How Can Online Schedules Improve Communication and Estimation Tradeoff?

Junfeng Wu, Ye Yuan, Huanshui Zhang, and Ling Shi

***Abstract*—We consider remote state estimation and investigate the tradeoff between the sensor-to-estimator communication rate and the remote estimation quality. It is well known that if the communication rate is one, e.g., the sensor communicates with the remote estimator at each time, then the remote estimation quality is the best. It degrades when the communication rate drops. We present one optimal offline schedule and two online schedules and show that the two online schedules provide better tradeoff between the communication rate and the estimation quality than the optimal offline schedule. Simulation examples demonstrate that significant communication savings can be achieved under the two online schedules which only introduce small increment of the estimation errors.**

***Index Terms*—Communication-estimation tradeoff, Kalman filter, online sensor schedules, remote state estimation.**

I. INTRODUCTION

THE past decade has seen a growing interest in the area of networked sensing and state estimation, which has a broad spectrum of applications in environmental monitoring, body sensor network, smart transportation and power grid, etc. In many of these applications, sensors that measure parameters of interest are battery-powered, and the amount of energy for communication with a remote data center is limited. Network bandwidth may be also limited and shared by many nodes. Therefore, it is important to see how the remote estimation performance degrades as the amount of communication reduces.

This paper considers a remote state estimation problem and investigates how the reduction of data communication between a sensor and a remote estimator affects the remote estimation quality. We present one optimal offline schedule and two online schedules and show that under the two online schedules, by tolerating a small increment of the estimation errors, a significant amount of communication savings can be achieved, which is impossible using the optimal offline schedule. Before we introduce the problem setup and present the main results, we briefly go over some related works in literature. More related works can be found in the references therein.

Shakeri *et al.* [1] considered sensor measurement scheduling with a fixed cost constraint. In their work, the measurement has a cost which is inversely proportional to its error covariance. Krishnamurthy [2] considered scheduling of noisy sensors which measure the state of a single Markov chain. They proposed some algorithms which aim to minimize the estimation errors as well as the measurement costs. Chhetri *et al.* [3] presented two sensor scheduling algorithms for target tracking. Chen *et al.*[4] also considered the optimal transmission scheduling to maximie the sensor network lifetime by making use of the channel information. Dong *et al.* [5] considered the data retrieval problem in a one-dimensional sensor network. The performance of deterministic and random schedules are compared. Mo *et al.* [6], [7] took a look at the sensor selection problems where a subset of sensors are selected at each time to maximize the network lifetime. Savage and La Scala [8] considered sensor measurement scheduling and provided the optimal schedule under the constraint that only $n < N$ measurements can be taken over a time horizon N. Arai *et al.* [9] considered a similar problem setting and proposed a fast sensor scheduling algorithm. Vitus *et al.* [10] considered multiple sensors scheduling where only one sensor is allowed to take a measurement at each time. Cohen and Lesham [11] presented a time-varying opportunistic protocol to maximize the network lifetime assuming that the sensors used are battery-powered and non-rechargeable. Yang and Shi [12] considered finite time-horizon sensor data scheduling under limited communication resource using both terminal and average error covariance as performance metrics.

The remainder of this paper is organized as follows. Section II introduces the mathematical problem. Section III contains the main results of this paper. Simulation examples are given in Section IV and some concluding remarks are given in the end.

Notations: $\mathbb{Z}$ is the set of nonnegative integers. $\mathbb{N}$ is the set of positive integers. $k \in \mathbb{Z}$ is the time index. $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{S}_+^n$ is the set of n by n positive semi-definite matrices. When $X \in \mathbb{S}_+^n$, it is written as $X \geq 0$. $X \geq Y$ if $X - Y \in \mathbb{S}_+^n$. $\mathbb{E}[\cdot]$ is the expectation of a random variable and $\mathbb{E}[\cdot|\cdot]$ is the conditional expectation. $\Pr(\cdot)$ is the probability of a random event. $\mathrm{Tr}[\cdot]$ is the trace of a matrix. $|x| = \sqrt{x'x}$ is the Euclidean norm of a vector x. I_n is the n by n identity matrix.

II. PROBLEM SETUP

Consider the following Gauss-Markov process (Fig. 1)

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

Manuscript received August 20, 2012; revised October 16, 2012; accepted November 26, 2012. Date of publication January 11, 2013; date of current version March 08, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ljubisa Stankovic. The work by J. Wu and L. Shi is supported by an RGC grant HKUST11/CRF/10. The work by H. Zhang is supported by National Natural Science Foundation for Distinguished Young Scholars of China (No. 60825304) and the National Basic Research Development Program of China (973 Program) (No. 2009cb320600).

