# Optimal Sensor Power Scheduling for State Estimation of Gauss–Markov Systems Over a Packet-Dropping Network

Ling Shi and Lihua Xie

***Abstract*—We consider sensor power scheduling for estimating the state of a general high-order Gauss–Markov system. A sensor decides whether to use a high or low transmission power to communicate its local state estimate or raw measurement data with a remote estimator over a packet-dropping network. We construct the optimal sensor power schedule which minimizes the expected terminal estimation error covariance at the remote estimator under the constraint that the high transmission power can only be used $m < T + 1$ times, given the time-horizon from $k = 0$ to $k = T$. We also discuss how to extend the result to cases involving multiple power levels scheduling. Simulation examples are the provided to demonstrate the results.**

***Index Terms*—Kalman filter, packet-dropping networks, power scheduling, remote state estimation.**

## I. Introduction

Remote state estimation has gained much interest in the past decade, and is found in a growing number of applications including sensor networks, smart grid, smart transportation systems, etc. In many of these applications, the available resources such as the communication energy and network bandwidth are limited. Furthermore, information flow across the network may be unreliable, e.g., data packets could be randomly delayed or dropped.

In this correspondence, we consider a remote state estimation problem subject to transmission energy constraint. A sensor measures the state of a process and sends its local state estimate or the measurement data over a packet-dropping network to a remote estimator. The sensor has limited communication energy and it decides whether to send the measurement data using a high transmission power or a low transmission power. We assume that using high transmission power leads to a higher packet arrival rate compared with using low transmission power. This assumption is motivated by two facts: most sensor nodes in the market have different transmission power to choose from [17], and higher transmission power leads to a higher signal-to-noise ratio at the remote estimator, which corresponds to a higher packet arrival rate [7].

Consider a time-horizon from $k = 0$ to $k = T$ and assume the sensor can only use the high transmission power $m < T + 1$ times due to the limited energy constraint. We are interested in how the sensor should schedule its transmission power at each time so that the expected terminal state estimate error covariance at the estimator is minimized. Before we state the main contributions of this correspondence, we briefly go over some related work from literature.

The majority of related work are concerned with sensor scheduling or sensor measurement scheduling. A nonlinear state estimation was studied by Baras and Bensoussan [2], where the authors considered scheduling a set of sensors so as to optimally estimate a function of an underlying parameter. Tiwari *et al.* [15] studied the problem of sensor scheduling for discrete-time state estimation using a Kalman filter. They considered two processes and one sensor and proposed schemes to determine which process that the sensor needs to observe in order to minimize the total estimation error. Shakeri *et al.* [12] studied the problem of sensor data scheduling subject to a fixed cost constraint, where the measurement data has a cost that is inversely proportional to its error covariance. Vitus *et al.* [16] considered sensor scheduling of a discrete-time system. Multiple sensors are employed, but only one sensor is allowed to take a measurement at each time. Cohen and Lesham [5] proposed a time-varying opportunistic protocol to maximize the network lifetime when the sensors used are battery-powered and non-rechargeable. Similar work has been carried out by Chen *et al.*[3], where the network lifetime is maximized by utilizing the channel information. Chhetri *et al.* [4] proposed two sensor scheduling algorithms for a target tracking problem. Krishnamurthy [9] proposed algorithms for scheduling noisy sensors for measuring the state of a single Markov chain. These algorithms aim to minimize a cost function consisting of the estimation error and the measurement cost. Dong *et al.* [6] considered the data retrieval problem in a 1-D sensor network. The performance of deterministic and random schedules are compared. A closely related work to this correspondence is by Savage and La Scala [11], where the authors considered the problem of optimal sensor measurement scheduling for first-order systems that minimizes the terminal error. Another work that is related to ours is given by the recent publication [10], where the authors formulated the data transmission scheduling problem as a finite horizon Markov decision process. The objective is to seek a transmission schedule which provides a tradeoff between transmission energy and packet loss rate subject to a delay constraint. It is proven that under some conditions, the optimal schedules is given in a threshold form which reduces the computational complexity significantly.

The main contributions of this correspondence are summarized as follows.

1) We consider power scheduling for remote state estimation of a general high-order Gauss–Markov system. To the best of our knowledge, the proposed framework is novel.
2) We consider natural constraints which are typical in wireless networks, e.g., sensor energy constraint and data packet drops.
3) We consider two scenarios in this correspondence: the sensor has sufficient or limited computation capability. For the first scenario, we derive the optimal power schedule (Theorem 3.3). For the second scenario, we give a sufficient condition under which, an optimal power schedule (Theorem 4.1) is given. Extension to scheduling of multiple power levels is also provided (Theorem 5.2).

