ITEA Journal 2009; 30: 549–556

Power Constrained Distributed Estimation with Cluster-Based Sensor Collaboration

Jun Fang, Ph.D. and Hongbin Li, Ph.D.
Stevens Institute of Technology,
Department of Electrical and Computer Engineering, Hoboken, New Jersey
Joseph Dorleus, Ph.D.
U.S. Army Program Executive Office for Simulation,
Training, and Instrumentation, Orlando, Florida
Hong-Liang Cui, Ph.D.
Stevens Institute of Technology,
Department of Physics, Hoboken, New Jersey

We consider the problem of distributed estimation in a power constrained collaborative wireless sensor network, where the network is divided into a set of sensor clusters, with collaboration allowed among sensors within the same cluster but not across clusters. Specifically, each cluster forms one or multiple local messages via sensor collaboration (in particular, linear operation is considered) and transmits the messages over noisy channels to a fusion center. The final estimate is constructed at the fusion center based on the noisy data received from all clusters. In this collaborative setup, we study the following fundamental problems. Given a total transmit power constraint, shall we transmit the raw data or some low-dimensional local messages for each cluster? What is the optimal collaboration scheme for each cluster? How do we optimally allocate the power among different clusters? These questions are addressed in this article. We will show that the optimum collaboration strategy is to compress the data into one local message that, depending on the channel characteristics, is transmitted using one or multiple available channels to the fusion center. The optimal power allocation among the clusters is also investigated, which yields a water-filling type of scheme.

Key words: Distributed estimation; wireless sensor network; sensor clusters; power allocation; collaboration strategy; data transmission; estimation distortion.

Distributed estimation has attracted much attention recently. One of the network architectures for distributed estimation involves a set of spatially distributed sensors linked with a fusion center (FC). Each sensor makes a noisy observation of the phenomena of interest and transmits its processed information to the FC, where a final estimate is formed. The problem of optimal power allocation among sensors given a total transmit power constraint was considered in Cui et al. (2007), Li and AlRegib (2007), Wu, Huang, and Lee (2008), and Xiao et al. (2006); the goal was to minimize the estimation distortion at the FC. For most of these works, intersensor communication is not considered. Intersensor collaboration can indeed be exploited to enhance transmission energy efficiency and improve system performance.

In this article, we consider distributed estimation in a hierarchical network architecture with localized collaboration. Specifically, we assume that the network is divided into a number of sensor clusters linked with a FC. The sensors within the same cluster have the communication resources to locally collaborate, whereas no collaboration is allowed across clusters. This might be the case for scenarios where multiple sets of sensors are spatially distributed, with each set of sensors within a small neighborhood. Each cluster

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **DEC 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Power Constrained Distributed Estimation with Cluster-Based Sensor Collaboration**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Program Executive Office for Simulation, Training, and Instrumentation,12350 Research Parkway,Orlando,FL,32826-3276**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

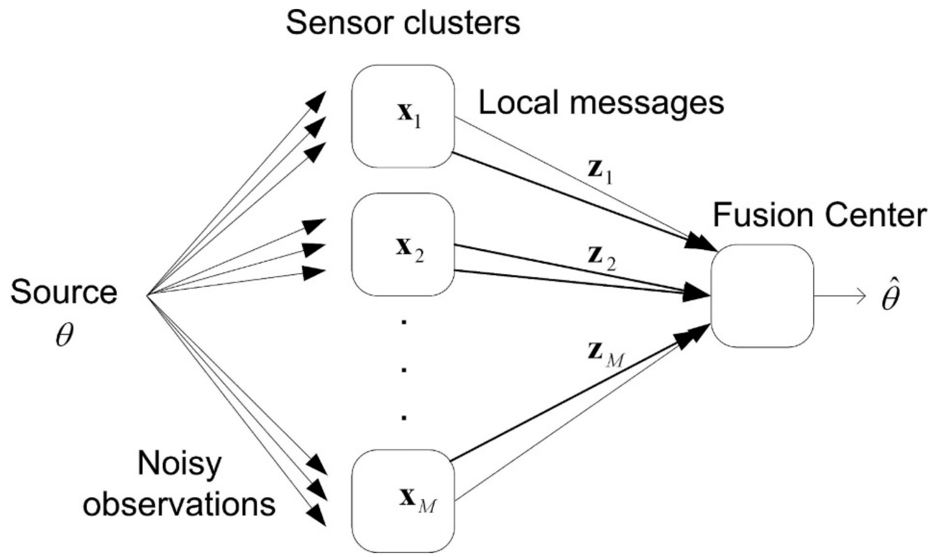

Figure 1. Collaborative setting: the network is divided into a number of sensor clusters. Sensors within each cluster can collaborate to convert their noisy observations {$\mathbf{x}_m$} into some local messages: {$\mathbf{z}_m$}.

then transmits one or multiple one-dimensional messages, which could be the raw data or obtained via sensor collaboration over noisy channels to the FC where a final estimate is formed based on the data received from all clusters. In this context, the following natural questions arise: Given a fixed amount of total transmit power, how should each cluster process its local measurements such that a minimum estimation distortion can be achieved at the FC? How should we allocate the power among the different clusters in an optimal power-distortion fashion? These questions are be addressed in this article, and we develop a fundamental understanding of this important hierarchical collaborative strategy for distributed estimation. Our work is closely related to the distributed compression-estimation approaches in Fang and Li (2008), Luo, Giannakis, and Zhang (2005), Schizas, Giannakis, and Luo (2007), Song, Zhu, and Zhou (2005), Zhang et al. (2003), and Zhu et al. (2005); their objective is to reduce the transmission requirements via dimensionality reduction. While sharing certain similarities with the distributed compression-estimation approaches, our work focuses on the optimal collaboration among sensors in a power constrained scenario.

