

Massive MIMO for Decentralized Estimation over Coherent Multiple Access Channels

Amirpasha Shirazinia¹, Subhrakanti Dey¹, Domenico Ciuonzo² and Pierluigi Salvo Rossi³

¹Signals & Systems Division, Uppsala University, Sweden, ²University of Naples, Italy, ³Dept. Electronics & Telecommunications, NTNU, Norway

Abstract—We consider a decentralized multisensor estimation problem where L sensor nodes observe noisy versions of a possibly correlated random source. The sensors amplify and forward their observations over a fading coherent multiple access channel (MAC) to a fusion center (FC). The FC is equipped with a large array of N antennas, and adopts a minimum mean square error (MMSE) approach for estimating the source. We optimize the amplification factor (or equivalently transmission power) at each sensor node in two different scenarios: 1) with the objective of total power minimization subject to mean square error (MSE) of source estimation constraint, and 2) with the objective of minimizing MSE subject to total power constraint. For this purpose, we apply an asymptotic approximation based on the massive multiple-input-multiple-output (MIMO) favorable propagation condition (when $L \ll N$). We use convex optimization techniques to solve for the optimal sensor power allocation in 1) and 2). In 1), we show that the total power consumption at the sensors decay as $1/N$, replicating the power savings obtained in Massive MIMO mobile communications literature. Through numerical studies, we also illustrate the superiority of the proposed optimal power allocation methods over uniform power allocation.

Index Terms—Decentralized estimation, Massive MIMO, Coherent MAC, Optimization, Power allocation.

I. INTRODUCTION

Wireless sensor networks have recently attracted much research interest due to their practical popularity in accomplishing autonomous tasks, such as monitoring, sensing, computation and communication. In this field, different variety of schemes for decentralized estimation of sources using multiple sensors have been proposed, e.g. [1]–[3]. One popular technique is analog amplify and forward [4], where sensors transmit over fading channels a scaled version of their analog measurements to a fusion center (FC), and it has been shown to be optimal in some situations [5]. Analog forwarding has been investigated under different multiple access schemes such as coherent multiple access [4], [6], which is the focus of the current work, and orthogonal multiple access [7], [8].

In the aforementioned work for analog forwarding of sensor measurements over coherent multiple access channels (MACs), the FC is typically equipped with a single antenna. It is, however, well-known that using multiple antennas can increase spectral efficiency of a wireless system through spatial multiplexing. Decentralized estimation over fading MAC where the FC is equipped with multiple antennas has been studied in [9]. Recently, there has been a vast interest to equip the FC (or the base station in cellular communication framework) with large arrays of antennas, also known as the massive multiple-input multiple output (MIMO) framework [10]–[12]. The use of arrays with massive number of antennas in wireless communication does not only increase spectral efficiency, but it can also improve energy efficiency of MIMO system.

With the assumption of employing massively many number of antenna arrays, known results in MIMO communication systems can be considerably simplified [12], [13]. As a result, it provides analytical solutions to problems that would otherwise be mathematically intractable. In wireless sensor networks, the massive MIMO framework has also been recently employed for decentralized detection and estimation [14]–[16]. While in [15], the authors have studied decentralized detection problem, they show similar results for decentralized estimation of a *scalar deterministic unknown source*. In the current paper, within the massive MIMO framework, we focus on the decentralized estimation problem in a more general case where the source is modeled as a *vector* comprised of *random* and *correlated* components. Our main objective is to optimally design the sensors' amplification factors with respect to minimizing sensor power consumptions or maximizing estimation accuracy subject to relevant constraints. In our setting, L sensor nodes observe noisy versions of a possibly correlated random source vector. The sensor nodes amplify the observations according to their power budget, and forward them over coherent fading MACs to an FC equipped with a large number of antennas, denoted by N . The FC estimates the source by adopting a minimum mean square error (MMSE) estimator. We optimize the amplification factor (or equivalently the transmission power allocation) at each sensor node 1) with the objective of total power minimization subject to a maximum MSE constraint (incurred by using the MMSE estimator), and 2) with the objective of minimizing MSE of source estimation subject to a total power constraint. For this purpose, we apply an asymptotic approximation based on the *favorable propagation conditions in the massive MIMO literature* in order to simplify the MSE expression when N is sufficiently large. Using the resulting asymptotic approximation, our contributions are as follows:

- We show that the optimization problems are convex, which, in general, can be solved numerically using well-known convex optimization techniques.
- We analyze the optimization problems, and under some conditions derive closed-form solutions to them. Our analysis reveals that, in the studied framework, as the number of antennas at the FC, N , increases, the sensors can decrease their total power consumption with a factor proportional to $1/N$ in order to satisfy a targeted MSE.
- It is also seen that using the massive MIMO approximation, the optimal sensor power allocation only depends on the distance based attenuation components and not on the randomly varying fading gains.

Notations: We denote vectors and matrices by bold lower-case and upper-case letters, respectively. The matrix trace is

denoted by $\text{Tr}(\cdot)$, and matrix/vector conjugate-transpose (and transpose) by $(\cdot)^H$ (and $(\cdot)^\top$). The notation $\text{diag}(a_1, a_2, \dots, a_n)$ is used for a diagonal matrix whose diagonal elements are a_1, \dots, a_n . We use $\mathbb{E}[\cdot]$ to denote the expectation operator. $[\mathbf{A}]_{ij}$ means the element of the matrix \mathbf{A} at i^{th} row and j^{th} column. The notation $\mathbf{X} \succeq \mathbf{0}$ means that the matrix \mathbf{X} is a positive semi-definite matrix. $[\cdot]^+$ denotes $\max\{0, \cdot\}$. The circularly-symmetric Gaussian distribution is denoted by \mathcal{CN} . We denote equality in an asymptotic sense by $\stackrel{a}{\approx}$.