The rest of the correspondence is organized as follows. The mathematical problem is introduced in Section II. The optimal power schedule for sensor with sufficient computation is provided in Section III and for sensor with limited computation is provided in Section IV. These results are extended to scheduling of multiple power levels in Section V. Simulation examples are provided in Section VI and some concluding remarks are given in the end.

Manuscript received June 17, 2011; revised November 20, 2011; accepted January 05, 2012. Date of publication January 16, 2012; date of current version April 13, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Amir Leshem. The work by L. Shi is supported in part by HKUST Grant RPC11EG34 and RGC Grant HKUST11/CRF/10. The work by L. Xie is supported by National Natural Science Foundation of China under Grant 61120106011.

L. Shi is with Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon 00852, Hong Kong (e-mail: eesling@ust.hk).

L. Xie is with EXQUISITUS, Centre for E-City, School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore, Singapore (e-mail: elhxie@ntu.edu.sg).

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

Digital Object Identifier 10.1109/TSP.2012.2184536

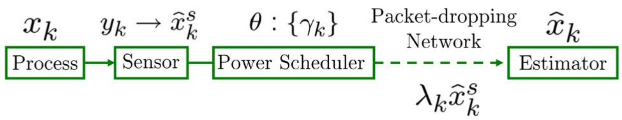


Fig. 1. Power scheduling for sensor with sufficient computation.

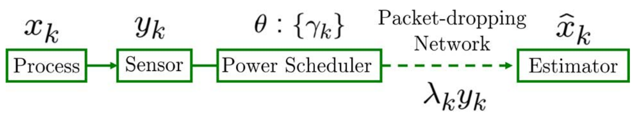


Fig. 2. Power scheduling for sensor with limited computation.

*Notations:* The non-negative integer $k$ 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 $M \in \mathbb{S}_+^n$, it is written as $M \geq 0$. $M \geq N$ if $MN \in \mathbb{S}_+^n$. $\mathbf{Pr}[\sigma]$ is the probability of a random event $\sigma$. $\mathbb{E}[X]$ is the expected value of a random variable $X$ and $\mathbb{E}[X|Y=y]$ is the conditional expectation of $X$ given that $Y=y$. For two functions $h$ and $g$ with the same domain and range, $hg$ denotes the function composition, i.e., $hg(x) = h(g(x))$. $f^t(x) \triangleq f(f^{t-1}(x))$ with $f^0(x) \triangleq x$.

## II. Problem Setup

Consider the following Gauss–Markov system:

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

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

where $x_k \in \mathbb{R}^{n_x}$ is the system state at $k$, $y_k \in \mathbb{R}^{n_y}$ is the sensor measurement, $w_k$ and $v_k$ are zero-mean white Gaussian noises with covariances $Q > 0$ and $R > 0$ respectively. The initial condition $x_0$ is zero-mean Gaussian with covariance $\Pi_0 \geq 0$. The pairs $(A, C)$ and $(A, \sqrt{Q})$ are observable and controllable. We consider two scenarios in this correspondence: 1) when the sensor has sufficient computation capability, it pre-processes $y_k$ to form the minimum mean-squared-error estimate $\hat{x}_k^s = \mathbb{E}[x_k|y_0, \ldots, y_k]$. It then sends $\hat{x}_k^s$ to a remote estimator over a packet-dropping network (Fig. 1); 2) when the sensor has limited computation, it sends the raw measurement $y_k$ to the remote estimator (Fig. 2).

Let $\hat{x}_k$ be the minimum mean-squared-error estimate of $x_k$ at the estimator based on all received data from the sensor and $P_k$ be the estimation error covariance, i.e.,

$$\hat{x}_k = \mathbb{E}[x_k|\text{all data received}] \tag{3}$$

$$P_k = \mathbb{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)'|\text{all data received}\right]. \tag{4}$$

Suppose the sensor has two transmission power levels to choose at each $k$: a high power $\Delta$ and a low power $\delta$. Let the binary variable $\gamma_k$ (1 or 0) indicate the sensor's decision that $\Delta$ or $\delta$ is chosen. The binary variable $\gamma_k$ is designed by the power scheduler at the sensor. Let another binary variable $\lambda_k$ (1 or 0) indicate that $\hat{x}_k^s$ or $y_k$ arrives at the estimator successfully or not. We assume $\{\lambda_k\}$ is an independent Bernoulli process and

$$\mathbb{E}[\lambda_k|\gamma_k = 1] = \lambda_\Delta > \lambda_\delta = \mathbb{E}[\lambda_k|\gamma_k = 0] \tag{5}$$

i.e., the packet arrival rate under the high transmission power $\Delta$ is higher than that under $\delta$.