J. Wu and L. Shi are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: jfwu@ust.hk; eesling@ust.hk).

Y. Yuan is with the Control Group, Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, U.K. (e-mail: yy311@cam.ac.uk).

H. Zhang is with the School of Control Science and Engineering, Shandong University, Jinan, 250061, Shandong, China (e-mail: hszhang@sdu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2238536

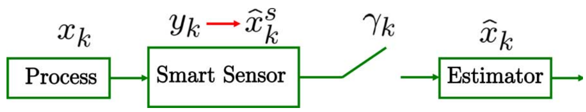


Fig. 1. System block diagram.

where $x_k \in \mathbb{R}^n$ is state of the process at time k, $w_k \in \mathbb{R}^n$ is zero-mean Gaussian disturbance with covariance $\mathbb{E}[w_k w_j'] = \delta_{kj} Q$ $(Q \geq 0)$, where δ_{kj} is the Kronecker delta function, i.e., $\delta_{kj} = 1$ if $k = j$ and $\delta_{kj} = 0$, otherwise. The initial state x_0 is also zero-mean Gaussian with covariance $\Pi_0 \geq 0$. A sensor measures x_k and obtains the following measurement

$$y_k = Cx_k + v_k, \tag{2}$$

where $v_k \in \mathbb{R}^l$ is zero-mean Gaussian measurement noise with covariance $\mathbb{E}[v_k v_j'] = \delta_{kj} R$ $(R > 0)$. We assume w_k's, v_k's, and the initial state x_0 are mutually uncorrelated. The pair $(A, \sqrt{Q})$ is assumed to be stabilizable and (A, C) is detectable.

After y_k is obtained, the sensor runs a local Kalman filter to compute $\hat{x}_k^s = \mathbb{E}[x_k | y_0, \ldots, y_k]$, the minimum mean-squared error estimate of x_k. Define the local estimation error e_k^s as

$$e_k^s = x_k - \hat{x}_k^s. \tag{3}$$

From standard Kalman filtering, $\hat{x}_k^s$ and its estimation error covariance matrix $P_k^s = \mathbb{E}[(e_k^s)(e_k^s)' | y_0, \ldots, y_k]$ are computed as follows:

$$\hat{x}_{k|k-1}^s = A\hat{x}_{k-1}^s, \tag{4}$$
$$P_{k|k-1}^s = AP_{k-1}^s A' + Q, \tag{5}$$
$$K_k = P_{k|k-1}^s C' \left[CP_{k|k-1}^s C' + R \right]^{-1}, \tag{6}$$
$$\hat{x}_k^s = \hat{x}_{k|k-1}^s + K_k \left[y_k - C\hat{x}_{k|k-1}^s \right], \tag{7}$$
$$P_k^s = (I - K_k C) P_{k|k-1}^s. \tag{8}$$

The sensor then decides whether or not it will send $\hat{x}_k^s$ to a remote estimator. The remote estimator has a built-in estimator that estimates x_k in case no data is received from the sensor. Let $\gamma_k \in \{0, 1\}$ be the indicator of such a decision, i.e., if $\gamma_k = 0$, $\hat{x}_k^s$ is not sent by the sensor; and if $\gamma_k = 1$, $\hat{x}_k^s$ is sent. Let θ be a sensor data schedule which specifies the value of γ_k for each $k \in \mathbb{Z}$. We sometime write γ_k as $\gamma_k(\theta)$ to indicate explicitly that the underlying sensor data schedule is θ. Define $\gamma(\theta)$ as the average sensor-to-estimator communication rate, i.e.,

$$\gamma(\theta) \triangleq \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}\left[\gamma_k(\theta) \right]. \tag{9}$$

Denote $\hat{x}_k$ as the minimum mean-squared error estimate of x_k computed at the remote estimator, and P_k as the corresponding estimation error covariance matrix. For a given θ, consider the following cost

$$J(\theta) = \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{Tr}\left(\mathbb{E}[P_k] \right). \tag{10}$$

In this paper, we are interested in finding out how does J degrades as a function of γ. In particular, we construct different θ's under which by tolerating a small increment of J, a significant deduction of γ can be achieved.

Standard Kalman filtering analysis [13] shows that P_k^s in (5) and (8) converges exponentially to a steady-state value $P \geq 0$. It is straightforward to show that P is the solution to

$$P = h(P) - h(P)C' \left[Ch(P)C' + R \right]^{-1} Ch(P), \tag{11}$$

where $h : \mathbb{S}_+^n \to \mathbb{S}_+^n$ is defined as $h(X) = AXA' + Q$. Since we are dealing with an infinite-time horizon, without loss of generality, in sequel, we will assume $P_k^s = P$ for all $k \in \mathbb{Z}$. As a result, $K_k = K$ for all $k \geq 0$ with $K = PC'[CPC' + R]^{-1}$.

Since

$$h(P)C' \left[Ch(P)C' + R \right]^{-1} Ch(P) \geq 0,$$

from (11), one can obtain that

$$h(P) \geq P$$

and hence

$$P \leq h(P) \leq h^2(P) \leq \cdots \leq h^t(P). \tag{12}$$

III. Sensor Data Schedules

A. Optimal Offline Sensor Data Schedule

In this subsection, we give an optimal offline sensor data schedule (Proposition 3.1). We restrict γ to be in the range of [1/2, 1] as we are interested in the reduction of communication rate which corresponds to only a small increment of the estimation errors.

Under any offline schedule, it is not difficult to show that the remote estimator calculates $\hat{x}_k$ as $\hat{x}_k = \hat{x}_k^s$ if $\gamma_k = 1$ and $\hat{x}_k = A\hat{x}_{k-1}$ if $\gamma_k = 0$. Intuitively, since $\hat{x}_k^s$ encodes all past measurements by the sensor, once it becomes available, the remote estimator should synchronize its own estimate with it; on the other hand, if $\hat{x}_k^s$ is not available, then the remote estimator simply predicts the value of x_k based on its previous optimal estimate. The corresponding error covariance P_k is computed as

$$P_k = \begin{cases} P, & \text{if } \gamma_k = 1, \\ AP_{k-1}A' + Q, & \text{if } \gamma_k = 0. \end{cases}$$

We now present an optimal offline sensor data schedule, which is very similar to the optimal sensor transmission energy power schedule given by Theorem 5.2 in [14]. With some modification of the notations, the proofs are also very similar, but we will present a simpler and alternative proof. Although the two results look similar, they deal with different problems: [14] considers power scheduling (i.e., should a high or low transmission energy be used) and the following result states whether should the sensor send data or not.