II. SYSTEM DESCRIPTION

We study the system with coherent MACs shown in Figure 1. Based on the studied system model in Figure 1, the received signal at the FC can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{D}\boldsymbol{\theta} + \underbrace{\mathbf{H}\mathbf{D}\mathbf{n}}_{\triangleq \mathbf{w}} + \mathbf{v}. \quad (1)$$

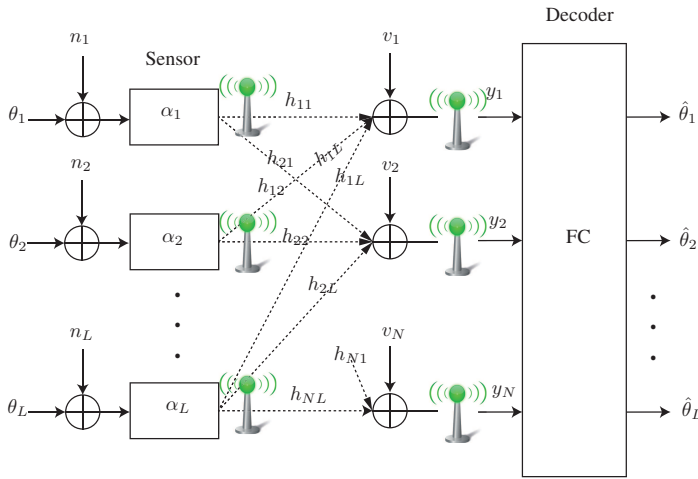


Fig. 1: System model for decentralized estimation over coherent MAC with analog forwarding.

In (1), the source is characterized by the vector $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^\top \in \mathbb{C}^L$, where $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_\theta)$ and $\mathbf{C}_\theta \in \mathbb{C}^{L \times L}$ is the positive definite source covariance matrix which is not necessarily diagonal. The l^{th} source-to-sensor noise component is denoted by n_l . Stacking up all the noise components, we denote the source-to-sensor noise vector by $\mathbf{n} = [n_1, \dots, n_L]^\top$, where $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_n)$ and $\mathbf{C}_n \triangleq \text{diag}(\sigma_{n_1}^2, \dots, \sigma_{n_L}^2)$ is the source-to-sensor noise covariance matrix. α_l is the complex amplification gain at the l^{th} sensor node, and $\mathbf{D} \triangleq \text{diag}(\alpha_1, \dots, \alpha_L)$. We denote the channel matrix by $\mathbf{H} \in \mathbb{C}^{N \times L}$, where $[\mathbf{H}]_{il}$ corresponds to the channel gain between the i^{th} antenna at the FC and the l^{th} sensor. The channel matrix \mathbf{H} models independent fast fading, geometric attenuation and log-normal shadow fading. The generic coefficient $[\mathbf{H}]_{il}$ is then expressed as

$$[\mathbf{H}]_{il} = \frac{1}{d_l^\beta} [\mathbf{G}]_{il}, \quad i = 1, \dots, N, \quad l = 1, \dots, L, \quad (2)$$

where d_l is the distance between the l^{th} sensor to the FC, and β is the pathloss exponent. Furthermore, $[\mathbf{G}]_{il}$ are independent and identically distributed (i.i.d.) random variables drawn from $\mathcal{CN}(0, 1)$. Based on the above assumptions, we have $\mathbf{H} = \mathbf{G}\boldsymbol{\Gamma}^{1/2}$, where $\boldsymbol{\Gamma} \in \mathbb{R}^{N \times N}$ is a diagonal matrix containing the attenuation coefficients $d_l^{-2\beta}$, $l = 1, \dots, L$, on its main diagonal.

The additive Gaussian noise at the FC is denoted by $\mathbf{v} = [v_1, \dots, v_N]^\top$ with distribution $\mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_N)$. Using N antennas, the FC provides an estimate of the source vector from the

received signal vector \mathbf{y} in (1). We assume that the FC has the perfect knowledge of the source and noise statistics as well as the channel gains. Hence, the minimum mean-square error (MMSE) estimator can be applied in order to give the lowest possible MSE. In the next section, we show the resulting MSE and discuss our design method for allocating optimal power to sensors.

