Distributed Massive MIMO for Estimation of a Correlated Source Vector in Sensor Networks

Jordi Serra, David Pubill and Christos Verikoukis, *Senior Member, IEEE*
Telecommunications Technological Center of Catalonia (CTTC), 08860 Castelldefels, Spain
E-mails: {jserra,dpubill,cveri}@cttc.es

Abstract—Distributed massive MIMO (DM-MIMO) systems are a key enabler to improve the energy efficiency (EE) in future wireless networks. Thereby, herein this architecture is considered for the estimation of a correlated source vector in wireless sensor networks (WSN), where each sensor node amplifies and forwards its observation through a coherent Multiple Access Channel (MAC) channel. Namely, the fusion centre (FC) consists of a large number of distributed single antenna access points (AP) connected through a backhaul network to a central processing unit (CPU), where a Linear Minimum Mean Square Error (LMMSE) estimation is computed. Within this setting the exact and an approximated MSE, obtained by the LMMSE estimation, are derived. Bearing in mind these results, we address the design of the optimal power allocation, at each sensor node, to minimize the total transmitted power subject to an MSE estimation constraint. The approximation of the MSE paves the way to cast the optimal allocation problem as a Semidefinite Programming Problem (SDP). Finally, the numerical simulations show that our system permits to reduce significantly the total transmitted power compared to related work architectures proposing a massive MIMO system where all the antennas are collocated at the FC.

Index Terms—Distributed Massive MIMO, power allocation, estimation, sensor networks.

I. INTRODUCTION

WSN have been widely considered for decentralized inference problems such as parameter estimation. Moreover, they are a key ingredient in disruptive technologies such as the Internet of Things (IoT). The sensor nodes of a WSN have a limited energy budget, as they are usually battery powered. Thereby, inference tasks in WSN are designed to be energy efficient. Herein, the energy efficient estimation of a source vector observed by a WSN is considered. Several methods have been proposed to address the energy efficient estimation task. [1] studied the quantization of the WSN data measurements. A sensor selection method was proposed in [2] to minimize the estimation error subject to a given energy budget. Data reduction methods based on a linear compression of sensor nodes' data were proposed in [3] and [4] taking into account its impact on the estimation error. [5] analyzed the power allocation at each sensor node to minimize the MSE subject to power constraints, in a multiple antenna FC.

In fact, EE is regarded as a key pillar in the design of future wireless communications systems [6]. The rationale is that the capacity has to increase dramatically without increasing the current power consumption to avoid economical and societal concerns. Thereby, new system architectures have been recently proposed to increase the EE through a network densification, i.e. increasing the number of access nodes or the communication links per unit area [6], [7]. One of them is collocated massive MIMO (CM-MIMO), which considers a large number of antennas at the base station (BS) [8]. This system architecture has been recently considered in WSN to reduce the transmitted power in estimation problems. Namely, considering a scalar parameter and a coherent MAC channel, [9] and [10] studied the ML and LMMSE estimation at the FC, respectively. Within this setting [11] tackled the LMMSE estimation of a correlated source vector.

Herein, the LMMSE estimation of a correlated source vector is studied within the context of a DM-MIMO system. Namely, a WSN sends the noisy measurements of the source vector through a coherent MAC channel to the FC. The FC consists of a large number of single antenna APs, which on its turn are connected to a CPU through a backhaul network. Thereby, the CPU obtains an LMMSE estimation of the source vector. Moreover, power allocation is considered at each WSN node to minimize the total transmitted power subject to an estimation error constraint. Thereby unlike the related works in [9], [10] and [11] a DM-MIMO system is considered for decentralized estimation. This type of systems, a.k.a. large distributed antenna systems (L-DAS), have been considered recently to increase dramatically the EE in 5G cellular networks, see e.g. [7] or [12]. Our contributions are stated next. We obtain the MSE associated to the LMMSE estimation as well as an approximation of it, bearing in mind the DM-MIMO context. This permits to obtain the optimal power allocation that minimizes the total transmitted power for a given estimation error as an SDP problem, which can be efficiently solved. The numerical simulations show that compared to [11], which considers a CM-MIMO, the total transmitted power in the WSN is reduced for a given MSE constraint, both in the uniform and the optimal power allocation problems. Moreover, our work is more general than [9] and [10] as a vector source is considered in the estimation rather than a scalar.