System model and problem formulation

We consider a wireless sensor network consisting of N spatially distributed sensors, with each sensor making a noisy observation of an unknown random parameter θ: $x = h_n\theta + w_n$, where h_n denotes the observation gain and w_n denotes the additive observation noise. The sensors in the network are divided into M sensor clusters *(Figure 1)*. Each cluster, say cluster m, consists of N_m closely located sensors. The sensors in each cluster are able to collaborate to form local messages that are sent to the FC, whereas no communication is allowed across different clusters. The objective is to obtain an estimate of the unknown parameter at the FC based on the information received from the clusters. In practice, the sensor collaboration can be easily implemented. For each cluster, we choose one sensor to be the cluster head whose task is to collect the data from other sensors within the same cluster and carry out the collaborative processing. The resultant local messages are then transmitted by the cluster head to the FC. We adopt the following assumptions for this collaborative setting.

A1: The links between sensors and the cluster head within each cluster are ideal. Sensor collaboration is confined to linear operations.

A2: An uncoded analog amplify-and-forward scheme is employed to transmit the local messages from the cluster heads to the FC over noisy, wireless channels.

For notational convenience, we use $x_{m,n}$ to denote the sensor measurement of sensor n in cluster m, where $n \in \{1,\ldots, N_m\}$, $m \in \{1,\ldots,M\}$, and

$$x_{m,n} = h_{m,n}\theta + w_{m,n} \quad (1)$$

in which $h_{m,n}$ and $w_{m,n}$ denote the corresponding observation gain and additive observation noise, respectively. To capture the cluster-based collaborative scenario, we write the measurements within a cluster in a vector form: $\mathbf{x}_m = [x_{m,1}\ x_{m,2}\ \ldots\ x_{m,N_m}]^T$, which is given by

$$\mathbf{x}_m = \mathbf{h}_m\theta + \mathbf{w}_m, \quad (2)$$

with $\mathbf{h}_m = [h_{m,1}\ h_{m,2}\ \ldots\ h_{m,N_m}]^T$ and $\mathbf{w}_m = [w_{m,1}\ w_{m,2}\ \ldots\ w_{m,N_m}]^T$. The local messages via sensor collaboration within each cluster can therefore be expressed as

$$\mathbf{z}_m = \mathbf{C}_m\mathbf{x}_m, \quad (3)$$

where $\mathbf{C}_m \in R^{p_m \times N_m}$ denotes the collaboration matrix for cluster m, $p_m \leq N_m$ is the dimensionality of the message vector $\mathbf{z}_m$ whose choice is discussed later. The signal received at the FC from the mth cluster is given by

$$\mathbf{y}_m = \mathbf{G}_m\mathbf{A}_m\mathbf{C}_m\mathbf{x}_m + \mathbf{v}_m, \quad (4)$$

where $\mathbf{G}_m \in R^{p_m \times p_m}$ denotes a fading multiplicative channel matrix, which can be diagonal or nondiagonal, depending on the transmission scheme (e.g., orthogonal vs. nonorthogonal channel access); $\mathbf{A}_m = \mathrm{diag}\{a_1, \ldots, a_{p_m}\}$ is the amplification matrix with a_i denoting the amplification factor used in transmitting the ith message of $\mathbf{z}_m$; $\mathbf{v}_m \in R^{p_m}$ denotes the additive channel noise vector. Without loss of generality, we assume $\mathbf{G}_m = \mathbf{I}$ and $\mathbf{G}_m = \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix, because the multiplicative effect of the channel matrix can be removed by carrying out a matrix inverse using an estimate of the channel matrix $\mathbf{G}_m$ at the receiver and the amplification matrix A_m can be absorbed into $\mathbf{C}_m$. We have the following assumption regarding observation noise $\{\mathbf{w}_m\}$ and channel noise $\{\mathbf{v}_m\}$.

A3: Noise $\{\mathbf{w}_m\}$ and $\{\mathbf{v}_m\}$ are zero mean with positive definite autocovariance $\{\mathbf{R}_{w,m}\}$ and $\{\mathbf{R}_{v,m}\}$, respectively, which are available at the FC. The noise across different clusters is mutually uncorrelated, i.e., $E[\mathbf{w}_i\mathbf{w}_j{}^T] = \mathbf{0}$ and $E[\mathbf{v}_i\mathbf{v}_j{}^T] = \mathbf{0}\ \forall i \neq j$.

