Wireless Power Transfer for Distributed Estimation in Sensor Networks
Vien V. Mai, Won-Yong Shin, Senior Member, IEEE, and Koji Ishibashi, Member, IEEE

Abstract—This paper studies power allocation for distributed estimation of an unknown scalar random source in sensor networks with a multiple-antenna fusion center (FC), where wireless sensors are equipped with radio-frequency-based energy harvesting technology. The sensors’ observation is locally processed by using an uncoded amplify-and-forward scheme. The processed signals are then sent to the FC, and are coherently combined at the FC, at which the best linear unbiased estimator (BLUE) is adopted for reliable estimation. We aim to solve the following two power allocation problems: 1) minimizing distortion under various power constraints; and 2) minimizing total transmit power under distortion constraints, where the distortion is measured in terms of mean-squared error of the BLUE. Two iterative algorithms are developed to solve the nonconvex problems, which converge at least to a local optimum. In particular, the above algorithms are designed to jointly optimize the amplification coefficients, energy beamforming, and receive filtering. For each problem, a suboptimal design, a single-antenna FC scenario, and a common harvester deployment for collocated sensors, are also studied. Using the powerful semidefinite relaxation framework, our result is shown to be valid for any number of sensors, each with different noise power, and for an arbitrarily number of antennas at the FC.

Index Terms—Amplify-and-forwarding, best linear unbiased estimator (BLUE), distributed estimation, mean-squared error (MSE), wireless power transfer (WPT).

I. INTRODUCTION

DISTRIBUTED inference in wireless sensor networks (WSNs) has been extensively studied for applications such as environmental monitoring, weather forecasts, health care, and home automation (see, e.g., [1]–[7] and references therein). Sensors in WSNs are powered typically by batteries, and hence the network lifetime is highly limited. In practice, periodically replacing or recharging batteries may be hard or even impossible (due to the fact that sensors are located inside toxic environments, building structures, or human bodies [8]). Therefore, although there have been many efforts in power management policies, the network lifetime remains a performance bottleneck and limits the wide-range deployment of WSNs.

A. Previous Work

The optimal power allocation strategies for distributed estimation in WSNs have received a great research interest both from analog and digital encoding perspectives [3]–[5], [9]–[13]. Among encoding schemes, the uncoded amplify-and-forward scheme has been extensively studied due to its simplicity and information-theoretic-optimality properties under certain conditions [14]. In particular, the authors in [9] studied power allocation for orthogonal multiple access channels (MACs), when the best linear unbiased estimator (BLUE) is adopted. The same problem was considered in [3] for a coherent MAC. The effects of channel estimation error were reported in [10] for orthogonal MACs adopting a linear minimum mean-squared error estimator, while in [13], the sensing noise uncertainty was investigated by adopting the BLUE. Recently, the optimal transmit strategy for cooperative linear estimation was studied in [4].

The tremendous performance gains achieved by multiple-antenna techniques highly motivate us to integrate this technology into future wireless systems including WSNs. The benefits of such technology in the context of WSNs have been recently studied for distributed inference [5], [7], [15]–[18]. For a large-scale fusion center (FC) over a Rayleigh fading channel, it has been shown in [17] that the detection/estimation performance remains asymptotically constant if the transmit power at each sensor decreases proportionally with increasing number of antennas at the FC. The benefits of the multiple-antenna FC in distributed detection were analyzed in terms of asymptotic error exponents in [7]. Power allocation strategies for distributed estimation were studied for the correlated source case [18] and for the correlated noise case [5].

Although the network life span can be prolonged by applying the aforementioned strategies, one needs a disruptive design that fundamentally changes the limitation of a WSN. One of the promising solution is the so-called energy harvesting (EH), in which sensors scavenge energy from the ambient environment (e.g., solar, wind, and vibration) that can guarantee an infinite life span in theory [19]. However, due to the unpredictable nature of energy sources, EH is typically uncontrolled,

Manuscript received June 16, 2016; revised November 25, 2016; accepted February 25, 2017. Date of publication March 3, 2017; date of current version April 11, 2017. This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (2014R1A1A2054577) and in part by the Ministry of Science, ICT & Future Planning (2015R1A2A1A15054248). The guest editor coordinating the review of this paper and approving it for publication was Daniel P. Palomar.

V. V. Mai was with Dankook University, Yongin 448-701, South Korea. He is now with the Department of Automatic Control, KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden (e-mail: maiv@kth.se).

W.-Y. Shin is with the Department of Computer Science and Engineering, Dankook University, Yongin 448-701, South Korea (e-mail: wyshin@dankook.ac.kr).

K. Ishibashi is with the Advanced Wireless and Communication Research Center, The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: koji@ieee.org).

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/JSTSP.2017.2678106
and thus can be critical for some reliable-sensitive applications. In addition to commonly used energy sources such as solar and wind, ambient radio-frequency (RF) signals can be a viable new source for energy scavenging. Most of the researches on wireless power transfer (WPT) have been focused on cellular networks, where user terminals replenish energy from the received signals sent by the base station via the far-field RF-based WPT [8], [20]–[25]. For example, the fundamental trade-off between the achievable rate and the transferred power was characterized in [20]. Several practical receiver architectures for simultaneous information and power transfer were investigated in [8], [21]. Exploiting multiple antenna technologies in WPT has been widely studied: multiple-input-multiple-output broadcast channels [8], beamforming designs for multiuser multiple-input-single-output (MISO) [23], physical-layer security problems for multiuser MISO [24], and multiple-antenna interference channels [25]. On the other hand, there are a relatively limited number of studies on WPT for WSNs; different WPT technologies for addressing energy/lifetime problems in WSNs were reviewed in [26], [27]; in [28], the authors studied a distributed estimation system in which some of the multiple-antenna sensors, named super sensors, are capable of WPT to its neighboring sensors via beamforming; and in [29], several multiple-antenna RF-based chargers were used to replenish the wireless sensors and then to switch to the information transmission phase, where each sensor sent a quantized version of its measurement to the FC for estimation.

B. Main Contributions

For distributed estimation in WSNs, an important question is how to intelligently exploit multiple-antenna technologies and WPT to improve both the inference performance and network lifetime. In this paper, we devote to studying the optimality of WPT and the optimal allocation of harvested energy for distributed estimation of an unknown random source in WSNs with a multiple-antenna FC. Our main contributions are summarized as follows:

1) When the BLUE is adopted at the FC for estimation, we jointly optimize the amplification coefficients, energy beamforming, and receiver filtering by adopting alternative minimization methods (see Algorithms 1 and 2). To that end, we first solve the mean-squared error (MSE) minimization problem under the total power constraint at the FC as well as the causal power constraint at each sensor. Then, we solve a converse problem where the total transmit power at the FC is minimized subject to an MSE requirement.

