Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Power-Efficient Estimation in Sensor Networks with Correlated Data: Perfect and Imperfect CSI

Muhammad Hafeez Chaudhary and Luc Vandendorpe
ICTEAM Institute, Université Catholique de Louvain, 1348 Louvain-La-Neuve, Belgium
{Muhammad.Chaudhary, Luc.Vandendorpe}@uclouvain.be

Abstract— **This paper considers power allocation problem in wireless sensor networks where distributed sensors amplify and forward their observations of a Gaussian random source to a remote fusion center (FC) which reconstructs the underlying source. The sensor networks are characterized by the availability of limited energy. Motivated by this fact, we design a power allocation scheme where our objective is to minimize the network power consumption such that the reconstruction distortion does not exceed a target value. The reconstruction distortion is quantified based on linear minimum mean-squared error (LMMSE) estimation rule. For power allocation, we propose a novel design based on successive approximation of the distortion function. The resulting algorithm turns out to be simple, computationally efficient and exhibits good convergence properties. The design is based on perfect knowledge of fading channel gains. We also address the case where only estimates of the channel gains are available. The simulation examples illustrate that the proposed design holds considerable performance gain compared to a uniform power allocation scheme.**

I. Introduction

Wireless sensor networking is an emerging technology which finds application in many fields including environment and habitat monitoring, health care, automation and military applications [1]. A wireless sensor network (WSN) consists of spatially distributed sensors that cooperatively observe the underlying physical phenomenon. Each sensor in the network is powered by a battery which has limited capacity.

The batteries are assumed to be non-rechargeable and irreplaceable because the nodes may be deployed in a hostile or inaccessible terrain. Our objective is to reconstruct the underlying source so that the total transmit power be minimized and the overall distortion (e.g. mean squared error) may not exceed a set target value. We consider a system in star topology where individual sensors transmit their noisy observations of a common source, via some orthogonal multiple access scheme, to a central processing unit called fusion center (FC), which produces a global picture of the physical phenomenon based on linear minimum mean squared error (LMMSE) estimation rule.

The sensors have partial and correlated observations of a common source. The correlation exists where sensors measure data in the same geographical location, e.g. 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 a uniform power allocation scheme is not an optimal strategy.

In a sensor network measuring a memoryless Gaussian source uncoded transmission, i.e. amplify and forward (AF), outperforms the separate coding and transmission over the multiple-access channel [2], [3]. Motivated by this result, Vuran et al. in [4] considered the estimation of a random source with distributed sensors and suggested a sensor selection procedure that exploits spatial correlation and minimizes the estimation error (based on LMMSE estimation criterion). The work assumes ideal communication channels to the FC. The sensor selection procedure suggests that 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 [5] which is based on the same topological settings as [4]. Bahceci and Khandani in [6] proposed a power allocation scheme for estimation where each sensor observes a separate source albeit correlated. Ref. [7] presented a power scheduling scheme for sensor networks to detect a source based on a binary hypothesis testing rule which exploits correlation in the observation noises at the sensors. Other works like [8]- [12] proposed power scheduling schemes for the sensor networks without considering the spatial correlation.

In this contribution, we present a novel framework which incorporates adaptive power allocation 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 allocations are jointly determined at the FC which are then conveyed to individual sensors via feedback channels. Due to the correlation among sensor observations, the design of power allocation scheme based on the given optimization problem presents a unique challenge because the LMMSE estimation error of the underlying source contains nonlinearly coupled optimization variables. To this end, we propose a novel design based on a successive approximation of the estimation distortion. The resulting power allocation algorithm is simple, computationally efficient and exhibits excellent convergence behavior. The proposed design holds considerable performance gain compared to a uniform power allocation scheme.

The gains of the communication channels from the sensors

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.

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 at the FC and the limited available power means that the channel estimation always incurs some estimation error. Consequently, the design of power allocation schemes should also take into account the channel estimation errors [13], [14]. Herein, first we design the power allocation scheme assuming perfect knowledge of the channel state information (CSI) and subsequently we extend the design to the case of imperfect CSI. 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 the spatial correlation, the observation quality, and the channel gains and their estimation errors.

The rest of the paper is organized as follows. Section II describes the setup of the sensor network and formulates the power allocation problem. Section III and Section IV respectively outlines solution to the optimization problem for perfect and imperfect CSI cases. Section V evaluates the performance of the power allocation designs with some simulation examples. Finally Section VI concludes the work.

II. System Model

Consider the system model shown in Fig. 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, e.g. 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

$$x_i(t) = s_i(t) + n_i(t),\ i = 1, \dots, N, \tag{1}$$

where $n_i, \forall i$ is independent across the sensors. The estimation of the source is done on a sample by sample bases, and its procedure is same for all samples. Therefore, in the subsequent formulation we drop the time index for clarity. 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 received signal at FC from sensor i is

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

where $\sqrt{P_i'}$ is a amplifying factor and $\bar{w}_i \sim \mathcal{CN}(0, \sigma_{\bar{w}_i}^2)$ is receiver noise which is independent across the sensors and is also independent of $\{n_i\}_{i=1}^N$. The channels $\{h_i\}_{i=1}^N$ between the sensors and the FC experience independent flat fading with $h_i \sim \mathcal{CN}(0, \sigma_{h_i}^2), \forall i$ and the gain factors $\{g_i = |h_i|\}_{i=1}^N$ which are Rayleigh distributed. Noting that $h_i = g_i e^{-j\theta_{h_j}}$, we can write from (2) as follows:

$$\bar{z}_i e^{j\theta_{h_j}} = g_i \sqrt{P_i'}\,(s_i + n_i) + \bar{w}_i e^{j\theta_{h_j}},$$

where the exponential term $e^{j\theta_{h_j}}$ can be merged into the Gaussian variable $\bar{w}_i$ without changing its statistical properties due to the circular-symmetry property of $\bar{w}_i$ [15]. 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

