Schedule Communication for Decentralized State Estimation

Chao Yang, Junfeng Wu, Wei Zhang, and Ling Shi

***Abstract*—This paper considers decentralized state estimation subject to communication constraints. A group of agents measure the state of a process and obtain their state estimates by exchanging data with each other. Due to the communication constraint, only a few communication channels are available. The main objective of this paper is to allocate these channels among the agents so as to minimize their average estimation errors. We provide optimal allocation strategies for agents having the homogeneous and heterogeneous sensing capabilities, respectively.**

***Index Terms*—Channel allocation, decentralized Kalman filtering, networked state estimation.**

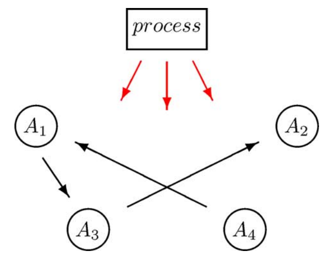


Fig. 1. Four agents observe a process with three available channels among them.

I. Introduction

THE beginning of the new century saw the burst of interest in the area of networked systems. In one networked system, sensors and computation center communicate with each other through shared communication networks. It has played an increasingly important role in the fields such as unmanned vehicle, surveillance, environment monitoring, and smart grid [1], [2]. New issues arise, however, despite the many advantages they offer. For example, the limited bandwidth of the shared channels may prevent communication between the sensors, thus leading to degraded estimation quality.

We consider in this paper a discrete linear time-invariant process with multiple agents, which estimate the process state based on their own measurement data and the received data from the other agents (Fig. 1).

If there are sufficient communication channels, each agent is able to receive the data from all the others, and the estimates made by these agents are identical and have the least estimation error. The HART (Highway Addressable Remote Transducer) Protocol [3], for example, is one popular global standard for sending and receiving digital information between devices and monitoring system. However, in most applications, the communication channels are limited and only a subset of the agents can communicate with another subset of the agents at each time, as a result of which, the estimation errors at each agent will be in general different and the overall estimation error depends on the allocation of these communication channels. The main objective of this work is to look for a channel allocation strategy which minimizes the average estimation error of these agents. The idea is applicable but not limited to the following real-world applications.

1) *Military Wireless Sensor Networks:* Wireless sensor network are often used for military purposes such as monitoring remote militant activities. The sensor network enables detection of enemy movement, identification of enemy forces and analysis of their movement and progress. How to share the information among the sensors are important to the overall performance of the military tasks. Due to limited bandwidth, not all sensor can communicate with each other at each time, thus it is important to schedule the communication among the sensors so as to achieve a desired performance (e.g., minimizing the tracking error, etc).
2) *Indoor Environmental Monitoring:* Consider one scenario of measuring the temperature and humidity level inside an office using a group of sensors. The sensors can get access to the measurements of others via a few different wireless channels to avoid potential signal interference (which may be, for example, several time slots under a TDMA protocol, or several different frequencies under a FDMA protocol.), and based on the measurements collected, each sensor can compute the estimate of the temperature and humidity locally. Again, here it is important to schedule who talks to whom at each time in order to minimize the estimation errors at each sensor node.

Estimation under limited resource constraints have been studied extensively in literature. We present a few that are mostly related to our current work. More related works can be found in the references therein.

Xu and Hespanha [4], [5] discussed the controlled communication problem over a network subject to bandwidth constraints

Manuscript received June 18, 2012; revised October 07, 2012 and January 22, 2013; accepted February 03, 2013. Date of publication February 08, 2013; date of current version April 23, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Y-W. Peter Hong. The work of C. Yang, J. Wu, and L. Shi was supported by a HKUST grant RPC11EG34.

C. Yang, 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: yangc@ust.hk; jfwu@ust.hk; eesling@ust.hk).

W. Zhang is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: zhang@ece.osu.edu).

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.2013.2246157

and communication delays. Krishnamurthy and Djonin [6] considered a model in which only a finite number of sensors can be selected to take measurements due to operational contraints, aiming at minimizing a cost function consisting of estimation errors and measurement costs. Chen *et al.* [7] addressed one lifetime maximization issue. To maximize the lifetime of a sensor network, they provided the schedule to choose the group of sensors to communicate directly with the access point. Chhetri *et al.* [8] proposed two sensor scheduling algorithms for a target tracking problem. Mo *et al.* [9] considered the problem to select a subset of sensors to communicate to a fusion center dynamically at each time step, with the purpose to minimize the asymptotic expected covariance matrix of the estimation error. They proposed a stochastic sensor selection algorithm that randomly selects a subset of sensors according to a certain probability distribution.

Due to the complexity of sensor scheduling problems, it is difficult to obtain an explicit solution in general, and most of the literature proposed different algorithms to tackle the problem. Explicit solutions, however, are always of special desire, attraction, and interest for scholars, driving them to explore all the possibilities for a nice form of optimal solutions. In the area of sensor scheduling, many researchers attempt to find an explicit solution for different system models and different objective objective functions. Savage and La Scala [10] investigated a sensor measurement scheduling problem, which aims at minimizing the terminal estimation error covariance under the constraint that only $n < N$ measurements could be taken with a finite time horizon N, and provided the optimal schedule in an explicit form. Although it is a relatively simple result, it may be viewed as a significant start. After that many other scholars moved on and made their contributions. In [11], Shi and Zhang looked at the periodic sensor scheduling problems involving two sensors measuring the state of a discrete-time linear process with limited resources. Yang and Shi [12] considered finite time-horizon sensor measurement scheduling using the average estimation error covariance as the performance metric, and proposed a necessary condition of the optimal schedules.

The work considered in this paper also follows this direction, and manages to give an explicit solution for the channel allocation problem. The main contributions of this paper are summarized as follows.

1) For a multi-dimensional system containing agents with identical sensing capabilities, we provide a sufficient and necessary condition for an channel allocation strategy to be optimal (Theorem 3.5). Based on this condition, an optimal allocation strategy is constructed.
2) For agents having different sensing capabilities, a first-order system is mainly considered and an optimal allocation is proposed.

The remainder of the paper is organized as follows. Section II presents the mathematical problem. Sections III and IV introduces the main results. Conclusions and future work are given in the end.

Notations: $\mathbb{Z}_+$ is the set of non-negative integers. $\mathbb{N}$ is the set of positive integers. $k \in \mathbb{Z}_+$ is the time index. $\mathbb{R}$ is the set of real numbers. $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{S}^n_+$ (and $\mathbb{S}^n_{++}$) is the set of n by n positive semi-definite matrices (and positive definite matrices). When $X \in \mathbb{S}^n_+$ (and $\mathbb{S}^n_{++}$), it is written as $X \geq 0$ (and $X > 0$). $X \geq Y$ if $X - Y \in \mathbb{S}^n_+$. $\boldsymbol{E}[\cdot]$ is the expectation of a random variable and $\boldsymbol{E}[\cdot\,|\,\cdot]$ is the conditional expectation. $\mathrm{Tr}(\,\cdot\,)$ is the trace of a matrix. $\lfloor x \rfloor$ denotes the largest integer which is smaller or equal to x. For functions f, f_1, f_2 with appropriate domains, $f_1 f_2(x)$ stands for the function composition $f_1(f_2(x))$, and $f^n(x) \triangleq f(f^{n-1}(x))$ with $f^0(x) \triangleq x$.