The rest of the paper is organized as follows. Section II describes the system model. Section III obtains the MSE associated to the LMMSE and the optimal power allocation for a given estimation error in the DM-MIMO setting. Section IV presents the numerical results and section V the conclusions.

II. SYSTEM MODEL

The proposed system model is depicted in figure 1. Namely, we consider a WSN composed of L sensor nodes, which

take noisy observations of a source vector $\boldsymbol{\theta}$. The vector of measurements $\mathbf{x}$ is expressed as

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

where $\mathbf{n}$ is the sensing noise. Both $\boldsymbol{\theta}, \mathbf{n} \in \mathbb{C}^L$ are modeled as circularly symmetric complex normal random vectors, i.e. $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_\theta)$ and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C_n})$, with $\mathbf{C_n} = \text{diag}(\sigma_{n_1}^2, \ldots, \sigma_{n_L}^2)$ [1]. Power allocation is considered at each sensor, i.e. the l-th node multiplies its observation by a complex gain α_l and forwards it through a coherent MAC channel to the FC.

Regarding the FC, it is composed of a large number N of geographically distributed single antenna APs, which are supposed to serve simultaneously all the sensor nodes. The APs are connected to a CPU via a backhaul network. On its turn, the CPU performs the estimation of $\boldsymbol{\theta}$ given the received WSN measurements stacked in $\mathbf{y}$. Herein it is assumed that $N \gg L$, thereby this is a DM-MIMO system, a.k.a. L-DAS or cloud radio access network (C-RAN) if the CPU is implemented within the cloud computing paradigm. DM-MIMO has been recently considered to increase the EE in future wireless communications systems, see e.g. [7] or [12]. Herein, DM-MIMO is considered within the context of wireless sensing to assess its potential to reduce the transmitted power of the WSN for a given estimation error constraint. Regarding $\mathbf{y}$, it is modeled as follows,

$$\mathbf{y} = \mathbf{HDx} + \mathbf{v}. \tag{1}$$

Where $\mathbf{D} = \text{diag}(\alpha_1, \ldots, \alpha_L)$ is the matrix of gains to be designed, i.e. the one implementing the power allocation. $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I})$ is an additive white gaussian noise (AWGN) and $\mathbf{H} \in \mathbb{C}^{N \times L}$ is the channel between the sensor nodes and the APs. Note that the l-th column of $\mathbf{H}$ is the channel between the l-th sensor node and each of the APs and can be modeled as,

$$\mathbf{H} = (\mathbf{h}_1, \ldots, \mathbf{h}_L), \quad \mathbf{h}_l = \boldsymbol{\Gamma}_l^{1/2} \mathbf{f}_l. \tag{2}$$

Where $\boldsymbol{\Gamma}_l^{1/2}$ is a diagonal matrix modeling the large-scale fading, i.e. $\boldsymbol{\Gamma}_l \triangleq \text{diag}(\gamma_{1l}, \ldots, \gamma_{Nl})$, $\gamma_{nl} = d_{nl}^{-2\beta}$, d_{nl} is the distance between the l-th sensor node and the n-th AP and β is the pathloss exponent. The small-scale fading is modeled by $\mathbf{f}_l$ whose components are i.i.d random variables, namely $[\mathbf{f}_l]_i \sim \mathcal{CN}(0, 1) \ \forall i = 1, \ldots, N$.

Remark 1: Several assumptions are assumed in the proposed DM-MIMO system. Firstly, as in e.g. [12], a perfect backhaul network is presumed between the APs and the CPU. Secondly, perfect synchronization is assumed between the WSN nodes and the APs, to allow coherent reception at the FC, similarly than [11]. Thirdly, the FC, namely the CPU, has perfect knowledge of the channel and noise statistics.

