Joint Transceiver Designs for MSE Minimization in MIMO Wireless Powered Sensor Networks

Naveen K. D. Venkategowda, *Member, IEEE*, Hoon Lee, *Member, IEEE*, and Inkyu Lee, *Fellow, IEEE*

***Abstract*— In this paper, we study vector parameter estimation in multiple-input multiple-output wireless-powered sensor networks (WPSNs) where sensor nodes operate by harvesting the radio frequency signals transmitted from energy access points (E-APs). We investigate a joint design of sensor data precoders, a fusion rule, and energy covariance matrices to minimize the mean square error (MSE) of the parameter estimate based on a non-linear energy harvesting model. First, we propose a centralized algorithm to solve the MSE minimization problem. Next, to reduce the computational complexity at the fusion center (FC) and feedback overhead from the sensors to the FC, we present a distributed algorithm to locally compute the precoders and the energy covariance matrices. We employ the alternating direction method of multipliers technique to minimize the MSE in a distributed manner without any coordination from the FC. In the proposed distributed algorithm, each sensor node calculates its own precoders and determines the local information of the fusion rule, and then messages are broadcast to other sensor nodes and E-APs. Simulation results demonstrate that the distributed algorithm performs close to the centralized algorithm with reduced complexity. Moreover, the proposed methods exhibit superior estimation performance over conventional techniques in WPSNs.**

***Index Terms*— Distributed estimation, energy harvesting, precoding, wireless power transfer, wireless sensor network.**

I. Introduction

WIRELESS sensor networks (WSN) consisting of distributed nodes are deployed in an environment for sensing, collecting information, and actuation. They are found in many applications such as surveillance, environmental monitoring, and automation [1]. An important use of the WSNs is to estimate a parameter of interest at a fusion center (FC) using the data received from multiple sensor nodes, which observe the source.

To enhance the estimation accuracy, the observations at the sensor nodes can be processed before transmitting them to the FC over the multiple access channel (MAC). Thus, the data transmission and the fusion of the data received from multiple sensors, which is termed as the fusion rule, can be intelligently designed to minimize the estimation error. This problem has been well researched in [2]–[5] and references therein. However, one of the major shortcomings of the WSNs is that the network lifetime is often short, since replacing batteries for the sensor nodes is expensive and the nodes may be inaccessible [6].

With the advent of wireless energy transfer (WET) technologies [7], a dedicated energy access point (E-AP) can remotely recharge the sensor nodes by transmitting radio frequency (RF) signals and prolong the lifetime of the WSN. The RF based WET is ideally suited for WSNs, since it can concurrently power a large number of sensors and also transmit information. In contrast, magnetic resonance and inductive coupling based WET methods suffer from short charging distances and a large form factor [8]. Further, the energy efficiency of the RF based WET can be enhanced by employing multiple-input multiple-output (MIMO) techniques to control the direction of the energy beams [9].

In wireless powered sensor networks (WPSN), the sensor nodes first harvest the energy of the RF signals from E-APs, and then utilize the collected energy for transmitting the data to the FC for parameter estimation. However, a rapid attenuation of RF waves with respect to the transmission distance reduces the energy transfer efficiency. Further, due to uneven deployment, energy harvested by the sensors vary drastically among the sensors and may be insufficient to sustain the operation at many sensors. Therefore, for the WPSN, in addition to the sensor processing and the fusion rule, an energy transmission strategy plays an important role in maximizing the network performance. Hence, designing the energy transmitting and harvesting policy for WSNs has attracted significant research interests recently [10]–[12].

Distributed energy beamforming with multiple chargers to maximize the transferred power was proposed in [13] and fairness-based energy beamforming to maximize the minimum energy harvested in the network was investigated in [14]. A joint design of energy beamformers and charging time was introduced to maximize the minimum throughput of users in [15] and the sum-rate in [16]. Developing on these results, the work in [12] studied the probability that the received power at the sensors exceeds a threshold. In [17], the authors presented the optimal power allocation to maximize the sum-throughput in a cooperative WSN where relays harvest energy by utilizing the signal from the sensor nodes. Similarly, [18] investigated energy beamforming and power allocation for

Manuscript received February 13, 2018; accepted May 7, 2018. Date of publication May 28, 2018; date of current version August 10, 2018. This work was supported by the National Research Foundation through the Ministry of Science, ICT, and Future Planning through the Korean Government under Grant 2017R1A2B3012316. The associate editor coordinating the review of this paper and approving it for publication was R. Brown. *(Corresponding author: Inkyu Lee.)*

N. K. D. Venkategowda and I. Lee are with the School of Electrical Engineering, Korea University, Seoul 02841, South Korea (e-mail: naveendv@korea.ac.kr; inkyu@korea.ac.kr).

H. Lee is with the Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore 487372 (e-mail: hoon_lee@sutd.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2018.2838589

the sensor nodes to maximize the received signal-to-noise ratio (SNR) at the FC. However, these works do not study a joint design of the WPSN for parameter estimation.

Power allocation between data acquisition and data transmission was examined in [19] for a WPSN that employs zero-forcing precoding for estimation. The authors in [20] proposed the optimal power transmission from the chargers and the optimal allocation of the harvested power for information transmission and local sensing at the sensor. Power allocation and energy beamforming were determined in [21] to minimize the mean square error (MSE) of the estimate at a FC for single antenna sensors. In [22], the optimal energy management when sensor nodes share energy was considered for the sub-optimal best linear unbiased estimator at the FC. Also, power allocation was introduced in [23] to minimize the MSE assuming chargers with a finite number of fixed energy-beam patterns. All these works were limited to scalar parameter estimation in a single-input single-output (SISO) WPSN and did not explicitly derive the energy transmit beamformers.

Moreover, all the existing works assume that a central coordinating entity such as a FC computes the precoders and the energy-beam directions, and forwards them to the sensor nodes and the E-APs. Also, they adopted a simple linear energy harvesting (EH) receiver model for power conversion. It has been shown that in practice, the EH circuits exhibit non-linear characteristics and a system design based on the conventional linear EH model may lead to inefficient utilization of the energy resources [24], [25].

In this paper, we consider vector parameter estimation in a general MIMO multi-sensor WPSN with a non-linear energy harvesting model. We investigate joint optimization of sensor data precoders, energy covariance matrices, and the fusion rule to minimize the MSE of the parameter estimate subject to power constraint at the E-APs. First, we propose a centralized algorithm in which the FC solves the MSE minimization problem. Since the optimal minimum mean square error (MMSE) estimation of vector parameters leads to a non-convex optimization problem, we employ an iterative alternating minimization algorithm to determine the optimal precoders, the energy covariance matrices, and the fusion rule.

However, the FC requires the knowledge of channel state information (CSI) among all the nodes. In addition, the computational complexity of the centralized algorithm scales with the number of sensors. To overcome these problems, we present a distributed algorithm based on the alternating direction method of multipliers (ADMM) technique to obtain the optimal precoders at each sensor node and the energy covariance matrices at the E-APs.

The distributed algorithm only requires the sensor nodes to broadcast local messages to other sensor nodes and the dual price updates to the E-APs instead of sharing the actual CSI and the observation statistical information to the FC. The proposed algorithm has low computational complexity independent of the number of sensors. Besides, we derive closed form expressions to calculate a solution.

Moreover, each sensor node computes the optimal MMSE fusion rule locally that can be fed back to the FC for parameter estimation. This relieves the FC from acquiring

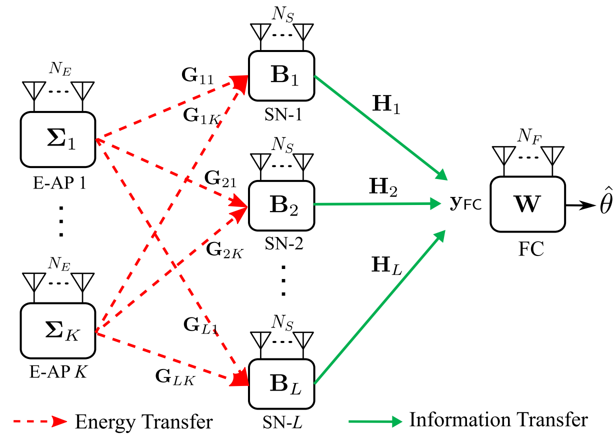

Fig. 1. System model of a WPSN.

CSI and all the local observation statistical information and evaluating the optimal fusion rule. This significantly decreases the feedback overhead from sensor nodes to the FC. Simulation results demonstrate that the proposed methods exhibit superior estimation performance over conventional techniques for parameter estimation in WPSNs. Also, we show that the proposed distributed algorithm performs close to the centralized technique while requiring low computational resources.

This paper is organized as follows: Section II describes the system model and formulates the parameter estimation problem in WPSN. In Section III, we present the centralized algorithm to optimize a WPSN. Section IV proposes the distributed algorithm to solve the MSE minimization problem. Simulation results are illustrated in Section V and the conclusions are summarized in Section VI.

The following notations are used throughout the paper. The operators $(\cdot)^T$, $(\cdot)^H$, and $\mathrm{tr}(\cdot)$ denote transpose, Hermitian, and trace of a matrix, respectively. $\|\cdot\|$ and $\|\cdot\|_F$ indicate the L_2 norm and the Frobenius norm, respectively. $\mathfrak{Re}(\cdot)$ represents the real part of a complex number. $\mathbf{I}_n$ defines an identity matrix of dimension n and $\otimes$ stands for the Kronecker product. The operator $\mathsf{vec}(\mathbf{A})$ forms a column vector from a matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$ by stacking the column vectors $\mathbf{a}_i$ and $\mathsf{ivec}(\mathbf{a}, m)$ rearranges the vector $\mathbf{a} \in \mathbb{C}^{mn\times 1}$ into a matrix with m rows and n columns.

II. System Model

We consider a MIMO WPSN comprising a FC with N_F antennas and L sensor nodes with N_S antennas as shown in Fig. 1. This network operates by harvesting energy from the RF signals emitted by K E-APs with N_E antennas each. The harvested energy is then utilized to transmit the sensor observations to the FC. During the WET phase, the jth E-AP broadcasts the energy signal $\mathbf{s}_j \in \mathbb{C}^{N_E\times 1}$ with covariance matrix $\mathbf{\Sigma}_j \in \mathbb{C}^{N_E\times N_E}$. The transmit power at the jth E-AP is limited to $\mathrm{tr}(\mathbf{\Sigma}_j) \le P_{T,j}$ for $j = 1, 2, \ldots, K$.

Let us denote the channel matrix between the ith sensor node and the jth E-AP by $\mathbf{G}_{ij} \in \mathbb{C}^{N_S\times N_E}$. The total received RF power at the ith sensor node is given by

$$P_{R,i} = \mathbb{E}\left[\left\|\sum_{j=1}^{K}\mathbf{G}_{ij}\mathbf{s}_j\right\|_F^2\right] = \sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{G}_{ij}\mathbf{\Sigma}_j\mathbf{G}_{ij}^H\right). \quad (1)$$

For simple analysis, the total harvested power at the ith sensor node is often modeled as $P_{H,i}^{\text{lin}} = \eta_i P_{R,i}$, where η_i is the energy harvesting efficiency. However, experimental results have shown that the EH circuits follow non-linear characteristics in practice [26]. Therefore, we adopt a non-linear parametric model for energy harvesting in the WPSN [24]. Adopting the model presented in [24], the harvested power at the ith node is expressed as

$$P_{H,i} = \frac{M_i}{1-\Omega_i}\left(\frac{1}{1+\exp\left[-a_i(P_{R,i}-b_i)\right]} - \Omega_i\right), \quad (2)$$

where Ω_i is defined as $\Omega_i = 1/(1+\exp(a_i b_i))$, a_i and b_i are obtained using a curve fitting algorithm from measurement results of the EH circuit, and M_i denotes the maximum harvested power by the sensor when the EH circuit is saturated.

