IEICE TRANS. COMMUN., VOL.E95–B, NO.2 FEBRUARY 2012

PAPER

Distributed Estimation for Vector Signal in Linear Coherent Sensor Networks*

Chien-Hsien WU[†a)], ***Member and*** **Ching-An LIN**[†], ***Nonmember***

SUMMARY We introduce the distributed estimation of a random vector signal in wireless sensor networks that follow coherent multiple access channel model. We adopt the linear minimum mean squared error fusion rule. The problem of interest is to design linear coding matrices for those sensors in the network so as to minimize mean squared error of the estimated vector signal under a total power constraint. We show that the problem can be formulated as a convex optimization problem and we obtain closed form expressions of the coding matrices. Numerical results are used to illustrate the performance of the proposed method.
key words: *distributed estimation, convex optimization, power allocation, wireless sensor network*

1. Introduction

Low power consumption and efficient bandwidth usage are two critical issues for distributed estimation in wireless sensor networks [1]. In the distributed estimation scenario, observation data is measured by spatially distributed sensors and transmitted to a fusion center (FC) to generate the final estimate. Due to power and bandwidth limitations, many research works (e.g. [2]–[6]) concentrated on parameter estimation by amplify-and-forward scheme and discussed the problems which restrict the amount of power and information sent from each sensor per observation period. Among these works, optimal power allocation schemes, in the form of optimal power gains, that minimize estimation distortion are proposed based on the orthogonal multiple access channel (MAC) model [2]–[4] as well as the coherent MAC model [5], [6].

All the power allocation schemes mentioned above consider single-input single-output (SISO) systems in which the sensor with a scalar measurement/transmitter is regarded as a unit. Recently, multiple-input multiple-output (MIMO) systems have attracted much attention due to their advantages of increased data rates and improved performance [7]. In MIMO wireless sensor networks, each sensor performs vector measurements and has multiple transmitters. Recent works on MIMO sensor networks addressed the dimensionality reduction problem based on the ideal channel assumption [8]–[10] and the optimal design of coding matrices [11].

Manuscript received April 9, 2011.
Manuscript revised October 14, 2011.
[†]The authors are with the Department of Electrical and Control Engineering, National Chiao Tung University, 1001 University Road, Hsinchu 300, Taiwan.
*Research sponsored by National Science Council under grant NSC 97-2221-E009-046-MY3.
a) E-mail: chwu.ece92g@nctu.edu.tw
DOI: 10.1587/transcom.E95.B.460

In this paper, we consider the distributed estimation of a vector signal using MIMO wireless sensor networks with the coherent MAC model. The measurement for each sensor is encoded by a coding matrix that forms a message for transmission to a FC, where the signal is estimated based on the linear minimum mean-squared error (LMMSE) rule. We study the problem of designing linear coding matrices that minimize the mean-squared error (MSE) under a total power constraint. We show that the problem can be formulated as a convex optimization problem. The solution, which is a water-filling type, gives closed form expressions for the coding matrices.

The organization of this paper is as follows. Section 2 is the network model and the problem statement of linear distributed estimation. In Sect. 3, the proposed method is given and closed form expressions for the coding matrices are shown. Section 4 is the simulation results. Finally, Sect. 5 is a brief conclusion.

Notations: Throughout this paper, the following notations are used. A lower case letter denotes a scalar, a boldface lower case letter denotes a vector, and a boldface uppercase letter denotes a matrix. In addition, $\mathbf{A}^T$ and $\mathbf{A}^H$ denote the transpose of $\mathbf{A}$ and the conjugate transpose of $\mathbf{A}$, respectively. The letter $\mathbf{I}_n$ denotes an identity matrix of size $n \times n$. $\mathbf{0}$ and $\mathbf{0}_{n\times m}$ denote, respectively, a zero vector and a zero matrix of size $n \times m$. The operator $\mathrm{diag}(x_1, \cdots, x_M)$ is a diagonal matrix with its mth diagonal element equal to x_m and $\mathrm{tr}(\mathbf{A})$ is the trace of $\mathbf{A}$.

2. System Model and Problem Formulation

We consider a wireless sensor network consisting of L sensors for estimating p random source signals, written in vector form $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T \in \mathbb{C}^p$, as shown in Fig. 1.

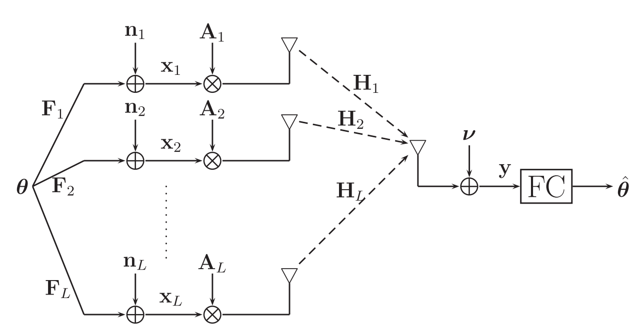

Fig. 1 Linear distributed estimation with coherent MAC.

