

Research Article

Adaptive Power Allocation in Wireless Sensor Networks with Spatially Correlated Data and Analog Modulation: Perfect and Imperfect CSI

Muhammad Hafeez Chaudhary and Luc Vandendorpe

ICTEAM Institute, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium

Correspondence should be addressed to Muhammad Hafeez Chaudhary, muhammad.chaudhary@uclouvain.be

Received 6 February 2010; Accepted 6 July 2010

Academic Editor: Carles Anton-Haro

Copyright © 2010 M. H. Chaudhary and L. Vandendorpe. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We address the problem of power allocation in a wireless sensor network where distributed sensors amplify and forward their partial and noisy observations of a Gaussian random source to a remote fusion center (FC). The FC reconstructs the source based on linear minimum mean-squared error (LMMSE) estimation rule. Motivated by the availability of limited energy in the sensor networks, we undertake the design of power allocation based on minimization of the reconstruction distortion subject to a constraint on the network transmit power. The design is based on the following two cases: (i) exact knowledge of the channel gains and (ii) the estimates of the channel gains. We show that the distortion can be represented as a convex function of the transmit powers of the sensors. Moreover, we show that the power allocation based on this distortion function does not bear any closed form solution. To this end, we propose a novel design based on the successive approximation of the LMMSE distortion, which turns out to be simple, computationally efficient, and exhibits excellent convergence properties. The simulation examples illustrate that the proposed design holds considerable performance gain compared to a uniform power allocation scheme.

1. Introduction

Wireless sensor networking is an emerging technology which finds application in many fields including environment and habitat monitoring, health care, automation, military applications such as battlefield monitoring and surveillance, and underwater wireless sensor networks (UWSNs) for marine environment monitoring [1, 2]. A wireless sensor network (WSN) consists of spatially distributed sensors that cooperatively monitor physical or environmental conditions, for example, temperature, vibration, pressure, motion, or pollutants.

We consider a system in star topology where sensors amplify and transmit their noisy observations of a common source, via some orthogonal multiple access scheme such as frequency division multiple-access (FDMA), to a central processing unit called fusion center (FC) which reconstructs the source in a way that the overall distortion (e.g., mean squared error) be minimized. Conceptually, the system is similar to the CEO problem [3, 4].

The sensors in the network have partial and spatially correlated observations of the underlying source. The correlation exists where sensors measure data in same geographical location, for example, acoustic sensors that are sensing a common event produce measurements that are correlated. In addition, observation noise and communication channel may not have same conditions across the sensors. Therefore, transmission of the observations based on uniform power allocation is not an optimal strategy.

In this paper, we study the problem of adaptive power allocation given a network power constraint with the objective to minimize the reconstruction MSE. The optimal power allocations are jointly determined at the FC which are then conveyed to the individual sensors via feedback channels. The communication channels from the sensors to the FC experience independent flat fading. The channel from the sensor to the FC is usually estimated using some training sequence. The receiver noise and the limited available power means that the channel estimation always incurs some estimation error. Consequently, the design of

power allocation scheme should also take into account the channel estimation errors [5, 6]. In this paper, first we design the power allocation scheme based on perfect knowledge of the channel state information (CSI) and subsequently, in the design, we incorporate the effect of imperfect CSI.

In a sensor network measuring a memoryless Gaussian source uncoded transmission, that is, amplify and forward (AF), outperforms the separate coding and transmission over the multiple-access channel [7–9]. Motivated by this result, Vuran et al. in [10] considered the estimation of a random source with distributed sensors and suggested a sensor selection procedure which exploits the spatial correlation to minimize the estimation error (based on the LMMSE estimation criterion). The sensor selection procedure suggests that the sensors with high correlation with the source and low cross-correlations should be selected. The procedure does not take into account the fact that even if a sensor has high correlation with the source and low cross-correlations with the other sensors, it can still be a bad selection in terms of energy efficiency if its observation noise is high and/or the communication channel to the FC is bad. A recent related work appears in [11]. Bahceci and Khandani in [12] proposed a power allocation scheme where each sensor observes a separate source albeit correlated. Reference [13] presented a power scheduling scheme for sensor networks to detect a source based on the binary hypothesis testing rule which exploits the correlation in the observation noises at the sensors. Other works like [14–18] proposed power allocation schemes for parameter estimation in wireless sensor networks without considering the spatial correlation. In this paper, we present a novel framework which incorporates adaptive power allocation (APA) in the network by taking into account the spatial correlation and cross-correlations of the observations, observation quality, and communication channel to the FC. The power allocation design also takes into account the channel estimation errors.

We assume that the FC reconstructs the underlying source using linear minimum mean squared error (LMMSE) estimation rule. The power allocation design is based on minimization of the reconstruction distortion subject to a constraint on the total transmit power of the sensors. Due to the spatial correlation among the sensor observations, the design of the power allocation scheme based on the given optimization problem presents a unique challenge because the LMMSE estimation/reconstruction error of the underlying source contains nonlinearly coupled optimization variables. Herein, first we prove that the estimation distortion can be represented as a convex function of the sensor transmit powers, then we show that the power allocation design based on this distortion function turns out to be complicated and does not bear a closed solution. Subsequently, we propose a novel design based on the successive approximation of the LMMSE estimation distortion. The resulting power allocation algorithm is simple, computationally efficient, and exhibits excellent convergence properties. The proposed designs hold considerable performance gain compared to a uniform power allocation scheme. To the best of our knowledge, in the present literature, there is no such work on the design of power allocation for the

sensor network under consideration which jointly exploits spatial correlation, observation noises, channel gains, and their estimation errors.

The rest of the paper is organized as follows. Section 2 describes the system set-up. The power allocation problems and their solutions are presented in Sections 3 and 4, respectively for perfect and imperfect knowledge of the CSI. Section 5 evaluates performance of the power allocation designs. Section 6 concludes the work.

2. System Model

Consider the system model shown in Figure 1 in which N spatially distributed sensors observe an unknown zero-mean real Gaussian random source $s \sim \mathcal{N}(0, \sigma_s^2)$, and communicate with the fusion center (FC) via orthogonal multiple-access channels. Each sensor has a partial and noisy observation of the source, and sends an amplified version of it to the FC. The FC collects the signals from all sensors and reconstructs the source according to a given fidelity criterion, for example, minimum mean-squared estimation error. The $s_i \sim \mathcal{N}(0, \sigma_{s_i}^2)$, and $n_i \sim \mathcal{N}(0, \sigma_{n_i}^2)$, respectively, denote the partial observation of the source s and the noise corrupting this observation such that the noisy observation at sensor i is

$$x_i(t) = s_i(t) + n_i(t), \quad i = 1, \dots, N. \quad (1)$$

The estimation of the source is done on a sample by sample basis, and its procedure is same for all samples. Therefore, for clarity, in the subsequent formulation we drop the time index. We assume that the sensors amplify and forward their observations to the FC via orthogonal channels where each channel experiences flat fading independent over time and across sensors.

The optimality of the AF scheme is established for the Gaussian network with nonorthogonal multiple-access channel from the sensors to the FC [7]. However, for the network with orthogonal multiple-access channel it has been shown in [19, 20] that the separate source channel coding outperforms the AF scheme. The optimality of the coded source-channel communication in general requires coding over long block lengths and will require some data processing at the sensors. This will increase the power consumption at the sensors and will lead to longer processing delays which may not be tolerable in many applications. Therefore, due to simplicity, low latency, and ease of implementation, in this paper we adopt the AF transmission strategy.

The received signal at FC from sensor i is

$$\bar{z}_i = h_i \sqrt{P'_i} x_i + \bar{w}_i = h_i \sqrt{P'_i} (s_i + n_i) + \bar{w}_i, \quad \forall i, \quad (2)$$

where $\sqrt{P'_i}$ is a amplifying factor and $\bar{w}_i \sim \mathcal{CN}(0, \sigma_{\bar{w}_i}^2)$ is a circularly-symmetric Gaussian receiver noise. The fading channels $\{h_i\}_{i=1}^N$ between the sensors and the FC are $h_i \sim \mathcal{CN}(0, \sigma_{h_i}^2)$, for all i with gain factors $\{g_i = |h_i|\}_{i=1}^N$ which are Rayleigh distributed. Noting that $h_i = g_i e^{j\theta_{h_i}}$, we can write (2) as

$$\bar{z}_i e^{-j\theta_{h_i}} = g_i \sqrt{P'_i} (s_i + n_i) + \bar{w}_i e^{-j\theta_{h_i}}, \quad (3)$$

where the exponential term $e^{-j\theta_{n_i}}$ can be merged into the variable \bar{w}_i without changing its statistical properties—due to the circular-symmetry property of \bar{w}_i [21]. Since the underlying source s and the noisy observation $x_i = s_i + n_i$ are real-valued, therefore, we only need to consider the component of the noise \bar{w}_i which is in-phase with the observation x_i , that is,

$$z_i = g_i \sqrt{P'_i} (s_i + n_i) + w_i, \quad \forall i, \quad (4)$$

where $w_i \sim \mathcal{N}(0, \sigma_{w_i}^2)$ and $\sigma_{w_i}^2 = 0.5\sigma_{\bar{w}_i}^2$.

For the analysis in this work, we assume that the observation noise n_i , $\forall i$ (similarly the receiver noise w_i , $\forall i$) is independent across the sensors and is also independent of w_i , $\forall i$ (n_i , $\forall i$). Moreover, we assume that the source s , the observation s_i at sensor i , the observation s_j at sensor j , the observation noise n_i at the sensor, and the receiver noise w_i at the FC are jointly Gaussian across sensors ($\forall i$ and $\forall j$) with zero mean and covariance ($\Lambda_{s, s_i, s_j, n_i, w_i}$) specified by

$$\Lambda = \begin{pmatrix} \sigma_s^2 & \sigma_s \sigma_{s_i} \rho_{si} & \sigma_s \sigma_{s_j} \rho_{sj} & 0 & 0 \\ \sigma_s \sigma_{s_i} \rho_{si} & \sigma_{s_i}^2 & \sigma_{s_i} \sigma_{s_j} \rho_{ij} & 0 & 0 \\ \sigma_s \sigma_{s_j} \rho_{sj} & \sigma_{s_i} \sigma_{s_j} \rho_{ij} & \sigma_{s_j}^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{n_i}^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{w_i}^2 \end{pmatrix}. \quad (5)$$

We also assume that the samples of s , s_i , n_i , and w_i are individually independent in time.