Proposition 3.1: Consider a given sensor communication rate $\gamma = \frac{p}{q} \geq \frac{1}{2}$, where $p, q \in \mathbb{N}$, p and q are co-prime and $p \leq q$.

An optimal offline schedule $\theta^\star$ can be constructed in terms of the values of $\gamma_k(\theta^\star)$ over a period q as follows:

$$\underbrace{(1\ 0)\cdots(1\ 0)}_{q-p\ times}\underbrace{(1)\cdots(1)}_{2p-q\ times} \tag{13}$$

The corresponding $J(\theta^\star)$ is given by

$$J(\theta^\star) = \frac{1}{q}\mathrm{Tr}\left[(pP + (q-p)h(P))\right].$$

Proof: First, an arbitrary schedule θ with sensor-to-estimator communication rate $\frac{p}{q}$ can be presented as a sequence of separated units as follows:

$$\underbrace{(1\ 0\cdots 0)}_{l_1}\cdots\underbrace{(1\ 0\cdots 0)}_{l_i}\cdots\underbrace{(1\ 0\cdots 0)}_{l_j}\cdots$$

where $l_i \geq 1$ and that

$$\limsup_{j\to\infty}\frac{j}{\sum_{i=1}^{j} l_i} = \frac{p}{q}.$$

We shall prove that $J(\theta^\star) \leq J(\theta)$ hence proving the optimality of $\theta^\star$.

Step 1) If there exists l_i and l_j with $l_j - l_i > 1$. Then construct the following different $\hat{\theta}$ based on θ as follows:

$$\underbrace{(1\ 0\cdots 0)}_{l_1}\cdots\underbrace{(1\ 0\cdots 0)}_{l_i+1}\cdots\underbrace{(1\ 0\cdots 0)}_{l_j-1}\cdots$$

As

$$\sum_{t=0}^{l_i-1} h^t(P) + \sum_{t=0}^{l_j-1} h^t(P) - \sum_{t=0}^{l_i} h^t(P) - \sum_{t=0}^{l_j-2} h^t(P) = h^{l_j-1}(P) - h^{l_i}(P) \geq 0,$$

where the last inequality is due to $l_j - 1 - l_i > 0$ and (12). Hence one concludes that $J(\hat{\theta}) \leq J(\theta)$.

Step 2) Repeat step 1 to $\hat{\theta}$ until

$$|l_j - l_i| \leq 1\ \forall i, j. \tag{14}$$

To meet the rate constraint $\frac{1}{2} \leq \frac{p}{q} \leq 1$, those l_i's in $\hat{\theta}$ can only be 1 or 2[1] and in particular, the portion of l_i's which equal to 2 is given by $\frac{q-p}{q}$ and the portion of l_i's which equal to 1 is given by $\frac{2p-q}{q}$. Otherwise, if there exists $l_i \geq s$ with $s \geq 3$, then from (14), all other l_i's have to be either $s+1$ or $s-1$ and the resulting rate will be strictly less than 1/2.

Step 3) It is straightforward to see that $J(\theta^\star) = J(\hat{\theta})$. Therefore for an arbitrary schedule θ, we have $J(\theta^\star) \leq J(\theta)$, i.e., $\theta^\star$ is indeed optimal.

■

Remark 3.2: Note that the optimal offline schedule $\theta^\star$ constructed in Proposition 3.1 is not unique. If we treat the $q-p$ copies of (1 0)'s as $q-p$ separate units, and also treat the $2p-q$ copies of (1)'s as $2p-q$ separate units, then any permutation of these p units generates one optimal schedule that has a cost equal to $J(\theta^\star)$.

From Proposition 3.1, $\frac{\partial J}{\partial \gamma} = -\mathrm{Tr}[h(P) - P]$, i.e., the rate of change in J is linear to the rate of change in γ. Hence if we can only tolerate a small degradation in the estimation quality, then the reduction of the communication rate is also small (this is illustrated in the example section). As we shall see from the next few sections, however, for the same amount of degradation in J, the reduction of the communication rate can be significant if online schedules are used.

B. Preliminaries for Online Schedules

Although $\theta^\star$ given by Proposition 3.1 is an *optimal* offline schedule, much more can be achieved by utilizing the real-time state information. Before we introduce these online schedules, we briefly go over some preliminaries which our main results are based upon.

Let

$$\Delta \triangleq K\left[Ch(P)C' + R\right]K$$

with

$$K = h(P)C'\left[Ch(P)C' + R\right]^{-1}.$$

From (11), it can be easily verified that

$$h(P) = P + \Delta. \tag{15}$$

Let the rank of Δ be r. Since $\Delta \geq 0$, we can find an orthonormal matrix F such that

$$F'\Delta F = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix},$$

where Λ is a diagonal matrix whose diagonal elements correspond to the non-zero eigenvalues of Δ. Let

$$E = F\begin{bmatrix} \Lambda^{-\frac{1}{2}} & 0 \\ 0 & I_{n-r} \end{bmatrix}.$$

Then we have

$$E'\Delta E = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}. \tag{16}$$

Let $\xi \in \mathbb{R}^r$ be a zero-mean Gaussian random variable with

$$\mathbb{E}[\xi\xi'] = T_r.$$

Let $\delta > 0$ be a positive real number. Define $\rho(\delta)$ and $\hat{\Gamma}(\delta)$ as

$$\rho(\delta) \triangleq \Pr\left(|\xi| \leq \delta\right), \tag{17}$$

$$\hat{\Gamma}(\delta) \triangleq \mathbb{E}\left[\xi\xi' \middle| |\xi| \leq \delta\right]. \tag{18}$$

Some useful properties of $\rho(\delta)$ and $\hat{\Gamma}(\delta)$ are given in the following Lemma. The proof follows from the definition of (conditional) probability density function of a Gaussian random variable and is omitted.