Define

$$\Theta_T = \underbrace{\{0,1\} \times \{0,1\} \times \cdots \times \{0,1\}}_{T+1 \text{ times}}$$

as the set of all power schedules for a given time-horizon $T$. In this correspondence, we wish to find an optimal power schedule $\theta \in \Theta_T$ under which the expected terminal error covariance $\mathbb{E}[P_T(\theta)]$ is minimized subject to the constraint that the high transmission power $\Delta$ can only be chosen for $m < T+1$ times,[1] i.e.,

*Problem 2.1:*

$$\min_{\theta \in \Theta_T} \mathbb{E}[P_T(\theta)]$$

$$\text{s.t.} \quad \sum_{k=0}^{T} \gamma_k(\theta) = m.$$

In the next two sections, we will give a closed-form solution to this problem when the sensor has sufficient or limited computation capability.

## III. Optimal Power Schedule for Sensor With Sufficient Computation

When the sensor has sufficient computation, from standard Kalman filtering analysis [1], $\hat{x}_k^s$ and its estimation error covariance matrix $P_k^s = \mathbb{E}[(x_k - \hat{x}_k^s)(x_k - \hat{x}_k^s)'|y_0, \ldots, y_k]$ are given by

$$\begin{aligned}
L_k^s &= P_{k|k-1}^s C' \left[C P_{k|k-1}^s C' + R\right]^{-1}, \\
\hat{x}_k^s &= \hat{x}_{k|k-1}^s + L_k^s \left(y_k - C\hat{x}_{k|k-1}^s\right), \\
P_k^s &= P_{k|k-1}^s - P_{k|k-1}^s C' \left[C P_{k|k-1}^s C' + R\right]^{-1} C P_{k|k-1}^s, \\
\hat{x}_{k+1|k}^s &= A\hat{x}_k^s, \\
P_{k+1|k}^s &= AP_k^s A' + Q
\end{aligned}$$

where the recursion starts from $\hat{x}_{0|-1}^s = 0$ and $P_{0|-1}^s = \Pi_0$. On the estimator side, from [13], $\hat{x}_k$ and $P_k$ have the following simple forms[2]:

$$\begin{aligned}
\hat{x}_k &= \lambda_k \hat{x}_k^s + (1 - \lambda_k)A\hat{x}_{k-1} \\
P_k &= \lambda_k P_k^s + (1 - \lambda_k)(AP_{k-1}A' + Q).
\end{aligned}$$

Define the functions $h$, $\tilde{g}$, and $g$: $\mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$\begin{aligned}
h(X) &\triangleq AXA' + Q \\
\tilde{g}(X) &\triangleq X - XC'[CXC' + R]^{-1}CX \\
g(X) &\triangleq \tilde{g}h(X).
\end{aligned}$$

Then, one can verify that $P_k^s = g^k(\tilde{g}(\Pi_0))$ and

$$P_k = \begin{cases} g^k\left(\tilde{g}(\Pi_0)\right), & \text{if } \lambda_k = 1 \\ h(P_{k-1}), & \text{otherwise.} \end{cases}$$

Given two power schedules $\theta_1$ and $\theta_2$, $\mathbb{E}[P_T(\theta_1)]$ and $\mathbb{E}[P_T(\theta_2)]$ are difficult to compare in general. However, they can be quickly compared when $\gamma_k(\theta_1)$ and $\gamma_k(\theta_2)$ only differ in two adjacent time instances.

*Lemma 3.1:* Let $\theta_1$ and $\theta_2$ be two identical power schedules except that $\gamma_i(\theta_1) = 0$ and $\gamma_i(\theta_2) = 1$, and $\gamma_{i+1}(\theta_1) = 1$ and $\gamma_{i+1}(\theta_2) = 0$ for some $0 \leq i \leq T-1$. Then, $\mathbb{E}[P_T(\theta_1)] \leq \mathbb{E}[P_T(\theta_2)]$.