$$z_i = g_i \sqrt{P_i'}\,(s_i + n_i) + w_i,\ \forall i, \tag{3}$$

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

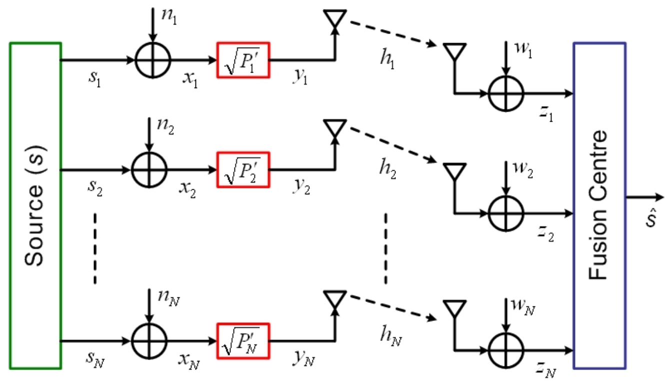

Fig. 1: Block diagram of the system.

For the analysis in this work, 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

$$\Lambda = \begin{pmatrix} \sigma_s^2 & \sigma_s\sigma_{s_i}\rho_{si} & \sigma_s\sigma_{s_j}\rho_{sj} & 0 & 0 \\ \sigma_s\sigma_{s_i}\rho_{si} & \sigma_{s_i}^2 & \sigma_{s_i}\sigma_{s_j}\rho_{ij} & 0 & 0 \\ \sigma_s\sigma_{s_j}\rho_{sj} & \sigma_{s_i}\sigma_{s_j}\rho_{ij} & \sigma_{s_j}^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{n_i}^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{w_i}^2 \end{pmatrix}. \tag{4}$$

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

In (4), 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 may 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

$$\rho_{si} = \frac{cov\{S, S_i\}}{\sigma_s\sigma_{s_i}} = e^{-\left(d_{si}/\theta_1\right)^{\theta_2}}, \tag{5a}$$

$$\rho_{ij} = \frac{cov\{S_i, S_j\}}{\sigma_{s_i}\sigma_{s_j}} = e^{-\left(d_{ij}/\theta_1\right)^{\theta_2}}, \tag{5b}$$

which is a power exponential model for correlation [4], [16]. In (5), d_{si} is the distance between the source s and the sensor i, 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 $\theta_2 \in (0, 2]$ is called a smoothness or roughness parameter. Eq. (5) shows that the correlation decays with distance with limiting values of 1 and 0 as $d_{si}(d_{ij}) \to 0$ and $d_{si}(d_{ij}) \to \infty$ respectively.

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 the 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 a two or three dimensional space. We assume that the relative positions of the source and the sensors are known and they do not change for at least one estimation cycle.

The sensor networks are characterized by the availability of limited power and often it is desired to achieve a specified distortion by spending as little power as possible. The sensor nodes are battery powered and the life of the nodes and consequently of the network depends on the battery life. Therefore, by minimizing the power consumption in each estimation cycle, we may prolong the life of the network. Moreover, minimizing the total power consumption corresponds to minimizing the contribution to the phenomenon of global warming [17]. To this purpose, we base our power allocation scheme on the following optimization problem:

$$(\textbf{Prob}): \textit{minimize the total power } P_{tot} = \sum_{i=1}^{N} P_i' \sigma_i^2$$

$$\begin{aligned} \textit{subject to} \quad & D \le D_{\max}, \\ & P_i' \ge 0, \quad \forall i, \end{aligned} \tag{6}$$

where $\sigma_i^2 := \sigma_{s_i}^2 + \sigma_{n_i}^2$, D defines the reconstruction distortion and $D_{\max}$ is its target value. Note that $P_i := P_i' \sigma_i^2$ denotes the total transmit power of sensor i. In the sequel, first we characterize the distortion D, which is measured as mean-squared estimation error, and subsequently we solve the optimization problem under the assumption of: (i) perfect knowledge of the channel gains and (ii) estimates of the channel gains with estimation errors.

III. 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.

At the FC the optimal estimate $\hat{s}$ of the source s in minimum mean-squared error (MMSE) sense is equal to the mean of s given the observations $\{z_i\}_{i=1}^N$, that is $\hat{s} = \mathbb{E}[s|z_i, \forall i]$, where $\mathbb{E}$ denotes the mathematical expectation. Under the jointly Gaussian assumption of s and $\{z_i\}_{i=1}^N$, the conditional mean estimator turns out to be linear and is called linear minimum mean-squared error (LMMSE) estimator [18] which can be written as follows:

$$\hat{s} = \mathbf{c}^T \mathbf{H} \left(\mathbf{HCH} + \mathbf{C_w}\right)^{-1} \mathbf{z}, \tag{7}$$

and the associated mean-squared estimation distortion can be written as follows:

$$\begin{aligned} D &= \sigma_s^2 - \mathbf{c}^T \mathbf{H} \left(\mathbf{HCH} + \mathbf{C_w}\right)^{-1} \mathbf{Hc}, \\ &= \sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \mathbf{c}^T \left(\mathbf{YC} + \mathbf{I}\right)^{-1} \mathbf{C}^{-1} \mathbf{c}, \end{aligned} \tag{8}$$

where $\mathbf{z} = [z_1, \ldots, z_N]^T$, $\mathbf{H} = \text{diag}\left(g_1\sqrt{P_1'}, \ldots, g_N\sqrt{P_N'}\right)$, $\mathbf{C} = \mathbb{E}[(\mathbf{s_i} + \mathbf{n})(\mathbf{s_i} + \mathbf{n})^T]$, $\mathbf{s_i} = [s_1, \ldots, s_N]^T$, $\mathbf{n} = [n_1, \ldots, n_N]^T$, $\mathbf{C_w} = \mathbb{E}[\mathbf{ww}^T]$, $\mathbf{w} = [w_1, \ldots, w_N]^T$ and $\mathbf{Y} = \mathbf{HC_w^{-1}H} = \text{diag}\left(\frac{g_1^2 P_1'}{\sigma_{w_1}^2}, \ldots, \frac{g_N^2 P_N'}{\sigma_{w_N}^2}\right)$ with $[\ldots]^T$ denoting the vector/matrix transpose operation [19].