[1]We only count those which appear infinitely often.

Lemma 3.3: Let $0 \le \delta_1 < \delta_2$. Then

$$\rho(\delta_1) < \rho(\delta_2) \quad \text{and} \quad \hat{\Gamma}(\delta_1) < \hat{\Gamma}(\delta_2). \tag{19}$$

Furthermore $\lim_{\delta \to \infty} \rho(\delta) = 1$ and $\lim_{\delta \to \infty} \hat{\Gamma}(\delta) = I$.

Some properties of e_k^s defined in (3) are summarized in the next lemma.

Lemma 3.4: The following statements on e_k^s hold:

1) e_k^s is zero-mean Gaussian and is independent of $\hat{x}_k^s$.
2) e_k^s is independent of w_{k_1} and v_{k_2} for any $k_1, k_2 \in \mathbb{N}$ and $k_1 \ge k, k_2 \ge k+1$.
3) e_k^s is independent of $\hat{x}_k^s - A\hat{x}_{k-1}^s$.

Proof:

1) Direct result from the orthogonality principle [15].
2) Write e_k^s as

$$e_k^s = (A - KCA)e_{k-1}^s + (I - KC)w_{k-1} - Kv_k. \tag{20}$$

Thus e_k^s is a linear function of $x_0, w_0, \ldots, w_{k-1}$, and $v_1, \ldots, v_k$. Since x_0, w_k's and v_k's are mutually independent, the statement holds.

3) From (20), we see that e_k^s is also a linear function of e_{k-1}^s, w_{k-1} and v_k. From (7), $\hat{x}_{k-1}^s$ only depends on x_0, $w_0, \ldots, w_{k-2}$, and $v_1, \ldots, v_{k-1}$, thus $\hat{x}_{k-1}^s$ is independent of $w_{k-1}, \ldots, w_{k-1}$ and v_k. From the first statement, $\hat{x}_{k-1}^s$ is independent of e_{k-1}^s. Therefore we conclude that $\hat{x}_{k-1}^s$ is independent of e_k^s. Together with the first statement, we arrive at the fact that e_k^s is independent of $\hat{x}_k^s - A\hat{x}_{k-1}^s$. ■

For convenience, we denote

$$\Gamma(\delta) = \Lambda^{\frac{1}{2}} \hat{\Gamma}(\delta) \Lambda^{\frac{1}{2}}. \tag{21}$$

Since $\Lambda > 0$, if $0 \le \delta_1 < \delta_2$, then from Lemma 3.3, we have

$$\Gamma(\delta_1) < \Gamma(\delta_2). \tag{22}$$

Define ϵ_k as

$$\epsilon_k = \hat{x}_k^s - A\hat{x}_{k-1}. \tag{23}$$

We have the following result on ϵ_k.

Lemma 3.5: Assume $\hat{x}_{k-1} = \hat{x}_{k-1}^s$. Then

1) ϵ_k is independent of e_k^s;
2) ϵ_k is zero-mean Gaussian with variance Δ;
3) $\mathbb{E}[\epsilon_k \epsilon_k' | |E'\epsilon_k| \le \delta] = F \begin{bmatrix} \Gamma(\delta) & 0 \\ 0 & 0 \end{bmatrix} F'$.

Proof:

1) Since $\hat{x}_{k-1} = \hat{x}_{k-1}^s$, we have

$$\epsilon_k = \hat{x}_k^s - A\hat{x}_{k-1} = \hat{x}_k^s - A\hat{x}_{k-1}^s.$$

From part 3) of Lemma 3.4, we conclude that ϵ_k is independent of e_k^s.

2) Note that

$$\begin{aligned} \epsilon_k &= \hat{x}_k^s - A\hat{x}_{k-1}^s \\ &= K\left[y_k - C(A\hat{x}_{k-1}^s)\right] \\ &= K\left[CAe_{k-1}^s + v_k + Cw_{k-1}\right]. \end{aligned}$$

From Lemma 3.4, e_{k-1}^s is independent of v_k and w_{k-1}, thus

$$\begin{aligned} \mathbb{E}\left[\epsilon_k \epsilon_k'\right] &= K\left[C(APA' + Q)C' + R\right]K' \\ &= K\left[Ch(P)C' + R\right]K' \\ &= \Delta. \end{aligned}$$

