpha(l)}$. Substituting this into (23) yields

$$r\nu_\alpha(l+1) = \left[e^{r\nu_\alpha(l)} - \gamma_\alpha(l)\right]^+ = e^{r\nu_\alpha(l)} - \gamma_\alpha(l) \geq \chi_\alpha(l+1) > 0,$$

since $e^{r\nu_\alpha(l)} - \gamma_\alpha(l) \geq \chi_\alpha(l+1) > 0$ by (33). Thus by induction $r\nu_\alpha(l) \geq \chi_\alpha(l) > 0$, $\forall l \geq l_\alpha$.

$2 \Rightarrow 4$.

Note that (23) and (24), take the same form as (27) and (29) in the proof of Thm. 4, and can be rewritten in the more cumbersome forms of (26) and (28). As $\nu_\alpha(l) \geq \chi_\alpha(l) > 0$, $\forall l \geq l_\alpha$ we can write

$$\nu_\alpha(l) = \frac{J_* - g(l)}{\lambda} - \frac{1}{r} + p^{-\sum_{i=0}^{l-1}\nu_\alpha(i)} \sum_{j=0}^{l-1}\left(\frac{J_*-g(j)}{\lambda} - \nu_\alpha(j)\right) p^{\sum_{i=0}^{j-1}\nu_\alpha(i)}.$$

$$\Rightarrow \nu_\alpha(l)p^{\sum_{i=0}^{l-1}\nu_\alpha(i)} = \left(\frac{\alpha - g(l)}{\lambda} - \frac{1}{r}\right)p^{\sum_{i=0}^{l-1}\nu_\alpha(i)} + \sum_{j=0}^{l-1}\left(\frac{\alpha - g(j)}{\lambda} - \nu_\alpha(j)\right)p^{\sum_{i=0}^{j-1}\nu_\alpha(i)}.$$

Splitting and collecting sums yields

$$\sum_{j=0}^{l}\nu_\alpha(j)p^{\sum_{i=0}^{j-1}\nu_\alpha(i)} = \frac{\alpha}{\lambda}\sum_{j=0}^{l}p^{\sum_{i=0}^{j-1}\nu_\alpha(i)} - \frac{1}{\lambda}\sum_{j=0}^{l}g(j)p^{\sum_{i=0}^{j-1}\nu_\alpha(i)} - \frac{1}{r}p^{\sum_{i=0}^{l-1}\nu_\alpha(i)}.$$

$$\Rightarrow \alpha\sum_{j=0}^{l}p^{\sum_{i=0}^{j-1}\nu_\alpha(i)} = \sum_{j=0}^{l}(g(j)+\lambda\nu_\alpha(j))p^{\sum_{i=0}^{j-1}\nu_\alpha(i)} + \frac{\lambda}{r}p^{\sum_{i=0}^{l-1}\nu_\alpha(i)}. \quad (36)$$

As $\forall l \geq l_\alpha$, $\nu_\alpha(l) \geq \chi(l)$, which is a positive, monotonically increasing function, it follows that $\sum_{i=0}^{j-1}\nu_\alpha(i)$ increases at least linearly with large j. Consequently the summands on the LHS of (36) decreases exponentially or faster with large j, since $p \in (0,1)$. Thus the sum on the LHS must converge as $l \to \infty$, and by definition its limit equals the denominator $D(\nu_\alpha)$ of the cost $J(\nu_\alpha)$ (19). By virtue of the same, the last term on the RHS must $\to 0$ as $l \to \infty$. Thus, the remaining sum on the RHS must also converge to a finite limit, which by definition is the just the cost numerator $N(\nu_\alpha)$. Thus we obtain