Let $\mathbf{y} = [\mathbf{y}_1\ \mathbf{y}_2\ \ldots\ \mathbf{y}_M]^T$ denote a column vector formed by stacking the data received from all clusters. We have

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{v} = \mathbf{C}(\mathbf{h}\theta + \mathbf{w}) + \mathbf{v}, \quad (5)$$

where $\mathbf{C} = \mathrm{diag}\{\mathbf{C}_1, \ldots, \mathbf{C}_M\}$ is a block diagonal matrix with its mth block-diagonal element equal to $\mathbf{C}_m$, $\mathbf{x} = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_M]^T$, $\mathbf{v} = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_M]^T$, $\mathbf{h} = [\mathbf{h}_1\ \mathbf{h}_2\ \ldots\ \mathbf{h}_M]^T$, and $\mathbf{w} = [\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_M]^T$. A natural question arising from this scenario is to find an overall optimal collaboration matrix $\mathbf{C}$, or equivalently, a set of individual collaboration matrices $\{\mathbf{C}_m\}_{m=1}^M$, to achieve a minimum estimation distortion at the FC. Also, because the amplification factors $\{\mathbf{A}_m\}$ are incorporated into the collaboration matrices $\{\mathbf{C}_m\}$, the overall collaboration matrix $\mathbf{C}$ has to satisfy a total transmit power constraint. Specifically, using a Linear Minimum Mean-Square Error (LMMSE) estimator (Kay 1993), it can be readily verified that we are faced with the following optimization problem:

$$\min_{\mathbf{C}}\ E[(\theta - \hat{\theta})^2] = \sigma_\theta^2 - \sigma_\theta^4\mathbf{h}^T\mathbf{C}^T(\mathbf{C}\mathbf{R}_x\mathbf{C}^T + \mathbf{R}_v)^{-1}\mathbf{C}\mathbf{h}$$
$$\text{s.t. } \mathrm{tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T) \leq P, \quad (6)$$

where σ_θ^2 denotes the signal variance; $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^T]$; $\mathrm{tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T)$ is the average transmit power required to send the local messages from all clusters to the FC; and P is a prespecified power budget for transmission.

Single cluster case

The development of the optimal collaboration matrix for the single cluster case is quite involved. Because of space limitations, we only present the main results without providing the proof.

THEOREM 1: Consider the optimal collaboration design problem formulated in (6) and described in *Figure 1*, where the sensor measurements $\mathbf{x}_m$, the local messages $\mathbf{z}_m$, and the received messages at the FC $\mathbf{y}_m$ are given by (2), (3), and (4), respectively. When $M = 1$, the optimal solution to (6) is

$$\mathbf{C}^* = \gamma\sqrt{P}\mathbf{U}_v[:, 1]\mathbf{h}^T\mathbf{R}_x^{-1}, \quad (7)$$

where $\mathbf{U}_v[:, 1]$ denotes the first column of $\mathbf{U}_v$; $\mathbf{U}_v$ is an orthonormal matrix obtained by carrying out the eigenvalue decomposition of $\mathbf{R}_v$, i.e., $\mathbf{R}_v = \mathbf{U}_v\mathbf{D}_v\mathbf{U}_v^T$; and $\gamma = 1/\sqrt{\mathbf{h}^T\mathbf{R}_x^{-1}\mathbf{h}}$. The associated estimation Mean Square Error (MSE), i.e., the value of the minimum objective function of (6), is given by

$$E\{[\theta - \hat{\theta}(\mathbf{C}^*)]^2\} = \sigma_\theta^2 - \sigma_\theta^4\frac{P}{P + \min(\mathbf{d}_v)}\mathbf{h}^T\mathbf{R}_x^{-1}\mathbf{h} \quad (8)$$

PROOF: A rigorous proof is provided in Fang and Li (in press).

The optimal solution (7) has very important implications that we shall explore in the following. Considering the scenario of independent channels, i.e., $\mathbf{R}_v$ is diagonal; $\mathbf{U}_v = \mathbf{I}$ and $\mathbf{U}_v[:, 1] = \mathbf{e}_1$, where $\mathbf{e}_i$ denotes the unit column vector with its ith entry equal to 1 and its other entries equal to 0. Therefore the optimal collaboration matrix becomes

$$\mathbf{C}^* = \begin{bmatrix} \gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1} \\ \mathbf{0}_{(p-1)\times N} \end{bmatrix}, \quad (9)$$

which is a matrix with its first row equal to

$\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1}$ and all other rows equal to 0. The solution suggests that we should compress the measurements into only one local message and transmit it via the best-quality channel (note that the first row corresponds to the first channel, which has the smallest noise variance because the diagonal elements of $\mathbf{R}_v$ are assumed in an ascending order) to the FC. If the channels have identical qualities, then we can use any of them to send out the local message. Also, by rewriting the collaboration weighting vector $\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1}$ as

$$\begin{aligned}\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1} &= \gamma\sqrt{P}\sigma_\theta^{-2}\sigma_\theta^2\mathbf{h}^T\mathbf{R}_x^{-1}\\ &= \gamma\sqrt{P}\sigma_\theta^{-2}\mathbf{R}_{\theta x}\mathbf{R}_x^{-1},\end{aligned} \tag{10}$$

where $\mathbf{R}_{\theta x}= E[\theta\mathbf{x}^T]$, we can immediately see that the local message is exactly the LMMSE estimate $\mathbf{R}_{\theta x}\mathbf{R}_x^{-1}\mathbf{x}$ multiplied by a scalar $\gamma\sqrt{P}\sigma_\theta^{-2}$. This means that when channels are independent, LMMSE estimation followed by an amplification factor is optimal in a power-distortion sense.

