Forty-Fourth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
Sept 27-29, 2006

A Constant Factor Approximation Algorithm for Event-Based Sampling

Randy Cogill, Sanjay Lall and João P. Hespanha

Abstract— **We consider a control system in which sensor data is transmitted from the plant to a receiver over a communication channel, and the receiver uses the data to estimate the state of the plant. Using a feedback policy to choose when to transmit data, the goal is to schedule transmissions to balance a trade-off between communication rate and estimation error. Computing an optimal policy for this problem is generally computationally intensive. Here we provide a simple algorithm for computing a suboptimal policy for scheduling state transmissions which incurs a cost within a factor of six of the optimal achievable cost.**

I. INTRODUCTION

We consider a control system in which sensor data is transmitted from the plant to a receiver over a communication channel, and the receiver uses the data to estimate the state of the plant. Sending data more frequently leads to increased use of limited communication resources, but also allows the average estimation error to be reduced. Conversely of course we may reduce the use of the channel if we are willing to allow larger estimation errors.

We consider feedback policies for choosing when to transmit data. That is, instead of simply choosing a transmission rate, at the plant measurements are used to decide whether to transmit data to the controller. This type of measurement is called *Lebesgue* or *event-based* sampling in [1]. Several other authors have considered both control and filtering problems using such sampling schemes, in particular [1]–[7].

The plant is modeled by a discrete-time linear system, and at each time step the channel allows exact transmission of the state. The cost function of interest in this prolem is a weighted sum of the estimation error and the transmission rate. The optimal controller for a given weight then lies on the Pareto-optimal trade-off curve, and choosing the weight allows one to select the trade-off between rate and error.

For this cost function, the problem of finding the optimal policy was considered in [8], where the authors show that the problem of computing an optimal scheduling policy can be addressed in the framework of Markov decision processes, and consequently the value iteration algorithm can be used to compute an optimal policy. Although this provides an algorithm for computing an optimal policy, the computation required to compute such a policy quickly becomes prohibitive as the system's state dimension increases.

Randy Cogill is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305, U.S.A. Email: rcogill@stanford.edu

Sanjay Lall is with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305, U.S.A. Email: lall@stanford.edu

João Hespanha is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, U.S.A. Email: hespanha@ece.ucsb.edu

Since the optimal policy is very difficult to compute, we consider *approximately optimal* policies. Specifically, the main result of this paper is to give a simple algorithm for computing a policy, and show that this policy is guaranteed to achieve a cost within a factor of six of the optimal achievable cost. This result is Theorem 2 below.

Approximation algorithms have been widely used for addressing computationally intractable problems [9]. While some NP-hard problems may be approximated to arbitrary accuracy, others may not be approximated within any constant factor. It is therefore extremely promising that the particular problem of rate-error trade-off considered in this paper is approximable within a constant factor of six. It is not currently known whether policies achieving better approximation ratios may be efficiently obtained.

II. PROBLEM FORMULATION

Here we will present the problem that will be considered throughout this paper. In the following subsection, it will be shown how this problem is a generalization of the problem of networked estimation.

We have dynamics

$$e_{t+1} = (1 - a_t)Ae_t + w_t \qquad e_0 = 0 \tag{1}$$

where for each $t \in \mathbb{N}$ the state $e_t \in \mathbb{R}^n$, and the action $a_t \in \{0,1\}$. Here $w_0, w_1, \dots$ is a sequence of independent identically distributed Gaussian random vectors, with $w_t \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \succ 0$. Define the function $r : \mathbb{R}^n \times \{0,1\} \to \mathbb{R}$ to be the cost at time t, given by

$$r(e_t, a_t) = (1 - a_t)e_t^T Q e_t + \lambda a_t \tag{2}$$

where $Q \succ 0$ and $\lambda > 0$. We would like to choose a state-feedback control policy $\mu : \mathbb{R}^n \to \{0,1\}$ to make the average cost incurred by the policy μ small. Here the average cost J is defined as

$$J(\mu) = \limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\big(r(e_t, \mu(e_t))\big) \tag{3}$$

See [10] for background on this choice of cost. Here, each a_t is determined according to the static state-feedback policy $a_t = \mu(e_t)$, and then the sequence $e_0, e_1, \dots$ is a Markov process. Therefore, the problem of choosing a policy which minimizes the cost J is can be addressed using the theory of Markov decision processes. The cost J given by equation (3) is called the *average per-period cost*, and we focus specifically on the problem of choosing a policy to minimize this. For convenience, define the space of policies

$$\mathcal{P} = \{\, f : \mathbb{R}^n \to \{0,1\} \mid f \text{ is measurable} \,\}$$

Then the above problem can be stated as follows.