II. Problem Setup

A. System Model

Consider a discrete linear time-invariant process observed by N agents:

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

$$y^i_k = C_i x_k + v^i_k, \quad i = 1, 2, \ldots, N, \tag{2}$$

where $x_k \in \mathbb{R}^{n_x}$ is the process state vector at time k, $y^i_k \in \mathbb{R}^{n_{y,i}}$ is the measurement taken by the ith agent, $w_k \in \mathbb{R}^{n_x}$ and $v^i_k \in \mathbb{R}^{n_{y,i}}$ are zero-mean Gaussian random vectors with $\boldsymbol{E}[w_k w'_j] = \delta_{kj} Q$ $(Q \geq 0)$, $\boldsymbol{E}[v^i_k (v^i_j)'] = \delta_{kj} R_i$ $(R_i > 0)$, $\boldsymbol{E}[w_k (v^i_j)'] = 0$ $\forall j, k$. The initial state x_0 is a zero-mean Gaussian random vector that is uncorrelated with w_k and v^i_k for all k and i and has covariance $\Pi_0 \geq 0$. The pair (A, C_i) is assumed to be observable and $(A, \sqrt{Q})$ is controllable.

B. Channel Model

The agents can share their data via a few communication channels, which are to be allocated. Define the *allocation variable* $\gamma_{ij} \in \{0, 1\}$ as follows: when $i \neq j$, $\gamma_{ij} = 1$ represents that the ith agent receives data from the jth one, and $\gamma_{ij} = 0$ indicates that the ith agent does not receive data from the jth one. Since an agent is always able to access its own measurements, we have $\gamma_{ii} = 1$. Define $\Gamma = [\gamma_{ij}]$ as the *channel allocation matrix*, which is to be designed.

We consider directed channels in this paper. A directed channel from agent i to agent j only allows the information from agent i to be sent to agent j.

C. Estimation Process

At each time k, all the agents first locally predict the state x_k. After that they transmit their local measurements to and receive those from their neighboring agents via the feasible channels. After the data communication, they update their local estimates by including their local measurement plus the newly received ones.

For the ith agent, denote $\hat{x}^i_{k\,|\,k-1}$ as the *a priori* estimate of x_k, which is the predicted state estimate, and $\hat{x}^i_k$ as the *a posteriori* estimate of x_k after updating both the measurements taken locally and sent by the other agents. Further denote $P^i_{k\,|\,k-1}$ and P^i_k as the estimation error covariance matrices of $\hat{x}^i_{k\,|\,k-1}$ and $\hat{x}^i_k$, respectively.

Computation of the aforementioned quantities is standard:

1) At time k, agent i first calculates $\hat{x}^i_{k\,|\,k-1}$ and $P^i_{k\,|\,k-1}$ according to the following:

$$\hat{x}^i_{k\,|\,k-1} = A\hat{x}^i_{k-1}, \tag{3}$$

$$P^i_{k\,|\,k-1} = AP^i_{k-1}A' + Q, \tag{4}$$

where the recursion starts from $\hat{x}_0^i = 0$ and $P_0^i = \Pi_0$.

2) After the local measurement y_k^i is taken, agent i transmit y_k^i through the assigned channels.
3) After the communication, the agents first do the fusion of measurements. For agent i, define

$$\tilde{y}_k^i \triangleq \left(\gamma_{i1}\left(y_k^1\right)', \gamma_{i2}\left(y_k^2\right)', \ldots, \gamma_{iN}\left(y_k^N\right)'\right)',$$
$$\tilde{C}_i \triangleq \operatorname{diag}\{\gamma_{i1}C_1, \gamma_{i2}C_2, \ldots, \gamma_{iN}C_N\}.$$

Agent i computes $\hat{x}_k^i$ and P_k^i as follows:

$$\left(P_k^i\right)^{-1} = \left(P_{k\,|\,k-1}^i\right)^{-1} + \sum_{j=1}^N \gamma_{ij} C_i' R_i^{-1} C_i, \tag{5}$$
$$K_k^i = P_k^i \tilde{C}_i' R_i^{-1}, \tag{6}$$
$$\hat{x}_k^i = \hat{x}_{k\,|\,k-1}^i + K_k^i\left(\tilde{y}_k^i - \tilde{C}_i \hat{x}_{k\,|\,k-1}^i\right). \tag{7}$$

From standard Kalman filtering analysis, P_k^i converges to a steady-state value exponentially fast. Define

$$P_i \triangleq \lim_{k\to\infty} P_k^i. \tag{8}$$

The computation (3)–(7) shows that P_i depends on the underlying channel allocation matrix Γ.

D. Problem Statement

If the available channels are limited, it is critical to design an allocation matrix Γ such that a certain objective is met. In this paper we consider the following cost function $J(\Gamma)$ for a channel allocation matrix Γ:

$$J(\Gamma) = \sum_{i=1}^N \operatorname{Tr}(P_i), \tag{9}$$

i.e., we consider the average steady-state estimation errors of the N agents.

The main objective of this paper is to design an allocation matrix Γ that minimizes the cost function (9), i.e.,

Problem 2.1:

$$\min_{\Gamma} \quad J(\Gamma),$$
$$\text{s.t.} \quad \sum_{i=1}^N \sum_{\substack{j=1\\ j\neq i}}^N \gamma_{ij} \le d,$$

where d ($d \in \mathbb{N}, d \le N^2 - N$) is the number of available channels.

III. Homogeneous Agents

Equation (5) suggests the information fusion in terms of the estimated error covariance. To reveal this, for the ith agent, define the *sensing accuracy matrix* S_i as follows:

$$S_i \triangleq C_i' R_i^{-1} C_i, \tag{10}$$

which serves as a whole in (5) and suggests the contribution to estimation quality from the ith agent.

Furthermore, define the *assimilated sensing accuracy matrix* as

$$\tilde{S}_i \triangleq \sum_{j=1}^N \gamma_{ij} S_j, \tag{11}$$

which indicates the total contribution to estimation quality of the ith agent.

In this section, we consider the scenario that all the agents have identical sensing capabilities.

$$S_1 = S_2 = \cdots = S_N \triangleq \bar{S}. \tag{12}$$

A. Preliminary

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

$$h(X) \triangleq AXA' + Q, \tag{13}$$
$$g(X, S) \triangleq (X^{-1} + S)^{-1}. \tag{14}$$

To simplify the notation in the following discussion we also use the notation $g_S(X)$ to denote $g(X, S)$.

Define

$$\tilde{g}(X, S) \triangleq g_S h(X) \tag{15}$$

and

$$W(S) \triangleq \lim_{k\to\infty} (g_S h)^k(X) \tag{16}$$

for an arbitrary $X \ge 0$. $\tilde{g}$ is the recursive update equation of the *a postoriori* estimated error covariance in standard Kalman filter and $W(S)$ is its limitation, i.e., the steady value of estimated error covariance after long run. Note that $W(S)$ is the fixed point of the function $g_S h(X)$, i.e.,

$$W(S) = g_S h(W(S)).$$

One has the following result.

Lemma 3.1: $\tilde{g}(X, S)$ and $W(S)$ are decreasing and convex in $\mathbb{S}_{++}^n$ with respect to S.

Proof: See Appendix. ■

Remark 3.2: The recursive update equation for estimated error covariance $\tilde{g}(X, S)$ is usually viewed as the function of only one independent variable X. Here the influence of S is also taken into consideration and investigated.

In this section, further define

$$W_n \triangleq W(n\bar{S}). \tag{17}$$

From Lemma 3.1, we have the following results on W_n.