The lth sensor has k_l measurements, which can be expressed in vector form as

$$\mathbf{x}_l = \mathbf{F}_l\boldsymbol{\theta} + \mathbf{n}_l, \quad 1 \le l \le L \tag{1}$$

where $\mathbf{F}_l \in \mathbb{C}^{k_l\times p}$ is the observation matrix and $\mathbf{n}_l \in \mathbb{C}^{k_l}$ is the additive noise. Let $K = \sum_{l=1}^{L} k_l$. In vector form, (1) can be written as

$$\mathbf{x} = \mathbf{F}\boldsymbol{\theta} + \mathbf{n}, \tag{2}$$

where $\mathbf{x} = [\mathbf{x}_1^T\,\mathbf{x}_2^T\,\cdots\,\mathbf{x}_L^T]^T \in \mathbb{C}^K$, $\mathbf{F} = [\mathbf{F}_1^T\,\mathbf{F}_2^T\,\cdots\,\mathbf{F}_L^T]^T \in \mathbb{C}^{K\times p}$, and $\mathbf{n} = [\mathbf{n}_1^T\,\mathbf{n}_2^T\,\cdots\,\mathbf{n}_L^T]^T \in \mathbb{C}^K$. The measurement of the lth sensor is encoded by a linear coding matrix $\mathbf{A}_l \in \mathbb{C}^{N\times k_l}$ to form the message vector $\mathbf{A}_l\mathbf{x}_l \in \mathbb{C}^N$, where N denotes the number of messages transmitted from the lth sensor. The message vector is then sent to the fusion center (FC) through the channel which is described by the gain matrix $\mathbf{H}_l \in \mathbb{C}^{N\times N}$. From (1), the received signal $\mathbf{y}$ at the FC can be expressed as

$$\mathbf{y} = \sum_{l=1}^{L} \mathbf{H}_l\mathbf{A}_l(\mathbf{F}_l\boldsymbol{\theta} + \mathbf{n}_l) + \boldsymbol{\nu} = \mathbf{B}(\mathbf{F}\boldsymbol{\theta} + \mathbf{n}) + \boldsymbol{\nu}, \tag{3}$$

where $\boldsymbol{\nu} \in \mathbb{C}^N$ is additive noise at the receiver and $\mathbf{B} = [\mathbf{B}_1\,\mathbf{B}_2\,\cdots\,\mathbf{B}_L] \in \mathbb{C}^{N\times K}$ with $\mathbf{B}_l = \mathbf{H}_l\mathbf{A}_l \in \mathbb{C}^{N\times k_l}$. We assume that i) $E[\boldsymbol{\theta}] = \mathbf{0}$ and $E[\boldsymbol{\theta}\boldsymbol{\theta}^H] = \sigma_\theta^2\mathbf{I}_p$, ii) $E[\mathbf{n}_l] = \mathbf{0}$, $E[\mathbf{n}_l\mathbf{n}_l^H] = \sigma_n^2\mathbf{I}_{k_l}$, and $E[\mathbf{n}_l\mathbf{n}_j^H] = \mathbf{0}$ for $j \ne l$, iii) $E[\boldsymbol{\nu}] = \mathbf{0}$ and $E[\boldsymbol{\nu}\boldsymbol{\nu}^H] = \sigma_\nu^2\mathbf{I}_N$, iv) $E[\boldsymbol{\theta}\mathbf{n}^H] = \mathbf{0}$, $E[\boldsymbol{\theta}\boldsymbol{\nu}^H] = \mathbf{0}$, and $E[\mathbf{n}\boldsymbol{\nu}^H] = \mathbf{0}$, v) $\mathbf{F}_l$ and $\mathbf{H}_l$ are known at the FC, vi) $K \ge N$, $K \ge p$, and $\text{rank}(\mathbf{F}) = p$, and vii) $\mathbf{H}_l = \text{diag}(h_1^l, \cdots, h_N^l)$ with $h_i^l \ne 0, \forall i, l$.

Remark 1: Assumption vii) simplifies the derivations to follow. In practice, if the N signals are transmitted using N different frequencies, the assumption is reasonable.

Using the received signal $\mathbf{y}$ in (3), the LMMSE estimate of $\boldsymbol{\theta}$ is [12, p.382]

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= E[\boldsymbol{\theta}\mathbf{y}^H]\left(E[\mathbf{y}\mathbf{y}^H]\right)^{-1}\mathbf{y}\\ &= \sigma_\theta^2\mathbf{F}^H\mathbf{B}^H\left[\mathbf{B}\mathbf{R}_x\mathbf{B}^H + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\mathbf{y},\end{aligned}$$

where $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^H] = \sigma_\theta^2\mathbf{F}\mathbf{F}^H + \sigma_n^2\mathbf{I}_K$, and the corresponding MSE is

$$\begin{aligned}J &= \text{tr}(E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^H])\\ &= \text{tr}\left(\sigma_\theta^2\mathbf{I}_p - \sigma_\theta^4\mathbf{F}^H\mathbf{B}^H\left[\mathbf{B}\mathbf{R}_x\mathbf{B}^H + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\mathbf{B}\mathbf{F}\right),\end{aligned} \tag{4}$$