*Problem 1 (*RATE-ERROR TRADE-OFF*):* Given A, $\Sigma \succ 0$, $Q \succ 0$, $\lambda > 0$, and $\gamma > 0$, find a state feedback policy $\mu \in \mathcal{P}$ such that

$$J(\mu) \leq \gamma$$

Minimizing the cost J balances a trade-off between the average size of e_t, as measured by the quadratic form defined by Q, and the frequency with which e_t is reset to the level of the noise by setting $a_t = 1$. The problem of computing an optimal policy was considered in [8], and a numerical procedure for finding such a policy was given. However, the computation required to compute an optimal policy increases rapidly with the state dimension. In the following section we present an easily computable and easily implementable policy for this problem which incurs a cost within a provable bound of the optimal achievable cost. Specifically, we focus our attention on the set of problem instances where Q and A are such that $A^TQA - Q \preceq 0$ and $Q \succ 0$. In particular, this implies that $\rho(A) \leq 1$ and the system is therefore at least marginally stable. We show that in this case there is a simple policy which always achieves a cost within a factor of six of the optimal cost. It is worth noting that, in general, both the policy which always transmits and the policy which never transmits may achieve cost arbitrarily far from optimal.

A. Application to Networked Estimation

Suppose we have the dynamical system

$$\begin{aligned} x_{t+1} &= Ax_t + w_t \quad x_0 = 0 \\ y_t &= a_t x_t \end{aligned}$$

where for each $t \in \mathbb{N}$ the state $x_t \in \mathbb{R}^n$ and $a_t \in \{0,1\}$. As above, $w_0, w_1, \dots$ is a sequence of independent identically distributed zero mean Gaussian random vectors with covariance $\Sigma \succ 0$. We have a per-period cost of

$$c(x_t, a_t, b_t) = (1 - a_t)(x_t - b_t)^T Q (x_t - b_t) + \lambda a_t \qquad (4)$$

and we would like to choose two controllers. The first is the function $\mu : \mathbb{R}^n \to \{0,1\}$, and the second is the sequence of functions ϕ_t indexed by t where $\phi_t : \{0,1\}^t \times \mathbb{R}^{nt} \to \mathbb{R}^n$. These are connected according to

$$\begin{aligned} a_t &= \mu(x_t) \\ b_t &= \phi_t(a_0, \dots, a_{t-1}, y_0, \dots, y_{t-1}) \end{aligned}$$

Again, we are interested in the cost

$$J(\mu, \phi_0, \phi_1, \dots) = \limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\big(r(x_t, a_t, b_t)\big)$$

The interpretation is shown in Figure 1, where the linear dynamics $x_{t+1} = Ax_t + w_t$ is denoted by G. The dashed lines indicate a communication channel. At each time t the transmitter μ chooses whether to transmit the signal x_t to the receiver ϕ. Each transmission costs λ. The receiver would like to estimate the state x_t of G, and choose b_t to minimize the error $x_t - b_t$ as measured by the quadratic form Q. The cost r is used to compute the trade-off, parametrized by λ, of estimation error against frequency of transmissions.

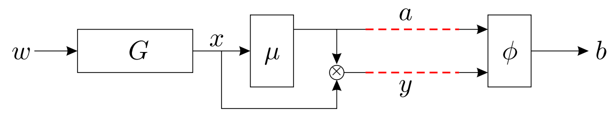

Fig. 1. Networked Estimation

The estimator ϕ considered in Xu and Hespanha [8] is as follows. Let $b_t = \phi_t(a_0, \dots, a_{t-1}, y_0, \dots, y_{t-1})$, and define ϕ by the realization

$$b_{t+1} = (1 - a_t)Ab_t + a_t A y_t \quad b_0 = 0$$

If the random variables $a_0, a_1, \dots$ are independent of $x_0, x_1, \dots$ then this is the time-varying Kalman filter, and b_t is the minimum mean square error estimate of x_t given measurements $y_0, \dots, y_{t-1}$.

We now have the dynamics

$$\begin{bmatrix} x_{t+1} \\ b_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ a_t A & (1-a_t)A \end{bmatrix} \begin{bmatrix} x_t \\ b_t \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} w_t$$

We change coordinates to

$$\begin{bmatrix} e_t \\ f_t \end{bmatrix} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix} \begin{bmatrix} x_t \\ b_t \end{bmatrix}$$

to give

$$\begin{bmatrix} e_{t+1} \\ f_{t+1} \end{bmatrix} = \begin{bmatrix} (1-a_t)A & 0 \\ a_t A & A \end{bmatrix} \begin{bmatrix} e_t \\ f_t \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} w_t$$

In these coordinates, the cost c specified in equation (4) is exactly equal to the cost (2), and e evolves according to the dynamics (1). With this choice of ϕ therefore the optimal choice of μ is found by solving the RATE-ERROR TRADE-OFF problem.

III. MAIN RESULTS

In this section we present the main result of this paper, which is that for a slightly restricted version of the RATE-ERROR TRADEOFF problem, there is a simple policy which achieves cost within a constant factor of optimal. Define for convenience

$$J_{\text{opt}} = \inf_{\mu\in\mathcal{P}} \left(\liminf_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\big(r(e_t, \mu(e_t))\big) \right)$$

The policy that we consider is a simple quadratic threshold policy. The main result of this paper is as follows.

