

# Optimal Communication Logics in Networked Control Systems

Yonggang Xu

João P. Hespanha

Dept. of Electrical and Computer Eng., Univ. of California, Santa Barbara, CA 93106

**Abstract**—This paper addresses the control of spatially distributed processes over a network that imposes bandwidth constraints and communication delays. Optimal communication policies are derived for an estimator-based Networked Control System architecture to reduce the communication load. These policies arise as solutions of an average cost optimization problem, which is solved using dynamic programming. The optimal policies are shown to be deterministic.

## I. INTRODUCTION

In recent years Networked Control Systems (NCS) emerged as a playground for classic information theory and control theory, in which feedback is realized by information flow over data networks. Information theory provides upper bounds on the amount of information a channel can transmit. However, the optimal coding schemes may introduce long delays, which are not acceptable for feedback control. To provide adequate control performance, in general there is a minimum requirement of information exchange among the elements of a control system [6, 7, 9]. For example, Nair and Evans [7] investigated the exponential stabilizability of LTI plants with limited feedback data rates and gave the minimum stabilizing bit-rate, which is a logarithm function of the plant poles.

In a previous paper [10], we considered an NCS structure for spatially distributed processes in continuous-time. In this control architecture, originally proposed by Yook et al. [11], local controllers interchange information from time to time. To reduce communication requirements, each local controller estimates the remote processes' states and coordinates its effort with others based on these estimates. Meanwhile, each controller broadcasts its local process' state to the others according to a protocol specified by a *communication logic*.

We proposed communication logics based on doubly stochastic Poisson processes (DSPP), in which the decision to broadcast data is triggered by increments of a DSPP with intensity (understood as a Poisson rate) related to the “necessity” to communicate. By choosing different DSPP intensity functions, we obtained different trade-offs between packet exchange rate and control performance. The threshold rule used by Yook et al. [11] can be viewed as a limiting case of this type of communication logics.

The objective of this paper is to find the *optimal* communication logic. We restrict our attention to the discrete-time domain to avoid some of the technicalities that arise

with continuous-time jump diffusion processes. We consider stochastic communication logics for which at each time instant, a probability  $u \in [0, 1]$  is assigned to whether or not data should be sent. These logics also include deterministic rules as a special case for which  $u \in \{0, 1\}$ . Since any data network induces delays, we assume that it takes  $\tau$  time steps for data from one local controller to reach another. The optimal communication problem is posed as a long term average cost (AC) minimization, and solved using dynamic programming (DP). The optimization index includes both communication and control (or estimation) performance penalties. Specifically, the per-stage cost of the AC optimization problem includes a quadratic term on the estimation error and a linear term on the communication cost. The optimal policy turns out to be deterministic, which is consistent with the observation in [10] that threshold rules appeared to be the most efficient ones.

The average cost criterion is mathematically more difficult to analyze than the discounted cost one, especially when the state and/or action spaces are Borel and the per-stage cost function is unbounded [1], which occurs in our case. We take advantage of some recurrent properties of the process and pose it as an essentially bounded cost per-stage problem, which makes results from [2, 4, 5] applicable.

In Section II, the control-communication architecture is briefly described with further details in [10]. Optimal communication logic are derived in Section III. Section IV contains conclusions and directions for future work.

## II. NCS ARCHITECTURE

In our setting, the dynamics of a spatially distributed processes are completely decoupled but the control objective is not. We view each local process with the associated local controller as a *node*. The overall control system consists of a certain number of nodes connected via a data network. Fig. 1 depicts the internal structure of the  $i$ th node. Each node consists of a *local process*, a *local controller*, a *bank of local estimators* that predicts both the local and remote process states, and a *communication logic* that schedules when to transmit data to the other nodes. The estimators run open-loop most of the time but are reset to the “correct” state when its value is received through the network. It is the responsibility of each node to broadcast to the network the state of its local process. The communication logic determines when this should occur. For simplicity of notation, we only consider two nodes but the results can be extended to any number of nodes.