The problem is to minimize J in (4) by designing the coding matrices $\mathbf{A}_l$ under a total power constraint. The transmitted power for the lth sensor is defined as $E[\mathbf{x}_l^H\mathbf{A}_l^H\mathbf{A}_l\mathbf{x}_l] = \text{tr}(\mathbf{A}_l\mathbf{R}_{x_l}\mathbf{A}_l^H)$, where $\mathbf{R}_{x_l} = E[\mathbf{x}_l\mathbf{x}_l^H] = \sigma_\theta^2\mathbf{F}_l\mathbf{F}_l^H + \sigma_n^2\mathbf{I}_{k_l}$. If P is the total power that the L sensors can use, then the constraint can be expressed as

$$\sum_{l=1}^{L} \text{tr}\left(\mathbf{H}_l^{-1}\mathbf{B}_l\mathbf{R}_{x_l}\mathbf{B}_l^H\mathbf{H}_l^{-H}\right) \le P, \tag{5}$$

where we use $\mathbf{A}_l = \mathbf{H}_l^{-1}\mathbf{B}_l$. Let $D = \text{tr}\left(\mathbf{F}^H\mathbf{B}^H\left[\mathbf{B}\mathbf{R}_x\mathbf{B}^H + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\mathbf{B}\mathbf{F}\right)$. Thus $J = p\sigma_\theta^2 - \sigma_\theta^4 D$. Since σ_θ and p are fixed, minimization of J subject to (5) can be expressed equivalently as

$$\begin{aligned}&\max_{\mathbf{B}_l,\,1\le l\le L}\ D\\ &\text{subject to} \sum_{l=1}^{L} \text{tr}\left(\mathbf{H}_l^{-1}\mathbf{B}_l\mathbf{R}_{x_l}\mathbf{B}_l^H\mathbf{H}_l^{-H}\right) \le P.\end{aligned} \tag{6}$$

Remark 2: In general, if $E[\boldsymbol{\theta}\boldsymbol{\theta}^H] = \mathbf{R}_\theta$ and $E[\mathbf{n}_l\mathbf{n}_l^H] = \mathbf{R}_{n_l}$ are Hermitian and positive definite, we can write $\mathbf{R}_\theta = \mathbf{R}_\theta^{1/2}\mathbf{R}_\theta^{1/2}$ and $\mathbf{R}_{n_l} = \mathbf{R}_{n_l}^{1/2}\mathbf{R}_{n_l}^{1/2}$, and introduce $\tilde{\boldsymbol{\theta}} = \mathbf{R}_\theta^{-1/2}\boldsymbol{\theta}$, $\tilde{\mathbf{n}}_l = \mathbf{R}_{n_l}^{-1/2}\mathbf{n}_l$, $\tilde{\mathbf{A}}_l = \mathbf{A}_l\mathbf{R}_{n_l}^{1/2}$, and $\tilde{\mathbf{F}}_l = \mathbf{R}_{n_l}^{-1/2}\mathbf{F}_l\mathbf{R}_\theta^{1/2}$. Then $\mathbf{A}_l$, $\mathbf{F}_l$, $\boldsymbol{\theta}$, and $\mathbf{n}_l$ in (3) can be replaced by $\tilde{\mathbf{A}}_l$, $\tilde{\mathbf{F}}_l$, $\tilde{\boldsymbol{\theta}}$, and $\tilde{\mathbf{n}}_l$, respectively, and the equivalent model satisfies assumption i) and ii). Hence, the optimization problem can still be formulated to have the same form as (6).

3. Proposed Approach

In this section, we first consider the objective function of (6) and determine its maximum through singular value decomposition technique. After obtaining the maximum objective function, we consider the power constraint and formulate the problem (6) as a convex optimization problem, which then yields a solution in closed form.

Since $\mathbf{R}_x = \sigma_\theta^2\mathbf{F}\mathbf{F}^H + \sigma_n^2\mathbf{I}_K$ is positive definite, it can be expressed as $\mathbf{R}_x = \mathbf{R}_x^{1/2}\mathbf{R}_x^{1/2}$, where $\mathbf{R}_x^{1/2}$ is Hermitian and positive definite. Since $\text{rank}(\mathbf{F}) = p$ by assumption vi), $\text{rank}(\mathbf{R}_x^{-1/2}\mathbf{F}\mathbf{F}^H\mathbf{R}_x^{-1/2}) = p$ and we have

$$\mathbf{R}_x^{-1/2}\mathbf{F}\mathbf{F}^H\mathbf{R}_x^{-1/2} = \mathbf{U}_C\boldsymbol{\Lambda}_C\mathbf{U}_C^H, \tag{7}$$