III. PROBLEM FORMULATION

Using the MMSE estimator at the FC, the estimated vector $\hat{\boldsymbol{\theta}} \triangleq [\hat{\theta}_1, \dots, \hat{\theta}_L]^\top$ is obtained as [17, Chapter 11]

$$\hat{\boldsymbol{\theta}} = \mathbb{E}[\boldsymbol{\theta} | \mathbf{H}, \mathbf{y}] = (\mathbf{C}_\theta^{-1} + \mathbf{D}^H \mathbf{H}^H \mathbf{C}_w^{-1} \mathbf{H} \mathbf{D})^{-1} \mathbf{D}^H \mathbf{H}^H \mathbf{C}_w^{-1} \mathbf{y}, \quad (3)$$

which gives the following MSE $\triangleq \mathbb{E}[\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|_2^2 | \mathbf{H}]$

$$\text{MSE} = \text{Tr} \left\{ (\mathbf{C}_\theta^{-1} + \mathbf{D}^H \mathbf{H}^H \mathbf{C}_w^{-1} \mathbf{H} \mathbf{D})^{-1} \right\}, \quad (4)$$

where $\mathbf{C}_w = \mathbb{E}[\mathbf{w}\mathbf{w}^H] = \mathbf{H}\mathbf{D}\mathbf{C}_n\mathbf{D}^H\mathbf{H}^H + \mathbf{C}_v$. By using matrix inversion lemma on \mathbf{C}_w^{-1} , (4) can be rewritten as

$$\text{MSE} = \text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + \sigma_v^{-2} \mathbf{D}^H \mathbf{H}^H \mathbf{H} \mathbf{D} - \sigma_v^{-4} \mathbf{D}^H \mathbf{H}^H \mathbf{H} \mathbf{D} (\mathbf{C}_n^{-1} + \sigma_v^{-2} \mathbf{D}^H \mathbf{H}^H \mathbf{H} \mathbf{D})^{-1} \mathbf{D}^H \mathbf{H}^H \mathbf{H} \mathbf{D} \right)^{-1} \right\}. \quad (5)$$

Now, since $[\mathbf{G}]_{il} \stackrel{\text{i.i.d.}}{\sim} \mathcal{CN}(0, 1)$, ($l = 1, \dots, L, i = 1, \dots, N$), then using massive MIMO framework, as $N \rightarrow \infty$ while L remains fixed ($L \ll N$) the so-called *favorable propagation conditions* hold [11], and it can be shown that

$$\mathbf{H}^H \mathbf{H} = \boldsymbol{\Gamma}^{1/2} \mathbf{G}^H \mathbf{G} \boldsymbol{\Gamma}^{1/2} \stackrel{a}{\approx} N \boldsymbol{\Gamma}. \quad (6)$$

Now, we define $\boldsymbol{\Lambda} \triangleq \mathbf{D}^H \boldsymbol{\Gamma} \mathbf{D} = \text{diag} \left(\frac{|\alpha_1|^2}{d_1^{2\beta}}, \dots, \frac{|\alpha_L|^2}{d_L^{2\beta}} \right)$ and $\lambda_l \triangleq \frac{|\alpha_l|^2}{d_l^{2\beta}}$, $\forall l$. Hence, the MSE in (5) can be asymptotically written as

$$\text{MSE} \stackrel{a}{\approx} \text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + N \sigma_v^{-2} \boldsymbol{\Lambda} - N^2 \sigma_v^{-4} \boldsymbol{\Lambda} (\mathbf{C}_n^{-1} + N \sigma_v^{-2} \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda} \right)^{-1} \right\}. \quad (7)$$

We note that from now on whenever we use the term MSE, we mean the asymptotic MSE shown by (7).

Remark 1. Using the massive MIMO approximation (6), the MSE formulation is simplified (cf. (7)). Hence, power allocation optimization problems, i.e., optimal design of the sensors' amplification factor α_l , (as we will see later in the subsequent sections) becomes mathematically more tractable and easier to solve.

The total power consumed by L sensors can be written as

$$P_{\text{tot}} = \mathbb{E}[\|\mathbf{H}\mathbf{D}\boldsymbol{\theta}\|_2^2] = \sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}. \quad (8)$$

Now, we pose the following two optimization problems dealing with sensor power allocation. The first problem minimizes the total power consumption by the sensors subject to reconstruction MSE constraint. Therefore, this optimization problem is desirable when power is limited, and there is not a tight restriction on MSE. In the second problem, the MSE is minimized subject to total power constraint. Hence, this problem formulation is posed when the estimation accuracy, in

terms of MSE, is more demanding. More specifically, the first optimization problem is stated as follows

$$\underset{\{\lambda_l \geq 0\}_{l=1}^L}{\text{minimize}} P_{tot}, \text{ subject to } \text{MSE} \leq \bar{d}. \quad (\text{P1})$$

where P_{tot} and MSE are specified by (8) and (7), respectively. Further, \bar{d} is a user-defined distortion threshold, and has to be chosen such that $\bar{d} \leq \text{Tr}\{\mathbf{C}_\theta\}$.

The second optimization problem is stated as follows

$$\underset{\{\lambda_l \geq 0\}_{l=1}^L}{\text{minimize}} \text{MSE}, \text{ subject to } P_{tot} \leq \bar{P}, \quad (\text{P2})$$

We note that after solving the optimization problems (P1) and (P2) for λ_l , one can recover the complex gain α_l with amplitude $|\alpha_l|$ and an arbitrary phase. It can be easily shown that the optimization problems (P1) and (P2) are convex in the variables λ_l , $l = 1, \dots, L$.