2) A key ingredient of our algorithms is the so-called semidefinite relaxation. We show that such a relaxation does not sacrifice the optimality of the relaxed problems. We derive the properties of the optimal solutions (see Theorems 1 and 2).

3) A special deployment of WPT in WSNs is also discussed, where a common energy harvester is used to collect energy from the FC. We show that the optimization problems are significantly simplified in this case. The optimal power-distortion trade-off is also characterized (see Theorem 3).

C. Organization

The rest of the paper is organized as follows. The system model and problem formulation are described in Section II. Section III studies the problem of minimizing the MSE subject to power constraints. In Section IV, the converse problem in Section III is studied. The numerical results are shown in Section V. Finally, we conclude the paper in Section VI.

D. Notations

The operators $(\cdot)^\top$, $(\cdot)^*$, $(\cdot)^\dagger$ are the transpose, complex conjugate, and transpose conjugate, respectively. The notation $I_n$ denotes the $n \times n$ identity matrix; $\text{tr}(A)$ denotes the trace of a matrix $A$; rank$(A)$ denotes the rank of a matrix $A$; diag$(a)$ denotes a diagonal matrix with vector $a$ being its diagonal, $A \succeq 0$ denotes the positive semidefinite $A$; $E\{\cdot\}$ denotes the expectation operator; $\dim(A)$ denotes a dimension of the subspace $A$. We use the Bachmann–Landau notation: $f(x) = O(g(x))$ if $\lim_{x \to \infty} \frac{f(x)}{g(x)} = c < \infty$. Finally, we use the notation $[n]$ to denote the set of positive natural numbers up to $n$, i.e., $[n] = \{i : i = 1, 2, \ldots, n\}$.
time-switching harvest-then-forward protocol [30] in which for each $\tau T$ amount of time, where $T$ is the length of one time slot, the FC transmits its energy signal to the sensors, and for the remaining $(1-\tau)T$ amount of time, the sensors observe and forward their observations to the FC for estimation while using the harvested energy from the RF signal. For analytical convenience, we set $\tau = 1/2$ in the sequel unless otherwise specified.\(^2\)

In the first phase (i.e., the energy harvesting phase) of a time slot, the FC broadcasts its energy signal to the sensors through energy beamforming. More precisely, $\nu_k \leq \nu_s$ energy beams are assigned to $\nu_s$ sensors without loss of generality. The energy signal received at the $k$th sensor is then given by

$$t_k = g_k^T x_e + m_k = g_k^T \sum_{i=1}^{\nu_1} w_i s_i + m_k,$$

where $x_e = \sum_{i=1}^{\nu_s} w_i s_i$ is the energy signal transmitted from the FC; $s_i$ is the energy-carrying signal for the $i$th energy beam fulfilling $\mathbb{E}\{|s_i|^2\} = 1$ and $\mathbb{E}\{s_i s_j\} = 0$ for $i \neq j$, which can be any arbitrary random signal provided that its power spectral density satisfies certain regulations on microwave radiation [23]; $w_i \in \mathbb{C}^{\nu_1 \times 1}$ is the $i$th energy beamforming vector; $g_k \in \mathbb{C}^{\nu_1 \times 1}$ is the channel between the FC and $k$th sensor; and $m_k$ is the additive noise for the $k$th sensor. By ignoring the background noise for the sake of simplicity, the harvested energy at the $k$th sensor in each slot is given by [8]

$$E_k = \frac{\zeta_k}{2} T \sum_{i=1}^{\nu_s} |w_i|^2 g_k^T |g_k|^2,$$

where $0 \leq \zeta_k \leq 1$ is the energy harvesting efficiency at the $k$th sensor. Then, the average power $P_k$ available for the information transmission phase at the $k$th sensor can be expressed as

$$P_k = \frac{2(E_k - E_{\text{cir} k})}{T} = \zeta_k \sum_{i=1}^{\nu_s} |w_i|^2 g_k^T |g_k|^2 - 2E_{\text{cir} k}^2,$$

where $E_{\text{cir} k}^2 \geq 0$ is the circuit energy consumption at the $k$th sensor, which is assumed to be constant over time slots. Similarly as in [25], [30], we simply assume $\zeta_k = 1$ and unit slot duration in the rest of this work (note that using an arbitrary $\zeta_k$ does not fundamentally change our power allocation problems). Similarly as in [30], [34], for easy of presentation, we also assume that $E_{\text{cir} k}^2 = 0$ to focus on the transmit power of the sensors.\(^3\) The FC has a total transmit power constraint $P$; we thus have

$$\mathbb{E}\{|x_e^T x_e|\} \leq P.$$ \hspace{1cm} (4)

Now, let us turn to describing the second phase (i.e., the information transmission phase) of a time slot. The observation at the $k$th sensor can be expressed as

$$x_k = \theta + u_k, \hspace{1cm} k = 1, \ldots, \nu_s,$$

where $u_k$ is the additive noise at the $k$th sensor with variance $\sigma_{u,k}^2$. The noise at each sensor is assumed to be independent of each other. In this paper, we adopt an analog uncoded amplify-and-forward scheme, i.e., the $k$th sensor just simply amplifies its observation by a factor $\alpha_k$. Therefore, by stacking the transmit signals from all sensors into a single vector $t$, it can be expressed as

$$t = A1\theta + Au,$$

where $A = \text{diag}(\alpha_1, \ldots, \alpha_{\nu_s}) \in \mathbb{C}^{\nu_1 \times \nu_s}$ is the amplification matrix; $u = [u_1, u_2, \ldots, u_{\nu_s}]^T \in \mathbb{C}^{\nu_1 \times 1}$ is the noise vector at the sensors with zero mean and covariance matrix $R_u = \text{diag}(\sigma_{u,1}^2, \sigma_{u,2}^2, \ldots, \sigma_{u,\nu_s}^2)$; and $1$ is the all one vector. Then, the received signal $z \in \mathbb{C}^{\nu_1 \times 1}$ at the FC can be written as

$$z = H A1\theta + HAu + n,$$

where $H \in \mathbb{C}^{\nu_1 \times \nu_s}$ is the channel between the sensors and FC; and $n \in \mathbb{C}^{\nu_1 \times 1}$ is the noise vector at the FC with zero mean and covariance matrix $R_n = \text{diag}(\sigma_{n,1}^2, \sigma_{n,2}^2, \ldots, \sigma_{n,\nu_s}^2)$. Here, the random quantities $\theta$, $u$, and $n$ are statistically independent.

Since we consider a coherent MAC, we assume that there is perfect synchronization between the sensors and the FC. All wireless channels are assumed to be quasi-static flat fading, i.e., once each channel is realized, it remains fixed during each time slot and changes independently between slots. We further assume that full channel state information (CSI) is available at the FC. In practice, the sensors-to-FC channel $H$ can be estimated at the FC via periodic pilot transmissions from the sensors, while the FC-to-sensors channels $[g_k]_{k=1}^{\nu_s}$ can be acquired owing to channel reciprocity between the sensors-to-FC and FC-to-sensors channels when the system operates in time-division-duplex mode.