Remark 2: The channel model in (2) shows that the proposed DM-MIMO generalizes the CM-MIMO system in [11]. Namely, assume that all the antennas are collocated at

---

a given AP, then the distance between the l-th sensor node and any of the antennas is approximately the same, i.e. the elements of $\boldsymbol{\Gamma}_l^{1/2}$ in (2) fulfill that $\gamma_{nl} = \gamma_l \ \forall n = 1, \ldots, N$. This leads to obtain the channel model for a CM-MIMO proposed in [11, eq.(3)].
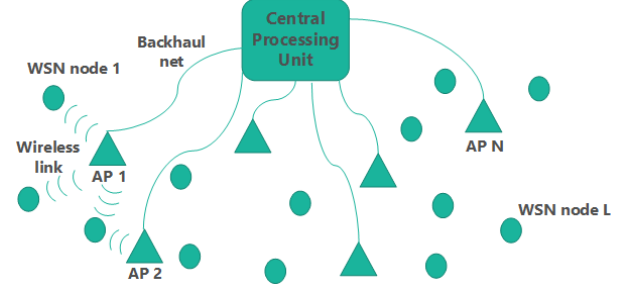

Figure 1. Proposed DM-MIMO system for decentralized estimation of a source.

III. OPTIMAL POWER ALLOCATION IN A DM-MIMO SYSTEM

A. LMMSE estimation and asymptotic MSE approximation

This section deals with the LMMSE estimation of the source of interest (SoI) $\boldsymbol{\theta}$ within the context of the system model proposed in section II. Namely, the data received at the CPU of the FC fulfills the model stated in (1) with the assumptions stated in *Remark 1*. Morevoer, recall that the LMMSE estimation of $\boldsymbol{\theta}$ is given by the conditional expectation $\hat{\boldsymbol{\theta}} = \mathbb{E}[\boldsymbol{\theta} \mid \mathbf{y}, \mathbf{H}]$, see [13, Ch.15]. Thereby, in a DM-MIMO setting $\hat{\boldsymbol{\theta}}$ takes the next expression,

$$\hat{\boldsymbol{\theta}} = \mathbb{E}[\boldsymbol{\theta} \mid \mathbf{y}, \mathbf{H}] = \left(\mathbf{C}_{\boldsymbol{\theta}}^{-1} + \mathbf{D}^H \mathbf{H}^H \mathbf{C_w}^{-1} \mathbf{HD} \right)^{-1}$$
$$\times \mathbf{D}^H \mathbf{H}^H \mathbf{C_w}^{-1} \mathbf{y}. \tag{3}$$

Moreover, after easy manipulations involving the matrix inversion lemma, one obtains that the MSE obtained by $\hat{\boldsymbol{\theta}}$ is given,

$$\text{MSE} = \text{Tr}\Big[\left(\mathbf{C}_{\boldsymbol{\theta}}^{-1} + \sigma_v^{-2} \mathbf{D}^H \mathbf{H}^H \mathbf{HD} - \sigma_v^{-4} \mathbf{D}^H \mathbf{H}^H \mathbf{HD} \right.$$
$$\times \left. \left(\mathbf{C_n}^{-1} + \sigma_v^{-2} \mathbf{D}^H \mathbf{H}^H \mathbf{HD} \right)^{-1} \mathbf{D}^H \mathbf{H}^H \mathbf{HD} \right)^{-1}\Big]. \tag{4}$$

Next, an asymptotic approximation of the MSE is derived. This has several benefits. Firstly, a more compact expression will be obtained. Secondly, the randomness of the channel due to the small-scale fading will vanish. This is interesting in terms of channel estimation, as the estimation must be updated less frequently. Namely, the large scale fading varies much more slowly than the small-scale fading. Thirdly, this MSE approximation will pave the way to formulate the design of the optimal power allocation within a convex optimization framework, see section III-B. To obtain the desired MSE approximation, let us introduce the next lemma,

Lemma 1. *Consider the asymptotic regime where $N \to \infty$ and L remains fixed. Then, $\mathbf{H}^H\mathbf{H}$, with $\mathbf{H}$ defined in (2), converges in probability to the next expression,*

$$\mathbf{H}^H\mathbf{H} \xrightarrow{p} \mathrm{diag}\left(\sum_{n=1}^{N}\gamma_{n1},\ldots,\sum_{n=1}^{N}\gamma_{nL}\right) \quad (5)$$

Proof: See Appendix.

Lemma 1 paves the way to obtain the desired approximation of the MSE. This is formally stated in the next theorem.