Remark-1: *The reconstruction distortion D is upper bounded by the variance of the underlying source and is lower bounded by the variances of the observation noises and the spatial correlation and cross-correlation values as given by*

$$\sigma_s^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} \le D \le \sigma_s^2. \tag{9}$$

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 $\frac{g_i^2 P_i'}{\sigma_{w_i}^2} = \infty$, $\forall i$ in (8). 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 $D_0 = \sigma_s^2 - \mathbf{c}^T \mathbf{C}_{\mathbf{s}_i}^{-1} \mathbf{c}$. The achieved distortion is equal to the upper bound value (i.e. $D = \sigma_s^2$) when either no signal is received at FC from the sensors or the observations are uncorrelated with the source s or both. Based on the Remark-1, following remark is in order

Remark-2: *If $D_{max} < D_0$ then no matter how high the transmit powers of the sensors are, we cannot achieve this distortion and the given optimization problem is strictly infeasible. Moreover, to achieve $D_{max} = D_0$ requires $\{P_i\}_{i=1}^N = \infty$ and consequently the problem is infeasible. On the other hand, if $D_{max} \ge \sigma_s^2$ then the problem is trivial and the distortion can be achieved by setting $\hat{s} = \mathbb{E}[s]$.*

With the distortion function given in (8) We can show that the optimization problem (6) is convex and to solve the problem we can use Lagrangian multipliers technique [20]. The Lagrangian function can be written as

$$f(P_i', \lambda, \mu_i) = \lambda(D - D_{\max}) + \sum_{i=1}^{N} \left(P_i' \sigma_i^2 - \mu_i P_i'\right), \tag{10}$$

where λ and $\{\mu_i\}_{i=1}^N$ are Lagrangian multipliers. Since the problem is convex, therefore following Karush-Kuhn-Tucker (KKT) conditions are sufficient to ensure the optimality of the solution.

$$\begin{aligned} \frac{\partial f}{\partial P_i'} = -\lambda \left(\frac{g_i^2}{\sigma_{w_i}^2} \mathbf{c}^T \left(\mathbf{YC} + \mathbf{I}\right)^{-1} \mathbf{J}_i \mathbf{C} \left(\mathbf{YC} + \mathbf{I}\right)^{-1} \mathbf{C}^{-1} \mathbf{c} \right) \\ + \sigma_i^2 - \mu_i = 0, \ \forall i, \end{aligned} \tag{11}$$

$$\lambda \left(D - D_{\max}\right) = 0, \ \lambda \ge 0, \ D \le D_{\max}, \tag{12}$$

$$\mu_i P_i' = 0, \ \mu_i \ge 0, \ P_i' \ge 0, \ \forall i, \tag{13}$$

where $\mathbf{J}_i$ is a diagonal matrix with unity at (i,i)th place and all other elements equal to zero. Unfortunately, to find the solution for $\{P_i'\}_{i=1}^N$, we have to numerically solve (11)-(13). This numerical approach may be computationally quite expensive due to the matrix inversions involved in (11). Next we seek answer to the following question. Can we do something different that is simpler and computationally efficient than the numerical approach? In the sequel, in answering this question, we propose a novel power allocation design and evaluate its performance.

The LMMSE estimate of the source can be written as $\hat{s} = \sum_{i=1}^{N} a_i z_i$ with a_i's denoting the LMMSE weighting coefficients. The resultant mean-squared distortion D can be written as follows:

$$\begin{aligned} D &= \mathbb{E}_{\{s,s_i,n_i,w_i|g_i,\forall i\}}\left[(s-\hat{s})^2\right], \\ &= \sigma_s^2 + \sum_{i=1}^{N} a_i^2\left(g_i^2 P_i' \sigma_i^2 + \sigma_{w_i}^2\right) - 2\sum_{i=1}^{N} a_i g_i \sqrt{P_i'}\sigma_s\sigma_{s_i}\rho_{si} \\ &\quad + \sum_{i=1}^{N}\sum_{j\neq i}^{N} a_i a_j g_i g_j \sqrt{P_i' P_j'}\sigma_{s_i}\sigma_{s_j}\rho_{ij}. \end{aligned} \quad (14)$$

By solving $\frac{\partial D}{\partial a_i} = 0$ we get the following expression for the LMMSE weighting coefficients:

$$a_i = \beta_i \gamma_i, \quad \forall i, \quad (15)$$

where β_i and γ_i is respectively defined as

$$\gamma_i = \frac{g_i\sqrt{P_i'}}{g_i^2 P_i' \sigma_i^2 + \sigma_{w_i}^2}, \quad \forall i, \quad (16)$$

$$\beta_i = \sigma_s\sigma_{s_i}\rho_{si} - \sum_{j\neq i}^{N} \beta_j\gamma_j g_j\sqrt{P_j'}\sigma_{s_i}\sigma_{s_j}\rho_{ij}, \quad \forall i. \quad (17)$$

With (16) and (17), the distortion in (14) simplifies to

$$\begin{aligned} D &= \sigma_s^2 - \sum_{i=1}^{N}\beta_i\gamma_i g_i\sqrt{P_i'}\sigma_s\sigma_{s_i}\rho_{si}, \\ &= \sigma_s^2 - \sum_{i=1}^{N}\frac{g_i^2 P_i'}{g_i^2 P_i'\sigma_i^2 + \sigma_{w_i}^2}\beta_i\sigma_s\sigma_{s_i}\rho_{si}. \end{aligned} \quad (18)$$