Theorem 2: Suppose $A \in \mathbb{R}^{n\times n}$, $Q \succ 0$, $\Sigma \succ 0$, and $A^TQA - A \preceq 0$. Then there exists a unique matrix $M \in \mathbb{S}^n$ satisfying

$$\frac{1}{1 + \operatorname{trace}(\Sigma M)} A^T M A - M + \frac{Q}{\lambda} = 0$$

Furthermore, define the policy μ by

$$\mu(e) = \begin{cases} 0 & \text{if } e^T M e \leq 1 \\ 1 & \text{otherwise} \end{cases} \qquad (5)$$

For this policy, the cost satisfies

$$J(\mu) \leq 6 J_{\text{opt}} \tag{6}$$

Proof: The result follows immediately from Theorems 7 and 12 which are proved below. ■

We determine an upper bound on $J(\mu)$ and a lower bound on J_{opt}. It is then easily shown that the upper and lower bounds differ by a factor of six. First we prove the required existence and uniqueness result.

Lemma 3: Suppose $A \in \mathbb{R}^{n\times n}$, $Q \succ 0$, $\Sigma \succ 0$ and $A^T QA - Q \preceq 0$. Then there exists a unique matrix $M \in \mathbb{S}^n$ such that

$$\frac{1}{1 + \operatorname{trace}(\Sigma M)} A^T MA - M + Q = 0 \tag{7}$$

and this solution satisfies $M \succ 0$.

Proof: Since $Q \succ 0$ and $A^T QA - Q \preceq 0$ the standard properties of Lyapunov equations imply that we have $\rho(A) \leq 1$, and hence for any α with $0 \leq \alpha < 1$ the equation

$$\alpha A^T MA - M + Q = 0$$

has a unique solution $M \succ 0$. Define the map $f : [0, 1) \to \mathbb{S}^n$ so that $f(\alpha)$ is this unique solution. Further define the map $h : [0, 1) \to \mathbb{R}$ by

$$h(\alpha) = \frac{1}{1 + \operatorname{trace}\big(\Sigma f(\alpha)\big)} - \alpha$$

Now $h(0) > 0$, since it is given by $h(0) = 1/(1 + \operatorname{trace}(\Sigma Q))$. Also, for all sufficiently small $\delta > 0$, we have $h(1 - \delta) < 0$. To see this, notice that $f(\alpha) \succeq Q \succ 0$ for all $\alpha \in [0, 1)$ and since $\Sigma \succ 0$ we have for all $\alpha \in [0, 1)$

$$\operatorname{trace}\big(\Sigma f(\alpha)\big) \geq \operatorname{trace}(\Sigma Q) > 0$$

Hence for $\delta > 0$ sufficiently small we have

$$\begin{aligned} h(1 - \delta) &= \frac{1}{1 + \operatorname{trace}\big(\Sigma f(1 - \delta)\big)} - 1 + \delta \\ &< \frac{1}{1 + \operatorname{trace}\big(\Sigma Q\big)} - 1 + \delta \\ &= \delta - \frac{\operatorname{trace}(\Sigma Q)}{1 + \operatorname{trace} \Sigma Q} \\ &< 0 \end{aligned}$$

Now the function h is continuous on $[0, 1 - \delta)$ and therefore there must exist some $\alpha_0 \in [0, 1 - \delta)$ such that $h(\alpha_0) = 0$. Now let $M = f(\alpha_0)$ and we immediately have $M \succ 0$ and M satisfies (7).

To show uniqueness, suppose $M_1 \in \mathbb{S}^n$ and $M_2 \in \mathbb{S}^n$ are two distinct solutions. Let

$$\beta_i = \frac{1}{1 + \operatorname{trace}(\Sigma M_i)}$$

and since M_i is a solution we have $f(\beta_i) = M_i$ and hence $h(\beta_i) = 0$. But equation $h(x) = 0$ has a unique root, since h is strictly decreasing, and hence $\beta_1 = \beta_2$ and hence $M_1 = M_2$. ■

Note that the above Lemma gives an algorithm for computing the unique solution M of equation (7). All that is needed is to perform a bisection search to find the root of h.

IV. BOUNDS FOR GENERAL MARKOV PROCESSES

The main tools that are used to determine the upper and lower bounds are the following two Lemmas. These are special cases of the more general results found in [11].

Lemma 4: Suppose $x_0, x_1, \ldots$ is a Markov process, with each $x_t \colon \Omega \to \mathcal{X}$. Suppose $r \colon \mathcal{X} \to \mathbb{R}$ and $h \colon \mathcal{X} \to \mathbb{R}$. Define

$$\overline{J} = \limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\big(r(x_t)\big)$$

If there exists $a \in \mathbb{R}$ such that

$$h(x) \geq a \quad \text{for all } x \in \mathcal{X}$$

then

$$\overline{J} \leq \sup_{q\in\mathcal{X}} \Big(r(q) + \mathrm{E}\big(h(x_{t+1}) \,|\, x_t = q\big) - h(q) \Big)$$