Corollary 3.3: Let $p, q, r, s \in \mathbb{N}$ with $p < r \le s < q$ and $p + q = r + s$, then

$$W_p + W_q \ge W_r + W_s.$$

Proof: See Appendix. ■

Proposition 3.4: For two matrices $X, Y > 0$, if $X \geq Y$ and $X \neq Y$, then

$$\mathrm{Tr}(X) > \mathrm{Tr}(Y).$$

Proof: See Appendix. ■

For agent i, from (4) and (5), one has

$$\begin{aligned}(P_k^i)^{-1} &= \left[h\left(P_{k-1}^i\right)\right]^{-1} + \sum_{j=1}^{N} \gamma_{ij} C_j' R_j^{-1} C_j \\ &= \left[h\left(P_{k-1}^i\right)\right]^{-1} + \tilde{S}_i,\end{aligned}$$

where $\tilde{S}_i$ is defined earlier in (11). Therefore,

$$P_k^i = \left(\left[h\left(P_{k-1}^i\right)\right]^{-1} + \tilde{S}_i\right)^{-1} = g_{\tilde{S}_i} h\left(P_{k-1}^i\right),$$

and further one has

$$P_i = W(\tilde{S}_i). \tag{18}$$

For a Γ, to count the number of sensors sending information to the ith agent, define the *connecting number* as

$$n_i = \sum_{j=1}^{N} \gamma_{ij}.$$

In this homogeneous scenario, from (12) one has

$$\tilde{S}_i = n_i \bar{S}.$$

As a result,

$$P_i = W(n_i \bar{S}) = W_{n_i}, \tag{19}$$

and

$$J(\Gamma) = \sum_{i=1}^{N} \mathrm{Tr}(P_i) = \mathrm{Tr}\left[\sum_{i=1}^{N} W_{n_i}\right]. \tag{20}$$

B. Channel Allocation

In this subsection, we present optimal solutions to Problem 2.1. Before proposing specific allocation solutions to it, first we give a necessary and sufficient condition for those feasible optimal allocations.

Theorem 3.5: A necessary and sufficient condition for an allocation Γ to be optimal to Problem 2.1 is that

$$n_i = \omega \text{ or } \omega + 1, \quad i = 1, 2, \ldots, N, \tag{21}$$

where $\omega = \lfloor \frac{d}{N} \rfloor + 1$.

Proof: For a feasible Γ, $\sum_{i=1}^{N} n_i \leq d + N$. If Γ is optimal, it is not difficult to show that

$$\sum_{i=1}^{N} n_i = d + N.$$

Intuitively, this means that all the feasible channels should be fully used.

To prove the necessity, assume that there is an integer p such that $n_p \neq \omega$ and $n_p \neq \omega + 1$. If $n_p \geq \omega + 2$, there must exist at least one q such that $n_q \leq \omega$. Since $n_p \geq n_q + 2$, agent p receives data from at least one agent r $(r \neq p, q)$ which does not send data to agent q. Then $\gamma_{pr} = 1$ and $\gamma_{qr} = 0$. Construct a new allocation $\Gamma' = [\gamma_{ij}']$ as

$$\gamma_{ij}' = \begin{cases} 0, & i = p, j = r, \\ 1, & i = q, j = r, \\ \gamma_{ij}, & \text{otherwise.} \end{cases}$$

It is easy to show that $n_p' = n_p - 1$, $n_q' = n_q + 1$, and $n_i' = n_i$ for $i \neq p, q$. From Corollary 3.3 and Proposition 3.4, we have

$$\mathrm{Tr}\left(W_{n_p} + W_{n_q}\right) > \mathrm{Tr}\left(W_{n_p'} + W_{n_q'}\right).$$

Hence,

$$\begin{aligned}J(\Gamma) - J(\Gamma') &= \mathrm{Tr}\left[\sum_{i=1}^{N} W_{n_i}\right] - \mathrm{Tr}\left[\sum_{i=1}^{N} W_{n_i'}\right] \\ &= \mathrm{Tr}\left[W_{n_p} + W_{n_q} - \left(W_{n_p'} + W_{n_q'}\right)\right] \\ &> 0,\end{aligned}$$

which contradicts the optimality of Γ.

Now let us prove the sufficiency. Construct an optimal Γ under which $n_i = \omega$ or $\omega + 1$ for all i. Denote the number of agents receiving ω and $\omega + 1$ measurements as N_ω and $N_{\omega+1}$, respectively. Then from

$$\begin{aligned}N_\omega + N_{\omega+1} &= N, \\ N_\omega \omega + N_{\omega+1}(\omega + 1) &= d + N,\end{aligned}$$

N_ω and $N_{\omega+1}$ can be determined. Therefore,

$$J(\Gamma) = N_\omega \mathrm{Tr}\left(W_{N_\omega}\right) + N_{\omega+1} \mathrm{Tr}\left(W_{N_{\omega+1}}\right).$$

On the other hand, for any other allocation Γ', if

$$n_i' = \omega \text{ or } \omega + 1, \quad \forall i,$$

then it has the same cost as Γ, which indicates that Γ' is also optimal. ■

Theorem 3.5 reveals the fact that the number of the channels connected to each agent should be *as uniform as possible* under an optimal allocation.

In the following part we present specific optimal allocation solutions. Though giving the allocation matrix is sufficient, to better understand the allocation topology among the agents, we also introduce the *channel allocation table* (One can refer to Table I). The row headings represent the agents receiving data and the column headings indicate the agents providing data. Each block of the numerical area indicates a feasible channel. If one block, whose row and column heading titles are 'Agent i' and 'Agent j' respectively, is assigned with a tick or a number, it means that agent i receives data from agent j.

The allocation strategy satisfying the sufficient and necessary condition given in Theorem 3.5 is not unique. We propose one optimal allocation strategy $\Gamma^\star$ interpreted by Table I. The channels are determined by the order shown in the table, i.e., this proposed optimal allocation consists of the channels marked with number 1 to d. We use an example to illustrate the proposed optimal allocation strategy.

TABLE I
A_i Represents Agent i

	A_1	A_2	$\cdots$	A_N
A_1	—	N	$\cdots$	$(N-1)^2+1$
A_2	1	—	$\cdots$	$(N-1)^2+2$
A_3	2	$N+1$	$\cdots$	$(N-1)^2+3$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
A_{N-1}	$N-2$	$2N-3$	$\cdots$	$(N-1)N$
A_N	$N-1$	$2N-2$	$\cdots$	—

TABLE II
Optimal Channel allocation

	Agent 1	Agent 2	Agent 3	Agent 4	Agent 5	n_i
Agent 1	—	5	9			3
Agent 2	1	—	10			3
Agent 3	2	6	—			3
Agent 4	3	7	11	—		4
Agent 5	4	8			—	3

Example 3.6: Consider a system with $N = 5$ and $d = 11$. The optimal allocation strategy proposed above in this case is illustrated in Table II. The number gives the order to set channels. The last column counts n_i of each agents, showing that the largest differences among them is one, which has been elaborated in Theorem 3.5. Therefore, the optimal allocation matrix $\Gamma^\star$ is

$$\Gamma^\star = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

IV. Heterogeneous Agents

In this section we discuss the agents with heterogeneous sensing capabilities. Since the sensing accuracy matrices are not comparable for general multi-dimensional process, here we only consider the first-order system.

We use lower case letters to denote the parameters and variables, i.e., the notations A, C_i, Q, R_i, P_i are replaced by a, c_i, q, r_i, p_i, respectively.