Eq. (17) forms a set of N coupled equations which constitute the Wiener-Hopf equation for the LMMSE filter coefficients (β_i, $\forall i$). If we know the transmit powers $\{P_i'\}_{i=1}^N$ then for given spatial covariance $\Lambda_{s,s_i,s_j,n_i,w_i}$ and the channel gains $\{g_i\}_{i=1}^N$, we can find the optimal coefficients $\{\beta_i\}_{i=1}^N$ by solving (17). To this end, it is convenient to employ the matrix-vector notation as follows:

$$\beta = \mathbf{R}^{-1}\mathbf{c}, \quad (19)$$

where $\beta = [\beta_1,\ldots,\beta_N]^T$, $\mathbf{c} = [\sigma_s\sigma_{s_1}\rho_{s1},\ldots,\sigma_s\sigma_{s_N}\rho_{sN}]^T$, $[\mathbf{R}]_{(ij,i\neq j)} = \gamma_j g_j\sqrt{P_j'}\sigma_{s_i}\sigma_{s_j}\rho_{ij}$ and $[\mathbf{R}]_{(ij,i=j)} = 1$. Note that $[\mathbf{R}]_{(i,j)}$ denotes the (i,j)th element of $\mathbf{R}$.

In the ensuing development, we outline an alternative solution to the problem of optimizing the transmit powers of the sensors such that the achieved distortion does not exceed the target value. Therein to derive an algorithm for the solution of P_i' the underlying idea is to assume β as given. Based on this assumption we derive an iterative algorithm for power allocation which computes β using the values of $\{P_n'\}_{n=1}^N$ from the previous iteration. This successive approximation (SA) of the distortion function in (18) makes the solution of the power allocation problem simple and easy to compute. Note that the resulting design for power allocation can be viewed as a joint optimization of $\{\beta_i\}_{i=1}^N$ and $\{P_i'\}_{i=1}^N$.

We use the outlined successive-approximation principle to solve the optimization problem (6) based on the distortion function D in (18). For given $\beta_i > 0$, it is easy to verify that the distortion constraint is a convex function of the optimization variables P_i', $i = 1,\ldots,N$ and since the objective function is linear, therefore the given optimization problem is convex. To solve the problem, we adopt the Lagrangian multipliers technique where KKT conditions are sufficient for optimality of the solution. The associated Lagrangian function can be written as follows:

$$f(P_i',\nu,\mu_i) = \nu(D - D_{\max}) + \sum_{i=1}^{N}\left(P_i'\sigma_i^2 - \mu_i P_i'\right),$$

where ν and $\{\mu_i\}_{i=1}^N$ are Lagrangian multipliers. The KKT conditions are

$$\frac{\partial f}{\partial P_i'} = -\nu\frac{\sigma_{w_i}^2 g_i^2\beta_i\sigma_s\sigma_{s_i}\rho_{si}}{\left(g_i^2 P_i'\sigma_i^2 + \sigma_{w_i}^2\right)^2} + \sigma_i^2 - \mu_i = 0, \forall i, \quad (20)$$

$$\nu(D - D_{\max}) = 0,\ \nu \geq 0,\ D \leq D_{\max}, \quad (21)$$

$$\mu_i P_i' = 0,\ \geq 0,\ P_i' \geq 0,\ \forall i. \quad (22)$$

For sensor i to be active $P_i > 0$, in which case solving (20) and (22) we get

$$P_i' = \frac{1}{\zeta_i\sigma_i^2}\left(\sqrt{\frac{\nu\zeta_i\beta_i\sigma_s\sigma_{s_i}\rho_{s_i}}{\sigma_i^2}} - 1\right)^+, \forall i, \quad (23)$$

where $\zeta_i := \frac{g_i^2}{\sigma_{w_i}^2}$ defines channel SNR for sensor i and $(x)^+ = \max(x,0)$. From (23), we can make the following observations about the power allocation policy for the sensors.

1) With respect to the channel SNR, in the range $\zeta_i \leq \zeta_i^{(o)}$, P_i' increases with increasing ζ_i and in the range $\zeta_i > \zeta_i^{(o)}$, P_i' decreases with increasing ζ_i, where $\zeta_i^{(o)} = \frac{4\nu\beta_i\sigma_s\sigma_{s_i}\rho_{si}}{\sigma_i^2}$.
2) Considering the observation noise variance, sensor i gets more power if $\sigma_{n_i}^2$ (i.e. σ_i^2) is less and vice versa.
3) Better correlation with the source results in higher power assigned to the sensor and vice versa. For example, if $\rho_{si} \to 0$ then $P_i' \to 0$.

These observations show that the final power allocation policy jointly depends on the correlation values, observation noise variances and the channel SNRs and depending on the values of these parameters some of the sensors may be switched-off. Note that for a sensor i to be active following condition

$$\frac{\zeta_i\beta_i\sigma_{s_i}\rho_{si}}{\sigma_i^2} > \frac{1}{\nu\sigma_s}, \quad (24)$$

must hold, otherwise it is switched-off. Let $\mathcal{L}$ defines the index-set of the active sensors as follows:

$$\mathcal{L} = \left\{l \middle| \frac{\zeta_l\beta_l\sigma_{s_l}\rho_{sl}}{\sigma_l^2} > \frac{1}{\nu\sigma_s}\right\}. \quad (25)$$

Since the given problem is convex, therefore the minimum of the objective function occurs at the distortion constraint boundary, which means that the constraint has to be tight, that is $D = D_{\max}$ giving

$$\nu = \left(\frac{\sum_{l\in\mathcal{L}} \sqrt{\frac{\beta_l \sigma_s \sigma_{s_l}}{\zeta_l \sigma_l^2}}}{D_{\max} - \sigma_s^2 + \sum_{l\in\mathcal{L}} \frac{\beta_l \sigma_s \sigma_{s_l} \rho_{sl}}{\sigma_l^2}} \right)^2. \tag{26}$$

The outlined solution for power allocation is similar to the waterfilling-like solution [9].