Proof: Define the function $\Delta : \mathcal{X} \to \mathbb{R}$ by

$$\Delta(q) = \mathrm{E}\big(h(x_{t+1}) \,|\, x_t = q\big) - h(q)$$

and let $\beta \in \mathbb{R}$ be the candidate bound

$$\beta = \sup_{x\in\mathcal{X}} \big(r(x) + \Delta(x)\big)$$

For convenience also define $y = \inf_{x\in\mathcal{X}} h(x)$. Then for any $z \in \mathcal{X}$ and $N > 0$, we have

$$\begin{aligned} &\frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t) \,|\, x_0 = z\big) \\ &\qquad = \frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t) + \Delta(x_t) \,|\, x_0 = z\big) \\ &\qquad\qquad + \frac{h(z) - y}{N} - \frac{\mathrm{E}\big(h(x_{t+1}) \,|\, x_0 = z\big) - y}{N} \end{aligned}$$

and using the hypothesis we have

$$\frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t) \,|\, x_0 = z\big) \leq \beta + \frac{h(z) - y}{N}$$

Taking the limit superior as $N \to \infty$ on both sides gives

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t) \,|\, x_0 = z\big) \leq \beta$$

and since this holds for all $z \in \mathcal{X}$ taking the expected value over initial states gives the desired result. ■

The above result will be used to give an upper bound on the cost incurred by a given policy μ. To find a lower bound on the cost achievable by any policy, we will use the following extension to Markov decision processes.

Lemma 5: Consider a Markov decision process such that for any policy $\mu \colon \mathcal{X} \to \mathcal{A}$ we have states $x_t \colon \Omega \to \mathcal{X}$ and actions $a_t \colon \Omega \to \mathcal{A}$. Suppose $r \colon \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ and $h \colon \mathcal{X} \to \mathbb{R}$. Define the cost

$$\underline{J}(\mu) = \liminf_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\big(r(x_t, \mu(x_t))\big)$$

Then if there exists $b \in \mathbb{R}$ such that

$$h(x) \le b \quad \text{for all } x \in \mathcal{X}$$

then for all $\mu\colon \mathcal{X} \to \mathcal{A}$ we have

$$\underline{J}(\mu) \ge \inf_{q\in\mathcal{X}, u\in\mathcal{A}} \Big(r(q,u) + E\big(h(x_{t+1}) \,|\, x_t = q, a_t = u\big) - h(q) \Big)$$

Proof: Define the function $\Delta : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ by

$$\Delta(q,u) = E\big(h(x_{t+1}) \,|\, x_t = q, a_t = u\big) - h(q)$$

and let β be

$$\beta = \inf_{q\in\mathcal{X}, u\in\mathcal{A}} \big(r(q,u) + \Delta(q,u)\big)$$

For convenience define $y = \sup\{h(x) \,|\, x \in \mathcal{X}\}$. Then for any $\mu \in \mathcal{P}$, $z \in \mathcal{X}$ and $N > 0$, we have

$$\begin{aligned} &\frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t, \mu(x_t)) \,|\, x_0 = z\big) \\ &= \frac{1}{N} \sum_{t=0}^{N-1} E\Big(r\big(x_t, \mu(x_t)\big) + \Delta\big(x_t, \mu(x_t)\big) \,\Big|\, x_0 = z\Big) \\ &\quad + \frac{h(z) - y}{N} - \frac{E\big(h(x_{t+1}) \,|\, x_0 = z, a_0 = \mu(z)\big) - y}{N} \end{aligned}$$

Using the hypothesis we have

$$\frac{1}{N} \sum_{t=0}^{N-1} E\big(r(x_t, \mu(x_t)) \,|\, x_0 = z\big) \ge \beta + \frac{h(z) - y}{N}$$

and taking the limit inferior as $N \to \infty$ we have

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathrm{E}\Big(r\big(x_t, \mu(x_t)\big) \,\Big|\, x_0 = z\Big) \ge \beta$$

for all $z \in \mathcal{X}$. Again, taking expectations over the initial state gives the desired result. ∎

V. Bounds for the Communication Cost

A. Upper bounds

We are now ready to prove the upper bound on $J(\mu)$ where μ is the policy in (5). The following lemma provides the upper bound and also shows that one may use semidefinite programming, combined with a line search, to find policies that minimize this upper bound.

Lemma 6: Suppose $M \succeq 0$ and $H \succeq 0$ are symmetric positive semidefinite matrices, and $\alpha \in \mathbb{R}$. If

$$\begin{aligned} A^T H A - H + Q - \alpha M &\preceq 0 \\ (\lambda - \alpha) M - H &\preceq 0 \\ \alpha - \lambda &\le 0 \\ \alpha &\ge 0 \end{aligned} \tag{8}$$

Then the policy

$$\mu(e) = \begin{cases} 0 & \text{if } e^T M e \le 1 \\ 1 & \text{otherwise} \end{cases}$$

achieves a cost which satisfies

$$J(\mu) \le \operatorname{trace}(\Sigma H) + \alpha$$

Proof: The proof makes use of Lemma 4 with the function

$$h(e) = e^T H e$$

Clearly this choice of h has $h(e) \ge 0$ for all e. With the above policy the dynamics are

$$e_{t+1} = \begin{cases} A e_t + w_t & \text{if } e_t^T M e_t \le 1 \\ w_t & \text{otherwise} \end{cases}$$

We now use the expected values computed in Lemma 8 to give

$$\begin{aligned} &\mathrm{E}\big(h(e_{t+1}) \,|\, e_t = q\big) \\ &\qquad = \begin{cases} \operatorname{trace}(\Sigma H) + q^T A^T H A q & \text{if } q^T M q \le 1 \\ \operatorname{trace}(\Sigma H) & \text{otherwise} \end{cases} \end{aligned}$$