Theorem 1. *Let define the diagonal matrix $\mathbf{B}$ with diagonal elements $b_l = |\alpha_l|^2\sum_{n=1}^{N}\gamma_{nl}$ for $l = 1,\ldots,L$. Then, when $N \to \infty$ for a fixed L, the MSE in (4) converges in probability to $\overline{\mathrm{MSE}}$, whose expression is given next,*

$$\mathrm{MSE} \xrightarrow{p} \overline{\mathrm{MSE}} = \mathrm{Tr}\Big[\big(\mathbf{C}_\theta^{-1} + \sigma_v^{-2}\mathbf{B} - \sigma_v^{-4}\mathbf{B}$$
$$\times \left(\mathbf{C_n}^{-1} + \sigma_v^{-2}\mathbf{B}\right)^{-1}\mathbf{B}\big)^{-1}\Big]. \quad (6)$$

Proof: (6) is obtained by applying Lemma 1 into (4).

B. Optimal Power allocation

The energy budget is a scarce resource in a WSN and the proposed DM-MIMO system permits to reduce drastically the WSN power consumption by having a large number of virtual antennas at the FC. Thereby, this section deals with the problem of minimizing the total transmitted power P_t of the WSN nodes for a given performance constraint in the estimation task, within the DM-MIMO context proposed herein. This problem is expressed as,

$$\begin{array}{ll}\underset{\{b_l\geqslant 0\}_{l=1}^{L}}{\text{minimize}} & P_t(b_l) \\ \text{subject to} & \overline{\mathrm{MSE}}(b_l) \leqslant \delta.\end{array} \quad (7)$$

Where δ is the maximum estimation error permitted in the estimation problem at hand and the approximation of the MSE, i.e. $\overline{\mathrm{MSE}}(b_l)$ in (6), is considered. This will pave the way to obtain a convex optimization problem. Moreover, according to section II, $P_t(b_l)$ has the next expression,

$$P_t(b_l) = \mathbb{E}[\|\mathbf{D}(\boldsymbol{\theta}+\mathbf{n})\|^2] = \mathrm{Tr}[\mathbf{D}^H\mathbf{D}(\mathbf{C}_\theta + \mathbf{C_n})]. \quad (8)$$

Next, let us define $\tilde{\boldsymbol{\Gamma}}$ as

$$\tilde{\boldsymbol{\Gamma}} = \mathrm{diag}\left(\sum_{n=1}^{N}\gamma_{n1},\ldots,\sum_{n=1}^{N}\gamma_{nL}\right).$$

Given this definition of $\tilde{\boldsymbol{\Gamma}}$ an recalling the expression of b_l in Theorem 1 leads to express $P_t(b_l)$ as follows,

$$P_t(b_l) = \mathrm{Tr}\left[\tilde{\boldsymbol{\Gamma}}^{-1}\mathbf{B}(\mathbf{C}_\theta + \mathbf{C_n})\right]. \quad (9)$$

Thus, the optimization problem (7) can be restated as follows after substituting (9) and (6) into it,

$$\begin{array}{ll}\underset{\mathbf{B}\succeq 0}{\text{minimize}} & \mathrm{Tr}\left[\tilde{\boldsymbol{\Gamma}}^{-1}\mathbf{B}(\mathbf{C}_\theta + \mathbf{C_n})\right] \\ \text{subject to} & \mathrm{Tr}\Big[\big(\mathbf{C}_\theta^{-1} + \sigma_v^{-2}\mathbf{B} - \sigma_v^{-4}\mathbf{B} \\ & \times \left(\mathbf{C_n}^{-1} + \sigma_v^{-2}\mathbf{B}\right)^{-1}\mathbf{B}\big)^{-1}\Big] \leqslant \delta.\end{array} \quad (10)$$