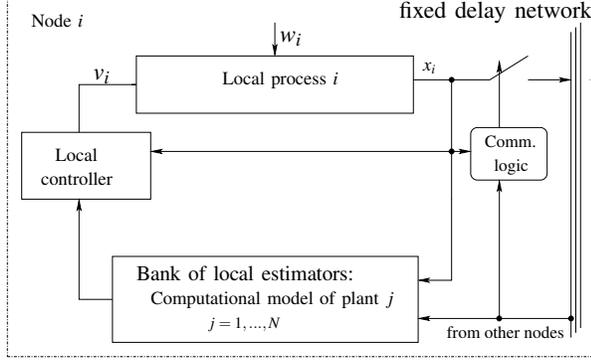


Fig. 1. One of the nodes in a networked control system.

The local processes are assumed linear time-invariant

$$\mathbf{x}_i(t+1) = A_i \mathbf{x}_i(t) + B_i \mathbf{v}_i(t) + \mathbf{w}_i(t) \quad \forall i \in \{1, 2\},$$

where  $\mathbf{x}_i \in \mathbb{R}^n$  denotes the state,  $\mathbf{v}_i \in \mathbb{R}^m$  the control input, and  $\mathbf{w}_i$   $n$ -dimensional zero-mean Gaussian white noise. The two noise processes are assumed independent and all matrices are real and of appropriate dimensions. For simplicity, we assume that all states can be measured locally, but local state observers could be included if this was not the case.

We take as given a state-feedback control laws

$$\mathbf{v}_i = K_{i1} \mathbf{x}_1 + K_{i2} \mathbf{x}_2, \quad \forall i \in \{1, 2\} \quad (1)$$

that would provide adequate closed-loop performance in a centralized configuration, resulting in the following closed-loop system

$$\mathbf{x}_1(t+1) = (A_1 + B_1 K_{11}) \mathbf{x}_1(t) + B_1 K_{12} \mathbf{x}_2(t) + \mathbf{w}_1(t), \quad (2a)$$

$$\mathbf{x}_2(t+1) = (A_2 + B_2 K_{22}) \mathbf{x}_2(t) + B_2 K_{21} \mathbf{x}_1(t) + \mathbf{w}_2(t). \quad (2b)$$

Since the state of the  $i$ th process is not directly available at the  $j$ th node ( $j \neq i$ ), we build at the node  $j$  the optimal estimate  $\hat{\mathbf{x}}_i$  of the state  $\mathbf{x}_i$ . The centralized laws (1) are then replaced by

$$\mathbf{v}_1 = K_{11} \mathbf{x}_1 + K_{12} \hat{\mathbf{x}}_2, \quad \mathbf{v}_2 = K_{21} \hat{\mathbf{x}}_1 + K_{22} \mathbf{x}_2 \quad (3)$$

in which the process state not available at a node is replaced by its estimate. The optimal estimators  $\hat{\mathbf{x}}_i$  are derived in the next section. Actually, *both nodes compute both estimates*: not only the estimate of the state of the other process used in (3), but also an estimate of its own local state. We will show that the latter is needed by the optimal communication logic to decide when to broadcast its local state to the other node.

The distributed control laws (3) result in the following closed-loop dynamics

$$\mathbf{x}_1(t+1) = (A_1 + B_1 K_{11}) \mathbf{x}_1(t) + B_1 K_{12} \hat{\mathbf{x}}_2(t) + \mathbf{w}_1(t), \quad (4a)$$

$$\mathbf{x}_2(t+1) = (A_2 + B_2 K_{22}) \mathbf{x}_2(t) + B_2 K_{21} \hat{\mathbf{x}}_1(t) + \mathbf{w}_2(t), \quad (4b)$$

to be contrasted with (2). To better understand the impact of the distributed architecture on the performance, we can re-write the closed-loop dynamics in (4) as

$$\begin{aligned} \mathbf{x}_1(t+1) &= (A_1 + B_1 K_{11}) \mathbf{x}_1(t) + B_1 K_{12} \mathbf{x}_2(t) + \mathbf{w}_1(t) \\ &\quad + B_1 K_{12} \mathbf{e}_2(t), \\ \mathbf{x}_2(t+1) &= (A_2 + B_2 K_{22}) \mathbf{x}_2(t) + B_2 K_{21} \mathbf{x}_1(t) + \mathbf{w}_2(t) \\ &\quad + B_2 K_{21} \mathbf{e}_1(t). \end{aligned}$$