Now let the function $f\colon \mathbb{R}^n \to \mathbb{R}$ be

$$f(q) = \mathrm{E}\big(h(e_{t+1}) \,|\, e_t = q\big) - h(q) + r\big(q, \mu(q)\big)$$

where μ is as in the hypothesis of the lemma. Then we have

$$\begin{aligned} &f(q) - \operatorname{trace}(\Sigma H) \\ &\qquad = \begin{cases} q^T (A^T H A - H + Q) q & \text{if } q^T M q \le 1 \\ \lambda - q^T H q & \text{otherwise} \end{cases} \end{aligned}$$

We now use Lagrange duality to express the above quadratic inequalities as follows. For any symmetric matrix $X \in \mathbb{R}^{n\times n}$, if there exists $\alpha \in \mathbb{R}$ such that

$$\begin{aligned} X &\preceq \alpha M \\ \alpha &\ge 0 \end{aligned}$$

then $q^T X q \le \alpha$ for all q such that $q^T M q \le 1$, since in that case $q^T X q \le \alpha q^T M q \le \alpha$. Similarly, if

$$\begin{aligned} (\lambda - \alpha) M &\preceq H \\ \alpha - \lambda &\le 0 \end{aligned}$$

then $\lambda - q^T H q \le \alpha$ for all q such that $q^T M q \ge 1$. Hence if conditions (8) hold, then Lemma 4 implies that

$$\begin{aligned} J(\mu) &\le \sup_{q\in\mathbb{R}^n} f(q) \\ &\le \operatorname{trace}(\Sigma H) + \alpha \end{aligned}$$

as desired. ∎

We now make use of this result to provide an explicit upper bound.

Theorem 7: Suppose $A \in \mathbb{R}^{n\times n}$, $Q \succ 0$, $\Sigma \succ 0$ and $A^T Q A - Q \preceq 0$. Let M be the unique solution to

$$\frac{1}{1 + \operatorname{trace}(\Sigma M)} A^T M A - M + Q/\lambda = 0$$

Then the policy

$$\mu(e) = \begin{cases} 0 & \text{if } e^T M e \le 1 \\ 1 & \text{otherwise} \end{cases}$$

achieves

$$J(\mu) \le \frac{2\lambda \operatorname{trace}(\Sigma M)}{1 + \operatorname{trace}(\Sigma M)}$$

Proof: We use Lemma 6 as follows. For convenience, let $d = 1 + \operatorname{trace}(\Sigma M)$ and pick $H = \lambda M/d$ and $\alpha = \lambda(d-1)/d$. Then it is immediately verified that H, M, α satisfy (8), and hence we have

$$\begin{aligned} J(\mu) &\le \operatorname{trace}(\Sigma H) + \alpha \\ &= 2\lambda \operatorname{trace}(\Sigma M)/d \end{aligned}$$

as desired. ∎

B. Lower bounds

For the class of instances of RATE-ERROR TRADEOFF with A and Q satisfying $A^T Q A - Q \preceq 0$, we can show that the policy μ of equation (5) achieves a cost within a constant factor of optimal. To complete the presentation of the main result of this paper, we now determine a lower bound on J_{opt} which guarantees that for this class of instances,

$$J(\mu) \le 6 J_{\text{opt}}$$

We first prove some preliminary results.

Lemma 8: Suppose $Y \succeq 0$ and $q \in \mathbb{R}^n$, and $w \sim \mathcal{N}(0, \Sigma)$ is a Gaussian random vector. Let f be the random variable

$$f = (q + w)^T Y (q + w)$$

Then

$$\operatorname{E} f = q^T Y q + \operatorname{trace}(\Sigma Y) \tag{9}$$

$$\begin{aligned} \operatorname{E}(f^2) = (q^T Y q)^2 + 4 q^T Y \Sigma Y q + \big(\operatorname{trace}(\Sigma Y)\big)^2 \\ + 2 \operatorname{trace}(\Sigma Y \Sigma Y) + 2 q^T Y q \operatorname{trace}(\Sigma Y) \end{aligned} \tag{10}$$

and further

$$\operatorname{E}(f^2) \le (q^T Y q)^2 + 6 q^T Y q \operatorname{trace}(\Sigma Y) + 3\big(\operatorname{trace}(\Sigma Y)\big)^2$$

Proof: Equation (9) holds by expanding the quadratic and using linearity of expectation, since

$$\begin{aligned} \operatorname{E}(w^T Y w) &= \operatorname{E} \operatorname{trace}(w^T Y w) \\ &= \operatorname{E} \operatorname{trace}(Y w w^T) \\ &= \operatorname{trace}\big(Y \operatorname{E}(w w^T)\big) \end{aligned}$$

and $\operatorname{trace}(AB) = \operatorname{trace}(BA)$ for all compatible matrices A, B. For equation (10), expanding gives

$$\begin{aligned} \operatorname{E}(f^2) = \operatorname{E}\Big(& (q^T Y q)^2 + 4 (q^T Y w)^2 \\ & + (w^T Y w)^2 + 4 (q^T Y q) q^T w \\ & + 4 (w^T Y w) q^T Y w + 2 (q^T Y q) w^T Y w \Big) \end{aligned}$$