$$\alpha D(\nu_\alpha) = N(\nu_\alpha) \Rightarrow J(\nu_\alpha) = \alpha.$$

Finally, the wrap-around' implication $4 \Rightarrow 1$ is trivial. □

This result is illuminating in several regards. Firstly, for $\alpha > 0$ such that the cost $J(\nu_\alpha)$ (19) and its denominator $D(\nu_\alpha)$ are finite (as required for the cost to be well-defined), it provides a lower bound on the growth of $\nu_\alpha(l)$ in terms of $\chi_\alpha(l)$. Thus if the system matrix A in (1) is unstable or marginally stable, then $\nu_\alpha(l) \to \infty$, since $g(l)$ and hence $\chi_\alpha(l)$ would increase unboundedly with l. Together with (23) and (24), this also implies that ν_α can eventually be given by the one-dimensional recursion

$$\nu_\alpha(l) = \frac{\alpha - g(l)}{\lambda} + \frac{\exp(r\nu_\alpha(l-1)) - 1}{r}, \quad \forall l > l_\alpha. \quad (37)$$

Secondly, if $D(\nu_\alpha) < \infty$ then the cost $J(\nu_\alpha)$ either equals α or is infinite. As Thm. 2 states that the global minimiser $\nu_* = \nu_\alpha$ with $\alpha = J_*$, the minimum cost can be shown to be expressible as

$$J_* = \min\{\alpha \in \mathbb{R}_{>0} : J(\nu_\alpha) < \infty, D(\nu_\alpha) < \infty\}. \quad (38)$$

This characterisation, which we give withou proof, provides a simple way of finding J_*. Beginning with an initial value $\alpha = J_0$ which upper-bounds J_*, we can reduce α in small steps, computing $\nu_\alpha, J(\alpha)$ and $D(\alpha)$ at each stage. As long as the latter two are finite, we are guaranteed that $J(\alpha)$, being equal α, will also decrease. However, once α drops below J_*, the cost J will increase rapidly, and indeed be infinite. The previous value of α is then within a step-size of J_0.

The simple procedure above suffers from the drawback of forcing us to compute the entire sequence ν_α as well as the infinite sums $D(\nu_\alpha)$ and $N(\nu_\alpha)$(at least, many terms thereof). at each stage. In this regard, statement 3 of Thm. 4 is important, because it states that checking for the finiteness of these infinite sums amounts to verifying that $r\nu_\alpha(l)$ is no smaller than $\chi_\alpha(l)$ at the *single point* $l = l_\alpha$. This leads to the following result.

Theorem 5. The minimum value J_* of the cost $J(\nu)$ (19) over all sequences ν in the sequence space $\mathbf{V} \subset \mathbb{R}^\infty_{\geq 0}$ (20) exists and is given by

$$J_* := \min\{\alpha \in \mathbb{R}_{>0} : r\nu_\alpha(l_\alpha) \geq \chi_\alpha(l_\alpha)\}, \quad (39)$$

where ν_α is recursively generated by (23)–(24), with l_α and χ_α given by (30)–(33). The minimum cost is uniquely achieved by the sequence ν_α with $\alpha = J_*$.

Proof: For reasons of limited space only a sketch of the proof is outlined here. It first uses the continuity w.r.t. α of $\gamma_\alpha(l)$ and $\chi_\alpha(l)$ (Thm. 2) and the right-continuity of l_α to show that the infimum of the set in (39) must be an element α_* of the same. Thus by Thm. 4, $J(\nu_{\alpha_*}) = \alpha_*$ with $\nu_{\alpha_*} \in \mathbf{V}$.

Furthermore, α_* must equal the infimum cost over all α-parametrised policies $\nu_\alpha \in \mathbf{V}$. Otherwise, if $\exists\beta >$ s.t. $J(\nu_\beta) < J(\nu_{\alpha_*}) \equiv \alpha_*$ with $\nu_\beta \in \mathbf{V}$, then Thm. 4 implies that we would have $J(\nu_\beta) = \beta$ and furthermore, β would also be an element of the set of (39). So $\beta < \alpha_*$, which would contradict the fact that α_* is the minimum over that set.

However, by Thm. 2, the global minimiser is uniquely given by ν_{J_*}, so $\alpha_* = J_*$. □.

This result is also useful because for each candidate value of α, we do not need to generate the entire sequence ν_α to search for the minimum.