where  $\mathbf{e}_j := \hat{\mathbf{x}}_j - \mathbf{x}_j$ ,  $\forall j \in \{1, 2\}$ . Comparing these equation with (2), we observe that the penalty paid for using a distributed architecture is the additive “disturbance” terms  $B_i K_{ij} \mathbf{e}_j$ ,  $i \neq j$ . Thus the performance of the closed-loop distributed architecture is to a great extent related to the properties of the estimation errors  $\mathbf{e}_j$ , which are the focus of investigation in this paper.

### III. OPTIMAL COMMUNICATION LOGIC

In this section, we first derive optimal state estimators using the data received from the network. Then an optimal communication logic for a long term average cost is derived using dynamic programming. Finally, we show that a value iteration scheme can provide an efficient way to solve the dynamic programming equation.

#### A. Optimal estimators

At each time instant  $t$ , a node  $i$  must decide whether or not it will send data. We denote by  $\mathbf{a}_i(t)$  the action it takes: either to broadcast data ( $\mathbf{a}_i(t) = 1$ ) or not ( $\mathbf{a}_i(t) = 0$ ).

In a  $\tau$ -step delay network, the information available from the network to build estimates at time  $t$  consists of all time-stamped pairs  $(s, \mathbf{x}_j(s))$  previously received, i.e.,

$$\mathcal{I}_{net}^\tau(t) := \{(s, \mathbf{x}_j(s)) \mid \mathbf{a}_j(s) = 1, s \leq t - \tau - 1, j = 1, 2\},$$

and the optimal estimate of  $\mathbf{x}_i(t)$  is given by

$$\hat{\mathbf{x}}_i^\tau(t) := \mathbb{E}[\mathbf{x}_i(t) \mid \mathcal{I}_{net}^\tau(t)].$$

We denote the corresponding estimation errors by  $\mathbf{e}_i^\tau(t) := \hat{\mathbf{x}}_i^\tau(t) - \mathbf{x}_i(t)$ . The  $\hat{\mathbf{x}}_i^\tau$  and  $\mathbf{e}_i^\tau$  just introduced correspond to the  $\hat{\mathbf{x}}_i$  and  $\mathbf{e}_i$  (without the  $\tau$  superscript) that appeared in (3) and (4). We use the superscript to emphasize the existence of a  $\tau$ -step delay. Note that every node can compute  $\hat{\mathbf{x}}_i^\tau(t)$  because  $\mathcal{I}_{net}^\tau(t)$  is available to all nodes.

To simplify the notations, we define  $A := A_{11} + B_1 K_{11} \in \mathbb{R}^{n \times n}$  and  $B := B_1 K_{12} \in \mathbb{R}^{n \times n}$ . Due to symmetry, we take the example of node 1. The optimal estimate can be constructed

recursively as follows,

$$\begin{aligned}
\hat{\mathbf{x}}_1^\tau(t+1) &= \mathbb{E}[\mathbf{x}_1(t+1) | \mathcal{S}_{net}^\tau(t+1)] \\
&= \mathbb{E}[\mathbf{A}\mathbf{x}_1(t) + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) + \mathbf{w}_1(t) | \mathcal{S}_{net}^\tau(t+1)] \\
&= A\mathbb{E}[\mathbf{x}_1(t) | \mathcal{S}_{net}^\tau(t+1)] + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) \\
&= \begin{cases} A\mathbb{E}[\mathbf{x}_1(t) | \mathcal{S}_{net}^\tau(t)] + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) & \text{if } \mathbf{a}_1(t-\tau) = 0 \\ A\mathbb{E}[\mathbf{x}_1(t) | \mathbf{x}_1(t-\tau), \mathcal{S}_{net}^\tau(t)] + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) & \text{if } \mathbf{a}_1(t-\tau) = 1 \end{cases} \\
&= \begin{cases} A\hat{\mathbf{x}}_1^\tau(t) + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) & \text{if } \mathbf{a}_1(t-\tau) = 0 \\ A^{\tau+1}\mathbf{x}_1(t-\tau) + \sum_{i=0}^{\tau} A^{\tau-i}\mathbf{B}\hat{\mathbf{x}}_2^\tau(t-\tau+i) & \text{if } \mathbf{a}_1(t-\tau) = 1 \end{cases} \quad (5)
\end{aligned}$$

