REFERENCES

[1] P. Antsaklis and J. Baillieul, "Special issue on technology of networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 5–8, Jan. 2007.

[2] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 457–462, Apr. 1969.

[3] J. Tugnait, "Asymptotic stability of the MMSE linear filter for systems with uncertain observations," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 2, pp. 247–250, Feb. 1981.

[4] M. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Autom. Control*, vol. AC-24, no. 6, pp. 944–948, Jun. 1979.

[5] O. Costa, "Linear minimum mean square error estimation for discrete-time Markovian jump linear systems," *IEEE Trans. Autom. Control*, vol. 39, no. 8, pp. 1685–1689, Aug. 1994.

[6] M. Sahebsara, T. Chen, and S. Shah, "Optimal H2 filtering in networked control systems with multiple packet dropout," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1508–1513, Aug. 2007.

[7] S. Sun, L. Xie, W. Xiao, and Y. Soh, "Optimal linear estimation for systems with multiple packet dropouts," *Automatica*, vol. 44, no. 5, pp. 1333–1342, 2008.

[8] Y. Liang, T. Chen, and Q. Pan, "Optimal linear state estimator with multiple packet dropouts," *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1428–1433, Jun. 2010.

[9] W. Zhang, L. Yu, and G. Feng, "Optimal linear estimation for networked systems with communication constraints," *Automatica*, vol. 47, no. 9, pp. 1992–2000, 2011.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[11] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

[12] E. Rohr, D. Marelli, and M. Fu, "Kalman filtering with intermittent observations: Bounds on the error covariance distribution," presented at the 50th IEEE Conf. Decision and Control and Eur. Control Conf., Orlando, FL, USA, 2011.

[13] S. Kar, B. Sinopoli, and J. M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 405–420, Feb. 2012.

[14] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," presented at the 43rd IEEE Conf. Decision and Control, 2004.

[15] A. Chiuso and L. Schenato, "Information fusion strategies and performance bounds in packet-drop networks," *Automatica*, vol. 47, no. 7, pp. 1304–1316, 2011.

[16] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[17] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, May 2008.

[18] H. Zhang, X. Song, and L. Shi, "Convergence and mean square stability of suboptimal estimator for systems with measurement packet dropping," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1248–1253, May 2012.

[19] E. Silva and M. Solis, "An approach to stationary state estimation with missing data," presented at the 9th IEEE Int. Conf. Control Automation, Santiago, Chile, Dec. 2011.

[20] E. Silva and S. Pulgar, "Control of LTI plants over erasure channels," *Automatica*, vol. 47, no. 8, pp. 1729–1736, August 2011.

[21] N. Elia, "Remote stabilization over fading channels," *Syst. Control Lett.*, vol. 54, no. 3, pp. 237–249, 2005.

[22] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[23] Q. Ling and M. Lemmon, "Power spectral analysis of networked control systems with data dropouts," *IEEE Trans. Autom. Control*, vol. 49, no. 6, pp. 955–960, Jun. 2004.

[24] D. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995.

[25] M. Poubelle, R. Bitmead, and M. Gevers, "Fake algebraic Riccati techniques and stability," *IEEE Trans. Autom. Control*, vol. AC-33, no. 4, pp. 379–381, Apr. 1988.

[26] A. Fioravanti, A. Goncalves, and J. Geromel, "Filter inputs with markovian lossy links: Zero or hold?," in *Proc. 9th IEEE Int. Conf. Control and Automation*, Santiago, Chile, 2011, pp. 656–661.

[27] L. Schenato, "To zero or to hold control inputs with lossy links?," *IEEE Trans. Autom. Control*, vol. 54, no. 5, pp. 1093–1099, May 2009.

Optimal Periodic Sensor Schedule for Steady-State Estimation Under Average Transmission Energy Constraint

Zhu Ren, Peng Cheng, Jiming Chen, Ling Shi, and Youxian Sun

***Abstract*—We consider periodic sensor scheduling for remote state estimation under average transmission energy constraint. The sensor decides whether or not to send its data to a remote estimator in order to meet the transmission energy constraint. The transmitted data are likely to be dropped due to the imperfect communication. An optimal periodic schedule is found via the tools from the Markov chain. Furthermore, a sufficient condition of the system dynamics, energy budget, and packet drop rate, under which the remote estimator is guaranteed to be stable, is derived. Examples are provided to show the effectiveness of results.**

***Index Terms*—Energy constraint, Markov chain, sensor scheduling, stability.**

I. INTRODUCTION

Networked control systems (NCSs) have attracted great research interest in the past decade. Typical applications can be found in autonomous vehicles, environmental monitoring, industrial automation, smart grids, etc. [1].

The sensors in wireless sensor networks (WSN) are usually powered by batteries, which can only provide limited energy for sensing, computation, and transmission, among which, the transmission energy dominates the total energy cost. Consequently, a sensor has to decide whether or not to send its current data. To save energy, the sensor may choose not to transmit. However, the estimation error of the underlying parameters, which depends on the raw sensor measurement data, may grow undesirably. Thus, it is of great interest to construct appropriate schedules of sensor data transmission so that the estimation error can be minimized under the energy constraint. Moreover, it is highly desirable that the proposed schedule can be implemented easily without much requirements of sensors.

Significant efforts have been devoted for sensor scheduling problems. For nonlinear state estimation, Baras and Bensoussan [2] considered how to schedule a set of sensors for estimating a function of an underlying parameter. For a number of processes, Walsh *et al.* [3] investigated when to schedule different processes to access the the network so that each process remains stable. Gupta *et al.* [4] considered stochastic sensor scheduling in which they defined

Manuscript received October 16, 2012; revised January 24, 2013, April 02, 2013, and May 04, 2013; accepted May 04, 2013. Date of publication May 17, 2013; date of current version November 18, 2013. This work was supported in part by the National Natural Science Foundation of China under Grants 61004060 and 61222305, the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 2011AA040101-1, the Research Grants Council (RGC) under Grants 20100101110066, 20120101110139, and NCET-11-0445, and the Fundamental Research Funds for the Central Universities under Grants 2013QNA5013 and 2013FZA5007. The work by L. Shi was supported in part by the RGC under Grant HKUST11/CRF/10. Recommended by Associate Editor L. Schenato.

Z. Ren, P. Cheng, J. Chen, and Y. Sun are with State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: zhuren@zju.edu.cn; pcheng@iipc.zju.edu.cn; jmchen@iipc.zju.edu.cn; yxsun@iipc.zju.edu.cn).

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

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/TAC.2013.2263651

a sensor selection distribution p_i (i.e., at each time, sensor i will be selected with probability p_i to observe the system) and gave the optimal values of p_i's which minimizes the upper bound of the expected steady-state estimation error covariance. Sandberg *et al.* [5] considered a heterogeneous sensor network and propose an optimal schedule by using a time-periodic Kalman filter. Similar problems were also considered by Arai *et al.* [6], [7]. In [5], the objective is to minimize the sum of the average energy and the trace of the estimation error covariance simultaneously, while in [6], [7], the objective is a quadratic cost function. Savage and Scala [8] considered the optimal sensor scheduling for scalar systems which aims to minimize the terminal estimation error covariance.

Due to the nature of wireless communication, the transmitted data packets are likely to be dropped; thus it is natural to investigate the relationship between the packet drop pattern and the estimation stability. In Sinopoli *et al.* [9], the sensor directly sends its raw measurement to the remote estimator. They proved the existence of a critical value for the packet loss rate below which the expected estimation error covariance diverges. The packet dropping was modeled as an i.i.d. Bernoulli process. Shi *et al.* [10] considered the same problem with a different performance index, the probability that the error covariance is less than an arbitrary bound. Huang and Dey [11] considered remote state estimation subject to Markovian packet drops. They defined the kth error covariance just before each successful packet reception as the the peak covariance M_k. A sufficient condition for the stability of peak covariance in the mean sense (i.e., $\sup_k \mathbb{E}[\|M_k\|] < \infty$) was provided. You *et al.* [12] extended the results to the mean square stability scenario.