Next, after introducing slack variables $\mathbf{X}, \mathbf{Z}$ and by considering the Schur complement of a block matrix, the optimization problem (10) can be cast as follows,

$$\begin{array}{ll}\underset{\mathbf{B},\mathbf{X},\mathbf{Z}\succeq 0}{\text{minimize}} & \mathrm{Tr}\left[\tilde{\boldsymbol{\Gamma}}^{-1}\mathbf{B}(\mathbf{C}_\theta + \mathbf{C_n})\right] \\ \text{subject to} & \mathrm{Tr}[\mathbf{Z}] \leqslant \delta \\ & \begin{pmatrix}\mathbf{C}_\theta^{-1} + \sigma_v^{-2}\mathbf{B} - \sigma_v^{-4}\mathbf{X} & \mathbf{I} \\ \mathbf{I} & \mathbf{Z}\end{pmatrix} \succeq 0 \\ & \begin{pmatrix}\mathbf{X} & \mathbf{B} \\ \mathbf{B} & \mathbf{C_n}^{-1} + \sigma_v^{-2}\mathbf{B}\end{pmatrix} \succeq 0.\end{array} \quad (11)$$

This is an SDP problem and thereby can be solved efficiently, i.e. in polynomial time, by software packages relying on interior point methods.

Remark 3: The comparison of the proposed algorithm in (11) to a uniform power allocation method is of paramount importance. This analysis is presented in section IV. In the latter case, $\mathbf{D} = \alpha\mathbf{I}$ and thereby $\mathbf{B} = |\alpha|^2\tilde{\boldsymbol{\Gamma}}$ and $P_t = \mathrm{Tr}[|\alpha|^2(\mathbf{C}_\theta + \mathbf{C_n})]$ in (11).

IV. NUMERICAL RESULTS

Next, we assess the power allocation method proposed in (11), which seeks to minimize the total transmitted power of a WSN in a DM-MIMO context, for a given estimation error constraint. The next simulation parameters are considered for the experiments. An exponential covariance matrix model is considered for the source covariance, i.e. $[\mathbf{C}_\theta]_{ij} = \rho_\theta^{|i-j|}$ with $0 \leqslant \rho_\theta \leqslant 1$. The sensing and AWGN noise variances, i.e. $\sigma_v^2, \sigma_{n_l}^2$, are specified below. The number of WSN nodes is $L = 15$. In the coefficients modeling the large-scale fading, i.e. $\gamma_{nl} = d_{nl}^{-2\beta}$, the pathloss exponent $\beta = 1$. The WSN nodes and the APs, of the DM-MIMO setting, are within a square area with side length $a = 160$, and their coordinates are generated randomly according to a uniform distribution between 0 and a at each iteration of the simulation, which leads to obtain d_{nl}. For the CM-MIMO setting, the same scenario is considered but the antennas of the FC are fixed at the center of the square area. This leads to obtain $\gamma_l = d_l^{-2\beta}$, see *Remark 2* above for further clarifications. Finally δ in (11) is set to the middle of the interval $(\mathrm{Tr}[(\mathbf{C}_\theta^{-1} + \mathbf{C_n}^{-1})^{-1}], \mathrm{Tr}[\mathbf{C}_\theta])$, which defines its possible values according to [11].

In Figure 2 the asymptotic approximation of the MSE proposed in (6) is studied. Namely, we consider a WSN with a total transmitted power of $P_t = 10^{-2}$ W, $\rho_\theta = 0.95$, $\sigma_{n_l}^2 = 10^{-3}$ W $\forall l$ and $\sigma_v^2 = 10^{-6}$ W. Moreover, a uniform power allocation (UPA) is implemented, i.e. $\mathbf{D} = \alpha\mathbf{I}$. Thereby, according to (8), $|\alpha|^2 = \frac{P_t}{L([\mathbf{C}_\theta+\mathbf{C_n}]_{ll})}\forall l$. Moreover, a Monte Carlo simulation with 100 iterations is considered. At each iteration the coordinates of the WSN nodes and APs are