We now investigate the case where the channels are correlated, i.e., $\mathbf{R}_v$ is nondiagonal. Each row of the optimal collaboration matrix can be readily expressed as follows by combining (7) and (10):

$$\mathbf{C}^*[i,:] = \mathbf{U}_v[i,1]\gamma\sqrt{P}\sigma_\theta^{-2}\mathbf{R}_{\theta x}\mathbf{R}_x^{-1}, \tag{11}$$

where $\mathbf{U}_v[i,1]$ is the $(i,1)$th entry of $\mathbf{U}_v$. Therefore the LMMSE estimate is transmitted by multiple channels with different amplification gains that are proportional to $\{\mathbf{U}_v[i,1]\}_{i=1}^p$. The number of local messages to be transmitted, p, $1 \le p \le N$, should be as large as possible because the more channels employed, the more diversity that can be provided. However, system complexity will also increase as more channels are involved.

Multiple cluster case

We now examine a general scenario where the network consists of multiple sensor clusters. In this case, the collaboration matrix $\mathbf{C}$ has a block diagonal structure because intercluster collaboration is not allowed. The approach described in previous subsection, therefore, cannot be directly applied here. To solve (6), we hope to decouple the optimization problem into a set of tractable subtasks. To this goal, we rewrite the estimation MSE as follows (Fang and Li in press).

$$\begin{aligned}E[(\theta-\hat{\theta})^2] &= \sigma_\theta^2 - \sigma_\theta^4\mathbf{h}^T\mathbf{C}^T(\mathbf{C}\mathbf{R}_x\mathbf{C}^T+\mathbf{R}_v)^{-1}\mathbf{C}\mathbf{h}\\ &= \left(\sigma_\theta^{-2} + \sum_{i=1}^{M}\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\right)^{-1},\end{aligned} \tag{12}$$

where we use the fact that $\mathbf{R}_x = \sigma_\theta^2\mathbf{h}\mathbf{h}^T + \mathbf{R}_w$, along with the block diagonal structures of $\mathbf{C}$, $\mathbf{R}_w$, and $\mathbf{R}_v$. Therefore the optimization problem (6) becomes

$$\begin{aligned}\max_{\{\mathbf{C}_i\}} \ & \sum_{i=1}^{M}\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ \text{s.t.} \ & \sum_{i=1}^{M}\text{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P\end{aligned} \tag{13}$$

in which the power constraint follows from

$$\text{tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T) = \sum_{i=1}^{M}\text{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T).$$

To use the theoretical results obtained for $M = 1$, we express the component $\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i$ in (13) as a function of $\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i$, which can be done by resorting to the Woodbury identity:

$$\begin{aligned}&\sigma_\theta^2 - \sigma_\theta^4\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ &\quad = (\sigma_\theta^{-2} + \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i)^{-1}.\end{aligned} \tag{14}$$

For notational convenience, let

$$\begin{aligned}\mu_i(\mathbf{C}_i) &= \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ \eta_i(\mathbf{C}_i) &= \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i.\end{aligned} \tag{15}$$

Therefore (14) can be rewritten as

$$\mu_i(\mathbf{C}_i) = \frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)} - 1\right). \tag{16}$$

Substituting (16) into (13), we arrive at the following optimization

$$\begin{aligned}\max_{\{\mathbf{C}_i\}} \ & \sum_{i=1}^{M}\frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)} - 1\right)\\ \text{s.t.} \ & \sum_{i=1}^{M}\text{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P.\end{aligned} \tag{17}$$

Clearly, (17) can be decoupled into two sequential subtasks, i.e., a power allocation (among clusters) problem and a set of collaboration matrix design problems that can be solved using the previous results. To see this, suppose $\{P_1^*, P_2^*, \ldots, P_M^*\}$ is an optimum power assignment with

$$\text{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P_i^* \qquad \forall i \in \{1,\ldots,M\}$$

$$\sum_{i=1}^{M}P_i^* \le P$$

then (17) is simplified into a set of identical problems as

$$\max_{\{\mathbf{C}_i\}} \frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)}-1\right)$$
$$\text{s.t. } \operatorname{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T)\le P_i^*. \tag{18}$$