We evaluate each of these terms. The second term is

$$\operatorname{E}\big((q^T Y w)^2\big) + \operatorname{E} \operatorname{trace}(q^T Y w w^T Y q) = q^T Y \Sigma Y q$$

For the third term, recall that if $y \sim \mathcal{N}(0, Q)$ then the fourth-order moments (see for example [12]) are given by

$$\operatorname{E}\big((y^T y)^2\big) = (\operatorname{trace} Q)^2 + 2 \operatorname{trace}(Q^2)$$

Now let $y = Y^{\frac{1}{2}} w$, then $y \sim \mathcal{N}(0, Y^{\frac{1}{2}} \Sigma Y^{\frac{1}{2}})$ and hence

$$\begin{aligned} \operatorname{E}\big((w^T Y w)^2\big) &= \operatorname{E}\big((y^T y)^2\big) \\ &= \big(\operatorname{trace}(\Sigma Y)\big)^2 + 2 \operatorname{trace}(\Sigma Y \Sigma Y) \end{aligned}$$

For the fourth term $\operatorname{E}(q^T Y q) q^T w = 0$ since $\operatorname{E} w = 0$. For the fifth term

$$\operatorname{E}\big((w^T Y w) q^T Y w\big) = \sum_{i,j,k} Y_{ij} (Y q)_k \operatorname{E}(w_i w_j w_k) = 0$$

since the Gaussian density is symmetric. Summing these terms gives equation (10).

For any square matrices $A \succeq 0$ and $B \succeq 0$ we have $\operatorname{trace}(AB) \le (\operatorname{trace} A)(\operatorname{trace} B)$, and also

$$\begin{aligned} q^T Y \Sigma Y q &= \operatorname{trace}(q^T Y \Sigma Y q) \\ &= \operatorname{trace}(q q^T Y \Sigma Y) \\ &\le q^T Y q \operatorname{trace}(\Sigma Y) \end{aligned}$$

The final inequality then follows. ∎

Lemma 9: Suppose there exists a positive semidefinite matrix $C \succeq 0$ and $s \in \mathbb{R}$ such that

$$\begin{aligned} \big(s - 6 \operatorname{trace}(C\Sigma)\big) A^T C A - sC + Q \succeq 0 \\ s^2 \le 4\lambda \\ A^T C A - C \preceq 0 \end{aligned} \tag{11}$$

Then for all policies $\mu \in \mathcal{P}$

$$J(\mu) \ge s \operatorname{trace}(C\Sigma) - 3\big(\operatorname{trace}(C\Sigma)\big)^2$$

Proof: Let the function h be

$$h(e) = s e^T C e - (e^T C e)^2$$

It is easily verified that $h(e)$ is bounded above. For convenience let $r = \operatorname{trace}(C\Sigma)$. Then since $C \succeq 0$ Lemma 8 gives

$$\begin{aligned} \operatorname{E}\big(h(e_{t+1} \,|\, e_t = q, a_t = 0\big) \ge sr + (s - 6r) q^T A^T C A q \\ - 3r^2 - (q^T A^T C A q)^2 \end{aligned}$$

and

$$\operatorname{E}\big(h(e_{t+1} \,|\, e_t = q, a_t = 1\big) = r(s - 3r)$$

Now let f be the function

$$f(q, u) = \operatorname{E}\big(h(e_{t+1} \,|\, e_t = q, a_t = u\big) - h(q) + r(q, u)$$

In order to apply Lemma 5 we need to compute a lower bound for f. First, we have

$$\begin{aligned} f(q, 1) &= s \operatorname{trace}(C\Sigma) - 3r^2 - s q^T C q + (q^T C q)^2 + \lambda \\ &= s \operatorname{trace}(C\Sigma) - 3r^2 + (q^T C q - s/2)^2 + \lambda - s^2/4 \end{aligned}$$

Since by hypothesis we have $s^2 \le 4\lambda$, we have

$$f(q, 1) \ge s \operatorname{trace}(C\Sigma) - 3r^2$$

Also we have

$$\begin{aligned} f(q, 0) \ge s \operatorname{trace}(C\Sigma) - 3r^2 + (s - 6r) q^T A^T C A q \\ - (q^T A^T C A q)^2 - s q^T C q + (q^T C q)^2 + q^T Q q \end{aligned}$$

Since by the hypothesis $A^TCA - C \preceq 0$ we have

$$(q^TA^TCAq)^2 - (q^TCq)^2 \leq 0 \quad \text{for all } q \in \mathbb{R}^n$$

and hence

$$\begin{aligned} f(q,0) \geq s\,\mathrm{trace}(C\Sigma) - 3r^2 \\ + q^T\big((s-6r)A^TCA - sC + Q\big)q \end{aligned}$$

and hence if inequalities (11) hold

$$f(q,u) \geq s\,\mathrm{trace}(C\Sigma) - 3r^2 \quad \text{for all } q \in \mathbb{R}^n, u \in \{0,1\}$$

as desired. ∎

Lemma 10: Suppose there exists $M \succeq 0$ such that

$$\frac{1}{1+\mathrm{trace}(\Sigma M)}A^TMA - M + Q/\lambda = 0$$
$$A^TMA - M \preceq 0$$