where the second equality follows from (4), the third from the fact that  $\hat{\mathbf{x}}_2^\tau(t)$  is a deterministic function of  $\mathcal{S}_{net}^\tau(t+1)$ , and the fifth from the Markovian property of equation (4).

### B. Optimal communication logic

The communication logic on each node decides whether to broadcast the local state information to the remote sites. The information available to node 1's communication logic at time  $t$  to decide whether or not the current state should be broadcast to the network is denoted by  $\mathcal{S}(t)$  and includes all the information  $\mathcal{S}_{net}^\tau(t)$  received from the network, the current and past local states, as well as the whole local broadcasting history:

$$\mathcal{S}(t) := \{\mathcal{S}_{net}^\tau(t), \mathbf{x}_1(t), \dots, \mathbf{x}_1(0), \mathbf{a}_1(t-1), \dots, \mathbf{a}_1(0)\}.$$

To consider the possibility that the optimal decision may be stochastic, the logic selects a probability  $\mathbf{u}(t) \in [0, 1]$  and chooses  $\mathbf{a}_1(t)$  by a random draw according to

$$\mathbf{a}_1(t) = \begin{cases} 0 & \text{w.p. } 1 - \mathbf{u}(t) \\ 1 & \text{w.p. } \mathbf{u}(t) \end{cases} \quad \forall t \geq 0.$$

Note that when  $\mathbf{u}(t) \in \{0, 1\}$ , the rule is deterministic. From an optimization perspective,  $\mathbf{u}(t)$  is a control to be generated by a feedback policy  $\gamma := \{\gamma_0, \gamma_1, \dots\}$ , i.e.,

$$\mathbf{u}(t) = \gamma_t(\mathcal{S}(t)) \quad \forall t.$$

A long term average cost (AC) is used as the criterion to be minimized

$$J(\gamma) := \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}^\gamma \sum_{t=0}^{N-1} \mathbf{e}'(t+\tau+1) \mathbf{Q} \mathbf{e}(t+\tau+1) + \lambda \mathbf{u}(t) \quad (6)$$

where  $\lambda > 0$ ,  $\mathbf{Q} > 0 \in \mathbb{R}^{n \times n}$ , and  $\mathbb{E}^\gamma$  denotes the expectation for a given feedback policy  $\gamma$ . This criterion penalizes both the *average square-error* of the estimator

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}^\gamma \sum_{t=0}^{N-1} \mathbf{e}'(t) \mathbf{Q} \mathbf{e}(t)$$

and the *communication rate*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}^\gamma \sum_{t=0}^{N-1} \mathbf{u}(t),$$

which equals the average number of broadcasts per unit of time.

To construct the optimal communication logic, it is convenient to define

$$\hat{\mathbf{x}}_1^k(t) := \mathbb{E}[\mathbf{x}_1(t) | \mathcal{S}_{net}^\tau(t), \mathcal{S}_{loc}^k(t)], \quad 0 \leq k \leq \tau - 1,$$

which can be viewed as an estimate of  $\mathbf{x}_1(t)$  based on the information  $\mathcal{S}_{net}^\tau(t)$  received from the network, as well as

$$\mathcal{S}_{loc}^k(t) := \{(s, \mathbf{x}_1(s)) \mid \mathbf{a}_1(s) = 1, s \leq t - k - 1\},$$

which is the data from node 1 that would be available to all nodes if the network delay was only  $k$ . In practice, only node 1 can compute this estimate at time  $t$  because  $\mathcal{S}_{loc}^k(t) \subset \mathcal{S}(t)$  but  $\mathcal{S}_{loc}^k(t)$  contains information not in  $\mathcal{S}_{net}^\tau(t)$ . Proceeding as in the derivation of (5), we conclude that, for  $0 \leq k \leq \tau - 1$ ,