*Proof:* For brevity, we write $\lambda_{T\cdots0}$ as the packet arrival sequence $(\lambda_T, \ldots, \lambda_0)$, and $\lambda_{T\ldots0\setminus i}$ as $(\lambda_T, \ldots, \lambda_{i+1}, \lambda_{i-1}, \ldots, \lambda_0)$. Let[3]

$$f_k(X) = \begin{cases} g^k\left(\tilde{g}(\Pi_0)\right), & \text{if } \lambda_k = 1 \\ h(X), & \text{otherwise.} \end{cases}$$

Further, write $f_{T\cdots0} = f_T \cdots f_0$. Notice that we write the packet arrival sequence in reverse order to facilitate the writing of function compositions. Let $\mathbf{Pr}(\lambda_{T\ldots0}|\theta)$ be the probability of the arrival sequence $(\lambda_T, \ldots, \lambda_0)$ for a given schedule $\theta$. Similarly $\mathbf{Pr}(\lambda_{T\ldots0\setminus i}|\theta)$ is the probability of the arrival sequence $(\lambda_T, \ldots, \lambda_{i+1}, \lambda_{i-1}, \ldots, \lambda_0|\theta)$.

[1]It is not difficult to show that the optimal schedule remains the same if the constraint is changed to $\sum_{k=0}^{T} \gamma_k(\theta) \leq m$. In other words, more higher power is always beneficial for reducing the estimation error.

[2]If $\lambda_0 = 0$, $P_0 = \Pi_0$.

[3]$f_0(\Pi_0) = \Pi_0$ if $\lambda_0 = 0$, and $f_0(\Pi_0) = \tilde{g}(\Pi_0)$ if $\lambda_0 = 1$.

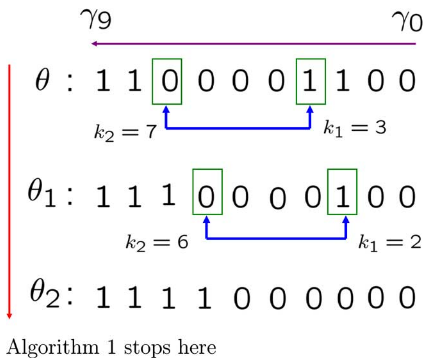


Fig. 3. Example of Algorithm 1.

Denoting $\overline{\lambda}_\Delta = 1 - \lambda_\Delta$ and $\overline{\lambda}_\delta = 1 - \lambda_\delta$ and using the above notations, we can write $\mathbb{E}[P_T(\theta_1)]$ and $\mathbb{E}[P_T(\theta_2)]$ as

$$\begin{aligned}\mathbb{E}[P_T(\theta_1)] =& \sum_{\lambda_{T\ldots 0}} \mathbf{Pr}(\lambda_{T\ldots 0}|\theta_1) f_{T\cdots 0}(\Pi_0)\\ =& \overline{\lambda}_\Delta \overline{\lambda}_\delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_1) f_{T\ldots i+2} h h f_{i-1\ldots 0}(\Pi_0)\\ & + \overline{\lambda}_\Delta \lambda_\delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_1) f_{T\ldots i+2} h g^i(\Pi_0)\\ & + \lambda_\Delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_1) f_{T\ldots i+2} g^{i+1}(\Pi_0)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[P_T(\theta_2)] =& \overline{\lambda}_\delta \overline{\lambda}_\Delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_2) f_{T\ldots i+2} h h f_{i-1\ldots 0}(\Pi_0)\\ & + \overline{\lambda}_\delta \lambda_\Delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_2) f_{\ldots i+2} h g^i(\Pi_0)\\ & + \lambda_\delta \sum_{\lambda_{T\ldots 0\setminus i+1 \setminus i}} \\ & \times \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_2) f_{T\ldots i+2} g^{i+1}(\Pi_0).\end{aligned}$$

Since $\overline{\lambda}_\delta \lambda_\Delta - \overline{\lambda}_\Delta \lambda_\delta = \lambda_\Delta - \lambda_\delta$, and $\theta_1$ and $\theta_2$ are identical for $k \in [0, i-1]$[4] and for $k \in [i+1, T]$, we conclude that

$$\begin{aligned}&\frac{\mathbb{E}[P_T(\theta_2)] - \mathbb{E}[P_T(\theta_1)]}{\lambda_\Delta - \lambda_\delta}\\ &= \sum_{\lambda_{T\ldots 0\setminus i+1\setminus i}} \mathbf{Pr}(\lambda_{T\ldots 0\setminus i+1\setminus i}|\theta_1)\\ &\quad \times \left[ f_{(\lambda_{T\ldots i+2})} h g^i(\Pi_0) - f_{(\lambda_{T\ldots i+2})} g^{i+1}(\Pi_0)\right]\\ &\geq 0\end{aligned}$$

where the last inequality is due to Lemma A.1. ■

[4] $[0, i-1] \triangleq \emptyset$ when $i = 0$.