generated randomly as specified above, this produces a given **H**. Given these simulation conditions, we plot the exact MSE in (4) and its approximation in (6) achieved by the LMMSE in a DM-MIMO setting. For comparison purposes, we plot the exact and approximated MSE achieved by the LMMSE in a CM-MIMO context as well. These MSE expressions were obtained in [11] and are a particular case of (4) and (6) due to the *Remark 2* explained above. Figure 2 highlights that $\overline{MSE}$ tends to MSE when N increases for a fixed L as predicted in Theorem 1 and for $N > 100$ on the approximation is tight. It can be observed, that for a UPA setting, the performance of the proposed LMMSE in a DM-MIMO outperforms the one of the LMMSE in a CM-MIMO, which was proposed in [11]. The rationale is that due to its distributed nature, a DM-MIMO setting permits to shorten the distance between the WSN nodes and the virtual antennas of the FC compared to a CM-MIMO scenario. Also there is more diversity in the links as it reflected in the corresponding channel models, see (2) and *Remark 2*.
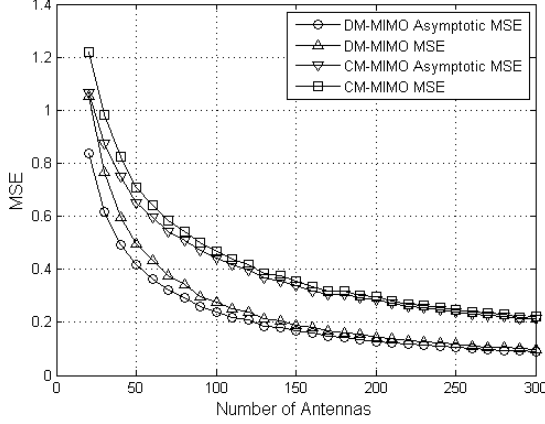

Figure 2. Comparison MSE in (4) and asymptotic MSE in (6) of a DM-MIMO and a CM-MIMO.

In Figures 3-5 we compare the Optimal Power Allocation (OPA) and UPA in a DM-MIMO scenario (see (11) and *Remark 3*) to the OPA and UPA in a CM-MIMO setting [11, eq.(10)]. Figure 3 considers that the SoI is an uncorrelated source vector, $\sigma_v^2 = 10^{-3}$ W and $\sigma_{n_l}^2 = 10^{-6}$ W $\forall l$. It can be observed that the total transmitted power of the WSN, for a given estimation error constraint, can be significantly reduced in the proposed DM-MIMO setting compared to a CM-MIMO setting. The rationale is analogous to the one of figure 2. Moreover, it can be observed that both the UPA and OPA obtain similar performance. The reason is that the source is uncorrelated, the sensing noise variance is the same among WSN nodes and the AWGN variance is the same at the APs. Thereby, in this case the optimal power allocation is a uniform one, because there is not redundant information measured by the sensors and the quality of the information arriving at the FC from different WSN nodes is the same.

In figure 4 the same type of simulation than figure 3 is carried out. Herein, $\sigma_v^2 = 10^{-3}$ W and the main difference is that the sensing noise variance is varied in the WSN
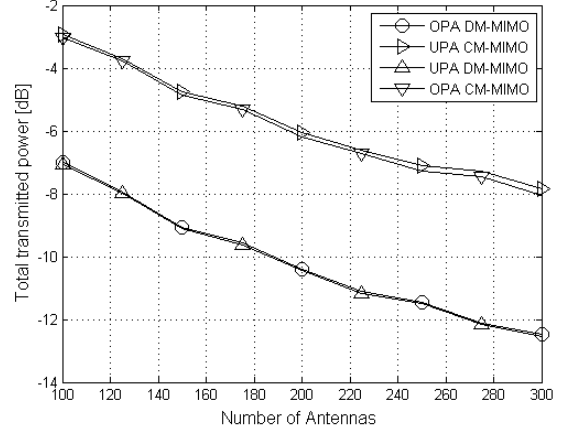

Figure 3. Comparison for uncorrelated source vector (i.e. $\rho_\theta = 0$).

nodes. Namely, it is generated randomly according to the next uniform distribution $\sigma_{n_l}^2 \sim \mathcal{U}(0.5 \cdot 10^{-3}, 0.9 \cdot 10^{-3})$. This variable sensing noise leads the OPA to obtain better performance than the UPA both for the DM-MIMO and CM-MIMO scenarios, as a higher weight is assigned to the nodes with better sensing information. Moreover, in this scenario the OPA DM-MIMO and UPA DM-MIMO permit to reduce significantly the total tranmstted power compared to the CM-MIMO case due to the same rationale than the previous figures.
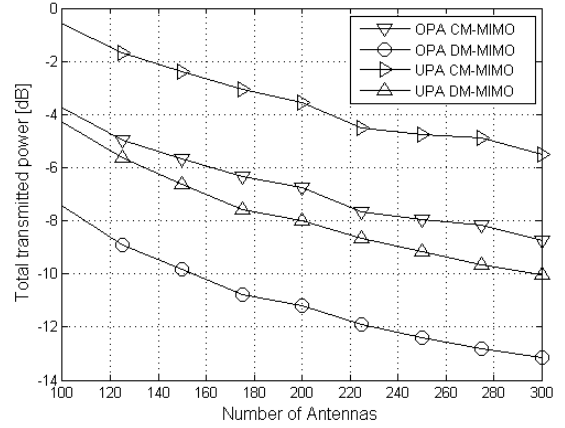