$$\begin{aligned}
\hat{\mathbf{x}}_1^k(t+1) &= \\
&\begin{cases} A\hat{\mathbf{x}}_1^k(t) + \mathbf{B}\hat{\mathbf{x}}_2^\tau(t) & \text{if } \mathbf{a}_1(t-k) = 0 \\ A^{k+1}\mathbf{x}_1(t-k) + \sum_{i=0}^k A^{k-i}\mathbf{B}\hat{\mathbf{x}}_2^\tau(t-k+i) & \text{if } \mathbf{a}_1(t-k) = 1. \end{cases} \quad (7)
\end{aligned}$$

We denote the corresponding estimation error by  $\mathbf{e}^k(t) := \hat{\mathbf{x}}_1^k(t) - \mathbf{x}_1(t)$ .

It will be shown that the optimal communication policy only depends on  $\mathbf{e}^0(t)$ . Note that  $\mathbf{e}^0(t)$  is completely determined by  $\mathcal{S}(t)$  and therefore it can be expressed as

$$\mathbf{e}^0(t) = \mathcal{F}(\mathcal{S}(t)), \quad \text{w.p.o.} \quad (8)$$

for an appropriately defined ‘‘deterministic’’ function  $\mathcal{F}$ .

From (4) and (7) with  $k = 0$ , we obtain

$$\mathbf{e}^0(t+1) = \begin{cases} A\mathbf{e}^0(t) - \mathbf{w}_1(t) & \text{if } \mathbf{a}_1(t) = 0 \\ -\mathbf{w}_1(t) & \text{if } \mathbf{a}_1(t) = 1, \end{cases} \quad (9)$$

which corresponds to a stationary transition probability from  $\mathbf{e} := \mathbf{e}^0(t)$  to  $\mathbf{y} := \mathbf{e}^0(t+1)$  given by

$$P(d\mathbf{y} | \mathbf{e}, u) = [u f(\mathbf{y}) + (1-u) f(\mathbf{y} - A\mathbf{e})] d\mathbf{y} \quad (10)$$

where

$$f(\mathbf{w}) := \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{-\frac{\mathbf{w}' \Sigma^{-1} \mathbf{w}}{2}\right\}, \quad \forall \mathbf{w} \in \mathbb{R}^n,$$

is the *p.d.f.* of the Gaussian white noise  $\mathbf{w}_1(t)$  with covariance matrix  $\Sigma$ .

To solve the original optimal communication problem, we first formulate an *auxiliary minimization problem* with the following AC criterion:

$$J(\pi) := \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}^\pi \sum_{t=0}^{N-1} c(\mathbf{e}^0(t), \mathbf{u}(t)),$$



The second equality above follows from Scheffe's Theorem [3], and  $p$  is the transition probability density function of  $P(\cdot|e, u)$  in (10). From (18) we have

$$\begin{aligned} & \|P(\cdot|e, u) - P(\cdot|e', u')\|_{TV} \\ &= \int p(y|e, u) + p(y|e', u') - 2 \min\{p(y|e, u), p(y|e', u')\} dy \\ &= 2 - 2 \int \min\{p(y|e, u), p(y|e', u')\} dy \quad (19) \end{aligned}$$

Since  $u \in \Pi_L(e)$  and  $u' \in \Pi_L(e')$ ,  $p(y|e, u) = f(y)$  when  $\|e\| > L$ , and  $p(y|e', u') = f(y)$  when  $\|e'\| > L$ . We consider three cases separately for (19),