IV. ANALYSIS OF POWER MINIMIZATION PROBLEM WITH MSE CONSTRAINT

A. Correlated Source

1) Exact Analysis

First, we consider the general case where the elements of the source vector are correlated, i.e., \mathbf{C}_θ is not diagonal. In order to solve for λ_l in (P1), we write the Karush-Kuhn-Tucker (KKT) conditions [18] by introducing the Lagrange multiplier $\mu \geq 0$. We first write the Lagrangian as

$$\begin{aligned} \mathcal{L}(\mathbf{A}, \mu) &= \sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} + \\ &\mu \left(\text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + N\sigma_v^{-2} \mathbf{A} - N^2 \sigma_v^{-4} \mathbf{A} (\mathbf{C}_n^{-1} + N\sigma_v^{-2} \mathbf{A})^{-1} \mathbf{A} \right)^{-1} \right\} - \bar{d} \right) \end{aligned} \quad (9)$$

Taking the partial derivative of (9) with respect to λ_l ($l = 1, \dots, L$), we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{A}, \mu)}{\partial \lambda_l} &= -\mu \text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + \text{diag} \left(\dots, \frac{N\lambda_l}{\sigma_v^2 + N\sigma_{n_l}^2 \lambda_l}, \dots \right) \right)^{-2} \times \right. \\ &\left. \text{diag} \left(0, \dots, 0, \frac{N\sigma_v^{-2} \sigma_{n_l}^{-4}}{(\sigma_{n_l}^{-2} + N\sigma_v^{-2} \lambda_l)^2}, 0, \dots, 0 \right) \right\} + d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} \end{aligned} \quad (10)$$

where we used the fact that for a positive-definite matrix \mathbf{A} , we have $\frac{d\mathbf{A}^{-1}}{dt} = -\mathbf{A}^{-1} \frac{d\mathbf{A}}{dt} \mathbf{A}^{-1}$ in which t is an element of the matrix \mathbf{A} . Letting (10) equal zero, and observing that $\lambda_l \geq 0$, it yields the following set of implicit non-linear equations for $l = 1, 2, \dots, L$

$$\begin{aligned} \lambda_l &= \left[\left\{ \frac{\mu \sigma_v^2 \sigma_{n_l}^{-4}}{N d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}} \times \right. \right. \\ &\left. \left. \left[\left(\mathbf{C}_\theta^{-1} + \text{diag} \left(\dots, \frac{N\lambda_l}{\sigma_v^2 + N\sigma_{n_l}^2 \lambda_l}, \dots \right) \right)^{-2} \right]_{ll} \right\}^{1/2} - \frac{\sigma_v^2}{N\sigma_{n_l}^2} \right] \end{aligned} \quad (11)$$

Since \mathbf{C}_θ is not diagonal, for an arbitrary Lagrange multiplier μ , (11) can be solved using non-linear equation solvers (e.g., `fsolve` in MATLAB). The optimal μ is determined such that the MSE constraint in (P1) is satisfied with equality.

2) Approximate Analysis

Here we propose an approach in order to asymptotically solve $\{\lambda_l\}_{l=1}^L$ when N is sufficiently large. For this purpose, we

first write the first-order Taylor series expansion of MSE in (7) around $1/N \rightarrow 0^1$ which yields

$$\begin{aligned} \text{MSE} &\approx \text{Tr} \{ (\mathbf{C}_\theta^{-1} + \mathbf{C}_n^{-1})^{-1} \} \\ &+ \frac{1}{N} \text{Tr} \left\{ (\mathbf{C}_\theta^{-1} + \mathbf{C}_n^{-1})^{-2} \text{diag} \left(\dots, \frac{\sigma_v^2}{\sigma_{n_l}^4 \lambda_l}, \dots \right) \right\}, \end{aligned} \quad (12)$$

where we have implicitly assumed that sensors always amplify their observations with positive (non-zero) gains, i.e., $\lambda_l > 0$.

Here, $\mathbf{C}_n = \text{diag}(\sigma_{n_1}^2, \dots, \sigma_{n_L}^2)$. For brevity, we define $\mathbf{Q} \triangleq (\mathbf{C}_\theta^{-1} + \mathbf{C}_n^{-1})^{-1}$. Then, by plugging (12) back into the constraint in (P1), we solve the following optimization problem for $\{\lambda_l\}_{l=1}^L$

$$\begin{aligned} \underset{\{\lambda_l > 0\}_{l=1}^L}{\text{minimize}} &\sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} \\ \text{subject to} &\text{Tr}\{\mathbf{Q}\} + \frac{1}{N} \text{Tr} \left\{ \mathbf{Q}^2 \text{diag} \left(\dots, \frac{\sigma_v^2}{\sigma_{n_l}^4 \lambda_l}, \dots \right) \right\} \leq \bar{d}. \end{aligned} \quad (\text{P3})$$

Theorem 2. *The optimal solution to Problem (P3) is given by*

$$\lambda_l^* = \frac{1}{N} \left[\frac{\frac{[\mathbf{Q}^2]_{ll}^{1/2} \sigma_v^2}{\sigma_{n_l}^2 [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} d_l^{2\beta}} \sum_{l=1}^L \frac{d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}^{1/2} [\mathbf{Q}^2]_{ll}^{1/2}}{\sigma_{n_l}^2}}{\bar{d} - \sum_{l=1}^L [\mathbf{Q}]_{ll}} \right]. \quad (13)$$

Proof. Please see Appendix A. \square

Remark 3. *By studying (13) in Theorem 2, it can be observed that the total power consumed by all sensors decays with a factor proportional to $1/N$, when N , the number of antennas at the FC, is sufficiently large.*

B. Uncorrelated Source

Here we assume that the elements of the source vector θ are uncorrelated, such that the source covariance matrix $\mathbf{C}_\theta \triangleq \text{diag}(\sigma_{\theta_1}^2, \dots, \sigma_{\theta_L}^2)$. In this case, the optimal solution for λ_l , $l = 1, 2, \dots, L$ can be obtained in closed-form analytically, and is stated in the following theorem.

