Title
Decentralized Estimation Under Correlated Noise

Journal
IEEE Transactions on Signal Processing, 62(21)

ISSN
1053-587X

Authors
Behbahani, Alireza S
Eltawil, Ahmed M
Jafarkhani, Hamid

Publication Date
2014

DOI
10.1109/tsp.2014.2356435

Peer reviewed

Decentralized Estimation Under Correlated Noise

Alireza S. Behbahani, *Member, IEEE*, Ahmed M. Eltawil, *Senior Member, IEEE*, and Hamid Jafarkhani, *Fellow, IEEE*

***Abstract*—In this paper, we consider distributed estimation of an unknown random scalar by using wireless sensors and a fusion center (FC). We adopt a linear model for distributed estimation of a scalar source where both observation models and sensor operations are linear, and the multiple access channel (MAC) is coherent. We consider a fusion center with multiple antennas and single antenna. In order to estimate the source, best linear unbiased estimation (BLUE) is adopted. Two cases are considered: Minimization of the mean square error (MSE) of the BLUE estimator subject to network power constraint, and minimization of the network power subject to the quality of service (QOS). For a fusion center with multiple antennas, iterative solutions are provided and it is shown that the proposed algorithms always converge. For a fusion center with single antenna, closed-form solutions are provided, and it is shown that the iterative solutions will reduce to the closed-form solutions. Furthermore, the effect of noise correlation at the sensors and fusion center is investigated. It is shown that knowledge of noise correlation at the sensors will help to improve the system performance. Moreover, if correlation exists and not factored in, the system performance might improve depending on the correlation structure. We also show, by simulations, that when noise at the fusion center is correlated, even with knowing the correlation structure, the system performance degrades. Finally, simulations are provided to verify the analysis and present the performance of the proposed schemes.**

***Index Terms*—Correlation, distributed estimation, multiple access channel, wireless sensor networks.**

I. Introduction

Wireless sensor networks (WSNs) have attracted significant attention recently due to their diverse applications, such as environmental monitoring, industrial monitoring, battlefield surveillance, agriculture, home applications, and smart energy to name a few [1], [2]. A wireless sensor network consists of a large number of spatially distributed sensors with limited power, limited processing and communication capabilities, and small size. However, such a distributed system can achieve numerous high level tasks [3]. Wireless sensor networks can be deployed to perform distributed processing techniques, including distributed data compression, tracking, classification and distributed detection [4]–[7]. One important characteristic of a wireless sensor network is its energy efficiency. There is a rich body of work in literature on the design of energy efficient wireless sensor networks [8]–[12], and references therein. In this paper, we focus on studying a distributed estimation problem with a fusion center where each sensor linearly encodes its observations and transmits the encoded observations simultaneously to the fusion center.

Manuscript received March 20, 2014; revised July 11, 2014; accepted September 01, 2014. Date of publication September 09, 2014; date of current version October 06, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chong-Yung Chi. This work was supported in part by the National Science Foundation under Awards ECCS-0955157 and CCF-1218771.

The authors are with the Center for Pervasive Communications and Computing, University of California, Irvine, CA 92697-2625 USA (e-mail: sshahanb@uci.edu; aeltawil@uci.edu; hamidj@uci.edu).

Digital Object Identifier 10.1109/TSP.2014.2356435

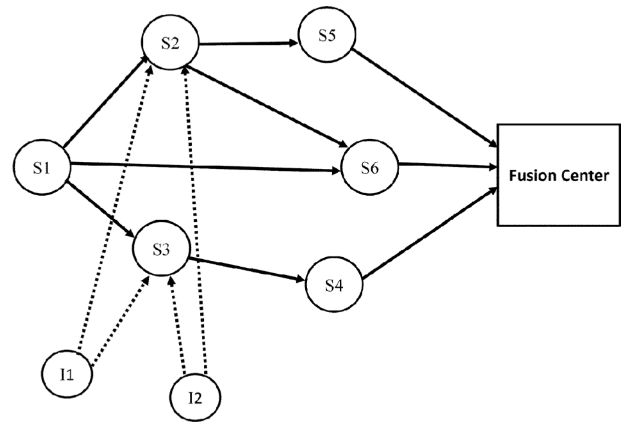

Fig. 1. Noise correlation model: common interference and noise propagation.

A. Prior Work

One of the main objective of a WSN is to reliably estimate event features from the collective information provided by distributed sensors [13]. Since typically sensors are sharing the same spectrum with possible interferers, for example WiFi access points, interference plays a major factor in the design of sensors functionality. Thus, while the distributed structure of WSNs creates opportunities to benefit from spatio-temporal phenomena correlations, the same structure also creates noise correlations that are unique. However most models in literature approach this problem in a simplified manner where noise is considered to be spatially uncorrelated among sensors with the same power [14]–[16], or it is assumed that noise observations are uncorrelated but with different powers [9], [17]. Noise at the sensors can be correlated possibly due to *common interference* or *noise propagation* as shown in Fig. 1 where common interferers (I1 and I2) are propagated through the network via member sensors and create spatially correlated noise among sensors. This scenario can be extrapolated to many practical applications such as multiple hop networks, mesh networks, Structure Health Monitoring (SHM), habitat monitoring, smart grid, etc., [18]–[20].

It is evident that spatially correlated noise has an impact on distributed estimation in a distributed wireless sensor network, but this impact is not well understood and has not been precisely quantified in the literature. Some early work in the area has already established that considering noise correlation, results in improved performance metrics and reduced network power. For example, in [11], [21], [22] the authors claim that when observation noise is considered to be correlated spatially, power savings can be achieved by considering the noise covariance matrix. In [11] authors consider an analog communication framework where a linear minimum mean-square error (LMMSE) estimator is deployed and show that power savings can be achieved by considering the noise covariance matrix in the design. In [21] authors consider a digital communication framework and introduce a distributed estimation technique where quantization levels at the sensors are jointly designed by the fusion center using the knowledge of the noise covariance matrix. In [22] authors consider a cluster based sensor network where sensors in the network are divided between different clusters. It is assumed that the communication between sensors of each cluster is allowed while there is no inter-cluster communication and noise within each cluster and the fusion center can be correlated. Each cluster sends out one or more messages to the fusion center through orthogonal MAC for the final estimation. However, in these works the authors assume that the noise covariance matrix is available at the fusion center and it is not clear how different correlation scenarios can affect system performance. Furthermore, the authors assume a fusion center with one antenna and sensors send information over orthogonal channels to the fusion center.

B. Contributions

In this paper, we adopt the model introduced by [16] for decentralized estimation where both observation models and sensors' operations are linear and a coherent MAC is considered[1]. We consider a case where sensors observe a common scalar source and each sensor has one observation and sends out one message per observation. First we provide solutions for the following two scenarios where the fusion center has multiple antennas: **A)** Minimizing best linear unbiased estimator (BLUE) subject to the total transmit power of sensors (network power), and **B)** Minimizing network power subject to the quality of service (QOS) where the MSE of BLUE is less than some predefined value. For both of these scenarios, we provide iterative solutions where the sensor factors and filter at the fusion center are designed jointly. Furthermore, the provided solutions consider a generalized noise covariance matrix at the sensors and fusion center where noise can be correlated. Finally, it is shown that when the fusion center is equipped with one antenna, a closed form solution can be derived for both scenarios.

We explore the effect of noise correlation for the following two cases:

- *Noise correlation at the sensors*. Due to the ambient noise, sensors distributed densly, noise propagation, common interference, and nonlinearity of sensors, observation noise can be correlated.
- *Noise correlation at the FC*. Due to the broadcast nature of wireless networks, the fusion center is exposed to a set of common interferers resulting in correlated noise at the FC.

Both of the above models are of considerable importance. The natural question will be if it is possible to exploit the noise correlation to improve performance. In practice, knowledge of correlation may result in overheads in the network. Whether such overheads are justified depends on the potential scenario.

The main results of the paper are:

1) For the case of a fusion center with multiple antennas :
   a) MSE of BLUE is minimized subject to network power. In this case sensor factors and filter at the fusion center are designed jointly and an iterative solution is provided. The optimality and convergence of the proposed solution is investigated.
   b) Network power is minimized subject to QOS where a minimum MSE is required to be achieved. An iterative solution is provided to design sensor factors and filter at the fusion center.
   c) It is shown by simulations that if noise at the fusion center is correlated, the MSE performance degrades and the knowledge of correlation does not help.
2) For the case of a fusion center with a single antenna:
   a) A closed form solution for the sensor factors and filter at the fusion center are provided for both optimization scenarios. It is shown that the iterative solution provided for the case of multiple antennas will reduce to the closed form solution for both scenarios.
   b) The effect of noise correlation at the sensor is considered for four different scenarios, where the noise is either correlated or uncorrelated and the FC either aware or unaware of the correlation. It is shown that while knowledge of noise correlation at the sensors will improve system performance (unlike knowledge of noise correlation at the fusion center), there are situations where correlation helps, even without knowing that correlation exists.

The remainder of the paper is organized as follows: Section II describes the system model and problem formulation. In Section III, we minimize MSE under network power constraint for a fusion center with multiple antennas. In Section IV, network power minimization under QOS is introduced for a fusion center with multiple antennas. In Section V, MSE is minimized under network power constraint for a single antenna fusion center and some asymptotic properties of such a network are provided. Section VI exploits noise correlation at the sensors. In Section VII, the solution for network power minimization for a single antenna fusion center is provided. Simulation results are provided in Section VIII and we conclude in Section IX.

Notation: We use bold lower case for vectors, while bold capital letters are used for matrices. Further $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ stand for conjugation, transposition and Hermitian transposition respectively. $\odot$ represents element-wise multiplication, $[\boldsymbol{A}]_{i,j}$ denotes the element in row i and column j of matrix $\boldsymbol{A}$, and $\mathrm{tr}[\boldsymbol{A}]$ is the trace of matrix $\boldsymbol{A}$. We further define $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{a})$ as a diagonal matrix $\boldsymbol{A}$ that contains the elements of $\boldsymbol{A}$ on its diagonal and $\boldsymbol{b} = \mathrm{diag}(\boldsymbol{F})$ as a vector whose elements are the diagonal elements of matrix $\boldsymbol{F}$. In addition, $\mathcal{P}(\boldsymbol{A})$ is the principle

[1]Note that while we adopt the model in [16] there are differences between the two models. In our proposed model, the fusion center has multiple antennas and the number of antennas at the fusion center can be arbitrary. In contrast, in [16], the authors consider a fusion center with the number of observations equal to the number of transmitted messages from the sensors. Furthermore, it is assumed that the noise covariance matrix at the sensors and the fusion center are identity matrices and therefore there is no discussion on the effect of noise correlation on the system performance in [16].