3) Let $\eta_k \triangleq E'\epsilon_k = \left[\begin{smallmatrix} \xi_k \\ \phi_k \end{smallmatrix}\right]$ with $\xi_k \in \mathbb{R}^r$ and $\phi_k \in \mathbb{R}^{n-r}$. Since $\mathbb{E}[\epsilon_k \epsilon_k'] = \Delta$, we have

$$\mathbb{E}\left[\eta_k \eta_k'\right] = E'\Delta E = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore $\phi_k = 0$ almost surely, which leads to the following

$$\mathbb{E}\left[\eta_k \eta_k' | |\eta_k| \le \delta\right] = \mathbb{E}\left[\eta_k \eta_k' | |\xi_k| \le \delta\right] = \begin{bmatrix} \hat{\Gamma}(\delta) & 0 \\ 0 & 0 \end{bmatrix}. \tag{24}$$

From (24), one has

$$\begin{aligned} \mathbb{E}\left[\epsilon_k \epsilon_k' | |E'\epsilon_k| \le \delta\right] &= (E')^{-1} \mathbb{E}\left[\eta_k \eta_k' | |\eta_k| \le \delta\right] (E)^{-1} \\ &= F \begin{bmatrix} \Gamma(\delta) & 0 \\ 0 & 0 \end{bmatrix} F', \end{aligned}$$

where we use the fact that $E^{-1} = \begin{bmatrix} \Lambda^{\frac{1}{2}} & 0 \\ 0 & I_{n-r} \end{bmatrix} F'$. ■

C. First Online Sensor Schedule

In this subsection, we introduce a simple online sensor data schedule θ_1, which assigns the value of γ_k according to the following rule

$$\gamma_k = \begin{cases} 0, & \text{if } k \text{ is even and } |E'\epsilon_k| \le \delta, \\ 1, & \text{otherwise,} \end{cases} \tag{25}$$

where ϵ_k and E_k are defined in the previous subsection. Under θ_1, when $\gamma_k = 1$, the remote estimator simply resets $\hat{x}_k = \hat{x}_k^s$; when $\gamma_k = 0$, $\hat{x}_k$ has the same form as that under the optimal offline schedule, i.e., $\hat{x}_k = A\hat{x}_{k-1}$. However, under schedule θ_1, if k is even and no update is sent, this means that $|E'\epsilon_k| \le \delta$, which enables a revision to the standard Kalman filtering update equation, yielding a smaller error covariance P_k.

The next theorem characterizes the tradeoff between the sensor-to-estimator communication rate γ and the corresponding cost function J under the online sensor data schedule θ_1.

Theorem 3.6: Under the online schedule θ_1, the expected sensor communication rate $\gamma(\delta)$ is given by

$$\gamma(\delta) = 1 - \frac{1}{2}\rho(\delta), \tag{26}$$

where $\rho(\delta)$ is defined in (17). The corresponding cost $J(\theta_1)$ is given by

$$J(\theta_1) = \mathrm{Tr}\left[P + \frac{1}{2}\rho(\delta) F \begin{bmatrix} \Gamma(\delta) & 0 \\ 0 & 0 \end{bmatrix} F'\right], \tag{27}$$

where $\Gamma(\delta)$ is given by (21).

Proof: When k is odd, from (25), $\gamma_k = 1$, thus $P_k = P_k^s = P$. When k is even, $k-1$ is odd, thus $\hat{x}_{k-1} = \hat{x}_{k-1}^s$. Note that $\epsilon_k = \hat{x}_k^s - A\hat{x}_{k-1}$. Assume $|E'\epsilon_k| \le \delta$ holds at this even k. Then $\hat{x}_k^s$ is kept by the sensor and the remote estimator computes $\hat{x}_k = A\hat{x}_{k-1}$. The corresponding P_k can be computed as

$$\begin{aligned} P_k &= \mathbb{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)' \| E'\epsilon_k| \le \delta\right] \\ &= \mathbb{E}\left[(e_k^s + \epsilon_k)(e_k^s + \epsilon_k)' \| E'\epsilon_k| \le \delta\right] \\ &= \mathbb{E}\left[(e_k^s)(e_k^s)' \| E'\epsilon_k| \le \delta\right] + \mathbb{E}\left[\epsilon_k\epsilon_k' \| E'\epsilon_k| \le \delta\right] \\ &= P + F\begin{bmatrix}\Gamma(\delta) & 0 \\ 0 & 0\end{bmatrix}F', \end{aligned}$$

where we use the independence of e_k^s and ϵ_k (Lemma 3.5). As a result,

$$\begin{aligned} J(\theta_1) &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\mathrm{Tr}\left[P_k(\theta)\right] \\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\left(1 - \Pr(k \text{ is even}, \gamma_k = 0)\right)\mathrm{Tr}[P_k] \\ &\quad + \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\Pr(k \text{ is even}, \gamma_k = 0)\mathrm{Tr}[P_k] \\ &= \mathrm{Tr}\left[\left(1 - \frac{1}{2}\rho(\delta)\right)P\right] \\ &\quad + \mathrm{Tr}\left[\frac{1}{2}\rho(\delta)\left(P + F\begin{bmatrix}\Gamma(\delta) & 0 \\ 0 & 0\end{bmatrix}F'\right)\right] \\ &= \mathrm{Tr}\left[P + \frac{1}{2}\rho(\delta)F\begin{bmatrix}\Gamma(\delta) & 0 \\ 0 & 0\end{bmatrix}F'\right]. \end{aligned}$$

The sensor communication rate γ under θ_1 is given by

$$\begin{aligned} \gamma &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\mathbb{E}[\gamma_k] \\ &= \lim_{N\to\infty}\frac{1}{N}\left(\sum_{k \text{ is odd}}\mathbb{E}[\gamma_k] + \sum_{k \text{ is even}}\mathbb{E}[\gamma_k]\right) \\ &= \frac{1}{2} + \lim_{N\to\infty}\frac{1}{N}\sum_{k \text{ is even}}\Pr\left(|E'\epsilon_k| > \delta\right) \\ &= 1 - \frac{1}{2}\rho(\delta). \end{aligned}$$