Lemma 3.1 can be relaxed to allow quick comparison of $\mathbb{E}[P_T(\theta_1)]$ and $\mathbb{E}[P_T(\theta_2)]$ when the two different time instances are not adjacent. The proof is straightforward and is omitted.

*Corollary 3.2:* Let $\theta_1$ and $\theta_2$ be two identical power schedules except that $\gamma_i(\theta_1) = 0$ and $\gamma_i(\theta_2) = 1$, and $\gamma_j(\theta_1) = 1$ and $\gamma_j(\theta_2) = 0$ for some $0 \leq i < j \leq T$. Then $\mathbb{E}[P_T(\theta_1)] \leq \mathbb{E}[P_T(\theta_2)]$.

**Algorithm 1:** Power Schedule Iteration

$t := 0$
$\theta_t := \theta$
**while** $\prod_{k=T-m+1}^{T} \gamma_k(\theta_t) = 0$
    $t := t + 1.$
    $\theta_t := \theta_{t-1}.$
    $k_1 := \max\{k \leq T - m : \gamma_k(\theta_t) = 1\}.$
    $k_2 := \max\{k > T - m : \gamma_k(\theta_t) = 0\}.$
    $\gamma_{k_1}(\theta_t) := 0.$
    $\gamma_{k_2}(\theta_t) := 1.$
**end while**

With the previous results, we present an optimal power schedule in the following theorem.

*Theorem 3.3:* When the sensor has sufficient computation and sends its local state estimate $\hat{x}_k^s$ to the remote estimator, an optimal power schedule $\theta^*$ to Problem 2.1 is given by

$$\theta^* = \{\gamma_0 = \cdots = \gamma_{T-m} = 0, \gamma_{T-m+1} = \cdots = \gamma_T = 1\} \tag{6}$$

i.e., under $\theta^*$, the $m$ high power levels are scheduled in the last $m$ time steps.

*Proof:* Consider a general power schedule $\theta$ different than $\theta^*$. Construct a sequence of power schedules $\{\theta_t : t = 1, 2, \ldots\}$ starting from $\theta$ according to Algorithm 1. For any feasible power schedule, there are exactly $m$ $\gamma_k$'s such that $\gamma_k = 1$. Therefore Algorithm 1 stops after $d \leq m$ iterations. It is easy to see that $\theta^* = \theta_d$. Fig. 3 shows an example with $T = 9$, $m = 4$, and $d = 2$. From Corollary 3.2, we obtain

$$\mathbb{E}[P_T(\theta)] \geq \mathbb{E}[P_T(\theta_1)] \geq \cdots \geq \mathbb{E}[P_T(\theta_d)] = \mathbb{E}[P_T(\theta^*)].$$

Thus, $\theta^*$ is indeed optimal. ■

## IV. Optimal Power Schedule for Sensor with Limited Computation

From [14], if the sensor has limited computation and sends the raw measurement $y_k$ to the remote estimator, $\hat{x}_k$ and $P_k$ are computed through a *modified Kalman filter*, e.g., when $y_k$ is received, a normal Kalman filter is implemented; otherwise, only the time update step of the Kalman filter is carried out. Using the same $h$ and $g$ defined in Section III, $P_k$ can be written as

$$P_k = \begin{cases} h(P_{k-1}), & \text{if } \lambda_k = 0\\ g(P_{k-1}), & \text{if } \lambda_k = 1.\end{cases}$$

We have the following main result on the optimal power schedule.

*Theorem 4.1:* When the sensor has limited computation and sends the raw measurement $y_k$ to the remote estimator, then the schedule $\theta^*$ given by (6) is still optimal for Problem 2.1 if the following holds:

$$\forall X \geq 0, \tilde{g}h(X) \leq h\tilde{g}(X). \tag{7}$$

*Proof:* Notice that we only need to prove

$$\mathbb{E}[P_T(\theta_1)] \leq \mathbb{E}[P_T(\theta_2)]$$

if $\theta_1$ and $\theta_2$ are two identical power schedules except that $\gamma_i(\theta_1) = 0$ and $\gamma_i(\theta_2) = 1$, and $\gamma_{i+1}(\theta_1) = 1$ and $\gamma_{i+1}(\theta_2) = 0$ for some $0 \leq i \leq T-1$. The remainder of the proof is analogous to that of Lemma 3.1, Corollary 3.2, and Theorem 3.3. Now using the same notations defined in the proof to Lemma 3.1 except that $f_k$ is redefined as[5]

$$f_k(X) = \begin{cases} g(X), & \text{if } \lambda_k = 1, \\ h(X), & \text{otherwise,} \end{cases}$$

we obtain the following:

$$\begin{aligned}
\mathbb{E}[P_T(\theta_1)] = & \sum_{\lambda_{T\dots 0}} \mathbf{Pr}(\lambda_{T\dots 0}|\theta_1) f_{T\cdots 0}(\Pi_0) \\
= & \overline{\lambda}_\Delta \overline{\lambda}_\delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}\theta_1) f_{T\cdots i+2} hh f_{i-1\cdots 0}(\Pi_0) \\
& + \overline{\lambda}_\Delta \lambda_\delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_1) f_{T\cdots i+2} hg f_{i-1\cdots 0}(\Pi_0) \\
& + \lambda_\Delta \overline{\lambda}_\delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_1) f_{T\cdots i+2} gh f_{i-1\cdots 0}(\Pi_0) \\
& + \lambda_\Delta \lambda_\delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_1) f_{T\cdots i+2} gg f_{i-1\cdots 0}(\Pi_0), \\
\mathbb{E}[P_T(\theta_2)] = & \overline{\lambda}_\delta \overline{\lambda}_\Delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_2) f_{T\cdots i+2} hh f_{i-1\cdots 0}(\Pi_0) \\
& + \overline{\lambda}_\delta \lambda_\Delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_2) f_{T\cdots i+2} hg f_{i-1\cdots 0}(\Pi_0) \\
& + \lambda_\delta \overline{\lambda}_\Delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_2) f_{T\cdots i+2} gh f_{i-1\cdots 0}(\Pi_0) \\
& + \lambda_\delta \lambda_\Delta \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \\
& \times \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_2) f_{T\cdots i+2} gg f_{i-1\cdots 0}(\Pi_0).
\end{aligned}$$

Since $\overline{\lambda}_\delta \lambda_\Delta - \overline{\lambda}_\Delta \lambda_\delta = \lambda_\Delta - \lambda_\delta$, and $\theta_1$ and $\theta_2$ are identical for $k \in [0, i-1]$ and for $k \in [i+1, T]$, we conclude that

$$\begin{aligned}
& \frac{\mathbb{E}[P_T(\theta_2)] - \mathbb{E}[P_T(\theta_1)]}{\lambda_\Delta - \lambda_\delta} \\
& = \sum_{\lambda_{T\dots 0\backslash i+1\backslash i}} \mathbf{Pr}(\lambda_{T\dots 0\backslash i+1\backslash i}|\theta_1) \\
& \quad \times [f_{T\cdots i+2} hg f_{i-1\cdots 0}(\Pi_0) - f_{T\cdots i+2} gh f_{i-1\cdots 0}(\Pi_0)] \\
& \geq 0
\end{aligned}$$

where the last inequality is due to (7) as

$$hg(X) = h\tilde{g}(h(X)) \geq \tilde{g}h(h(X)) = gh(X), \; \forall X \geq 0.$$

■

[5] $\boldsymbol{f}_0(\boldsymbol{\Pi}_0) = \boldsymbol{\Pi}_0$ if $\boldsymbol{\lambda}_0 = \mathbf{0}$ and $\boldsymbol{f}_0(\boldsymbol{\Pi}_0) = \tilde{\boldsymbol{g}}(\boldsymbol{\Pi}_0)$ otherwise.

*Remark 4.2:* Notice that (7) is only a sufficient condition for $\theta^*$ to be optimal. We will provide an example in Section VI which shows that even if this condition is violated, $\theta^*$ may still be optimal. Finding a necessary condition of the optimal power schedule, however, is difficult due to the huge solution space when $T$ is large as well as the nonlinearity of the function $g$.

We now give two conditions for (7) to hold: 1) the system is of first order, i.e., $A \in \mathbb{R}$, with $|A| \geq 1$; 2) the system is of higher order with $(C'R^{-1}C)^{-1} \leq Q$. When $A \in \mathbb{R}$ with $|A| \geq 1$, straightforward computation shows that (7) holds. When $(C'R^{-1}C)^{-1} \leq Q$, then

$$\begin{aligned}
h\tilde{g}(X) &= A\tilde{g}(X)A' + Q \\
&\geq Q \geq (C'R^{-1}C)^{-1} \\
&\geq \left[(h(X))^{-1} + C'R^{-1}C\right]^{-1} = \tilde{g}h(X)
\end{aligned}$$

where the last equality is from the well-known matrix inversion lemma [8]. The condition $(C'R^{-1}C)^{-1} \leq Q$ means intuitively that the sensor has a smaller noise covariance $R$ (weighted by $C$) than the process noise covariance $Q$, i.e., the sensor provides a relatively accurate measurement. Then, it is optimal to assign the $m$ high power levels in the last $m$ steps to minimize the terminal error covariance.