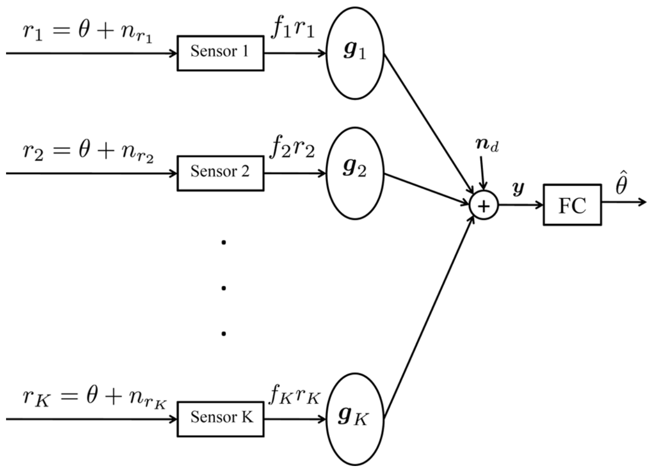

Fig. 2. Block diagram of a general coherent MAC linear decentralized estimation.

eigenvector of matrix $\boldsymbol{A}$ associated to its highest eigenvalue and $\lambda_{\max}(\boldsymbol{A})$ is the maximum eigenvalue of matrix $\boldsymbol{A}$. Also $\mathbf{E}$ stands for expectation operator, $\mathcal{C}$ denotes the set of complex scalars, and $\mathcal{C}^{m\times n}$ represents the set of $m \times n$ matrices with complex entries.

II. System Model and Problem Formulation

We consider a distributed wireless sensor network (WSN) with a fusion center as shown in Fig. 2. Suppose there are K sensors, each observing a common unknown random source θ, where θ has zero mean and variance of σ_θ^2. We assume there is no inter-sensor communication since the sensors are distributed. In this work, we consider a linear decentralized estimation where both observation model and sensor operations are linear. The K sensors transmit their information to a fusion center through a coherent multiple access channel (MAC).

The observation at sensor i can be described as

$$r_i = \theta + n_{r_i}, \qquad 1 \le i \le K \tag{1}$$

where $r_i \in \mathcal{C}$ is the observation at the ith sensor, and n_{r_i} is the additive noise at the ith sensor with the variance $\sigma_{n_{r_i}}^2$. By stacking the observations from all sensors into a single vector, the collection of sensor observations can be expressed as

$$\boldsymbol{r} = \mathbf{1}\theta + \boldsymbol{n}_r, \tag{2}$$

where $\boldsymbol{r} = [r_1, r_2, \ldots, r_K]^T$ is the $K \times 1$ observations at the sensors. Also $\boldsymbol{n}_r = [n_{r_1}, n_{r_2}, \ldots, n_{r_K}]^T$ is the $K \times 1$ additive noise vector at the sensors which has zero mean and covariance matrix $\boldsymbol{R}_{\boldsymbol{n}_r} = \mathbf{E}(\boldsymbol{n}_r \boldsymbol{n}_r^H)$, and $\mathbf{1} = [1\ 1\ \cdots\ 1]^T$ is the $K \times 1$ vector with all elements equal to one.

Considering the assumption of a linear model at the sensors, Sensor i encodes its observations by multiplying it with f_i. Thus, the transmitted signal, $t_i \in \mathcal{C}$, at Sensor i can be written as

$$t_i = f_i r_i \tag{3}$$

where $f_i \in \mathcal{C}$ is the encoder coefficient for Sensor i. Again, by stacking the transmitted signals from all sensors into a single vector, the collection of the transmitted sensor signals, $\boldsymbol{t} \in \mathcal{C}^{K\times 1}$, can be expressed as

$$\boldsymbol{t} = \boldsymbol{F}\mathbf{1}\theta + \boldsymbol{F}\boldsymbol{n}_r \tag{4}$$

where $\boldsymbol{t} = [t_1, t_2, \ldots, t_K]^T$, $\boldsymbol{F} = \mathrm{diag}(\boldsymbol{f})$ is $K \times K$ block diagonal matrix such that the ith diagonal element is the encoder matrix of the ith sensor which is f_i, and $\boldsymbol{f} = [f_1, f_2, \ldots, f_K]^T$.

Assuming a coherent MAC between sensors and the fusion center, all the sensors transmit simultaneously by using e.g., the same time slot in TDMA or the same frequency band in FDMA. The transmitted signals from all sensors will reach the fusion center coherently under the assumption of perfect synchronization between sensors and the fusion center[2]. Assuming a fusion center which is equipped with N_{FC} antennas, the received signal at the fusion center can be expressed as

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{F}\mathbf{1}\theta + \boldsymbol{G}\boldsymbol{F}\boldsymbol{n}_r + \boldsymbol{n}_d \tag{5}$$

where $\boldsymbol{y} \in \mathcal{C}^{N_{FC}\times 1}$, $\boldsymbol{G} \in \mathcal{C}^{N_{FC}\times K}$ is the $N_{FC} \times K$ channel matrix between the sensors and fusion center and $\boldsymbol{n}_d$ is an additive noise with covariance matrices $\boldsymbol{R}_{\boldsymbol{n}_d} = \mathbf{E}(\boldsymbol{n}_d^H \boldsymbol{n}_d)$. The channel matrix $\boldsymbol{G}$ is further defined as $\boldsymbol{G} = [\boldsymbol{g}_1 \boldsymbol{g}_2 \cdots \boldsymbol{g}_K]$ where $\boldsymbol{g}_i \in \mathcal{C}^{N_{FC}\times 1}$ is the $N_{FC} \times 1$ channel from the ith sensor to the FC as shown in Fig. 2. Furthermore, we assume that the noise at the sensors, $\boldsymbol{n}_r$, and the noise at the fusion center, $\boldsymbol{n}_d$, are independent.

III. MSE Minimization Under Network Power for a Multiple Antennas FC

In this section, we jointly design the optimum sensor encoder coefficients and fusion center filter in order to minimize the MSE of BLUE estimator subject to total transmit power constraints at the sensors. We use the total transmit power since it allows tractable analysis leading to useful insights. The fusion center linearly processes the received signal, $\boldsymbol{y}$, by multiplying it with the filter $\boldsymbol{z} \in \mathcal{C}^{N_{FC}\times 1}$. Thus, the decision statistic $y_{FC} \in \mathcal{C}$ is given by

$$y_{FC} = \boldsymbol{z}^H\boldsymbol{y} = \boldsymbol{z}^H\boldsymbol{G}\boldsymbol{F}\mathbf{1}\theta + \boldsymbol{z}^H\boldsymbol{G}\boldsymbol{F}\boldsymbol{n}_r + \boldsymbol{z}^H\boldsymbol{n}_d. \tag{6}$$

Furthermore, the total noise power after processing at the fusion center would be $\sigma_{FC}^2 = \boldsymbol{z}^H\left[\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H + \boldsymbol{R}_{\boldsymbol{n}_d}\right]\boldsymbol{z}$.

A. Sensor Optimizations

We use a BLUE estimator [26] for estimating the source, θ, which can be expressed as

$$\hat{\theta} = \left[\boldsymbol{f}^H\boldsymbol{G}^H\boldsymbol{z}\sigma_{FC}^{-2}\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}\right]^{-1}\boldsymbol{f}^H\boldsymbol{G}^H\boldsymbol{z}\sigma_{FC}^{-2}y_{FC}, \tag{7}$$

[2]Note that a coherent MAC is considered over an orthogonal MAC since it is K times more bandwidth efficient. Furthermore, for a coherent MAC, the achievable MSE decreases in the order of $1/K$ while for an orthogonal MAC the achievable MSE is finite even when the number of sensors goes to infinity.

From an information theoretical point of view, it has been shown that an asynchronous MAC can provide a higher capacity [28]. From a communication/networking point of view, asynchronous transmission is still an important open research topic. However some recent results show that it is possible to handle asynchronous MAC transmission and get results that are as good as synchronous MAC (see [29] and references therein).

and therefore the MSE of the BLUE estimator can be written as

$$\begin{aligned}\mathrm{MSE}(\boldsymbol{f},\boldsymbol{z}) &= \mathbf{E}|\theta-\hat{\theta}|^2 = \left[\boldsymbol{f}^H\boldsymbol{G}^H\boldsymbol{z}\sigma_{FC}^{-2}\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}\right]^{-1}\\ &= \left[\frac{|\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H+\boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z}}\right]^{-1}. \quad (8)\end{aligned}$$

Note that the expectation is taken over random variables θ and $\hat{\theta}$. However, since BLUE estimator is an unbiased estimator, $\theta-\hat{\theta}$ is not a function of the random variable θ anymore and the expectation is only needed to be taken over the distribution of noise at the sensors and the fusion center.

Now, the optimization can be expressed as

$$\begin{aligned}&\min_{\boldsymbol{f},\boldsymbol{z}} \mathrm{MSE}(\boldsymbol{f},\boldsymbol{z}) \equiv \max_{\boldsymbol{f},\boldsymbol{z}} \frac{|\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H+\boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z}}\\ &\text{s.t. } \boldsymbol{f}^H\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2+\boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]\boldsymbol{f}\le P_T \quad (9)\end{aligned}$$

where P_T is the total transmit power of sensors. The above optimization can be solved iteratively if we can find $\boldsymbol{z}$ in terms of $\boldsymbol{f}$ and $\boldsymbol{f}$ in terms of $\boldsymbol{z}$.