In this technical note, we consider the problem of scheduling sensor data in order to provide the optimal estimation quality with the transmission energy constraint. We focus on the periodic scheduling schemes which are robust and practical to implement. Note that Hovareshti *et al.* [13] has considered a special case (the channel is perfect), while here we assume the channel introduces data packet drops which is modeled as a Bernoulli process. Shi *et al.* [14] has considered a problem with two transmission power levels, i.e., high level corresponds to perfect communication while low level results in random packet drops. Different from existing works in [13], [14], here, when the sensor decides to transmit, the packet is still possible to be dropped due to the imperfect communication. Such a problem is much more difficult as the transmitted data is never guaranteed to arrive. Moreover, such a setting is also more practical as high transmission power does not necessarily promise perfect communication.

The main contributions of this technical note are as follows.

1) For any arbitrary periodic schedule, we derive the expected average error covariance as a function of the steady-state error covariance $\overline{P}$ by constructing a corresponding Markov chain.
2) We then construct an optimal periodic schedule which minimizes the estimation error at the estimator and satisfies the energy constraint simultaneously.
3) We present a sufficient condition under which the stability of the estimator is guaranteed with the proposed optimal periodic schedule.

The remainder of the technical note is organized as follows. In Section II, we introduce the system models and problem setup along with some preliminaries on Kalman filter. In Section III, we derive the expected error covariance for any given periodic schedule. Section IV provides an optimal periodic schedule under any energy constraint. In Section V, the estimation stability condition is investigated. Examples are provided to demonstrate the results. Section VI concludes the note.

Notations. $\mathbb{Z}$ is the set of integers. $\mathbb{Z}^+$ is the set of positive integers; $\mathbb{N}$ is the set of nonnegative integers; $k \in \mathbb{N}$ is the time index. Given $m, n \in \mathbb{N}$, if $m = ln + r$, $l \in \mathbb{Z}$, $1 \le r \le n$, we use $m \bmod n$ to denote r. $\mathbb{R}^n$ is the n-dimensional Euclidian space. $\mathcal{S}_+^n$ is the set of $n \times n$ positive semidefinite matrices. We simply write $X \ge 0$, when $X \in \mathcal{S}_+^n$; write $X > 0$, when X is positive definite, and $X \ge Y$, when $X - Y \in \mathcal{S}_+^n$. For a vector $a = (a_i)$, we always mean a row vector; $|a|$ denotes its dimension. For matrix $A = (a_{ij})$, A' denotes its transpose. $\mathbf{1}_n$ is an $1 \times n$ row vector $(1, 1, \cdots, 1)$ (n may be equal to ∞). $\mathbf{0}_n$ is an $1 \times n$ row vector $(0, 0, \cdots, 0)$. $\mathbf{I}_n$ is an $n \times n$ identity matrix. $\lfloor x \rfloor$ ($x \in \mathbb{R}$) denotes the largest integer that is not larger than x, e.g., $\lfloor 0.75 \rfloor = 0$, $\lfloor 4.2 \rfloor = 4$. For functions $f, f_1, f_2 \colon \mathcal{S}_+^n \to \mathcal{S}_+^n$, $f_1 \circ f_2$ is defined as $f_1 \circ f_2 \triangleq f_1(f_2(X))$, and f^t is defined as $f^t(X) \triangleq \underbrace{f \circ f \circ \cdots \circ f}_{t \text{ times}}(X)$ with $f^0(X) = X$.

II. Problem Formulation

A. System Model

We consider a discrete linear time-invariant system

$$x_{k+1} = Ax_k + \omega_k, \quad y_k = Cx_k + \nu_k$$

where $x_k \in \mathbb{R}^n$ is the state of system, $y_k \in \mathbb{R}^m$ is the measurement obtained by the sensor, $\omega_k \in \mathbb{R}^n$ and $\nu_k \in \mathbb{R}^m$ are both zero-mean Gaussian random noises with covariances satisfying $\mathbb{E}[\omega_k \omega_j'] = \Delta_{kj} Q$, $Q \ge 0$, $\mathbb{E}[\nu_k \nu_j'] = \Delta_{kj} R$, $R > 0$, and $\mathbb{E}[\omega_k \nu_j'] = 0$, $\forall j, k$, where $\Delta_{kj} = 1$ if $k = j$ and $\Delta_{kj} = 0$ otherwise. The initial state x_0 is also a zero-mean Gaussian random vector which is uncorrelated with ω_k or ν_k and has covariance $P_0 \ge 0$. The pair $(A, \sqrt{Q})$ is controllable and (C, A) is observable.

Assume that the sensor communicates its data packet with the remote estimator via a wireless channel. Denote

$$Y_k = \{y_1, \cdots, y_k\}$$

as all the measurement data collected by the sensor from time 1 to time k. The sensor is able to estimate the state as $\hat{x}_k^s$ according to the Kalman filter[1] which results in

$$\begin{aligned} \hat{x}_k^s &= \mathbb{E}[x_k | Y_k], \\ P_k^s &= \mathbb{E}\left[(x_k - \hat{x}_k^s)(x_k - \hat{x}_k^s)' | Y_k \right]. \end{aligned}$$

We assume that the sensor is able to decide whether to transmit its $\hat{x}_k^s$ or not at each time step. When the sensor transmits $\hat{x}_k^s$ at time k (we call, this time, the sensor's decision variable is $\gamma_k = 1$), it will cost an energy Ψ. We assume the communication channel from the sensor to the remote estimator follows a Bernoulli process $\mu_k \in \{0, 1\}$: for any k, $\mathbf{Pr}(\mu_k = 1) = p \in (0, 1]$ and $\mathbf{Pr}(\mu_k = 0) = 1 - p$, where "1" denotes that a packet can be successfully received by the estimator if a packet is transmitted by the sensor and "0" means otherwise. On the other hand, the sensor is also allowed not to transmit $\hat{x}_k^s$ (we call, this time, the sensor's decision variable is $\gamma_k = 0$) for saving the energy. Typically, wireless sensor nodes always has a limited energy budget. Since the transmission power occupies a significant part of the total energy consumption (e.g., [16] shows that the transmission power is 15 mW while the CPU power is less than 2.5 mW), it is necessary to schedule the transmission of sensor data so that the performance will not degrade much while the total energy constraint is satisfied.

B. Problem of Interest

Let θ denote a schedule for the sensor's decisions γ_k's at each time. We can see the complete scheduling space is exponentially large: from time 1 to k, it consists of 2^k different policies, so it will be very difficult to analyze within the complete space when the time horizon is infinite. Moreover, in many applications, it is desirable to propose simple but

[1]Different types of sensors in the market have such computational capability, e.g., MicaZ [15].

effective scheduling schemes which will be easy to implement in sensors with limited resources. Thus, throughout this technical note, we only focus on periodic scheduling policies. A notational convenient way of expressing a periodic schedule is to use a binary vector,[2]: $\theta = (\theta_1, \theta_2, \ldots, \theta_N)$, where $\theta_1 = 1$ and $\theta_j = 0$ or 1 for all $2 \leq j \leq N$. Since the sensor's decisions under one periodic schedule can be defined by $\gamma_k = \theta_{k \bmod N}$, we call the vector θ the sensor's periodic schedule.