5. SIMULATION RESULTS

In this section, we briefly present numerical MATLAB studies that illustrate the results of the preceding section. For reasons of space and to enable comparisons, we consider only the scalar system that was solved by value iteration in Mesquita et al. [2009]. The

To determine the minimum cost J_* we first selected a candidate value J_0 that exceeds the desired J_*, then implemented the recursive equations (23)–(24) with $\alpha = J_0$, and then decremented J_0 in small steps of $\varepsilon = 0.01$ until the defining condition of the set in (39) was violated. The previous value of J_0 was then guaranteed to be within ε of J_*. (The initial value of J_0 was selected by considering a constant candidate sequence $\nu(l) = c > 0$, deriving a simple explicit upper bound on $J(\nu)$ in terms of c and the system parameters, and then minimising this analytically w.r.t. c.)

We first considered an unstable scalar system, with dynamical constant $A = 2$, noise variance $\Sigma_W = 3$, error weight $Q = 1$ and communication rate coefficient $\lambda = 2.2$. In Figs. 1 and 2 we have plotted the cost J and communication rate R for the global optimal sequence ν_* as well as for the rounded, integer-valued version of ν_*. We note that the cost for ν_* is smaller than that of the optimal simplified policy of Mesquita et al. [2009], which was derived using MDP-methods. This is because we have relaxed the integer requirement here and have also let ν_* be unbounded. On the other hand, after rounding ν_* to yield an implementable policy, we obtain slightly worse performance.

Figures 3 and 4 depict plots of the transmission redundancy functions and stationary distributions for ν_* and its rounded version, with $p = 0.05$. We remark that in general ν_* grows extremely quickly (note the logarithmic scale) and indeed is too

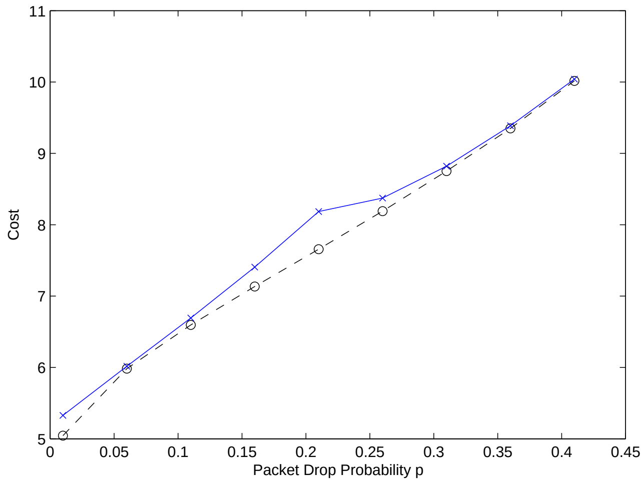

Fig. 1. J vs. p plots for ν_* (circles) and rounded ν_* (crosses).

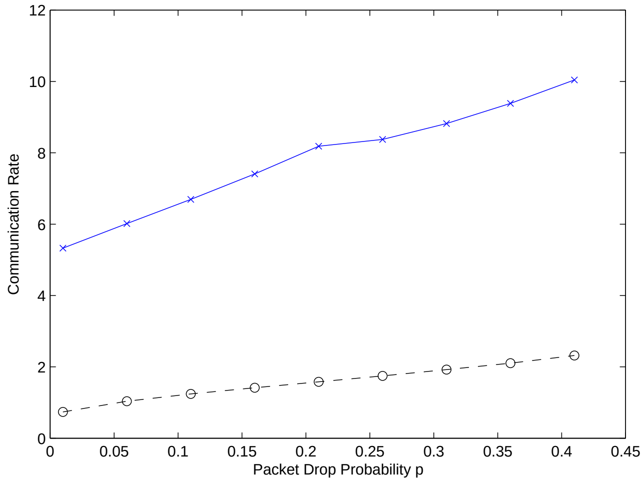

Fig. 2. R vs. p plots for ν_* (circles) and rounded ν_* (crosses).