In particular, the sensing accuracy matrix S_i now becomes s_i, which is renamed as *the sensing accuracy index*:

$$s_i = \frac{c_i^2}{r_i}, \quad i = 1, 2, \ldots, N. \tag{22}$$

Without loss of generality, assume

$$s_1 \geq s_2 \geq \cdots \geq s_N. \tag{23}$$

Also $\tilde{S}_i$ is renamed as the *assimilated sensing accuracy index*:

$$\tilde{s}_i = \sum_{j=1}^{N} \gamma_{ij} s_j. \tag{24}$$

The cost function (9) becomes

$$J(\Gamma) = \sum_{i=1}^{N} p_i. \tag{25}$$

TABLE III
A_i Represents Agent i

	A_1	A_2	$\cdots$	A_{N-1}	A_N
A_1	—	—	$\cdots$	—	—
A_2	$N-1$	—	$\cdots$	—	—
A_3	$N-2$	$2N-2$	$\cdots$	—	—
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
A_{N-1}	2	$N+2$	$\cdots$	—	—
A_N	1	$N+1$	$\cdots$	$(N-1)^2$	—
A_1	—	N	$\cdots$	$(N-1)^2-1$	$N(N-1)$
A_2	—	—	$\cdots$	$(N-1)^2-2$	$N(N-1)-1$
A_3	—	—	$\cdots$	$(N-1)^2-3$	$N(N-1)-2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
A_{N-1}	—	—	$\cdots$	—	$(N-1)^2+1$
A_N	—	—	$\cdots$	—	—

A. Preliminary

We re-define the scalar version of functions h and g defined in (13) and (14) as

$$h(x) \triangleq a^2 x + q, \tag{26}$$
$$g_s(x) \triangleq (x^{-1} + s)^{-1}, \tag{27}$$

where $x, s \in \mathbb{R}$. Denote $w(s)$ as

$$w(s) = \lim_{k\to\infty} (g_s h)^k (x),$$

hence $w(s)$ is the solution to the equation

$$x = g_s(h(x)).$$

Similarly as the homogeneous scenario, one has

$$p_i = w(\tilde{s}_i). \tag{28}$$

In this first-order setting, Lemma 3.1 can be strengthened as the following lemma.

Lemma 4.1: $w(s)$ is monotonically decreasing with s and is strictly convex.

Proof: See Appendix. ■

Moreover, one has the following property.

Lemma 4.2: For any $z_1, z_2 \in \mathbb{R}$, $z_1 < z_2$, and $z_3, z_4 \in (z_1, z_2)$, if $z_3 - z_1 \geq z_2 - z_4$, then

$$w(z_1) + w(z_2) > w(z_3) + w(z_4).$$

Proof: See Appendix. ■

B. Channel Allocation

We propose an optimal allocation strategy $\Gamma^\star = [\gamma_{ij}^\star]$, which is obtained by a greedy algorithm. The procedure to allocate the channels is shown in Table III. To facilitate the demonstration, the form of the table is slightly modified. The number in the table is the order to allocate the channel resources. The following example helps to illustrate it.

Example 4.3: Consider a system with $N = 5$, $d = 11$. The optimal allocation procedure in terms of the order for setting

TABLE IV
5 AGENTS WITH 11 CHANNELS OFFERED. A_i REPRESENTS AGENT i

	A_1	A_2	A_3	A_4	A_5	$\tilde{s}_i$
A_1	—	5	10			$s_1+s_2+s_3$
A_2	4	—	9			$s_1+s_2+s_3$
A_3	3	8	—			$s_1+s_2+s_3$
A_4	2	7		—		$s_1+s_2+s_4$
A_5	1	6	11		—	$s_1+s_2+s_3+s_5$

TABLE V
A DIFFERENT OPTIMAL CHANNEL ALLOCATION

	A_1	A_2	A_3	A_4	A_5
A_1	—	5	$\star$		
A_2	4	—	$\star$		
A_3	3	8	—		
A_4	2	7		—	
A_5	1	6			—

TABLE VI
STAGE 1

	A_1	A_2	A_3	$\cdots$	A_i	$\cdots$
A_i	$\surd$		$\surd$		—	

$\Downarrow$

	A_1	A_2	A_3	$\cdots$	A_i	$\cdots$
A_i	$\surd$	$\surd$			—	

channels is given in Table IV. The optimal allocation matrix is

$$\Gamma^\star = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

Note that homogeneous agents can be viewed as a special scenario of heterogeneous ones. Therefore, Table III provides an alternative schedule for homogeneous case in addition to Table I.

Remark 4.4: Some of the allocating orders given in Table III are not unique. In Table V, the channels marked by stars are equivalent as agent 1 and 2 have full information of each other. The cost is exactly the same when agent 3 sends its data to either of the two, i.e., the 9th channel can be set in either of the star places.

Theorem 4.5: The schedule $\Gamma^\star$, which allocates the channels according to the procedure shown in Table III, is optimal.

Proof: Take an arbitrary $\Gamma = [\gamma_{ij}]$, under which the steady state error variance of agent i is denoted as p_i and the assimilated sensing accuracy index as $\tilde{s}_i$. To prove the theorem, we construct a series of intermediate matrices $\Gamma_1, \Gamma_2, \ldots, \Gamma_m$ by two types of operations on Γ, and show that

$$J(\Gamma) \geq J(\Gamma_1) \geq J(\Gamma_2) \geq \cdots \geq J(\Gamma_m) \geq J(\Gamma^\star).$$

The proof is divided into three stages as follows.

1) Operation 1 is to rearrange one agent, say, agent i, to receive data from the most accurate n_i neighboring agents if feasible, which is illustrated in Table VI.
After applying operation 1 to every such agent, a new allocation $\Gamma_1 = [\gamma_{ij}^1]$ is constructed. Under Γ_1, the steady state error covariance of the agent i is denoted as $p_{1,i}$

TABLE VII
STAGE 2

	A_1	A_2	A_3	$\cdots$	A_{i_1}	$\cdots$	A_{i_2}	$\cdots$
A_{i_1}	$\surd$	$\surd$	$\surd$		—			
A_{i_2}	$\surd$						—	

$\Downarrow$

	A_1	A_2	A_3	$\cdots$	A_i	$\cdots$	A_{i_2}	$\cdots$
A_{i_1}	$\surd$	$\surd$			—			
A_{i_2}	$\surd$	$\surd$					—	

and the assimilated sensing accuracy variable as $\tilde{s}_i^1$. Since $s_1 \geq s_2 \geq \cdots \geq s_N$, then

$$\tilde{s}_i^1 = \sum_{j=1}^{N} \gamma_{ij}^1 s_j \geq \sum_{j=1}^{N} \gamma_{ij} s_j = \tilde{s}_i.$$

From (28) and that $w(s)$ is decreasing with s, one concludes

$$p_{1,i} \leq p_i,$$

which leads to $J(\Gamma) \geq J(\Gamma_1)$.

2) In Γ_1, define $n_i^1 = \sum_{j=1}^{N} \gamma_{ij}^1$. If there are two row indices i_1 and i_2, such that $n_{i_1} - n_{i_2} \geq 2$, then apply operation 2: first remove the channel providing the least accurate data to agent i_1, and then add a channel sending the available most accurate information to agent i_2. The operation is interpreted in Table VII.
To show the improvement operation 2 brings, we only consider the case $i_1 > i_2$ and $n_{i_1} > i_1, n_{i_2} \geq i_2$ for simplicity. According to the manner Γ_1 is constructed, for row i_1 and i_2, one knows

$$\gamma_{i_1 j}^1 = \begin{cases} 1, & j = 1, 2, \ldots, n_{i_1}, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\gamma_{i_2 j}^1 = \begin{cases} 1, & j = 1, 2, \ldots, n_{i_2}, \\ 0, & \text{otherwise}. \end{cases}$$

Then

$$\tilde{s}_{i_1}^1 = \sum_{j=1}^{n_{i_1}} s_j, \quad \tilde{s}_{i_2}^1 = \sum_{j=1}^{n_{i_2}} s_j.$$

By operation 2, $\Gamma_2 = [\gamma_{ij}^2]$ is constructed as

$$\gamma_{ij}^2 = \begin{cases} 0, & i = i_1, j = n_{i_1}, \\ 1, & i = i_2, j = n_{i_2} + 1, \\ \gamma_{ij}^1, & \text{otherwise}. \end{cases}$$

Under Γ_2, the steady state error covariance of the agent i is denoted as $p_{2,i}$ and the assimilated sensing accuracy variable as $\tilde{s}_i^2$. Then

$$\tilde{s}_{i_1}^2 = \sum_{j=1}^{n_{i_1}-1} s_j, \quad \tilde{s}_{i_2}^2 = \sum_{j=1}^{n_{i_2}+1} s_j.$$