Theorem 4. *Provided $\frac{\sigma_{\theta_l}^2}{d_l^{2\beta} (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}$ is ordered decreasingly in $l \in \{1, \dots, L\}$, \exists a unique M^* such that*

$$\begin{aligned} M^* &= \max \left\{ M \in \{1, \dots, L\} : \frac{\sigma_{\theta_l}^2}{d_l^{2\beta} (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}} > \right. \\ &\left. \frac{\bar{d} - \sum_{l=1}^M \frac{1}{\sigma_{\theta_l}^2 + 1/\sigma_{n_l}^2} - \sum_{l=M+1}^L \sigma_{\theta_l}^2}{\sum_{l=1}^M \frac{\sigma_{\theta_l}^2 d_l^{2\beta}}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}} \right\}. \end{aligned} \quad (14)$$

Then, the optimal solution is given by

$$\begin{aligned} \lambda_l^* &= \frac{1}{N} \left[\frac{\frac{\sigma_{\theta_l}^2 \sigma_v^2}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{3/2} d_l^{2\beta}} \sum_{l=1}^{M^*} \frac{\sigma_{\theta_l}^2 d_l^{2\beta}}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}}{\bar{d} - \sum_{l=1}^{M^*} 1/[1/\sigma_{\theta_l}^2 + 1/\sigma_{n_l}^2] - \sum_{l=M^*+1}^L \sigma_{\theta_l}^2} \right. \\ &\left. - \frac{\sigma_v^2}{\sigma_{\theta_l}^2 + \sigma_{n_l}^2} \right], \end{aligned} \quad (15)$$

¹The reason for expanding the Taylor series around $1/N \rightarrow 0$ is due to massive MIMO framework with a large number of antennas N , and for a tight approximation, we can neglect higher-order terms in the series.

for $l = 1, \dots, M^*$, and $\lambda_l^* = 0$ for $l = M^* + 1, \dots, L$. Further, the amplitude of the optimal amplification gain for sensor l becomes $|\alpha_l^*| = d_l^\beta \lambda_l^*$.

Proof. Please see Appendix B. \square

Remark 5. From (15) in Theorem 4, it can be realized that the total power consumed by all sensors decreases exactly with a factor proportional to $1/N$.

V. ANALYSIS OF MSE MINIMIZATION PROBLEM WITH POWER CONSTRAINT

A. Correlated Source

1) Exact Analysis

Similar to the analysis in Section IV-A1, by introducing the Lagrange multiplier $\mu \geq 0$, we write the Lagrangian as

$$\begin{aligned} \mathcal{L}(\mathbf{\Lambda}, \mu) = & \mu \left(\sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} - \bar{P} \right) \\ & + \text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + N\sigma_v^{-2} \mathbf{\Lambda} - N^2 \sigma_v^{-4} \mathbf{\Lambda} (\mathbf{C}_n^{-1} + N\sigma_v^{-2} \mathbf{\Lambda})^{-1} \mathbf{\Lambda} \right)^{-1} \right\}. \end{aligned} \quad (16)$$

Taking the partial derivative of (16) with respect to λ_l ($l = 1, \dots, L$), we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{\Lambda}, \mu)}{\partial \lambda_l} = & -\text{Tr} \left\{ \left(\mathbf{C}_\theta^{-1} + \text{diag} \left(\dots, \frac{N\lambda_l}{\sigma_v^2 + N\sigma_{n_l}^2 \lambda_l}, \dots \right) \right)^{-2} \times \right. \\ & \left. \text{diag} \left(0, \dots, 0, \frac{N\sigma_v^{-2} \sigma_{n_l}^{-4}}{(\sigma_{n_l}^{-2} + N\sigma_v^{-2} \lambda_l)^2}, 0, \dots, 0 \right) \right\} + \mu d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} \end{aligned} \quad (17)$$

and letting (17) equal zero, it yields (since $\lambda_l \geq 0$)

$$\begin{aligned} \lambda_l = & \left[\left\{ \frac{\sigma_v^2 \sigma_{n_l}^{-4}}{N\mu d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}} \times \right. \right. \\ & \left. \left. \left[\left(\mathbf{C}_\theta^{-1} + \text{diag} \left(\dots, \frac{N\lambda_l}{\sigma_v^2 + N\sigma_{n_l}^2 \lambda_l}, \dots \right) \right)^{-2} \right]_{ll} \right\}^{-1} - \frac{\sigma_v^2}{N\sigma_{n_l}^2} \right]^{1/2} \end{aligned} \quad (18)$$

Finally, μ is determined in order to satisfy the power constraint in (P2) with equality.