To solve (9), we first solve $\boldsymbol{z}$ in terms of $\boldsymbol{f}$. In this case, the optimization problem becomes

$$\max_{\boldsymbol{z}} \frac{|\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H+\boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z}}. \quad (10)$$

The above optimization problem can be recast as

$$\min_{\boldsymbol{z}} \boldsymbol{z}^H\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H+\boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z} \quad \text{s.t.} \quad \boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}=1. \quad (11)$$

This is the minimum variance distortionless response (MVDR) problem [23] and the solution to this problem is

$$\boldsymbol{z}=\alpha\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H+\boldsymbol{R}_{\boldsymbol{n}_d}\right)^{-1}\boldsymbol{G}\boldsymbol{f}, \quad (12)$$

where α is a constant which satisfies the equality condition in (11) and can be ignored since it has no effect on the original optimization problem (9). Note that, if the cost function in (11) equals to another number other than 1, it will change the constant scalar α in (11). However, α in (12) which satisfies the equality condition in the constraint can be ignored since it has no effect on the original optimization problem in (9). In fact any scaling factor for the vector $\boldsymbol{z}$ does not change the value of the cost function as can be seen from (9). This means that no matter what value we choose for the constraint in (11), it does not change the cost function and consequently it does not affect the value of vector $\boldsymbol{f}$. Therefore, we have chosen an arbitrary number, i.e., 1. Now that we have solved $\boldsymbol{z}$ in terms of $\boldsymbol{f}$, we solve $\boldsymbol{f}$ in terms of $\boldsymbol{z}$ from the original optimization problem in (9). In order to do that we substitute $\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{F}=\boldsymbol{f}^T\boldsymbol{E}$ where $\boldsymbol{E}=\mathrm{diag}(\boldsymbol{e})$ and $\boldsymbol{e}=[\boldsymbol{z}^H\boldsymbol{g}_1,\boldsymbol{z}^H\boldsymbol{g}_2,\ldots,\boldsymbol{z}^H\boldsymbol{g}_K]^T$. Therefore, for fixed $\boldsymbol{z}$, the vector $\boldsymbol{f}$ can be obtained from the following optimization problem

$$\begin{aligned}&\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^*+\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}\\ &\qquad \text{s.t.} \quad \boldsymbol{f}^T\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2+\boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]\boldsymbol{f}^*\le P_T. \quad (13)\end{aligned}$$

It is shown in Appendix A that the inequality power constraint in the above optimization can be replaced with equality. Therefore, the optimization in (13) can be recast as

$$\begin{aligned}&\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^*+\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}\\ &\qquad \text{s.t.} \quad \boldsymbol{f}^T\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2+\boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]\boldsymbol{f}^*= P_T. \quad (14)\end{aligned}$$

Now, as shown in Appendix B, the optimum solution is

$$\boldsymbol{f}=\alpha(\boldsymbol{B}^{-1}\boldsymbol{e})^*, \quad (15)$$

where $\boldsymbol{B}=\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H P_T+\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2+\boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]$ and $\alpha^2=\frac{P_T}{\boldsymbol{e}^H\boldsymbol{B}^{-H}[(\mathbf{1}\mathbf{1}^T\sigma_\theta^2+\boldsymbol{R}_{\boldsymbol{n}_r})\odot\boldsymbol{I}]\boldsymbol{B}^{-1}\boldsymbol{e}}$.

Now that we have solved $\boldsymbol{z}$ in terms of $\boldsymbol{f}$ and $\boldsymbol{f}$ in terms of $\boldsymbol{z}$, we can iteratively solve for $\boldsymbol{z}$ and $\boldsymbol{f}$ such that the total MSE monotonically decreases. The algorithm starts with initializing the vector $\boldsymbol{f}$, then the fusion center filter, $\boldsymbol{z}$, and sensors' encoder vector, $\boldsymbol{f}$, can be updated iteratively. Note that the objective function, namely MSE, will be nonincreasing in every iteration step. Therefore, it will converge [24].

B. High Observation SNR

At high observation SNR at the sensors, i.e., $\frac{\sigma_\theta^2}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{ii}}\gg 1\forall i$, noise at the sensors can be neglected. Then, the received signal can be expressed as

$$\boldsymbol{y}=\boldsymbol{G}\boldsymbol{F}\mathbf{1}\theta+n_d \quad (16)$$

and the total network transmit power will be $\boldsymbol{f}^H\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2\right)\odot\boldsymbol{I}\right]\boldsymbol{f}=P_T$.

In this case, it is possible to find an exact closed form solution instead of having an iterative algorithm. This solution is in fact the matched filter. In this case, the optimizations can be expressed as

$$\max_{\boldsymbol{f},\boldsymbol{z}} \frac{|\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \quad \text{s.t.} \quad \boldsymbol{f}^H\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2\right)\odot\boldsymbol{I}\right]\boldsymbol{f}=P_T. \quad (17)$$

One can first solve $\boldsymbol{z}$ in terms of $\boldsymbol{f}$ similar to (12) and obtain

$$\boldsymbol{z}=\beta\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G}\boldsymbol{f}, \quad (18)$$

where β is a scalar which will be determined by the constraint $\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f}=1$. However, this scalar can be ignored since it has no effect on the cost function in (17). Substituting the above solution into the cost function in (17), the optimization problem becomes

$$\max_{\boldsymbol{f}} \boldsymbol{f}^H\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G}\boldsymbol{f} \quad \text{s.t.} \quad \boldsymbol{f}^H\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2\right)\odot\boldsymbol{I}\right]\boldsymbol{f}=P_T. \quad (19)$$

By using the methodology presented in Appendix B and Rayleigh quotient [25], the solution can be expressed as

$$\boldsymbol{f}=\alpha\mathcal{P}(\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G}), \quad (20)$$

where α is determined in order to satisfy the power constraint and is $\alpha^2=\frac{P_T}{\mathcal{P}^H(\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G})[(\mathbf{1}\mathbf{1}^T\sigma_\theta^2)\odot\boldsymbol{I}]\mathcal{P}(\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G})}$. Furthermore, the maximum cost function in (19) can be expressed as $\lambda_{\max}[\frac{P_T}{\sigma_\theta^2}\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G}]=\frac{P_T}{\sigma_\theta^2}\lambda_{\max}[\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{n}_d}^{-1}\boldsymbol{G}]$. Therefore, the minimum MSE for the case of high observation SNR for a WSN with a FC with multiple antennas can be expressed as

$$\mathrm{MSE} = \frac{\sigma_\theta^2}{P_T \lambda_{\max}[\boldsymbol{G}^H \boldsymbol{R}_{\boldsymbol{n}_d}^{-1} \boldsymbol{G}]}. \tag{21}$$

IV. Network Power Minimization Under QOS Constraint for a Multiple Antennas FC

In this section we design sensor factors by minimizing the total sensor transmit power (network power), P_T, subject to the fusion center QOS constraint. The QOS is given by the BLUE MSE at the FC, therefore, this problem can be written as

$$\min_{\boldsymbol{f}} P_T \quad \text{s.t.} \quad \mathrm{MSE} \le \frac{1}{\gamma}. \tag{22}$$

By using (7) to (9), the above optimization can be written as

$$\min_{\boldsymbol{f},\boldsymbol{z}} \boldsymbol{f}^T \underbrace{\left[\left(\boldsymbol{1}\boldsymbol{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right) \odot \boldsymbol{I}\right]}_{\boldsymbol{X}} \boldsymbol{f}^*$$
$$\text{s.t.} \quad \frac{|\boldsymbol{z}^H \boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H \left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H + \boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z}} \ge \gamma, \tag{23}$$

where $\boldsymbol{X} = \left[\left(\boldsymbol{1}\boldsymbol{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right) \odot \boldsymbol{I}\right]$. Note that in the optimization problem in (23) the constraint function is a function of $\boldsymbol{z}$, and $\boldsymbol{f}$, however the cost function is only a function of $\boldsymbol{f}$. Consequently, we first construct a non constraint optimization and find $\boldsymbol{z}$ in terms of $\boldsymbol{f}$, then we construct another optimization to find $\boldsymbol{f}$ in terms of $\boldsymbol{z}$. To solve $\boldsymbol{z}$ in terms of $\boldsymbol{f}$, we formulate the following optimization problem

$$\max_{\boldsymbol{z}} \frac{|\boldsymbol{z}^H \boldsymbol{G}\boldsymbol{f}|^2}{\boldsymbol{z}^H \left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H + \boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z}}. \tag{24}$$

The above optimization problem can be recast as

$$\min_{\boldsymbol{z}} \boldsymbol{z}^H \left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H + \boldsymbol{R}_{\boldsymbol{n}_d}\right)\boldsymbol{z} \quad \text{s.t.} \quad \boldsymbol{z}^H\boldsymbol{G}\boldsymbol{f} = 1. \tag{25}$$

This is the minimum variance distortionless response (MVDR) problem [23] with the following solution

$$\boldsymbol{z} = \alpha\left(\boldsymbol{G}\boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H\boldsymbol{G}^H + \boldsymbol{R}_{\boldsymbol{n}_d}\right)^{-1}\boldsymbol{G}\boldsymbol{f}, \tag{26}$$

where α is a constant which satisfies the equality condition in (25). Note that α can be ignored since it has no effect on the original optimization problem (24).

Now that we have solved $\boldsymbol{z}$ in terms of $\boldsymbol{f}$, we solve $\boldsymbol{f}$ in terms of $\boldsymbol{z}$ from the original optimization problem in (23). In order to do that, we substitute $\boldsymbol{z}^H\boldsymbol{G}\boldsymbol{F} = \boldsymbol{f}^T\boldsymbol{E}$ where $\boldsymbol{E} = \mathrm{diag}(\boldsymbol{e})$ and $\boldsymbol{e} = [\boldsymbol{z}^H\boldsymbol{g}_1, \boldsymbol{z}^H\boldsymbol{g}_2, \ldots, \boldsymbol{z}^H\boldsymbol{g}_K]^T$. For fixed $\boldsymbol{z}$, the vector $\boldsymbol{f}$ can be obtained from the following optimization problem

$$\min_{\boldsymbol{f}} \boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^* \quad \text{s.t.} \quad \max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \ge \gamma. \tag{27}$$

Introducing

$$\bar{\boldsymbol{f}}^* = \boldsymbol{X}^{1/2}\boldsymbol{f}^*,$$
$$\boldsymbol{Q} = \boldsymbol{X}^{-1/2}\left[\boldsymbol{e}\boldsymbol{e}^H - \gamma\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\right]\boldsymbol{X}^{-1/2} \tag{28}$$

we can reformulate the optimization problem in (27) as

$$\min_{\bar{\boldsymbol{f}}} \|\bar{\boldsymbol{f}}^*\|^2 \quad \text{s.t.} \quad \bar{\boldsymbol{f}}^T\boldsymbol{Q}\bar{\boldsymbol{f}}^* \ge \gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}. \tag{29}$$

The constraint function in (29) can be used for checking the feasibility of the problem for any given value of γ. In particular, for all the values of γ that lead to a negative semidefinite $\boldsymbol{Q}$, the problem in (29) is infeasible. In addition, the constraint in (29) can be replaced by the equality constraint $\bar{\boldsymbol{f}}^T\boldsymbol{Q}\bar{\boldsymbol{f}}^* = \gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}$. Hence, the problem in (29) is equivalent to

$$\min_{\bar{\boldsymbol{f}}} \|\bar{\boldsymbol{f}}^*\|^2 \quad \text{s.t.} \quad \bar{\boldsymbol{f}}^T\boldsymbol{Q}\bar{\boldsymbol{f}}^* = \gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}. \tag{30}$$