Note that

$$\tilde{s}_{i_2}^1 < \tilde{s}_{i_2}^2 < \tilde{s}_{i_1}^1, \quad \tilde{s}_{i_2}^1 < \tilde{s}_{i_1}^2 < \tilde{s}_{i_1}^1,$$

and

$$\tilde{s}_{i_1}^1 - \tilde{s}_{i_1}^2 = s_{n_{i_1}},$$
$$\tilde{s}_{i_2}^2 - \tilde{s}_{i_2}^1 = s_{n_{i_2}+1}.$$

TABLE VIII
STAGE 3

	A_1	A_2	A_3	$\cdots$	$A_{i_{2l-1}}$	$\cdots$	$A_{i_{2l}}$	$\cdots$
$A_{i_{2l-1}}$	$\surd$	$\surd$			—			
$A_{i_{2l}}$	$\surd$						—	

$\Downarrow$

	A_1	A_2	A_3	$\cdots$	$A_{i_{2l-1}}$	$\cdots$	$A_{i_{2l}}$	$\cdots$
$A_{i_{2l-1}}$	$\surd$				—			
$A_{i_{2l}}$	$\surd$	$\surd$					—	

Since $n_{i_1} - n_{i_2} \geq 2$, $n_{i_1} > n_{i_2} + 1$. Therefore $s_{n_{i_1}} < s_{n_{i_2}+1}$, i.e.,

$$\tilde{s}_{i_2}^2 - \tilde{s}_{i_2}^1 > \tilde{s}_{i_1}^1 - \tilde{s}_{i_1}^2.$$

Γ_2 satisfies

$$\tilde{s}_{i_2}^1 < \tilde{s}_{i_2}^2 < \tilde{s}_{i_1}^1, \quad \tilde{s}_{i_2}^1 < \tilde{s}_{i_1}^2 < \tilde{s}_{i_1}^1,$$

and

$$\tilde{s}_{i_2}^2 - \tilde{s}_{i_2}^1 \geq \tilde{s}_{i_1}^1 - \tilde{s}_{i_1}^2.$$

Applying Lemma 4.2 one has

$$w\left(\tilde{s}_{i_1}^1\right) + w\left(\tilde{s}_{i_2}^1\right) > w\left(\tilde{s}_{i_1}^2\right) + w\left(\tilde{s}_{i_2}^2\right).$$

Therefore,

$$\begin{aligned} J(\Gamma_1) - J(\Gamma_2) &= \sum_{i=1}^{N} p_{1,i} - \sum_{i=1}^{N} p_{2,i} \\ &= p_{1,i_1} + p_{1,i_2} - p_{2,i_1} - p_{2,i_2} \\ &= w\left(\tilde{s}_{i_1}^1\right) + w\left(\tilde{s}_{i_2}^1\right) - w\left(\tilde{s}_{i_1}^2\right) - w\left(\tilde{s}_{i_2}^2\right) \\ &> 0, \end{aligned}$$

which proves Γ_2 is better than Γ_1.

After obtaining Γ_2, we continue to check whether there are another two row indices i_3 and i_4 with $n_{i_3} - n_{i_4} \geq 2$. If so we continue to apply operation 2 and construct Γ_3. By doing this iteratively, we obtain a Γ_l, and for any i_{2l-1} and i_{2l}, $|n_{i_{2l-1}} - n_{i_{2l}}| \leq 1$.

3) In Γ_l, if there are two row indices i_{2l-1} and i_{2l}, $i_{2l-1} < i_{2l}$, $n_{i_{2l}} < i_{2l}$, and $n_{i_{2l-1}} - n_{i_{2l}} = 1$, then apply operation 2 and construct Γ_{l+1}, illustrated in Table VIII. The verification of $J(\Gamma_l) > J(\Gamma_{l+1})$ is similar to that in stage 2.

Repeat doing this step and one obtains $\Gamma^\star$, which is optimal.

■

The procedure of the proof is demonstrated by a simple example in Table IX.

Remark 4.6: We can extend the first-order case to a particular vector case, where

$$S_i = t_i S, \tag{29}$$

in which $S > 0$ and $t_1 \geq t_2 \geq \cdots \geq t_N$. Then the optimal allocation is the same as the first-order case.

TABLE IX
FROM AN ARBITRARY ALLOCATION Γ TO THE OPTIMAL ONE $\Gamma^\star$

Γ:

	Agent 1	Agent 2	Agent 3	Agent 4
Agent 1	—		$\surd$	
Agent 2		—		
Agent 3	$\surd$	$\surd$	—	
Agent 4	$\surd$		$\surd$	—

$\Downarrow$ *operation* 1

Γ_1:

	Agent 1	Agent 2	Agent 3	Agent 4
Agent 1	—	$\surd$		
Agent 2		—		
Agent 3	$\surd$	$\surd$	—	
Agent 4	$\surd$	$\surd$		—

$\Downarrow$ *operation* 2

Γ_2:

	Agent 1	Agent 2	Agent 3	Agent 4
Agent 1	—	$\surd$		
Agent 2	$\surd$	—		
Agent 3	$\surd$	$\surd$	—	
Agent 4	$\surd$			—

$\Downarrow$ *operation* 2

$\Gamma^\star$:

	Agent 1	Agent 2	Agent 3	Agent 4
Agent 1	—	$\surd$		
Agent 2	$\surd$	—		
Agent 3	$\surd$		—	
Agent 4	$\surd$	$\surd$		—

V. SIMULATION

Example 5.1: Consider the system with homogeneous agents. Let $N = 10$, then the maximum channel number is $N(N-1) = 90$. Let $A = \begin{bmatrix} 3 & 1 \\ 0.5 & 2 \end{bmatrix}$, $Q = \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}$, and $C_i = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, $R_i = \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}$, $i = 1, 2, \ldots, N$. Given d available channels, let d vary from 1 to 90. We compare

1) the cost of the optimal schedule presented in Table I, and
2) the expected cost of a schedule whose channels are randomly picked, computing by the Monte Carlo method.

Fig. 2 shows this comparison. From the figure one can see that the optimal schedule has lower cost than that of the random one. Note that when $d = 1$ and $d = 90$, which is easy to understand. Moreover, the curve of the optimal schedule is segmented by the multiples of N, and within each segment the curve is linear, due to the homogeneity of the agents as well as the uniform property of channel distributing.

Example 5.2: Consider a system with heterogeneous agents. Still let $N = 10$, $A = \begin{bmatrix} 3 & 1 \\ 0.5 & 2 \end{bmatrix}$, $Q = \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}$. Also $C_i = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, $R_i = (N+1-i)\begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}$, $i = 1, 2, \ldots, N$. Similar as what is done in Example 5.1, we compare the cost of the optimal schedule given in Table III and the expected cost of a random one. Fig. 3 shows the result. Note that when $d = 1$, the cost of the two schedules is different. That is because the optimal schedule under this scenario is unique, due to the heterogeneity of the agents.

VI. CONCLUSION

In this paper, we consider communication channel allocation for a group of agents which measure the state of an underlying process. Optimal channel allocation strategies are provided for

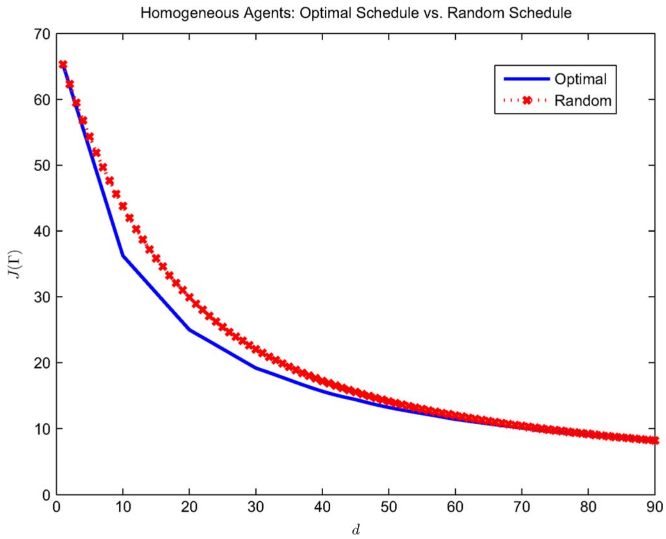


Fig. 2. The concrete curve represents the optimal schedule and the dotted one indicates random schedules.