Based on (19) and (23)-(26), to find power allocation for the sensors, an iteratively algorithm can be proposed as outlined under "Algorithm 1". The algorithm successively optimizes $\{P_i'\}_{i=1}^N$ and $\{\beta_i\}_{i=1}^N$ while minimizing the total power subject to the distortion constraint. If during iterations any sensor does not fulfill the condition in (24), it is switched off and the algorithm continues with the remaining sensors until the convergence criterion is fulfilled. Note that for given $\{P_i'\}_{i=1}^N$, optimal $\{\beta_i\}_{i=1}^N$ can be computed by the simple equation (19) and for given $\{\beta_i\}_{i=1}^N$, optimal $\{P_i'\}_{i=1}^N$ is given by the waterfilling-like solution (23)-(26).

Algorithm 1

```
Initialize P'_i^[0] for i = 1, ..., N
Calculate β_i^[0] for i = 1, ..., N
Set κ = 0
while ( |P_tot^[κ] − P_tot^[κ−1]| ≥ ϵ ) do   ▷ where κ denotes the while loop iteration index
    κ = κ + 1
    For i = 1, ..., N determine transmit power as follows:
    if Condition in (24) is true then
        Determine P'_i^[κ] from (23)
    else
        P'_i^[κ] = 0
    end if
    For i = 1, ..., N update β_i^[κ] from (19)
    Calculate P_tot^[κ]
end while
```

Regarding the convergence properties of the algorithm consider the following points. Since in each iteration we are minimizing a convex function over the convex-set of the transmit powers $\{P_i' | D \le D_{\max} \text{ and } P_i' \ge 0,\ i = 1, \dots, N\}$, 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 the $\{\beta_i\}_{i=1}^N$ for given $\{P_i'\}_{i=1}^N$ mean that the algorithm achieves monotonic decrease in P_{tot} and consequently it does converge to a unique minimum point. Moreover, the algorithm may achieve convergence in as few as two or three iterations. In Section V, we illustrate the convergence properties of the algorithm with several simulation examples. It is quite interesting that the algorithm exhibits such convergence properties which illustrates that the proposed successive approximation strategy works quite well. Finally the ease of computation and simplicity of the design based on the successive approximation principle can be appreciated from the simple and concise structure of (23)-(26).

IV. 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 denotes 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 MMSE estimate $\hat{h}_i$ of h_i is

$$\hat{h}_i = \frac{\mathbb{E}_{\{h_i, \bar{w}_i\}}[h_i r_i^*]}{\mathbb{E}_{\{h_i, \bar{w}_i\}}[|r_i|^2]} r_i = \frac{\sigma_{h_i}^2 t_i^*}{\sigma_{h_i}^2 |t_i|^2 + \sigma_{\bar{w}_i}^2} r_i, \tag{27}$$

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

$$\bar{\delta}_i^2 = \mathbb{E}_{\{h_i, \bar{w}_i\}} \left[|h_i - \hat{h}_i|^2 \right] = \frac{\sigma_{h_i}^2 \sigma_{\bar{w}_i}^2}{\sigma_{\bar{w}_i}^2 + \sigma_{h_i}^2 |t_i|^2}, \tag{28}$$

wherein $|t_i|^2$ denotes the 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

$$h_i = \hat{h}_i + \Delta\bar{h}_i, \tag{29}$$

where $\Delta\bar{h}_i \sim \mathcal{CN}(0, \bar{\delta}_i^2)$.

One way to design the power-scheduling scheme is by replacing h_i by $\hat{h}_i$ and g_i by $\hat{g}_i$ in the formulation in the foregoing section. This constitutes the naive-approach because it ignores the error in the channel estimate. An alternative design originates by substituting (29) in (2) as follows:

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

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

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

where the exponential term $e^{-j\theta_{\hat{h}_i}}$ can be absorbed into $\bar{u}_i$, i.e. 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 part of the noise $\bar{u}_i$ term in-phase with the sensor observation is relevant for estimation of the source s . Therefore, we can write the received signal from sensor i as

$$z_i = \hat{g}_i \sqrt{P_i'}\,(s_i + n_i) + u_i, \tag{30}$$

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 III, the mean-squared reconstruction distortion of the estimate $\hat{s} = \sum_{i=1}^{N} a_i z_i$ with respect to s can be written as follows:

$$\tilde{D} = \mathbb{E}_{\{s,s_i,n_i,w_i,\Delta h_i,|\hat{g}_i,\forall i\}}\left[(s-\hat{s})^2\right], \\ = \sigma_s^2 - \sum_{i=1}^{N} \beta_i \gamma_i \hat{g}_i \sqrt{P_i'} \sigma_s \sigma_{s_i} \rho_{si}, \quad (31)$$

where

$$a_i = \beta_i \gamma_i, \\ \beta_i = \sigma_s \sigma_{s_i} \rho_{si} - \sum_{j \neq i}^{N} \beta_j \gamma_j \hat{g}_j \sqrt{P_j'} \sigma_{s_i} \sigma_{s_j} \rho_{ij}, \\ \gamma_i = \frac{\hat{g}_i \sqrt{P_i'}}{\hat{g}_i^2 P_i' \sigma_i^2 + P_i' \sigma_i^2 \delta_i^2 + \sigma_{w_i}^2}, \quad \forall i. \quad (32)$$

The solution to the optimization problem in (6) with the objective to minimize the total network power subject to the constraint on the distortion $\tilde{D}$ can be obtained by using the method of Lagrangian multipliers. The resulting solution is outlined as follows.

1) The power allotted to sensor k for $k = 1 \ldots, N$ is

$$P_k' = \frac{1}{\left(1+\frac{\delta_k^2}{\hat{g}_k^2}\right)\hat{\zeta}_k \sigma_k^2}\left(\sqrt{\frac{\nu \hat{\zeta}_k \beta_k \sigma_s \sigma_{s_k} \rho_{sk}}{\sigma_k^2}} - 1\right)^+, \quad (33)$$

where $\hat{\zeta}_k := \frac{\hat{g}_k^2}{\sigma_{w_i}^2}$.