## V. Discussion

In the previous two sessions, we investigate the case when the sensor has two power levels. In this section, we extend the results to multiple power levels. Consider the sensor has $L$ power levels $\{\delta_1, \dots, \delta_L\}$ which satisfy $0 < \delta_1 < \cdots < \delta_L$. Let $\gamma_k = \{1, \dots, L\}$ denote the sensor's decision on which power level to use at time $k$, and $\lambda_k \in \{0,1\}$ be the indicator whether or not the data from the sensor arrives at the remote estimator at time $k$. Define $\lambda_{\delta_i} \triangleq \mathbb{E}[\lambda_k|\gamma_k = i]$ which satisfy $0 \leq \lambda_{\delta_1} < \cdots < \lambda_{\delta_L} \leq 1$. Assume within time horizon $T$, each power level $\delta_i$ can only be used for $m_i$ times. A power schedule $\theta \in \Theta$ specifies the value of $\gamma_k$ for $k \in [0, T]$. Denote $\mathrm{Count}_i(\theta)$ as the number of $\gamma_k$'s such that $\gamma_k = i$ under $\theta$. Consider the following power scheduling problem:

*Problem 5.1:*

$$\begin{aligned}
& \min_{\theta \in \Theta_T} \mathbb{E}[P_T(\theta)] \\
& \text{s.t.} \quad \mathrm{Count}_i(\theta) = m_i, i = 1, \dots, L.
\end{aligned}$$

Notice that Problem 2.1 is a special case of Problem 5.1 when $L = 2$, $\delta_1 = \delta$, $\delta_2 = \Delta$, $\lambda_{\delta_1} = \lambda_\delta$, $\lambda_{\delta_2} = \lambda_\Delta$, and $m_2 = m$. We have the following result based on Theorem 3.3 and Theorem 4.1.

*Theorem 5.2:* Consider the following $\theta^*$ under which the $\gamma_k$'s take the following values: $\gamma_0 = \cdots = \gamma_{m_1-1} = 1$, $\gamma_{m_1} = \cdots = \gamma_{m_1+m_2-1} = 2, \dots$, and $\gamma_{T-m_L+1} = \cdots = \gamma_T = L$. Then

1) when the sensor has sufficient computation and sends its local state estimate $\hat{x}_k^s$ to the remote estimator, $\theta^*$ is an optimal power schedule to Problem 5.1;
2) when the sensor has limited computation and sends its raw measurement $y_k$ to the remote estimator, $\theta^*$ is an optimal power schedule to Problem 5.1 if (7) holds.

The proof is similar to that of Theorem 3.3 and 4.1. Given a general power schedule $\theta$, we can always construct a sequence of power schedules $\theta_t$, $t = 1, \dots, d$ with decreasing terminal error covariance such that $\theta_1 = \theta$ and $\theta_d = \theta^*$, where $\theta_j$ and $\theta_{j+1}$ are the same except that for some $k_1 < k_2$, $\gamma_{k_1}(\theta_j) = \gamma_{k_2}(\theta_{j+1}) > \gamma_{k_2}(\theta_j) = \gamma_{k_1}(\theta_{j+1})$. The comparison of $\theta_j$ and $\theta_{j+1}$ reduces to comparing two power levels only, thus all previously developed results can be used. Due to the space limitation, we omit the proof.

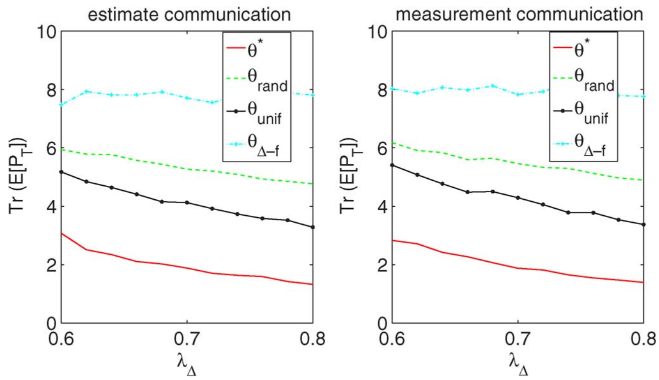


Fig. 4. Trace of $\mathbb{E}[\boldsymbol{P}_k]$ as a function of $\boldsymbol{\lambda}_\Delta$. Here condition (7) is satisfied.