### B. Problem Formulation

The received signal $z$ is constructively combined at the FC by a filtering vector $v \in \mathbb{C}^{\nu_1 \times 1}$. Then, by adopting the well-known BLUE [35, Th. 6.1], the FC estimates the parameter $\theta$ based on the minimal sufficient statistic $y = v^T z$ as follows:\(^4\)

$$\hat{\theta} = \left[a^H \sigma_{\text{tot},1}^2 v^H H a\right]^{-1} a^H \sigma_{\text{tot},1}^2 v^H y,$$

where $\sigma_{\text{tot}}^2 = v^H [H A a] a^H + R_{\text{u}}$ is the total noise power after post-processing at the FC; and $a = [\alpha_1 \alpha_2 \ldots \alpha_{\nu_s}]^T$. The minimal sufficient statistic is defined in the sense that we no longer need the individual sample since all the information has been captured by the sufficient statistic [35].
MSE of the BLUE can be written as
\[
\text{mse} = E \left\{ |\theta - \hat{\theta}|^2 \right\} = \left[ a' H \sigma_{\text{tot}}^{-2} v' H a \right]^{-1}
\]
\[
= \left[ \frac{|v' H a|^2}{v' \left( H A R_s A' H^{'^2} + R_n \right) v} \right]^{-1}. \tag{9}
\]

Since the three quantities \( a, \{w_k\}_{k=1}^{n_h}, \) and \( v \) critically affect both the power requirement and estimation performance of the entire system, we jointly design the optimum sensor amplification coefficients \( a \), receive filtering vector \( v \), and energy beamforming \( \{w_k\}_{k=1}^{n_h} \) under practical constraints. To that end, we solve two types of minimization problems: 1) minimizing the MSE of the BLUE under causal individual power constraints at the sensors and a total power constraint at the FC; and 2) minimizing the total power consumed at the FC given a minimum requirement of the MSE. In particular, we aim to find the solution to the first problem, named (P1), by solving the following optimization problem.

\[\text{(P1)} : \]
\[
\text{maximize} \quad \frac{|v' H a|^2}{v' \left( H A R_s A' H^{'^2} + R_n \right) v}
\]
\[
\text{subject to} \quad |\alpha_k|^2 \left( \delta_{\hat{\theta}}^2 + \sigma_{u,k}^2 \right) \leq \sum_{i=1}^{n_h} |w_i'| g_k|^2, \quad \forall k \in [n_s]
\]
\[
\sum_{i=1}^{n_h} \|w_i\|^2 \leq P.
\]

As a counterpart of (P1), for a given MSE threshold \( \text{mse} = 1/\gamma \), the second optimization problem is stated as follows.

\[\text{(P2)} : \]
\[
\text{minimize} \quad \sum_{i=1}^{n_h} \|w_i\|^2
\]
\[
\text{subject to} \quad \frac{|v' H a|^2}{v' \left( H A R_s A' H^{'^2} + R_n \right) v} \geq \gamma,
\]
\[
|\alpha_k|^2 \left( \delta_{\hat{\theta}}^2 + \sigma_{u,k}^2 \right) \leq \sum_{i=1}^{n_h} |w_i'| g_k|^2, \quad \forall k \in [n_s].
\]

Note that closed-form solutions to the global optimization of these two problems are generally unknown. Indeed, both problems are non-convex due to the coupled amplification vector \( a \) and receive filtering \( v \). Therefore, we turn to a simple approach—alternative minimization—which guarantees convergence, at least to a local optimum.

### III. Minimizing MSE Under Power Constraints

In this section, we propose an alternative minimization algorithm to obtain the minimum solution to problem (P1). We also study the MSE performance for a large-scale antenna FC as well as a single-antenna FC.

#### A. Proposed Solution to Problem (P1)

Since problem (P1) is non-convex due to a non-concave objective function, we solve (P1) by using the alternative minimization method. Our goal is to progressively increase the objective function in (P1) by iteratively optimizing (P1) over \( a \) and \( \{w_i\}_{i=1}^{n_h} \) for given \( v \), and then over \( v \) for given \( a \). In order to find \( v \), we first fix \( a \) and solve the following unconstrained optimization problem:

\[
\text{maximize} \quad \frac{|v' H a|^2}{v' \left( H A R_s A' H^{'^2} + R_n \right) v}
\]
\[
\text{subject to} \quad v' H a = 1. \tag{11}
\]

Solving the above problem, we obtain

\[
v^* = \kappa \left( H A R_s A' H^{'^2} + R_n \right)^{-1} H a. \tag{12}
\]

Note that the value of \( \kappa \) is chosen to guarantee the equality constraint in (11). However, any selected value of \( \kappa \) will not affect the objective function in (P1), and thus we simply choose \( \kappa = 1 \) without loss of optimality. For a given \( v \) in (12), we are now ready to find an update of \( a \) and \( \{w_i\}_{i=1}^{n_h} \) in (P1). To facilitate the calculations, we define \( f = [v' h_1, v' h_2, \ldots v' h_{n_r}]^\top \) and \( F = \text{diag}(f) \), where \( h_i \) is the \( i \)-th column of the matrix \( H \). Then, for a fixed receive filtering \( v \), problem (P1) can be expressed as

\[
\text{maximize} \quad \frac{|a' f|^2}{a' R_F F' a + v' R_n v}
\]
\[
\text{subject to} \quad |\alpha_k|^2 \left( \delta_{\hat{\theta}}^2 + \sigma_{u,k}^2 \right) \leq \sum_{i=1}^{n_h} |w_i'| g_k|^2, \quad \forall k \in [n_s]
\]
\[
\sum_{i=1}^{n_h} \|w_i\|^2 \leq P. \tag{13}
\]

We remark that even with a fixed receive filtering \( v \), problem (P1) is still non-convex, and thus needs to be transformed to a simple form. We further introduce \( Q = a' a^{-1}, W = \sum_{i=1}^{n_h} w_i w_i^\top, \Sigma = f f^\top, \Psi = F R_F F' \), and \( G_k = g_k g_k^\top \), and \( D_k = \text{diag}(0, \ldots, \delta_{\hat{\theta}}^2 + \sigma_{u,k}^2, \ldots, 0) \). Then, we can rewrite the optimization problem (13) as

\[
\text{maximize} \quad \frac{\text{tr} \left( Q \Sigma \right)}{\text{tr} (Q \Psi) + v' R_n v}
\]
\[
\text{subject to} \quad \text{tr} \left( D_k Q \right) - \text{tr} (G_k W) \leq 0, \quad \forall k \in [n_s]
\]
\[
\text{tr} \left( W \right) \leq P
\]
\[
W \succeq 0, Q \succeq 0
\]
\[
\text{rank} (Q) = 1. \tag{14}
\]