$$\begin{aligned} & \|P(\cdot|e, u) - P(\cdot|e', u')\|_{TV} \\ &= 2 - 2 \begin{cases} \int \min\{f(y), f(y)\} dy & \text{if } \|e\|, \|e'\| \geq L \\ \int \min\{uf(y) + (1-u)f(y-Ae), f(y)\} dy & \text{if } \|e\| < L, \|e'\| \geq L \\ \int \min\{uf(y) + (1-u)f(y-Ae), \\ u'f(y) + (1-u')f(y-Ae')\} dy & \text{if } \|e\|, \|e'\| < L \end{cases} \\ &\leq 2 - 2 \begin{cases} 1 & \text{if } \|e\|, \|e'\| \geq L \\ \int \min\{f(y-Ae), f(y)\} dy & \text{if } \|e\| < L, \|e'\| \geq L \\ \int \min\{f(y), f(y-Ae)\} dy & \text{if } \|e\|, \|e'\| < L \end{cases} \\ &\leq 2 - 2 \int \min\{f(y), f(y-Ae)\} dy \end{aligned}$$

For  $\|e\| < L$ , there exists an  $\varepsilon > 0$  and a measurable set  $B$  with Lebesgue measure  $\mu(B) > 0$ , such that for all  $y \in B$ ,  $f(y) > \varepsilon$ ,  $f(y-Ae) > \varepsilon$ , and  $\varepsilon\mu(B) < 1$ . Therefore

$$\begin{aligned} & \|P(\cdot|e, u) - P(\cdot|e', u')\|_{TV} \\ &\leq 2 - 2 \int_B \min\{f(y), f(y-Ae)\} dy \leq 2 - 2\varepsilon\mu(B) \end{aligned}$$

which completes the proof by setting  $\beta = 1 - \varepsilon\mu(B)$ . ■

The *span semi-norm* of a bounded function  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  is defined by

$$\text{span}\{h\} := \sup_y h(y) - \inf_y h(y).$$

The following result follows from the ergodic property (17).

*Lemma 2 ([4]):* The dynamic programming operator  $T$  in (12) is a span-contraction, i.e., for any bounded functions  $g(\cdot)$  and  $h(\cdot)$ ,  $\text{span}\{Tg - Th\} \leq \beta \text{span}\{g - h\}$ , where  $\beta \in [0, 1)$  is given by Lemma 1.

We are now ready to prove Theorem 1.

*Proof:* [Theorem 1] From Lemma 2 we conclude that the operator  $T$  is a span-contraction and therefore, by Banach's Fixed-Point Theorem, there exists a bounded function  $\tilde{h}(\cdot)$  that is a fixed-point in the span sense, i.e.,  $\text{span}\{(T\tilde{h})(e) - \tilde{h}(e)\} = 0$ . Eq (14) holds for  $h(e) := \tilde{h}(e) - \tilde{h}(0)$ , which proves part one.

To finish the proof it remains to show that the minimum cost is equal to  $\bar{J}$  (part two) and that (16) is optimal (part three). To this effect, we re-write (14) as

$$\begin{aligned} & h(e) + \bar{J} \\ &= \min_{u \in \Pi_L(e)} u(\lambda - e'A'QAe - E[h(Ae + \mathbf{w})] + E[h(\mathbf{w})]) \\ &\quad + e'A'QAe + \rho^2 + E[h(Ae + \mathbf{w})], \end{aligned} \quad (20)$$

and note that the minimum is achieved for

$$u = \pi^*(e). \quad (21)$$

From (4) and (7), after straightforward algebraic derivations, we conclude that

$$\mathbf{e}^{k+1}(t+1) = A\mathbf{e}^k(t) - \mathbf{w}_1(t), \quad \forall 0 \leq k \leq \tau - 1,$$

from which, together with (9), we obtain that the estimation error at time  $t + \tau + 1$  satisfies

$$\mathbf{e}^\tau(t + \tau + 1) = \begin{cases} A^{\tau+1}\mathbf{e}^0(t) - \sum_{i=0}^{\tau} A^{\tau-i}\mathbf{w}_1(t+i) & \mathbf{a}(t) = 0 \\ -\sum_{i=0}^{\tau} A^{\tau-i}\mathbf{w}_1(t+i) & \mathbf{a}(t) = 1 \end{cases}$$