Under a given θ, the estimator will also calculate a state estimate $\hat{x}_k(\theta)$ and its associated error covariance $P_k(\theta)$, which will be defined in the next subsection. We use $\hat{x}_k$ instead of $\hat{x}_k(\theta)$, etc., for short, when the underlying schedule θ is clear. Let $L \in \mathbb{Z}^+$ be the time horizon. The average energy cost on the sensor side is defined as

$$J_a(\theta) \triangleq \limsup_{L \to \infty} \frac{1}{L} \sum_{k=1}^{L} \gamma_k \Psi. \tag{1}$$

And the average estimation error covariance on the estimator side in the infinite horizon is[3]

$$P_a(\theta) \triangleq \limsup_{L \to \infty} \frac{1}{L} \sum_{k=1}^{L} P_k(\theta). \tag{2}$$

Consider an energy budget $\delta \leq \Psi$. Assume that δ and Ψ are rational numbers. In this technical note, we are interested in the following problem

Problem Optimal periodic schedule (OPS)

$$\begin{aligned} &\min\nolimits_\theta P_{a(\theta)} \\ &s.t.\ J_a(\theta) \leq \delta. \end{aligned}$$

Since $P_a(\theta) \in \mathcal{S}_+^n$ (see Section II-C), we wish to find a periodic schedule θ^*, whose average energy cost on the sensor side is not greater than δ and at the estimator side $P_a(\theta) \geq P_a(\theta^*)$ for any θ with $J_a(\theta) \leq \delta$. In addition, since for symmetric semidefinite matrices, $X \geq Y$ implies $trace(X) \geq trace(Y)$, the solution of our problem θ^* is also an optimal schedule for problem $\min_{J_a(\theta) \leq \delta} trace(P_a(\theta))$. Note that since the transmitted data is never guaranteed to arrive, we have to examine $P_a(\theta)$ over the entire infinite time horizon instead of focusing on one period as in [14]. In our case, the state estimation at the estimator side may diverge if we schedule improperly or have no sufficient energy, which is also thoroughly different from the problem considered in [14]. Later in this technical note, in Section III and IV, we assume that $P_a(\theta) < \infty$ and derive an optimal schedule θ^*. In Section V, we present a condition under which $P_a(\theta^*) < \infty$.

C. Kalman Filter Preliminaries

We define the functions h and g as $h(X) \triangleq AXA' + Q$, $g(X) \triangleq X - XC'[CXC' + R]^{-1}CX$. It can be proved that if $0 \leq X \leq Y$, then $h(X) \leq h(Y)$, $g(X) \leq g(Y)$ and $g(X) \leq X$, e.g., see [10].

At the sensor's side, $\hat{x}_k^s$ and P_k^s are calculated by a Kalman filter (KF). Denote $\overline{P}$ as the steady-state error covariance, i.e., $g \circ h(\overline{P}) = \overline{P}$ with $\overline{P} \geq 0$ and $\overline{P} \neq 0$ ([17]). Then $\overline{P}$ has the following property [14]:

Property 1: If $t_1 \leq t_2$, then $h^{t_1}(\overline{P}) \leq h^{t_2}(\overline{P})$ and $h(\overline{P}) \neq \overline{P}$. Since P_k^s converges to $\overline{P}$ exponentially fast, we assume that the Kalman filter enters steady-state at the sensor side. Then $(\hat{x}_k, P_k)$ at the estimator side is simply given as

$$(\hat{x}_k, P_k) = \begin{cases} \left(\hat{x}_k^s, \overline{P}\right), & \text{if } \hat{x}_k^s \text{ is received,} \\ \left(A\hat{x}_{k-1}, h(P_{k-1})\right), & \text{otherwise.} \end{cases}$$

Hence, at time k_2, if the latest time that the estimator has received a packet is at time $k_1 (\leq k_2)$, then the estimator's error covariance $P_{k_2} = h^{k_2 - k_1}(\overline{P})$ (note that $h^0(\overline{P}) = \overline{P}$).

III. Average Error Covariance Under Any Periodic Scheme

Given a schedule θ, we use $N = |\theta|$ to denote θ's period and $M = \sum_{j=1}^{N} \theta_j$ to denote the number of 1's in the θ. Let $c = (c_1, c_2, \ldots, c_M)$ denote a row vector with M positive integer entries such that $\sum_{i=1}^{M} c_i = N$. Then every θ can be defined precisely as a function of c:

$$\theta(c) = (1, \underbrace{0, \ldots, 0}_{c_1 - 1 \text{ times}}, 1, \underbrace{0, \ldots, 0}_{c_2 - 1 \text{ times}}, \ldots, 1, \underbrace{0, \ldots, 0}_{c_M - 1 \text{ times}}) \tag{3}$$

where $c_i - 1$ denotes the number of 0's after the ith 1 in the vector θ. For example, given $c = (2, 3, 4)$, we have $\theta(c) = (1, 0, 1, 0, 0, 1, 0, 0, 0)$. Later, we write θ as $\theta(c)$ if no confusion arises. We introduce some notations that will be used throughout the rest of this technical note.

- $\Theta(\delta) = \{\theta | J_a(\theta) \leq \delta\}$;
- $\theta_\delta^* = \arg\min_{\theta \in \Theta(\delta)} P_a(\theta)$: an optimal periodic schedule of **Problem** OPS;
- $\Theta(M, N) = \{\theta | |\theta| = N, \sum_{j=1}^{N} \theta_j = M\}$[4]: the set of periodic schedules with M number of 1's over a period N;
- $\theta_{M,N}^* = \arg\min_{\theta \in \Theta(M,N)} P_a(\theta)$.

Then **Problem** OPS becomes to find the optimal vector c such that $\theta(c) = \theta_\delta^*$. Note that in order to construct an optimal periodic schedule, it is highly desirable that we can calculate an expression of $P_a(\theta)$ for any $\theta \in \Theta(M, N)$. However, such calculation becomes much more difficult in our problem setting since the transmitted data is never guaranteed to arrive at the estimator.

A. Markov Chain

Now, we will show how to construct a recursive Markov chain by properly defining and clustering the possible state of the packet sending and receiving process. Define a state $S_k \triangleq (P_k, k \bmod N)$, where the first entry represents the estimator's error covariance, and the second entry is an index which represents that P_k is caused by the sensor's decision $\gamma_k = \theta_{k \bmod N}$. Since at next time, the sensor's decision is $\gamma_{k+1} = \theta_{(k+1) \bmod N}$, we get a stochastic dynamic system:

$$S_{k+1} = f(S_k, \gamma_{k+1}, \mu_{k+1}), \qquad k = 1, 2, \ldots \tag{4}$$

where

- S_k state of the stochastic system at time k and $S_k \in \mathbf{S} = \{(h^i(\overline{P}), j) | i \in \mathbb{N}, 1 \leq j \leq N\}$.
- γ_{k+1} decision variable to be selected at time $k+1$ and $\gamma_{k+1} = \theta_{(k+1) \bmod N}$ where $\theta \in \Theta(\delta)$.
- μ_{k+1} channel's state at time $k+1$ which follows the Bernoulli process given in Section II-A.

[2]Here we require that $\theta_1 = 1$. This is because, given arbitrary θ, if we redefine $\hat{\theta} = (\theta_t, \theta_{t+1}, \ldots, \theta_N, \theta_1, \ldots, \theta_{t-1})$ as the new schedule, where $t = \min(j | \theta_j = 1)$, we have not changed the values of $J_a(\theta)$ and $P_a(\theta)$ given later.

[3]Note that the average estimation error covariance is a function of both θ and the Bernoulli process μ_k. To simplify the expressions, we will use $P_a(\theta)$ wherever no confusion will arise.

[4]We always assume that M and N are coprime, which is for simplicity of analysis. In fact, in **Algorithm** OPS which generates an optimal periodic schedule, we can see the optimal schedule is independent of M, N's common divider.

Note that (4) is a Markov chain, which represents the evolution of the error covariance estimation and $\mathbf{S}$ is the state space for this stochastic system. If $S_k = (h^i(\overline{P}), j)$, S_k will evolve to S_{k+1} caused by decision γ_{k+1} as follows.