The sensor nodes measure the parameter of interest $\boldsymbol{\theta} \in \mathbb{C}^{M_\theta \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\theta)$ with covariance matrix $\mathbf{R}_\theta$. The observations $\mathbf{m}_i \in \mathbb{C}^{M\times 1}$ at the ith node can be written using the linear observation model as

$$\mathbf{m}_i = \mathbf{A}_i \boldsymbol{\theta} + \mathbf{n}_i, \quad (3)$$

where $\mathbf{A}_i \in \mathbb{C}^{M\times M_\theta}$ represents the observation matrix and $\mathbf{n}_i \in \mathbb{C}^{M\times 1}$ indicates the zero mean Gaussian observation noise uncorrelated at different sensors with covariance matrix $\mathbb{E}[\mathbf{n}_i\mathbf{n}_i^H]=\mathbf{R}_i$.

We assume that the sensors transmit linearly precoded observations to the FC over a MAC [2]. Defining $\mathbf{B}_i \in \mathbb{C}^{N_S\times M}$ as the precoding matrix at sensor node i, the received signal $\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L) \in \mathbb{C}^{N_F\times 1}$ at the FC can be expressed as

$$\begin{aligned}\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L) &= \sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{m}_i + \mathbf{n}_{\mathsf{FC}} \\ &= \sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\boldsymbol{\theta} + \sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{n}_i + \mathbf{n}_{\mathsf{FC}}, \quad (4)\end{aligned}$$

where $\mathbf{H}_i \in \mathbb{C}^{N_F\times N_S}$ is the channel matrix between the ith node and the FC and $\mathbf{n}_{\mathsf{FC}} \in \mathbb{C}^{N_F\times 1}$ equals the noise at the FC distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathsf{FC}})$.

The energy available at the ith sensor for information transmission is limited by

$$\begin{aligned}&\tau P_{C,i} + \tau_I \mathbb{E}\left[\|\mathbf{B}_i\mathbf{m}_i\|^2\right] \\ &= \tau P_{C,i} + \tau_I \text{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H + \mathbf{R}_i)\mathbf{B}_i^H) \le \tau_E \alpha_i P_{H,i}, \quad (5)\end{aligned}$$

where $P_{C,i}$ refers to the circuit power consumption for sensor operations, α_i indicates the fraction of the harvested energy available for data transmission, τ_I and τ_E represent the time duration for information transmission and energy harvesting, respectively, and $\tau = \tau_I + \tau_E$.

Then, the FC estimates the parameter $\boldsymbol{\theta}$ employing the received data $\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)$ and the fusion rule $\mathbf{W} \in \mathbb{C}^{N_F\times M_\theta}$ as $\hat{\boldsymbol{\theta}} = \mathbf{W}^H\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)$. In a WPSN for the parameter estimation, the main objective is to determine the precoders $\{\mathbf{B}_i\}_{i=1}^L$, the fusion rule $\mathbf{W}$, and the energy covariance matrices $\{\boldsymbol{\Sigma}_j\}_{j=1}^K$ that minimize the MSE of the estimate. This problem can be modeled as

$$\begin{aligned}\min_{\{\mathbf{B}_i\},\{\boldsymbol{\Sigma}_j\},\mathbf{W}} \quad & \mathbb{E}[\|\boldsymbol{\theta} - \mathbf{W}^H\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)\|^2] \\ \text{s. t.} \quad & \tau P_{C,i} + \tau_I \text{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H + \mathbf{R}_i)\mathbf{B}_i^H) \\ & \le \tau_E \alpha_i P_{H,i}(\{\boldsymbol{\Sigma}_j\}),\ i=1,\ldots,L, \\ & \text{tr}(\boldsymbol{\Sigma}_j) \le P_{T,j},\ \text{and}\ \boldsymbol{\Sigma}_j \succeq \mathbf{0},\ j=1,\ldots,K. \quad (6)\end{aligned}$$

In the subsequent sections, we derive centralized and distributed algorithms to solve the above problem.

III. Centralized Algorithm for MSE Minimization

Substituting $\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)$ from (4), the MSE $\mathbb{E}[\|\boldsymbol{\theta} - \mathbf{W}^H\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)\|^2]$ can be written as

$$\begin{aligned}&\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i\}_{i=1}^L, \mathbf{W}) \\ &= \text{tr}\bigg(\mathbf{W}^H\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)^H\mathbf{W} \\ &\quad -\mathbf{W}^H\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta - \mathbf{R}_\theta\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)^H\mathbf{W} \\ &\quad +\sum_{i=1}^L \mathbf{W}^H\mathbf{H}_i\mathbf{B}_i\mathbf{R}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{W} + \mathbf{W}^H\mathbf{R}_{\mathsf{FC}}\mathbf{W} + \mathbf{R}_\theta\bigg). \quad (7)\end{aligned}$$

As $\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i\}_{i=1}^L, \mathbf{W})$ is non-convex in terms of $\{\mathbf{B}_i\}$ and $\mathbf{W}$, the problem in (6) is non-convex and difficult to solve. However, it can be divided into two convex sub-problems where $\mathbf{W}$ and $\{\{\mathbf{B}_i\}_{i=1}^L, \{\boldsymbol{\Sigma}_j\}_{j=1}^K\}$ are alternately optimized. For a given $\{\mathbf{B}_i\}_{i=1}^L$, finding the derivative of the MSE in (7) with respect to $\mathbf{W}$ and equating it to zero yields the optimal fusion rule as

$$\begin{aligned}\mathbf{W}^\star &= \arg\min_{\mathbf{W}} \mathcal{E}_{\mathsf{C}}(\mathbf{W}|\{\mathbf{B}_i\}) \\ &= \bigg(\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)^H + \mathbf{R}_{\mathsf{T}}\bigg)^{-1} \\ &\quad \times\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta, \quad (8)\end{aligned}$$

where $\mathbf{R}_{\mathsf{T}} = \sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{R}_i\mathbf{B}_i^H\mathbf{H}_i^H + \mathbf{R}_{\mathsf{FC}}$ is the total noise covariance matrix.

Next, defining $\mathbf{b}_i \triangleq \mathsf{vec}(\mathbf{B}_i) \in \mathbb{C}^{N_S M\times 1}$ and employing the relations $\text{tr}\left(\mathbf{A}\mathbf{A}^H\right) = \mathsf{vec}(\mathbf{A})^H\mathsf{vec}(\mathbf{A}) = \|\mathsf{vec}(\mathbf{A})\|^2$ and $\mathsf{vec}(\mathbf{AXB}) = \left(\mathbf{B}^T\otimes\mathbf{A}\right)\mathsf{vec}(\mathbf{X})$ [27], the first term on the right hand side of (7) can be expressed as

$$\begin{aligned}&\text{tr}\bigg(\mathbf{W}^H\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)^H\mathbf{W}\bigg) \\ &\qquad = \left\|\sum_{i=1}^L \Big((\mathbf{R}_\theta^{\frac{1}{2}})^T\mathbf{A}_i^T \otimes \mathbf{W}^H\mathbf{H}_i\Big)\mathbf{b}_i\right\|^2.\end{aligned}$$

Further, applying the matrix identities $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB})$ and $\text{tr}\left(\mathbf{A}^H\mathbf{BCD}^H\right) = \mathsf{vec}(\mathbf{A})^H((\mathbf{D}^H)^T\otimes\mathbf{B})\mathsf{vec}(\mathbf{C})$ [27]

on the second and third terms in (7), the MSE for a given $\mathbf{W}$ can be rewritten as

$$\begin{aligned}&\mathcal{E}_{\mathsf{C}}(\{\mathbf{b}_i\}_{i=1}^L|\mathbf{W})\\&=\left\|\sum_{i=1}^L\left((\mathbf{R}_\theta^{\frac{1}{2}})^T\mathbf{A}_i^T\otimes\mathbf{W}^H\mathbf{H}_i\right)\mathbf{b}_i\right\|^2\\&\quad+\sum_{i=1}^L\mathbf{b}_i^H(\mathbf{R}_i^T\otimes\mathbf{H}_i^H\mathbf{W}\mathbf{W}^H\mathbf{H}_i)\mathbf{b}_i\\&\quad-2\mathfrak{Re}\big(\mathsf{vec}((\mathbf{A}_i\mathbf{R}_\theta\mathbf{W}^H\mathbf{H}_i)^H)^H\mathbf{b}_i\big)+\mathrm{tr}(\mathbf{W}^H\mathbf{R}_{\mathsf{FC}}\mathbf{W}+\mathbf{R}_\theta).\end{aligned}\tag{9}$$

Similarly, applying the vectorization operator on (5), the optimization problem to compute $\{\mathbf{b}_i\}$ and $\{\mathbf{\Sigma}_j\}$ for a given $\mathbf{W}$ can be formulated as

$$\begin{aligned}\min_{\substack{\{\mathbf{b}_i\},\{\theta_i\}\\\{\mathbf{\Sigma}_j\}}}\ &\mathcal{E}_{\mathsf{C}}(\{\mathbf{b}_i\}_{i=1}^L|\mathbf{W})\\\text{s. t. }&\tau P_{C,i}+\tau_I\mathbf{b}_i^H\mathbf{Q}_i\mathbf{b}_i\\&\quad\le\frac{\tau_E\alpha_iM_i}{1-\Omega_i}\Big(\frac{1}{1+e^{-a_i(\theta_i-b_i)}}-\Omega_i\Big),i=1,\ldots,L,\\&\theta_i\le\sum\nolimits_{j=1}^K\mathrm{tr}(\mathbf{G}_{ij}\mathbf{\Sigma}_j\mathbf{G}_{ij}^H),\ i=1,\ldots,L,\\&\mathrm{tr}(\mathbf{\Sigma}_j)\le P_{T,j},\ \text{and }\mathbf{\Sigma}_j\succeq\mathbf{0},\ j=1,\ldots,K,\end{aligned}\tag{10}$$

where $\mathbf{Q}_i\triangleq(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H+\mathbf{R}_i)^T\otimes\mathbf{I}_{N_S}$.

It can be easily verified that (10) is a convex problem [25]. However, it does not have a closed form solution and must be solved by interior point algorithms [28]. Now, $\mathbf{W},\{\mathbf{b}_i\}$, and $\{\mathbf{\Sigma}_j\}$ can be obtained by an iterative alternating minimization algorithm in Algorithm 1, which is shown above. The algorithm can be terminated when the MSE $\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^{(n)}\}_{i=1}^L,\mathbf{W}^{(n)})$ converges, where n is the iteration index. The individual precoders are calculated as $\mathbf{B}_i=\mathsf{ivec}(\mathbf{b}_i,N_S)$.