The solution to (30) can be found by using the Lagrangian method. The Lagrangian function of the optimization problem in (30) can be written as

$$L(\bar{\boldsymbol{f}}, \mu) = \bar{\boldsymbol{f}}^T\bar{\boldsymbol{f}}^* + \mu\left[\gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z} - \bar{\boldsymbol{f}}^T\boldsymbol{Q}\bar{\boldsymbol{f}}^*\right] \tag{31}$$

where $\mu \ge 0$ is the Lagrange multiplier. Now, in order to find the optimum sensor factors, $\bar{\boldsymbol{f}}$, we differentiate (31) with respect to $\bar{\boldsymbol{f}}^T$ and equate it to zero, resulting in,

$$\boldsymbol{Q}\bar{\boldsymbol{f}}^* = \frac{1}{\mu}\bar{\boldsymbol{f}}^*. \tag{32}$$

Multiplying both sides of (32) by $\mu\bar{\boldsymbol{f}}^T$ yields

$$\mu\bar{\boldsymbol{f}}^T\boldsymbol{Q}\bar{\boldsymbol{f}}^* = \bar{\boldsymbol{f}}^T\bar{\boldsymbol{f}}^* = \|\bar{\boldsymbol{f}}^*\|^2 = \mu\gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}. \tag{33}$$

By looking at (33) and the optimization problem in (30), it can be seen that minimizing $\|\bar{\boldsymbol{f}}\|^2$ leads to the smallest positive μ since γ and $\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}$ are fixed positive values. This is equivalent to finding the largest $\frac{1}{\mu}$ in (32). Now, the optimal solution of the problem in (29) can be written as

$$\bar{\boldsymbol{f}}^* = \alpha\mathcal{P}(\boldsymbol{Q}) \tag{34}$$

where α ensures that the power constraint is satisfied and can be written as

$$\alpha = \sqrt{\frac{\gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}{\mathcal{P}(\boldsymbol{Q})^H\boldsymbol{Q}\mathcal{P}(\boldsymbol{Q})}}. \tag{35}$$

Finally, the optimal solution and the minimum total power can be expressed as

$$\boldsymbol{f} = \sqrt{\frac{\gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}{\mathcal{P}(\boldsymbol{Q})^H\boldsymbol{Q}\mathcal{P}(\boldsymbol{Q})}}\boldsymbol{X}^{-1/2}\mathcal{P}(\boldsymbol{Q})^* \tag{36}$$

$$P_{T_{\min}} = \frac{\gamma\boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}{\lambda_{\max}[\boldsymbol{Q}]}. \tag{37}$$

Now that we have solved $\boldsymbol{z}$ in terms of $\boldsymbol{f}$ and $\boldsymbol{f}$ in terms of $\boldsymbol{z}$, similar to the algorithm proposed in Section III-A, we can iteratively solve for $\boldsymbol{z}$ and $\boldsymbol{f}$ such that the total MSE monotonically decreases. The algorithm starts with initializing the vector $\boldsymbol{f}$, then the fusion center filter, $\boldsymbol{z}$, and sensors' encoder vector, $\boldsymbol{f}$, can be updated iteratively. Note that the objective function, namely MSE, will be nonincreasing in every iteration step and the algorithm converges.

V. MSE Minimization Under Network Power for a Single Antenna FC

In this section we consider a fusion center with one antenna. The problem formulation in the case of single antenna is the same as the one introduced in Section III except that there is no need to have a filter at the fusion center. The received signal at the fusion center, which is equipped with one antenna, can be expressed as

$$y = \boldsymbol{g}^T \boldsymbol{F}\mathbf{1}\theta + \boldsymbol{g}^T \boldsymbol{F}\boldsymbol{n}_r + n_d \tag{38}$$

where $\boldsymbol{g} = [g_1 g_2 \cdots g_K]^T$[3] is the $1 \times K$ channel vector between the sensors and the fusion center and g_i is the channel between Sensor i and the fusion center. Also, n_d is an additive noise with zero mean and variance $\sigma_{n_d}^2$, where we assume there is no correlation between the fusion center noise, n_d, and the sensor noise vector $\boldsymbol{n}_r$. Furthermore, the total noise variance at the fusion center is $\sigma_{FC}^2 = \boldsymbol{g}^T \boldsymbol{F}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{F}^H \boldsymbol{g}^* + \sigma_{n_d}^2$.

The BLUE estimator for the case of single antenna at the fusion center can be expressed as

$$\hat{\theta} = \left[\boldsymbol{f}^H \boldsymbol{g}^* \sigma_{FC}^{-2} \boldsymbol{g}^T \boldsymbol{f}\right]^{-1} \boldsymbol{f}^H \boldsymbol{g}^* \sigma_{FC}^{-2} y, \tag{39}$$

and the MSE of the BLUE estimator is

$$\mathrm{MSE} = \left[\frac{|\boldsymbol{f}^T \boldsymbol{g}|^2}{\boldsymbol{f}^T \boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{R}_{\boldsymbol{n}_r} \boldsymbol{D}_{\boldsymbol{g}}^H \boldsymbol{f}^* + \sigma_{n_d}^2}\right]^{-1}, \tag{40}$$

where $\boldsymbol{D}_{\boldsymbol{g}} = \mathrm{diag}(\boldsymbol{g})$.

A. Sensor Optimization

In this section, similar to Section III-A, we design sensor gains in such a way as to minimize BLUE MSE subject to the total transmit power of sensors. Therefore, the optimization problem can be formulated as

$$\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T \boldsymbol{g}|^2}{\boldsymbol{f}^T \boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{R}_{\boldsymbol{n}_r} \boldsymbol{D}_{\boldsymbol{g}}^H \boldsymbol{f}^* + \sigma_{n_d}^2} \quad \text{s.t.} \quad \boldsymbol{f}^T\left[\left(\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right) \odot \boldsymbol{I}\right] \boldsymbol{f}^* = P_T. \tag{41}$$

Note that for the case of single antenna fusion center there is no need to have a filter at the fusion center. This also can be verified by looking at (9), where $\boldsymbol{z}$ is a scalar and is canceled out from the numerator and denominator of the cost function. Therefore a closed form solution can be found.

By using the same methodology provided in Appendix B, the minimum MSE for BLUE estimator for a distributed wireless sensor network with one antenna at the fusion center and the sensor noise covariance matrix $\boldsymbol{R}_{\boldsymbol{n}_r}$ is given by

$$\mathrm{MSE} = \frac{1}{P_T \boldsymbol{g}^H \boldsymbol{B}^{-1} \boldsymbol{g}} \tag{42}$$

and the optimal sensor gain factors are

$$\boldsymbol{f} = \alpha (\boldsymbol{B}^{-1} \boldsymbol{g})^*, \tag{43}$$

where $\boldsymbol{B} = \boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{R}_{\boldsymbol{n}_r} \boldsymbol{D}_{\boldsymbol{g}}^H P_T + \sigma_{n_d}^2 \left[\left(\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right) \odot \boldsymbol{I}\right]$ and $\alpha^2 = \frac{P_T}{\boldsymbol{g}^H \boldsymbol{B}^{-H}[(\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}) \odot \boldsymbol{I}]\boldsymbol{B}^{-1} \boldsymbol{g}}$.

B. Asymptotic Properties of the Network

1) Case I: $P_T \to \infty$ (or $\sigma_{n_d}^2 \to 0$): At high total network power P_T, for any $\boldsymbol{f}$, the noise at the FC is negligible. Then, the received signal can be written as

$$\begin{aligned} y &= \boldsymbol{g}^T \boldsymbol{F}\mathbf{1}\theta + \boldsymbol{g}^T \boldsymbol{F}\boldsymbol{n}_r \\ &= \boldsymbol{g}^T \boldsymbol{F}(\mathbf{1}\theta + \boldsymbol{n}_r) \end{aligned} \tag{44}$$

and the optimal solution would be

$$\boldsymbol{f} = \alpha \left[\boldsymbol{D}_{\boldsymbol{g}}^{-H} \boldsymbol{R}_{\boldsymbol{n}_r}^{-1} \mathbf{1}\right]^*. \tag{45}$$

Then, the MSE can be expressed as

$$\mathrm{MSE} = \frac{1}{\mathbf{1}^T \boldsymbol{R}_{\boldsymbol{n}_r}^{-1} \mathbf{1}}. \tag{46}$$

By looking at (44), one can see that the network looks like a single input multiple output (SIMO) system. Also, MSE is inversely proportional with the number of sensors in the network. Furthermore, noise power at the senors is the main limiting factor of MSE.

2) Case II: High Observation SNR: At high observation SNR at the sensors, i.e., $\frac{\sigma_\theta^2}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{ii}} \gg 1 \forall i$, noise at the sensors can be neglected. Then, the received signal can be expressed as

$$\begin{aligned} y &= \boldsymbol{g}^T \boldsymbol{F}\mathbf{1}\theta + n_d \\ &= \boldsymbol{f}^T \boldsymbol{D}_{\boldsymbol{g}} \mathbf{1}\theta + n_d \end{aligned} \tag{47}$$

and the total network transmit power would be $\boldsymbol{f}^T\left[\left(\mathbf{1}\mathbf{1}^T \sigma_\theta^2\right) \odot \boldsymbol{I}\right] \boldsymbol{f}^* = P_T$. In this case, the optimal solution can be expressed as

$$\boldsymbol{f} = \alpha \left[\left(\sigma_{n_d}^2 \left(\mathbf{1}\mathbf{1}^T \sigma_\theta^2\right) \odot \boldsymbol{I}\right)^{-1} \boldsymbol{g}\right]^*, \tag{48}$$

and the MSE will be

$$\mathrm{MSE} = \frac{\sigma_{n_d}^2 \sigma_\theta^2}{P_T} \frac{1}{\boldsymbol{g}^H \boldsymbol{g}}. \tag{49}$$

This network behaves like a multiple input single output (MISO) system. Again, like the previous case, MSE is inversely proportional with the number of sensors.

VI. Exploiting Correlation

In this section we address different correlation scenarios and their effect on the MSE. The natural question to ask is whether correlation improves or degrades performance given that the noise covariance matrix is known or unknown. In order to answer the above question, we consider the following four scenarios:

[3]Note that $\boldsymbol{g}^T = \boldsymbol{G}$ for the case of fusion center with one antenna. This is due to the fact that all vectors are defined as column vectors

- The noise is *uncorrelated* and FC is *aware* of that (Scheme-00).
- The noise is *i.i.d.* and FC is *aware* of that (Scheme-iid). (Note: This is a special case of Scheme-00.)
- The noise is *correlated* and FC is *unaware* of the correlation (Scheme-10).
- The noise is *correlated* and FC is *aware* of the correlation (Scheme-11).

Note that the trace of the noise covariance matrix, $\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]$, represents the total sensors' noise power in the network. Therefore, it is fair to compare different correlation scenarios as long as the total noise power in the network is the same. As a result, in order to be fair, we assume that the trace of the noise covariance matrix, $\boldsymbol{R}_{\boldsymbol{n}_r}$, is the same for all schemes. This means that for Scheme-11 and Scheme-10, $\boldsymbol{R}_{\boldsymbol{n}_r}$ is the general noise covariance matrix. For Scheme-00, $\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I}$ is the noise covariance matrix and finally for Scheme-iid, $\frac{\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]}{K}\boldsymbol{I}$ is the noise covariance matrix.