If $\gamma_{k+1} = 1$ (transmit the packet), then

$$\left.\begin{array}{l}\mathbf{Pr}\left(S_{k+1} = \left(\overline{P}, (j+1) \bmod N\right) | S_k, \gamma_{k+1}\right) = p \\ \mathbf{Pr}\left(S_{k+1} = \left(h^{i+1}(\overline{P}), (j+1) \bmod N\right) | S_k, \gamma_{k+1}\right) = q \\ \mathbf{Pr}(S_{k+1} = \text{others}|S_k, \gamma_{k+1}) = 0\end{array}\right\} \quad (5)$$

where $q \triangleq 1 - p$. If $\gamma_{k+1} = 0$ (drop the packet)

$$\left.\begin{array}{l}\mathbf{Pr}\left(S_{k+1} = \left(h^{i+1}(\overline{P}), (j+1) \bmod N\right) | S_k, \gamma_{k+1}\right) = 1 \\ \mathbf{Pr}(S_{k+1} = \text{others}|S_k, \gamma_{k+1}) = 0.\end{array}\right\} \quad (6)$$

Note that in the stochastic system above, when $S_k = (h^i(\overline{P}), j)$, the decision $\gamma_{k+1} = \theta_{(k+1) \bmod N} = \theta_{(j+1) \bmod N}$. Let $(\cdot, j) = \{h^0(\overline{P}, j), (h^1(\overline{P}), j), (h^2(\overline{P}), j), \ldots\}$ be a subset of $\mathbf{S}$, and $\mathbb{T}_j$ be the transition probability matrix from $(\cdot, (j-1) \bmod N)$ to $(\cdot, j)$ caused by decision θ_j. That is the (i_1, i_2)th entry in the matrix $\mathbb{T}_j$ is $\mathbf{Pr}((h^{i_2-1}(\overline{P}), j)|(h^{i_1-1}(\overline{P}), (j-1) \bmod N))$. From (5) and (6), we obtain that $\mathbb{T}_j$ only has two types. When $\theta_j = 1$ [see (5)], $\mathbb{T}_j = (p\mathbf{1}'_\infty, q\mathbf{I}_\infty)$; when $\theta_j = 0$ [see (6)], $\mathbb{T}_j = (\mathbf{0}'_\infty, \mathbf{I}_\infty)$. We can see $S_k, k = 1, 2, \ldots$ follow a Markov chain with state space $\mathbf{S}$, and its transition probability matrix $\mathbb{T}$ is given as follows:

$$\mathbb{T} = \begin{array}{c} \\ (\cdot,1) \\ (\cdot,2) \\ \vdots \\ (\cdot,N-1) \\ (\cdot,N) \end{array} \begin{array}{c} \begin{array}{ccccc} (\cdot,1) & (\cdot,2) & (\cdot,3) & \cdots & (\cdot,N) \end{array} \\ \begin{pmatrix} & \mathbb{T}_2 & & & \\ & & \mathbb{T}_3 & & \\ & & & \ddots & \\ & & & & \mathbb{T}_N \\ \mathbb{T}_1 & & & & \end{pmatrix} \end{array}$$

where $\mathbb{T}_j$ is defined according to the underlying θ.

B. Average Error Covariance

In this part, we show how to calculate $P_a(\theta)$ based on the constructed Markov chain. Write $(h^i(\overline{P}), j)$ as (i, j) for short. Without loss of generality, we assume the sensor's initial decision is $\gamma_0 = \theta_N$. This implies that the initial state $S_0 \in (\cdot, N)$. Note that $\mathbb{E}[\text{total number of } (i,j) \text{ occurred in } (0, L]] = \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i,j)}$, where $\mathbb{T}^k_{S_0,(i,j)} = \mathbf{Pr}(S_k = (i,j)|S_0)$, $k \in \mathbb{N}$ (see [18, Sec. 4.3]). Define $\pi_{i,j}$ as the long-run proportion of time that the chain is in state (i, j), i.e.,

$$\pi_{i,j} = \lim_{L\to\infty} \frac{\mathbb{E}\left[\text{total number of } (i,j) \text{ occurred in } (0, L]\right]}{L}.$$

Then $\pi_{i,j} = \lim_{L\to\infty} \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i,j)}/L$. Let $\pi_{i,\cdot} = (\pi_{i,1}, \pi_{i,2}, \pi_{i,3}, \ldots, \pi_{i,N})$ $(i \in \mathbb{N})$. From the matrix $\mathbb{T}$, we can calculate all the values of $\pi_{i,j}$'s and find that, in every row vector $\pi_{i,\cdot}$, only M entries are nonzero (see Appendix A). Define an $1 \times M$ row vector $\hat{\pi}_{i,\cdot}$ whose entries are equal to the nonzero entries in the $\pi_{i,\cdot}$ with their original order. With the definition of $\pi_{i,j}$, we have

$$\sum_{j=1}^{N} \pi_{i,j} = \lim_{L\to\infty} \frac{\mathbb{E}\left[\text{total number of } h^i(\overline{P}) \text{ occurred in } (0, L]\right]}{L}.$$

Thus $P_a(\theta) = \sum_{i=0}^{\infty} h^i(\overline{P}) \sum_{j=1}^{N} \pi_{i,j}$, provided (2) converges absolutely. We now give $P_a(\theta)$ in an explicit form.

Theorem 1: For $\theta \in \Theta_{M,N}$, define two variables:

$$\mathbf{a} = \left(\underbrace{\overline{P}, \cdots, \overline{P}}_{M \text{ times}}, \underbrace{h(\overline{P}), \cdots, h(\overline{P})}_{M \text{ times}}, \cdots, \underbrace{h^i(\overline{P}), \cdots, h^i(\overline{P})}_{M \text{ times}}, \cdots\right)$$

$$\boldsymbol{\pi} = (\hat{\pi}_{0,\cdot}, \hat{\pi}_{1,\cdot}, \cdots, \hat{\pi}_{N-1,\cdot}, \hat{\pi}_{N,\cdot}, \cdots)$$

where $\mathbf{a}$ is a constant row vector with matrix $h^i(\overline{P})$ as its entries, and $\boldsymbol{\pi}$ denotes the vectorization of $\hat{\pi}_{i,\cdot}$ formed by stacking all of them into a single row vector. Then, $P_a(\theta) = \mathbf{a}\boldsymbol{\pi}'$.

IV. Structure of Optimal Periodic Schedules

In this section, we will first prove that, if θ^*_δ is an optimal schedule for Problem OPS, then $J_a(\theta^*_\delta) = \delta$, i.e., an optimal schedule must consume all the available energy.

Note that for any $\theta \in \Theta(M, N)$, from (1), $J_a(\theta) = \Psi(M/N)$. In turn, if we have a periodic schedule $\bar{\theta}$ such that $J_a(\bar{\theta}) = \delta$, then the coprime integers M, N for $\bar{\theta}$ can be obtained from the equation: $(M/N) = (\delta/\Psi)$, i.e., $\bar{\theta} \in \Theta_{M,N}$.

Algorithm Optimal periodic schedule (OPS)

Require rational numbers δ, Ψ

1: get coprime $M, N \in \mathbb{Z}^+$ such that $(M/N) = (\delta/\Psi)$

2: $l_0 = \lfloor (N - M)/M \rfloor$, $r_0 = N - M - l_0 M$

3: If $M - r_0 > r_0$, then $m_1 = M - r_0$, $n_1 = r_0$, $k_1 = l_0 + 1$, $k'_1 = l_0 + 2$; else $m_1 = r_0$, $n_1 = M - r_0$, $k_1 = l_0 + 2$, $k'_1 = l_0 + 1$.

4: $A_1 = (k_1)$, $B_1 = (k'_1)$, $i = 2$

5: **if** $n_1 = 0$ **then** goto 9 **endif**

6: $l_{i-1} = \lfloor m_{i-1}/n_{i-1} \rfloor$, $r_{i-1} = m_{i-1} - l_{i-1}n_{i-1}$. If $n_{i-1} - r_{i-1} > r_{i-1}$, then $m_i = n_{i-1} - r_{i-1}$, $n_i = r_{i-1}$, $k_i = l_{i-1} + 1$, $k'_i = l_{i-1} + 2$; else $m_i = r_{i-1}$, $n_i = n_{i-1} - r_{i-1}$, $k_i = l_{i-1} + 2$, $k'_i = l_{i-1} + 1$.

7: $A_i = (\underbrace{A_{i-1}, \cdots, A_{i-1}}_{k_{i-1} \text{ times}}, B_{i-1})$, $B_i = (\underbrace{A_{i-1}, \cdots, A_{i-1}}_{k_{i'-1} \text{ times}}, B_{i-1})$

8: **if** $n_i \neq 0$ **then** $i \leftarrow i + 1$, goto 6 **endif**

9: $\theta^*_\delta = \theta(A_i)$, $c^* = A_i$, $\bar{n} = i$

A. Necessary Condition for Optimal Periodic Schedule

We now present a necessary condition for a schedule θ^*_δ to be optimal. Since this proof is not difficult, for saving space, we omit the details.