Algorithm 1 Centralized Algorithm

Initialize $n=0$ and $\mathbf{B}_i^{(0)}$ for $i=1,\ldots,L$.
repeat
  Compute $\mathbf{W}^{(n+1)}$ from (8) with $\mathbf{B}_i=\mathbf{B}_i^{(n)}$, $\forall i$.
  Calculate $\{\mathbf{b}_i^{(n+1)}\}$ and $\{\mathbf{\Sigma}_j^{(n+1)}\}$ by solving (10) with $\mathbf{W}=\mathbf{W}^{(n+1)}$.
  Obtain $\mathbf{B}_i^{(n+1)}=\mathsf{ivec}(\mathbf{b}_i^{(n+1)},N_s)$.
  Set $n=n+1$.
until convergence

A. Convergence of Centralized Algorithm

Due to convexity of the problems in (8) and (10), we can observe that

$$\begin{aligned}\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^{(n)}\}_{i=1}^L,\mathbf{W}^{(n)})&\ge\min_{\mathbf{W}}\mathcal{E}_{\mathsf{C}}(\mathbf{W}|\{\mathbf{B}_i^{(n)}\}_{i=1}^L)\\&\ge\min_{\{\mathbf{b}_i\},\{\mathbf{\Sigma}_j\}}\mathcal{E}_{\mathsf{C}}(\{\mathbf{b}_i\}_{i=1}^L|\mathbf{W}^{(n+1)})\\&=\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^{(n+1)}\}_{i=1}^L,\mathbf{W}^{(n+1)}).\end{aligned}$$

Thus, $\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^{(n)}\}_{i=1}^L,\mathbf{W}^{(n)})$ monotonically decreases with the iteration index n and is bounded from below.

The objective function in (7) is block-convex, i.e. $\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i\}_{i=1}^L,\mathbf{W})$ is convex with the variables $\{\mathbf{B}_i\}_{i=1}^L$ or $\mathbf{W}$ fixed. Also, it can be seen that the problems (8) and (10) have an unique solution for a given $\{\mathbf{B}_i\}_{i=1}^L$ and $\mathbf{W}$, respectively. Further, for every feasible point $(\{\mathbf{B}_i^0\},\{\mathbf{\Sigma}_j^0\},\mathbf{W}^0)$, let us define a set

$$\begin{aligned}\mathcal{B}^0=\Big\{&(\{\mathbf{B}_i\},\{\mathbf{\Sigma}_j\},\mathbf{W})|\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i\}_{i=1}^L,\mathbf{W})\le\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^0\},\mathbf{W}^0),\\&0\le\tau_I\mathrm{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H+\mathbf{R}_i)\mathbf{B}_i^H)\le\big(\tau_E\alpha_iP_{H,i}(\{\mathbf{\Sigma}_j^0\})\\&-\tau P_{C,i}\big),\forall i,0\le\mathrm{tr}(\mathbf{\Sigma}_j)\le P_{T,j},\mathbf{\Sigma}_j\succeq\mathbf{0},\forall j\Big\}.\end{aligned}$$

This set is closed since it is a Cartesian product of the preimage of the continuous functions $\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i\}_{i=1}^L,\mathbf{W})$, $h_i:\mathbb{R}^{N_S\times M}\to\mathbb{R}_+,h_i(\mathbf{B}_i)=\tau_I\mathrm{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H+\mathbf{R}_i)\mathbf{B}_i^H)$, and $g_j:\mathbb{S}_+^{N_E\times N_E}\to\mathbb{R}_+,g_j(\mathbf{\Sigma}_j)=\mathrm{tr}(\mathbf{\Sigma}_j)$ on the closed sets $[0,\mathcal{E}_{\mathsf{C}}(\{\mathbf{B}_i^0\},\mathbf{W}^0)]$, $[0,\tau_E\alpha_iP_{H,i}(\{\mathbf{\Sigma}_j^0\})-\tau P_{C,i}]$, and $[0,P_{T,j}]$, $\forall i,j$, respectively. As $\mathcal{B}^0$ is bounded, from the Borel-Heine Theorem, $\mathcal{B}^0$ is a compact set. Hence, all the conditions in [29, Th. 1] for the convergence of alternating minimization algorithm are fulfilled. Thus, we can conclude that Algorithm 1 converges to the set of stationary points.

B. Computational Complexity

The optimization problem in (10), which comprises a quadratic objective function with LMN_S variables and positive semidefinite (PSD) constraint on K variables of dimension $N_E\times N_E$, can be reformulated as a semidefinite program (SDP). Then, the worst-case complexity to solve (10) can be determined as $\mathcal{O}((LMN_S+KN_E^2)^{3.5})$ [28]. Thus, for a large L, the computational requirements become prohibitively high.

The FC executes the algorithm and feeds back $\{\mathbf{B}_i\}$ to the sensors and $\{\mathbf{\Sigma}_j\}$ to the E-APs. Consequently, it requires the knowledge of the observation statistical information $\mathbf{A}_i$ and $\mathbf{R}_i$ along with global CSI to implement the algorithm. This leads to significant communication overheads between the nodes and FC, especially when the channels vary often. Therefore, it is desired to develop a distributed algorithm that has low computational complexity and communication overhead at every entity of the network. This is particularly important in WPSNs, since the nodes have a limited computational power and the communication between the nodes and the FC is expensive. In the next section, we propose a low complex distributed algorithm to find $\mathbf{W},\{\mathbf{B}_i\}_{i=1}^L$, and $\{\mathbf{\Sigma}_j\}_{j=1}^K$ without exchanging actual CSI and observation matrices.

IV. Distributed Algorithm for MSE Minimization

In this section, we present a distributed algorithm in which the network of sensors and E-APs collectively solve the global MSE minimization problem in (6). To this end, the sensor nodes need the knowledge of the optimal $\mathbf{W}$ to evaluate the precoders. Hence, first we begin with developing an algorithm to determine the fusion rule at every sensor nodes.

A. Local Computation of MMSE Fusion Rule

We assume that the FC does not have global CSI and observation statistical information. Therefore, it cannot calculate and feedback $\mathbf{W}$ to the nodes. In that case, the network of sensors collectively identify $\mathbf{W}$ by exchanging messages among themselves. Let $\mathbf{W}_i$ be the fusion rule calculated locally at node i. Then, the objective function (7) is redefined as

$$\begin{aligned}&\mathcal{E}_{\mathsf{D}}(\{\mathbf{B}_i\},\{\mathbf{W}_i\})\\&=\mathrm{tr}(\mathbf{R}_\theta)+\Big\|\sum_{i=1}^{L}\mathbf{W}_i^H\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\mathbf{R}_\theta^{\frac{1}{2}}\Big\|_F^2\\&\quad+\sum_{i=1}^{L}\mathrm{tr}\Big(\mathbf{W}_i^H\Big(\mathbf{H}_i\mathbf{B}_i\mathbf{R}_i\mathbf{B}_i^H\mathbf{H}_i^H+\frac{1}{L}\mathbf{R}_{\mathsf{FC}}\Big)\mathbf{W}_i\Big)\\&\quad-2\mathfrak{Re}\left(\mathrm{tr}(\mathbf{W}_i^H\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\mathbf{R}_\theta)\right).\end{aligned}\tag{11}$$

Now solving (8) is equivalent to addressing the optimization problem

$$\begin{aligned}&\min_{\{\mathbf{W}_i\},\mathbf{W}}\ \mathcal{E}_{\mathsf{D}}(\{\mathbf{B}_i\},\{\mathbf{W}_i\})\\&\ \text{s. t. } \mathbf{W}_i=\mathbf{W},\ i=1,\ldots,L,\end{aligned}\tag{12}$$

since both (8) and (12) have an identical solution. The constraint in (12) is known as the consensus constraint which forces the local information $\mathbf{W}_i$ at the ith node to be equal to the actual $\mathbf{W}$. Let us define $\mathbf{w}_i=\mathrm{vec}(\mathbf{W}_i)$ and $\mathbf{w}=\mathrm{vec}(\mathbf{W})$. Then, for a given $\{\mathbf{B}_i\}$, the MSE in (11) can be further modified as

$$\begin{aligned}\mathcal{E}_{\mathsf{D}}(\{\mathbf{w}_i\}|\{\mathbf{B}_i\})&=\mathrm{tr}(\mathbf{R}_\theta)+\Big\|\sum_{i=1}^{L}\mathbf{D}_i\mathbf{w}_i\Big\|^2\\&\quad+\sum_{i=1}^{L}\mathbf{w}_i^H\mathbf{Y}_i\mathbf{w}_i-2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{h}_i),\end{aligned}\tag{13}$$

where $\mathbf{h}_i\triangleq\mathrm{vec}(\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\mathbf{R}_\theta)\in\mathbb{C}^{N_FM_\theta}$, $\mathbf{D}_i\triangleq\mathbf{I}_{N_F}\otimes(\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\mathbf{R}_\theta^{\frac{1}{2}})^H\in\mathbb{C}^{N_F^2\times N_FM_\theta}$, and $\mathbf{Y}_i\triangleq\mathbf{I}_{M_\theta}\otimes(\mathbf{H}_i\mathbf{B}_i\mathbf{R}_i\mathbf{B}_i^H\mathbf{H}_i^H+\frac{1}{L}\mathbf{R}_{\mathsf{FC}})\in\mathbb{C}^{N_FM_\theta\times N_FM_\theta}$.

Subsequently, using the vectorization operator on the constraint in (12) and introducing auxiliary variables $\{\mathbf{f}_i\}$, the optimization problem to find $\mathbf{w}_i$ is recast as

$$\begin{aligned}&\min_{\{\mathbf{w}_i\},\{\mathbf{f}_i\},\mathbf{w}}\ \Big\|\sum_{i=1}^{L}\mathbf{f}_i\Big\|^2+\sum_{i=1}^{L}\mathbf{w}_i^H\mathbf{Y}_i\mathbf{w}_i-2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{h}_i)\\&\quad\text{s. t. } \mathbf{D}_i\mathbf{w}_i=\mathbf{f}_i,\ i=1,\ldots,L,\\&\qquad\quad \mathbf{w}_i=\mathbf{w},\ i=1,\ldots,L.\end{aligned}\tag{14}$$

In the above problem, the auxiliary variables $\{\mathbf{f}_i\}$ enable us to decompose the objective function into a sum of the local cost functions $l_i(\mathbf{w}_i)=\mathbf{w}_i^H\mathbf{Y}_i\mathbf{w}_i-2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{h}_i))$ and the shared global loss function $\left\|\sum_{i=1}^{L}\mathbf{f}_i\right\|^2$. The ith sensor node should independently design $\mathbf{w}_i$ and $\mathbf{f}_i$ to minimize the local cost as well as the global cost function. We can adopt the ADMM technique to solve this problem in a distributed manner without any coordination from the FC [30]. The ADMM enforces the consensus constraints in a scalable and robust manner, and achieves a faster convergence to the distributed solution.