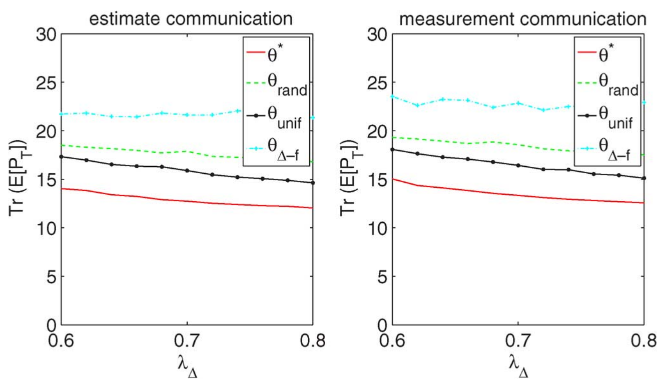


Fig. 5. Trace of $\mathbb{E}[\boldsymbol{P}_k]$ as a function of $\boldsymbol{\lambda}_\Delta$. Here condition (7) is not satisfied.

## VI. Example

Consider system (1) and (2) with

$$A = \begin{bmatrix} 1.2 & 0.1 \\ 0 & 1 \end{bmatrix},\ C = Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ R = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

Notice that (7) holds for this system. We run a Monte Carlo simulation for $T = 30$, $m = 5$, $\lambda_\delta = 0.4$ and $\lambda_\Delta \in \{0.6, 0.8\}$ and compare the performance of $\theta^*$ with the following three common schedules:

1) Random schedule $\theta_{\mathrm{rand}}$: At each $k$, $\Delta$ power level is selected with probability $\frac{m}{T}$ until at certain time it has been selected $m$ times (or $\delta$ power has been selected $T + 1 - m$ times), in which case, for all remaining times, $\delta$ (or $\Delta$) power will be selected.
2) Uniform schedule $\theta_{\mathrm{unif}}$: $\Delta$ power level is scheduled uniformly between 0 to $T$. For example, when $T = 30$ and $m = 5$, $\Delta$ power is chosen at times 5, 10, 15, 20, and 25.
3) $\Delta$-first schedule $\theta_{\Delta - f}$: $\Delta$ power level is scheduled during the first $m$ times.

Fig. 4 plots the trace of $\mathbb{E}[P_T]$ as a function of $\lambda_\Delta$ under the four power schedules when the sensor has sufficient (estimate communication) or limited computation (measurement communication). Clearly, the trace of $\mathbb{E}[P_T]$ under $\theta^*$ is the smallest among the four schedules.

We also run a simulation for the scenario when $y_k = [1\ 0]x_k + v_k$ with $R = 0.5$. For this system, the condition (7) is not satisfied. However as Fig. 5 shows, the trace of $\mathbb{E}[P_T]$ under $\theta^*$ is still the smallest among the four schedules.

## VII. Conclusion

Sensor power scheduling for remote state estimation is considered in this correspondence. An optimal power schedule is derived when the sensor has sufficient or limited computation capability. Under this optimal power schedule, the expected terminal error covariance at the remote estimator is minimized.

There are many interesting directions to explore along the line of this work: find the optimal schedule when the packet drops are governed by a Markov chain; consider optimal power schedule to minimize the average estimation error; investigate power schedule in a multi-sensor scenario.

## Appendix

*Lemma A.1.:* The functions $h$ and $g$ have the following properties: $\forall 0 \le X \le Y$, $h(X) \le h(Y)$, $g(X) \le g(Y)$, and $g(X) \le h(X)$.

*Proof:* See Lemma 1 in [14]. ■

## References


[1] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
[2] J. S. Baras and A. Bensoussan, "Sensor scheduling problems," presented at the IEEE Conf. Decision Control, Austin, TX, 1988.
[3] 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.
[4] 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, Jun. 2007.
[5] 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, Oct. 2010.
[6] 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, Dec. 2007.
[7] *Sensor Networks and Configuration: Fundamentals, Standards, Platforms, and Applications*, N. P. Mahalik, Ed. New York: Springer, 2007.
[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[9] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden markov model sensors," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1382–1397, Jun. 2002.
[10] M. Ngo and V. Krishnamurthy, "Optimality of threshold policies for transmission scheduling in correlated fading channels," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2474–2483, 2009.
[11] 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, 2009.
[12] 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.
[13] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 594–604, 2010.
[14] 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.
[15] A. Tiwari, M. Jun, D. E. Jeffcoat, and R. M. Murray, "Analysis of dynamic sensor coverage problem using Kalman filters for estimation," in *Proc. 16th IFAC World Congr.*, 2005.
[16] 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, 2010.
[17] J. Xiao, S. Cui, Z. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.