In (14), if there exist a rank one solution of the optimal \( Q = Q^* \) and a rank \( n_h \) solution of the optimal \( W = W^* \), then one can recover the optimal \( a^* \) and \( \{w_i^*\}_{i=1}^{n_h} \) by taking the eigenvalue
decomposition of the matrices $Q'$ and $W'$, respectively. Note that problem (14) is non-convex due to the linear fractional structure of its objective function. However, we can use the Charnes-Cooper transformation [36] to reformulate the quasi-convex objective function in (14) into a simpler form as follows:\footnote{Here, we use the transformations $\eta^{-1} = \text{tr}((Q\Psi) + v^H R_v v)$, $\bar{Q} = \eta Q$, and $\bar{W} = \eta W$.}

$$\begin{align*}
\text{maximize} & \quad \text{tr} (Q\Sigma) \\
\text{subject to} & \quad \text{tr} (Q\Psi) + \eta v^H R_v v = 1 \\
& \quad \text{tr} (D_k \bar{Q}) - \text{tr} (G_k \bar{W}) \leq 0, \quad \forall k \in [n_s] \\
& \quad \text{tr} (\bar{W}) \leq \eta P \\
& \quad \bar{W} \succeq 0, \quad \bar{\eta} \geq 0, \quad \eta > 0 \\
& \quad \text{rank} (\bar{Q}) = 1.
\end{align*}$$

(15)

Note that $\eta = 0$ is not feasible because from the third constraint, we must have $\bar{W} = 0$ if $\eta = 0$. Thus, from the second constraint for any $k$, it follows that $\bar{Q} = 0$, which however violates the first constraint in (15).

Remark 1 (The Equivalence of Problems (14) and (15)): If $(Q', W', \eta')$ is the optimal solution to problem (15), then $(Q'/\eta', W'/\eta')$ is feasible to problem (14) and achieves the same objective value as that of problem (15). On the other hand, let $t^* = \text{tr}(Q'\Psi) + v^H R_v v$. Then, if $(Q', W')$ is the optimal solution to problem (14), then $(Q'/t^*, W'/t^*, 1/t^*)$ is feasible to problem (15) and achieves the same objective value as that of problem (14). This implies that the Charnes-Cooper transform is a one-to-one mapping between the feasible sets of problems (14) and (15). We can thus obtain the optimal solution to problem (14) by solving problem (15), which has a simpler form in the sense that the non-convexity of the objective function in problem (14) is eliminated.

Note that problem (15) is still non-convex due to the rank constraint, which makes problem (15) intractable in general. Hence, we will solve a relaxed version of (15) by ignoring the rank constraint on $Q$, which leads to the semidefinite relaxation (SDR) of problem (15).

$$(\text{SDR1}):$$

$$\begin{align*}
\text{maximize} & \quad \text{tr} (Q\Sigma) \\
\text{subject to} & \quad \text{tr} (Q\Psi) + \eta v^H R_v v = 1 \\
& \quad \text{tr} (D_k \bar{Q}) - \text{tr} (G_k \bar{W}) \leq 0, \quad \forall k \in [n_s] \\
& \quad \text{tr} (\bar{W}) \leq \eta P \\
& \quad \bar{W} \succeq 0, \quad \bar{\eta} \geq 0, \quad \eta > 0.
\end{align*}$$

The relaxed problem (SDR1) is now convex—indeed semidefinite program (SDP)—whose optimal solution can be found, for example, by using the interior-point method (e.g., CVX [37]). The following theorem characterizes the properties of the optimal solution to problem (SDR1).

Theorem 1 (Properties of Optimal Solution): Let $\nu^*$ and $\beta^*$ be the optimal dual solutions associated with the first and third constraint in (SDR1), respectively. We also let $Q'$ and $W'$ be the optimal primal solutions to problem (SDR1). Then, the following three properties are fulfilled:

1. $\nu^* > 0$, $\beta^* > 0$;
2. $\text{rank}(W') \leq \min(n_s, n_r)$;
3. $\text{rank}(Q') = 1$.

Proof: See Appendix A.

Remark 2: The condition $\beta^* > 0$ implies that the total power constraint at the FC must be satisfied with equality, while property 2) implies that at most $n_v = \min(n_s, n_r)$ energy beams are required for the optimal solution of problem (SDR1). It is worth noting that for fixed $v$, at the optimal solution $(\bar{Q}', \bar{W}', \eta')$, the individual power constraints in (SDR1) are not necessarily all tight, i.e., there may exist some $k$ such that $\text{tr}(D_k \bar{Q}) - \text{tr}(G_k \bar{W}) < 0$. This fact reveals that the sensors do not always transmit all the power budget harvested from the energy harvesting phase, but power control is required to guarantee the MSE optimality. A similar observation was made in throughput optimization for multiple-antenna multiuser cellular systems in [30].

Remark 3 (The Equivalence of Problems (15) and (SDR1)): We remark that since problem (SDR1) is a relaxed version of problem (15), in general, the solution to problem (SDR1) provides an upper bound on the optimal solution to problem (15), or equivalently, an upper bound on problem (P1) for a given $v$. Fortunately, we can show that the optimal solution to (SDR1) is also optimal to (15). To do that, let $\Phi_\eta(Q, W)$ be the objective function of problem (15) or (SDR1) for a given feasible $\eta$, and $(Q', W')$ and $(Q, W)$ be the optimal solutions to problems (SDR1) and (15), respectively. Since the optimization problem (SDR1) is a relaxation of problem (15), we must have

$$\Phi_\eta(Q', W') \geq \Phi_\eta(Q, W).$$

(16)

On the other hand, since $\text{rank}(Q') = 1$, the solution $(Q', W')$ is also a feasible solution to problem (15). Therefore, we have

$$\Phi_\eta(Q', W') \leq \Phi_\eta(Q, W).$$

(17)

From (16) and (17), it follows that $\Phi_\eta(Q', W') = \Phi_\eta(Q, W)$. In other words, $(Q', W')$ is also optimal solution to problem (15). Note that the above equivalence holds for any feasible $\eta$, and hence it holds for the optimal $\eta$.

Remark 3 suggests that we can solve the original problem (P1) for a given $v$ by equivalently solving the relaxed problem (SDR1) without loss of optimality. Finally, we summarize the overall procedure for solving problem (P1) in Algorithm 1 below. In this algorithm, the FC iteratively updates $v, a$ and $\{w_i\}_{i=1}^{n_v}$ in Step 3 and 4, respectively. The convergence and complexity of Algorithm 1 are analyzed in the following remark.

Remark 4 (Convergence and Complexity): Note that the objective function in (P1) is increased in each step of Algorithm 1. Moreover, the objective function is upper-bounded by a certain value due to the finite total power at the FC, which implies that
Algorithm 1: Proposed Algorithm to Solve (P1).