Theorem 2: Problem OPS $\min_{J_a(\theta)\leq\delta} P_a(\theta)$ is equivalent to $\min_{J_a(\theta)=\delta} P_a(\theta)$. Thus $\theta^*_\delta = \theta^*_{M,N}$, where $(M/N) = (\delta/\Psi)$.

Therefore, when δ and Ψ are given, the original problem becomes to find the schedule $\theta^*_{M,N}$ in $\Theta(M, N)$, where $(M/N) = (\delta/\Psi)$. This reduces the search region.

B. Optimal Periodic Schedule

Lemma 1: Given one positive semidefinite matrix sequence $0 \leq X_1 \leq X_2 \leq \cdots \leq X_n$ and real sequences $0 \leq v_1 \leq v_2 \leq \cdots \leq v_n$ $(v_i \in \mathbb{R})$. If $v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}$ is an arbitrary permutation of $v_1, v_2, \ldots, v_n$, let $\Xi = v_{\sigma(1)}X_1 + v_{\sigma(2)}X_2 + \cdots + v_{\sigma(n)}X_n$. Then Ξ is least when $v_{\sigma(i)}$ and X_i are monotonic in opposite order, i.e.,

$$\Xi \geq v_n X_1 + v_{n-1} X_2 + \cdots + v_1 X_n.$$

This lemma is a matrix version of the classical Hardy–Littlewood–Pòlya rearrangement inequality [19]. It says that the lower bound of Ξ is attained only for the permutation which reverses the order of X_i. Since the entries in the vector $\mathbf{a}$ are increasing, $P_a(\theta) = \mathbf{a}\boldsymbol{\pi}'$ is like the Ξ in Lemma 1. Thus, if we can prove that the vector $\boldsymbol{\pi}$ for each $\theta \in \Theta(M, N)$ is a permutation of a const vector, and find a schedule $\theta^* \in \Theta(M,N)$ such that the corresponding entries in $\boldsymbol{\pi}$ are decreasing, then $\theta^* = \theta^*_{M,N}$. Based on this intuition, we give a condition which guarantees a schedule $\theta^*_{M,N}$ is optimal.

Lemma 2: If the underlying vector $c = (c_1, c_2, \ldots, c_M)$ in $\theta(c) \in \Theta(M,N)$ satisfies that for any $d_1, d_2 \in \{1,2,\ldots,M\}$ and $u \in \{0,1,\ldots,M-1\}$ (when $t > M$, let $c_t = c_{t \bmod M}$)

$$\left|\sum_{t=d_1}^{d_1+u} c_t - \sum_{t=d_2}^{d_2+u} c_t\right| = 0 \; or \; 1 \tag{7}$$

then $\theta(c)$ is an optimal schedule in $\Theta(M,N)$.

Proof: See Appendix B. ■

Now we are ready to present an optimal periodic schedule for Problem OPS.

Theorem 3: For any δ, Ψ, the schedule θ^*_δ generated by Algorithm OPS is an optimal periodic schedule.

Proof: See Appendix C. ■

In the 1–5 steps, we have m_1 A_1's and n_1 B_1's. Steps 6–8 aim to divide the A_i's by B_i's as "uniform" as possible. Since the algorithm will stop when $n_i = 0$ and $M > n_1 > \cdots > n_{i-1} > n_i > \cdots$, at most M iterations are needed for the calculation. What's more, the input of every iteration (i.e., steps 6–8) consists two integers m_{i-1}, n_{i-1} and two set of numbers A_{i-1}, B_{i-1}. Only two times of addition, two times of subtraction and one time of multiplication (division) are involved in each iteration. Thus Algorithm OPS has a low computational complexity. On the other hand, we can see $m_{\bar{n}}$ in the last step is the greatest common divider of N and M. This reflects that $\theta^*_{dM,dN} = \theta^*_{M,N}$, i.e., we need not consider the common divider of M and N.

Example 1: If $\delta = 8.5, \Psi = 10$, then from $M/N = \delta/\Psi = 0.85$, we have $M = 17, N = 20$. The algorithm returns the optimal schedule in 3 steps as follows. We have $\theta^*_{8.5} = \theta(A_3)$.

$m_1 = 14$ $n_1 = 3$ $k_1 = 1$ $k'_1 = 2$ $A_1 = (1), B_1 = (2)$
$m_2 = 2$ $n_2 = 1$ $k_2 = 6$ $k'_2 = 5$ $A_2 = (\mathbf{1}_5, 2), B_2 = (\mathbf{1}_4, 2)$
$m_3 = 1$ $n_3 = 0$ $k_3 = 3$ $k'_3 = 4$ $A_3 = (\mathbf{1}_5, 2, \mathbf{1}_5, 2, \mathbf{1}_4, 2)$

It is interesting to note that in our problem, "as uniform as possible" is not sufficient to guarantee the optimality. In fact, the sequence order will also directly affect the estimation performance. From the algorithm, it is clear that the solution constructed by our method is an optimal solution for [14], but the solution provided there may not be optimal here, which will also be shown in later sections.

V. Convergence Conditions

As mentioned above, the result of Section IV is based on the convergence of $P_a(\theta)$. In this section, we provide conditions under which the expected error covariance converges.

Lemma 3: If $\rho(A) < 1$, where $\rho(A) = \max_i |\sigma_i|$ and σ_i is the ith eigenvalue of matrix A, series $\sum_{i=1}^{\infty} A^i X (A')^i$ (X is an $n \times n$ matrix) converges absolutely. The condition $\rho(A) < 1$ is also necessary if X is positive definite.

Now, one sufficient condition for the convergence of $P_a(\theta^*_\delta)$ can be derived as follows, which considers a special class of the energy budget.

Theorem 4: Assume $(\delta/\Psi) = (1/l)$. When the system is controllable and observable, if $p > 1 - \rho(A)^{-2l}$, then $P_a(\theta^*_\delta)$ converges absolutely. If $Q > 0$ or $\overline{P} > 0$, the condition $p > 1 - \rho(A)^{-2l}$ is also necessary.

Proof: From Algorithm OPS, an optimal schedule is $\theta^*_\delta = \theta^*_{1,l} = (1, \mathbf{0}_{l-1})$. Then, using $P_a(\theta^*_\delta) = \mathbf{a}\boldsymbol{\pi}'$, we get $P_a(\theta^*_\delta)$

$$\begin{aligned}
&= \frac{p}{l}\sum_{i=0}^{\infty} q^i \left[h^{il}(\overline{P}) + h^{il+1}(\overline{P}) + \cdots + h^{(i+1)l-1}(\overline{P})\right] \\
&= \frac{p(l-1)}{l}\sum_{i=1}^{\infty} q^i \sum_{u=0}^{il-1} A^u Q A'^u + \frac{p}{l}\sum_{i=0}^{\infty} q^i \\
&\quad \times \left[\sum_{u=il}^{(i+1)l-1} A^u \overline{P} A'^u + (l-1)A^{il} Q A'^{il} \right. \\
&\qquad + (l-2)A^{il+1} Q A'^{il+1} + \cdots \\
&\qquad + 2A^{(i+1)l-3} Q A'^{(i+1)l-3} \\
&\qquad \left. + A^{(i+1)l-2} Q A'^{(i+1)l-2}\right]
\end{aligned} \tag{8}$$

where $q = 1 - p$. Since $\sum_{i=1}^{\infty} q^i \sum_{u=0}^{il-1} A^u Q A'^u$

$$\begin{aligned}
&= qX + q^2(X + A^l X A'^l) \\
&\quad + q^3(X + A^l X A'^l + A^{2l} X A'^{2l}) + \cdots \\
&= \frac{1}{1-q}(qX + q^2 A^l X A'^l + q^3 A^{2l} X A'^{2l} + \cdots) \\
&= \frac{q}{1-q}\sum_{i=0}^{\infty} (\sqrt{q}A^l)^i X (\sqrt{q}A'^l)^i
\end{aligned}$$

where $X = Q + AQA' + \cdots + A^{l-1}QA'^{l-1}$, when $\rho(\sqrt{q}A^l) < 1$, i.e., $p > 1 - \rho(A)^{-2l}$, from Lemma 3, the first term of (8) converges absolutely. In addition, when $p > 1 - \rho(A)^{-2l}$, the other terms of (8) will also converge. And if $Q > 0$ (indicates $X > 0$) or $\overline{P} > 0$, $p > 1 - \rho(A)^{-2l}$ is a necessary condition. ■

For the general case $(\delta/\Psi) = (M/N)$, we have the following statement.