The optimizing criterion in (6) then becomes

$$\begin{aligned} J(\gamma) &= \lim_{N \rightarrow \infty} \frac{1}{N} E^\gamma \sum_{t=0}^{N-1} [(1 - \mathbf{u}(t))\mathbf{e}^0(t)'Q\mathbf{e}^0(t) + \lambda\mathbf{u}(t) + \rho^2] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} E^\gamma \sum_{t=0}^{N-1} [c(\mathbf{e}^0(t), \mathbf{u}(t))]. \end{aligned} \quad (22)$$

Since  $\mathbf{e}^0(t)$  is completely determined by  $\mathcal{I}(t)$ ,

$$\begin{aligned} & E[h(\mathbf{e}^0(t+1)) | \mathcal{I}(t), \mathbf{u}(t)] \\ &= E[h(\mathbf{e}^0(t+1)) | \mathbf{e}^0(t), \mathcal{I}(t), \mathbf{u}(t)] \\ &= \int h(y)P(dy | \mathbf{e}^0(t), \mathbf{u}(t)) \\ &\geq \bar{J} + h(\mathbf{e}^0(t)) - c(\mathbf{e}^0(t), \mathbf{u}(t)). \end{aligned} \quad (23)$$

where the second equality and the last inequality follow from (9) and (14), respectively. Moreover, because of (21), the inequality holds with equality when

$$\mathbf{u}(t) = \pi^*(\mathbf{e}^0(t)). \quad (24)$$

From (23), we then conclude that

$$\begin{aligned} & \sum_{t=0}^{N-2} h(\mathbf{e}^0(t+1)) - E[h(\mathbf{e}^0(t+1)) | \mathcal{I}(t), \mathbf{u}(t)] \\ &\leq -N\bar{J} + \sum_{t=0}^{N-1} c(\mathbf{e}^0(t), \mathbf{u}(t)) \\ &\quad + E[h(\mathbf{e}^0(N)) | \mathcal{I}(N-1), \mathbf{u}(N-1)] - h(\mathbf{e}^0(0)). \end{aligned}$$

Taking expectation for an arbitrary policy  $\gamma$ , we conclude that

$$0 \leq -N\bar{J} + E^\gamma \sum_{t=0}^{N-1} c(\mathbf{e}^0(t), \mathbf{u}(t)) + E^\gamma [h(\mathbf{e}^0(N)) - h(\mathbf{e}^0(0))].$$

Since  $h(\cdot)$  is bounded, we obtain from this and (22) that

$$J(\gamma) = \lim_{N \rightarrow \infty} \frac{1}{N} E^\gamma \left[ \sum_{t=0}^{N-1} c(\mathbf{e}^0(t), \mathbf{u}(t)) \right] \geq \bar{J}$$

with equality when (24) holds. This completes the proof because it shows that  $\gamma^*$  is optimal and that the minimum cost is indeed  $\bar{J}$ . By setting  $e = 0$  in (20), we obtain (15). ■

#### D. Value iteration

Eq. (14) can be solved efficiently using a *value iteration algorithm* defined as follows

- 1)  $h^0(e) = 0, \forall e$ ;
- 2) The sequence of functions  $h^k : \mathbb{R}^n \rightarrow \mathbb{R}, J^k : \mathbb{R}^n \rightarrow \mathbb{R}, \forall t \geq 0$  are defined by the following iteration:
  - $\bar{h}^{t+1} := Th^k$
  - $J^{t+1} := \bar{h}^{t+1} - h^k$
  - $h^{t+1}(e) := \bar{h}^{t+1}(e) - \bar{h}^{t+1}(0), \forall e$

*Theorem 2:* The sequence  $(J^k, h^k)$  converges exponentially fast to the pair  $(\bar{J}, h)$ , whose existence is guaranteed by Theorem 1. In particular,

$$\|h^k - h\|_\infty \leq \beta^k \cdot \text{span}\{h\} \quad (25)$$

$$\|J^k - \bar{J}\|_\infty \leq 3\beta^{k-1} \cdot \text{span}\{h\}, \quad (26)$$

where  $\|\cdot\|_\infty$  denotes the sup-norm of a function.