1: Initialization: set $n := 0$, and generate $\mathbf{a}^{(0)}$ and $\mathbf{A}^{(0)}$.
2: repeat
3: \hspace{2mm} $\mathbf{v}^{(n)} = (\mathbf{H} \mathbf{A}^{(n)} \mathbf{R}_s \mathbf{A}^{(n)}\mathbf{H} + \mathbf{R}_n)^{-1} \mathbf{H} \mathbf{a}^{(n)}$
4: \hspace{2mm} Solve problem (SDR1) with $\mathbf{v} = \mathbf{v}^{(n)}$ to obtain the optimal solution $(\mathbf{Q}^{*}, \mathbf{W}^{*}, \eta^{*})$
5: \hspace{2mm} Set $(\mathbf{Q}^{(n+1)}, \mathbf{W}^{(n+1)}, \eta^{(n+1)}) := (\mathbf{Q}^{*}, \mathbf{W}^{*}, \eta^{*})$
6: \hspace{2mm} Construct $(\mathbf{a}^{(n+1)}, \mathbf{A}^{(n+1)})$ from $\mathbf{Q}^{(n+1)}/\eta^{(n+1)}$
7: Update $n := n + 1$
8: until convergence
9: Output: $(\mathbf{Q} = \mathbf{Q}^{(n)}/\eta^{(n)}, \mathbf{W} = \mathbf{W}^{(n)}/\eta^{(n)}, \mathbf{v} = \mathbf{v}^{(n)})$

Therefore, it suffices to prove that

\[
a' \mathbf{H} (\mathbf{H} \mathbf{R}_s \mathbf{A}' \mathbf{H} + \mathbf{R}_n)^{-1} \mathbf{H} \mathbf{a} \xrightarrow{\text{as}} \mathbf{1}' \mathbf{R}_s^{-1} \mathbf{1}
\]  

(20)

as $n_r$ tends to infinity. Using the matrix inversion lemma, we can show that

\[
(\mathbf{H} \mathbf{R}_s \mathbf{A}' \mathbf{H} + \mathbf{R}_n)^{-1} = \mathbf{R}_n^{-1} - \mathbf{R}_n^{-1} \mathbf{H} (\mathbf{K}^{-1} + \mathbf{H}' \mathbf{R}_n^{-1} \mathbf{H})^{-1} \mathbf{H}' \mathbf{R}_n^{-1},
\]  

(21)

where $\mathbf{K} = \mathbf{A} \mathbf{R}_s \mathbf{A}'$. Substituting (21) into (19), we obtain

\[
\text{mse}^{-1} = a' \mathbf{H} \mathbf{R}_n^{-1} \mathbf{H} a - a' \mathbf{H} \mathbf{R}_n^{-1} \mathbf{H} (\mathbf{K}^{-1} + \mathbf{H}' \mathbf{R}_n^{-1} \mathbf{H})^{-1} \mathbf{H}' \mathbf{R}_n^{-1} \mathbf{H} a.
\]  

(22)

Note that as $n_r \to \infty$, we have [42]

\[
\frac{1}{n_r} \mathbf{H} \mathbf{R}_n^{-1} \mathbf{H} \xrightarrow{\text{as}} \mathbf{R}_n^{-1}.
\]  

(23)

Using this identity, we obtain

\[
\frac{\text{mse}^{-1}}{n_r} \xrightarrow{\text{as}} a' \mathbf{R}_n^{-1} a - a' \mathbf{R}_n^{-1} \left( \frac{\mathbf{K}^{-1} + \mathbf{R}_n^{-1}}{n_r} \right)^{-1} \mathbf{R}_n^{-1} a
\]

\[= a' \mathbf{R}_n^{-1} \left( \mathbf{I} - \left( \frac{\mathbf{R}_s \mathbf{K}^{-1} + \mathbf{I}}{n_r} \right) \right)^{-1} a
\]

\[= \frac{1}{n_r} a' \left( \mathbf{K} + \frac{\mathbf{R}_s}{n_r} \right)^{-1} a.
\]  

(24)

where the second equality follows from the matrix inversion lemma. From (24) and the definitions of the matrices $\mathbf{K}$ and $\mathbf{A}$, as $n_r \to \infty$, we finally have

\[
\text{mse} \xrightarrow{\text{as}} \left[ \mathbf{1}' \mathbf{R}_s^{-1} \mathbf{1} \right]^{-1}
\]  

(25)

which concludes the proof of the proposition.

Proposition 1 implies that as the number of antennas grows large, the effects of fading and noise at the FC disappear, and hence the performance benchmark is determined by the sensing quality. From (18), if the sensing noise at the sensors is equal to $\mathbf{R}_s = \alpha_n^2 \mathbf{I}_{n_s}$, then it follows that $\left[ \mathbf{1}' \mathbf{R}_s^{-1} \mathbf{1} \right]^{-1} = \frac{\alpha_n^2}{n_s}$. This means that the MSE linearly decreases according to $n_s$.

C. Single-Antenna FC

It is of importance to study the single-antenna FC scenario separately not only for comparison but also because the problem is remarkably simplified. Specifically, for a single-antenna FC, the design of energy beamforming and receive filtering is neglected, and thus we aim to simply find the optimal amplification coefficients $\mathbf{a}$ that minimize the MSE of the BLUE. During the energy transmission phase, we assume that the FC transmits an energy signal $\mathbf{s}$ such that $\mathbb{E} \{ ||\mathbf{s}||^2 \} = P$. In this case, the harvested energy at the $k$th sensor is given by

\[
E_k = \frac{P \mathbf{g}_k^2}{2}.
\]  

(26)
The MSE of the BLUE is boiled down to
\[
\text{mse} = \left( \frac{a^\top hh^\top a^*}{a^\top FR_aF^\top a^* + \sigma_n^2} \right)^{-1},
\tag{27}
\]
where \( h^\top \in \mathbb{C}^{1 \times n_s} \) is the channel between the sensors and the FC; \( F = \text{diag}(h) \); and \( n \in \mathbb{C} \) is the additive noise at the FC. Given the MSE in (27), we aim to solve the following problem
\[
\text{maximize } \begin{array}{c}
\frac{a^\top hh^\top a^*}{a^\top FR_aF^\top a^* + \sigma_n^2}
\end{array}
\]
subject to \( |\alpha_\ell| \left( \delta_\ell^2 + \sigma_n^2 \right) \leq P \|g_\ell\|^2 \), \( \forall \ell \in [n_s] \).
\tag{28}

The above problem—quadratically constrained ratio of quadratic functions (QCRO)—has been studied for parameter tracking using the Kalman filter at the FC [6], where the optimal solution is given by
\[
a^* = \frac{1}{\sqrt{(P^*)_{n_s+1,n_s+1}}} \bar{a}^*.
\tag{29}
\]