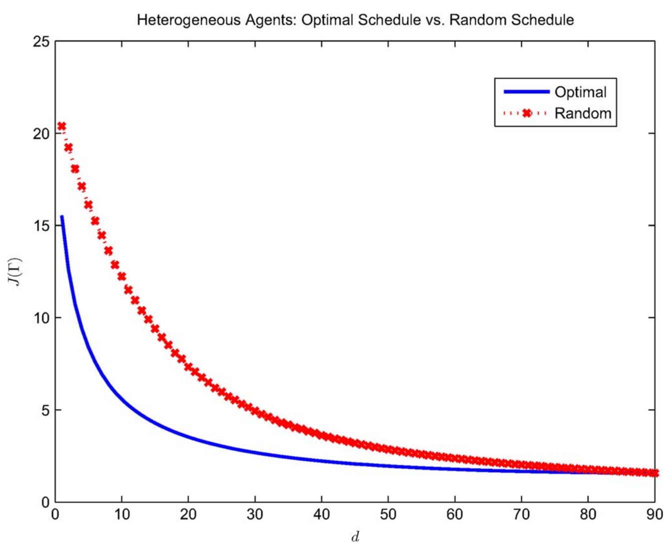


Fig. 3. The concrete curve represents the optimal schedule and the dotted one indicates random schedules.

agents with homogeneous and heterogeneous sensing capabilities respectively such that the average estimation error of these agents is minimized.

There are many interesting future research directions along the line of this work. In particular, we will consider the following aspects.

1) In this paper, the communication channels are set to be reliable ones. Unreliable channels (e.g., which introduce data packet drops or random delays) will be considered.
2) This paper does not consider the cost of using a particular communication channel. It is reasonable to consider non-homogenous channel usage costs. For example, two sensors that are closer typically spend less communication energy than if they are far away from each other.
3) This paper also does not consider sensing cost at each sensor. Variable sensing costs will result in a drastically different channel allocation. For example, those sensors with smaller sensing cost will be allocated with more communication channels, and vice verse.

Appendix

Proposition 7.1: The matrix function $f(X) = X^{-1}$ is convex in $\mathbb{S}^n_{++}$.

Proof: We need to verify that, given two matrices $X > 0$ and $Y > 0$, and two real numbers $\alpha \geq 0$ and $\beta \geq 0$ satisfying $\alpha + \beta = 1$, the following inequality holds:

$$(\alpha X + \beta Y)^{-1} \leq \alpha X^{-1} + \beta Y^{-1}.$$

Since $X, Y > 0$, then $X^{-1} > 0, Y^{-1} > 0$.

Consider the matrix $\Phi = \begin{bmatrix} X & I_n \\ I_n & X^{-1} \end{bmatrix}$, where I_n is the $n \times n$ identity matrix. Since $X > 0$, and the Schur complement of the block X of Φ is $X - I_n X I_n = 0$, $\Phi \geq 0$. Similarly $\Psi = \begin{bmatrix} Y & I_n \\ I_n & Y^{-1} \end{bmatrix} \geq 0$. Thus the matrix $\Theta = \alpha\Phi + \beta\Psi \geq 0$. Now

$$\begin{aligned} \Theta &= \begin{bmatrix} \alpha X + \beta Y & \alpha I_n + \beta I_n \\ \alpha I_n + \beta I_n & \alpha X^{-1} + \beta Y^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \alpha X + \beta Y & I_n \\ I_n & \alpha X^{-1} + \beta Y^{-1} \end{bmatrix} \geq 0, \end{aligned}$$

and $\alpha X^{-1} + \beta Y^{-1} > 0$, then the Schur complement of $\alpha X^{-1} + \beta Y^{-1}$ of Θ is given by

$$\alpha X^{-1} + \beta Y^{-1} - I_n(\alpha X + \beta Y)^{-1} I_n \geq 0,$$

which leads to

$$(\alpha X + \beta Y)^{-1} \leq \alpha X^{-1} + \beta Y^{-1}.$$

■

Lemma 7.2: (Matrix Inverse Lemma): Let $X > 0$. If $X = B^{-1} + CD^{-1}C'$, then

$$X^{-1} = B - BC(D + C'BC)^{-1}C'B$$

Proposition 7.3: Define matrix function

$$\psi(X) \triangleq (h(X^{-1}))^{-1}. \tag{30}$$

It is concave in $\mathbb{S}^n_{++}$ and is monotonically increasing with X.

Proof: The monotonicity is easy to see. We only prove the concavity.

If Q is invertible, then by Lemma 7.2,

$$\begin{aligned} \psi(X) &= (AX^{-1}A' + Q)^{-1} \\ &= Q^{-1} - Q^{-1}A(A'Q^{-1}A + X)^{-1}A'Q^{-1}. \end{aligned}$$

For two arbitrary matrices $X_1, X_2 > 0$, and two real numbers $\alpha, \beta > 0$, $\alpha + \beta = 1$,

$$\begin{aligned} &\alpha\psi(X_1) + \beta\psi(X_2) \\ &\quad = \alpha(Q^{-1} - Q^{-1}A(A'Q^{-1}A + X_1)^{-1}A'Q^{-1}) \\ &\qquad + \beta(Q^{-1} - Q^{-1}A(A'Q^{-1}A + X_2)^{-1}A'Q^{-1}) \\ &\quad = Q^{-1} - Q^{-1}A(\alpha(A'Q^{-1}A + X_1)^{-1} \\ &\qquad + \beta(A'Q^{-1}A + X_2)^{-1})A'Q^{-1} \\ &\quad \leq Q^{-1} - Q^{-1}A(\alpha(A'Q^{-1}A + X_1) \\ &\qquad + \beta(A'Q^{-1}A + X_2))^{-1}A'Q^{-1} \\ &\quad = Q^{-1} - Q^{-1}A(A'Q^{-1}A + \alpha X_1 + \beta X_2)^{-1}A'Q^{-1} \\ &\quad = \psi(\alpha X_1 + \beta X_2), \end{aligned}$$

where the inequality follows from the convexity of X^{-1} verified in Proposition 7.1. Hence, $\psi(X)$ is concave.

When Q is not invertible, for an arbitrary $\Delta > 0$, define

$$\tilde{\psi}(X) \triangleq (AX^{-1}A' + Q + \Delta)^{-1}.$$

Since $Q + \Delta$ is invertible, $\tilde{\psi}(X)$ is concave in X. Note that

$$\begin{aligned}\psi(X) &= (AX^{-1}A' + Q + \Delta - \Delta)^{-1}\\ &= ((-\Delta) + (\tilde{\psi}(X))^{-1})^{-1}\\ &= -\Delta^{-1} - \Delta^{-1}(\tilde{\psi}(X) - \Delta^{-1})^{-1}\Delta^{-1}.\end{aligned}$$

Further define

$$\Omega(X) = -\Delta^{-1} - \Delta^{-1}(X - \Delta^{-1})^{-1}\Delta^{-1}.$$

It can be easily verified that the function $\Omega(X)$ is concave by a similar way for $\psi(X)$ with invertible Q, and is monotonically increasing with X. Now

$$\psi(X) = \Omega(\tilde{\psi}(X)).$$

For two arbitrary matrices $X_1, X_2 > 0$, and two positive numbers α, β, $\alpha + \beta = 1$,

$$\begin{aligned}\psi(\alpha X_1 + \beta X_2) &= \Omega(\tilde{\psi}(\alpha X_1 + \beta X_2))\\ &\geq \Omega(\alpha\tilde{\psi}(X_1) + \beta\tilde{\psi}(X_2))\\ &\geq \alpha\Omega(\tilde{\psi}(X_1)) + \beta\Omega(\tilde{\psi}(X_2))\\ &= \alpha\psi(X_1) + \beta\psi(X_2),\end{aligned}$$

where the first inequality holds because of the concavity of $\tilde{\psi}$ and monotonic increasing of $\Omega(X)$, and the second inequality is the result of the concavity of $\Omega(X)$. Therefore, the concavity of $\psi(X)$ is proved. ■

We are now ready to prove Lemma 3.1.