where $\boldsymbol{\Lambda}_C = \text{diag}(c_1, \cdots, c_p, 0, \cdots, 0)$ with $c_1 \ge \cdots \ge c_p > 0$ and $\mathbf{U}_C \in \mathbb{C}^{K\times K}$ is unitary. Multiplying (7) by $\mathbf{R}_x^{-1/2}$ on the right and $\mathbf{R}_x^{1/2}$ on the left shows that the nonzero eigenvalues c_i, $1 \le i \le p$, are also the nonzero eigenvalues of $\mathbf{F}\mathbf{F}^H(\sigma_\theta^2\mathbf{F}\mathbf{F}^H + \sigma_n^2\mathbf{I}_K)^{-1}$ and it follows that $c_i \le 1/\sigma_\theta^2$, $1 \le i \le p$. Let $\bar{\mathbf{B}} = \mathbf{B}\mathbf{R}_x^{1/2}$ and use (7), the objective function D in (6) can be expressed as

$$\begin{aligned}D &= \text{tr}(\bar{\mathbf{B}}^H\left[\bar{\mathbf{B}}\bar{\mathbf{B}}^H + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\bar{\mathbf{B}}\mathbf{R}_x^{-1/2}\mathbf{F}\mathbf{F}^H\mathbf{R}_x^{-1/2})\\ &= \text{tr}(\bar{\mathbf{B}}^H\left[\bar{\mathbf{B}}\bar{\mathbf{B}}^H + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\bar{\mathbf{B}}\mathbf{U}_C\boldsymbol{\Lambda}_C\mathbf{U}_C^H).\end{aligned} \tag{8}$$

Express $\bar{\mathbf{B}}$ as a singular value decomposition

$$\bar{\mathbf{B}} = \mathbf{U}_{\bar{B}}\boldsymbol{\Lambda}_{\bar{B}}\mathbf{V}_{\bar{B}}^H, \tag{9}$$

where $\mathbf{U}_{\bar{B}} \in \mathbb{C}^{N\times N}$ is unitary, $\boldsymbol{\Lambda}_{\bar{B}} = \text{diag}(\sqrt{b_1}, \cdots, \sqrt{b_N})$, $b_1 \ge \cdots \ge b_N \ge 0$, and $\mathbf{V}_{\bar{B}} \in \mathbb{C}^{K\times N}$ has orthonormal columns. With (9), D in (8) can be further simplified as

$$D = \text{tr}(\boldsymbol{\Lambda}_{\bar{B}}\left[\boldsymbol{\Lambda}_{\bar{B}}^2 + \sigma_\nu^2\mathbf{I}_N\right]^{-1}\boldsymbol{\Lambda}_{\bar{B}}\mathbf{V}_{\bar{B}}^H\mathbf{U}_C\boldsymbol{\Lambda}_C\mathbf{U}_C^H\mathbf{V}_{\bar{B}}). \tag{10}$$

To find an upper bound on (10), we need the follow fact.

Fact. *[13, p.326] Let* $\mathbf{X}$, $\mathbf{Y} \in \mathbb{C}^{n\times n}$ *be positive semidefinite matrices with eigenvalues* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ *and* $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n \geq 0$ *respectively. Then*

$$\mathrm{tr}(\mathbf{XY}) \leq \sum_{i=1}^{n} \lambda_i \delta_i.$$

Applying the fact to (10) with $\mathbf{X} = \mathbf{\Lambda}_{\bar{B}} \left[\mathbf{\Lambda}_{\bar{B}}\mathbf{\Lambda}_{\bar{B}} + \sigma_\nu^2 \mathbf{I}_N\right]^{-1} \mathbf{\Lambda}_{\bar{B}}$ and $\mathbf{Y} = \mathbf{V}_{\bar{B}}^H \mathbf{U}_C \mathbf{\Lambda}_C \mathbf{U}_C^H \mathbf{V}_{\bar{B}}$, we have

$$D \leq \sum_{i=1}^{p} \frac{c_i b_i}{\sigma_\nu^2 + b_i}. \tag{11}$$

We note that since $\mathbf{X}$ is diagonal, if we choose $\mathbf{V}_{\bar{B}}$ so that $\mathbf{V}_{\bar{B}}^H \mathbf{U}_C = [\mathbf{I}_N \ \mathbf{0}] \in \mathbb{C}^{N\times K}$, where $N \geq p$, then equality in (11) holds. Hence the upper bound in (11) is achieved if we choose

$$\mathbf{V}_{\bar{B}} = \mathbf{U}_C(:, 1:N), \tag{12}$$

where $\mathbf{U}_C(:, 1:N)$ denotes the first N columns of $\mathbf{U}_C$. Since the upper bound is the same for all $N > p$, we choose $N = p$. This keeps the number of transmitters for each sensor at a minimum.

With the choice $\mathbf{V}_{\bar{B}} = \mathbf{U}_C(:, 1:p)$, we obtain

$$D = \sum_{i=1}^{p} \frac{c_i b_i}{\sigma_\nu^2 + b_i} \tag{13}$$

where $c_i > 0$ and σ_ν^2 are fixed. Hence, the problem of choosing $\mathbf{B}_l$ in $\mathbf{B}$ to maximize D amounts to choosing the singular values $\sqrt{b_i}$ of $\bar{\mathbf{B}}$ in (9) since the choice of $\mathbf{U}_{\bar{B}}$ is irrelevant and $\mathbf{B} = \bar{\mathbf{B}}\mathbf{R}_x^{-1/2}$. It is clear from (13) that to maximize D, we should choose each $b_i > 0$ as large as possible; however, it can not be chosen arbitrarily large due to the total power constraint.