Theorem 5: Assume $(\delta/\Psi) = (M/N)$ (M and N are coprime). Let $l = \lfloor N/M \rfloor + 1$ when $M \neq 1$, and let $l = N$ when $M = 1$. When the system is controllable and observable, if $p > 1 - \rho(A)^{-2l}$, then $P_a(\theta^*_\delta)$ will converge.

Proof: From the definition of l in the theorem, we have $(1/l) \leq (M/N)$, then $P_a(\theta(c_{1,l})) \geq P_a(\theta^*_{M,N})$ (since the schedule $\theta^*_{M,N}$ consumes more energy than $\theta(c_{1,l})$). From Theorem 4, when $p > 1 - \rho(A)^{-2l}$, $P_a(\theta(c_{1,l})) < \infty$, which indicates $P_a(\theta^*_\delta) = P_a(\theta^*_{M,N}) \leq P_a(\theta(c_{1,l})) < \infty$. This proves the theorem. ■

Example 2: Assume the system's parameters are

$$A = \begin{pmatrix} 0.1 & 0.2 \\ 1.2 & 1 \end{pmatrix} \qquad C = (1 \quad 2) \qquad Q = \begin{pmatrix} 0.3 & 0.3 \\ 0.3 & 0.3 \end{pmatrix}$$

and $R = 0.1$. Let $(M/N) = (5/12)$. We obtain an optimal periodic scheduling $\theta^*_\delta = (1,0,1,0,0,1,0,1,0,0,1,0)$. From Theorem 5, we get $l = 3$. Therefore when the arrival probability $p > 1 - \rho(A)^{-2l} = 0.69$, θ^*_δ will converge. Fig. 1 depicts the trace of $P_a(\theta^*_\delta)$ under two different arrival probabilities. It can be observed that, when $p = 0.7$, θ^*_δ converges rapidly, while if $p = 0.6$, θ^*_δ indeed diverges. In Fig. 2, we show the performance (with $p = 0.75$) of θ^*_δ against two other policies: $\theta^{(1)} = (1,1,1,1,1,0,0,0,0,0,0,0)$, $\theta^{(2)} = (1,0,1,0,1,0,1,0,0,1,0,0)$. The trace of $P_a(\theta^*_\delta)$ outperforms the others. In addition, $\theta^{(2)}$ and θ^*_δ are both optimal for the scenario in [14], but, for the scenario considered in this technical note, only θ^*_δ is the optimal one.

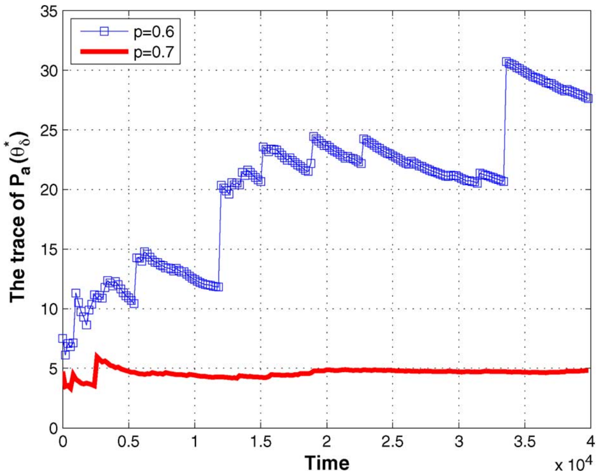

Fig. 1. Performance of $P_a(\theta_\delta^*)$ under two arrival probabilities.

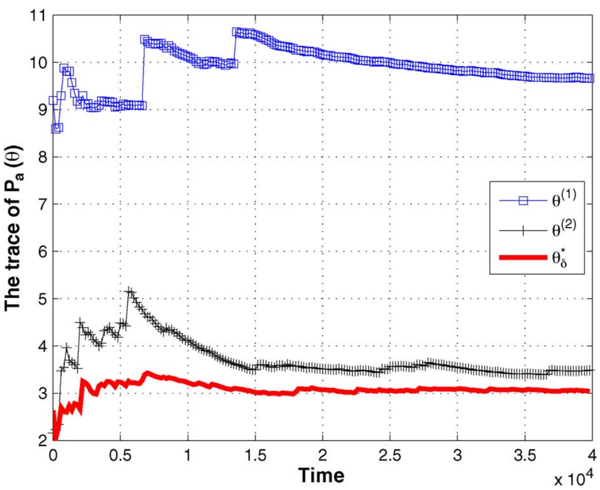

Fig. 2. Performance of different policies ($p = 0.75$).

VI. Conclusion and Future Work

In this technical note, we investigate the sensor data scheduling problem where the sensor decides whether to send its data to a remote estimator so that the estimation error is minimized while the energy constraint is satisfied. Different from previous results, the data transmission is still prone to be dropped due to the unreliable communication. We construct an optimal periodic schedule that minimizes the estimation error at the estimator side while satisfying the energy constraint. Furthermore, we are able to provide a sufficient condition under which the estimation is guaranteed to be stable. Future works along the line of this work include using the estimate value in the schedule calculation, comparing periodic schedules with other nonperiodic ones and finding the necessary and sufficient condition for the stability of the remote estimator.

Appendix A

Property 2: For a given periodic schedule $\theta(c) \in \Theta(M, N)$,

1) $\pi_{0,\cdot} = (1/N)(p, \mathbf{0}_{c_1-1}, p, \mathbf{0}_{c_2-1}, \cdots, p, \mathbf{0}_{c_M-1})$.

2) If $i \in \mathbb{Z}^+$,

$$\pi_{i,j} = \begin{cases} q\pi_{i-1,(j-1) \bmod N}, & \theta_j = 1; \\ \pi_{i-1,(j-1) \bmod N}, & \theta_j = 0. \end{cases}$$

3) For all $i = lN + r$ ($l, r \in \mathbb{N}$ and $0 \le r < N$), $\pi_{i,j} = q^{lM}\pi_{r,j}$, i.e., $\pi_{i,\cdot} = q^{lM}\pi_{r,\cdot}$. In every row vector $\pi_{i,\cdot}$, only M entries are nonzero.

Proof: 1) Assume $\theta_j = 1$. Since at time k, state S_k must go to set $(\cdot, k \bmod N)$, we have $\mathbb{T}^k_{S_0,s} = 0, \forall s \notin (\cdot, k \bmod N)$, $\sum_{s\in\mathbb{S}} \mathbb{T}^k_{S_0,s} = \sum_{s\in(\cdot,k \bmod N)} \mathbb{T}^k_{S_0,s} = 1$. Thus $\mathbb{T}^k_{S_0,(0,j)} = 0$ when $k \bmod N \ne j$. From $\mathbb{T}^k_{s,(0,j)} = p, \forall s \in (\cdot, (j-1) \bmod N)$, we will find, for any $S_0 \in (\cdot, N)$,

$$\begin{aligned} \mathbb{T}^k_{S_0,(0,j)} &= \sum_{s\in\mathbb{S}} \mathbb{T}^{k-1}_{S_0,s}\mathbb{T}_{s,(0,j)} \\ &= \sum_{s\in(\cdot,(k-1) \bmod N)} \mathbb{T}^{k-1}_{S_0,s}\mathbb{T}_{s,(0,j)} \\ &= \sum_{s\in(\cdot,(j-1) \bmod N)} \mathbb{T}^{k-1}_{S_0,s}\mathbb{T}_{s,(0,j)} \\ &= p \sum_{s\in(\cdot,(j-1) \bmod N)} \mathbb{T}^{k-1}_{S_0,s} = p \end{aligned}$$

when $k \bmod N = j$. Thus $\pi_{0,j} = \lim_{L\to\infty} \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(0,j)}/L = p/N$ when $\theta_j = 1$.