Note that $\sigma_\theta^2\eta_i(\mathbf{C}_i)$ must lie within the interval (0, 1) because we have $\eta(\mathbf{C}_i) > 0$ and $\mu_i(\mathbf{C}_j) > 0$ from their definitions. Hence (18) is equivalent to

$$\max_{\mathbf{C}_i} \ \eta_i(\mathbf{C}_i)$$
$$\text{s.t. } \operatorname{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T)\le P_i^*, \tag{19}$$

which is exactly the optimization problem discussed in the previous section. The optimal solution to (19) is given in Theorem 1. The key problem, therefore, is to determine the optimum power assignment $\{P_1^*, P_2^*, \ldots, P_M^*\}$. To meet this goal, we need to find out the relationship between the maximum objective function value $\eta_i(\mathbf{C}_i^*)$ and P_i^*. Recalling Theorem 1, more precisely, (8), we have

$$\eta_i(\mathbf{C}_i^*)=\frac{P_i^*}{P_i^*+\min(\mathbf{d}_{v,i})}\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i=\frac{\alpha_iP_i^*}{\beta_i+P_i^*}, \tag{20}$$

where we define $\alpha_i = \mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$, $\beta_i = \min(\mathbf{d}_{v,j})$, and $\mathbf{d}_{v,j}$ is a column vector consisting of the eigenvalues of $\mathbf{R}_{v,i}$ (note that $\mathbf{R}_{v,j}$ can be nondiagonal). Substituting (20) into the objective function of (17), we get

$$\sum_{i=1}^{M}\frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i^*)}-1\right)=\sum_{i=1}^{M}\frac{\alpha_iP_i^*}{(1-\sigma_\theta^2\alpha_i)P_i^*+\beta_i}. \tag{21}$$

Clearly, the optimal power allocation $\{P_1^*, P_2^*, \ldots, P_M^*\}$ must be the one, among all feasible power assignments, that maximizes (21). Therefore, it can be found out by

$$\min_{\{P_1,\ldots,P_M\}} -\sum_{i=1}^{M}\frac{\alpha_iP_i}{(1-\sigma_\theta^2\alpha_i)P_i+\beta_i}$$
$$\text{s.t. } \sum_{i=1}^{M}P_i\le P$$
$$P_i\ge 0 \quad \forall i\in\{1,\ldots,M\}. \tag{22}$$

It is easy to verify that the optimization problem (22) is convex because its Hessian matrix, which is a diagonal matrix in this case, is positive semidefinite on the convex set defined by the linear constraints. Although (22) is efficiently solvable by numerical methods, it can also be solved analytically by resorting to the Lagrangian function and Karush-Kuhn-Tucker conditions, which leads to a water-filling type power allocation scheme. The details are omitted here because of space limitations. Briefly speaking, for a threshold λ, we have

$$P_i=\begin{cases}\frac{1}{\varphi_i}\left(\sqrt{\frac{\phi_i}{\lambda}}-1\right) & \phi_i\ge\lambda\\ 0 & \text{otherwise}\end{cases} \tag{23}$$

where $\phi_i = \alpha_i/\beta_i$, $\varphi_i=(1-\sigma_\theta^2\alpha_i)/\beta_i$. It is easy to see that each cluster can decide whether to transmit or keep silent by the criterion $\phi_i \ge \lambda$. Note that ϕ_i is the ratio of $\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$ to min $(\mathbf{d}_{v,i})$, with the former is a measure of the cluster's estimation quality (a larger value indicates a better estimation accuracy) and the latter a measure of the cluster's channel quality (a smaller value indicates a better channel quality).

So far we have developed an analytical approach that leads to an optimal solution to (6). For clarity, we now summarize the steps of our proposed method.

1. Given the prior knowledge of the autocorrelation matrices $\{\mathbf{R}_{v,i}\}_{i=1}^M$, $\{\mathbf{R}_{w,i}\}_{i=1}^M$, and the observation gain vectors $\{\mathbf{h}_i\}_{i=1}^M$, compute $\{\alpha_i\}_{i=1}^M$, and $\{\beta_i\}_{i=1}^M$, where $\alpha_i=\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$ and $\beta_i = \min(\mathbf{d}_{v,\mathrm{i}})$.
2. Given the total power constraint P, find the optimal power allocation among clusters via (22).
3. With the optimal power assignment $\{P_1^*, P_2^*, \ldots, P_M^*\}$ derived in the previous step, determine the optimal collaboration matrices $\{\mathbf{C}_i\}_{i=1}^M$ via (19), whose solution is detailed in Theorem 1.