In (5), the correlation coefficient ρ_{si} represents the correlation between s and s_i and the coefficient ρ_{ij} denotes the correlation between s_i and s_j . The values of these correlation coefficients depend on the distance of the sensors w.r.t. the position of the source s and w.r.t. each other, respectively, and can be characterized as follows:

$$\rho_{si} = \frac{\text{Cov}\{S, S_i\}}{\sigma_s \sigma_{s_i}} = e^{-(d_{si}/\theta_1)^{\theta_2}},$$

$$\rho_{ij} = \frac{\text{Cov}\{S_i, S_j\}}{\sigma_{s_i} \sigma_{s_j}} = e^{-(d_{ij}/\theta_1)^{\theta_2}}, \quad (6)$$

which is a power exponential model for correlation [10, 22]. In (6), d_{si} is the distance between the source s and sensor i , and d_{ij} is the distance between the sensors i and j . The parameter $\theta_1 > 0$ controls how fast the correlation decays with distance and is called range parameter. The other parameter θ_2 is called a smoothness or roughness parameter which is $0 < \theta_2 \leq 2$. Equation (6) shows that the correlation decays with distance with limiting values of 1 and 0 as $d_{si}(d_{ij}) \rightarrow 0$ and $d_{si}(d_{ij}) \rightarrow \infty$, respectively. Therefore, the correlation changes with the change in the relative positions of the source and the sensors. The change may happen due to movement of either the sensors or the source or both, for example, an animal may kick a sensor node to a different location. We assume that the relative positions of the sensors with respect to each other and the underlying source are perfectly known. Moreover, we assume that the positions remain unchanged for at least one estimation cycle.

Based on the correlation model, therefore, we can say that the FC in essence is interested to reconstruct the source s which is located at a specific location by collecting observations from spatially distributed sensors where correlation of the observations with the source and among the sensors, respectively, depends on the spatial location of the sensors w.r.t. the source and w.r.t. each other. Note that the location of the source and the sensors can be in two or three dimensional space.

The equation (4) can be written equivalently in matrix-vector notation as follows:

$$\mathbf{z} = \mathbf{H}(\mathbf{s}_i + \mathbf{n}) + \mathbf{w}, \quad \text{where}$$

$$\mathbf{H} = \text{diag}\left(g_1 \sqrt{P'_1}, \dots, g_N \sqrt{P'_N}\right), \quad (7)$$

$$\mathbf{z} = [z_1, \dots, z_N]^T, \quad \mathbf{s}_i = [s_1, \dots, s_N]^T,$$

$$\mathbf{n} = [n_1, \dots, n_N]^T, \quad \mathbf{w} = [w_1, \dots, w_N]^T,$$

where $[\dots]^T$ denotes the matrix-vector transpose operation. At the FC, the optimal estimator in minimum mean-squared error (MMSE) sense is the conditional mean of s given the observation \mathbf{z} , that is, $\hat{s} = \mathbb{E}[s | \mathbf{z}]$, where \mathbb{E} denotes the mathematical expectation. Under the jointly Gaussian assumption of the s and \mathbf{z} , the conditional mean estimator turns out to be linear and is called linear minimum mean-squared error estimator (LMMSEE). Therefore, we seek the estimate of the source like $\hat{s} = \sum_{i=1}^N a_i z_i = \mathbf{a}^T \mathbf{z}$ which can be written as [23]

$$\hat{s} = \mathbf{c}_{sz} \mathbf{C}_z^{-1} \mathbf{z}, \quad (8)$$

where $\mathbf{a}^T = [a_1, \dots, a_N] = \mathbf{c}_{sz} \mathbf{C}_z^{-1}$ is a row-vector of LMMSEE weighting coefficients. The resultant distortion of the estimate \hat{s} in comparison to the original signal s is measured by the mean-squared error and is given by [23]

$$D = \mathbb{E}_{\{s, s_i, n_i, w_i, g_i, \forall i\}} \left[(s - \hat{s})^2 \right] = C_s - \mathbf{c}_{sz} \mathbf{C}_z^{-1} \mathbf{c}_{zs}$$

$$= \sigma_s^2 - \mathbf{c}^T \mathbf{H} (\mathbf{H} \mathbf{C} \mathbf{H} + \mathbf{C}_w)^{-1} \mathbf{H} \mathbf{c}, \quad (9)$$

where $\mathbf{C} = \mathbf{C}_{s_i} + \mathbf{C}_n$, $\mathbf{c} = \mathbb{E}[s_i s]$, $\mathbf{C}_{s_i} = \mathbb{E}[s_i s_i^T]$, $\mathbf{C}_n = \mathbb{E}[\mathbf{n} \mathbf{n}^T]$, and $\mathbf{C}_w = \mathbb{E}[\mathbf{w} \mathbf{w}^T]$. The estimation distortion D can also be written as follows:

$$D = \sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \mathbf{c}^T (\mathbf{H} \mathbf{C}_w^{-1} \mathbf{H} \mathbf{C} + \mathbf{I})^{-1} \mathbf{C}^{-1} \mathbf{c}, \quad (10)$$

which is obtained by using the Woodbury identity for matrix inversion [24] (see Appendix A). Let $\mathbf{Y} = \mathbf{H} \mathbf{C}_w^{-1} \mathbf{H}$, where $\mathbf{H} \mathbf{C}_w^{-1} \mathbf{H} = \text{diag}(g_1^2 P'_1 / \sigma_{w_1}^2, \dots, g_N^2 P'_N / \sigma_{w_N}^2)$. Now we can write (10) as

$$D = \sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \mathbf{c}^T (\mathbf{Y} \mathbf{C} + \mathbf{I})^{-1} \mathbf{C}^{-1} \mathbf{c}. \quad (11)$$

Theorem 1. *The estimation distortion function in (11) is convex over P'_i , for $i = 1, \dots, N$.*

Let $\mathbf{p}' = [P'_1, \dots, P'_N]^T$ be a vector of the transmit power of the sensors. The proof of the theorem consists in showing that the Hessian of the distortion function in (11) with respect to \mathbf{p}' is positive semidefinite. To this end, a detailed proof is given in Appendix B.

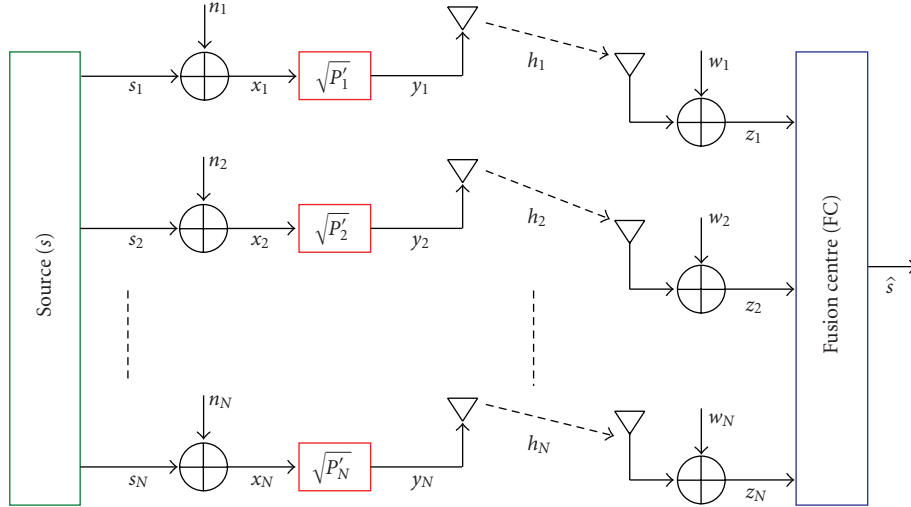


FIGURE 1: Block diagram of the system.

Remark 1. The reconstruction distortion is upper bounded by the variance of the source σ_s^2 , and lower bounded by the variances of the observation noises and spatial correlation and cross-correlation values as given by

$$\sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} \leq D \leq \sigma_s^2. \quad (12)$$

The lower-bound distortion is achieved when the observations of the sensors are received at the FC via ideal communication channels which can be verified by setting $g_i^2 P_i' / \sigma_{w_i}^2 = \infty$, for all i in (11). Note that it is not possible to achieve distortion less than $D_0 = \sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c}$. Moreover, when observation noise variances are $\{\sigma_{n_i}^2\}_{i=1}^N = 0$ then the lower bound distortion reduces to $D_0 = \sigma_s^2 - \mathbf{c}^T \mathbf{C}_s^{-1} \mathbf{c}$. The achieved distortion is equal to the upper bound value (i.e., $D = \sigma_s^2$) when either no signal is received at the FC from the sensors or the observations are uncorrelated with the source s or both, which can be verified from (11).

3. Power Allocation with Perfect CSI

In this section, we assume that the channel state information (CSI), that is, the channel gains $\{h_i\}_{i=1}^N$ are perfectly known at the FC. The case of imperfect CSI is considered in the next section.

3.1. Minimization of the Distortion Subject to the Power Constraint. We base our adaptive power allocation design on the following optimization problem.

Prob. Minimize the distortion $D(P'_1, \dots, P'_N)$ subject to

$$\sum_{i=1}^N P_i = \sum_{i=1}^N P_i' \sigma_i^2 \leq P_{\text{tot}}, \quad (13)$$

$$P_i' \geq 0, \quad \forall i,$$

where $P_i = P_i'(\sigma_{s_i}^2 + \sigma_{n_i}^2) = P_i' \sigma_i^2$ denotes the total power in the transmitted signal of sensor i . The sum power constraint

in (13) enables a fair comparison between the networks of different sizes. Moreover, for a sensor network which forms part of a bigger network where each subnetwork performs different sensing task but share the same frequency band to transmit observations, to limit the interference between the subnetworks, the total power emitted from each subnetwork is upper bounded. Furthermore, recent studies have shown that the ICT (Information and Communication Technology) power consumption is a significant contributor to the global warming [25]. Therefore, in the context of sensor networks, putting cap on the total power consumption conserves energy and limits the contribution to the global warming.