Here, \( (P^*)_{i,j} \) is the \((i,j)\)th element of the matrix \( P^* \); \( \bar{a} \) is the vector satisfying \( \bar{a}^\top \bar{a} = P^*_{n_s,n_s} \); \( P^*_{n_s,n_s} \) is the \(n_s\)th order leading principal submatrix of \( P^* \) obtained by excluding the \((n_s+1)\)th row and column; and \( P^* = \mathbb{C}^{(n_s+1) \times (n_s+1)} \) is the optimal solution to the following problem
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top X a^*}
\end{array}
\]
subject to \( a^\top X a^* \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to
\[
\text{maximize } \begin{array}{c}
\frac{|a^\top f|^2}{a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v}}
\end{array}
\]
subject to \( a^\top FR_aF^\top a^* + \bar{v}^\top R_n \bar{v} \leq \frac{1}{\alpha} Y a^* \), \( \alpha > 0 \), \( \frac{1}{\alpha} Y a^* \) and \( f = (f_1, f_2, \ldots, f_{n_s})^\top \) and \( F = \text{diag}(f) \). The problem is equivalent to

\[\text{IV. MINIMIZING POWER UNDER AN MSE CONSTRAINT}\]

In this section, we study the power minimization for distributed estimation with an MSE constraint.

\[\text{A. Proposed Solution to Problem (P2)}\]

Similarly as in problem (P1), we adopt an alternative minimization method to iteratively solve problem (P2). Specifically, we first solve problem (P2) over \( \bar{v} \) for given \( a \) by finding a solution to the following feasibility problem:
\[
\text{minimize } \begin{array}{c}
0
\end{array}
\]
subject to \( \frac{|\bar{v}|^2}{\bar{v}^\top \left( H A R_a A^\top + R_n \right) \bar{v}} \geq \gamma. \tag{35}\n\]

Since the left-hand side (LHS) of the constraint in (35) is increasing with the norm of \( a \), one should choose \( \bar{v} \) such that the LHS term is as large as possible. Hence, problem (35) can be
rewritten as an unconstrained optimization problem as follows:

$$\max_v \frac{|v' H a|^2}{v' (H A R_n A' H^\dagger + R_n) v}.$$  (36)

Solving the above problem, we obtain

$$v^+ = (H A R_n A' H^\dagger + R_n)^{-1} H a.$$  (37)

For fixed $v$ given in (37), we now solve problem (P2) over $a$ and \{\$w_i\}_i=1^{n_s}$ as in the following:

$$\min_{a, \{w_i\}_i=1^{n_s}} \sum_{i=1}^{n_s} \|w_i\|^2$$

subject to

$$|a^\dagger f|^2 \frac{a^\dagger F R_n F' a^\dagger + v' R_n v}{a^\dagger F R_n F' a^\dagger + v' R_n v} \geq \gamma,$$

$$|\alpha_k|^2 (\delta_k^2 + \sigma^2_{n,k}) \leq \sum_{i=1}^{n_s} \|w_i g_k\|^2, \forall k \in [n_s],$$  \hspace{1cm} (38)

where $f = [v' h_1, v' h_2, \cdots, v' h_{n_s}]^\dagger$; where $h_i$ is the $i$th column of the matrix $H$; and $F = \text{diag}(f)$. We remark that for a fixed $v$, the MSE constraint (i.e., the first constraint) at the optimal solution to problem (38) must be fulfilled with equality. We prove it by contradiction. Assume that the MSE constraint is satisfied with a strict inequality at the optimal solution $(a^\star, \{w_i^\star\}_i=1^{n_s})$. By letting $\tilde{a} = t a^\star$ for $0 < t < 1$, we can choose a sufficient large $t$ such that

$$|a^\dagger f|^2 \frac{a^\dagger F R_n F' a^\dagger + v' R_n v}{a^\dagger F R_n F' a^\dagger + v' R_n v} > \gamma.$$  \hspace{1cm} (39)

When $\tilde{w}_i = t w_i^\star, (\tilde{a}, \{\tilde{w}_i\}_i=1^{n_s})$ can also be a feasible solution to problem (38) with the new objective value $t^2 \sum_{i=1}^{n_s} \|w_i^\star\|^2$, which is definitely smaller than the original objective value when the optimal solution is $(a^\star, \{w_i^\star\}_i=1^{n_s})$. This contradicts to the assumption that $(a^\star, \{w_i^\star\}_i=1^{n_s})$ is optimal. Therefore, the MSE constraint must hold with equality. Let $Q = a^\dagger a^\star \geq 0$, $W = \sum_{i=1}^{n_s} w_i w_i^\dagger$, $f = \sum_{i=1}^{n_s} w_i f_i$, $\Psi = F R_n F'$, $G_k = g_k g_k^\dagger$, and $D_k = \text{diag}(0, \ldots, \delta_k^2 + \sigma^2_{n,k}, \ldots, 0)$. Then, as in problem (15), we will omit the rank constraint on $Q$ and solve a relaxed version of (38), which leads to

**Algorithm 2: Proposed Algorithm to Solve (P2).**

1. **Initialization:** set $n := 0$, and generate $a^{(0)}$ and $A^{(0)}$.
2. **repeat**
3. $v^{(n)} = (H A R_n A' H^\dagger + R_n)^{-1} H a^{(n)}$
4. Solve problem (SDR2) with $v = v^{(n)}$ to obtain the optimal value $(Q^*, W^*)$.
5. Set $(Q^{(n+1)}, W^{(n+1)}) := (Q^*, W^*)$.
6. Construct $(a^{(n+1)}, A^{(n+1)})$ from $Q^{(n+1)}$.
7. Update $n := n + 1$.
8. **until convergence**
9. **Output:** $(Q = Q^{(n)}, W = W^{(n)}, v = v^{(n)})$

The following theorem characterizes the properties of the optimal solution to problem (SDR2).

**Theorem 2 (Properties of Optimal Solution):** Let $\beta^\star$ be the dual optimal solutions associated with the equality constraint in (SDR2). We also let $Q^*$ and $W^*$ be the primal optimal solutions to (SDR2). Then the following three properties are fulfilled:

1. $\beta^\star > 0$;
2. $\text{rank}(W^*) \leq \text{min}(n_s, n_r)$;
3. $\text{rank}(Q^*) = 1$.

**Proof:** The proof can be found using the similar steps to the proof for Theorem 1.

Similarly as in Section III, we summarize the overall procedure for solving problem (P2) in Algorithm 2. In this algorithm, the objective value is monotonically reduced in each step, and for a given feasible threshold $\gamma$, it is lower-bounded by a certain value. As a result, the algorithm converges at least to a local optimum. Finally, it can be verified that the computational complexity of Algorithm 2 is same as that of Algorithm 1.