2) For sensor k to have $P_k' > 0$ following condition must hold

$$\frac{\hat{\zeta}_k \beta_k \sigma_{s_k} \rho_{sk}}{\sigma_k^2} > \frac{1}{\nu \sigma_s}, \quad (34)$$

otherwise $P_k' = 0$.

3) The index-set of active sensors is

$$\mathcal{L} = \left\{ l \middle| \frac{\hat{\zeta}_l \beta_l \sigma_{s_l} \rho_{sl}}{\sigma_l^2} > \frac{1}{\nu \sigma_s} \right\}. \quad (35)$$

4) The Lagrangian multiplier ν is determined by satisfying the distortion constraint with equality and is given by

$$\nu = \left(\frac{\sum_{k\in\mathcal{L}} \sqrt{\frac{\beta_k \sigma_s \sigma_{s_k} \rho_{sk}}{\left(1+\frac{\delta_k^2}{\hat{g}_k^2}\right)^2 \hat{\zeta}_k \sigma_k^2}}}{D_{\max} - \sigma_s^2 + \sum_{k\in\mathcal{L}} \frac{\beta_k \sigma_s \sigma_{s_k} \rho_{sk}}{\left(1+\frac{\delta_k^2}{\hat{g}_k^2}\right)\sigma_k^2}} \right)^2. \quad (36)$$

Based on (33)-(36), power allocation for the sensors can be obtained using the procedure outlined under Algorithm 1. Note that the convergence properties of the algorithm with perfect CSI also applies to the imperfect CSI case.

Remark-3: *We can observe that as $\delta_k^2 \rightarrow 0$ then $\hat{g}_k \rightarrow g_k$ for $k = 1, \ldots, N$ and the power allocation design with imperfect CSI* (33)-(36) *converges to the design with perfect CSI* (23)-(26) *respectively.*

V. Performance Evaluation and Discussion

This section corroborates the analytical findings with some simulation examples. In order to show the effectiveness 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, for the UPA design, we have $\{P_i\}_{i=1}^N = P_u$. In the figures, we assume $\log(.) = \log_{10}(.)$.

We consider two sensor networks respectively comprising $N = 3$ and $N = 500$ sensors that are uniformly distributed in a 100×100 planar region with the source s at its center. Fig. 2 plots the total power $P_{tot} = \sum_{i=1}^N P_i$ consumed by the APA design as a percentage of the power consumed by the UPA design to achieve the given target distortion $D_{\max}$. The power P_{tot} is averaged over 10^4 realizations of the sensors deployment. To ensure the feasibility of the problem, i.e. $\sigma_s^2 < D_{\max} < D_0$, and to enable a fair comparison, we set the target distortion according to the following rule

$$D_{\max}(m) = \sigma_s^2 - \frac{m}{M}\left\{\sigma_s^2 - (1+\epsilon)D_0\right\}, \; 1 \leq m \leq M, \quad (37)$$

in each deployment of the sensors. In (37), M is an integer greater than or equal to 1 and $\epsilon \in (0, 1)$. The value of ϵ should be sufficiently small such that $(1+\epsilon)D_0 < \sigma_s^2$.

The figure shows that our proposed design outperforms the UPA scheme and the APA design may require transmit power as little as 0.01 percent of that needed by the UPA scheme to achieve the target distortion. For given θ_1, we can observe that the performance gap between the APA and UPA designs increases with increasing the number of sensors. Moreover, we can see that increasing the value of θ_1 for given N, the performance gap between the APA and UPA scheme decreases. 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. (5)] and consequently the total power consumed by the APA design approaches to the power consumed by the UPA design.

Next for the sake of illustration, we focus on the network with three sensors, i.e. $N = 3$, and we consider the following example: $(d_{X_1}, d_{X_2}, d_{X_3}) = (-0.3, 0, 0.8)$ and $(d_{Y_1}, d_{Y_2}, d_{Y_3}) = (0, 1.6, 0)$, where (d_{X_i}, d_{Y_i}) specifies position of sensor i with respect to the origin in a XY-plane. Note that we can view this example as a specific realization of the deployment of the sensors. Assuming the source to be at the origin of the XY-plane and $\theta_1 = \theta_2 = 1$. From (5) we obtain the following spatial correlation values: $(\rho_{s_1}, \rho_{s_2}, \rho_{s_3}) = (0.7408, 0.2019, 0.4493)$ and $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.1963, 0.3329, 0.1671)$. The simulations in the sequel are based on this example. We have taken the example for purely illustrative purpose.

The simulation examples in the sequel assume $\{\sigma_{n_i}^2\}_{i=1}^N = 0.01$, $\{h_i^2\}_{i=1}^N = 1$ and $\{\sigma_{w_i}^2\}_{i=1}^N = 10$ unless stated otherwise. The spatial correlation values and the observation noise variances lower bounds the reconstruction distortion at $D_0 = 0.4086$. The target distortion D_{max} in the subsequent simulation examples cannot be below this value, otherwise the optimization problem will be strictly infeasible [c.f. Remark-2].

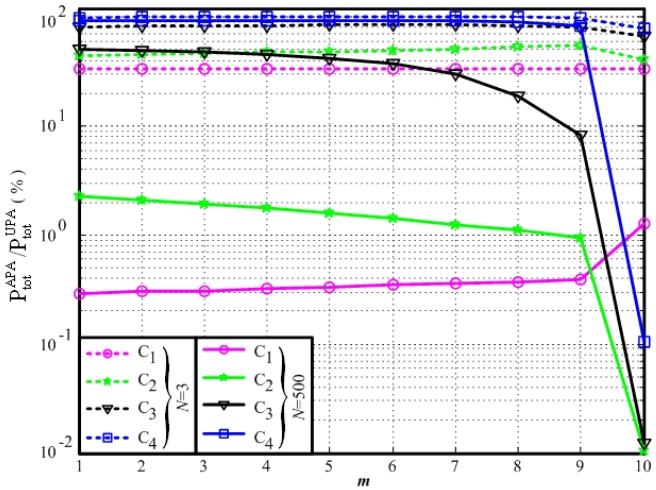

Fig. 2: $\epsilon = 10^{-5}$, $\theta_2 = 1$, $\left(\sigma_{n_1}^2,\, h_i,\, \sigma_{w_i}^2\right)_{i=1}^{N} = (0.01,\, 1,\, 10)$, and (C$_1$) $\theta_1 = 1$, (C$_2$) $\theta_1 = 10$, (C$_3$) $\theta_1 = 100$ and (C$_4$) $\theta_1 = 1000$.