Since the optimization problem in (13) is convex (the objective is convex and the constraints are linear), therefore we can use the Lagrangian method of multipliers to find the optimal P_i' 's [26]. The Lagrangian cost function is

$$f(P_i', \lambda) = D + \lambda \left(\sum_{i=1}^N P_i' \sigma_i^2 - P_{\text{tot}} \right) - \sum_{i=1}^N \mu_i P_i', \quad (14)$$

where λ and $\{\mu_i\}_{i=1}^N$ are dual variables or Lagrange multipliers. The associated Karush-Kuhn-Tucker (KKT) conditions are

$$\frac{\partial f}{\partial P_i'} = -\frac{g_i^2}{\sigma_{w_i}^2} \mathbf{c}^T (\mathbf{Y}\mathbf{C} + \mathbf{I})^{-1} \mathbf{J}_i \mathbf{C} (\mathbf{Y}\mathbf{C} + \mathbf{I})^{-1} \mathbf{C}^{-1} \mathbf{c} + \lambda \sigma_i^2 - \mu_i = 0, \quad (15)$$

$$\lambda \left(\sum_{i=1}^N P_i' \sigma_i^2 - P_{\text{tot}} \right) = 0, \quad \lambda \geq 0, \quad \sum_{i=1}^N P_i' \sigma_i^2 \leq P_{\text{tot}}, \quad (16)$$

$$\mu_i P_i' = 0, \quad \mu_i \geq 0, \quad P_i' \geq 0, \quad \forall i, \quad (17)$$

where \mathbf{J}_i is a diagonal matrix with unity at (i, i) th place and all other elements are equal to zero.

The expression in (15) is a complicated function of the optimization variables. Therefore, a closed form solution for this problem is not tractable. However, we can resort

1: Initialize $\lambda^{[0]}$ and $P_i^{[0]}$ for $i = 1, \dots, N$

2: Set $\kappa = 0$

3: **While** $(|D^{[\kappa]} - D^{[\kappa-1]}| \geq \epsilon)$ **do** \triangleright where κ denotes the while loop iteration index

4: $\kappa = \kappa + 1$

5: Find $P_i^{[\kappa]}$, $i = 1, \dots, N$, by numerically solving (15), for example, using bisectional search method

6: Update λ using gradient method as follows:

$$\lambda^{[\kappa+1]} = \max\{\lambda^{[\kappa]} + \alpha^{[\kappa]}(\sum_{i=1}^N P_i^{[\kappa]} \sigma_i^2 - P_{\text{tot}}), 0\} \quad *$$

7: Calculate $D^{[\kappa]}$

8: **end while**

ALGORITHM 1

to numerical methods (e.g., bisectional search over P_i' 's in (15) and gradient method to update λ) to find the optimal P_i' , $i = 1, \dots, N$ in an iterative manner as outlined under Algorithm 1. From (17), note that the active sensors $P_i' > 0$ have corresponding Lagrangian multipliers $\mu_i = 0$. The sensors with $P_i' = 0$ are removed from the system. The parameter $\alpha^{[\kappa]}$ in * denotes the step-size. Since the objective function of the optimization problem is convex and bounded and the constraints are linear, therefore the algorithm can achieve convergence to the absolute minimum (the KKT point) of the problem provided the step-size $\alpha^{[\kappa]}$ is selected properly [27]. Unfortunately, Algorithm 1 will be computationally quite expensive (unless the network size is small) due to

- (i) a number of matrix inversions involved in (15) while numerically searching for P_i' , $i = 1, \dots, N$ in each iteration;
- (ii) the dependence of the convergence properties on the step-size $\alpha^{[\kappa]}$ [27, 28].

In the sequel, based on a successive approximation (SA) principle, we present a novel quasianalytical solution of the optimization problem which is simple and the associated algorithm is computationally efficient compared to Algorithm 1, exhibits remarkable convergence properties, and achieves distortion very close to the global optimum of the Algorithm 1 with no appreciable performance gap. This so-called SA-based design can be viewed as the joint optimization of the transmit powers and the modified LMMSE coefficients as we will see in the subsequent development.

According to the idea of successive approximation, a modified function is constructed from the given function in some special way [29–31]. Then that modified function is solved iteratively/successively to find the solution for the underlying problem. The solution obtained by the SA approach can be viewed as quasianalytical solution. We apply the idea of successive approximation to the reconstruction distortion function and solve the problem of power allocation in the sensor network. To this end, at the FC, to form the estimate $\hat{s} = \sum_{i=1}^N a_i z_i$ of the source s and to characterize the

resultant mean-squared distortion D , we proceed as follows. We can write the distortion D as

$$\begin{aligned} D &= \mathbb{E}_{\{s, s_i, n_i, w_i\}_{g_i, \forall i}} [(s - \hat{s})^2] \\ &= \sigma_s^2 + \sum_{i=1}^N a_i^2 (g_i^2 P_i' \sigma_i^2 + \sigma_{w_i}^2) - 2 \sum_{i=1}^N a_i g_i \sqrt{P_i'} \sigma_s \sigma_{s_i} \rho_{si} \\ &\quad + \sum_{i=1}^N \sum_{j \neq i}^N a_i a_j g_i g_j \sqrt{P_i' P_j'} \sigma_{s_i} \sigma_{s_j} \rho_{ij}, \end{aligned} \quad (18)$$

and by solving $\partial D / \partial a_i = 0$, we get the following expression for the LMMSE weighting coefficients:

$$a_i = \beta_i \gamma_i, \quad \forall i. \quad (19)$$

The variables β_i and γ_i are, respectively, defined as follows:

$$\gamma_i = \frac{g_i \sqrt{P_i'}}{g_i^2 P_i' \sigma_i^2 + \sigma_{w_i}^2}, \quad \forall i, \quad (20)$$

$$\beta_i = \sigma_s \sigma_{s_i} \rho_{si} - \sum_{j \neq i}^N \beta_j \gamma_j g_j \sqrt{P_j'} \sigma_{s_i} \sigma_{s_j} \rho_{ij}, \quad \forall i, \quad (21)$$

where $\sigma_i^2 = \sigma_{s_i}^2 + \sigma_{n_i}^2$. With (20) and (21), the distortion in (18) simplifies to

$$\begin{aligned} D &= \sigma_s^2 - \sum_{i=1}^N \beta_i \gamma_i g_i \sqrt{P_i'} \sigma_s \sigma_{s_i} \rho_{si} \\ &= \sigma_s^2 - \sum_{i=1}^N \frac{g_i^2 P_i'}{g_i^2 P_i' \sigma_i^2 + \sigma_{w_i}^2} \beta_i \sigma_s \sigma_{s_i} \rho_{si}. \end{aligned} \quad (22)$$

Equation (21) forms a set of N coupled equations which constitute the Wiener-Hopf equation for the LMMSE filter coefficients (β_i , for all i). If we know the transmit powers $\{P_i'\}_{i=1}^N$ then for given covariance $\Lambda_{s, s_i, s_j, n_i, w_i}$ and the channel gains $\{g_i\}_{i=1}^N$, we can find the coefficients $\{\beta_i\}_{i=1}^N$ by solving (21). For the solution, it is convenient to employ the matrix-vector form as follows:

$$\boldsymbol{\beta} = \mathbf{R}^{-1} \mathbf{c}, \quad (23)$$

where $\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]^T$, $\mathbf{c} = [\sigma_s \sigma_{s_1} \rho_{s1}, \dots, \sigma_s \sigma_{s_N} \rho_{sN}]^T$, $[\mathbf{R}]_{ij} = \gamma_j g_j \sqrt{P_j'} \sigma_{s_i} \sigma_{s_j} \rho_{ij}$ for $i \neq j$, and $[\mathbf{R}]_{ij} = 1$ for $i = j$. Note that $[\mathbf{R}]_{ij}$ denotes the (i, j) th element of the matrix \mathbf{R} . However, the point here is that we do not know the transmit powers $\{P_i'\}_{i=1}^N$. The following subsection presents an alternative solution to the problem of optimizing the transmit powers of the sensors under the network-wide power constraint such that the distortion D be minimized. Therein, to derive an algorithm for the solution of P_i' , the underlying idea is to assume $\boldsymbol{\beta}$ as constant. Based on this assumption, we derive an iterative algorithm which computes $\boldsymbol{\beta}$ using the values of $\{P_n'\}_{n=1}^N$ from the previous iteration. This successive approximation (SA) of the distortion function in (22) makes the solution of the

power allocation problem simple and easy to compute as will be seen in the ensuing development. Note that the resulting design for power allocation can be viewed as a joint optimization of $\{\beta_i\}_{i=1}^N$ and $\{P'_i\}_{i=1}^N$.

3.2. Minimization of the Distortion Subject to the Power Constraint-SA. Herein, we solve the optimization problem in (13) based on the distortion function D in (22) and using the successive-approximation principle outlined in the preceding subsection. For given $\beta_i \geq 0$, it is easy to verify that the distortion function is convex with respect to the optimization variables P'_i , $i = 1, \dots, N$. Therefore, the KKT conditions are sufficient for optimality [26] which are given as follows:

$$-\frac{\sigma_{w_i}^2 g_i^2 \beta_i \sigma_s \sigma_{s_i} \rho_{s_i}}{(g_i^2 P'_i \sigma_i^2 + \sigma_{w_i}^2)^2} + \lambda \sigma_i^2 - \mu_i = 0, \quad (24)$$

$$\lambda \left(\sum_{i=1}^N P'_i \sigma_i^2 - P_{\text{tot}} \right) = 0, \quad \lambda \geq 0, \quad (25)$$

$$\sum_{i=1}^N P'_i \sigma_i^2 \leq P_{\text{tot}},$$

$$\mu_i P'_i = 0, \quad \mu_i \geq 0, \quad P'_i \geq 0, \quad \forall i. \quad (26)$$

Solving (24) for active sensor i (i.e., $P'_i > 0$, $\mu_i = 0$ from (26)), we get

$$P'_i = \frac{1}{\zeta_i \sigma_i^2} \left(\sqrt{\frac{\zeta_i \beta_i \sigma_s \sigma_{s_i} \rho_{s_i}}{\lambda \sigma_i^2}} - 1 \right)^+, \quad (27)$$

for $i = 1, \dots, N$, where $\zeta_i = g_i^2 / \sigma_{w_i}^2$ denotes the channel SNR for sensor i and $(x)^+ = \max(x, 0)$. Based on (27), following observations are in order.