To this end, the augmented Lagrangian for problem (14) with the quadratic penalty function for the constraint violations is formed as

$$\begin{aligned}&\mathcal{J}_w(\{\mathbf{w}_i\},\{\mathbf{f}_i\},\mathbf{w},\{\mathbf{y}_i\},\{\mathbf{x}_i\})\\&=\Big\|\sum_{j=1}^{L}\mathbf{f}_j\Big\|^2+\sum_{i=1}^{L}\Big(\mathbf{w}_i^H\mathbf{Y}_i\mathbf{w}_i-2\mathfrak{Re}\left(\mathbf{w}_i^H\mathbf{h}_i\right)\\&\quad+\mathfrak{Re}\left(\mathbf{y}_i^H(\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i)\right)+\frac{\rho_u}{2}\|\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i\|^2\\&\quad+\mathfrak{Re}\left(\mathbf{x}_i^H(\mathbf{w}_i-\mathbf{w})\right)+\frac{\rho_w}{2}\|\mathbf{w}_i-\mathbf{w}\|^2\Big),\end{aligned}\tag{15}$$

where ρ_u and ρ_w are penalty parameters, and $\mathbf{y}_i$ and $\mathbf{x}_i$ equal the dual variables for the first and second constraint in (14), respectively. We employ the augmented Lagrangian because the intermediate ascent step to update $\mathbf{w}$ is affine, which may yield unbounded solutions.

It can be seen that optimization for $\mathbf{w}_i$ is separable in (15) and the terms can be simplified as $\mathfrak{Re}(\mathbf{y}_i^H(\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i))+\frac{\rho_u}{2}\|\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i\|^2=\frac{\rho_u}{2}\|\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i+\mathbf{u}_i\|^2-\frac{\rho_u}{2}\|\mathbf{u}_i\|^2$, where $\mathbf{u}_i=\mathbf{y}_i/\rho_u$. Hence, relying on the ADMM technique [30], we obtain the following iterative steps for problem (14) as

$$\begin{aligned}\mathbf{w}_i^{(k+1)}&=\arg\min_{\mathbf{w}_i}\Big(\mathbf{w}_i^H\mathbf{Y}_i\mathbf{w}_i-2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{h}_i)\\&\quad+\mathfrak{Re}(\mathbf{x}_i^{(k)H}(\mathbf{w}_i-\mathbf{w}^{(k)}))+\frac{\rho_w}{2}\|\mathbf{w}_i-\mathbf{w}^{(k)}\|^2\\&\quad+\frac{\rho_u}{2}\|\mathbf{D}_i\mathbf{w}_i-\mathbf{f}_i^{(k)}+\mathbf{u}_i^{(k)}\|^2\Big),\end{aligned}\tag{16}$$

$$\mathbf{w}^{(k+1)}=\arg\min_{\mathbf{w}}\sum_{i=1}^{L}\frac{\rho_w}{2}\|\mathbf{w}_i^{(k+1)}-\mathbf{w}\|^2-\mathfrak{Re}(\mathbf{x}_i^{(k)H}\mathbf{w}),\tag{17}$$

$$\begin{aligned}\{\mathbf{f}_i^{(k+1)}\}_{i=1}^{L}&=\arg\min_{\{\mathbf{f}_j\}}\Big(\frac{\rho_u}{2}\sum_{j=1}^{L}\left\|\mathbf{D}_j\mathbf{w}_j^{(k+1)}-\mathbf{f}_j+\mathbf{u}_j^{(k)}\right\|^2\\&\quad+\Big\|\sum_{j=1}^{L}\mathbf{f}_j\Big\|^2\Big),\end{aligned}\tag{18}$$

$$\mathbf{x}_i^{(k+1)}=\mathbf{x}_i^{(k)}+\rho_w(\mathbf{w}_i^{(k+1)}-\mathbf{w}^{(k+1)}),\tag{19}$$

$$\mathbf{u}_i^{(k+1)}=\mathbf{u}_i^{(k)}+(\mathbf{D}_i\mathbf{w}_i^{(k+1)}-\mathbf{f}_i^{(k+1)}),\tag{20}$$

where the superscript k is the iteration index.

Next, we derive closed form solutions to address the subproblems in (16)–(20). It is apparent that (16) is an unconstrained quadratic optimization problem. Thus, by computing the gradient and equating it to zero, the optimal point at the ith node is given by

$$\begin{aligned}\mathbf{w}_i^{(k+1)}&=\left(\mathbf{Y}_i+\tfrac{\rho_w}{2}\mathbf{I}_{N_FM_\theta}+\tfrac{\rho_u}{2}\mathbf{D}_i^H\mathbf{D}_i\right)^{-1}\Big(\mathbf{h}_i+\tfrac{\rho_w}{2}\mathbf{w}^{(k)}\\&\quad-\tfrac{1}{2}\mathbf{x}_i^{(k)}+\tfrac{\rho_u}{2}\mathbf{D}_i^H\left(\mathbf{f}_i^{(k)}-\mathbf{u}_i^{(k)}\right)\Big).\end{aligned}\tag{21}$$

This step is executed independently in parallel at all nodes $i=1,\ldots,L$.

Differentiating the objective function in (17) and setting it to zero, it follows

$$\mathbf{w}^{(k+1)} = \frac{1}{L}\sum_{i=1}^{L}\left(\mathbf{w}_i^{(k+1)} + \frac{1}{\rho_w}\mathbf{x}_i^{(k)}\right). \tag{22}$$

Summing up the update expression in (19) overall i and substituting $\mathbf{w}^{(k+1)}$, we find that

$$\frac{1}{L}\sum_{i=1}^{L}\mathbf{x}_i^{(k+1)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{x}_i^{(k)} + \frac{\rho_w}{L}\sum_{i=1}^{L}\left(\mathbf{w}_i^{(k+1)} - \mathbf{w}^{(k+1)}\right) = \mathbf{0}$$

which indicates that the sum of dual variables $\mathbf{x}_i$ is zero. Therefore, the optimal consensus variable can be simplified as

$$\mathbf{w}^{(k+1)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{w}_i^{(k+1)}. \tag{23}$$

To solve the sub-problem in (18), let us introduce an auxiliary variable $\bar{\mathbf{f}} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{f}_i$. From the procedure outlined in Appendix, $\mathbf{f}_i^{(k+1)}$ and $\bar{\mathbf{f}}^{(k+1)}$ are given by

$$\mathbf{f}_i^{(k+1)} = \bar{\mathbf{f}}^{(k+1)} + \mathbf{D}_i\mathbf{w}_i^{(k+1)} + \mathbf{u}_i^{(k)} - \frac{1}{L}\sum_{i=1}^{L}\left(\mathbf{D}_i\mathbf{w}_i^{(k+1)} + \mathbf{u}_i^{(k)}\right) \tag{24}$$

and

$$\bar{\mathbf{f}}^{(k+1)} = \frac{\rho_u}{2L+\rho_u}\sum_{j=1}^{L}(\mathbf{D}_j\mathbf{w}_j^{(k+1)} + \mathbf{u}_j^{(k)}). \tag{25}$$

Now, substituting $\mathbf{f}_i^{(k+1)}$ in the update step in (20) yields $\mathbf{u}_i^{(k+1)} = \frac{1}{L}\sum_{j=1}^{L}(\mathbf{D}_j\mathbf{w}_j^{(k+1)} + \mathbf{u}_j^{(k)}) - \bar{\mathbf{f}}^{(k+1)}$, which implies that the dual variables $\mathbf{u}_i^{(k+1)}$ are equal across all the sensor nodes and can be replaced with a single variable $\mathbf{u}^{(k+1)}$.

Hence, the expression for $\mathbf{f}_i^{(k+1)}$ reduces to $\mathbf{f}_i^{(k+1)} = \bar{\mathbf{f}}^{(k+1)} + \mathbf{D}_i\mathbf{w}_i^{(k+1)} - \frac{1}{L}\sum_{i=1}^{L}\mathbf{D}_i\mathbf{w}_i^{(k+1)}$ and the dual variable $\mathbf{u}^{(k+1)}$ can be determined as

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \bar{\mathbf{f}}^{(k+1)} + \frac{1}{L}\sum_{j=1}^{L}(\mathbf{D}_j\mathbf{w}_j^{(k+1)}). \tag{26}$$

This simplifies the update step in (25) to

$$\bar{\mathbf{f}}^{(k+1)} = \frac{\rho_u}{2L+\rho_u}\left(L\mathbf{u}^{(k)} + \sum_{j=1}^{L}\mathbf{D}_j\mathbf{w}_j^{(k+1)}\right). \tag{27}$$

The update steps in (23) and (26) require the nodes to gather $\mathbf{D}_i\mathbf{w}_i$ and $\mathbf{w}_i$ to form the averages. These updates are also carried out in parallel by all the nodes. Finally, inserting the optimal $\mathbf{f}_i^{(k+1)}$ and $\mathbf{w}^{(k+1)}$ in (21), $\mathbf{w}_i$ at node i can be updated as

$$\mathbf{w}_i^{(k+1)} = \left(\mathbf{Y}_i + \frac{\rho_w}{2}\mathbf{I}_{N_F M_\theta} + \frac{\rho_u}{2}\mathbf{D}_i^H\mathbf{D}_i\right)^{-1}\left(\mathbf{h}_i + \frac{\rho_w}{2}\mathbf{w}^{(k)} - \frac{\mathbf{x}_i^{(k)}}{2} + \frac{\rho_u}{2}\mathbf{D}_i^H\left(\bar{\mathbf{f}}^{(k)} + \mathbf{D}_i\mathbf{w}_i^{(k)} - \bar{\mathbf{d}}_w^{(k)} - \mathbf{u}^{(k)}\right)\right), \tag{28}$$

where $\bar{\mathbf{d}}_w^{(k)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{D}_i\mathbf{w}_i^{(k)}$.

B. Distributed Computation of Precoders and Energy Covariance Matrices

With the fusion rule known at the sensor nodes, it remains to find the optimal precoders and the energy covariance matrices. In this section, we present an algorithm to locally calculate $\mathbf{B}_i$ at the ith sensor node and $\mathbf{\Sigma}_j$ at the jth E-AP to minimize the MSE. Let us denote $\mathbf{b}_i \triangleq \text{vec}(\mathbf{B}_i)$ and $\mathbf{W}_i$ as the fusion rule obtained locally at node i. Replacing the objective function in (10) with (11), we can recast the precoder optimization problem as

$$\begin{aligned}
\min_{\substack{\{\mathbf{b}_i\},\{\mathbf{z}_i\}\\ \{\mathbf{\Sigma}_j\}}} \ & \left\|\sum_{i=1}^{L}\mathbf{z}_i\right\|^2 + \sum_{i=1}^{L}\mathbf{b}_i^H\mathbf{\Xi}_i\mathbf{b}_i - 2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{H}_{A,i}\mathbf{b}_i)\\
\text{s. t. } & \mathbf{C}_i\mathbf{b}_i = \mathbf{z}_i,\ i = 1,\ldots,L,\\
& \tau P_{C,i} + \tau_I\mathbf{b}_i^H\mathbf{Q}_i\mathbf{b}_i \\
& \quad \le \frac{\tau_E\alpha_i M_i}{1-\Omega_i}\Big(\frac{1}{1+e^{-a_i(\theta_i - b_i)}} - \Omega_i\Big), i = 1,\ldots,L\\
& \theta_i \le \sum\nolimits_{j=1}^{K}\text{tr}(\mathbf{G}_{ij}\mathbf{\Sigma}_j\mathbf{G}_{ij}^H),\ i = 1,\ldots,L,\\
& \mathbf{\Sigma}_j \succeq 0 \text{ and } \text{tr}(\mathbf{\Sigma}_j) \le P_{T,j},\ j = 1,\ldots,K,
\end{aligned} \tag{29}$$

where $\{\mathbf{z}_i\}$ and $\{\theta_i\}$ are the auxiliary variables, and we define $\mathbf{\Xi}_i \triangleq \mathbf{R}_i^T \otimes \mathbf{H}_i^H\mathbf{W}_i\mathbf{W}_i^H\mathbf{H}_i \in \mathbb{C}^{MN_S\times MN_S}$, $\mathbf{H}_{A,i} \triangleq \mathbf{R}_\theta^T\mathbf{A}_i^T \otimes \mathbf{H}_i \in \mathbb{C}^{N_F M_\theta \times MN_S}$, and $\mathbf{C}_i \triangleq (\mathbf{R}_\theta^{\frac{1}{2}})^T\mathbf{A}_i^T \otimes \mathbf{W}_i^H\mathbf{H}_i \in \mathbb{C}^{M_\theta^2\times MN_S}$.