By taking into account the power constraint in (6), we set $\mathbf{U}_{\bar{B}} = \mathbf{I}_N$ to simplify our analysis and thus we have $\mathbf{B} = \Lambda_{\bar{B}}\mathbf{V}_{\bar{B}}\mathbf{R}_x^{-1/2}$, or equivalently,

$$\mathbf{B}_l = \mathbf{\Lambda}_{\bar{B}}\hat{\mathbf{U}}_{C,l}, \quad 1 \leq l \leq L, \tag{14}$$

where $\hat{\mathbf{U}}_{C,l} \in \mathbb{C}^{p\times k_l}$ is the lth block matrix in $\mathbf{V}_{\bar{B}}^H \mathbf{R}_x^{-1/2} = [\hat{\mathbf{U}}_{C,1}, \cdots, \hat{\mathbf{U}}_{C,L}]$ with $\mathbf{V}_{\bar{B}} = \mathbf{U}_C(:, 1:p)$. Substituting (14) into (6), since $\mathbf{H}_l$ are diagonal matrices from assumption vii), the power constraint can be further simplified as

$$\mathrm{tr}\left(\bar{\mathbf{R}}\mathbf{\Lambda}_{\bar{B}}^2\right) = \sum_{i=1}^{p} r_i b_i \leq P, \tag{15}$$

where $\bar{\mathbf{R}} = \sum_{l=1}^{L} \mathbf{H}_l^{-1}\hat{\mathbf{U}}_{C,l}\mathbf{R}_{x_l}\hat{\mathbf{U}}_{C,l}^H\mathbf{H}_l^{-H}$ with diagonal entries r_i, $1 \leq i \leq p$. Note that since $\mathbf{R}_{x_l} = \sigma_\theta^2 \mathbf{F}_l\mathbf{F}_l^H + \sigma_n^2 \mathbf{I}_{k_l}$, we see that the diagonal entries of $\mathbf{H}_l^{-1}\hat{\mathbf{U}}_{C,l}\mathbf{R}_{x_l}\hat{\mathbf{U}}_{C,l}^H\mathbf{H}_l^{-H}$ are positive and thus $r_i > 0$. From (13) and (15), the problem in (6) with $\mathbf{V}_{\bar{B}} = \mathbf{U}_C(:, 1:p)$ and $\mathbf{U}_{\bar{B}} = \mathbf{I}_N$ can be written as

$$\min_{b_i, 1\leq i\leq p} -\sum_{i=1}^{p} \frac{c_i b_i}{\sigma_\nu^2 + b_i}$$

$$\text{subject to} \sum_{i=1}^{p} r_i b_i \leq P, \tag{16}$$

$$b_i \geq 0.\ i = 1, \cdots, p.$$

This is a convex optimization problem since the cost function is convex and the constraints are linear.

To solve the problem (16), we form the Lagrangian as

$$L(b_i, \mu_0, \mu_i) = -\sum_{i=1}^{p} \frac{c_i b_i}{\sigma_\nu^2 + b_i} + \mu_0\left(\sum_{i=1}^{p} r_i b_i - P\right) - \sum_{i=1}^{p} \mu_i b_i,$$

where $\mu_0 \geq 0$ and $\mu_i \geq 0$, and the associated KKT conditions [14] are

$$-\frac{c_i\sigma_\nu^2}{(\sigma_\nu^2 + b_i)^2} + \mu_0 r_i - \mu_i = 0 \tag{17}$$

$$\mu_0\left(\sum_{i=1}^{p} r_i b_i - P\right) = 0 \tag{18}$$

$$\mu_i b_i = 0 \tag{19}$$

From (17), we obtain

$$b_i = \sigma_\nu^2\left(\sqrt{\frac{c_i}{\sigma_\nu^2(\mu_0 r_i + \mu_i)}} - 1\right). \tag{20}$$

From (19), if $b_i > 0$, we have $\mu_i = 0$ and thus (20) can be written as

$$b_i = \sigma_\nu^2\left(\sqrt{\frac{c_i}{\sigma_\nu^2\mu_0 r_i}} - 1\right)^+ \tag{21}$$

where $(x)^+ = \max(0, x)$. From (17) with $b_i > 0$ and $\mu_i = 0$, we have $\mu_0 > 0$ otherwise we have a contradiction $\mu_0 < 0$. Hence, from (18), we get

$$\sum_{i=1}^{p} r_i b_i = P. \tag{22}$$

Let $w_{m_i} = \sqrt{c_{m_i}/r_{m_i}}$ and assume $w_{m_1} \geq \cdots \geq w_{m_p}$, where $m_i \in \{1, \cdots, p\}$. We define a function

$$f(m_n) = w_{m_n} \times \frac{\frac{P}{\sigma_\nu^2} + \sum_{i=m_1}^{m_n} r_i}{\sum_{i=m_1}^{m_n} \sqrt{r_i c_i/\sigma_\nu^2}}.$$