Simulation results

We consider the single cluster case and carry out a simple performance analysis to corroborate our theoretical results (more analysis and simulation results are available in Fang and Li [in press]). We compare our optimal collaboration strategy with the scheme proposed in Cui et al. (2007), where there is no intersensor collaboration and each sensor transmits its observation to the FC with optimally assigned power. For simplicity, we consider a homogeneous environment with identical observation and channel qualities, where σ_w^2 denotes the observation noise variance and σ_v^2 represents the channel noise variance. Also, all observation and channel gains are assumed to be unitary, i.e., =1, throughout all examples in the article. Clearly, an equal power allocation is optimum for Cui et al. (2007) and the corresponding estimation MSE can be shown to be

$$\mathrm{MSE}_{\mathrm{NC}}=\frac{P\sigma_w^2\sigma_\theta^2+N\sigma_v^2\sigma_\theta^4+N\sigma_v^2\sigma_\theta^2\sigma_w^2}{PN\sigma_\theta^2+P\sigma_w^2+N\sigma_v^2\sigma_\theta^2+N\sigma_v^2\sigma_w^2}, \tag{24}$$

where the subscript NC denotes noncollaboration. For our collaboration strategy, the estimation MSE can be

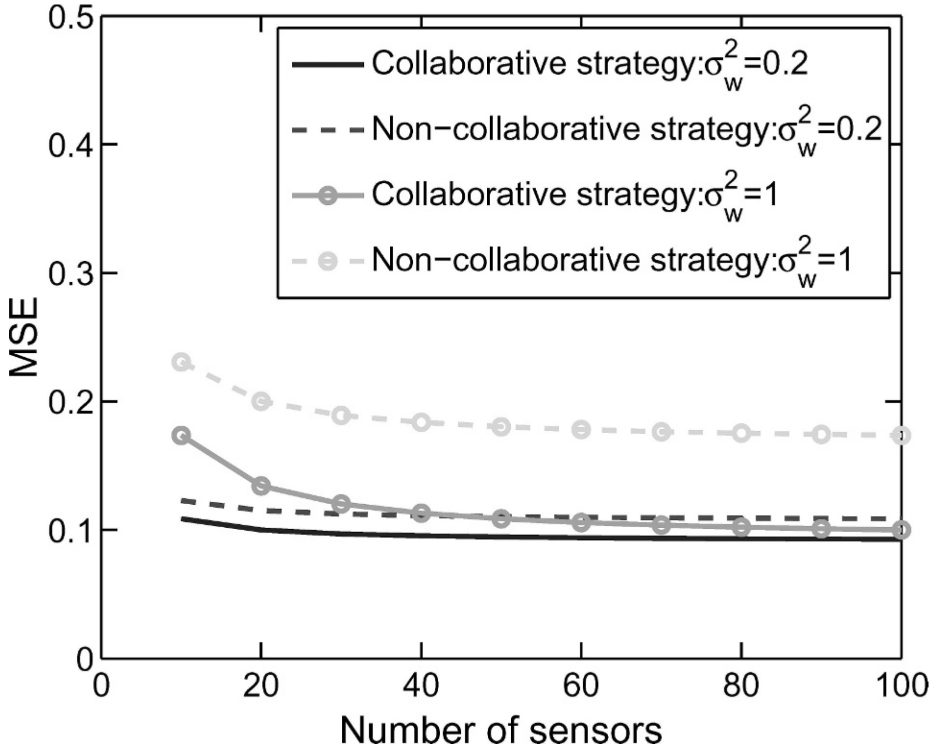

Figure 2. MSEs of collaborative and noncollaborative strategies versus number of sensors $\sigma_v^2=0.1$, $\sigma_\theta^2=1$, and P = 1.

computed by using (8), which reduces to

$$\mathrm{MSE}_{\mathrm{OC}} = \frac{P\sigma_w^2\sigma_\theta^2 + N\sigma_v^2\sigma_\theta^4 + \sigma_v^2\sigma_\theta^2\sigma_w^2}{PN\sigma_\theta^2 + P\sigma_w^2 + N\sigma_v^2\sigma_\theta^2 + \sigma_v^2\sigma_w^2}, \qquad (25)$$

where the subscript OC denotes optimal collaboration. For notational convenience, let $a = P\sigma_w^2\sigma_\theta^2 + N\sigma_v^2\sigma_\theta^4$ and $b = PN\sigma_\theta^2 + P\sigma_w^2 + N\sigma_v^2\sigma_\theta^2$. It can be easily verified that

$$(a + N\sigma_v^2\sigma_\theta^2\sigma_w^2)(b + \sigma_v^2\sigma_w^2) \geq (a + \sigma_v^2\sigma_\theta^2\sigma_w^2)(b + N\sigma_v^2\sigma_w^2), \qquad (26)$$

where (26) becomes an equality only when $N = 1$. Hence as expected, the relationship MSENC ≥ MSEOC holds, which means that the optimal collaboration scheme should always outperform the noncollaboration scheme.

Figure 2 depicts the estimation MSEs of the two schemes as a function of N under a total transmit power constraint, with $\sigma_w^2 = 0.2$ and $\sigma_w^2 = 1$, respectively. From *Figure 2*, we see that both schemes benefit from an increasing number of sensors; as N increases, the estimation MSEs will asymptotically approach certain values that, however, are nonzero. This observation can be readily verified from (24)–(25). Also, it can be seen that the noncollaborative scheme is sensitive to the value of σ_w^2; as the observation quality deteriorates, its performance degrades considerably. In contrast, the collaborative strategy demonstrates a certain degree of robustness against the observation quality deterioration. In *Figure 3*, we plot the estimation MSE versus the total transmit power. We see that the performance gap between the two strategies shrinks as the transmit power increases. In fact, from (24)–(25) we observe that as the transmit power goes to infinity, these two strategies approach identical performance. This suggests that the collaborative strategy should be preferred especially when the sensor observation qualities are bad and transmit power is severely constrained.