2) Approximate Analysis

Similar to the analysis in Section IV-A2, we can derive approximate closed-form solution for Problem (P2) when N is sufficiently large. To do so, we follow the approximation of MSE in (12). Hence, the problem can be shown to be simplified into

$$\begin{aligned} & \underset{\{\lambda_l > 0\}_{l=1}^L}{\text{minimize}} \quad \text{Tr} \left\{ \mathbf{Q}^2 \text{diag} \left(\dots, \frac{\sigma_v^2}{\sigma_{n_l}^4 \lambda_l}, \dots \right) \right\} \\ & \text{subject to} \quad \sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} \leq \bar{P}, \end{aligned} \quad (P5)$$

where $\mathbf{Q} = (\mathbf{C}_\theta^{-1} + \mathbf{C}_n^{-1})^{-1}$, and $\mathbf{C}_n = \text{diag}(\sigma_{n_1}^2, \dots, \sigma_{n_L}^2)$.

Note that unlike (P3), the optimal solution to (P5) does not depend on the number of antennas N . We have the following result.

Theorem 6. The optimal solution to (P5) is given by

$$\lambda_l^* = \frac{\bar{P} [\mathbf{Q}^2]_{ll}^{1/2}}{\sigma_{n_l}^2 d_l^\beta [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}^{1/2} \sum_l \frac{[\mathbf{Q}^2]_{ll}^{1/2} d_l^\beta [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}}{\sigma_{n_l}^2}}. \quad (19)$$

Proof. The proof can be closely followed by the proof of Theorem 2. \square

B. Uncorrelated Source

In the case the elements of the source vector $\boldsymbol{\theta}$ are uncorrelated as in Section IVB, the optimal solution to the sensor power allocation can be found in closed-form as follows.

Theorem 7. Provided $\frac{\sigma_{\theta_l}^2}{d_l^\beta (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}$ is ordered decreasingly in $l \in \{1, \dots, L\}$, \exists a unique M^* such that

$$\begin{aligned} M^* = & \max \left\{ M \in \{1, \dots, L\} : \frac{d_l^\beta (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}{\sigma_{\theta_l}^2} < \right. \\ & \left. \frac{N\bar{P}/\sigma_v^2 + \sum_{l=1}^M d_l^\beta}{\sum_{l=1}^M \frac{\sigma_{\theta_l}^2 d_l^\beta}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}} \right\}. \end{aligned} \quad (20)$$

Then, the optimal λ_l is given by

$$\begin{aligned} \lambda_l^* = & \frac{1}{N} \left[\frac{\sigma_v^2 \sigma_{\theta_l}^2 \sum_{l=1}^{M^*} d_l^{2\beta}}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{3/2} d_l^\beta \sum_{l=1}^{M^*} \frac{\sigma_{\theta_l}^2 d_l^\beta}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}} - \frac{\sigma_v^2}{\sigma_{\theta_l}^2 + \sigma_{n_l}^2} \right] \\ & + \frac{\bar{P} \sigma_{\theta_l}^2}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{3/2} d_l^\beta \sum_{l=1}^{M^*} \frac{\sigma_{\theta_l}^2 d_l^\beta}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}} \end{aligned} \quad (21)$$

for $l = 1, \dots, M^*$, and $\lambda_l^* = 0$ for $l = M^* + 1, \dots, L$. Further, the amplitude of the optimal amplification gain for sensor l becomes $|\alpha_l^*| = d_l^\beta \lambda_l^*$.

Proof. The proof can be closely followed by the proof of Theorem 4. \square

Corollary 8. As $N \rightarrow \infty$, the asymptotic MSE incurred by using the optimal power allocation derived in (21) becomes

$$\lim_{N \rightarrow \infty} \text{MSE} = \sum_{l=1}^L \frac{1}{1/\sigma_{\theta_l}^2 + 1/\sigma_{n_l}^2}. \quad (22)$$

Proof. First, it can be verified from (20) that as $N \rightarrow \infty$, $M^* \rightarrow \infty$. Hence, $\forall l$, we obtain $\lambda_l > 0$. Then, we find the limit of the MSE as $N \rightarrow \infty$, which yields (22). \square

VI. NUMERICAL EXPERIMENTS

In this section, we evaluate the performance of the proposed optimization methods. In all simulation studies, we assume $L = 15$ number of sensors. We also assume a homogenous scenario, where $\forall l \in \{1, \dots, L\}$, the source-to-sensor noise variances $\sigma_{n_l}^2 = 10^{-4}$. Further, $\sigma_v^2 = 10^{-4}$. In the uncorrelated source case, the variance of source entries is chosen the same and equal to $\sigma_{\theta_l}^2 = 1$ for each node. We also set the pathloss exponent to $\beta = 2$, and sensor-to-FC distances d_l are uniformly distributed ranging from 2 to 7.

In the correlated source case, we consider the exponential covariance matrix model [19] for the source, where each entry at row i and column j of the source covariance matrix \mathbf{C}_θ is chosen as $\rho^{|i-j|}$ in which $0 < \rho < 1$ is known as correlation coefficient.

We first consider the uncorrelated case. In Figure 2, we plot total power consumed by sensors as a function of number of antennas N (varying from 50 to 200 at a step size of 10) using the optimized power allocation (15) and uniform power

allocation, for varying distortion threshold \bar{d} . For the uniform power allocation, we assume that all sensor nodes consumes equal power. It can be observed that the total power in log-log scale decays linearly in N for both methods. However, in all setups, the optimal power allocation outperforms the uniform power allocation by almost 2 dB.