Let $\{\mathbf{y}_i\}$, $\{\lambda_i\}$ and $\{\beta_i\}$ be the dual variables associated with the first constraint, power constraint on $\mathbf{b}_i$, and third constraint in (29), respectively. Further, let us denote $\{\nu_j\}$ and $\{\mathbf{Z}_j\}$ as the dual variables corresponding to power constraint at E-APs and PSD constraints on the energy covariance matrices, respectively. The augmented Lagrangian of problem (29) is now given by

$$\begin{aligned}
\mathcal{J}_b = & \left\|\sum_{i=1}^{L}\mathbf{z}_i\right\|^2 + \sum_{i=1}^{L}\Big(\mathbf{b}_i^H\mathbf{\Xi}_i\mathbf{b}_i - 2\mathfrak{Re}(\mathbf{w}_i^H\mathbf{H}_{A,i}\mathbf{b}_i)\\
& + \mathfrak{Re}(\mathbf{y}_i^H(\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i)) + \frac{\rho_v}{2}\|\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i\|^2 + \lambda_i\tau P_{C,i} + \beta_i\theta_i\\
& + \lambda_i\tau_I\mathbf{b}_i^H\mathbf{Q}_i\mathbf{b}_i - \frac{\tau_E\alpha_i M_i\lambda_i}{1-\Omega_i}\Big(\frac{1}{1+e^{-a_i(\theta_i-b_i)}} - \Omega_i\Big)\Big)\\
& + \sum_{j=1}^{K}\text{tr}\left(\left(\nu_j\mathbf{I}_{N_E} - \mathbf{Z}_j - \sum_{i=1}^{L}\beta_i\mathbf{G}_{ij}^H\mathbf{G}_{ij}\right)\mathbf{\Sigma}_j\right) - \nu_j P_{T,j},
\end{aligned} \tag{30}$$

where $\frac{\rho_v}{2}\|\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i\|^2$ represents the penalty function on constraint in (29) and ρ_v indicates the penalty parameter. Defining $\mathbf{v}_i = \mathbf{y}_i/\rho_v$, the term $\mathfrak{Re}(\mathbf{y}_i^H(\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i)) + \frac{\rho_v}{2}\|\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i\|^2$ is simplified as $\frac{\rho_v}{2}\|\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i + \mathbf{v}_i\|^2 - \frac{\rho_v}{2}\|\mathbf{v}_i\|^2$.

From (30), it can be seen that the optimization for $\mathbf{b}_i$ is separable. Hence, adopting the ADMM technique, $\{\mathbf{b}_i\}$ and

$\{\boldsymbol{\Sigma}_j\}$ are determined as

$$\mathbf{b}_i^{(k+1)} = \arg\min_{\mathbf{b}_i}\bigg(\mathbf{b}_i^H\boldsymbol{\Xi}_i\mathbf{b}_i - 2\Re\mathfrak{e}(\mathbf{w}_i^H\mathbf{H}_{A,i}\mathbf{b}_i) + \lambda_i^{(k+1)}\tau_I\mathbf{b}_i^H\mathbf{Q}_i\mathbf{b}_i + \frac{\rho_v}{2}\|\mathbf{C}_i\mathbf{b}_i - \mathbf{z}_i^{(k)} + \mathbf{v}_i^{(k)}\|^2\bigg), \quad (31)$$

$$\{\mathbf{z}_i^{(k+1)}\}_{i=1}^L = \arg\min_{\{\mathbf{z}_j\}}\bigg(\frac{\rho_v}{2}\sum_{j=1}^L\big\|\mathbf{C}_j\mathbf{b}_j^{(k+1)} - \mathbf{z}_j + \mathbf{v}_j^{(k)}\big\|^2 + \Big\|\sum_{j=1}^L\mathbf{z}_j\Big\|^2\bigg), \quad (32)$$

$$\mathbf{v}_i^{(k+1)} = \mathbf{v}_i^{(k)} + (\mathbf{C}_i\mathbf{b}_i^{(k+1)} - \mathbf{z}_i^{(k+1)}), \quad (33)$$

$$\theta_i^{(k+1)} = \arg\min_{\theta_i}\bigg(\beta_i^{(k+1)}\theta_i - \lambda_i^{(k+1)}\frac{\tau_E\alpha_i M_i}{1-\Omega_i} \times\Big(\frac{1}{1+e^{-a_i(\theta_i - b_i)}} - \Omega_i\Big)\bigg), \quad (34)$$

$$\{\boldsymbol{\Sigma}_j^{(k+1)}, \nu_j^{(k+1)}\}_{j=1}^K = \arg\max_{\substack{\{\nu_l\geq 0\}\\ \{\mathbf{Z}_l\succeq\mathbf{0}\}}}\min_{\{\boldsymbol{\Sigma}_l\}}\bigg(\sum_{l=1}^K \mathrm{tr}\Big(\Big(\nu_l\mathbf{I}_{N_E} - \mathbf{Z}_l - \sum_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{il}^H\mathbf{G}_{il}\Big)\boldsymbol{\Sigma}_l\Big) - \nu_l P_{T,l}\bigg). \quad (35)$$

Next, we derive closed form expressions to solve the above steps in parallel at different nodes. As (31) is a quadratic minimization problem, the first order optimality conditions lead to a solution

$$\mathbf{b}_i^{(k+1)} = \Big(\boldsymbol{\Xi}_i + \lambda_i^{(k+1)}\tau_I\mathbf{Q}_i + \frac{\rho_v}{2}\mathbf{C}_i^H\mathbf{C}_i\Big)^{-1} \times\Big(\mathbf{H}_{A,i}^H\mathbf{w}_i^{(k)} + \frac{\rho_v}{2}\mathbf{C}_i^H(\mathbf{z}_i^{(k)} - \mathbf{v}_i^{(k)})\Big). \quad (36)$$

The precoder $\mathbf{b}_i$ should satisfy the complementary slackness conditions $\lambda_i^{(k+1)}(\tau_I\mathbf{b}_i^{(k+1)H}\mathbf{Q}_i\mathbf{b}_i^{(k+1)} - \tau_E\alpha_i P_{H,i}^{(k)} + \tau P_{C,i}) = 0, \forall i$.

Therefore, from (36), we can evaluate the Lagrange multipliers $\lambda_i^{(k+1)}$ from the equation [31, Appendix A]

$$\Big\|\mathbf{Q}_i^{\frac{1}{2}}\Big(\boldsymbol{\Xi}_i + \lambda_i^{(k+1)}\tau_I\mathbf{Q}_i + \frac{\rho_v}{2}\mathbf{C}_i^H\mathbf{C}_i\Big)^{-1}\mathbf{t}_i^{(k)}\Big\|^2 = \frac{\tau_E\alpha_i P_{H,i}^{(k)} - \tau P_{C,i}}{\tau_I}, \quad (37)$$

where $\mathbf{t}_i^{(k)} = \mathbf{H}_{A,i}^H\mathbf{w}_i^{(k)} + \frac{\rho_v}{2}\mathbf{C}_i^H(\mathbf{z}_i^{(k)} - \mathbf{v}_i^{(k)})$. It should be noted that the sensor nodes need not know the actual energy covariance matrices, but only estimate the harvested power $P_{H,i}^{(k)}$.

Next, from the procedure described in Appendix, a solution to the problem in (32) is obtained as

$$\mathbf{z}_i^{(k)} = \bar{\mathbf{z}}^{(k)} + \mathbf{C}_i\mathbf{b}_i^{(k)} + \mathbf{v}_i^{(k)} - \frac{1}{L}\sum_{i=1}^L\mathbf{C}_i\mathbf{b}_i^{(k)} + \mathbf{v}_i^{(k)}, \quad (38)$$

where $\bar{\mathbf{z}}^{(k+1)} = \frac{\rho_v}{2L+\rho_v}\frac{1}{L}\sum_{j=1}^L(\mathbf{C}_j\mathbf{b}_j^{(k+1)} + \mathbf{v}_j^{(k)})$. By applying arguments similar to Section IV-A, we can prove that the dual variable $\mathbf{v}_i^{(k+1)}$ is the same for all the sensor nodes. Hence, replacing $\mathbf{v}_i^{(k+1)}$ for $i = 1,\ldots,L$ with $\mathbf{v}^{(k+1)}$, the dual variable is written as

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} - \bar{\mathbf{z}}^{(k+1)} + \frac{1}{L}\sum_{j=1}^L(\mathbf{C}_j\mathbf{b}_j^{(k+1)}). \quad (39)$$

By computing the derivative of the objective function in (34) with respect to θ_i and equating it to zero, β_i can be determined as

$$\beta_i^{(k+1)} = a_i\lambda_i^{(k+1)}\frac{\tau_E\alpha_i M_i}{1-\Omega_i}\frac{\exp\big[-a_i(\theta_i^{(k+1)} - b_i)\big]}{\Big(1+\exp\big[-a_i(\theta_i^{(k+1)} - b_i)\big]\Big)^2}. \quad (40)$$

It can be shown that for the optimal solution of (29), the third constraint will be satisfied with equality. Therefore, $\theta_i^{(k+1)}$ can be updated as $\theta_i^{(k+1)} = \sum_{j=1}^K \mathrm{tr}(\mathbf{G}_{ij}\boldsymbol{\Sigma}_j^{(k)}\mathbf{G}_{ij}^H)$.

Next, to obtain $\{\boldsymbol{\Sigma}_j^{(k+1)}\}$, the dual function in (35) is expressed as

$$g(\{\nu_j\},\{\mathbf{Z}_j\}) = \inf_{\{\boldsymbol{\Sigma}_j\}}\bigg(\sum_{j=1}^K -\nu_j P_{T,j} + \mathrm{tr}\Big(\Big(\nu_j\mathbf{I}_{N_E} - \mathbf{Z}_j - \sum_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\Big)\boldsymbol{\Sigma}_j\Big)\bigg). \quad (41)$$

For a given optimal $\{\beta_i^{(k+1)}\}$, $g(\{\nu_j\},\{\mathbf{Z}_j\})$ can be decomposed into K independent functions. Moreover, the objective function in (41) is linear in $\boldsymbol{\Sigma}_j$, which reduces to [32]

$$g(\{\nu_j\},\{\mathbf{Z}_j\}) = \begin{cases} -\nu_j P_{T,j}, \text{ if } \nu_j\mathbf{I}_{N_E} - \sum_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\succeq\mathbf{0}, \\ -\infty, \text{ otherwise.}\end{cases}$$

Thus, a solution $\nu_j^{(k+1)}$ of the dual optimization problem in (35) can be derived as

$$\min_{\nu_j}\ \nu_j P_{T,j} \quad \text{s. t. } \nu_j\geq 0 \text{ and } \nu_j\mathbf{I}_{N_E}\succeq\sum\nolimits_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}. \quad (42)$$

It is known that a solution of the above problem equals $\nu_j^{(k+1)} = \lambda_{\max}\Big(\sum_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\Big)$, where $\lambda_{\max}(\mathbf{X})$ is the dominant eigenvalue of a matrix $\mathbf{X}$ [32].