Then for all policies $\mu \in \mathcal{P}$ we have

$$J(\mu) \geq \frac{\lambda\,\mathrm{trace}(\Sigma M)}{3\big(1+\mathrm{trace}(\Sigma M)\big)}$$

Proof: Denote for convenience $d = 1 + \mathrm{trace}(\Sigma M)$. We will use Lemma 9 with

$$C = \frac{\sqrt{\lambda}M}{3d}$$
$$s = 2\sqrt{\lambda}$$

We also denote $r = \mathrm{trace}(C\Sigma)$. Then some algebra gives

$$\begin{aligned} &\big(s - 6\,\mathrm{trace}(C\Sigma)\big)A^TCA - sC + Q \\ &\quad = \frac{2}{3d}\left(\frac{1}{d}A^TMA - M + \frac{Q}{\lambda}\right) + \left(1 - \frac{2}{3d}\right)\frac{Q}{\lambda} \\ &\quad = \left(1 - \frac{2}{3d}\right)\frac{Q}{\lambda} \\ &\quad \succeq 0 \end{aligned}$$

since $d \geq 1$. Hence inequalities 11 are satisfied, and Lemma 9 implies that for all policies $\mu \in \mathcal{P}$ we have

$$\begin{aligned} J(\mu) &\geq s\,\mathrm{trace}(C\Sigma) - 3\big(\mathrm{trace}(C\Sigma)\big)^2 \\ &= \frac{2\lambda(d-1)}{3d} - \frac{\lambda(d-1)^2}{3d^2} \\ &= \frac{\lambda(d-1)}{3d^2} \\ &\geq \frac{\lambda(d-1)}{3d} \end{aligned}$$

as desired, since $d \geq 1$. ∎

Lemma 11: Suppose $Q \succ 0$ and $A^TQA - Q \preceq 0$, and $\alpha \in \mathbb{R}$ satisfies $0 \leq \alpha < 1$. Then there exists a unique $M \in \mathbb{S}^n$ such that

$$\alpha A^TMA - M + Q = 0 \tag{12}$$

and the matrix M is positive definite and satisfies

$$A^TMA - M \preceq 0$$

Proof: We assume $\alpha > 0$, since otherwise the result is trivially true. The conditions $Q \succ 0$ and $A^TQA - Q \preceq 0$ imply by the standard properties of Lyapunov equations that $\rho(A) \leq 1$, and hence $\rho(\sqrt{\alpha}A) < 1$, and hence (12) has a unique solution M, given by

$$M = \sum_{i=0}^{\infty} \alpha^i (A^T)^i QA^i$$

and since $Q \succ 0$ we have $M \succ 0$. Since by assumption $Q \succeq A^TQA$, we know by induction that

$$Q \succeq (A^T)^iQA^i \quad \text{for all } i \in \mathbb{N}$$

and hence

$$\begin{aligned} M &\preceq \sum_{i=0}^{\infty} \alpha^i Q \\ &= \frac{1}{1-\alpha}Q \end{aligned}$$

Therefore

$$\begin{aligned} \alpha A^TMA &= M - Q \\ &\preceq M - (1-\alpha)M \\ &= \alpha M \end{aligned}$$

and since $\alpha > 0$ this implies $M \succeq A^TMA$ as desired. ∎

Theorem 12: Suppose $A \in \mathbb{R}^{n\times n}$, $Q \succ 0$, $\Sigma \succ 0$ and $A^TQA - Q \preceq 0$. Let M be the unique solution to

$$\frac{1}{1+\mathrm{trace}(\Sigma M)}A^TMA - M + Q/\lambda = 0$$

Then for all policies $\mu \in \mathcal{P}$ we have

$$J(\mu) \geq \frac{\lambda\,\mathrm{trace}(\Sigma M)}{3\big(1+\mathrm{trace}(\Sigma M)\big)}$$

Proof: By Lemma 11, $A^TQA - Q \preceq 0$ implies $A^TMA - M \preceq 0$. Then applying Lemma 10 gives the result. ∎

VI. EXAMPLE

Here we present a simple example to illustrate the results of this paper. In this example, the system state evolves as a random walk with Gaussian increments. In order to plot the computed policy and a sequence of estimation errors, we consider a system with state in $\mathbb{R}^2$.

The system in this example has an A matrix and covariance Σ given by

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 0.03 & -0.02 \\ -0.02 & 0.04 \end{bmatrix}$$

The error cost and transmission cost are specified by

$$Q = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad \lambda = 20$$

Note that for this A, any Q satisfies $A^TQA - Q \preceq 0$. To determine the quadratic threshold policy, we find M satisfying

$$\left(\frac{1}{1+\mathrm{trace}(\Sigma M)}\right)A^TMA - M + Q/\lambda = 0$$

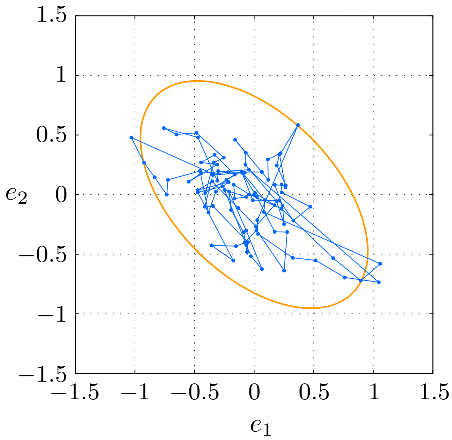

Fig. 2. The threshold used by the policy μ and a trajectory of the error $e(t)$.