■

D. Second Online Sensor Schedule

We now present an improved schedule θ_2, which assigns the values of γ_k as follows:

$$\gamma_k = \begin{cases} 0, & \text{if } \gamma_{k-1} = 1 \text{ and } |E'\epsilon_k| \le \delta, \\ 1, & \text{otherwise.} \end{cases} \tag{28}$$

Notice that the event-triggering condition is slightly complicated than the first one given by (25), as it also needs to remember the past decision γ_{k-1}. This additional light complexity, however, will improve the remote estimation quality as shown in the subsequent analysis.

Under θ_2, the remote estimator computes $\hat{x}_k$ in the same way as that under θ_1, i.e., when $\gamma_k = 1$, $\hat{x}_k = \hat{x}_k^s$; when $\gamma_k = 0$, $\hat{x}_k = A\hat{x}_{k-1}$. Similar to θ_1, the error covariance P_k under θ_2 is different than that under the optimal offline schedule. The tradeoff between γ and J under θ_2 is given as follows.

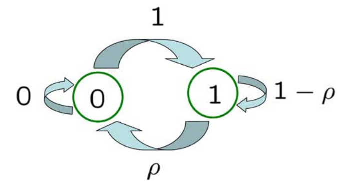


Fig. 2. A two-state Markov chain that represents the possible values taken by γ_k given γ_{k-1}.

Theorem 3.7: Under the online schedule θ_2, the expected sensor communication rate is given by

$$\gamma(\delta) = \frac{1}{1+\rho(\delta)}. \tag{29}$$

The corresponding cost $J(\theta_2)$ is given by

$$J(\theta_2) = \mathrm{Tr}\left[P + \frac{\rho(\delta)}{1+\rho(\delta)}F\begin{bmatrix}\Gamma(\delta) & 0 \\ 0 & 0\end{bmatrix}F'\right]. \tag{30}$$

Proof: From (28), γ_k depends on γ_{k-1}, but not on γ_{k-t} for $t \ge 2$. Thus we can use a two-state Markov chain (Fig. 2) to represent the possible values taken by γ_k. In Fig. 2, an arrow starts from a possible value taken by γ_{k-1} and the arrow ends at a possible value taken by γ_k. The probability transition matrix $\mathbf{P}$ of this Markov chain is easily seen to be

$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ \rho(\delta) & 1-\rho(\delta) \end{bmatrix}.$$

Let $[\pi_0\ \pi_1]$ be the steady-state distribution of the two states. Simple calculation reveals that

$$\pi_0 = \frac{\rho(\delta)}{1+\rho(\delta)}, \quad \pi_1 = \frac{1}{1+\rho(\delta)}.$$

The theorem is proved by noticing that π_1 corresponds to the portion of times when $\gamma_k = 1$, which equals the sensor communicate rate γ. The remaining part of the theorem is straightforward to show. ■

E. Comparison of Different Schedules

We now compare the performance of $\theta^\star$, θ_1 and θ_2, and we have the following main result.

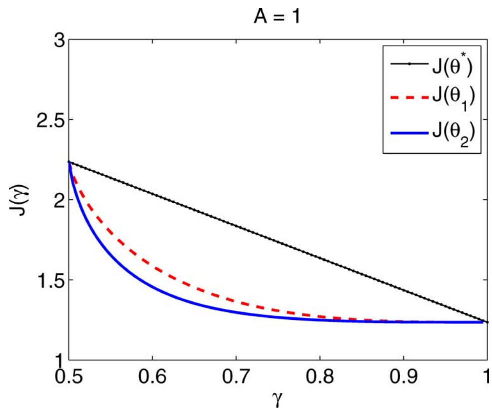
Proposition 3.8: If the sensor-to-estimator communication rate $\gamma \in [\frac{1}{2}, 1]$ is the same under $\theta^\star$, θ_1 and θ_2, then

$$J(\theta_2) \le J(\theta_1) \le J(\theta^\star)$$

with equality iff $\gamma = \frac{1}{2}$ or $\gamma = 1$.

Proof: We first compare $\theta^\star$ with θ_1. Assume $\gamma(\theta^\star) = \gamma(\theta_1)$, i.e.,

$$\frac{p}{q} = 1 - \frac{1}{2}\rho(\delta).$$



(a)

ΔJ	$\Delta\gamma(\theta^\star)$	$\Delta\gamma(\theta_1)$	$\Delta\gamma(\theta_2)$
+1%	-1%	-14%	-19.5%
+5%	-3%	-23.5%	-30%
+10%	-6%	-29.5%	-35%

(b)

Fig. 3. Comparison of $J(\theta^\star)$, $J(\theta_1)$ and $J(\theta_2)$ under the same sensor-to-estimator communication rate γ.