In addition, the dual of the above problem in (42) is given by

$$\max_{\mathbf{X}_j}\ \mathrm{tr}\Big(\sum\nolimits_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\mathbf{X}_j\Big) \quad \text{s. t. } \mathbf{X}_j\succeq\mathbf{0} \text{ and } \mathrm{tr}(\mathbf{X}_j)\leq P_{T,j}. \quad (43)$$

The problem is optimized by a solution $\mathbf{X}_j^* = P_{T,j}\mathbf{q}_j\mathbf{q}_j^H$ for $j=1,\ldots,K$, where $\mathbf{q}_j$ is obtained as $\mathbf{q}_j = \nu_{\max}\Big(\sum_{i=1}^L\beta_i^{(k+1)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\Big)$ and $\nu_{\max}(\mathbf{X})$ represents the eigenvector corresponding to the dominant eigenvalue of a matrix $\mathbf{X}$ [32]. Since the Slater's conditions are met in the problem in (42), strong duality is satisfied by the solutions of (42) and (43). Therefore, the optimal energy covariance matrices are rank-one and are determined by

$$\mathbf{\Sigma}_j^{(k+1)} = P_{T,j}\mathbf{q}_j\mathbf{q}_j^H,\ j = 1,\ldots,K.$$

Finally, substituting (38) and (39) in (36), $\{\mathbf{b}_i\}$ are updated as

$$\mathbf{b}_i^{(k+1)} = \left(\mathbf{\Xi}_i + \lambda_i^{(k+1)}\tau_I\mathbf{Q}_i + \frac{\rho_v}{2}\mathbf{C}_i^H\mathbf{C}_i\right)^{-1}\mathbf{t}_i^{(k)}, \quad (44)$$

$i = 1,\ldots,L$, where $\mathbf{t}_i^{(k)} = \mathbf{H}_{A,i}^H\mathbf{w}_i^{(k)} + \frac{\rho_v}{2}\mathbf{C}_i^H(\mathbf{C}_i\mathbf{b}_i^{(k)} + \bar{\mathbf{z}}^{(k)} - \bar{\mathbf{c}}_b^{(k)} - \mathbf{v}^{(k)})$ and $\bar{\mathbf{c}}_b^{(k)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{C}_i\mathbf{b}_i^{(k)}$. Collating the derivations presented in the previous sections, the distributed algorithm to find $\{\mathbf{b}_i\}$, $\{\mathbf{w}_i\}$, and $\{\mathbf{\Sigma}_j\}$ is summarized in Algorithm 2 below. In Algorithm 2, the inner loop allows the distributed computation of precoders and the MMSE fusion rule, whereas the outer loop is needed for alternating minimization of $\{\mathbf{b}_i\}$ and $\{\mathbf{w}_i\}$.

Algorithm 2 Distributed Algorithm for MSE Minimization

Initialize $n = 0$ and $\mathbf{B}_i^{(0)}$ for $i = 1,\ldots,L$.
repeat
  repeat
    Compute $\mathbf{w}_i^{(k)}$ from (28).
    Broadcast $\mathbf{D}_i\mathbf{w}_i^{(k)}$ and $\mathbf{w}_i^{(k)}$ to other nodes.
    Calculate local variables $\bar{\mathbf{d}}_w^{(k)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{D}_i\mathbf{w}_i^{(k)}$ and $\mathbf{w}^{(k)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{w}_i^{(k)}$.
    Update local variable $\bar{\mathbf{f}}^{(k)}$ using (27).
    Update dual variables $\mathbf{u}^{(k)}$ using (26) and $\mathbf{x}_i^{(k)}$ using (19).
  until convergence
  Update $\mathbf{\Xi}_i$ and $\mathbf{C}_i$ for $i = 1,\ldots,L$.
  repeat
    Compute $\mathbf{b}_i^{(k)}$ from (44), $\lambda_i^{(k)}$ from (37), and $\beta_i^{(k)}$ from (40).
    Broadcast $\beta_i^{(k)}$ to E-APs and $\mathbf{C}_i\mathbf{b}_i^{(k)}$ to other nodes.
    Compute local variable $\bar{\mathbf{c}}_b^{(k)} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{C}_i\mathbf{b}_i^{(k)}$.
    Update $\bar{\mathbf{z}}^{(k)}$ using (38) and $\mathbf{v}^{(k)}$ using (39).
    Obtain $\mathbf{q}_j = \nu_{\max}\left(\sum_i \beta_i^{(k)}\mathbf{G}_{ij}^H\mathbf{G}_{ij}\right)$.
    Broadcast $\mathbf{s}_j$ with $\mathbf{\Sigma}_j^{(k)} = P_{T,j}\mathbf{q}_j\mathbf{q}_j^H$ for energy transfer.
  until convergence
  Update $\mathbf{Y}_i, \mathbf{D}_i$, and $\mathbf{h}_i$ for $i = 1,\ldots,L$.
  Set $n = n + 1$.
until $\{\mathbf{b}_i\}$, $\{\mathbf{w}_i\}$, $\{\mathbf{\Sigma}_j\}$ converge

Since the dual variables converge to the optimal point and are the same at all the nodes in the network, the nodes can simultaneously switch between updating the block variables when $\|\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}\| \le \epsilon_u$ and $\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\| \le \epsilon_v$ without requiring a central controller to coordinate the alternating minimization process. Here, the values for ϵ_u and ϵ_v can be chosen prior based on the desired accuracy.

C. Convergence of Distributed Algorithm

It is evident that strong duality holds for problems in (14) and (29) since they satisfy the Slater's conditions. Further, the objective functions in (14) and (29) are closed, proper, and quadratic convex functions. In addition, they are a sum of Lipschitz continuous functions[1] since $\mathbf{Y}_i \succ \mathbf{0}$, $\mathbf{R}_i \succ \mathbf{0}$, $\mathbf{\Xi}_i = \mathbf{R}_i^T \otimes \mathbf{H}_i^H\mathbf{W}_i\mathbf{W}_i^H\mathbf{H}_i \succ \mathbf{0}$, $\forall i$. As a result, from [33, Table 2], it can be concluded that $\{\mathbf{w}_i^{(k)}\}, \{\mathbf{f}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}$ and $\{\mathbf{b}_i^{(k)}, \mathbf{\Sigma}_j^{(k)}\}$, $\{\mathbf{z}_i^{(k)}\}$, $\{\mathbf{v}_i^{(k)}\}$ from the recursions in (16)–(20) and (31)–(35) converge to the optimal points of (14) and (29), respectively. The outer iterations are nothing but the alternating minimization of (14) and (29), which are equivalent to the optimization problems (8) and (10) by definition. Therefore, by the arguments presented in Section III-A for the convergence of the alternating minimization procedure to solve (6), Algorithm 2 is guaranteed to converge to the set of stationary points.

D. Computational Complexity and Message Exchange Overhead

For the calculations in (21), (36) and (43), the computational requirement at each sensor node is $\mathcal{O}((MN_S)^3 + (M_\theta N_F)^3)$ and that at the each E-AP is $\mathcal{O}(N_E^3)$. Therefore, the proposed distributed algorithm has much lower complexity in contrast to the centralized algorithm, which has the complexity of $\mathcal{O}((LMN_S + KN_E)^{3.5})$. For large L, the distributed algorithm is computationally efficient by a factor of $\mathcal{O}(L^{2.5})$. For example, in a WPSN consisting of two E-APs with $N_E = 4$, 10 sensors with $M_\theta = 2$, $M = 2$, $N_S = 2$, and a FC with $N_F = 4$, it costs 7.662×10^5 floating point operations (flops) at the FC to execute the centralized algorithm. In contrast, the distributed algorithm requires only 6.4×10^2 flops which is less than 0.1% compared to the centralized algorithm.

As for the message exchange overhead, node i computes $\mathbf{b}_i, \lambda_i, \mathbf{C}_i\mathbf{b}_i$, $\mathbf{w}_i$ and $\mathbf{D}_i\mathbf{w}_i$, and broadcasts $\mathbf{C}_i\mathbf{b}_i$, $\mathbf{w}_i$, and $\mathbf{D}_i\mathbf{w}_i$ to other sensor nodes and λ_i to the E-APs. Hence, the number of messages shared by a node in each iteration is $\mathcal{O}(M_\theta^2 + N_F M_\theta + N_F^2)$, which is independent of the number of observations and the number of sensors. The sensors can obtain CSI using the training signals transmitted by the FC. Finally, the nodes may transmit the MMSE fusion rule to the FC in order to relieve the FC from acquiring the CSI and the observation statistical information.

V. Simulation Results

In this section, we present the MSE performance of the centralized and the distributed algorithms proposed in Sections III and IV. To benchmark the MSE performance, we include the Bayesian Cramer-Rao bound (BCRB) on the MSE of the estimate for the ideal scenario where the observations $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_L^T]^T$ are available perfectly at the FC for parameter estimation. The BCRB is given by [34]

$$\mathsf{BCRB} = \mathrm{tr}((\mathbf{R}_\theta^{-1} + \mathbf{A}^H\mathbf{R}^{-1}\mathbf{A})^{-1}), \quad (45)$$

where $\mathbf{A} \triangleq [\mathbf{A}_1^T, \mathbf{A}_2^T, \ldots, \mathbf{A}_L^T]^T$ and $\mathbf{R} \triangleq \mathsf{diag}(\mathbf{R}_1, \ldots, \mathbf{R}_L)$ defines a block diagonal matrix with matrices $\mathbf{R}_i$ on the ith diagonal. For comparison, we also consider a conventional two-step framework for parameter estimation in WPSNs.

[1]A function f from $\mathbb{R}^n$ into $\mathbb{R}$ is called Lipschitz continuous if there is a constant $\alpha > 0$ such that $\|\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})\| \le \alpha\|\mathbf{y} - \mathbf{x}\|$, for all $\mathbf{y}, \mathbf{x} \in \mathbb{R}^n$. It can be easily seen that $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ is Lipschitz continuous for $\mathbf{A} \succ \mathbf{0}$.

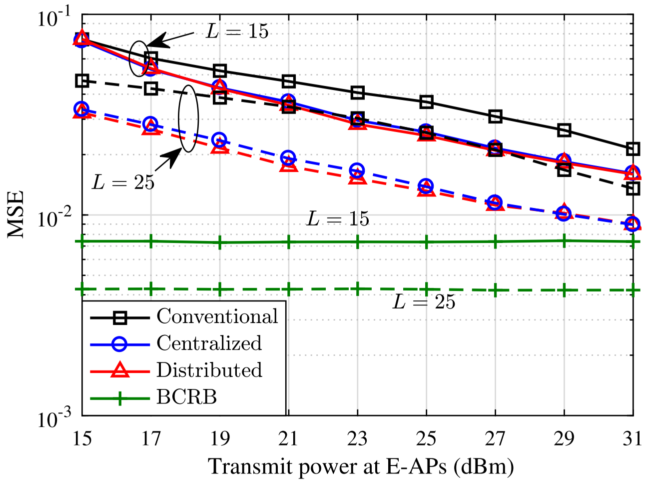

Fig. 2. MSE as a function of transmit power at E-APs with L sensors and $K = 2$.