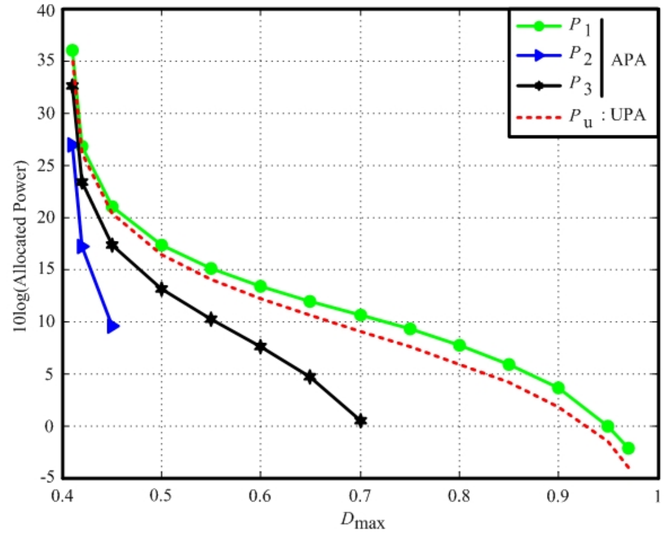

Fig. 3: Power Allocation.

Fig. 3 illustrates the power allotted to the sensors by the APA and UPA schemes. We can see that, to achieve the target distortion, the UPA scheme gives equal power to all sensors regardless of their correlation structure and thereby wasting power whereas the APA design allocates more power to the sensors with good correlation properties and vice versa. Moreover, depending on the target distortion, some of the sensors may be switched off. The corresponding comparison of the total powers of the two schemes is given in Fig. 4. The figure shows that, to achieve the target distortion, the APA design may require as little as fifty-percent of the power needed by the UPA scheme. Note that the difference in the powers of the two schemes depends on the correlation structure and the target distortion. Nevertheless, the proposed design substantially outperforms the UPA scheme in the required total power to achieve the given distortion.

Fig. 5 illustrates the convergence behavior of the Algorithm 1 for different values of $D_{\max}$. In each case, the algorithm is initialized with uniform power allocation such that the distortion is equal to $D_{\max}$. From the figure we can see that there is a monotonic decreases in P_{tot} and in two iterations it approaches fairly close to the minimum value. Nevertheless, after the third or forth iterations there is no appreciable change in the power.

Fig. 6 compares the total power required by the APA design as a percentage of the UPA scheme when we have: (i) perfect CSI, i.e. $\left\{\delta_i^2\right\} = \delta^2 = 0$ and (ii) the estimates of the channels with estimation error variance $\delta^2 = 0.01$ and $\delta^2 = 0.1$. In the simulations, we assume that the communication channels $\{h_i\}_{i=1}^{N}$ from the sensors to the FC undergo independent Rayleigh fading with $\left\{\sigma_{h_i}^2\right\}_{i=1}^{N} = 1$. Moreover, we assume that $\left\{\sigma_{n_i}^2, \sigma_{w_i}^2\right\}_{i=1}^{N} = (0.01,\, 10)$. The total power P_{tot} is averaged over 5×10^6 independent realizations of the channels for the APA and UPA designs respectively. The figure shows that the APA design outperforms the UPA scheme in terms of power efficiency for both perfect and imperfect CSI cases.

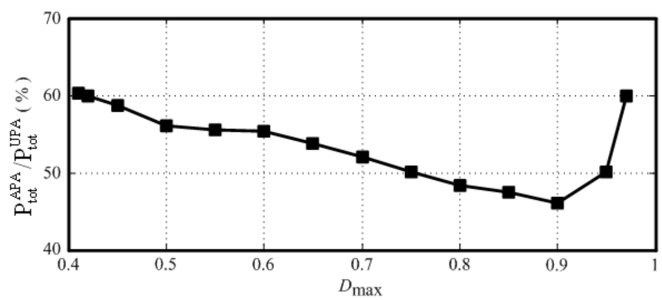

Fig. 4: Power efficiency comparison.

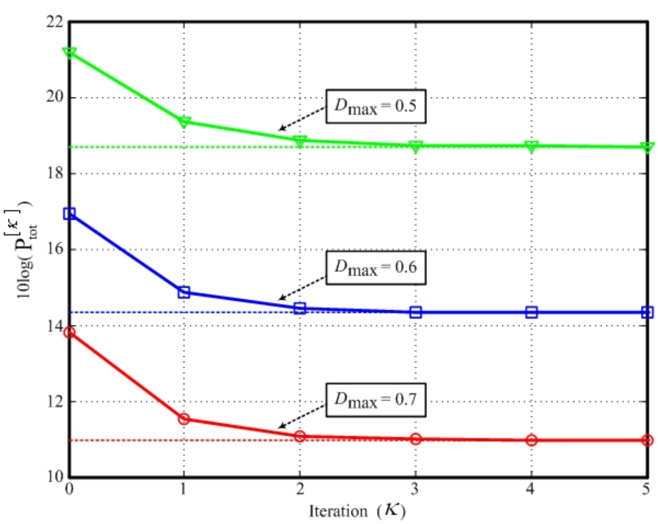

Fig. 5: Convergence behavior of Algorithm 1.

Moreover, the performance gap decreases with increasing the variance of the channel errors.