- (1) There exists a cut-off value $\zeta_i^{(o)} = (4\beta_i \sigma_s \sigma_{s_i} \rho_{s_i}) / (\lambda \sigma_i^2)$ such that for $\zeta_i \leq \zeta_i^{(o)}$, the power allocation policy follows waterfilling on channel SNR, that is, P'_i increases with increasing ζ_i ; and for $\zeta_i > \zeta_i^{(o)}$, the power allocation is according to inversion in the channel SNR, that is, increasing ζ_i decreases P'_i .
- (2) The sensors with higher observation noise variances are given less power. For sensor i , in the limiting case $\sigma_{n_i}^2 \rightarrow \infty$ (i.e., $\sigma_i^2 \rightarrow \infty$) then $P'_i \rightarrow 0$.
- (3) The sensors with weak correlation with the source are allotted less power. For instance, if $\rho_{s_i} \rightarrow 0$ then $P'_i \rightarrow 0$.

Combining the aforementioned points we can see that the final power allocation policy for the sensors depends on the spatial correlations, variance of the observation noises, and the channel SNRs. Moreover, we see that depending on the values of these system parameters some of the sensors may

be switched-off altogether. For sensor i to be active, following condition must hold:

$$\frac{\zeta_i \beta_i \sigma_s \sigma_{s_i} \rho_{s_i}}{\sigma_i^2} > \frac{\lambda}{\sigma_s}, \quad (28)$$

which stems from the fact that $P'_i > 0$ if sensor i is active. Let \mathcal{K} denotes the set of active sensors defined as follows:

$$\mathcal{K} = \left\{ k \mid \frac{\zeta_k \beta_k \sigma_s \sigma_{s_k} \rho_{s_k}}{\sigma_k^2} > \frac{\lambda}{\sigma_s} \right\}. \quad (29)$$

Since the problem is convex, the minimum of the objective function occurs at the sum power constraint boundary, that is, the constraint is active. Therefore, the transmit powers P'_k , $k \in \mathcal{K}$ must satisfy the power constraint with equality, that is, $\sum_{k \in \mathcal{K}} P'_k \sigma_k^2 = P_{\text{tot}}$, which gives

$$\lambda = \left(\frac{\sum_{k \in \mathcal{K}} \sqrt{\beta_k \sigma_s \sigma_{s_k} \rho_{s_k} / (\zeta_k \sigma_k^2)}}{P_{\text{tot}} + \sum_{k \in \mathcal{K}} 1 / \zeta_k} \right)^2. \quad (30)$$

Based on the solution from (27) through (30), Algorithm 2 can be proposed which iteratively optimizes the transmit powers $\{P'_i\}_{i=1}^N$ and the variables $\{\beta_i\}_{i=1}^N$, while minimizing the reconstruction distortion subject to the power constraint. If during iterations any sensor does not fulfill the condition in (28), it is switched off and the algorithm continues with the remaining sensors until the convergence criterion is fulfilled. Regarding the convergence properties of the algorithm, consider the following.

- (i) Since in each iteration (successive approximation) we are minimizing a convex function over the convex-set of the transmit powers $\{P'_i \mid \sum_{i=1}^N P'_i \sigma_i^2 \leq P_{\text{tot}} \text{ and } P'_i \geq 0, i = 1, \dots, N\}$, therefore the optimality of the transmit powers $\{P'_i\}_{i=1}^N$ in each approximation (for given $\{\beta_i\}_{i=1}^N$) combined with the optimality of $\{\beta_i\}_{i=1}^N$ for given $\{P'_i\}_{i=1}^N$ as per (23) means that the algorithm achieves monotonic decrease in the distortion, that is,

$$\begin{aligned} & D(P_i^{[k+1]}; \beta_i^{[k+1]}, i = 1, \dots, N) \\ & \leq D(P_i^{[k+1]}; \beta_i^{[k]}, i = 1, \dots, N) \\ & \leq D(P_i^{[k]}; \beta_i^{[k]}, i = 1, \dots, N) \end{aligned} \quad (31)$$

and consequently it does converge to a unique minimum point.

- (ii) The algorithm consistently arrives at the same combination of the transmit power tuple (P'_1, \dots, P'_N) and achieves the same minimum distortion for a wide range of different initialization points. In other words, we can say that the algorithm exhibits start point independence (for a wide range of initialization points). Moreover, the algorithm asymptotically

```

1: Initialize  $P_i^{[0]}$  for  $i = 1, \dots, N$ 
2: Calculate  $\beta_i^{[0]}$  for  $i = 1, \dots, N$ 
3: Set  $\kappa = 0$ 
4: while ( $|D^{[\kappa]} - D^{[\kappa-1]}| \geq \epsilon$ ) do where  $\kappa$  denotes the
   while loop iteration index
5:    $\kappa = \kappa + 1$ 
6:   For  $i = 1, \dots, N$  determine transmit power as follows:
7:   if Condition in (28) is true then
8:     Determine  $P_i^{[\kappa]}$  from (27)
9:   else
10:     $P_i^{[\kappa]} = 0$ 
11:   end if
12:   For  $i = 1, \dots, N$  update  $\beta_i^{[\kappa]}$  from (23)
13:   Calculate  $D^{[\kappa]}$ 
14: end while
    
```

ALGORITHM 2

achieves the lower-bound distortion D_0 with increasing transmit power P_{tot} . In Section 5, we illustrate the monotonic decrease, start point insensitivity, and the asymptotic convergence to D_0 with several simulation examples. There we also show that the convergence may be achieved in as few as two or three iterations.

- (iii) We have shown that the original problem is convex and therefore the objective function has a global minima under the power constraint which can be achieved by Algorithm 1. Now the question is how closely does the successive approximation-based algorithm converge to the global minimum value? The simulation examples in Section 5 show that the distortion achieved by both algorithms are extremely close and the performance gap between the full-optimization and the successive approximation based algorithms is virtually negligible.

It is a quite remarkable that the algorithm exhibits such excellent convergence properties which illustrates that the proposed successive approximation strategy works quite well. Finally, compared to the power allocation Algorithm 1, the ease of computation and simplicity of the design based on the successive approximation principle can be appreciated from the simple and elegant structure of (27)–(30).

4. Power Allocation with Imperfect CSI

Heretofore, we have assumed perfect knowledge of the channel gains $\{h_i\}_{i=1}^N$. However, in practice, we have estimates $\{\hat{h}_i\}_{i=1}^N$ of the actual channel gains. One way to estimate the channel is by a training sequence whereby each sensor transmits a known sequence of data symbols called pilots. Then based on the received data, the FC estimates the channel. Let t_i denote the pilot symbol transmitted by sensor i in the channel estimation phase. The corresponding

received signal is $r_i = h_i t_i + \bar{w}_i$ and based on which the LMMSE estimate \hat{h}_i of h_i is

$$\hat{h}_i = \frac{\mathbb{E}_{\{h_i, \bar{w}_i\}} [h_i r_i^*]}{\mathbb{E}_{\{h_i, \bar{w}_i\}} [|r_i|^2]}, \quad r_i = \frac{\sigma_{h_i}^2 t_i^*}{\sigma_{h_i}^2 |t_i|^2 + \sigma_{\bar{w}_i}^2} r_i, \quad (32)$$

where $(\cdot \cdot \cdot)^*$ denotes the complex conjugate operation. The variance of the estimation error $\Delta \bar{h}_i := h_i - \hat{h}_i$ is

$$\bar{\delta}_i^2 = \mathbb{E}_{\{h_i, \bar{w}_i\}} \left[|h_i - \hat{h}_i|^2 \right] = \frac{\sigma_{h_i}^2 \sigma_{\bar{w}_i}^2}{\sigma_{\bar{w}_i}^2 + \sigma_{h_i}^2 |t_i|^2}, \quad (33)$$

wherein $|t_i|^2$ is power of the transmitted pilot. Note that the variance of channel estimation error is finite for finite $|t_i|^2$ and $\sigma_{\bar{w}_i}^2$. The actual channel can be represented as a sum of the estimate and the estimation error, that is,

$$h_i = \hat{h}_i + \Delta \bar{h}_i, \quad (34)$$

where $\Delta \bar{h}_i \sim \mathcal{CN}(0, \bar{\delta}_i^2)$. Such an approach to model the channel estimation error can be viewed as the Bayesian approach [6].

One way to design the power-scheduling scheme is by replacing h_i and g_i , respectively, by \hat{h}_i and \hat{g}_i in the formulations of the foregoing section. This constitutes a naive-approach because it ignores the error in the channel estimate. An alternative design originates by substituting (34) in (2) as follows:

$$\bar{z}_i = \hat{h}_i \sqrt{P_i'}(s_i + n_i) + \underbrace{\sqrt{P_i'}(s_i + n_i) \Delta \bar{h}_i + \bar{w}_i}_{:= \bar{u}_i}, \quad (35)$$

in which \bar{u}_i can be viewed as total receiver noise corresponding to sensor i with $\mathbb{E}_{\{s_i, n_i, \bar{w}_i, \Delta \bar{h}_i\}} [\bar{u}_i] = 0$, $\mathbb{E}_{\{s_i, n_i, \bar{w}_i, \Delta \bar{h}_i\}} [|\bar{u}_i|^2] = P_i'(\sigma_{s_i}^2 + \sigma_{n_i}^2) \bar{\delta}_i^2 + \sigma_{\bar{w}_i}^2$, for all i , and $\mathbb{E}_{\{s_i, s_j, n_i, n_j, \bar{w}_i, \bar{w}_j, \Delta \bar{h}_i, \Delta \bar{h}_j\}} [\bar{u}_i \bar{u}_j^*] = 0$, for all $i \neq j$. Noting that $\hat{h}_i = \hat{g}_i e^{j\theta_{\hat{h}_i}}$, we can write

$$\bar{z}_i e^{-j\theta_{\hat{h}_i}} = \hat{g}_i \sqrt{P_i'}(s_i + n_i) + \bar{u}_i e^{-j\theta_{\hat{h}_i}}, \quad (36)$$

where the exponential term $e^{-j\theta_{\hat{h}_i}}$ can be absorbed into \bar{u}_i , that is, into the Gaussian variables $\Delta \bar{h}_i$ and \bar{w}_i without changing their statistical properties—thanks to their circular symmetry. Since the underlying source s and the observation $s_i + n_i$ are real-valued, as a consequence only the part of the noise \bar{u}_i in-phase with the sensor observation is relevant for estimation of the source s . Therefore, we can write

$$z_i = \hat{g}_i \sqrt{P_i'}(s_i + n_i) + u_i, \quad (37)$$

where $u_i = \Re\{\bar{u}_i\} = \sqrt{P_i'}(s_i + n_i) \Delta h_i + w_i$, $\Delta h_i \sim \mathcal{N}(0, \delta_i^2)$ and $w_i \sim \mathcal{N}(0, \sigma_{w_i}^2)$. Following a procedure similar to Section 3, it can be shown that the mean-squared reconstruction distortion of the estimate $\hat{s} = \sum_{i=1}^N a_i z_i$ with respect to s is given by

$$\begin{aligned} \tilde{D} &= \mathbb{E}_{\{s, s_i, n_i, w_i, \Delta h_i, \hat{g}_i, \forall i\}} \left[(s - \hat{s})^2 \right] \\ &= \sigma_s^2 - \sum_{i=1}^N \beta_i \gamma_i \hat{g}_i \sqrt{P_i'} \sigma_s \sigma_{s_i} \rho_{s_i}, \end{aligned} \quad (38)$$

where

$$\begin{aligned} a_i &= \beta_i \gamma_i, \\ \beta_i &= \sigma_s \sigma_{s_i} \rho_{si} - \sum_{j \neq i}^N \beta_j \gamma_j \hat{g}_j \sqrt{P'_j} \sigma_{s_i} \sigma_{s_j} \rho_{ij}, \\ \gamma_i &= \frac{\hat{g}_i \sqrt{P'_i}}{\hat{g}_i^2 P'_i \sigma_i^2 + P'_i \sigma_i^2 \delta_i^2 + \sigma_{w_i}^2}, \quad \forall i. \end{aligned} \quad (39)$$