*Proof:* We start by noting for any given function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \|g\|_\infty &\leq \|g - g(x)\|_\infty + |g(x)| \\ &\leq \text{span}\{g\} + |g(x)|, \quad \forall x \in \mathbb{R}^n. \end{aligned} \quad (27)$$

From (14),  $\text{span}\{h\} = \text{span}\{Th\}$ .

$$\begin{aligned} \text{span}\{h^k - h\} &= \text{span}\{Th^{k-1} - Th\} \leq \beta \text{span}\{h^{k-1} - h\} \\ &\leq \beta^k \text{span}\{h^0 - h\} = \beta^k \text{span}\{h\}, \end{aligned}$$

from which the inequality (25) follows using (27) and the fact that  $h^k(0) = h(0) = 0$ . To prove (26), we use the fact that

$$\begin{aligned} \text{span}\{J^k - \bar{J}\} &= \text{span}\{Th^{k-1} - h^{k-1} - (Th - h)\} \\ &= \text{span}\{Th^{k-1} - Th - (h^{k-1} - h)\} \\ &\leq \text{span}\{Th^{k-1} - Th\} + \text{span}\{h^{k-1} - h\} \\ &\leq (\beta^k + \beta^{k-1}) \text{span}\{h\}. \end{aligned} \quad (28)$$

Moreover, we conclude from the definitions of  $J^k$  and  $T$ , and (15) that

$$\begin{aligned} J^k(0) &= (Th^{k-1})(0) - h^{k-1}(0) \\ &= \rho^2 + E[h^{k-1}(\mathbf{w})] = \bar{J} + E[h^{k-1}(\mathbf{w}) - h(\mathbf{w})] \end{aligned}$$

and therefore

$$|J^k(0) - \bar{J}| = |E[h^{k-1}(\mathbf{w}) - h(\mathbf{w})]| \leq \beta^{k-1} \text{span}\{h(\cdot)\}.$$

Inequality (26) follows using (27), (28), and this. ■

## IV. CONCLUSIONS AND FUTURE WORK

This paper formulates optimal communication in the context of NCS as a long term average cost optimization problem that can be solved using dynamic programming. For discrete-time LTI systems with fixed network delays, we prove that the best logic is deterministic and provide the optimal rules. Future work includes the study of optimal logics for continuous systems.

## ACKNOWLEDGMENTS

Part of the material in this paper was in the final project for the course *Stochastic Control* taught by Professor Karl Åström in Fall of 2003 at the Univ. of California, Santa Barbara.

## REFERENCES

- [1] A. Araphstathis and et al. Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM J. Contr. Optimization*, 31(2):282–344, Mar. 1993.
- [2] D. P. Bertsekas. *Dynamic programming and optimal control*, volume 2, pages 191–192. Athena Scientific, Belmont, Mass, 1995.
- [3] L. Devroye and L. Györfi. *Nonparametric density estimation: The  $L_1$  view*. Wiley, New York, 1985.
- [4] O. Hernández-Lerma. *Adaptive Markov Control Processes*, volume 79 of *Applied mathematical sciences*. Springer-Verlag, New York, 1989.
- [5] O. Hernández-Lerma and J. B. Lasserre. *Discrete-time Markov Control Processes*, volume 30 of *Applications of Mathematics*. Springer, Berlin, 1996.
- [6] J. P. Hespanha, A. Ortega, and L. Vasudevan. Towards the control of linear systems with minimum bit-rate. In *Proc. of the Int. Symp. on the Mathematical Theory of Networks and Syst.*, Aug. 2002.
- [7] G. N. Nair and R. J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39(4):585–593, Apr. 2003.
- [8] W. Rudin. *Real and complex analysis*. McGraw-Hill, 3 edition, 1987.
- [9] S. Tatikonda. *Control Under Communication Constraints*. PhD thesis, MIT, Cambridge, MA, 2000.
- [10] Y. Xu and J. P. Hespanha. Communication logics for networked control systems. In *Proc. of the 2004 Amer. Contr. Conf.*, pages 572–577, June 2004.
- [11] J. K. Yook, D. M. Tilbury, and N. R. Soparkar. Trading computation for bandwidth: Reducing communication in distributed control systems using state estimators. *IEEE Trans. Contr. Syst. Technol.*, 10(4):503–518, July 2002.