Conclusion

We studied an optimal collaboration and power allocation problem for distributed estimation in a power-constrained collaborative sensor network, where the network consists of a number of sensor clusters, and collaboration is allowed within the same cluster but not across clusters. Our theoretical results showed that, given a specified total transmit power, the power should be assigned among the clusters in a water-filling manner, with each cluster deciding whether to transmit or keep silent by comparing with a threshold the ratio of a measure of the cluster's estimation quality to a measure of the cluster's channel quality. Also, for each cluster, if the channels from this cluster to the FC are independent, then an optimal collaboration yields only one local message, which is sent from the best channel

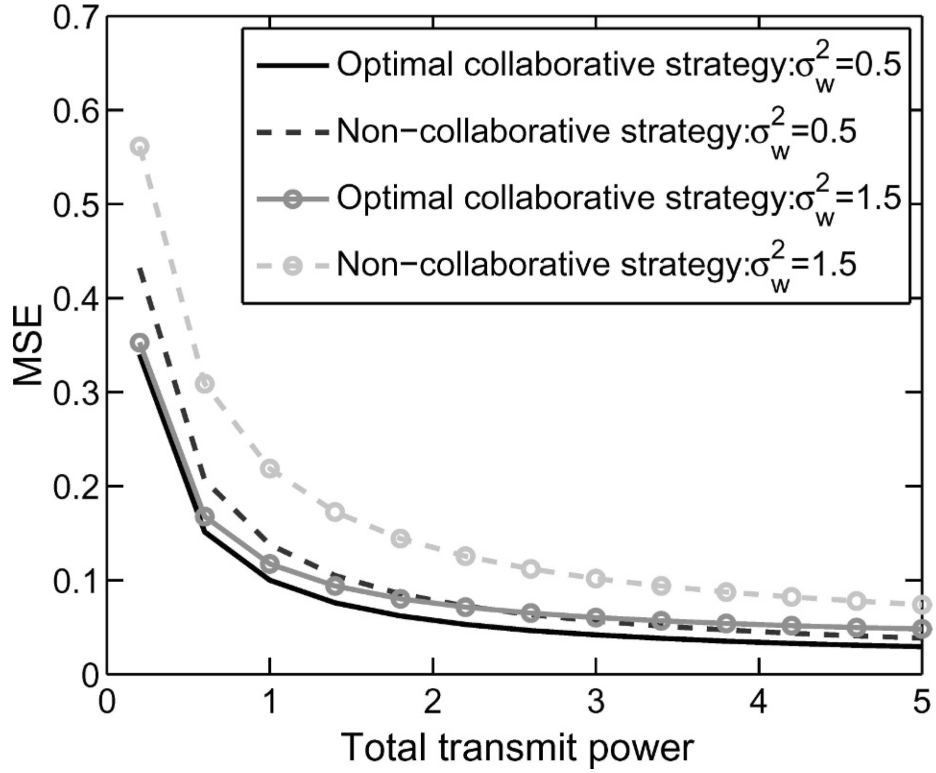

Figure 3. MSEs of collaborative and noncollaborative strategies versus total transmit power. $\sigma_v^2 = 0.1$, $\sigma_\theta^2 = 1$, *and N = 50.*

within the cluster to the FC; otherwise the local message has to be sent across all channels within the cluster at different power levels matched to their channel quality. Specifically, in either case, the compressed local message is exactly the local LMMSE estimate multiplied by an amplification factor. Simulation results have been presented to corroborate our theoretical analysis. ❑

JUN FANG received a bachelor's of science and a master's of science degree in electrical engineering from Xidian University, Xi'an, China, in 1998 and 2001, respectively, and earned a doctor of philosophy degree in electrical engineering from the National University of Singapore, Singapore, in 2006. During 2006 he was with the Department of Electrical and Computer Engineering, Duke University, as a postdoctoral research associate. Currently he is a postdoctoral research associate with the Department of Electrical and Computer Engineering at the Stevens Institute of Technology. His research interests include statistical signal processing, wireless communications, and distributed estimation and detection with their applications on wireless sensor networks. E-mail: Jun.Fang@stevens.edu

JOSEPH DORLEUS is currently a lead telecommunication/systems engineer at PEO STRI, Orlando, Florida. He has worked and held both technical and managerial positions in the private sector as well as in the government. He holds bachelor's of science and master's of science degrees in electrical engineering from Polytechnic University (formerly Polytechnic Institute of New York), Brooklyn, New York, and a doctor of philosophy degree in electrical engineering from Stevens Institute of Technology, Hoboken, New Jersey. His research interests include optical networks, all-optical network management and monitoring, and modeling and simulation of wireless networks. He is a member of the International Test and Evaluation (ITEA), the Institute of Electrical, Electronics Engineering (IEEE), the Defense Technical Information Center (DTIC), the Army Acquisition Corps (AAC), and the International Society for Optical Engineering (SPIE). He has authored, coauthored, and presented numerous technical papers that are published in technical journals, conferences, and proceedings such as IEEE, SPIE, ITSEC, and ITEA. Dr. Dorleus is the recipient of the Army Achievement Medal for Exceptional Civilian Service. He was also the Army Materiel Command's nominee for Black Engineer of Year Award in 2001. E-mail: Joseph.dorleus@us.army.mil