The solution of the optimization problem in (13) with the objective to minimize the distortion \tilde{D} defined in (38) subject to the constraint on the total network power can be obtained by using the method of Lagrangian multipliers and is outlined as follows.

$$P'_k = \frac{1}{\left(1 + \delta_k^2 / \hat{g}_k^2\right) \hat{\zeta}_k \sigma_k^2} \left(\sqrt{\frac{\hat{\zeta}_k \beta_k \sigma_s \sigma_{sk} \rho_{sk}}{\lambda \sigma_k^2}} - 1 \right)^+, \quad (40)$$

- (1) The power allotted to sensor k is for $k = 1, \dots, N$, where $\hat{\zeta}_k := \hat{g}_k^2 / \sigma_{w_k}^2$ defines the channel SNR based on the channel estimate.
- (2) For sensor k to be active, that is, $P'_k > 0$, the following condition:

$$\frac{\hat{\zeta}_k \beta_k \sigma_s \sigma_{sk} \rho_{sk}}{\sigma_k^2} > \frac{\lambda}{\sigma_s} \quad (41)$$

must hold, otherwise it is switched-off.

- (3) The index-set \mathcal{K} of the active sensors is

$$\mathcal{K} = \left\{ k \mid \frac{\hat{\zeta}_k \beta_k \sigma_s \sigma_{sk} \rho_{sk}}{\sigma_k^2} > \frac{\lambda}{\sigma_s} \right\}. \quad (42)$$

- (4) The Lagrangian multiplier λ is

$$\lambda = \left(\frac{\sum_{k \in \mathcal{K}} \sqrt{\beta_k \sigma_s \sigma_{sk} \rho_{sk} / \left((1 + \delta_k^2 / \hat{g}_k^2) \hat{\zeta}_k \sigma_k^2 \right)}}{P_{\text{tot}} + \sum_{k \in \mathcal{K}} 1 / \left((1 + \delta_k^2 / \hat{g}_k^2) \hat{\zeta}_k \right)} \right)^2, \quad (43)$$

which is determined such that the power constraint be satisfied with equality.

Based on (40)–(43), the power allocation for the sensors can be obtained using the procedure outlined under Algorithm 2. (The power allocation design under Algorithm 1 can similarly be extended to the imperfect CSI case.) Note that the convergence properties of the algorithm with perfect CSI also applies to the imperfect CSI case. Moreover, the above power allocation design exhibits robustness to the channel estimation errors compared to the naive approach as shown in the subsequent section.

Remark 2. We can observe that as $\delta_k^2 \rightarrow 0$ then $\hat{g}_k \rightarrow g_k$ for $k = 1, \dots, N$ and (40)–(43), respectively, converges to the power allocation design with perfect CSI in (27)–(30).

5. Performance Evaluation and Discussion

Through simulation examples, this section corroborates the analytical findings and illustrates the effectiveness of the proposed adaptive power allocation (APA) designs under Algorithms 1 and 2 for the perfect and imperfect CSI cases. We assume without any loss of generality that $\sigma_s^2 = \{\sigma_{s_i}^2\}_{i=1}^N = 1$. In the simulations, the distortion is calculated from 10^5 realizations of the underlying source, partial observations, and observation and receiver noises according to the covariance $\Lambda_{s, s_i, s_j, n_i, w_i}$. The simulation examples focus on the successive approximation-based power allocation design unless stated otherwise. In the figures, $\log(\cdot \cdot \cdot)$ denotes the logarithm with base 10.

5.1. Spatial Correlation. In order to show the efficacy of our design, we compare its performance with a uniform power allocation-based design. In the figures, the designs are, respectively, denoted as APA (Adaptive Power Allocation) and UPA (Uniform Power Allocation). Moreover, $\{P_i\}_{i=1}^N = P_u = P_{\text{tot}}/N$ for the UPA design.

We consider two sensor networks, respectively, comprising $N = 3$ and $N = 500$ sensors which are uniformly distributed in a 100×100 grid with the source s at its center. Figure 2 plots the distortion achieved by the SA-based APA design and compares it with the UPA design. The distortion is averaged over 10^4 independent realizations (drawn from a uniform distribution) of the sensors deployment. The figure shows that our proposed design outperforms the UPA scheme and the achieved distortion monotonically approaches the lower-bound distortion value D_0 with increasing P_{tot} . For given θ_1 , we can observe that the distortion decreases with increasing the number of sensors and the performance gap between the APA and UPA designs also increases. Moreover, we can see that increasing the value of θ_1 decreases the distortion. This is because, for given deployment, the spatial correlation of the sensors with the source (and with each other) improves with increasing θ_1 [c.f. (6)]. Note that at low value of θ_1 , the distortion is high and the performance gap between the APA and UPA designs is very small. However, as the value of θ_1 increases the distortion decreases and the performance gap between the APA and UPA increases. However, we can see that with increasing the value of θ_1 further the performance gap starts decreasing. This is because for very small value of θ_1 the sensors have very low correlation with the source and for very large value of θ_1 the correlation is high for all sensors, and in these extreme cases the UPA scheme is as good as the APA scheme.

Next, for the sake of illustration, we focus on the network with three sensors, that is, $N = 3$, and we consider the following two examples:

- (i) Ex1: $(d_{X_1}, d_{X_2}, d_{X_3}) = (-0.3, 0, 0.8)$ and $(d_{Y_1}, d_{Y_2}, d_{Y_3}) = (0, 1.6, 0)$,
- (ii) Ex2: $(d_{X_1}, d_{X_2}, d_{X_3}) = (-0.1, 0, 1.5)$ and $(d_{Y_1}, d_{Y_2}, d_{Y_3}) = (0, 5, 0)$,

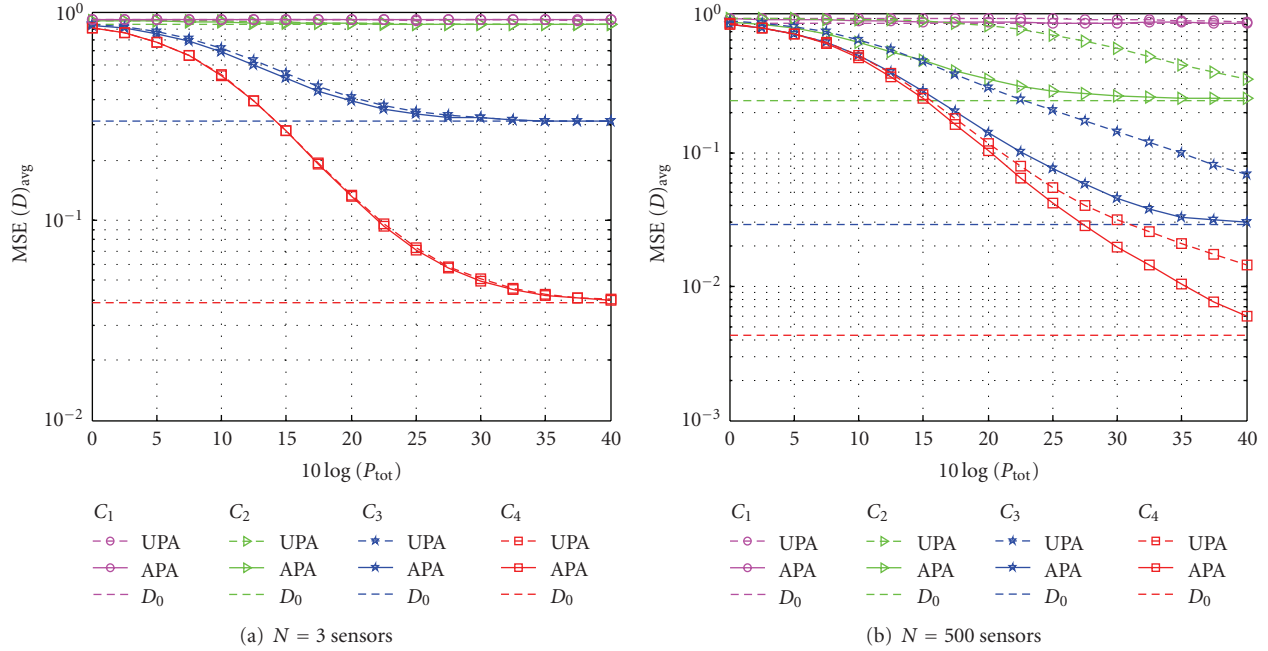


FIGURE 2: $\theta_2 = 1$, $(\sigma_{n_i}^2, h_i, \sigma_{w_i}^2)_{i=1}^N = (0.01, 1, 10)$, and $(C_1)\theta_1 = 1$, $(C_2)\theta_1 = 10$, $(C_3)\theta_1 = 10^2$, and $(C_4)\theta_1 = 10^3$.

where (d_{X_i}, d_{Y_i}) gives the position of sensor i with respect to the origin in the XY -plane. Note that we can view these examples as specific realizations of the deployment of the sensors. Assuming the source at the origin and for $\theta_1 = \theta_2 = 1$ in (6), we obtain the following spatial correlation values:

- (i) Ex1: $(\rho_{s_1}, \rho_{s_2}, \rho_{s_3}) = (0.7408, 0.2019, 0.4493)$ and $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.1963, 0.3329, 0.1671)$,
- (ii) Ex2: $(\rho_{s_1}, \rho_{s_2}, \rho_{s_3}) = (0.9048, 0.0067, 0.2231)$ and $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.0067, 0.2019, 0.0054)$.