Figure 4. Comparison for uncorrelated source vector (i.e. $\rho_\theta = 0$) and variable sensing noise variance ($\sigma_{n_l}^2 \sim \mathcal{U}(0.5 \cdot 10^{-3}, 0.9 \cdot 10^{-3})$).

Finally, in figure 5 a similar simulation than figure 3 is presented. The difference is that now the source vector is correlated with correlation factor $\rho_\theta = 0.95$. It can be observed that the OPA methods obtain better performance than the UPA ones. The rationale is that in this case there is redundant information among sensors. Thereby, one can discard the information or lower the weight at some of the WSN nodes to reduce significantly the total transmitted power without affecting significantly the resulting MSE at the FC. As in the previous figures, the transmitted power in the proposed DM-MIMO setting is reduced notably compared to the CM-MIMO scenario. Again this is because in a DM-MIMO setting the

virtual antennas lower the distance to the WSN nodes thanks to their distributed nature compared to CM-MIMO where all the antennas are collocated at the FC.
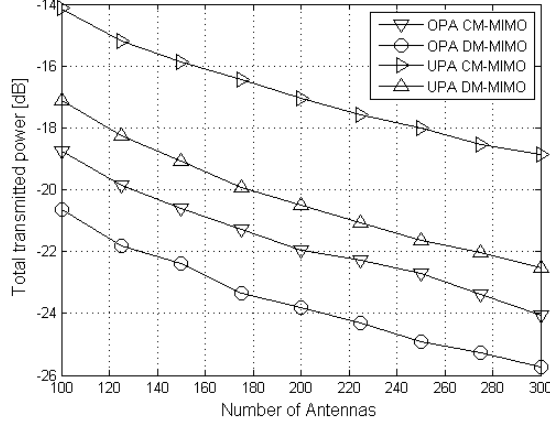

Figure 5. Comparison for correlated source vector with $\rho_\theta = 0.95$.

V. Conclusions

This paper has considered a distributed massive MIMO scenario to reduce significantly the total transmitted power of a WSN in the estimation of a correlated source vector. Namely, a WSN takes noisy measurements of the SoI and transmit them via a coherent MAC channel to a FC. This consists of a large number of distributed APs, or virtual antennas, which are connected to a CPU via a backhaul network. We have presented the exact MSE associated to an LMMSE estimation at the CPU of the FC. Then, leveraging the massive MIMO setting, an approximated MSE has been obtained, which converges to the exact MSE when L remains fixed and N grows without bound. This approximated MSE has paved the way to design an optimal power allocation method to minimize the total transmitted power of the WSN subject to an estimation error constraint at the FC. The obtained method can be cast as an SDP problem. The numerical results show that the proposed method leads to reduce notably the total transmitted power for a given MSE constraint compared to the related work where all the antennas are collocated at the FC, i.e. a collocated massive MIMO setting.

Appendix: Proof of Lemma 1

First, according to (2), we obtain the next equality,

$$\left[\mathbf{H}^H\mathbf{H}\right]_{ll} = \mathbf{f}_l^H \mathbf{\Gamma}_l \mathbf{f}_l.$$

Recalling the expression of $\mathbf{\Gamma}_l$ in (2) and defining $\mathbf{f}_l = (f_{1l}, \ldots, f_{Nl})^T$, $\left[\mathbf{H}^H\mathbf{H}\right]_{ll}$ ca be rewritten as

$$\left[\mathbf{H}^H\mathbf{H}\right]_{ll} = \gamma_{1l}|f_{1l}|^2 + \ldots + \gamma_{Nl}|f_{Nl}|^2. \tag{12}$$