**B. Single-Antenna FC**

It would also be of interest to study problem (P2) for the single-antenna FC scenario. In this case, problem (P2) can be rewritten as

$$\min_{a, b} P$$

subject to

$$a^\dagger h h^\dagger a^\star + \sigma_n^2 = \gamma,$$

$$|\alpha_k|^2 (\delta_k^2 + \sigma^2_{n,k}) \leq P |g_k|^2, \forall k \in [n_s].$$  \hspace{1cm} (40)

Define the matrices $\Omega = (\ast a' \ast_0); P = (0_0 01)_D k = (\ast \ast_0 0 -1); E = h h^\dagger - \gamma F R_n F'$, Then, problem (40) can be recast as

$$\min_{\Omega} \text{tr}(\Omega)$$

subject to

$$\text{tr}(\Omega E) = \gamma \sigma_n^2$$

$$\text{tr}(\Omega D_k) \leq 0, \forall k \in [n_s]$$

$$\text{rank}(\Omega) = 1.$$  \hspace{1cm} (41)

By dropping the rank constraint on $\Omega$, problem (41) is a SDP and thus can be solved efficiently. If we denote $\Omega^\star$ by the optimal solution to the relaxed problem of (41), then $\text{rank}(\Omega^\star) = 1$ and
the optimal \( \mathbf{a} \) and \( P \) can be found from \( \Omega^* \). Particularly,
\[
P^* = (\Omega^*)^1_{n_1+1,n_2+1} \tag{42}
\]
\[
\mathbf{a}^* = \sqrt{\text{tr}(\Omega^*_0)} \mathbf{u}_1^1, \tag{43}
\]
where \( \Omega^*_0 \) is the \( n_k \times n_k \) order leading principal submatrix of \( \Omega^* \) obtained by excluding the \( (n_k + 1) \)th row and column and \( \mathbf{u}_1 \) is the eigenvector associated with the largest eigenvalue of \( \Omega^*_0 \). Similarly as in problem (P1), in this case, the optimal solution \((P^*, \mathbf{a}^*)\) is indeed a global optimum.

C. A Common Energy Harvester

Now, we consider the converse problem of (P1–Sum), in which we aim to minimize the transmit power at the FC subject to a minimum requirement of the MSE performance,
\[
\text{(P2–Sum)}: \quad \min_{\mathbf{v}, \mathbf{a}, \mathbf{w}} \quad \|\mathbf{w}\|^2
\]
subject to
\[
\mathbf{v}^\dagger \mathbf{H} \mathbf{a}^2 \geq \gamma 
ad^\dagger \mathbf{D} \mathbf{a} \leq \|\mathbf{w}\|^2. \tag{46}
\]
If we multiply \( \mathbf{w} \) and \( \mathbf{a} \) by a scalar \( \alpha > 1 \) and \( \beta < 1 \), respectively, then the left-hand side of the MSE constraint (i.e., the first constraint) is strictly increased while the right-hand side of the sum power constraint (i.e., the second constraint) as well as the objective function are strictly decreased. Thus, the optimality for (P2–Sum) is achieved when all the above constraints are satisfied with equality. Problem (P2–Sum) can be formulated as a SDP, and hence solved efficiently by CVX. In the following, we establish a fundamental relationship between two problems (P1–Sum) and (P2–Sum).

Theorem 3 (Power–Distortion Trade-off): For a distributed estimation system using the BLUE with a common energy harvester, if we assume that the alternative algorithms solving (P1–Sum) and (P2–Sum) are initialized with \( \mathbf{a}^{(0)} \), then the optimal power–distortion trade-off is given by
\[
\frac{1}{\text{mse}} = \mathbf{f}^\dagger \left( \mathbf{F} \mathbf{R} \mathbf{F}^\dagger + \frac{\mathbf{v}^\dagger \mathbf{R}_n \mathbf{v}}{\mathbf{P} \|\mathbf{h}_n\|^2} \mathbf{D} \right)^{-1} \mathbf{f}. \tag{44}
\]

Proof: See Appendix B.

Theorem 3 is important since it enables to (numerically) find the power–distortion trade-off for distributed estimation in the cumulative power constraint case.

V. Numerical Results

In this section, we provide numerical examples by evaluating our proposed algorithms in Sections III and IV. In the simulations, we consider the widely used 915 MHz frequency band in WSNs [45] for both energy and information transmissions. For energy transmission, we consider the use of both the commercially available power transmitter (Powercast TX91501) with transmit power \( P = 1 \) W (30 dBm) and the RF power harvester (Powercast P2110). The detailed system parameters are summarized in Table I. To model a small-scale fading, we assume that the elements of the channel matrices are drawn independently from the Gaussian distribution with zero mean and unit variance. To further evaluate the effectiveness of the proposed algorithms, we also perform comparisons to low-complexity baseline schemes specified below.

A. Baseline Schemes

1) Suboptimal Design for (P1): We divide the optimization procedure into two phases. In the first phase, the energy beamforming vectors \( \{\mathbf{w}_i\}_{i=1}^{n_1} \) are designed such that the total harvested energy is maximized, which leads to the following maximization problem:
\[
\max_{\mathbf{w}_i} \sum_{k=1}^{n_1} \beta_k \left( \sum_{i=1}^{n_1} |\mathbf{w}_i^\dagger \mathbf{g}_k|^2 \right)
\]
subject to \( \sum_{i=1}^{n_1} \|\mathbf{w}_i\|^2 \leq P. \tag{45}\n\]
Here, \( \{\beta_k\}_{k=1}^{n_2} \) denote the energy weights indicating the priority (e.g., sensors with weaker channels can be assigned to a higher weight to guarantee fairness) of the corresponding sensors. It has been shown in [8] that the optimal strategy is to allocate all the power budget to the direction of \( \mathbf{\eta} \)—the eigenvector associated with the largest eigenvalue of the matrix \( \sum_{k=1}^{n_2} \beta_k \mathbf{g}_k \mathbf{g}_k^\dagger \). The optimal value in problem (45) is achieved when \( \mathbf{w}_i^\dagger = \sqrt{\mathbf{\eta}} \) with \( p_i \geq 0 \) such that \( \sum_{i=1}^{n_1} p_i = P \).

In the second phase, we find the amplification vector \( \mathbf{a} \) and the receive filtering \( \mathbf{v} \) in terms of minimizing the MSE subject to the energy harvested in the first phase. In particular, we solve the following problem:
\[
\max_{\mathbf{w}, \mathbf{a}} \quad \|\mathbf{a}^\dagger \mathbf{f}\|^2
\]
subject to \( \|\mathbf{a}^\dagger \mathbf{F} \mathbf{R}_n \mathbf{F}^\dagger \mathbf{a} + \mathbf{w}^\dagger \mathbf{R}_n \mathbf{v} \|^2 \leq P \), \( \forall k \in [n_2] \), \( \|\mathbf{h}_n\|^2 \)
\[
\]  where \( \mathbf{f} = [\mathbf{w}_1^\dagger \mathbf{v}_1^\dagger \ldots \mathbf{w}_{n_1}^\dagger \mathbf{v}_{n_1}^\dagger]^\dagger \); \( \mathbf{F} = \text{diag}(\mathbf{f}) \); and \( P_k = |\mathbf{g}_i^\dagger \mathbf{w}_i|^2 \). Problem (46) corresponds to problem (P1) without the total power constraint and can be solved by iteratively updating \( \mathbf{v} \) and \( \mathbf{a} \).