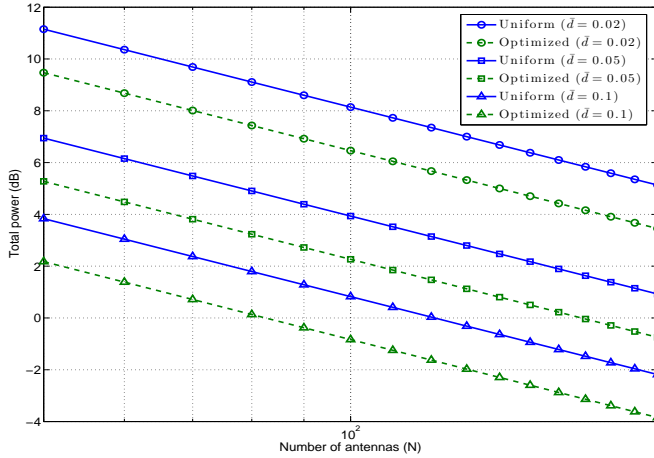


Fig. 2: Total power consumed by sensors as a function of number of antennas N for different values of distortion threshold \bar{d} using optimized design and uniform power allocation.

Next, in Figure 3, we plot the MSE as a function of number of antennas N using the optimized power allocation (21) and uniform power allocation. As can be observed, the optimal power allocation provides a lower MSE compared to the uniform power allocation.

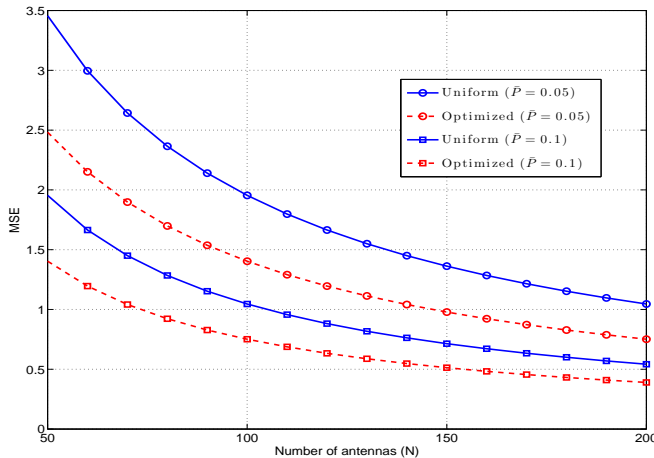


Fig. 3: MSE as a function of number of antennas N at the FC for different values of power threshold \bar{P} using optimized design and uniform power allocation.

In our last experiment, we study the correlated case, and plot, in Figure 4, MSE as function of number of antennas N for different values of the correlation coefficient ρ and for fixed $\bar{P} = 0.1$. The curves in Figure 4 are associated with the exact analysis (by solving (P2) using CVX solver or equivalently by solving the KKT conditions in (18)) shown in solid line, and the approximate analysis (by solving (19)) shown in dashed line, respectively. We observe that the approximate solution is tight in all numerical setups. As can be also expected, higher correlation leads to lower MSE.

VII. EXTENSIONS

The optimal power allocation algorithms studied here can also be extended to cases where the MIMO channels are correlated, or the sensor-to-FC additive noise elements are correlated.

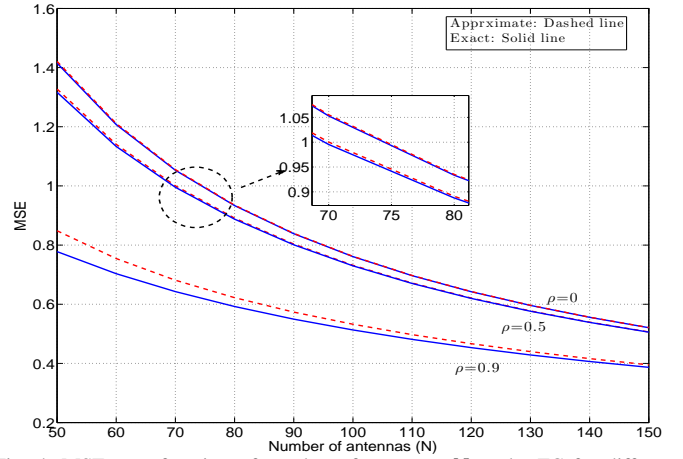


Fig. 4: MSE as a function of number of antennas N at the FC for different values of correlation parameter ρ .

Extension to the case of imperfect channel estimates at the FC can also be readily addressed. These extensions will be reported in a longer version of this paper.

APPENDIX A

PROOF OF THEOREM 2

It can be easily shown that the above optimization problem is convex, therefore, the optimization can be solved by the standard KKT condition technique. By introducing the Lagrange multiplier $\mu \geq 0$, the Lagrangian can be written as

$$\mathcal{L}(\{\lambda_l\}_{l=1}^L, \mu) = \sum_{l=1}^L \lambda_l d_l^{2\beta} [\mathbf{C}_\theta + \mathbf{C}_n]_{ll} + \mu \left(\sum_{l=1}^L [\mathbf{Q}]_{ll} + \frac{1}{N} \sum_{l=1}^L \frac{\sigma_v^2}{\lambda_l \sigma_{n_l}^4} [\mathbf{Q}^2]_{ll} - \bar{d} \right).$$

Taking the derivative of the Lagrangian with respect to $\lambda_l > 0$, and letting it equal zero, we have

$$\lambda_l = \frac{\sqrt{\frac{\mu}{N}} \frac{\sigma_v}{\sigma_{n_l}^2} [\mathbf{Q}^2]_{ll}^{1/2}}{d_l^\beta [\mathbf{C}_\theta + \mathbf{C}_n]_{ll}^{1/2}}. \quad (23)$$

Then, plugging (23) back into the constraint in (P3), the Lagrange multiplier μ is determined, which yields the optimal λ_l^* in (13).