HONGBIN LI received his bachelor's of science and master's of science degrees from the University of Electronic Science and Technology of China, Chengdu, in 1991 and 1994, respectively, and a doctor of philosophy degree from the University of Florida, Gainesville, Florida, in 1999, all in electrical engineering. From July 1996 to May 1999, he was

a research assistant in the Department of Electrical and Computer Engineering at the University of Florida. He was a visiting summer faculty member at the Air Force Research Laboratory, Rome, New York, in the summers of 2003 and 2004. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey, where he is an associate professor. His current research interests include statistical signal processing, wireless communications, and radars. Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. He is/has been an editor or associate editor for the IEEE Transactions on Wireless Communications, IEEE Signal Processing Letters, and IEEE Transactions on Signal Processing, and served as a guest editor for EURASIP Journal on Applied Signal Processing, Special Issue on Distributed Signal Processing Techniques for Wireless Sensor Networks. E-mail: Hongbin.Li@stevens.edu

HONG-LIANG CUI is a professor of physics at Stevens Institute of Technology, where he directs the Applied Electronics Laboratory. He received his undergraduate education in applied physics with a concentration in laser optics from the Changchun Institute of Optics and Fine Mechanics in Changchun, China, with a bachelor's of engineering degree. In 1981 he came to the United States for graduate study as one of the first group of Chinese physics students in the CUSPEA program, obtaining a doctor of philosophy degree in theoretical condensed matter physics in 1987 from Stevens Institute of Technology, where he has been on the faculty ever since. His research efforts have been concentrated in the areas of solid-state electronics/nanoelectronics, optical communications and sensing, electromagnetic wave propagation, and interaction with matters such as chemical and bioagents, and high-performance computing approach to modeling of physical devices and phenomena. His work has been funded by NSF, ARO, ONR, and DARPA. He has published more than 190 research papers in peer-reviewed scientific journals, holds 9 U.S. patents, and guided more than 30 doctoral dissertations to completion. He holds membership in the American Physical Society, the Institute of Electrical and Electronics Engineers, the Optical Society of America, and Sigma Xi. E-mail: Hong-Liang.Cui@stevens.edu

References

Cui, S., J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor. 2007. Estimation diversity and energy efficiency in distributed sensing. *IEEE Transactions on Signal Processing*, no. 9 (55): 4683–4695.

Fang, J., and H. Li. 2008. Joint dimension assignment and compression for distributed multi-sensor estimation. *IEEE Signal Processing Letters*. 15, 174–177.

Fang, J., and H. Li. 2009. Power constrained distributed estimation with cluster-based sensor collaboration. *IEEE Transactions on Wireless Communications*. 8: 3822–3832.

Kay, S. M. 1993. *Fundamentals of statistical signal processing: estimation theory*. Upper Saddle River, New Jersey: Prentice Hall.

Li, J., and G. AlRegib. 2007. Rate-constrained distributed estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*. 55 (5): 1634–1643.

Luo, Z.-Q., G. B. Giannakis, and S. Zhang. 2005. Optimal linear decentralized estimation in a bandwidth constrained sensor network. In *Proceedings of the International Symposium on Information Theory*. Adelaide, Australia.

Schizas, I. D., G. B. Giannakis, and Z.-Q. Luo. 2007. Distributed estimation using reduced dimensionality sensor observations. *IEEE Transactions on Signal Processing*, no. 8, 55: 4284–4299.

Song, E., Y. Zhu, and J. Zhou. 2005. Sensors' optimal dimensionality compression matrix in estimation fusion. *Automatica*. 41, 2131–2139.

Wu, J.-Y., Q.-Z. Huang, and T.-S. Lee. 2008. Minimal energy decentralized estimation via exploiting the statistical knowledge of sensor noise variance. *IEEE Transactions on Signal Processing*. 56 (5): 2171–2176.

Xiao, J.-J., S. Cui, Z.-Q. Luo, and A. J. Goldsmith. 2006. Power scheduling of universal decentralized estimation in sensor networks. *IEEE Transactions on Signal Processing*. 54 (2): 413–422.

Zhang, K., X. R. Li, P. Zhang, and H. Li. 2003. Optimal linear estimation fusion—Part VI: sensor data compression. In *Proceedings of the International Conference on Information Fusion*. Queensland, Australia.

Zhu, Y., E. Song, J. Zhou, and Z. You. 2005. Optimal dimensionality reduction of sensor data in multisensor estimation fusion. *IEEE Transactions on Signal Processing*. 53 (5): 1631–1639.