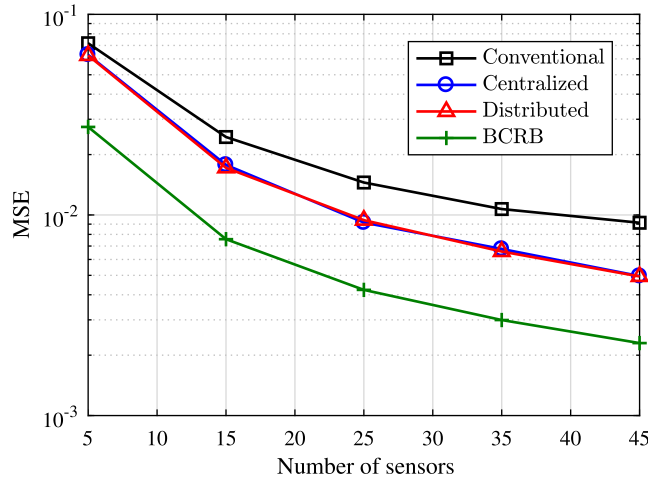

Fig. 3. MSE as a function of number of sensors with $K = 2$ and $P_{T,1} = P_{T,2} = 30$ dBm.

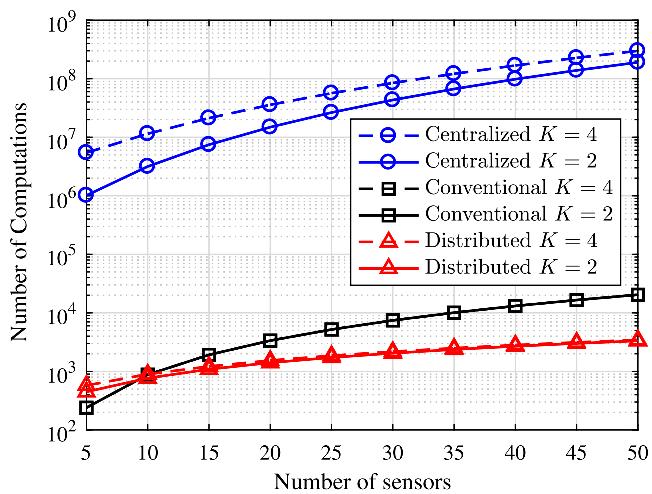

Fig. 4. Computational complexity versus number of sensors for K E-APs, $M = 2$, $N_S = 2$, $N_F = 2$ and $N_E = 4$.

In the first step, the energy covariance matrices $\{\boldsymbol{\Sigma}_j^{\mathsf{conv}}\}$ are designed to maximize the total harvested power $P_{H,\mathsf{sum}}^{\mathsf{conv}} = \sum_{i=1}^{L} P_{H,i}$ employing the procedure proposed in [24], where $P_{H,i}$ is given by (1) and (2). In the next step, the optimal joint MMSE precoder and the fusion rule are obtained as in [3] with the data transmission power at the ith sensor set as $P_{H,i}^{\mathsf{conv}}$.

For the simulations, we consider a WPSN with L sensors and K E-APs distributed randomly where the distance between the sensor nodes and the FC is 100 m and that between the sensor nodes and the E-APs is 10 m. The number of antennas at the sensors, the FC, and the E-APs are $N_S = 2$, $N_F = 2$, and $N_E = 4$, respectively. The circuit energy $P_{C,i}$ is set to 20 μW, $\forall i$, time durations τ_I and τ_E are set to 0.5. The parameters for EH circuit are taken as $a_i = 1500$, $b_i = 0.0014$, and $M_i = 24$ mW [25]. The FC noise covariance matrix is set as $\mathbf{R}_{\mathsf{FC}} = \sigma_{\mathsf{FC}}^2 \mathbf{I}_{N_F}$ with the noise variance $\sigma_{\mathsf{FC}}^2 = -90$ dBm [35], [36]. All the channel coefficients are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance. The path-loss is set to $c_0 \left(\frac{d}{d_0}\right)^{-\omega}$, where $c_0 = -20$ dB is a constant attenuation at a reference distance $d_0 = 1$ m, d represents the distance between two terminals, and $\omega = 3$ indicates the path-loss exponent [15], [37]. The parameter is assumed to be Gaussian with $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_\theta})$ and $M_\theta = 2$. The sensor nodes acquire observations of dimension $M = 2$. The elements of the observation matrix $\mathbf{A}_i$ are generated as i.i.d. Gaussian random variables with zero mean and unit variance. The observation noise covariance matrix at the sensor nodes $\mathbf{R}_i$ equals $0.1\mathbf{I}_M$. For all simulations, the initialization is chosen as $\mathbf{B}_i^{(0)} = \sqrt{\frac{\sum_{j=1}^{K} P_{T,j} \|\mathbf{G}_{ij}\|_F^2}{N_S M}} \mathbf{1}_{N_S \times M}$, where $\mathbf{1}_{N_S \times M}$ stands for a $N_S \times M$ matrix of all ones. The number of iterations is fixed to $n_{\max} = 20$, while the number of consensus iterations for Algorithm 2 is set to $k_{\max} = 20$ and the ADMM penalty parameters are taken as $\rho_w = \rho_u = \rho_v = 4$.

Fig. 2 shows the MSE of the estimate as a function of the transmit power at the E-APs for $K = 2$ and $L = 15$ or 20. From the figure, we can observe that the proposed joint designs approach the BCRB quickly and have superior estimation accuracy in comparison to a conventional design. It also confirms that the distributed algorithm performs very close to the centralized algorithm with much reduced complexity. The proposed algorithms yield as much as 5 dB gain for $L = 15$ and 7 dB gain for $L = 25$ over the conventional method.

In Fig. 3, we plot the MSE in terms of the number of sensors with the transmit power at the E-APs $P_{T,1} = P_{T,2} = 30$ dBm. It should be noted that as the number of sensors increases, the performance gain becomes more pronounced in comparison to the conventional techniques. This is because the proposed methods recharge the sensors taking into account local CSI and observation statistics at the nodes, thereby increasing the estimation diversity [38]. Moreover, the centralized and distributed algorithms exhibit almost the same performance. This is important because, in contrast to the centralized algorithm, the computational complexity of the distributed algorithm is independent of the number of sensors.

In Fig. 4, we plot the total computational complexity as a function of number of sensors in a WPSN with $M = 2$,

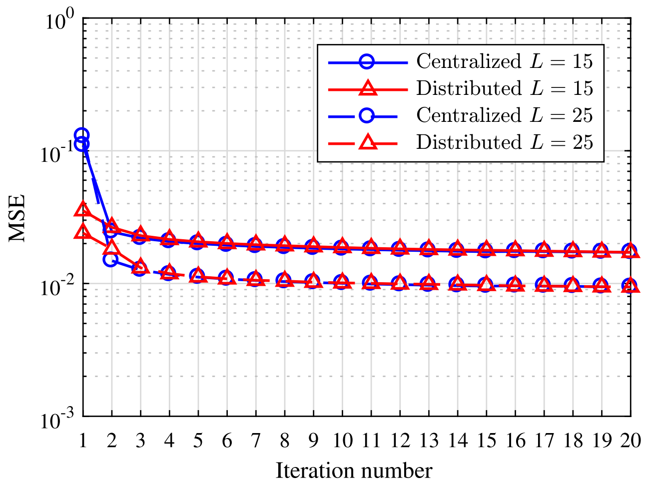

Fig. 5. MSE as a function of iteration number with L sensors, $K = 2$, $P_{T,1} = P_{T,2} = 30$ dBm and $k_{\max} = 20$.

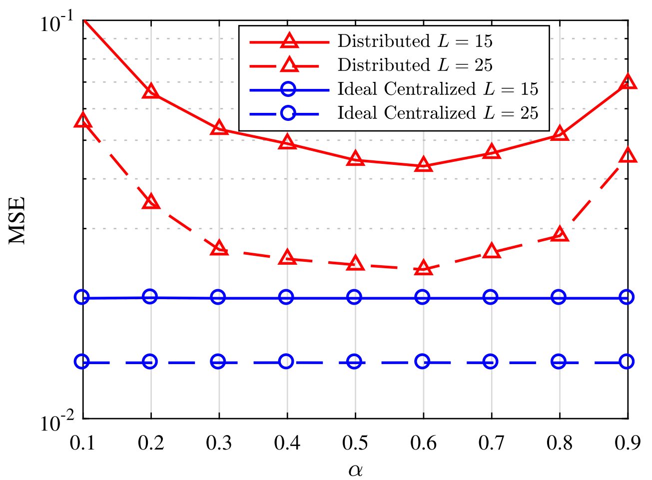

Fig. 6. MSE as a function of α with $K = 2$, $L = 15, 25$ and $P_{T,1} = P_{T,2} = 25$ dBm.

$N_S = 2$, $N_F = 2$ and $N_E = 4$. The centralized algorithm has a complexity of $\mathcal{O}((LMN_S + KN_E)^{3.5})$ at the FC. As derived in Section IV-D, the complexity of the distributed algorithm is $\mathcal{O}((MN_S)^3 + (M_\theta N_F)^3)$ at each sensor node and $\mathcal{O}(N_E^3)$ at each E-AP. On the other hand, the conventional scheme in [3], which is an iterative algorithm, carries out $L\mathcal{O}(N_S^3 + LMN_S^2)$ computations in each iteration. The graph shows that as the number of sensors increases, the computational requirement to execute the centralized algorithm becomes prohibitively large. In contrast, the distributed algorithm offers a significant reduction in computational costs over both the centralized algorithm and the conventional method.

In Fig. 5, we demonstrate the convergence behavior of Algorithms 1 and 2. The number of times that the sensors exchange the messages is set to $k_{\max} = 20$. The simulation shows that the MSE monotonically decreases with the iteration number and converges in a few iterations. It can be seen that even for a larger number of sensors the distributed algorithm converges within 10 iterations.

Fig. 6 presents the MSE performance of the distributed algorithm for varying levels of $\alpha = \alpha_i$. We consider a WPSN where the distance between sensors is equal to 10m, and assume the channel between sensors is Gaussian with the noise variance $\sigma_{\mathsf{FC}}^2 = -90$ dBm and path-loss exponent $\omega = 2$. The performance gain decreases when α is close to zero or one. This because as $\alpha \to 1$, the accuracy of the messages exchanged reduces due to a lower SNR at nodes, and as $\alpha \to 0$, the resource available for data transmission gets smaller. From Fig. 6, it can be observed that the MSE is more sensitive to the availability of energy resources for data transmission. In the distributed algorithm, each node shares a total of $n_{\mathsf{iter}}\mathcal{O}(M_\theta^2 + N_F M_\theta + N_F^2)$ messages, where n_{iter} is the number of exchanges over short distances. This consumes only a fraction of the energy used for data transmission by the nodes to the FC, which is located at a much farther distance.