Proof to Lemma 3.1: Define

$$W_k(S) \triangleq (g_S h)^k(X).$$

Note that

$$\tilde{g}(X, S) = W_1(S)$$

and

$$\lim_{k\to\infty} W_k(S) = W(S).$$

To prove the monotonicity, one needs to show

$$W_k(S_1) \leq W_k(S_2)$$

for $S_1 \geq S_2 > 0$ and all $k > 0$. One can verify this by mathematical induction. When $k = 1$,

$$\begin{aligned}g_{S_1}h(X) &= ([h(X)]^{-1} + S_1)^{-1} \leq ([h(X)]^{-1} + S_2)^{-1}\\ &= g_{S_2}h(X).\end{aligned}$$

When $k = n - 1$, assume $(g_{S_1}h)^{n-1}(X) \geq (g_{S_2}h)^{n-1}(X)$ holds. When $k = n$,

$$\begin{aligned}(g_{S_1}h)^n(X) &= ([(g_{S_1}h)^{n-1}(X)]^{-1} + S_1)^{-1}\\ &\geq ([(g_{S_2}h)^{n-1}(X)]^{-1} + S_2)^{-1}\\ &= (g_{S_2}h)^n(X).\end{aligned}$$

Therefore,

$$W_k(S_1) \leq W_k(S_2).$$

By letting $k \to \infty$ one has

$$W(S_1) \leq W(S_2).$$

The monotonicity is proved of both $\tilde{g}(X, S)$ and $W(S)$ in S.

For arbitrary two real numbers $\alpha, \beta > 0$, $\alpha + \beta = 1$, and two matrices $S_1, S_2 > 0$, to show the convexity, it is needed to verify

$$\alpha W_k(S_1) + \beta W_k(S_2) \geq W_k(\alpha S_1 + \beta S_2)$$

for all $k > 0$. First claim

$$\alpha[W_k(S_1)]^{-1} + \beta[W_k(S_2)]^{-1} \leq [W_k(\alpha S_1 + \beta S_2)]^{-1}. \quad (31)$$

We still verify this by mathematical induction. When $k = 1$,

$$\begin{aligned}&\alpha[W_1(S_1)]^{-1} + \beta[W_1(S_2)]^{-1}\\ &\quad= \alpha[h(X)]^{-1} + \alpha S_1 + \beta[h(X)]^{-1} + \beta S_2\\ &\quad= [h(X)]^{-1} + \alpha S_1 + \beta S_2\\ &\quad= [W_1(\alpha S_1 + \beta S_2)]^{-1}.\end{aligned}$$

Suppose the argument holds when $k = n - 1$. When $k = n$,

$$\begin{aligned}&\alpha[W_n(S_1)]^{-1} + \beta[W_n(S_2)]^{-1}\\ &\quad= \alpha[h(W_{n-1}(S_1))]^{-1} + \alpha S_1\\ &\qquad+ \beta[h(W_{n-1}(S_2))]^{-1} + \beta S_2\\ &\quad= \alpha\psi([W_{n-1}(S_1)]^{-1}) + \beta\psi([W_{n-1}(S_2)]^{-1})\\ &\qquad+ \alpha S_1 + \beta S_2\\ &\quad\leq \psi(\alpha[W_{n-1}(S_1)]^{-1} + \beta[W_{n-1}(S_2)]^{-1}) + \alpha S_1 + \beta S_2\\ &\quad\leq \psi([W_{n-1}(\alpha S_1 + \beta S_2)]^{-1}) + \alpha S_1 + \beta S_2\\ &\quad= [h(W_{n-1}(\alpha S_1 + \beta S_2))]^{-1} + \alpha S_1 + \beta S_2\\ &\quad= [W_n(\alpha S_1 + \beta S_2)]^{-1},\end{aligned}$$

where the first inequality holds because of the concavity of $\psi(X)$ and the second is due to the assumption of $k = n - l$ and the monotonically increasing property of $\psi(X)$. Therefore, from the mathematical induction, for all $k > 0$,

$$\alpha[W_k(S_1)]^{-1} + \beta[W_k(S_2)]^{-1} \leq [W_k(\alpha S_1 + \beta S_2)]^{-1}.$$

Furthermore,

$$\begin{aligned}&\alpha W_k(S_1) + \beta W_k(S_2)\\ &\quad= \alpha[[W_k(S_1)]^{-1}]^{-1} + \beta[[W_k(S_2)]^{-1}]^{-1}\\ &\quad\geq [\alpha[W_k(S_1)]^{-1} + \beta[W_k(S_2)]^{-1}]^{-1}\\ &\quad\geq [[W_k(\alpha S_1 + \beta S_2)]^{-1}]^{-1}\\ &\quad= W_k(\alpha S_1 + \beta S_2),\end{aligned}$$

where the inequalities hold due to the monotonicity and convexity of X^{-1}. Let k approaches infinity, one has

$$\alpha W(S_1) + \beta W(S_2) \geq W(\alpha S_1 + \beta S_2).$$

By now it has both been proved that $\tilde{g}(X, S)$ and $W(S)$ are convex with respect to S in $\mathbb{S}_{++}$. ■

Proof to Corollary 3.3: From Lemma 3.1, $W(S)$ is convex. Therefore, for arbitrary two real numbers $\alpha, \beta > 0$, $\alpha + \beta = 1$, and two matrices $S_1, S_2 > 0$, one has

$$\alpha W(S_1) + \beta W(S_2) \geq W(\alpha S_1 + \beta S_2).$$

Let $S_1 = m\bar{S}$, $S_2 = (m+2)\bar{S}$, $m \in \mathbb{N}$, $\alpha = \beta = \frac{1}{2}$, it becomes

$$\frac{1}{2}W(m\bar{S}) + \frac{1}{2}W((m+2)\bar{S}) \geq W\left(\frac{1}{2}m\bar{S} + \frac{1}{2}(m+2)\bar{S}\right),$$

i.e.,

$$W_m + W_{m+2} \geq 2W_{m+1}.$$

Rearrange the above inequality, yielding

$$W_m - W_{m+1} \geq W_{m+1} - W_{m+2}.$$

Therefore,

$$W_p - W_{p+1} \geq W_{p+1} - W_{p+2} \geq \cdots \geq W_{q-1} - W_q,$$

from which one has

$$W_p + W_q \geq W_{p+1} + W_{q-1}.$$

As a result,

$$\begin{aligned} W_p + W_q &\geq W_{p+1} + W_{q-1} \\ &\geq W_{p+2} + W_{q-2} \\ &\geq \cdots \\ &\geq W_r + W_s. \end{aligned}$$

■

Proof to Proposition 3.4: Assume $X = [x_{ij}]$ and $Y = [y_{ij}]$. Define $D = [d_{ij}]$ as $D \triangleq X - Y$.