When $\theta_j = 0$, since state S_k will never go to $(0, j)$, we have $\mathbb{T}^k_{S_0,(0,j)} = 0, \forall k \in \mathbb{Z}^+$. Hence, $\pi_{0,j} = \lim_{L\to\infty} \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(0,j)}/L = 0$ when $\theta_j = 0$.

2) For any state (i, j), $i \in \mathbb{Z}^+$, $\mathbb{T}_{s,(i,j)} = 0$ when $s \ne (i-1, j-1)$. It follows that $\sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i,j)} = \sum_{k=1}^{L}\sum_{s\in\mathbb{S}} \mathbb{T}^{k-1}_{S_0,s}\mathbb{T}_{s,(i,j)} = \mathbb{T}_{(i-1,j-1),(i,j)} \sum_{k=1}^{L} \mathbb{T}^{k-1}_{S_0,(i-1,j-1)} = \mathbb{T}_{(i-1,j-1),(i,j)}[\mathbf{Pr}(S_0 = (i-1, j-1)) - \mathbb{T}^L_{S_0,(i-1,j-1)} + \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i-1,j-1)}]$. Since $\pi_{i-1,j-1} = \lim_{L\to\infty} \sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i-1,j-1)}/L$, taking limit on $\sum_{k=1}^{L} \mathbb{T}^k_{S_0,(i,j)}/L$, we have

$$\pi_{i,j} = \pi_{i-1,j-1}\mathbb{T}_{(i-1,j-1),(i,j)}, i \in \mathbb{Z}^+.$$

The proof is completed by noting that $\mathbb{T}_{(i-1,j-1),(i,j)} = q$ when $\theta_j = 1$ and $\mathbb{T}_{(i-1,j-1),(i,j)} = 1$ when $\theta_j = 0$.

3) Combining by 1) and 2), we can calculate all the values of $\pi_{i,j}$'s. Then for any $\pi_{i,j}$, assume $i = lN + r$, $0 \le r < N$. We can get $\pi_{i,j} = q^{lM}\pi_{r,j}$ by using property 2) lN times.

What's more, since $\pi_{0,\cdot}$ has M nonzero entries, from property 2), it can be derived that, in every $\pi_{i,\cdot}$ ($i \in \mathbb{Z}^+$), only M entries are nonzero by mathematical induction. This completes the proof.

Appendix B

Assume $\pi_{0,j_1}, \pi_{0,j_2}, \ldots, \pi_{0,j_M}$ are the M nonzero entries in the $\pi(0,\cdot)$. Define matrix $\Phi = (\varphi_{i,d}) = (\Phi[1], \Phi[2], \ldots, \Phi[M])$, where the dth column $\Phi[d] = (\pi_{0,j_d}, \pi_{1,j_d+1}, \ldots, \pi_{N-j_d,N}, \pi_{N-j_d+1,1}, \ldots, \pi_{i,j}, \ldots)'$, $i \ge 0$, $j = (i + j_d) \bmod N$. We have the next property.

Property 3: Given $\theta(c) \in \Theta(M, N)$, we obtain

$$\Phi = \frac{p}{N}\begin{pmatrix} \mathbf{1}'_{c_1} & \mathbf{1}'_{c_2} & \cdots & \mathbf{1}'_{c_M} \\ q\mathbf{1}'_{c_2} & q\mathbf{1}'_{c_3} & \cdots & q\mathbf{1}'_{c_1} \\ \cdots & \cdots & \cdots & \cdots \\ q^{M-3}\mathbf{1}'_{c_{M-2}} & q^{M-3}\mathbf{1}'_{c_{M-1}} & \cdots & q^{M-3}\mathbf{1}'_{c_{M-3}} \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \quad (9)$$

and $\Phi(i) = q^{lM}\Phi(r)$ when $i = lN + r$ ($1 \le r \le N$), where $\Phi(i)$ is the ith row of Φ. Φ contains all the nonzero entries in $\pi_{i,j}$'s, and $\Phi(i)$ is transformed from $\pi_{i-1,\cdot}$.

Proof: We will first prove it for the first column $\Phi[1] = (\pi_{0,1}, \pi_{1,2}, \ldots, \pi_{N-1,N}, \pi_{N,1}, \pi_{N+1,2}, \ldots)$. Consider $\pi_{0,1} = p/N$ and use the Property 2.2) repeatedly by $c_1 - 1$ times. We get $\pi_{0,1} = \pi_{1,2} = \pi_{2,3} = \cdots = \pi_{c_1-1,c_1}$, since $\theta_i = 0$ for $i = 2, 3, \ldots, c_1$. Then $(\pi_{0,1}, \pi_{1,2}, \ldots, \pi_{c_1-1,c_1})' = (1/N)p\mathbf{1}'_{c_1}$.

Now consider π_{c_1,c_1+1} and use the Property 2.2) by $c_2 - 1$ times. We have $(\pi_{c_1,c_1+1}, \pi_{c_1+1,c_1+2}, \ldots, \pi_{c_1+c_2-1,c_1+c_2})' = (1/N)pq\mathbf{1}'_{c_2}, \cdots$. Repeat this procedure until the first column $\Phi[1]$ is calculated. Other columns can be calculated in a similar way with the only difference being their start entries.

Since all the nonzero entries in $\pi_{i,j}$'s are calculated from $\pi_{0,j} = p/N$, from the preceding steps, we can see that Φ contains all the nonzero entries in $\pi_{i,j}$'s, and $\Phi(i)$ is transformed from $\pi_{i-1,\cdot}$, i.e., $\Phi(i) = q^{lM}\Phi(r)$. ■

The entries in the first N rows of (9) are

$$\varphi_{i,d} = \begin{cases} \frac{p}{N}, & \text{if } 1 \le i \le c_d, \\ \frac{pq^u}{N}, & \text{if } \sum_{t=d}^{d+u-1} c_t < i \le \sum_{t=d}^{d+u} c_t \end{cases} \tag{10}$$

where we set that, when $t > M$, $c_t = c_{t \bmod M}$. And $\Phi(i) = \hat{\pi}_{i-1,\cdot}$, so $\boldsymbol{\pi} = (\Phi(1), \Phi(2), \Phi(3), \ldots)$.

Let $\overline{\boldsymbol{\pi}} = (\hat{\pi}_{0,\cdot}, \hat{\pi}_{1,\cdot}, \ldots, \hat{\pi}_{N,\cdot})$ denote the row vectorization of the first N rows of Φ.

Lemma 4: For every schedule $\theta(c) \in \Theta(M,N)$, the vector $\boldsymbol{\pi}$ is decreasing if and only if $\overline{\boldsymbol{\pi}}$ is decreasing.

Proof: By the definition of Φ, we observe that, for any $\theta(c)$, $\Phi(N) = (1/N)pq^{M-1}\mathbf{1}_M$ and $\Phi(N+1) = (1/N)pq^M\mathbf{1}_M$. Thus vector $(\Phi(N), \Phi(N+1))$ is decreasing. Then $(\overline{\boldsymbol{\pi}}, q^M\overline{\boldsymbol{\pi}})$ is also decreasing, if $\overline{\boldsymbol{\pi}}$ is decreasing. Using this procedure repeatedly and since $\boldsymbol{\pi} = (\overline{\boldsymbol{\pi}}, q^M\overline{\boldsymbol{\pi}}, q^{2M}\overline{\boldsymbol{\pi}}, \ldots)$, the statement holds. ■

Based on these results, we get the proof of Lemma 2.