Let $1 \leq p_1 \leq p$ be such that $f(m_{p_1}) > 1$ and $f(m_{p_1+1}) \leq 1$. Then we have

$$b_{m_i} = \begin{cases} \sqrt{\frac{\sigma_\nu^2}{\mu_0}} w_{m_i} - \sigma_\nu^2, & i \leq p_1 \\ 0, & i > p_1 \end{cases} \tag{23}$$

where

$$\mu_0 = \left(\frac{\sum_{i=m_1}^{m_{p_1}} \sqrt{r_i c_i/\sigma_\nu^2}}{\frac{P}{\sigma_\nu^2} + \sum_{i=m_1}^{m_{p_1}} r_i}\right)^2$$

is obtained by substituting (23) into (22).

With the choice $\mathbf{\Lambda}_{\bar{B}} = \mathrm{diag}(\sqrt{b_1}, \cdots, \sqrt{b_p})$ according to (23), the coding matrix can be written from (14) and $\mathbf{A}_l = \mathbf{H}_l^{-1}\mathbf{B}_l$, and is given by

$$\mathbf{A}_l = \mathbf{H}_l^{-1}\mathbf{\Lambda}_{\bar{B}}\hat{\mathbf{U}}_{C,l}, \quad 1 \le l \le L; \tag{24}$$

moreover, from (4) and (13), the corresponding MSE can be written as

$$J = p\sigma_\theta^2 - \sigma_\theta^4 \sum_{i=1}^{p} \frac{c_i b_i}{\sigma_\nu^2 + b_i}. \tag{25}$$

As $P \to \infty$, from (15), we have $b_i \to \infty$ and thus the lower bound of the MSE is

$$J_{low} = p\,\sigma_\theta^2 - \sigma_\theta^4 \sum_{i=1}^{p} c_i, \tag{26}$$

Note that since $c_i \le 1/\sigma_\theta^2$, we have $J_{low} \ge 0$.

4. Numerical Results

In this section, we use numerical simulations to illustrate the analytical results established in the previous section. In all simulations, the random vectors $\boldsymbol{\theta}$, $\mathbf{n}_l$, and $\boldsymbol{\nu}$, are complex Gaussian. Specifically, each entry in the vectors is set as a complex Gaussian random variable with zero mean and unit variance. We also take the entries of $\mathbf{F}_l$ and the channels h_i^l as complex Gaussian random variables with zero mean and unit variance. The number of source signals is set as $p = 5$.

The average MSE versus N with different power levels $P = 5$ dB, 10 dB, and 20 dB is shown in Fig. 2, in which we take $L = 10$ and $k_l = 8$, $\forall l$. The dash-line denotes the MSE lower bound in (26). We can see that the MSE decreases as N increases and remains a constant as $N \ge p$. That is, there is no improvement in network performance as the number of transmitters for each sensor is greater than the number of source signals. The result is consistent with our analysis in Sect. 3 that under a total power constraint, the minimum achieved MSE is (25) for all $N > p$. From the figure we can also see that the MSE decreases as the power level increases as is expected.

Since the performance is the same for all $N > p$, we set $N = p$ in all the simulations to follow. To see the effect of the number of measurements, we consider $L = 10$ and assume that k_l are equal for all sensors. Figure 3 shows the average MSE versus k_l with different power levels $P = 5$ dB, 10 dB, and 20 dB. We note that as k_l increases, the MSEs decreases; moreover, the gap between the lower bound and the power constraint case becomes large as k_l increases. This is because although the increase of k_l leads to the increase of measurement power, the transmitted power for each sensor is restricted if there is a power constraint. Hence, the performance of the power constraint case has no significant improvement as k_l increases compared with that of the unconstraint power case (the lower bound).

The comparison of the MSE for the proposed method and the equal power method is plotted in Fig. 4, in which we take $L = 10$ and $k_l = 8$, $\forall l$. The equal power method is to set the transmitted powers for all sensors to be equal, that is, we choose the coding matrix $\mathbf{A}_l = \alpha_l \cdot [\mathbf{I}_5\ \mathbf{0}]$, where

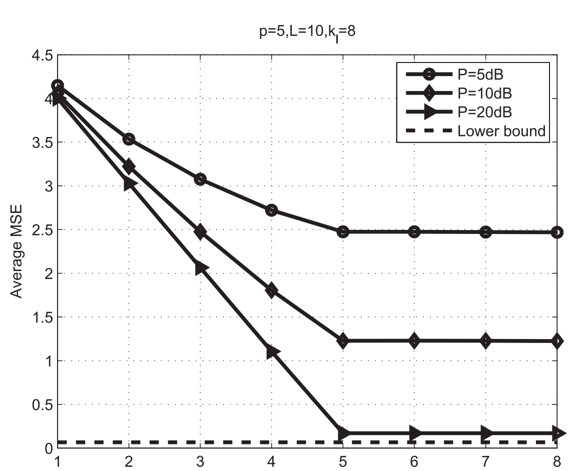

Fig. 2 MSE versus N with different power levels.