Since $X \geq Y$ and $X \neq Y$, $D \geq 0$ and $D \neq 0$. Note that D is diagonalizable under a similarity transformation since it is symmetric, i.e.,

$$D = P\Lambda P^{-1},$$

where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ and $\lambda_1, \lambda_2, \ldots, \lambda_N$ are the eigenvalues of D. Since $D \geq 0$, $\Lambda \geq 0$ which implies that $\lambda_i \geq 0, \forall i$. Assume $\text{Tr}(D) = 0$, then $\text{Tr}(\Lambda) = \text{Tr}(D) = 0$. Therefore, $\lambda_i = 0, \forall i$ and $D = P\Lambda P^{-1} = 0$, which contradicts that $D \neq 0$. Therefore, $\text{Tr}(D) > 0$. ■

Proof to Lemma 4.1: The monotonicity and convexity have been proved in Lemma 3.1. To show $w(s)$ is strictly convex, one can further calculate $w(s)$:

$$\begin{aligned} w(s) = \frac{1}{2a^2}&\sqrt{(a^2-1)^2\left(\frac{1}{s}\right)^2 + 2(a^2+1)q\left(\frac{1}{s}\right) + q^2} \\ &+\frac{1}{2a^2}\left((a^2-1)\left(\frac{1}{s}\right) - q\right). \end{aligned}$$

Direct calculation shows that $\frac{dw}{ds} < 0$. Further calculation reveals that $\frac{d^2w}{ds^2} > 0$. Since $\frac{dw}{ds} \neq 0$, $w(s)$ is strictly convex. ■

We state one property for the strictly convex real function as follows.

Proposition 7.4: Given a strictly convex function $\varphi(z)$, for any $z_1, z_2 \in \mathbb{R}$, $z_1 < z_2$, and $z \in (z_1, z_2)$, there is

$$\frac{\varphi(z) - \varphi(z_1)}{z - z_1} < \frac{\varphi(z_2) - \varphi(z_1)}{z_2 - z_1} < \frac{\varphi(z_2) - \varphi(z)}{z_2 - z}. \tag{32}$$

Proposition 7.4 has a corollary as follows.

Corollary 7.5: Given a strictly convex function $\varphi(z)$, for any $z_1, z_2 \in \mathbb{R}$, $z_1 < z_2$, and $z_3, z_4 \in (z_1, z_2)$, there is

$$\frac{\varphi(z_3) - \varphi(z_1)}{z_3 - z_1} < \frac{\varphi(z_2) - \varphi(z_4)}{z_2 - z_4}. \tag{33}$$

Proof: Given $z \in (z_1, z_2)$, from Proposition 7.4, (32) holds.

To prove (33), consider the following three cases:

1) $z_3 = z_4$.
   Let $z_3 = z_4 = z$. Then (33) becomes

$$\frac{\varphi(z) - \varphi(z_1)}{z - z_1} < \frac{\varphi(z_2) - \varphi(z)}{z_2 - z},$$

   which is directly verified by (32).

2) $z_3 < z_4$.
   Since $z_1 < z_3 < z_4$, (32) shows

$$\frac{\varphi(z_3) - \varphi(z_1)}{z_3 - z_1} < \frac{\varphi(z_4) - \varphi(z_3)}{z_4 - z_3}.$$

   Since $z_3 < z_4 < z_2$, by applying (32),

$$\frac{\varphi(z_4) - \varphi(z_3)}{z_4 - z_3} < \frac{\varphi(z_2) - \varphi(z_4)}{z_2 - z_4}.$$

   Combine the above two inequalities and obtain

$$\frac{\varphi(z_3) - \varphi(z_1)}{z_3 - z_1} < \frac{\varphi(z_4) - \varphi(z_3)}{z_4 - z_3} < \frac{\varphi(z_2) - \varphi(z_4)}{z_2 - z_4},$$

   which verifies (33).

3) $z_3 > z_4$.
   Since $z_1 < z_4 < z_3$, from (32) there is

$$\frac{\varphi(z_3) - \varphi(z_1)}{z_3 - z_1} < \frac{\varphi(z_3) - \varphi(z_4)}{z_3 - z_4}.$$

   Since $z_4 < z_3 < z_2$, (32) verifies

$$\frac{\varphi(z_3) - \varphi(z_4)}{z_3 - z_4} < \frac{\varphi(z_2) - \varphi(z_4)}{z_2 - z_4}.$$

   The above two inequalities lead to

$$\frac{\varphi(z_3) - \varphi(z_1)}{z_3 - z_1} < \frac{\varphi(z_3) - \varphi(z_4)}{z_3 - z_4} < \frac{\varphi(z_2) - \varphi(z_4)}{z_2 - z_4},$$

   which proves (33).

■

Proof to Lemma 4.2: Since $w(s)$ is strictly convex, from Corollary 7.5, there is

$$\frac{w(z_3) - w(z_1)}{z_3 - z_1} < \frac{w(z_2) - w(z_4)}{z_2 - z_4}.$$

Since $z_1 < z_3$, $z_4 < z_2$ and $w(s)$ is decreasing with s,

$$\begin{aligned} w(z_1) - w(z_3) > 0, \\ w(z_4) - w(z_2) > 0. \end{aligned}$$

Then

$$\frac{w(z_1) - w(z_3)}{z_3 - z_1} > \frac{w(z_4) - w(z_2)}{z_2 - z_4},$$

i.e.,

$$w(z_1) - w(z_3) > \frac{z_3 - z_1}{z_2 - z_4}(w(z_4) - w(z_2)).$$

From the condition $z_3 - z_1 \geq z_2 - z_4 > 0$, there is

$$w(z_1) - w(z_3) > w(z_4) - w(z_2),$$

i.e.,

$$w(z_1) + w(z_2) > w(z_3) + w(z_4).$$

■

ACKNOWLEDGMENT

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

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, 2002.

[2] K.-D. Kim and P. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, 2012.

[3] WirelessHART [Online]. Available: http://www.hartcomm.org/

[4] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," in *Proc. 43rd IEEE Conf. Decision Control*, 2004, vol. 4, pp. 3527–3532.

[5] Y. Xu and J. P. Hespanha, "Estimation under uncontrolled and controlled communications in networked control systems," in *Proc. 44th IEEE Conf. Decision Control/ Eur. Control Conf.*, 2005, pp. 842–847.

[6] V. Krishnamurthy and D. V. Djonin, "Structured threshold policies for dynamic sensor scheduling-A partially observed Markov decision process approach," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4938–4957, 2007.

[7] 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.

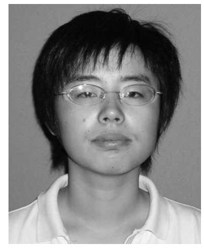
[8] A. S. Chhetri, D. Morrell, and A. Papandreou-Suppappola, "On the use of binary programming for sensor scheduling," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2826–2839, 2007.

[9] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "Stochastic sensor scheduling for energy constrained estimation in multi-hop wireless sensor networks," *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2489–2495, 2011.

[10] 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, May 2009.

[11] L. Shi and H. Zhang, "Scheduling two Gauss-Markov systems: An optimal solution for remote state estimation under bandwidth constraint," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2038–2042, 2012.

[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, 2011.



Chao Yang received the B.S. degree in mechanics and aerospace engineering from Peking University, Peking, China, in 2009.

She 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. Her research interests include sensor scheduling, wireless sensor networks and remote state estimation.

Junfeng Wu received the B.Eng. degree from 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. His research interests include networked control systems, distributed state estimation, and wireless sensor networks.

Wei Zhang received the B.E. degree in automatic control from the University of Science and Technology of China, Hefei, China, in 2003 and the M.S. degree in mathematical statistics and Ph.D. degree in electrical engineering both from Purdue University, West Lafayette, IN, USA, in 2009.

Between January 2010 and August 2011, he was a Postdoctoral Researcher in the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA. His research interests include control and estimation of hybrid and stochastic dynamical systems, and their applications in various engineering fields, especially power systems, robotics, and intelligent transportation.

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 at 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.