A. Correlation Scenarios

1) Sensors With Uncorrelated Noise (Scheme-00, and Scheme-iid): In this case, the noise covariance matrix is given by $\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I}$. The optimum solution can be obtained by replacing $\boldsymbol{R}_{\boldsymbol{n}_r}$ in (43) with $\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I}$, i.e.,

$$\boldsymbol{f}_{00} = \alpha \left([\boldsymbol{B} \odot \boldsymbol{I}]^{-1} \boldsymbol{g} \right)^{*}. \tag{50}$$

Thus, the corresponding MSE in this case will be

$$\mathrm{MSE}_{00} = \frac{1}{P_T \boldsymbol{g}^H [(\boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{R}_{\boldsymbol{n}_r} \boldsymbol{D}_{\boldsymbol{g}}^H P_T + \sigma_{n_d}^2 [\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}]) \odot \boldsymbol{I}]^{-1} \boldsymbol{g}}. \tag{51}$$

A special case of Scheme-00 is where noise at the sensors are i.i.d. Considering the requirement that the trace of noise covariance matrix has to be the same for all schemes, this means that the noise covariance matrix for Scheme-iid will be $\frac{\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]}{K}\boldsymbol{I}$. The MSE for this case is

$$\mathrm{MSE}_{\mathrm{iid}} = \frac{1}{P_T \boldsymbol{g}^H [(\boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{D}_{\boldsymbol{g}}^H \frac{\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]}{K} P_T + \sigma_{n_d}^2 [\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \frac{\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]}{K} \boldsymbol{I}]) \odot \boldsymbol{I}]^{-1} \boldsymbol{g}}. \tag{52}$$

2) Sensors With Correlated Noise When FC is not Aware of the Correlation (Scheme-10): In this case, since the FC is not aware of the correlation, the correlation is ignored in designing the sensor amplification factor. Therefore, the solution for sensor factors is the same as that of Scheme-00 i.e.,

$$\boldsymbol{f}_{10} = \boldsymbol{f}_{00} = \alpha \left([\boldsymbol{B} \odot \boldsymbol{I}]^{-1} \boldsymbol{g} \right)^{*}, \tag{53}$$

and the MSE can be written as

$$\begin{aligned} \mathrm{MSE}_{10} &= \left[\frac{P_T \boldsymbol{f}_{10}^T \boldsymbol{g} \boldsymbol{g}^H \boldsymbol{f}_{10}^*}{\boldsymbol{f}_{10}^T [P_T \boldsymbol{D}_{\boldsymbol{g}} \boldsymbol{R}_{\boldsymbol{n}_r} \boldsymbol{D}_{\boldsymbol{g}}^H + \sigma_{n_d}^2 (\mathbf{1}\mathbf{1}^T \sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}) \odot \boldsymbol{I}] \boldsymbol{f}_{10}^*} \right]^{-1} \\ &= \left[\frac{P_T |\boldsymbol{g}^H (\boldsymbol{B} \odot \boldsymbol{I})^{-1} \boldsymbol{g}|^2}{\boldsymbol{g}^H (\boldsymbol{B} \odot \boldsymbol{I})^{-1} \boldsymbol{B} (\boldsymbol{B} \odot \boldsymbol{I})^{-1} \boldsymbol{g}} \right]^{-1}. \end{aligned} \tag{54}$$

3) Sensors With Correlated Noise When FC is Aware of the Correlation (Scheme-11): In this case, since the FC is aware of the correlation, the correlation can be exploited in the sensor design. The solution for this case is already provided in Section V-A.

B. Asymptotic Analysis

1) Case I: $P_T \to \infty$ (or $\sigma_{n_d}^2 \to 0$):

- Scheme-00:

$$\mathrm{MSE}_{00} = \frac{1}{\mathbf{1}^T \left(\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I} \right)^{-1} \mathbf{1}} = \frac{1}{\sum_{i=1}^{K} \frac{1}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}}. \tag{55}$$

- Scheme-iid:

$$\mathrm{MSE}_{\mathrm{iid}} = \frac{\mathrm{tr}\left[\boldsymbol{R}_{\boldsymbol{n}_r}\right]/K}{\mathbf{1}^T \mathbf{1}} = \frac{\mathrm{tr}\left[\boldsymbol{R}_{\boldsymbol{n}_r}\right]}{K^2}. \tag{56}$$

- Scheme-10:

$$\mathrm{MSE}_{10} = \frac{\mathbf{1}^T \left(\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I} \right)^{-1} \boldsymbol{R}_{\boldsymbol{n}_r} \left(\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I} \right)^{-1} \mathbf{1}}{|\mathbf{1}^T \left(\boldsymbol{R}_{\boldsymbol{n}_r} \odot \boldsymbol{I} \right)^{-1} \mathbf{1}|^2}. \tag{57}$$

- Scheme-11:

$$\mathrm{MSE}_{11} = \frac{1}{\mathbf{1}^T \boldsymbol{R}_{\boldsymbol{n}_r}^{-1} \mathbf{1}}. \tag{58}$$

2) Case II: High Observation SNR: At high observation SNR at the sensors, for all schemes, MSE can be expressed as

$$\mathrm{MSE} = \frac{\sigma_{n_d}^2 \sigma_\theta^2}{P_T} \frac{1}{\boldsymbol{g}^H \boldsymbol{g}}. \tag{59}$$

C. Performance Comparison

In this section we compare the performance of different schemes introduced earlier. Since Scheme-11 has complete knowledge of correlation, it is easy to see that always

$$\mathrm{MSE}_{11} \leq \mathrm{MSE}_{10}. \tag{60}$$

For comparison between Scheme-10 and Scheme-00 one can show that

$$\mathrm{MSE}_{10} = \mathrm{MSE}_{00} + \frac{b}{P_T a^2} \tag{61}$$

where

$$a = \boldsymbol{g}^H \left(\boldsymbol{B} \odot \boldsymbol{I} \right)^{-1} \boldsymbol{g} \tag{62}$$

and

$$b = \sum_{i=1}^{K} \sum_{j \neq i} \frac{P_T |g_i|^2 |g_j|^2 [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j}}{\left(P_T |g_i|^2 [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i} + \sigma_{n_d}^2 \left(\sigma_\theta^2 + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i} \right) \right)} \times \frac{1}{\left(P_T |g_j|^2 [\boldsymbol{R}_{\boldsymbol{n}_r}]_{j,j} + \sigma_{n_d}^2 \left(\sigma_\theta^2 + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{j,j} \right) \right)}. \tag{63}$$

It is easy to see that scalar a^2 is always positive, however b can be positive or negative depending on negative or positive correlated noise. So the difference between the MSE of Scheme-10 and Scheme-00 can be positive or negative. In other words,

- If all the noise components are correlated negatively then

$$\mathrm{MSE}_{10} \leq \mathrm{MSE}_{00}. \tag{64}$$

- If all the noise components are correlated positively then

$$\mathrm{MSE}_{10} \geq \mathrm{MSE}_{00}. \tag{65}$$

- If some of the noise components are correlated positively and some negatively then the relationship between MSE_{10} and MSE_{00} depends on the correlation values and the channels from sensors to the fusion center and is governed by (63).

One should note that if the source is real and thus the channels and noise, then $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j}$ is a real scalar. In the case of a complex source, where we assume the source and thus channels and noise are complex random variables, this makes $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j}$ for $i \neq j$ a complex number. However, it is easy to see that if one changes the index of i and j, the term that is multiplied by $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j}$ and $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{j,i}$ is the same which implies that b is a function of $\Re([\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j}) = \Re([\boldsymbol{R}_{\boldsymbol{n}_r}]_{j,i})$ since $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,j} = [\boldsymbol{R}_{\boldsymbol{n}_r}]_{j,i}^H$ for $i \neq j$. In this case, positive or negative correlation implies the real part of the cross correlation between noise.

1) Example 1: In order to clarify the above discussion, we consider a network with two sensors. In this case one can write

$$\mathrm{MSE}_{10} = \underbrace{\frac{1}{P_T\left(\frac{|g_1|^2}{B_1} + \frac{|g_2|^2}{B_2}\right)}}_{\mathrm{MSE}_{00}} + \underbrace{\frac{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2}}{\frac{B_2}{B_1}\frac{|g_1|^2}{|g_2|^2} + \frac{B_1}{B_2}\frac{|g_2|^2}{|g_1|^2} + 2}}_{\frac{b}{P_T a^2}} \tag{66}$$

where $B_1 = [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,1}\left(P_T|g_1|^2 + \sigma_{n_d}^2\right) + \sigma_{n_d}^2\sigma_\theta^2$, $B_2 = [\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,2}\left(P_T|g_2|^2 + \sigma_{n_d}^2\right) + \sigma_{n_d}^2\sigma_\theta^2$, g_1 is the channel from sensor one to the fusion center, and g_2 is the channel from sensor two to the fusion center. Note that since B_1 and B_2 are always positive numbers, the denominator of the right hand side of (66) is always a positive real number. Therefore, the numerator of (66) which is $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2}$ defines the relationship between MSE_{10} and MSE_{10}.

If the source and thus channels and noise are real then $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} = [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2}$ and therefore $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2} = 2[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} = 2[\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2}$

If the source and thus channels and noise are complex then $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} = [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2}^H$ and therefore $[\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1} + [\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2} = 2\Re([\boldsymbol{R}_{\boldsymbol{n}_r}]_{2,1}) = 2\Re([\boldsymbol{R}_{\boldsymbol{n}_r}]_{1,2})$

It is easy to see from (66) that if noise at the sensors are correlated negatively then the right hand side of (66) which is $\frac{b}{P_T a^2}$, is negative and $MSE_{10} \leq MSE_{00}$ and if noise is positively correlated then $MSE_{10} \geq MSE_{00}$. □

The only remaining schemes to compare are Scheme-00 and Scheme-iid. The MSE for these two schemes are provided in (51) and (52). Since it is hard to compare these two schemes in the general form, we consider the case where $P_T \to \infty$ (or $\sigma_{n_d}^2 \to 0$). In this case, one can show that

$$\begin{aligned}\mathrm{MSE}_{00} - \mathrm{MSE}_{\mathrm{iid}} &= \frac{1}{\sum_{i=1}^K \frac{1}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}} - \frac{\mathrm{tr}[\boldsymbol{R}_{\boldsymbol{n}_r}]}{K^2} \\ &= \frac{1}{\sum_{i=1}^K \frac{1}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}} - \frac{\sum_{i=1}^K [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}{K^2} \\ &= \frac{1}{K}\left(\frac{K}{\sum_{i=1}^K \frac{1}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}} - \frac{\sum_{i=1}^K [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}{K}\right).\end{aligned} \tag{67}$$

The term $\frac{K}{\sum_{i=1}^K \frac{1}{[\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}}$ is the harmonic mean (HM) and the term $\frac{\sum_{i=1}^K [\boldsymbol{R}_{\boldsymbol{n}_r}]_{i,i}}{K}$ is the arithmetic mean (AM). It is known that $\mathrm{AM} \geq \mathrm{HM}$. This means that for high network power

$$\mathrm{MSE}_{00} \leq \mathrm{MSE}_{\mathrm{iid}}. \tag{68}$$

Note that if noise is uncorrelated with the same marginal distributions (or uncorrelated and with different marginal distributions but with the same power), then the same results as in Scheme-iid can be drawn as we are only dealing with the noise covariance matrix. This can be seen from (51) and (52) (also from (67) and (68) for high network power), where if noise is uncorrelated with the same marginal distributions (or uncorrelated and with different marginal distributions but with the same power) then $\mathrm{MSE}_{00} = \mathrm{MSE}_{\mathrm{iid}}$.