APPENDIX B

PROOF OF THEOREM 4

The optimization can be solved by the standard Karush-Kuhn-Tucker (KKT) condition technique [18] since the objective and the constraint are both convex. By introducing the Lagrange multiplier $\mu \geq 0$, the Lagrangian can be written as

$$\mathcal{L}(\{\lambda_l\}_{l=1}^L, \mu) = \sum_{l=1}^L \lambda_l d_l^{2\beta} (\sigma_{\theta_l}^2 + \sigma_{n_l}^2) + \mu \left(\sum_{l=1}^L \frac{\sigma_{\theta_l}^2 \sigma_{n_l}^2 \lambda_l + \frac{\sigma_{\theta_l}^2 \sigma_v^2}{N}}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2) \lambda_l + \frac{\sigma_v^2}{N}} - \bar{d} \right).$$

Taking the first derivative of the Lagrangian $\mathcal{L}(\{\lambda_l\}_{l=1}^L, \mu)$ with respect to λ_l ($l = 1, \dots, L$), and letting the result equal zero, we obtain

$$\lambda_l = \left[\frac{\sqrt{\frac{\mu}{N}} \sigma_{\theta_l}^2 \sigma_v}{d_l^\beta (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{3/2}} - \frac{\sigma_v^2}{N (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)} \right]^+. \quad (24)$$

Now, we define $z_l \triangleq \frac{\sigma_{\theta_l}^2}{d_l^\beta (\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}} - \frac{\sigma_v}{\sqrt{N\mu}}$ such that $\lambda_l = \frac{\sigma_v \sqrt{\mu/N}}{\sigma_{\theta_l}^2 + \sigma_{n_l}^2} [z_l]^+$. Hence, since $z_1 \geq z_2 \geq \dots \geq z_L$, we assume that the first M^* values of z_l , $l = 1, \dots, L$, are positive, and the

remaining $L - M^*$ values are zero, where $M^* \leq L$. Due to the above assumption, M^* is unique, since the threshold $\frac{\sigma_v}{\sqrt{N}\mu}$ for comparing z_l only depends on N , σ_v and μ , which are constant for a given system.

Plugging (24) back into the MSE constraint, we determine the Lagrange multiplier as

$$\mu = \frac{1}{N} \left[\frac{\sum_{l=1}^{M^*} \frac{\sigma_v \sigma_{\theta_l}^2 d_l^\beta}{(\sigma_{\theta_l}^2 + \sigma_{n_l}^2)^{1/2}}}{\bar{d} - \sum_{l=1}^{M^*} 1/[1/\sigma_{\theta_l}^2 + 1/\sigma_{n_l}^2] - \sum_{l=M^*+1}^L \sigma_{\theta_l}^2} \right]^2, \quad (25)$$

and inserting μ back into (24) yields the optimal values λ_l^* in (15). Now, in order to determine M^* , we let it be equal to the maximum integer in the interval $[1, L]$ such that $z_l > 0$, which yields (14).

REFERENCES

- [1] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part i: Gaussian case," *IEEE Trans. Sig. Proc.*, vol. 54, no. 3, pp. 1131–1143, March 2006.
- [2] I. Schizas, G. Giannakis, and Z.-Q. Luo, "Distributed estimation using reduced-dimensionality sensor observations," *IEEE Trans. Sig. Proc.*, vol. 55, no. 8, pp. 4284–4299, Aug. 2007.
- [3] J.-J. Xiao and Z.-Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. Inf. Theo.*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.
- [4] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Lect. Notes in Comput. Science.* Springer, 2003, pp. 162–177.
- [5] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inf. Theo.*, vol. 49, no. 5, pp. 1147–1158, May 2003.
- [6] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Sig. Proc.*, vol. 56, no. 2, pp. 757–770, Feb. 2008.
- [7] S. Cui, J.-J. Xiao, A. Goldsmith, Z.-Q. Luo, and H. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. Sig. Proc.*, vol. 55, no. 9, pp. 4683–4695, Sep. 2007.
- [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, July 2008.
- [9] A. Smith, M. Banavar, C. Tepedelenlioglu, and A. Spanias, "Distributed estimation over fading macs with multiple antennas at the fusion center," in *Asilomar Conf. Sig., Syst. and Comput.*, Nov. 2009, pp. 424–428.
- [10] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.
- [11] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, April 2013.
- [12] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [13] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Select. Topics Sig. Proc.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [14] F. Jiang, J. Chen, and A. Swindlehurst, "Phase-only analog encoding for a multi-antenna fusion center," in *IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, March 2012, pp. 2645–2648.
- [15] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. Lopez-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *CoRR*, vol. abs/1408.2192, 2014. [Online]. Available: <http://arxiv.org/abs/1408.2192>
- [16] D. Ciuonzo, P. Salvo Rossi, and S. Dey, "Massive MIMO channel-aware decision fusion," *IEEE Trans. Sig. Proc.*, vol. 63, no. 3, pp. 604–619, Feb. 2015.
- [17] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [19] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.