VI. Conclusion

This paper has proposed centralized and distributed algorithms to jointly optimize the precoders, energy covariance matrices and MMSE fusion rule in MIMO WPSNs. Employing the ADMM technique, we have developed an iterative distributed algorithm that allows local computation of the precoders and the fusion rule at the sensor nodes, and the energy covariance matrices at the E-APs. Further, we have derived low complex closed form expressions to determine the optimal solutions. Finally, numerical simulations have validated that the proposed techniques perform superior to a conventional design. Also, the distributed algorithm exhibits almost the same performance as the centralized algorithm with much reduced computational complexity. It will be worthwhile to investigate the performance improvement with a transceiver design with both power and data exchange between the sensor nodes, and joint transceiver design and path planning for mobile data collection and energy transfer.

Appendix
Solution to Problem (18)

Introducing an auxiliary variable $\bar{\mathbf{f}} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{f}_i$, the optimization problem in (18) can be rewritten as

$$\begin{aligned} \min_{\{\mathbf{f}_i\},\bar{\mathbf{f}}} \; & L^2\left\|\bar{\mathbf{f}}\right\|^2 + \frac{\rho_u}{2}\sum_{j=1}^{L}\|\mathbf{D}_j\mathbf{w}_j^{(k+1)} - \mathbf{f}_j + \mathbf{u}_j^{(k)}\|^2 \\ \text{s. t.} \; & \frac{1}{L}\sum_{j=1}^{L}\mathbf{f}_j = \bar{\mathbf{f}}. \end{aligned} \tag{46}$$

For a fixed $\bar{\mathbf{f}}^{(k+1)}$ in the above problem, employing the Karush-Kuhn-Tucker (KKT) conditions, one can obtain a solution as

$$\mathbf{f}_i^{(k+1)} = \mathbf{D}_i\mathbf{w}_i^{(k+1)} + \mathbf{u}_i^{(k)} + \frac{\boldsymbol{\mu}_f}{\rho_u L}, \tag{47}$$

where $\boldsymbol{\mu}_f$ is the dual variable corresponding to the constraint in (46).

Now, inserting the solution in the constraint $\frac{1}{L}\sum_{j=1}^{L}\mathbf{f}_j^{(k+1)} = \bar{\mathbf{f}}^{(k+1)}$, the optimal dual variable is given by $\frac{\boldsymbol{\mu}_f}{\rho_u L} = \bar{\mathbf{f}}^{(k+1)} - \frac{1}{L}\sum_{i=1}^{L}(\mathbf{D}_i\mathbf{w}_i^{(k+1)} + \mathbf{u}_i^{(k)})$.

Substituting $\frac{\boldsymbol{\mu}_f}{\rho_u L}$, the auxiliary variables are expressed by

$$\mathbf{f}_i^{(k+1)} = \bar{\mathbf{f}}^{(k+1)} + \mathbf{D}_i \mathbf{w}_i^{(k+1)} + \mathbf{u}_i^{(k)} - \frac{1}{L} \sum_{j=1}^{L} (\mathbf{D}_j \mathbf{w}_j^{(k+1)} + \mathbf{u}_j^{(k)}). \tag{48}$$

Replacing $\mathbf{f}_i^{(k+1)}$ in (18) with (48) simplifies the problem to

$$\min_{\bar{\mathbf{f}}} L^2 \left\|\bar{\mathbf{f}}\right\|^2 + \frac{L\rho_u}{2} \sum_{i=1}^{L} \left\| \bar{\mathbf{f}} - \frac{1}{L} \sum_{j=1}^{L} \left(\mathbf{D}_j \mathbf{w}_j^{(k+1)} + \mathbf{u}_j^{(k)} \right) \right\|^2,$$

which is a quadratic minimization problem. Therefore, obtaining the gradient and equating it to zero, a closed form solution is determined as

$$\bar{\mathbf{f}}^{(k+1)} = \frac{\rho_u}{2L + \rho_u} \frac{1}{L} \sum_{j=1}^{L} \left(\mathbf{D}_j \mathbf{w}_j^{(k+1)} + \mathbf{u}_j^{(k)} \right).$$

REFERENCES

[1] C.-Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[2] 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.

[3] A. S. Behbahani, A. M. Eltawil, and H. Jafarkhani, "Linear decentralized estimation of correlated data for power constrained wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 16, no. 11, pp. 6003–6016, Nov. 2012.

[4] F. Jiang, J. Chen, and A. L. Swindlehurst, "Estimation in phase-shift and forward wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3840–3851, Aug. 2013.

[5] X. Guo, A. S. Leong, and S. Dey, "Distortion outage minimization in distributed estimation with estimation secrecy outage constraints," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 3, no. 1, pp. 12–28, Mar. 2017.

[6] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embedded Comput. Syst.*, vol. 6, no. 4, p. 38, Sep. 2007.

[7] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless charging technologies: Fundamentals, standards, and network applications," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1413–1452, 2nd Quart., 2016.

[8] L. Xie, Y. Shi, Y. T. Hou, and W. Lou, "Wireless power transfer and applications to sensor networks," *IEEE Wireless Commun. Mag.*, vol. 20, no. 4, pp. 140–145, Aug. 2013.

[9] Z. Ding *et al.*, "Application of smart antenna technologies in simultaneous wireless information and power transfer," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86–93, Apr. 2015.

[10] X. Lu, *Sensor Networks With Wireless Energy Harvesting*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[11] M. Nourian, S. Dey, and A. Ahlén, "Distortion minimization in multisensor estimation with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 524–539, Mar. 2015.

[12] Z. Wang, L. Duan, and R. Zhang, "Adaptively directional wireless power transfer for large-scale sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1785–1800, May 2016.

[13] S. Lee and R. Zhang, "Distributed wireless power transfer with energy feedback," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1685–1699, Apr. 2017.

[14] A. Thudugalage, S. Atapattu, and J. Evans, "Beamformer design for wireless energy transfer with fairness," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[15] L. Liu, R. Zhang, and K.-C. Chua, "Multi-antenna wireless powered communication with energy beamforming," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4349–4361, Dec. 2014.

[16] H. Lee, K.-J. Lee, H.-B. Kong, and I. Lee, "Sum-rate maximization for multiuser MIMO wireless powered communication networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 9420–9424, Nov. 2016.

[17] S. Guo, F. Wang, Y. Yang, and B. Xiao, "Energy-efficient cooperative transmission for simultaneous wireless information and power transfer in clustered wireless sensor networks," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4405–4417, Nov. 2015.

[18] J. Xu, Z. Zhong, and B. Ai, "Wireless powered sensor networks: Collaborative energy beamforming considering sensing and circuit power consumption," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 344–347, Aug. 2016.

[19] N. K. D. Venkategowda, H. Lee, and I. Lee, "Data precoding and energy transmission for parameter estimation in MIMO wireless powered sensor networks," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.

[20] G. Yang, W. P. Tay, and Y. L. Guan, "Optimal wireless power transfer and harvested power allocation for diffusion LMS in wireless sensor networks," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2015, pp. 1–5.

[21] V. V. Mai, W.-Y. Shin, and K. Ishibashi, "Wireless power transfer for distributed estimation in sensor networks," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 3, pp. 549–562, Apr. 2017.

[22] S. Knorn, S. Dey, A. Ahlén, and D. E. Quevedo, "Distortion minimization in multi-sensor estimation using energy harvesting and energy sharing," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2848–2863, Jun. 2015.

[23] Y.-W. P. Hong, T.-C. Hsu, and P. Chennakesavula, "Wireless power transfer for distributed estimation in wireless passive sensor networks," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5382–5395, Oct. 2016.

[24] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2082–2085, Dec. 2015.

[25] E. Boshkovska, D. W. K. Ng, N. Zlatanov, A. Koelpin, and R. Schober, "Robust resource allocation for MIMO wireless powered communication networks based on a non-linear EH model," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 1984–1999, May 2017.

[26] T. Le, K. Mayaram, and T. Fiez, "Efficient far-field radio frequency energy harvesting for passively powered sensor networks," *IEEE J. Solid-State Circuits*, vol. 43, no. 5, pp. 1287–1302, May 2008.

[27] K. B. Petersen and M. S. Pedersen. (Nov. 2012). *The Matrix Cookbook*. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?3274

[28] Y. Ye, *Interior Point Algorithms: Theory and Analysis*. New York, NY, USA: Wiley, 1997.

[29] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, "A unified algorithmic framework for block-structured optimization involving big data: With applications in machine learning and signal processing," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 57–77, Jan. 2016.

[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[31] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[33] T. Lin, S. Ma, and S. Zhang, "On the global linear convergence of the ADMM with multiblock variables," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1478–1497, 2015.

[34] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Hoboken, NJ, USA: Wiley, 2007.

[35] J. Kim, H. Lee, C. Song, T. Oh, and I. Lee, "Sum throughput maximization for multi-user MIMO cognitive wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 913–923, Feb. 2017.

[36] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[37] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan, "Throughput optimization for massive MIMO systems powered by wireless energy transfer," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1640–1650, Aug. 2015.

[38] 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.

Naveen K. D. Venkategowda (S'12–M'17) received the B.E. degree in electronics and communication engineering from Bangalore University, Bengaluru, India, in 2008, and the Ph.D. degree in electrical engineering from IIT Kanpur, Kanpur, India, in 2016. At IIT Kanpur, he was a Tutor, a Teaching Assistant, and a Research Associate. From 2009 to 2010, he was a Senior Engineer with Mercedes-Benz Research and Development India, India. From 2016 to 2017, he was with Korea University as a Research Professor. His research interests include statistical signal processing and its applications in wireless networks. He has received the TCS Research Fellowship for graduate studies from 2011 to 2015 and the ERCIM Alan Bensoussan Fellowship in 2017.

Hoon Lee (S'14–M'18) received the B.S. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2012 and 2017, respectively. In 2014, he visited Imperial College London, London, U.K., to conduct a collaborative research. In 2017, he was a Post-Doctoral Fellow with Korea University. In 2018, he joined the Singapore University of Technology and Design, Singapore, where he is currently a Post-Doctoral Fellow. His research interests include machine learning and signal processing for wireless communications, such as visible light communications, wireless energy transfer communication systems, and secure wireless networks.

Inkyu Lee (S'92–M'95–SM'01–F'16) received the B.S. degree (Hons.) in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1992 and 1995, respectively. From 1995 to 2001, he was a Member of Technical Staff with Bell Laboratories, Lucent Technologies, where he studied high-speed wireless system designs. From 2001 to 2002, he was with Agere Systems, Murray Hill, NJ, USA, as a Distinguished Member of Technical Staff. Since 2002, he has been with Korea University, Seoul, where he is currently a Professor with the School of Electrical Engineering. In 2009, he was a Visiting Professor with the University of Southern California, Los Angeles, CA, USA. He has authored over 150 journal papers in the IEEE. He has 30 U.S. patents granted or pending. His research interests include digital communications, signal processing, and coding techniques applied for next generation wireless systems. He was elected as a member of the National Academy of Engineering of Korea (NAEK) in 2015. He was a recipient of the IT Young Engineer Award at the IEEE/IEEK Joint Award in 2006 and the Best Paper Award at APCC in 2006, the IEEE VTC in 2009, and ISPACS in 2013. He was also a recipient of the Best Research Award from the Korea Information and Communications Society in 2011, the Best Young Engineer Award from NAEK in 2013, and the Korea Engineering Award from the National Research Foundation of Korea in 2017. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2001 to 2011 and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011. He was the Chief Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on 4G Wireless Systems in 2006. He currently serves as an Editor for the IEEE ACCESS. He is an IEEE Distinguished Lecturer.