Then

$$\begin{aligned}J(\theta^\star)-J(\theta_1)&=\frac{1}{2}\rho(\delta)\left(\mathrm{Tr}[h(P)]-\mathrm{Tr}\left[P+F\begin{bmatrix}\Gamma(\delta)&0\\0&0\end{bmatrix}F'\right]\right)\\&\geq\frac{1}{2}\rho(\delta)\left(\mathrm{Tr}[h(P)]-\mathrm{Tr}\left[P+F\begin{bmatrix}\Lambda&0\\0&0\end{bmatrix}F'\right]\right)\\&=\frac{1}{2}\rho(\delta)\left(\mathrm{Tr}[h(P)]-\mathrm{Tr}[P+\Delta]\right)=0,\end{aligned}$$

where the inequality is from Lemma 3.3, the second last equality is from (16) and the last equality is from (15). Clearly, $J(\theta^\star)=J(\theta_1)$ iff $\Gamma(\delta)=I$ or $\rho(\delta)=0$, which correspond to $\gamma=\frac{1}{2}$ or $\gamma=1$, respectively.

We now compare θ_1 with θ_2. Note that

$$\frac{1}{1+\rho(\delta)}\leq 1-\frac{1}{2}\rho(\delta)$$

with equality iff $\rho(\delta)=0$ or $\rho(\delta)=1$. Therefore if

$$\frac{1}{1+\rho(\delta_2)}=1-\frac{1}{2}\rho(\delta_1),$$

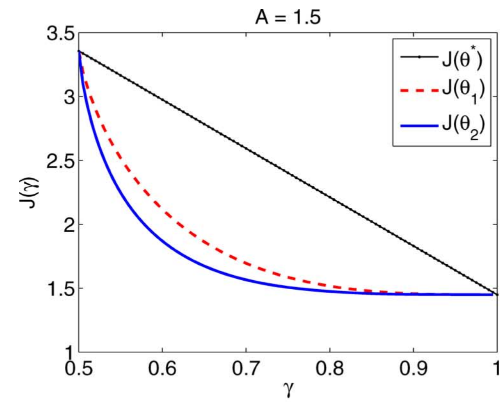
i.e., the sensor-to-estimator communication rate is the same under θ_1 and θ_2, then

$$\frac{1}{1+\rho(\delta_2)}=1-\frac{1}{2}\rho(\delta_1)\leq 1-\frac{1}{2}\rho(\delta_2),$$

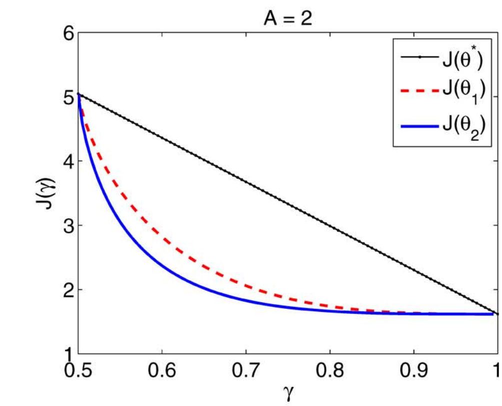
i.e., $\rho(\delta_2)\leq\rho(\delta_1)$ with equality iff $\delta_1=\delta_2=0$ or $\delta_1=\delta_2=\infty$, which correspond to $\gamma=1$ or $\gamma=\frac{1}{2}$, respectively. Thus if $\delta_1\in(0,\infty)$ and $\delta_2\in(0,\infty)$, then $\rho(\delta_2)<\rho(\delta_1)$, from which, one has $\delta_2<\delta_1$. Now (22) implies that $\Gamma(\delta_2)<\Gamma(\delta_1)$. One then has

$$\begin{aligned}J(\theta_1)-J(\theta_2)&=\mathrm{Tr}\left[\left(P+\frac{1}{2}\rho(\delta_1)F\begin{bmatrix}\Gamma(\delta_1)&0\\0&0\end{bmatrix}F'\right)\right]\\&\quad-\mathrm{Tr}\left[\left(P+\frac{\rho(\delta_2)}{1+\rho(\delta_2)}F\begin{bmatrix}\Gamma(\delta_2)&0\\0&0\end{bmatrix}F'\right)\right].\\&=\frac{1}{2}\rho(\delta_1)\mathrm{Tr}\left[F\begin{bmatrix}\Gamma(\delta_1)-\Gamma(\delta_2)&0\\0&0\end{bmatrix}F'\right]>0.\end{aligned}$$

■



(a)



(b)

Fig. 4. J as a function of γ under the three schedules.

IV. Examples

We consider the process (1) with different A's and $C=1$, $Q=2$, $R=2$. We plot the cost J as a function of γ in Fig. 3(a) and Fig. 4(a), 4(b) for the three schedules when $A=1$, 1.5 and 2, respectively. From the figures, we can see that for the same $\gamma\in(0.5,1)$, $J(\theta^\star)>J(\theta_1)>J(\theta_2)$, which agrees with Proposition 3.8. Furthermore, J is an affine function of γ under $\theta^\star$ which agrees with Proposition 3.1.

Fig. 3(b) also summarizes the percentage changes in J and the corresponding percentage changes in γ for the three schedules when $A=1$. In the figure, ΔJ represents the estimation quality degradation compared with the perfect communication case (e.g., $\gamma=1$), and $\Delta\gamma$ represents the amount of communication rate reduction corresponding to such a estimation quality degradation. From the figure we can see that the online sensor data schedules offer much better tradeoff between the sensor-to-estimator communication rate and the estimation quality.

V. Conclusion

In this paper, we propose two online sensor-to-estimator communication strategies, both of which demonstrate a tradeoff between the sensor-to-estimator communication rate and the estimation quality. As shown both theoretically and via examples, for a small degradation in the estimation quality, a significant communication rate reduction can be achieved using these two online schedules which is impossible using offline schedules.