While it was shown in the above that when noise at all sensors is negatively correlated then $\mathrm{MSE}_{10} \leq \mathrm{MSE}_{00}$, it is not possible to have noise at all sensors negatively correlated simultaneously. It is easy to see that only two sensors at the same time can have negative correlated noise. However, as mentioned before, if some of noise components are correlated positively and some negatively then the relationship between MSE_{10} and MSE_{00} depends on the correlation value and channels from sensors to the fusion center and is governed by (63). This means that different network configuration will control the relationship between MSE_{10} and MSE_{00}.

VII. Network Power Minimization Under QOS Constraint for a Single Antenna FC

In this section, similar to Section IV, we design sensor factors by minimizing the total sensor transmit power (network power), P_T, subject to the fusion center QOS constraint for the case of one antenna. The QOS is given by the BLUE MSE at the FC, therefore, this problem can be written as

$$\min_{\boldsymbol{f}} P_T \quad \text{s.t.} \quad \mathrm{MSE} \leq \frac{1}{\gamma}, \tag{69}$$

which can be expressed as

$$\min_{\boldsymbol{f}} \boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^* \quad \text{s.t.} \quad \frac{|\boldsymbol{f}^T\boldsymbol{g}|^2}{\boldsymbol{f}^T\boldsymbol{D_g}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{D_g}^H\boldsymbol{f}^* + \sigma_{n_d}^2} \geq \gamma, \tag{70}$$

where $X = \left[\left(\boldsymbol{1}\boldsymbol{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right) \odot \boldsymbol{I}\right]$. Note that this problem formulation is similar to the case of multiple antennas defined in (23) except that there is no need to have a filter at the fusion center as it is canceled out from the numerator and denominator of the cost function in (23). Therefore, following the same methodology introduced in Section IV, the optimal solution and the minimum total power can be expressed as

$$\boldsymbol{f} = \sqrt{\frac{\gamma\sigma_{n_d}^2}{\mathcal{P}(\boldsymbol{Q})^H\boldsymbol{Q}\mathcal{P}(\boldsymbol{Q})}}\boldsymbol{X}^{-1/2}\mathcal{P}(\boldsymbol{Q})^* \tag{71}$$

$$P_{T_{\min}} = \frac{\gamma\sigma_{n_d}^2}{\lambda_{\max}[\boldsymbol{Q}]}, \tag{72}$$

where $\boldsymbol{Q} = \boldsymbol{X}^{-1/2}\left[\boldsymbol{g}\boldsymbol{g}^H - \gamma\boldsymbol{D_g}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{D_g}^H\right]\boldsymbol{X}^{-1/2}$.

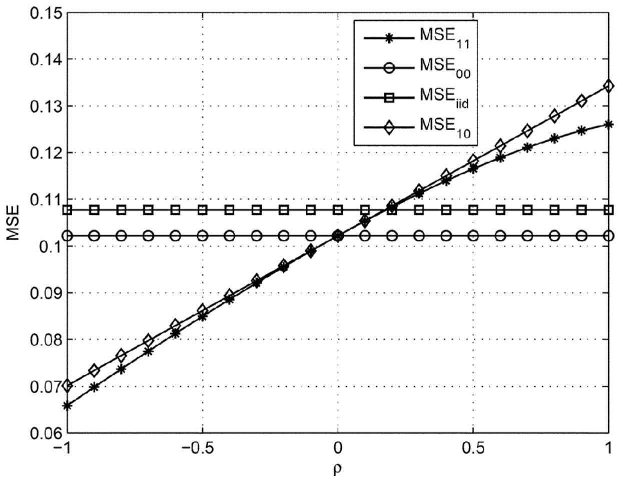

Fig. 3. MSE for the four schemes presented in Section VI-C versus ρ for a network with two sensors. Here there is one antenna at the fusion center, only one specific channel realization is used, $N_{FC} = 1$, $P_T = 10$, and $\mathrm{SNR} = 10$ dB.

VIII. Simulation Results

In this section, we provide simulation results to verify analytical calculations. We assume the elements of $\boldsymbol{G}$ (and $\boldsymbol{g}^T$ for the single antenna fusion center) are generated as zero-mean and unit-variance i.i.d complex Gaussian random variables. Note that while the iterative algorithm provided for the case of a FC with multiple antennas does not guarantee global optimality and might converge to a local optimum, this does not change the conclusions drawn. In order to have a fair comparison, different scenarios are compared under the same simulation setup.

First we consider a network with two sensors. The noise at the sensors is assumed to be correlated with the covariance matrix defined as $\boldsymbol{R}_{\boldsymbol{n}_r} = \sigma^2 \begin{bmatrix} 1+\alpha_1 & \rho \\ \rho & 1+\alpha_2 \end{bmatrix}$ where $-1 \leq \rho \leq 1$ and α_1 and α_2 are two random variables to create a heterogenous system (sensors with different noise power). Furthermore, we assume $\sigma_\theta^2 = 1$ and $\mathrm{SNR} = \frac{1}{\sigma_{n_d}^2} = \frac{1}{\sigma^2}$.

Fig. 3 shows MSE for the four schemes discussed in Section VI-C, for positive and negative correlation values. The figure is for a network with two sensors and a fusion center with one antenna where $P_T = 10$, $\mathrm{SNR} = 10$ dB and for a fixed channel realization. One can see that for negative correlation $\mathrm{MSE}_{11} \leq \mathrm{MSE}_{10}$ and for positive correlation $\mathrm{MSE}_{11} \geq \mathrm{MSE}_{10}$. While in this figure $\mathrm{MSE}_{00} \leq \mathrm{MSE}_{\mathrm{iid}}$, one can find channels where $\mathrm{MSE}_{00} \geq \mathrm{MSE}_{\mathrm{iid}}$. However, as the total network power is increased, it can be verified by simulations that $\mathrm{MSE}_{00} \leq \mathrm{MSE}_{\mathrm{iid}}$ which is consistent with the discussion in Section VI-C.

Fig. 4 shows MSE versus different number of antennas at the fusion center for a network with two sensors and averaged over different channel realizations where $P_T = 10$, $\mathrm{SNR} = 10$ dB, $\rho = -1$. As it can be seen, by increasing the number of antennas at the fusion center, N_{FC}, the MSE performance improves and gets closer to the benchmark. Also, the performance benchmark is defined for all decentralized estimators where all sensor observations are directly available to the fusion center. The MSE in this case is $\left[\boldsymbol{1}^T \boldsymbol{R}_{\boldsymbol{n}_r}^{-1} \boldsymbol{1}\right]^{-1}$. Note that due to the coherent MAC channel between the sensors and the fusion center, intersymbol interference is created at the fusion center. Having multiple antennas at the fusion center provides more degrees of freedom and therefore it is possible to remove intersymbol interference better and have a performance closer to the benchmark.

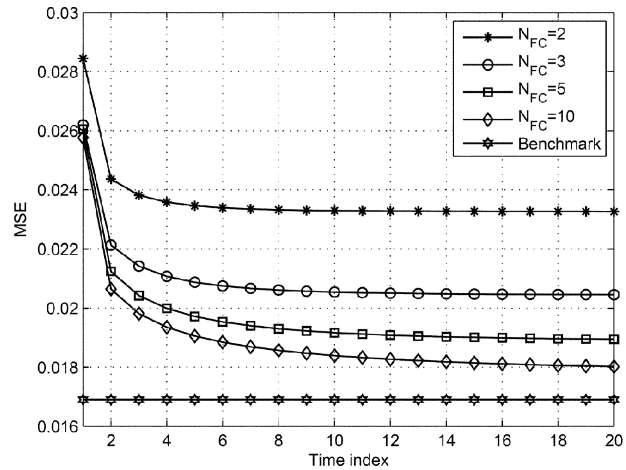

Fig. 4. MSE versus time index where $\rho = -1$ for a network with two sensors and averaged over different channel realizations. Here there are two antennas at the fusion center, $N_{FC} = 2$, $P_T = 10$, and $\mathrm{SNR} = 10$ dB.

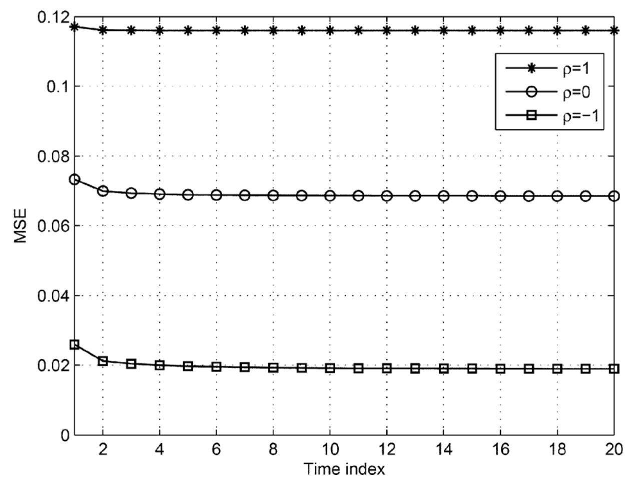

Fig. 5. MSE versus time index for $\rho = -1, 0, 1$. Here $K = 2$, $N_{FC} = 5$, $P_T = 10$, and $\mathrm{SNR} = 10$ dB. Also, MSE is averaged over different channel realizations.

Fig. 5 shows MSE versus time index for $\rho = -1, 0, 1$. The figure is for a network with two sensors and a fusion center with 5 antennas, $N_{FC} = 5$, where $P_T = 10$, $\mathrm{SNR} = 10$ dB and MSE is averaged over different channel realizations. It can be seen that as the correlation becomes less positive and more negative, the MSE performance improves which is consistent with the results provided in Section VI-C.