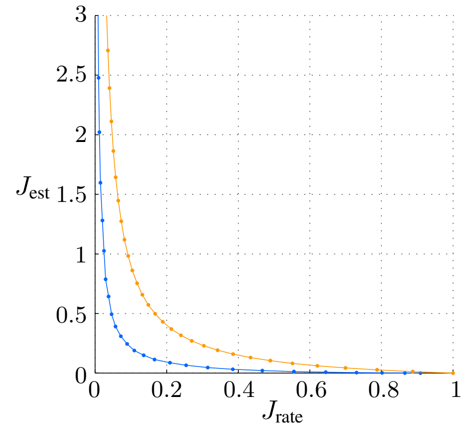

Fig. 3. The trade-off between average communication rate J_{rate} and weighted average error cost J_{est} when using quadratic threshold policies for the example system. The upper curve shows the error costs achieved when a_t is IID Bernoulli with probability J_{rate}. The lower curve shows the error costs achieved using the proposed quadratic threshold.

In this case we compute

$$M = \begin{bmatrix} 1.47 & 0.73 \\ 0.73 & 1.47 \end{bmatrix}$$

and use the policy of equation (5). This threshold is shown by the ellipse in Figure 2.

For this problem instance, we have the lower bound on the optimal cost

$$\begin{aligned} J_{\text{opt}} &\geq \frac{\lambda \ \operatorname{trace}(\Sigma M)}{3(1+\operatorname{trace}(\Sigma M))} \\ &\approx 0.45 \end{aligned}$$

We obtain the sequence of estimation errors shown in Figure 2 by simulating this system for 100 time steps under the policy μ. For this simulation, we obtain the empirical average cost

$$\begin{aligned} J_{\text{avg}}(\mu) &= \frac{1}{N}\sum_{t=0}^{N-1}\Big(\big(1-\mu(e_t)\big)e_t^T Q e_t + 20\mu(e_t)\Big) \\ &\approx 1.6 \text{ for large } N \end{aligned}$$

As expected, we observe $J_{\text{avg}}(\mu) \leq 6J_{\text{opt}}$.

For this system we also compute the curve showing the trade-off between average communication rate and average estimation error, as shown in Figure 3. This curve is computed by considering a series of values of λ, and for each λ computing the quadratic threshold policy μ, as well as a corresponding empirical average transmission rate J_{rate} and an empirical estimation error J_{est} given by

$$J_{\text{rate}} = \frac{1}{N}\sum_{t=0}^{N-1} a_t \qquad J_{\text{est}} = \frac{1}{N}\sum_{t=0}^{N-1}(1-a_t)e_t^T Q e_t$$

The plot shows the average error cost associated with each average rate.

VII. Conclusions

In this paper we considered a simple, yet fundamental estimation problem involving balancing the trade-off between communication rate and estimation error in networked linear systems. This paper extended the work of [8], where it was shown that this problem can be posed as a Markov decision process. Here we show that there is a simple, easily computable suboptimal policy for scheduling state transmissions which incurs a cost within a factor of six of the optimal achievable cost.

References

[1] K. J. Åström and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proc. IEEE Conference on Decision and Control*, 2002.

[2] D. Hristu-Varsakelis and P. R. Kumar, "Interrupt-based feedback control over a shared communication medium," in *Proc. IEEE Conference on Decision and Control*, 2002.

[3] 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. Sys. Technol.*, vol. 4, no. 10, pp. 503–518, July 2002.

[4] J. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," *Proc. Int. Symp. Mathematical theory of Networks and Systems*, August 2002.

[5] Y. Xu and J. P. Hespanha, "Communication logics for networked control systems," in *Proc. American Control Conference*, 2004.

[6] M. Rabi and J. S. Baras, "Sampling of diffusion processes for real-time estimation," in *Proc. IEEE Conference on Decision and Control*, 2004.

[7] O. C. Imer and T. Başar, "Optimal estimation with limited measurements," in *Proc. IEEE Conference on Decision and Control*, 2005, pp. 1029–1034.

[8] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," *Proc. IEEE Conference on Decision and Control*, pp. 3527–3532, 2004.

[9] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties*. Springer Verlag, 1999.

[10] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled markov processes with average cost criterion: a survey," *Optimization*, vol. 31, no. 2, pp. 282–344, 1993.

[11] R. Cogill and S. Lall, "Suboptimality bounds in stochastic control: A queueing example," *Proc. American Control Conference*, pp. 1642 – 1647, 2006.

[12] K. Triantafyllopoulos, "On the central moments of the multidimensional Gaussian distribution," *The Mathematical Scientist*, vol. 28, pp. 125–128, 2003.