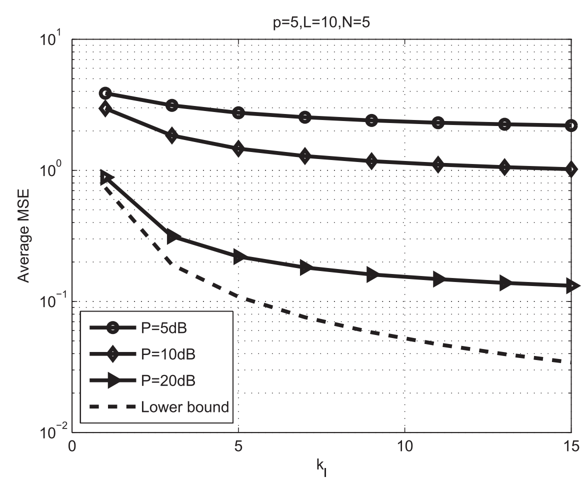

Fig. 3 MSE versus k_l with different power levels.

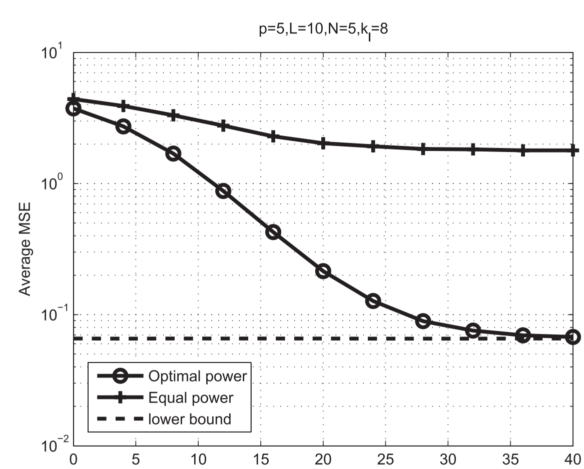

Fig. 4 MSE comparison between the proposed method and the equal power method.

$\alpha_l = \sqrt{P/[L\mathrm{tr}(\mathbf{R}_{x_l}(1:5,1:5))]}$, so that $\mathrm{tr}(\mathbf{A}_l\mathbf{R}_{x_l}\mathbf{A}_l^H) = P/L$, $1 \le l \le L$, here $\mathbf{R}_{x_l}(1:5,1:5)$ denotes the first 5 rows and columns of $\mathbf{R}_{x_l}$ and $[\mathbf{I}_5\ \mathbf{0}]$ is a 5×8 matrix with its diagonal entries equal to 1 and other entries equal to 0. As we can see from the figure, the proposed method performs better than the equal power method. Moreover, the MSE of the proposed method goes to the lower bound as P increases while the MSE of the equal power method approaches a constant MSE for $P > 25$ dB. This is because the proposed method takes into account the effects of the observation and channel matrices in the design of coding matrices, while the equal power method does not use the information of observation matrices and channel matrices and thus the performance improvement is limited as P increases.

To estimate a vector signal, [5] proposed an amplify-and-forward scheme based on the SISO sensor network. For comparison, we simulate the scheme in [5], where a scalar measurement x_l at the lth sensor is multiplied by an amplified factor $a_l(n)$ at the nth time instant before it is transmitted to the receiver through a fading channel h_l, $1 \le l \le Q$, here Q denotes the total number of sensors in the SISO network. At time n, the received signal is $y(n) = \sum_{l=1}^{Q} h_l a_l(n) x_l + \nu_l(n)$, where $\nu_l(n)$ is an additive noise. After collecting p received signals $\{y(1), \cdots, y(p)\}$ at the FC, the source vector $\boldsymbol{\theta}$ is then estimated based on the LMMSE fusion rule. The proposed MIMO scheme is modelled in (3) with $L = 1$, a single vector sensor. To compare these two schemes based on the equal condition, we assume the number of measurements in the proposed scheme is equal to the scheme in [5], that is, $k_l = Q$; moreover, the total network powers, consumed by the sensors, are set equal for both schemes. Figure 5 shows that the average MSE of the proposed scheme with $k_l = 10$ is lower than that of the scheme in [5] with $Q = 10$ when power P is small. This is because in our MIMO sensor network, the antenna diversity are used and thus each received signal comes from individual channel; while in [5], the received signal is a linear combination of transmitted signals without using antenna diversity. From the figure, we can also see that as the power increases, the performance of two schemes are very close. Moreover, for the proposed scheme with $k_l = 15$ and the scheme in [5] with $Q = 15$, the MSEs are almost the same for over all P. In this simulation, we see that as total power less than 20 dB and the number of measurements less than 15, the proposed method has a noticeable improvement compared with the method in [5].