Proof: We first prove that, if $\min(\Phi(i)) \ge \max(\Phi(i+1))$ for every i, then $\theta(c)$ is optimal. From (9), we can see, for any $\theta(c) \in \Theta(M,N)$, $\boldsymbol{\pi}$ is a permutation of the const vector $(p/N)(\mathbf{1}_N, q\mathbf{1}_N, q^2\mathbf{1}_N, \ldots)$. We rearrange every row of Φ in the decreasing order: $\varphi_{i,d_1} \ge \varphi_{i,d_2} \ge \cdots \ge \varphi_{i,d_M}$, and expand them to a new row vector $\boldsymbol{\pi}_{re}$. By such rearrangement, we have not changed the value $P_a(\theta)$. Since $\boldsymbol{\pi}_{re}$ is decreasing, using Lemma 1, the schedule $\theta(c)$ is optimal.

Next, we prove that, when a schedule $\theta(c)$ satisfies condition (7), its matrix Φ will satisfy $\min(\Phi(i)) \ge \max(\Phi(i+1))$ for all i. We prove it by contradiction. If the statement does not hold, from Lemma 4, we can find two entries: $\varphi_{i,d_1} < \varphi_{i+1,d_2}$, $1 \le i \le N-1$. Without loss of generality, assume $\varphi_{i,d_1} = pq^u/N$, $\varphi_{i+1,d_2} = pq^{u-r}/N$, then $r \ge 1$ and $M - 1 \ge u \ge 1$. From (10), we have

$$\sum_{t=d_1}^{d_1+u-1} c_t < i, \quad \sum_{t=d_2}^{d_2+u-r} c_t \ge i+1,$$
$$\sum_{t=d_2}^{d_2+u-1} c_t \ge \sum_{t=d_2}^{d_2+u-r} c_t \ge i+1 > i > \sum_{t=d_1}^{d_1+u-1} c_t,$$
$$\sum_{t=d_2}^{d_2+u-1} c_t - \sum_{t=d_1}^{d_1+u-1} c_t > i+1-i = 1.$$

It contradicts with (7). Thus the proof is completed. ■

Appendix C

Proof: First, let $c^{(\bar{n})} = (k_{\bar{n}})$ and $c^{(\bar{n}-1)} = (k_{n-1}\mathbf{1}_{k_{\bar{n}}-1}, k'_{\bar{n}-1})$. Then replace every k_i in the $c^{(i)}$ by $(k_{i-1}\mathbf{1}_{k_i-1}, k'_{i-1})$ and k'_i by $(k_{i-1}\mathbf{1}_{k'_{i-1}}, k'_{i-1})$ from n to 1, iteratively. We get a series of vectors $c^{(\bar{n})}, c^{(\bar{n}-1)}, \ldots, c^{(2)}, c^{(1)}$. Note tat $c^* = c^{(1)}$.

Since $\theta((\underbrace{c^{(1)}, \ldots, c^{(1)}}_{m_{\bar{n}} \text{ times}})) \in \Theta(M,N)$, from the property of periodic schedule, we can only focus on schedule $c^{(1)}$. Then if we can prove that the vectors $c^{(i)}$ defined above satisfy the condition (7), then $c^* = c^{(1)}$ is indeed an optimal periodic schedule.

Clearly, $c^{(\bar{n})}$ and $c^{(\bar{n}-1)}$ obtained in the Algorithm OPS satisfy the condition (7). Assume $c^{(i)}$ satisfies condition (7), i.e., $|\sum_{t=d_1}^{d_1+u} c_t^{(i)} - \sum_{t=d_2}^{d_2+u} c_t^{(i)}| = 0 \ or \ 1$. We will show $c^{(i-1)}$ also holds by contradiction.

If not, without loss of generality, assume we have d_1, d_2, u such that $\sum_{t=d_1}^{d_1+u} c_t^{(i-1)} - \sum_{t=d_2}^{d_2+u} c_t^{(i-1)} = 2$. Let $\Upsilon_1 = (c_{d_1}^{(i-1)}, c_{d_1+1}^{(i-1)}, \ldots, c_{d_1+u}^{(i-1)})$ and $\Upsilon_2 = (c_{d_2}^{(i-1)}, c_{d_2+1}^{(i-1)}, \ldots, c_{d_2+u}^{(i-1)})$. Since $c^{(i-1)}$ is a vector constructed by k_{i-1} and k'_{i-1}, we have 2 more k'_{i-1} in Υ_1 than Υ_2, and $k'_{i-1} = k_{i-1} + 1$. Assume $c_{d_1+a_1}^{(i-1)}$ and $c_{d_1+b_1}^{(i-1)}$ are the first and last k'_{i-1} in the $c^{(i-1)}$, respectively; $c_{d_2-a_2}^{(i-1)}$ and $c_{d_2+u+b_2}^{(i-1)}$ are the first k'_{k-1} before $c_{d_2}^{(i-1)}$ and the first k'_{k-1} after $c_{d_2+u}^{(i-1)}$ in the $c^{(i-1)}$, respectively. Let $\overline{\Upsilon}_1 = (c_{d_1+a_1+1}^{(i-1)}, c_{d_1+a_1+2}^{(i-1)}, \ldots, c_{d_1+b_1}^{(i-1)})$ and $\overline{\Upsilon}_2 = (c_{d_2-a_2+1}^{(i-1)}, c_{d_2-a_2+2}^{(i-1)}, \ldots, c_{d_2}^{(i-1)}, \ldots, c_{d_2+u}^{(i-1)}, \ldots c_{d_2+u+b_2}^{(i-1)})$. We can see $\overline{\Upsilon}_1$ and $\overline{\Upsilon}_2$ have the same number of k'_{i-1} (assume the number is v). Then $|\overline{\Upsilon}_1| = b_1 - a_1 \le u - 1$ is a sum of v successive elements in $c^{(i)}$, and $|\overline{\Upsilon}_2| = b_2 + u + a_2 \ge u + 1$ is another sum. Since $|\overline{\Upsilon}_1| - |\overline{\Upsilon}_2| \ge 2$, it contradicts with the assumption of $c^{(i)}$. Therefore, all $c^{(i)}$ satisfy the condition (7). ■

References

[1] J. P. Hespanha, P. Naghshtabrizi, and X. Yonggang, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[2] J. S. Baras and A. Bensoussan, "Sensor scheduling problems," presented at the IEEE Conf. Decision and Control, 1988.

[3] G. C. Walsh, Y. Hong, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 3, pp. 438–446, May 2002.

[4] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[5] H. Sandberg, M. Rabi, M. Skoglund, and K. H. Johansson, "Estimation over heterogeneous sensor networks," presented at the IEEE Conf. Decision and Control, 2008.

[6] S. Arai, Y. Iwatani, and K. Hashimoto, "Fast and optimal sensor scheduling for networked sensor systems," presented at the IEEE Conf. Decision and Control, 2008.

[7] S. Arai, Y. Iwatani, and K. Hashimoto, "Fast sensor scheduling for spatially distributed heterogeneous sensors," presented at the Amer. Control Conf., 2009.

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

[9] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[10] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packetdropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, Mar. 2010.

[11] M. Huang and S. Dey, "Stability of kalman filtering with markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[12] K. You, M. Fu, and L. Xie, "Mean square stability for kalman filtering with markovian packet losses," *Automatica*, vol. 47, no. 12, pp. 2647–2657, 2011.

[13] P. Hovareshti, V. Gupta, and J. S. Baras, "Sensor scheduling using smart sensors," presented at the IEEE Conf. Decision and Control, 2007.

[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] H. Medeiros, J. Park, and A. Kak, "Distributed object tracking using a cluster-based Kalman filter in wireless camera networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 4, pp. 448–463, Aug. 2008.

[16] D. Estrin, "Wireless sensor networks tutorial part iv: Sensor network protocols," presented at the 8th Annu. Int. Conf. Mobile Computing and Networking (Mobicom), 2002.

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

[18] S. M. Ross, *Stochastic Processes*. Hoboken, NJ: Wiley, 1996.

[19] G. H. Hardy, J. E. Littlewood, and G. Pòlya, *Inequalities*. Cambridge, MA: Cambridge Univ. Press, 1952.