The simulations in the sequel are based on these two examples. We have taken these examples for purely illustrative purpose and they in no way limit the generality of our designs. For zero-observation noise variances, the spatial correlation lower bounds the reconstruction distortion at $D_0(0.4038, 0.1796)$, where first value is for Ex1 and the second value is for Ex2. The distortion in the subsequent simulation examples cannot be below this value no matter how high the transmit power becomes.

Figure 3 shows that the proposed design gives reconstruction distortion which is less than that achieved by the uniform power allocation. This superior performance originates from the reason that the proposed design assigns all or more power to the sensor(s) with better correlation properties. This is contrary to the UPA scheme which gives equal importance to all sensors regardless of the correlation structure and thereby wasting power. For both examples, note that the achieved distortion decreases monotonically with increasing P_{tot} but is never less than the lower-bound value $D_0(0.4086, 0.1875)$.

For Ex1 and Ex2, Figure 4 shows that with different initial values of $\{P'_k\}_{k=1}^N$, Algorithm 2 converges to the

same distortion value (0.5817 for Ex1 and 0.2646 for Ex2). Moreover, at the convergence, the power distribution among the sensors is $10 \log(P_1, P_2, P_3)_{\text{Ex1}} = (13.8913, 0, 8.5279)$ and $10 \log(P_1, P_2, P_3)_{\text{Ex2}} = (19.8403, 0, 5.5743)$, respectively, for Ex1 and Ex2 in all cases irrespective of the initialization point of the algorithm. Note that in one iteration the distortion reaches fairly close to the minimum value. Nevertheless, after the second or third iteration there is virtually no appreciable change in these values. This observation extends to all simulation examples presented herein.

Figure 5 compares performance of the proposed APA designs under Algorithms 1 and 2 which shows that the distortion curves produced by the two algorithms are extremely close and the performance gap is negligible. This is quite remarkable result especially when viewed in combination with the simplicity and computational efficiency of Algorithm 2 compared to Algorithm 1.

The simulation examples in the sequel only treat the APA design based on the successive approximation (SA) without including comparison with the APA design under Algorithm 1 and the UPA scheme. Nevertheless, in all the instances, the SA-based design closely achieves the performance of the design in Algorithm 1 and outperforms the UPA scheme except in a symmetric case where both designs (APA and UPA) converge. In the symmetric case, the correlations, variances of the observation and the receiver noises, and the channel gains are same across all sensors.

5.2. Channel SNR. Assuming fixed channel gains (no fading), $\{h_i\}_{i=1}^N = 1$ and $\{\sigma_{n_i}^2\}_{i=1}^N = 0.01$, we consider the following cases: $(C_{w1})\{\zeta_i\}_{i=1}^N = -10$ dB, $(C_{w2})\{\zeta_i\}_{i=1}^N = -3.98$ dB,

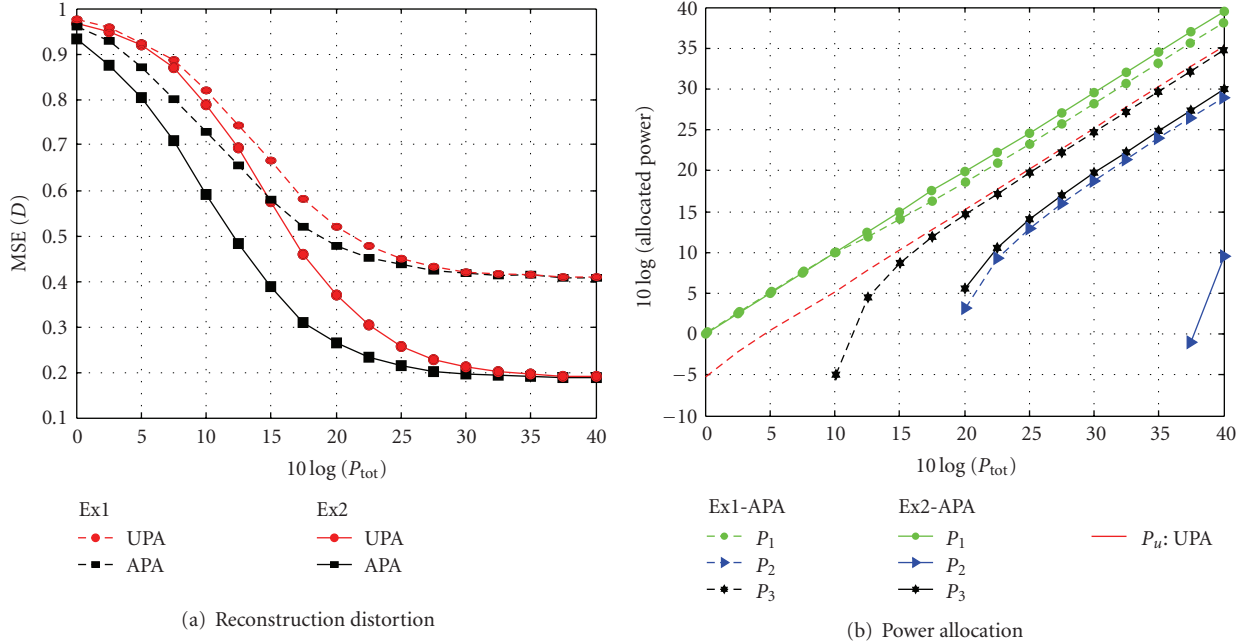


FIGURE 3: $(\sigma_{n_i}^2, h_i, \sigma_{w_i}^2)_{i=1}^N = (0.01, 1, 10)$.

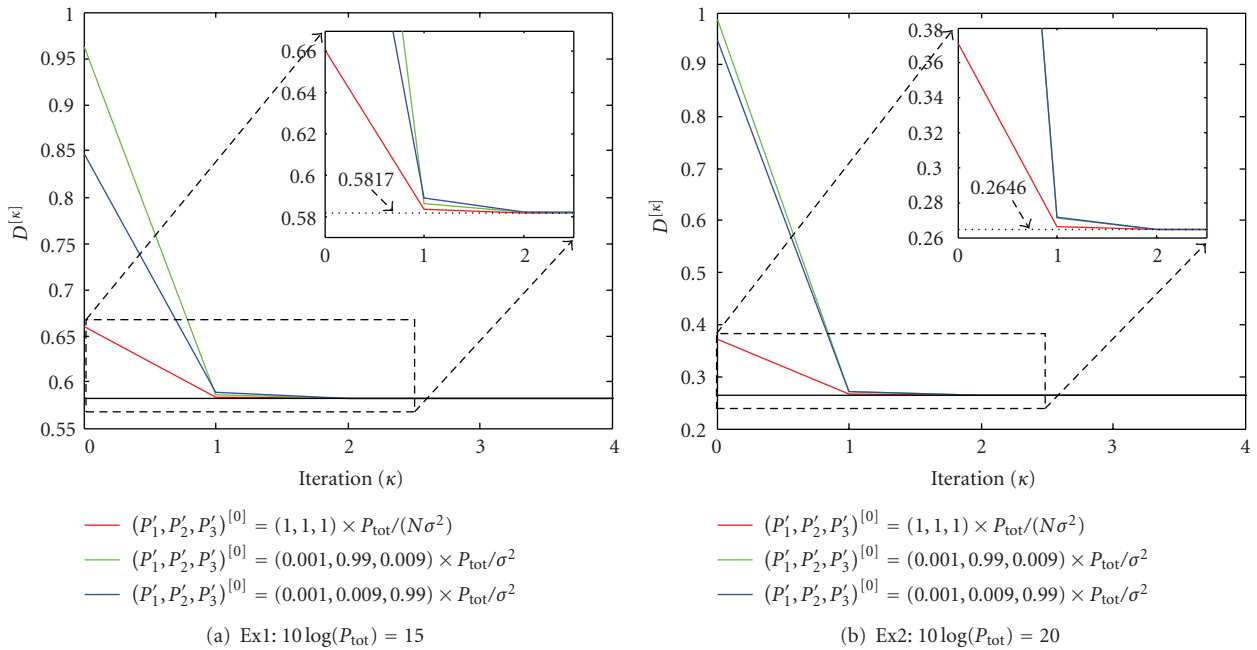


FIGURE 4: Algorithm convergence behavior for different initialization points $(P'_1, P'_2, P'_3)^{[0]}$.

$(C_{w3})\{\zeta_i\}_{i=1}^N = 0$ dB, $(C_{w4})\{\zeta_i\}_{i=1}^N = 10$ dB, $(C_{w5})\{\zeta_i\}_{i=1}^N = -20$ dB, and $(C_{w6})\{\zeta_1, \zeta_2, \zeta_3\} = \{-30, -10, -20\}$ dB. Figure 6 shows that for given P_{tot} the distortion decreases with increase in the so-called channel SNR and vice versa. Note that in each case the achieved distortion monotonically approaches the lower-bound value ($D_0 = 0.4086$) with increasing P_{tot} . Figure 6(b) shows how the total power P_{tot} is distributed among the sensors in C_{w5} and C_{w6} . The

figure shows that in C_{w5} , the sensors with better correlation properties are given more power, which is due to the fact that, in this case, the system is symmetric with respect to all other system parameters. However, the case C_{w6} is different where the channel SNRs are not same across the sensors. For this case the figure shows that the power allocation policy follows sensor selection and waterfilling with respect to the SNR until the next sensor is turned on, after which point the

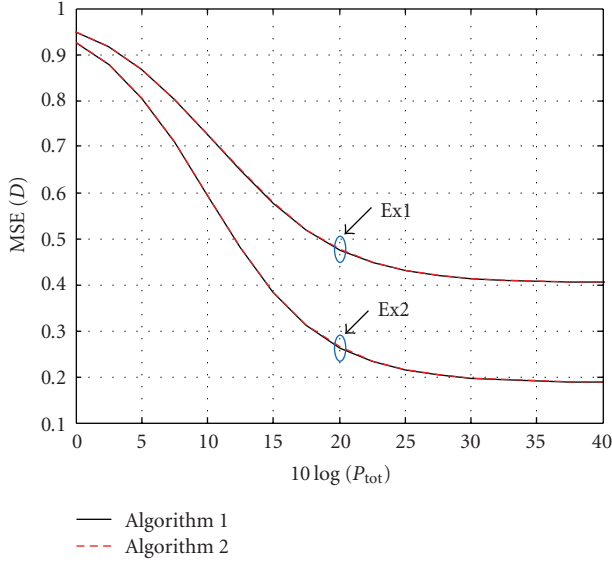


FIGURE 5: Performance comparison of the APA designs.

power allocation is according to the channel inversion. These examples show that the power allocation policy is jointly determined by the spatial correlation values, variance of the observation noises, and the channel SNRs.