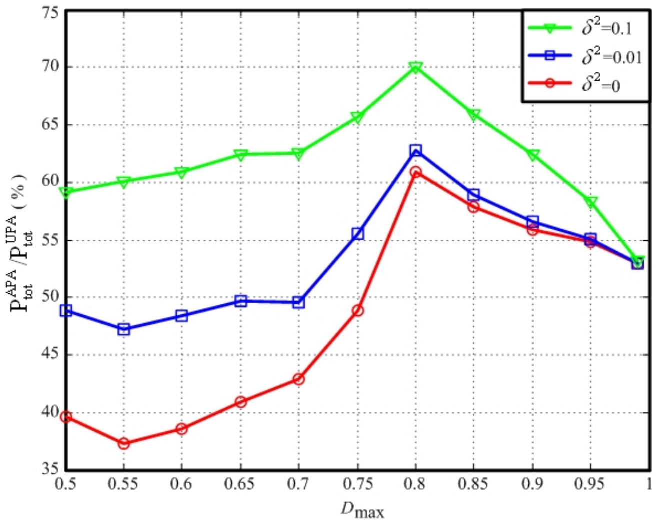

Fig. 6: Comparison of perfect and imperfect CSI.

VI. Conclusions

We have developed an adaptive power allocation design to minimize the network power consumption such that the fidelity criterion given by the maximum tolerable estimation distortion is satisfied. The proposed design simultaneously exploits the spatial correlation, the observation quality and the communication channel gains by incorporating these parameters in (quasi-)analytical expressions. Regarding knowledge of the channel state information, we considered two cases: perfect knowledge of the instantaneous channel gains is available or the estimates of the gains with estimation errors are known. The proposed design is novel and the associated algorithm is simple and easy to compute, and exhibits good convergence properties. We demonstrated the effectiveness of our proposed design with a few examples and we showed that the design outperforms the uniform power allocation scheme. The performance gain may be large for a relatively large sensor network with high heterogeneity, across the sensors, in the spatial correlation, the observation quality and the channel gains.

References

[1] I. F. Akyildiz, W. Su, A. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks 38, Elsevier Science B.V.*, pp. 393-422, 2002.

[2] M. Gastpar, "Uncoded Transmission is Exactly Optimal for a Simple Gaussian Sensor Network," *IEEE Trans. on Information Theory*, Vol. 54, No. 11, Novermber 2008.

[3] P. L. Dragotti and M. Gastpar, "Distributed Source Coding: Theory, Algorithms and Applications," *Elsevier Inc.*, 2009.

[4] M. C. Vuran and O. B. Akan and I. F. Akylidiz, "Spatio-Temporal Correlation: Theory and Applications for Wireless Sensor Networks," *Elsevier, Computer Networks*, Vol. 45, pp. 245-259, March 2004.

[5] F. Davoli, M. Marchese and M. Mongelli, "A Decision Theoretic Approach to Gaussian Sensor Networks," in *Proceedings of IEEE International Conference on Communications (ICC'09)*, June 2009.

[6] I. bahceci, and A. K. Khandani, "Linear Estimation of Correlated Data in Wireless Sensor Networks with Optimum Power Allocation and Analog Modulation," *IEEE Trans. on Communications*, Vol. 56, No. 7, pp. 1146-1156,July 2008.

[7] T. Wimalajeewa, S. K. Jayaweera "Optimal Power Scheduling for Correlated Data Fusion in Wireless Sensor Networks via Constrained PSO," *IEEE Trans. on Wireless Communications*, Vol. 7, No. 9, pp. 3608-3618, September 2008.

[8] J.-J. Xiao, S. Cui and A. J. Goldsmith "Power-efficient Analog Forwarding Transmission in an Inhomogenious Gaussian Sensor Network," in *Proceedings of 6th IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 121-125, June 2005.

[9] S. Cui, J-J Xiao and A. J. Goldsmith "Estimation Diversity and Energy Efficiency in Distributed Sensing," *IEEE Trans. on Signal Processing*, Vol. 55, No. 9, pp. 4683-4695, September 2007.

[10] K. Bai, C. Tepedelenlioğlu and H. Senol "Outage Diversity for Distributed Estimation over Parallel Fading Channels," in *Proceedingsof 42nd Asilomar Conference on Signals, Systems and Computers*, pp- 863-867, October 2008.

[11] J. Y. Wu, Q. Z. Huang and T. S. Lee, "Minimal Energy Decentralized Estimation via Exploiting the Statistical Knowledge of Sensor Noise Variance," *IEEE Trans. on Signal Processing*, Vol. 56, No. 5, pp. 2171-2176, May 2008.

[12] G. Thatte and U. Mitra "Sensor Selection and Power Allocation for Distributed Estimation in Sensor Networks: Beyond the Star Topology," *IEEE Trans. on Signal Processing*, Vol. 56, No. 7, pp. 2649-2661, July 2008.

[13] J. Duplicy and L. Vandendorpe, "Robust MMSE Precoding for the MIMO Compex Gaussian Broadcast Channel," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'07)*, pp. 421-424, 2007.

[14] D. P. Palomar, *A Unified Framework for Communications Through MIMO Channels,* PhD Dissertation, Universitat Politechnica de Catalunya, Spain, May 2003.

[15] D. Tse and P. Viswanath *"Fundamentals of Wireless Communication,"* Cambridge University Press, 2005.

[16] J. O. Berger, V. D. Oliveira and B. Sansó, "Objective Bayesian Analysis of Spatially Correlated Data," *Journal of American Statistical Association*, Vol. 96, No. 456, pp. 1361-1374, December 2001.

[17] "Special Theme: Towards Green ICT," *ERCIM News*, No. 79, October 2009.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory,* Prentice-Hall, 1993.

[19] M. H. Chaudhary and L. Vandendorpe "Adaptive Power Allocation in Wireless Sensor Networks with Spatially Correlated Data and Analog Modulation: Perfect and Imperfect CSI," *EURASIP Journal on Wireless Communications and Networking*, Vol. 2010.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization,* Cambridge University Press, 2008.