In order to generate a valid noise covariance matrix at the sensors, for any number of sensors, we consider a common interference based model. In such a model each sensor in addition to its local noise which includes thermal and components noise, observes received signals from different common interferers. The total noise at the sensors under such a model can be expressed as

$$\boldsymbol{n}_r = \sum_{i=1}^{L} \boldsymbol{h}_i s_i + \bar{\boldsymbol{n}}_r \tag{73}$$

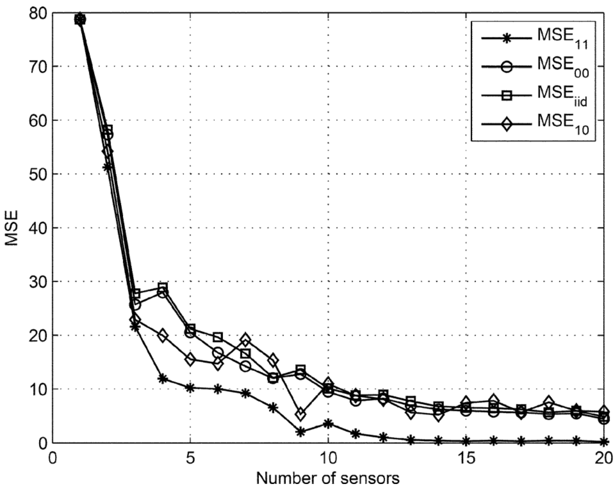

Fig. 6. MSE versus the number of sensors, K. Here $N_{FC} = 1$, $P_T = 10$, $L = 10$, $P_{s_i} = 10 \forall i$, and $\mathrm{SNR} = 10$ dB. Also, MSE is averaged over different channel realizations.

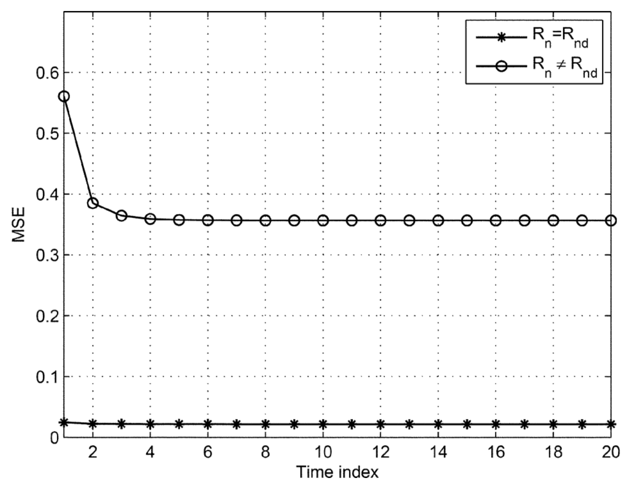

Fig. 7. MSE versus time index. Here $K = 5$, $N_{FC} = 1$, $P_T = 10$, $L = 10$, $P_{s_i} = 10 \forall i$, and $\mathrm{SNR} = 10$ dB. Also, MSE is averaged over different channel realizations and $\boldsymbol{R}_{\boldsymbol{n}_r} = \sigma^2 \boldsymbol{I}$.

where L is the number of interferer, s_i is the transmitted signal from the ith interferer with zero mean and the power of P_{s_i}, and $\boldsymbol{h}_i$ is the $K \times 1$ channel from the ith interferer to the sensors. $\bar{\boldsymbol{n}}_r$ is additive noise with the covariance matrix of $\boldsymbol{R}_{\bar{\boldsymbol{n}}_r} = \mathbf{E}(\bar{\boldsymbol{n}}_r \bar{\boldsymbol{n}}_r^H) = \sigma^2 \boldsymbol{I}$. It is also assumed that different interferers are independent of each other. As previously mentioned, we define $\mathrm{SNR} = \frac{1}{\sigma^2}$. Given this model, the covariance matrix of noise at the sensors can be expressed as

$$\boldsymbol{R}_{\boldsymbol{n}_r} = \mathbf{E}(\boldsymbol{n}_r \boldsymbol{n}_r^H) = \sum_{i=1}^{L} \boldsymbol{h}_i \boldsymbol{h}_i^H P_{s_i} + \sigma^2 \boldsymbol{I} \tag{74}$$

where P_{s_i} is the interference power from the ith interferer.

Fig. 6 shows MSE versus different number of sensors, K, for a fusion center with one antenna, $N_{FC} = 1$. It is assumed that there are 10 interferers, $L = 10$, and each has the power of $P_{s_i} = 10 \forall i$. Furthermore, $\mathrm{SNR} = 10$ dB, $P_T = 10$, and the MSE is averaged over different channel realizations from the sensors to the fusion center. Since the noise covariance matrix is generated randomly, there is no control on how noise at the sensors are correlated. This is the reason why the curves are not smooth. Also, as expected, $\mathrm{MSE}_{11} \le \mathrm{MSE}_{10}$ for any number of sensors. Note that in Fig. 6 noise correlation is fixed for each K, after being generated randomly. If further the curves are averaged over different interferes (different noise correlation matrices), then the curves would be smooth. In this case, as expected, MSE_{11} outperforms the other scenarios with a very good margin. MSE_{10} and MSE_{00} have a very similar performance and $\mathrm{MSE}_{\mathrm{iid}}$ slightly performs worse than MSE_{10} and MSE_{00}.

Fig. 7 shows MSE versus time index for a network with 5 sensors, $K = 5$, $N_{FC} = 5$, and $P_T = 10$. Furthermore, $\mathrm{SNR} = 10$ dB, and the MSE is averaged over different channel realizations from the sensors to the fusion center. Two different cases for noise covariance matrix at the sensors and fusion center are considered. In one case, $\boldsymbol{R}_{\boldsymbol{n}_r} = \boldsymbol{R}_{\boldsymbol{n}_d} = \sigma^2 \boldsymbol{I}$ which means that noise at the sensors and fusion center are i.i.d. with the same power. In the other case, $\boldsymbol{R}_{\boldsymbol{n}_r} = \sigma^2 \boldsymbol{I} \neq \boldsymbol{R}_{\boldsymbol{n}_d}$, and $\boldsymbol{R}_{\boldsymbol{n}_d}$ is generated based on the interference model. As it can be seen, when noise at the fusion center is correlated due to the common interference, the MSE performance degrades compared to the case that noise at the fusion center is i.i.d. This suggests that correlated noise can be beneficial at the sensors, while it degrades the MSE performance if it exists at the fusion center.

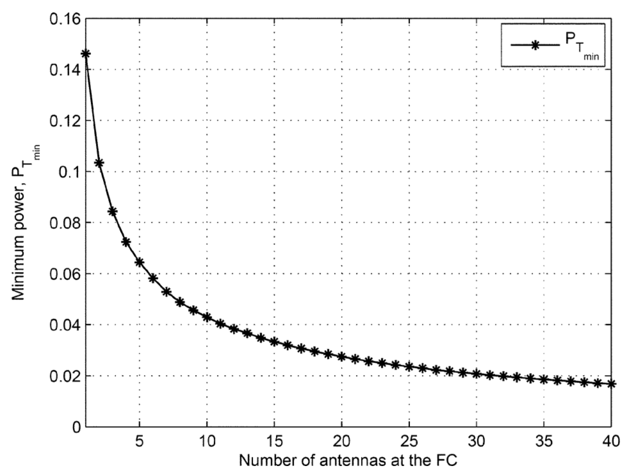

Fig. 8. Minimum power versus number of antennas at the fusion center. Here $K = 10$, $\mathrm{SNR} = 10$ dB, $\gamma = 10$, and $\boldsymbol{R}_{\boldsymbol{n}_r} = \boldsymbol{R}_{\boldsymbol{n}_d} = \sigma^2 \boldsymbol{I}$. Also, MSE is averaged over different channel realizations.

Finally, Fig. 8 shows the minimum network power, $P_{\min}$, versus the number of antennas at the fusion center. In this figure $K = 10$, $\mathrm{SNR} = 10$ dB, $\gamma = 10$, $\boldsymbol{R}_{\boldsymbol{n}_r} = \boldsymbol{R}_{\boldsymbol{n}_d} = \sigma^2 \boldsymbol{I}$, and the MSE is averaged over different channel realizations from the sensors to the fusion center. It is clear that as the number of antennas at the fusion center increases, the minimum power required to achieve the upper bound on MSE, decreases. In other words, increasing the number of antennas at the fusion center decreases the network power.

IX. Conclusions

In this paper, a WSN is considered in which sensor observations and operations are linear and a coherent MAC is used. We design sensor precoders and filter at the fusion center jointly in order to estimate an unknown scalar random source for the two cases of: Minimization of the mean square error (MSE) of BLUE estimator subject to the network power constraint, and minimization of the network power subject to the quality of service (QOS). For a fusion center with multiple antennas, iterative solutions are provided and it is proved that the proposed algorithms always converge. Also, closed form solutions are provided for a fusion center with single antenna, and it is shown that the iterative solutions will reduce to the closed form solutions. Furthermore, the effect of noise correlation at the sensors and fusion center is investigated. It is shown that knowledge of noise correlation at the sensors will help to improve the system performance. Moreover, if correlation exists and we are not aware of that, the system performance might improve depending on the correlation structure. It is also shown, by simulations, that when noise at the fusion center is correlated, even with knowing the structure of the correlation, the system performance degrades. Finally, simulations are provided to verify the analysis and present the performance of the proposed schemes.

Appendix A
Power Constraint Proposition

In this appendix we show that the optimization in (13) is achieved with the equality power constraint. In order to show this we start with the optimization in (13) which is

$$\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \quad \text{s.t.} \quad \boldsymbol{f}^T \underbrace{\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot \boldsymbol{I}\right]}_{\boldsymbol{X}} \boldsymbol{f}^* \leq P_T. \tag{75}$$

We prove that if $\boldsymbol{f}_o$ is the solution of the optimization problem

$$\max_{\boldsymbol{f}} \frac{P_T|\boldsymbol{f}^T\boldsymbol{e}|^2}{P_T\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}\boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^*} \tag{76}$$

then $\boldsymbol{f}_{oo} \triangleq \sqrt{\frac{P_T}{\boldsymbol{f}_o^T\boldsymbol{X}\boldsymbol{f}_o^*}}\boldsymbol{f}_o$ solves the original optimization in (75). This means that $\boldsymbol{f}_{oo}^T\boldsymbol{X}\boldsymbol{f}_{oo}^* = P_T$ and the optimum solution to the original optimization problem is achieved with equal power constraint.