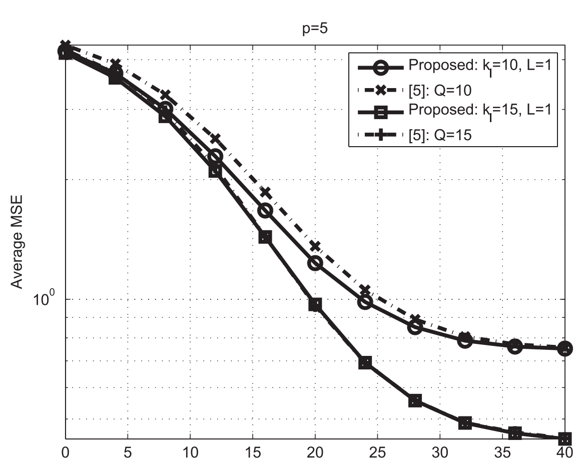

Fig. 5 MSE comparison between the proposed scheme and that in [5].

5. Conclusion

We study distributed estimation of a random vector signal in power-constrained MIMO sensor networks. We find an upper bound of the objective function and show that the minimum number of transmitters to reach the upper bound is equal to the number of the source signals. By choosing specific singular vectors, we formulate the original problem as a convex optimization problem. The solution then yields closed form expressions for the coding matrices. The proposed MIMO scheme can be viewed as an extension of the SISO scheme proposed in [5]. From simulation results, we see that the improvement in performance of the proposed MIMO scheme over the SISO scheme in [5] is more significant when the transmitted power is not too high and the number of measurements is not too large, while in other cases, they have almost identical performance.

Acknowledgement

We thank the reviewers for insightful comments that improve the paper.

References

[1] J.J. Xiao, A. Ribeiro, Z.Q. Luo, and G.B. Giannakis, “Distributed compression-estimation using wireless sensor networks,” IEEE Signal Process. Mag., vol.23, no.4, pp.27–41, July 2006.
[2] S. Cui, J.J. Xiao, A.J. Goldsmith, Z.Q. Luo, and H.V. Poor, “Estimation diversity and energy efficiency in distributed sensing,” IEEE Trans. Signal Process., vol.55, no.9, pp.4683–4695, Sept. 2007.
[3] I. Bahceci and A.J. Khandani “Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation,” IEEE Trans. Commun., vol.56, no.7, pp.1146–1156, July 2008.
[4] H. Şenol and C. Tepedelenlioğlu, “Performance of distributed estimation over unknown parallel fading channels,” IEEE Trans. Signal Process., vol.56, no.12, pp.6057–6068, Dec. 2008.
[5] W. Guo, J.-J. Xiao, and S. Cui, “An efficient water-filling solution for linear coherent joint estimation,” IEEE Trans. Signal Process., vol.56, no.10, pp.5301–5305, Oct. 2008.
[6] C.H. Wu and C.A. Lin, “Linear coherent distributed estimation over unknown channels,” Signal Process., vol.91, no.4, pp.1000–1011, April 2011.
[7] E.G. Larsson and P. Stoica, Space-Time Block Coding for Wireless Communication, Cambridge Univ. Press, Cambridge, U.K., 2003.
[8] Y. Zhu, E. Song, J. Zhou, and Z. You, “Optimal dimensionality reduction of sensor data in multisensor estimation fusion,” IEEE Trans. Signal Process., vol.53, no.5, pp.1631–1639, May 2005.
[9] I.D. Schizas, G.B. Giannakis, and Z.-Q. Luo, “Distributed estimation using reduced-dimensionality sensor observations,” IEEE Trans. Signal Process., vol.55, no.8, pp.4284–4299, Aug. 2007.
[10] J. Fang and H. Li, “Optimal/near-optimal dimensionality reduction for distributed estimation in homogeneous and certain inhomogeneous scenarios,” IEEE Trans. Signal Process., vol.58, no.8,

pp.4339–4353, Aug. 2010.
[11] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A.J. Goldsmith, "Linear coherent decentralized estimation," IEEE Trans. Signal Process., vol.56, no.2, pp.757–770, Feb. 2008.
[12] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall PTR, 1993.
[13] D.S. Bernstein, Matrix Mathematics: Theory, Facts, and Formulas with Allpication to Linear Systems Theory, Princeton University Press, 2005.
[14] S. Boyd and L. Vanderberghe, Convex Optimization, Cambridge Univ. Press, Cambridge, U.K., 2003.

Chien-Hsien Wu received the B.S. degree in electrical engineering from the National Taiwan Ocean University, Keelong, Taiwan, in 2001, and the M.S. degree in electrical and control engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 2003. He is currently a Ph.D. student with the Department of Electrical and Control Engineering, National Chiao Tung University. His current research interests are in signal processing.

Ching-An Lin received the B.S. degree from the National Chiao Tung University, Hsinchu, Taiwan, in 1977, the M.S. degree from the University of New Mexico, Albuquerque, in 1980, and the Ph.D. degree from the University of California, Berkeley, in 1984, all in electrical engineering. He was with the Chung Shan Institute of Science and Technology from 1977 to 1979 and with Integrated Systems Inc. from 1984 to 1986. Since June, 1986, he has been with the Department of Electrical and Control Engineering, National Chiao Tung University, where he is a Professor. His current research interests are in control and signal processing.