Future work along the line of this work include considering multi-sensor data scheduling and more online sensor data schedules which can provide even better tradeoff. Another interesting direction is to consider data packet drops which are frequently seen in wireless communications. The extra packet drops make the problem more intriguing since if no update is received by the estimator, it may get confused whether the lack of measurement is due to the failure of the communication link or due to the event-triggering at the sensor. How to modify the online schedules or the estimation procedures at the remote estimator to counteract and minimize the effect of data packet drops will be investigated.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the anonymous reviewers for many constructive comments and suggestions that help improve the presentation of the paper.

REFERENCES

[1] M. Shakeri, K. R. Pattipati, and D. L. Kleinman, “Optimal measurement scheduling for state estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 716–729, 1995.

[2] V. Krishnamurthy, “Algorithms for optimal scheduling and management of hidden Markov model sensors,” *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1382–1397, 2002.

[3] A. S. Chhetri, D. Morrell, and A. Papandreou-Suppappola, “On the use of binary programming for sensor scheduling,” *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2826–2839, 2007.

[4] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, “Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach,” *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2294–2309, May 2007.

[5] M. Dong, L. Tong, and B. M. Sadler, “Information retrieval and processing in sensor networks: Deterministic scheduling versus random access,” *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5806–5820, 2007.

[6] Y. Mo, L. Shi, R. Ambrosino, and B. Sinopoli, “Network lifetime maximization via sensor selection,” in *Proc. Asian Control Conf.*, 2009, pp. 441–446.

[7] Y. Mo, R. Ambrosino, and B. Sinopoli, “Sensor selection strategies for state estimation in energy constrained wireless sensor networks,” *Automatica*, vol. 47, no. 7, pp. 1330–1338, 2011.

[8] C. O. Savage and B. F. L. Scala, “Optimal scheduling of scalar gauss-markov systems with a terminal cost function,” *IEEE Trans. Autom. Control*, vol. 54, no. 5, pp. 1100–1105, 2009.

[9] S. Arai, Y. Iwatani, and K. Hashimoto, “Fast sensor scheduling for estimation of networked sensor systems,” presented at the ICROS-SICE Int. Joint Conf., Fukuoka, Japan, 2009.

[10] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, “On efficient sensor scheduling for linear dynamical systems,” presented at the Amer. Control Conf., Baltimore, MD, USA, 2010.

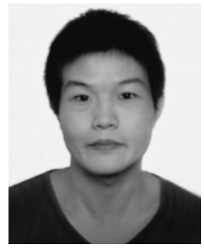
[11] K. Cohen and A. Leshem, “A time-varying opportunistic approach to lifetime maximization of wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5307–5319, 2010.

[12] C. Yang and L. Shi, “Deterministic sensor data scheduling under limited communication resource,” *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5050–5056, May 2011.

[13] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

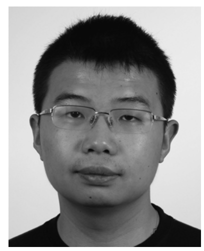
[14] L. Shi, P. Cheng, and J. Chen, “Sensor data scheduling for optimal state estimation with communication energy constraint,” *Automatica*, vol. 47, no. 8, pp. 1693–1698, 2011.

[15] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ, USA: Prentice Hall, 2000.



Junfeng Wu was born in Tianjin City, China, in 1987. He received the B.Eng. degree in the Department of Automatic Control from Zhejiang University, Hangzhou, China, in 2009.

He is currently working towards the Ph.D. degree in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong, China. His current research interests include networked control systems, distributed state estimation and wireless sensor networks.



Ye Yuan was born on 1986. He received his B.Eng. degree (Valedictorian) from the Department of Automation, Shanghai Jiao Tong University in September 2008, M. Phil. and Ph.D. from the Department of Engineering, Cambridge University in October 2009 and February 2012 respectively. His research interest lies in the mathematical control theory with applications to network and biology. He is the recipient of Dorothy Hodgkin Postgraduate Awards, Microsoft Research Ph.D. Scholarship, Cambridge Overseas Scholarship (twice), Chinese Government Award for Outstanding Students Studying Abroad and Henry Lester Scholarship (four times), Best Paper Finalist in IEEE ICIA and a number of travel awards from Pembroke College, Cambridge Engineering Department, Microsoft Research, IEEE, EECI, Los Alamos National Lab etc.

Huanshui Zhang graduated in mathematics from the Qufu Normal University in 1986 and received the M.Sc. degree from Heilongjiang University, China, in 1991 and the Ph.D. degree from Northeastern University, China, in 1997, both in control theory.

He was a Postdoctoral Fellow at Nanyang Technological University from 1998 to 2001 and a Research Fellow at Hong Kong Polytechnic University from 2001 to 2003. He is currently a Changjiang Professor at Shandong University, China. He was a Professor in Harbin Institute of Technology from 2003 to 2006, and he also held visiting appointments as a Research Scientist and a Fellow with Nanyang Technological University, Curtin University of Technology, and Hong Kong City University from 2003 to 2006. His interests include optimal estimation and control, time-delay systems, stochastic systems, signal processing, and wireless sensor networked systems. He is the author and coauthor of more than 100 publications.

Dr. Zhang is an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the *Journal of Industrial and Management Optimization*, and the *Journal of Control Theory and Applications*.

Ling Shi received the B.S. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology in 2002 and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2008.

He is currently an Assistant Professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks, and distributed control.