Proof: For any vector $\boldsymbol{f}$ satisfying the power constraint $\boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^* \leq P_T$, one can show that

$$\begin{aligned}
&\frac{|\boldsymbol{f}_{oo}^T\boldsymbol{e}|^2}{\boldsymbol{f}_{oo}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}_{oo}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \\
&= \frac{P_T|\boldsymbol{f}_o^T\boldsymbol{e}|^2}{P_T\boldsymbol{f}_o^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}_o^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}\boldsymbol{f}_o^T\boldsymbol{X}\boldsymbol{f}_o^*} \\
&\geq \frac{P_T|\boldsymbol{f}^T\boldsymbol{e}|^2}{P_T\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}\boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^*} \\
&\geq \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}}.
\end{aligned} \tag{77}$$

The equality holds by just substituting $\boldsymbol{f}_{oo} \triangleq \sqrt{\frac{P_T}{\boldsymbol{f}_o^T\boldsymbol{X}\boldsymbol{f}_o^*}}\boldsymbol{f}_o$ in the equation. The first inequality holds because of the definition of $\boldsymbol{f}_o$ which is the optimal solution to the optimization in (76). The second equality holds since any other vector $\boldsymbol{f}$ is such that $\frac{\boldsymbol{f}^T\boldsymbol{X}\boldsymbol{f}^*}{P_T} \leq 1$. Therefore $\boldsymbol{f}_{oo}$ is the optimal solution to the optimization problem in (75) and satisfies the power constraint with equality which is $\boldsymbol{f}_{oo}^T\boldsymbol{X}\boldsymbol{f}_{oo}^* = P_T$.

Appendix B
Proof of the Optimization

The goal is to provide the solution to the optimization introduced in (14) which is,

$$\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \quad \text{s.t.} \quad \boldsymbol{f}^T\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]\boldsymbol{f}^* = P_T. \tag{78}$$

To solve the above optimization problem, we incorporate the power constraint in the denominator of the cost function, where we have

$$\begin{aligned}
&\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} \\
&= \max_{\boldsymbol{f}} \frac{\boldsymbol{f}^T\boldsymbol{e}\boldsymbol{e}^H\boldsymbol{f}^*P_T}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^*P_T + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}P_T} \\
&= \max_{\boldsymbol{f}} \frac{\boldsymbol{f}^T\boldsymbol{e}\boldsymbol{e}^H\boldsymbol{f}^*P_T}{\boldsymbol{f}^T[\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^HP_T + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}[(\mathbf{1}\mathbf{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r})\odot\boldsymbol{I}]]\boldsymbol{f}^*} \\
&= \max_{\boldsymbol{f}} \frac{\boldsymbol{f}^T\boldsymbol{A}\boldsymbol{f}^*}{\boldsymbol{f}^T\boldsymbol{B}\boldsymbol{f}^*},
\end{aligned} \tag{79}$$

where $\boldsymbol{A} = \boldsymbol{e}\boldsymbol{e}^HP_T$ and $\boldsymbol{B} = \boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^HP_T + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}\left[\left(\mathbf{1}\mathbf{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r}\right)\odot\boldsymbol{I}\right]$. Note that the matrix $\boldsymbol{A}$ is Hermitian and positive semi-definite and matrix $\boldsymbol{B}$ is Hermitian and positive definite.

By using generalized eigenvalue problem [27], the solution to the optimization in (78) can be expressed as

$$\begin{aligned}
\boldsymbol{f}^* &\propto \mathcal{P}(\boldsymbol{B}^{-1}\boldsymbol{A}) \\
&= \alpha(\boldsymbol{B}^{-1}\boldsymbol{e})
\end{aligned} \tag{80}$$

where $\mathcal{P}$ is the principle eigenvector associated to the highest eigenvalue and α ensures that the power constraint is satisfied. Therefore, the optimal sensor gains can be expressed as

$$\boldsymbol{f} = \alpha(\boldsymbol{B}^{-1}\boldsymbol{e})^*, \tag{81}$$

where $\alpha^2 = \frac{P_T}{\boldsymbol{e}^H\boldsymbol{B}^{-H}[(\mathbf{1}\mathbf{1}^T\sigma_\theta^2 + \boldsymbol{R}_{\boldsymbol{n}_r})\odot\boldsymbol{I}]\boldsymbol{B}^{-1}\boldsymbol{e}}$.

Accordingly, the maximum value of the cost function can be expressed as

$$\begin{aligned}
\max_{\boldsymbol{f}} \frac{|\boldsymbol{f}^T\boldsymbol{e}|^2}{\boldsymbol{f}^T\boldsymbol{E}\boldsymbol{R}_{\boldsymbol{n}_r}\boldsymbol{E}^H\boldsymbol{f}^* + \boldsymbol{z}^H\boldsymbol{R}_{\boldsymbol{n}_d}\boldsymbol{z}} &= \lambda_{\max}[\boldsymbol{B}^{-1}\boldsymbol{A}] \\
&= P_T\lambda_{\max}[\boldsymbol{e}^H\boldsymbol{B}^{-1}\boldsymbol{e}] \\
&= P_T\boldsymbol{e}^H\boldsymbol{B}^{-1}\boldsymbol{e}.
\end{aligned} \tag{82}$$

Finally the minimum MSE for the BLUE estimator introduced in (8) for a given $\boldsymbol{z}$ can be expresses as

$$\mathrm{MSE}_{|\boldsymbol{z}} = \frac{1}{P_T\boldsymbol{e}^H\boldsymbol{B}^{-1}\boldsymbol{e}}. \tag{83}$$

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarsubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, pp. 393–422, Mar. 2002.

[2] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, pp. 2292–2330, Aug. 2008.

[3] L. J. Guibas, "Sensing, tracking and reasoning with relations," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 73–85, Mar. 2002.

[4] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.

[5] D. Li, K. D. Wang, Y. H. Hu, and A. M. Sayeed, "Detection, classification and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.

[6] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors. Part I: Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[7] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag, 1994.

[8] I. Bahceci and A. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1146–1156, Jul. 2008.

[9] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4683–4695, Sep. 2007.

[10] J. Li and G. AlRegib, "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3746–3758, Oct. 2009.

[11] J. Fang and H. Li, "Power constrained distributed estimation with correlated sensor data," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3292–3297, Aug. 2009.

[12] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.

[13] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 27–41, Jul. 2006.

[14] M. F. A. Ahmed, T. Y. Al-Naffouri, and M.-S. Alouini, "On the effect of correlated measurements on the performance of distributed estimation," *IEEE Trans. Inf. Theory*, Apr. 2013, submitted for publication.

[15] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Comput. Netw.*, vol. 45, pp. 245–259, 2004.

[16] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 757–770, Feb. 2008.

[17] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Joint estimation in sensor networks under energy constraints," *IEEE Sensor and Ad Hoc Commun. and Networks (SECON)*, pp. 264–271, 2004.

[18] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," *Comput. Netw.*, vol. 50, no. 7, pp. 877–897, May 2006.

[19] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4258–4265, Oct. 2009.

[20] "U.S. Department Energy Assessment study on sensors and automation in the industries of the future," Office of Energy and Renewable Energy Rep., 2004.

[21] A. Krasnopeev, J.-J. Xiao, and Z.-Q. Luo, "Minimum energy decentralized estimation in sensor network with correlated sensor noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2005, vol. 3, pp. 673–676.

[22] J. Fang and H. Li, "Power constrained distributed estimation with cluster-based sensor collaboration," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3822–3832, Jul. 2009.

[23] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, pp. 1684–1696, May 2005.

[24] J. Yeh, *Real Analysis: Theory of Measure and Integration*. Singapore: World Scientific, 2006.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[26] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood cliffs, NJ, USA: Prentice-Hall, 1993.

[27] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. San Diego, CA, USA: Academic, 1979.

[28] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 733–751, Jul. 1989.

[29] S. Barghi, H. Jafarkhani, and H. Yousefi'zadeh, "MIMO-assisted MPR-aware MAC design for asynchronous WLANs," *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1652–1665, Dec. 2011.

Alireza S. Behbahani (S'07–M'10) received the B.Sc degree in electrical engineering from Tehran Polytechnic, Tehran, Iran, in 1992, the M.S. degree with honors in electrical engineering from K.N.Toosi University of Technology, Tehran, Iran, in 1996, and the Ph.D. degree in electrical engineering from University of California, Irvine, in 2009. He is currently a postdoctoral scholar at the department of Electrical Engineering and Computer Science, University of California, Irvine, and a member of the Center for Pervasive Communications and Computing (CPCC).

His research interests lie in the general areas of signal processing, wireless communications, wireless sensor networks, and estimation/detection. His main interests are applications that require an interdisciplinary approach, where expertise in a range of subjects is an advantage.

Dr. Behbahani has been on the technical program committees and reviewer for several symposia, conferences, and journals in the area of signal processing and wireless communications.

Ahmed M. Eltawil (S'97–M'03–SM'14) received his Ph.D. from the University of California, Los Angeles, in 2003. Since 2005, he has been with the Department of Electrical Engineering and Computer Science, University of California, Irvine. He is the founder and director of the Wireless Systems and Circuits Laboratory (http://newport.eecs.uci.edu/aeltawil/). His current research interests are in low power digital circuit and signal processing architectures for wireless communication systems. He has been on the technical program committees and steering committees for numerous workshops, symposia and conferences in the area of VLSI, and communication system design. He received several distinguished awards, including the NSF CAREER award in 2010 supporting his research in low power systems.

Hamid Jafarkhani (F'06) is a Chancellor's Professor at the Department of Electrical Engineering and Computer Science, University of California, Irvine, where he is also the Director of Center for Pervasive Communications and Computing and the Conexant-Broadcom Endowed Chair.

Dr. Jafarkhani ranked first in the nationwide entrance examination of Iranian universities in 1984. He was a co-recipient of the American Division Award of the 1995 Texas Instruments DSP Solutions Challenge. He received an NSF Career Award in 2003. He received the UCI Distinguished Mid-Career Faculty Award for Research in 2006 and the School of Engineering Fariborz Maseeh Best Faculty Research Award in 2007. Also, he was a co-recipient of the 2002 best paper award of ISWC, the 2006 IEEE Marconi Best Paper Award in Wireless Communications, the 2009 best paper award of the *Journal of Communications and Networks*, the 2012 IEEE Globecom best paper award (Communication Theory Symposium), the 2013 IEEE Eric E. Sumner Award, and the 2014 IEEE Communications Society Award for Advances in Communication.

He is listed as a highly cited researcher in http://www.isihighlycited.com. According to the Thomson Scientific, he is one of the top 10 most-cited researchers in the field of "computer science" during 1997–2007. He is a Fellow of AAAS and the author of the book *Space-Time Coding: Theory and Practice*.