5.3. Observation Noise. Herein, we assume that $\{h_i\}_{i=1}^N = 1$, $\{\sigma_{w_i}^2\}_{i=1}^N = 10$, and consider the following cases: $(C_{n1})\{\sigma_{n_i}^2\}_{i=1}^N = 0.01$ and $D_0 = 0.4086$, $(C_{n2})\{\sigma_{n_i}^2\}_{i=1}^N = 0.1$ and $D_0 = 0.4484$, $(C_{n3})\{\sigma_{n_i}^2\}_{i=1}^N = 1$ and $D_0 = 0.6652$, and $(C_{n4})\{\sigma_{n_1}^2, \sigma_{n_2}^2, \sigma_{n_3}^2\} = \{1, 0.01, 0.1\}$ and $D_0 = 0.6160$. Figure 7 shows that increasing the variances of the observation noise at the sensors, increases the reconstruction distortion. Moreover, in each case, the achieved distortion D decreases monotonically with the transmit power P_{tot} and approaches D_0 at high transmit power.

5.4. Imperfect CSI. We assume that the communication channels $\{h_i\}_{i=1}^N$ from the sensors to the FC undergo independent Rayleigh fading such that $\{\sigma_{h_i}^2\}_{i=1}^N = 1$. Moreover, we assume that $\{\sigma_{n_i}^2\}_{i=1}^N = 0.01$ and $\{\sigma_{w_i}^2\}_{i=1}^N = 10$. The simulations in Figure 8 illustrate the achieved distortion when we have (i) perfect knowledge of the CSI, that is, $\{\delta_i^2\}_i = \delta^2 = 0$ and (ii) the estimates of the channels with estimation error variances $\delta^2 = 0.01$ and $\delta^2 = 0.1$. The distortion is averaged over 2×10^6 realizations of the source and the corresponding partial observations, the observation and receiver noises, and the channel gains. From the figure we can see that the channel errors increase the reconstruction distortion. Moreover, the proposed design (Rob in the figure) shows robustness to the channel estimation errors compared to the naive approach (Nav in the figure).

6. Conclusions

In this paper we have investigated the joint adaptive power allocation design for the sensors to transmit their observations to the FC with the goal to reconstruct the underlying source with minimum distortion subject to a constraint on the total network power. We demonstrated the effectiveness of our proposed design with a few simple examples where we vary one parameter while keeping all others the same across sensors. In practice, the correlation, observation noises, and the SNRs of the communication channels may simultaneously vary across sensors. Consequently, the power allocation policy is a nontrivial function of all these factors. Our proposed design incorporates them in (quasi)analytical expressions. We showed that the power allocation problem is convex which does not bear closed-form analytical solution. We showed that the problem can be solved by an iterative numerical procedure which is computationally expensive unless the network size is small. Then, based on the successive approximation principle, we proposed a novel power allocation design which is simple and easy to compute, and exhibits excellent convergence properties. We demonstrated that our proposed design outperforms a scheme based on the uniform power allocation except in a symmetric case where both designs converge. The performance gain may be large for a relatively large sensor network with high heterogeneity in the spatial correlation, observation quality, and the channel gains. In the proposed design, we also incorporated the case when imperfect knowledge of the channel gains is available. We demonstrated that the proposed design exhibits robustness to the channel estimation errors relative to a naive approach. The future work may consider the extension to the nonorthogonal channels from the sensors to the FC. Moreover, optimizing the performance such as reconstruction distortion or the power consumption over a priori given network life can be investigated. We only considered the power required for transmitting the observations to the FC, where as the future work may also include the power consumed in the sensing, preparing the observation for transmission and in the receiving processes.

Appendices

A. Proof of Reconstruction Distortion in (10)

From (9), we can write

$$\begin{aligned} D &= \sigma_s^2 - \mathbf{c}^T \mathbf{H} (\mathbf{H} \mathbf{C} \mathbf{H} + \mathbf{C}_w)^{-1} \mathbf{H} \mathbf{c} \\ &= \sigma_s^2 - \mathbf{c}^T \left[\mathbf{H} (\mathbf{H} \mathbf{C} \mathbf{H} + \mathbf{C}_w)^{-1} \mathbf{H} \mathbf{C} \right] \mathbf{C}^{-1} \mathbf{c}. \end{aligned} \quad (\text{A.1})$$

Let $\mathbf{A}^{-1} = \mathbf{I}$ (where \mathbf{I} denotes identity matrix), $\mathbf{V} = \mathbf{H} \mathbf{C}$, $\mathbf{U} = \mathbf{H}$, and $\mathbf{B}^{-1} = \mathbf{C}_w$, then the Woodbury identity [24]

$$(\mathbf{A} + \mathbf{U} \mathbf{B} \mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{V} \mathbf{A}^{-1} \mathbf{U} + \mathbf{B}^{-1})^{-1} \mathbf{V} \mathbf{A}^{-1} \quad (\text{A.2})$$

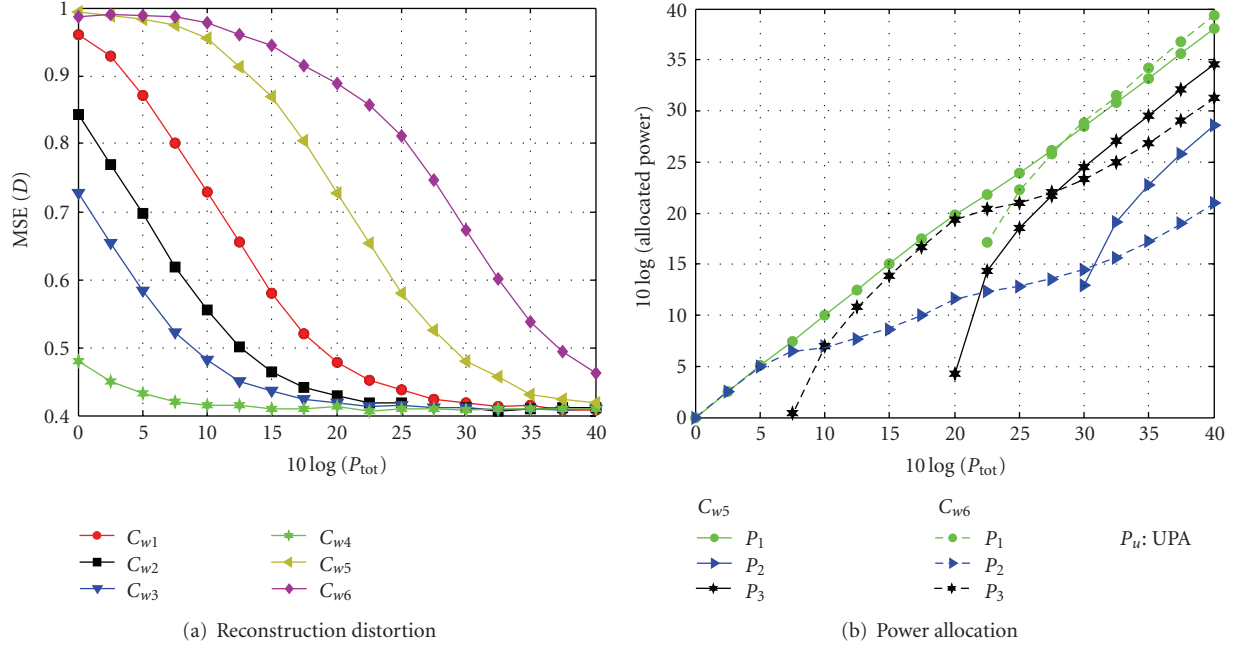


FIGURE 6: Effect of channel SNR—Ex1.

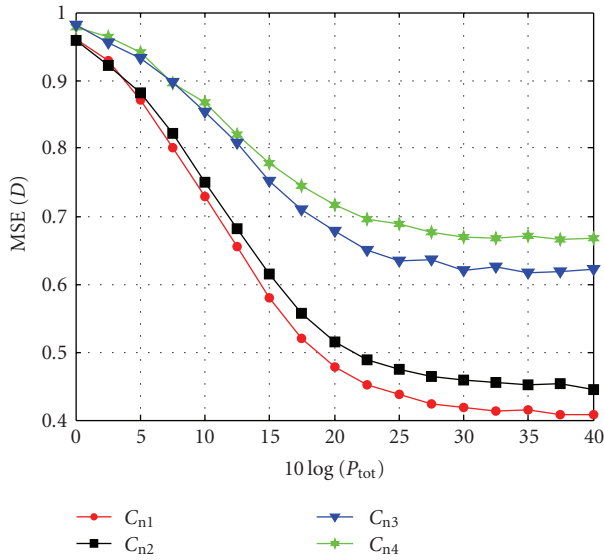


FIGURE 7: Effect of observation noise—Ex1.

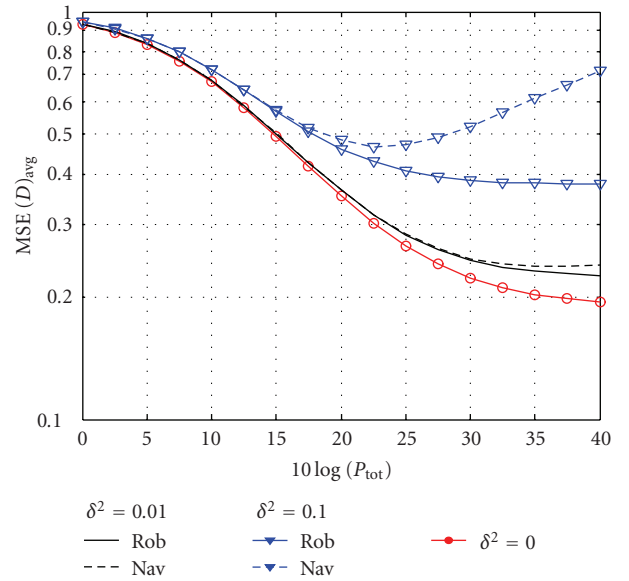


FIGURE 8: Comparison of perfect and imperfect CSI—Ex2.

gives

$$(\mathbf{I} + \mathbf{H}\mathbf{C}_w^{-1}\mathbf{H}\mathbf{C})^{-1} = \mathbf{I} - \mathbf{I}\mathbf{H}(\mathbf{H}\mathbf{C}\mathbf{I}\mathbf{H} + \mathbf{C}_w)^{-1}\mathbf{H}\mathbf{C}\mathbf{I}, \quad (\text{A.3})$$

$$\mathbf{H}(\mathbf{H}\mathbf{C}\mathbf{H} + \mathbf{C}_w)^{-1}\mathbf{H}\mathbf{C} = \mathbf{I} - (\mathbf{I} + \mathbf{H}\mathbf{C}_w^{-1}\mathbf{H}\mathbf{C})^{-1},$$

which after substitution in (A.1) gives (10).