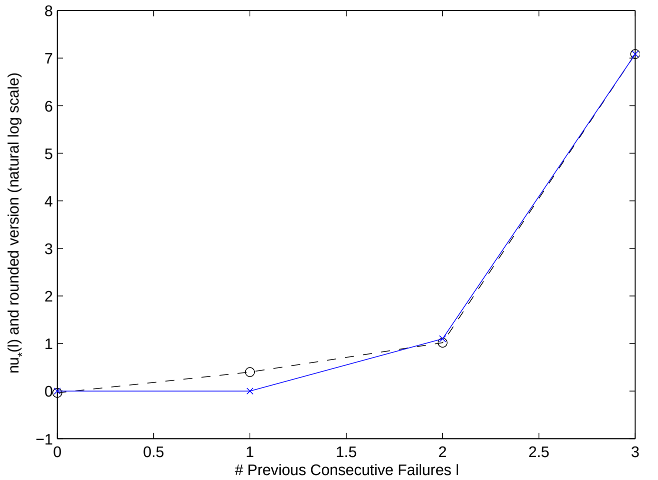

Fig. 3. Redundandt transmission functions ν_* (circles) and rounded version (crosses).

large for MATLAB to handle after about $l = 3$ or 4. However, its stationary distribution decays with corresponding rapidity, so that the average communication rate remains finite.

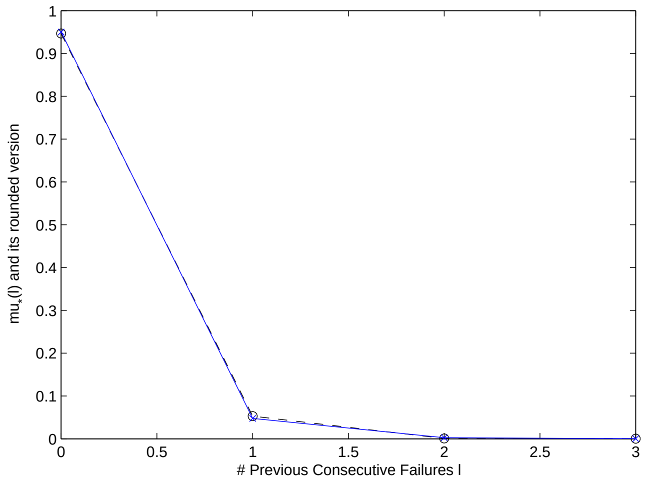

Fig. 4. Stationary distributions μ for ν_* (circles) and rounded version (crosses).

6. CONCLUSION

In this paper, we have presented and analysed an alternative approach to designing redundant transmission policies for state estimation of unstable stochastic linear systems over packet-dropping channels. By relaxing the integer requirement on the number of packets transmitted during each sampling interval, we explicitly obtained a real-time recursion for the optimal transmission function. We then analysed the properties of this recursion to propose a simple numerical procedure for finding the minimum over all real-valued transmission functions, which yields an implementable policy when discretised. Our results were supported by numerical studies in MATLAB.

Several important directions for future research include the incorporation of bounds on the number of packets that can be transmitted in each sampling interval, the study of discretisation strategies other than simple rounding, and the extension of these results to feedback control systems and costs.

REFERENCES

Antsaklis, P. and Baillieul, J. (eds.) (2007). *Special Issue on the Technology of Networked Control Systems, in Proc. IEEE*, volume 95. IEEE.

Gupta, V., Hassibi, B., and Murray, R. (2007). Optimal lqg control across packet-dropping links. *Sys. Contr. Lett.*, 56, 439–446.

Mesquita, A., Hespanha, J., and Nair, G. (2009). Redundant data transmission in control/estimation over wireless networks. In *Proc. American Control Conf.*, 3378–3383. IEEE.

Sinopoli, B., Schenato, L., Franceschetti, M., Poola, K., Jordan, M., and Shastry, S. (2004). Kalman filtering with intermittent aboservations. *IEEE Trans. Autom. Contr.*

Preprint submitted to 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems. Received May 3, 2010.