Now, for ease of notation, we introduce S_{ll}, x_n and μ_n, which have the next definitions,

$$S_{ll} = \left[\mathbf{H}^H\mathbf{H}\right]_{ll}, \quad S_{ll} = \sum_{n=1}^{N} x_n, \quad \mu_n = \mathbb{E}[x_n].$$

Thus, assuming that $\mathrm{var}(x_n) \leqslant c < \infty$ and x_n are independent $\forall n$, by the Chebyshev's inequality the next relation holds for any real number $\epsilon > 0$,

$$\mathrm{Pr}[|N^{-1}S_{ll} - N^{-1}\sum_{n=1}^{N}\mu_n| \geqslant \epsilon] \leqslant \frac{\sum_{n=1}^{N}\mathrm{var}(x_n)}{N^2\epsilon^2} \leqslant \frac{c}{N\epsilon^2}.$$

Note that $\mu_n = \mathbb{E}[\gamma_{nl}|f_{nl}|^2] = \gamma_{nl}$, as $f_{nl} \sim \mathcal{CN}(0,1)$. Thereby, taking $N \to \infty$ in both sides of the last equation, the next convergence in probability is obtained,

$$\lim_{N\to\infty} \mathrm{Pr}[|N^{-1}S_{ll} - N^{-1}\sum_{n=1}^{N}\gamma_{nl}| \geqslant \epsilon] = 0 \tag{13}$$

The same procedure can be applied to show that the off diagonal elements of $\mathbf{H}^H\mathbf{H}$ converge to 0. To this end, note that $\left[\mathbf{H}^H\mathbf{H}\right]_{lk} = \mathbf{f}_l^H \mathbf{\Gamma}_l^{1/2} \mathbf{\Gamma}_k^{1/2} \mathbf{f}_k$ and recall that $\mathbf{f}_l, \mathbf{f}_k$ are independent with zero mean elements.

References

[1] Z. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Info. Th.*, vol. 51, pp. 2210–2219, June 2005.

[2] S. Liu, S. Chepuri, M. Fardad, E. Masazade, G. Leus, and P. Varshney, "Sensor selection for estimation with correlated measurement noise," *IEEE Trans. Signal Proc.*, vol. 64, pp. 3509–3522, July 2016.

[3] I. Schizas, G. Giannakis, and Z. Luo, "Distributed estimation using reduced-dimensionality sensor observations," *IEEE Trans. Signal Proc.*, vol. 55, pp. 4284–4299, August 2007.

[4] A. Shirazinia and S. Dey, "Power-constrained sparse gaussian linear dimensionality reduction over noisy channels," *IEEE Trans. Signal Proc.*, vol. 63, pp. 5837–5852, Nov. 2015.

[5] A. Behbahani, A. Eltawil, and H. Jafarkhani, "Decentralized estimation under correlated noise," *IEEE Trans. Signal Proc.*, vol. 62, pp. 5603–5614, Nov. 2014.

[6] E. Bjrnson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: smmall cells meet massive MIMO," *IEEE journal of selected areas in communications*, vol. 34, pp. 832–847, April 2016.

[7] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: a survey," *IEEE Communications surveys & tutorials*, vol. 18, pp. 2522–2545, Fourth quarter 2016.

[8] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Comm.*, vol. 9, pp. 3590–3600, Nov. 2010.

[9] F. Jiang, J. Chen, and A. Swindlehurst, "Phase-only analog encoding for a multi-antenna fusion center," in *Proc. IEEE ICASSP*, pp. 2645–2648, Mar. 2012.

[10] F. Jiang, J. Chen, A. Swindlehurst, and J. Lopez-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *IEEE Trans. Signal Proc.*, vol. 63, pp. 3005–3017, Jun. 2015.

[11] A. Shirazinia, S. Dey, D. Ciuonzo, and P. S. Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Trans. Signal Proc.*, vol. 64, pp. 2499–2512, May 2016.

[12] H. Q. Ngo, A. Ashikhmin, H. Yang, E. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Trans. Wirel. Comm.*, vol. 16, pp. 1834–1850, March 2017.

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