B. Proof of Theorem 1

The vector of the transmit powers is $\mathbf{p}' = [P'_1, \dots, P'_N]^T$ and we know that the estimation distortion is given by

$$D = \sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \mathbf{c}^T (\mathbf{Y}\mathbf{C} + \mathbf{I})^{-1} \mathbf{C}^{-1} \mathbf{c}, \quad (\text{B.4})$$

which can be written as

$$D = \sigma_s^2 - \text{Tr}\{\tilde{\mathbf{C}}\} + \text{Tr}\{\tilde{\mathbf{C}}(\mathbf{Y}\mathbf{C} + \mathbf{I})^{-1}\}, \quad (\text{B.5})$$

where $\tilde{\mathbf{C}} = \mathbf{C}^{-1} \mathbf{c}\mathbf{c}^T$ and $\text{Tr}\{\mathbf{A}\}$ denotes the trace of matrix \mathbf{A} . In moving from (B.4) to (B.5) we have used the invariance

property of Tr under cyclic permutations, that is $\text{Tr}(\mathbf{ABC}) = \text{Tr}(\mathbf{BCA}) = \text{Tr}(\mathbf{CAB})$. [24].

In order to prove that the function D is convex over \mathbf{p}' , we derive its Hessian and then show that it is positive semidefinite. To this end, first we calculate the first-order derivative of D with respect to each component of \mathbf{p}' , that is, $\partial D/\partial P'_k$ for $k = 1, \dots, N$ as follows:

$$\begin{aligned} \frac{\partial D}{\partial P'_k} &= \text{Tr} \left\{ \tilde{\mathbf{C}} \left(\frac{\partial(\mathbf{YC} + \mathbf{I})^{-1}}{\partial P'_k} \right) \right\} \\ &= -\frac{g_k^2}{\sigma_{w_k}^2} \text{Tr} \left\{ \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \right\}. \end{aligned} \quad (\text{B.6})$$

In (B.6), we have used $\partial \mathbf{Z}^{-1}/\partial [\mathbf{Z}]_{kk} = -\mathbf{Z}^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{Z}^{-1}$, where $\mathbf{e}_k = (e_1, \dots, e_N)^T$ with $e_i = 1$ for $i = k$ and $e_i = 0$ for $i \neq k$, $i = 1, \dots, N$ [24].

Now we take the derivative of (B.6) with respect to P'_l for $l = 1, \dots, N$ as follows:

$$\begin{aligned} \frac{\partial^2 D}{\partial P'_k \partial P'_l} &= -\frac{g_k^2}{\sigma_{w_k}^2} \text{Tr} \left\{ \tilde{\mathbf{C}} \frac{\partial(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1}}{\partial P'_l} \right\} \\ &= \frac{g_k^2 g_l^2}{\sigma_{w_k}^2 \sigma_{w_l}^2} \text{Tr} \left\{ \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \right. \\ &\quad \left. + \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \right\} \\ &= \frac{g_k^2 g_l^2}{\sigma_{w_k}^2 \sigma_{w_l}^2} \text{Tr} \left\{ \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \right. \\ &\quad \left. + \mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \right\} \\ &= \frac{g_k^2 g_l^2}{\sigma_{w_k}^2 \sigma_{w_l}^2} \left\{ \left(\mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \right) \right. \\ &\quad \left(\mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \right) + \left(\mathbf{e}_l^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_k \right) \right. \\ &\quad \left. \left(\mathbf{e}_k^T \mathbf{C}(\mathbf{YC} + \mathbf{I})^{-1} \tilde{\mathbf{C}}(\mathbf{YC} + \mathbf{I})^{-1} \mathbf{e}_l \right) \right\} \end{aligned} \quad (\text{B.7})$$

where we have used $\partial \mathbf{UV}/\partial \mathbf{z} = \mathbf{U} \partial \mathbf{V}/\partial \mathbf{z} + \partial \mathbf{U}/\partial \mathbf{z} \mathbf{V}$ and $\text{Tr}(\mathbf{ABC}) = \text{Tr}(\mathbf{BCA})$. Now let $\mathbf{R} = \text{diag}(g_1^2/\sigma_{w_1}^2, \dots, g_N^2/\sigma_{w_N}^2)^T$ and $\tilde{\mathbf{Y}} = (\mathbf{YC} + \mathbf{I})^{-1}$ then the Hessian of D can be written in the following form:

$$\frac{\partial^2 D}{\partial \mathbf{p}'^2} = (\mathbf{RC}\tilde{\mathbf{Y}}\mathbf{R}) \circ (\mathbf{C}\tilde{\mathbf{Y}}\tilde{\mathbf{C}})^T + (\mathbf{RC}\tilde{\mathbf{Y}}\mathbf{R})^T \circ (\mathbf{C}\tilde{\mathbf{Y}}\tilde{\mathbf{C}}), \quad (\text{B.8})$$

where $\mathbf{A} \circ \mathbf{B}$ denotes the Hadamard or Schur product of \mathbf{A} and \mathbf{B} . Note that \mathbf{R} , \mathbf{C} , $\tilde{\mathbf{C}}$, and $\tilde{\mathbf{Y}}$ are positive semidefinite matrices. We know that the product of the positive semidefinite matrices is also a positive semidefinite matrix and the Hadamard product of positive semidefinite matrices is also a positive semidefinite matrix. Therefore, we can conclude

that the Hessian $\partial^2 D/\partial \mathbf{p}'^2$ is also positive semidefinite which establish, the convexity of the function in (11) with respect to the transmit powers P'_i , $i = 1, \dots, N$.

Acknowledgments

The authors would like to thank the Walloon Region Ministry DGTRE framework program COSMOS/Tsarine and EU project FP7 NEWCOM++ for the financial support and the scientific inspiration.

References

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.
- [2] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Elsevier Journal on Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, 2005.
- [3] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 887–902, 1996.
- [4] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. 43, no. 5, pp. 1549–1559, 1997.
- [5] J. Duplicy and L. Vandendorpe, "Robust MMSE precoding for the MIMO complex Gaussian broadcast channel," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, pp. 421–424, Honolulu, Hawaii, USA, April 2007.
- [6] D. P. Palomar, *A unified framework for communications through MIMO channels*, PhD Dissertation, Universitat Politècnica de Catalunya, Spain, May 2003.
- [7] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Transactions on Information Theory*, vol. 54, no. 11, 2008.
- [8] M. Gastpar, M. Vetterli, and P. L. Dragotti, "Sensing reality and communicating bits: a dangerous liaison," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 70–83, 2006.
- [9] P. L. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms and Applications*, Elsevier, Amsterdam, The Netherlands, 2009.
- [10] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Elsevier Journal on Computer Networks*, vol. 45, no. 3, pp. 245–259, 2004.
- [11] F. Davoli, M. Marchese, and M. Mongelli, "A decision theoretic approach to Gaussian sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '09)*, Genova, Italy, June 2009.
- [12] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1146–1156, 2008.
- [13] T. Wimalajeewa and S. K. Jayaweera, "Optimal power scheduling for correlated data fusion in wireless sensor networks via constraint PSO," *IEEE Transaction on Wireless Communications*, vol. 7, no. 9, pp. 3608–3618, 2008.
- [14] J.-J. Xiao, Z.-Q. Luo, S. Cui, and A. J. Goldsmith, "Power-efficient analog forwarding transmission in an inhomogeneous gaussian sensor network," in *Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless*

- Communications (SPAWC '05)*, pp. 121–125, Minneapolis, Minn, USA, June 2005.
- [15] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor, “Estimation diversity and energy efficiency in distributed sensing,” *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.
- [16] H. Senol and C. Tepedelenlioglu, “Performance of distributed estimation over unknown parallel fading channels,” *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 6057–6068, 2008.
- [17] K. Bai, C. Tepedelenlioglu, and H. Şenol, “Outage diversity for distributed estimation over parallel fading channels,” in *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, pp. 863–867, October 2008.
- [18] G. Thatte and U. Mitra, “Sensor selection and power allocation for distributed estimation in sensor networks: beyond the star topology,” *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2649–2661, 2008.
- [19] J.-J. Xiao and Z.-Q. Luo, “Multiterminal source-channel communication over an orthogonal multiple-access channel,” *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3255–3264, 2007.
- [20] H. Behroozi and M. R. Soleymani, “Source-channel communication in one-helper problem,” in *Proceedings of the 64th IEEE Vehicular Technology Conference (VTC '06)*, pp. 1748–1752, Quebec, Canada, September 2006.
- [21] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, 2005.
- [22] J. O. Berger, V. D. Oliveira, and B. Sansó, “Objective bayesian analysis of spatially correlated data,” *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1361–1374, 2001.
- [23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 1993.
- [24] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*, Technical University of Denmark, 2008.
- [25] “Special theme: towards green ICT,” *ERCIM News*, no. 79, October 2009.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2008.
- [27] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass, USA, 2nd edition, 1999.
- [28] D. P. Palomar and M. Chiang, “A tutorial on decomposition methods for network utility maximization,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, 2006.
- [29] J.-F. M. Barthelemy and R. T. Haftka, “Approximation concepts for optimum structural design—a review,” *Springer Journal on Structural and Multidisciplinary Optimization*, vol. 5, no. 3, pp. 129–144, 1993.
- [30] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, Belmont, Mass, USA, 1997.
- [31] M. Ronó, “Numerical-analytic successive approximation method for non-linear boundary value problems,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 30, no. 5, pp. 3179–3188, 1997.