2) Suboptimal Design for (P2): To reduce the computational burden of the joint optimization for (P2), we propose a suboptimal design, in which the optimization procedure is divided into two phases. In the first phase, we aim to solve the
following problem:

\[
\begin{align*}
\text{minimize} & \quad a^\top Da \\
\text{subject to} & \quad \frac{|v^\top H a|^2}{v^\top \left(H A R a A^\top H^\top + R_n\right) v} \geq \gamma.
\end{align*}
\]  
(47)

We note that the objective function in problem (47) is the total transmit power of the sensors. Since the receive filtering \(v\) appears only in the constraint, we can iteratively solve problem (47) for \(a\) and \(v\). Since the left-hand side of the constraint is non-decreasing with the norm of \(a\), the constraint must be satisfied with equality. For a fixed \(v\), the above problem can be expressed as follows:

\[
\begin{align*}
\text{minimize} & \quad a^\top D a^* \\
\text{subject to} & \quad a^\top E a^* = \gamma v^\top R_n v,
\end{align*}
\]  
(48)

where \(E = \frac{f f^\top - \gamma F R a A^\top H^\top}{\gamma}. \) To guarantee the feasibility of problem (48), the value of \(\gamma\) must be chosen such that \(|a|^2 \geq \gamma \alpha^\top F R a A^\top a^*\). Since the quantities \(a^\top D a^* \geq 0\) and \(a^\top E a^* \geq 0\) are positive, problem (48) can be rewritten as

\[
\begin{align*}
\text{maximize} & \quad a^\top E a^* \\
\text{subject to} & \quad a^\top D a^* \\
& \quad = a^\top E a^* = \gamma v^\top R_n v,
\end{align*}
\]  
(49)

which is a Rayleigh quotient. Thus, the optimal solution to problem (49) is given by

\[
a^* = \sqrt{\frac{\gamma v^\top R_n v}{u_1^\top D^{1/2} E D^{-1/2} u_1}},
\]  
(50)

where \(u_1\) denotes the unit-norm eigenvector associated with the largest eigenvalue of the matrix \(D^{1/2} E D^{-1/2}\). It follows that the minimum total transmit power of the sensors, \(P^*_s\), required to achieve the MSE of \(1/\gamma\) is given by

\[
P^*_s = \frac{\gamma v^\top R_n v}{\lambda_{\max}(D^{1/2} E D^{-1/2})^2}.
\]  
(51)

In the second phase, we aim to minimize the total transmit power at the FC with the amplification coefficients \(\{\alpha_k\}_{k=1}^n\) that are the solutions to problem (47). In other words, we find the optimal solution to the following minimization problem:

\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{n_t} \|w_i\|^2 \\
\text{subject to} & \quad \sum_{i=1}^{n_t} |w_i g_i|^2 \geq |\alpha_k|^2 \left(\delta^2 + \alpha^2_{\alpha,k}\right), \quad \forall k \in [n_s].
\end{align*}
\]  
(52)

Problem (52) can be effectively solved by CVX. In the following subsections, we use these suboptimal designs as the baseline schemes to assess the effectiveness of our proposed algorithms.

**B. MSE Minimization**

Fig. 2 shows the average MSE for distributed estimation versus iteration index when \(n_s = 5, R_s = 0.1 I_{n_s}, P = 30\) dBm, \(\delta = 1\), and \(n_r = 5, 10, 15, 20\). One can see that the average MSE monotonically decreases while the algorithm converges within a few iterations. It can be obviously seen that the MSE performance is improved with the increasing number of antennas at the FC, \(n_r\). In this figure, we also plot a benchmark ideal case for distributed estimation, where all the observations at the sensors are assumed to be directly available at the FC, which will give a lower bound on the MSE performance. One can see that the average MSE evaluated via our simulation tends to approach the benchmark value, \([1/R_n^2]^{-1}\), as \(n_r\) increases. In Fig. 3, the average MSE for distributed estimation is shown as a function of \(n_s\) for the optimal and suboptimal solutions when \(P = 30\) dBm, \(n_r = 5, R_s = 0.1 I_{n_s}, R_n = 0.5 I_{n_s}, \) and \(\delta = 1\).
TABLE II

| Sensor index | Harvested power [dBm] | Transmit power [dBm] |
|--------------|-----------------------|----------------------|
| 1            | −31.449               | −31.449              |
| 2            | −27.687               | −34.347              |
| 3            | −30.737               | −30.737              |
| 4            | −32.865               | −32.865              |
| 5            | −13.847               | −48.067              |
| 6            | −29.886               | −31.999              |
| 7            | −28.307               | −32.964              |

As expected, the MSE performance is improved as \( n_s \) increases. In this example, we can see that the suboptimal solution shows a reasonable performance compared to the optimal one.

In Table II, in order to elaborate on the attributes of the optimal solution to the MSE minimization problem, we present the values of the harvested and transmit power of each sensor at the optimal solution to problem (P1). In this example, we set \( P = 30 \text{ dBm}, R_s = 0.1I_{n_s}, \) \( \gamma^{-1} = 0.015, \) \( \delta_\theta = 1, \) and \( n_r = 5, 10, 15, 20. \) One can see that some of the sensors do not use their maximum power harvested from the FC, which implies that power control is needed to guarantee the optimal solution. In other words, some of individual power constraints (i.e., the first constraint) in problem (P1) may not be fully utilized, or equivalently, the corresponding dual variables may be zero. This is attributed from the fact that transmission with the full power may increase the interference level at the FC, which in turn reduces the estimation reliability. In this example, sensors 2, 5, 6, and 7 use only a fraction of their harvested power.

C. Total Power Minimization

Fig. 4 illustrates the average minimum transmit power at the FC for distributed estimation versus iteration index when \( n_s = 10, R_s = 0.1I_{n_s}, \gamma^{-1} = 0.015, \) \( \delta_\theta = 1, \) and \( n_r = 5, 10, 15, 20. \) As shown in this figure, the proposed algorithm converges quickly, and the transmit power is reduced as the number of antennas, \( n_r, \) increases. In Fig. 5, the average minimum transmit power at the FC for distributed estimation is shown as a function of the distortion target \( \gamma^{-1} \) when \( n_s = 10, R_s = 0.1I_{n_s}, \delta_\theta = 1, \) and \( n_r = 5. \) It is clear that the more strict distortion requirement is, the more power is needed. Note that the distortion target must be no smaller than the benchmark MSE value such that the optimization problem is feasible. One can also see that the optimal scheme should be used for power saving. In this example, we can save the amount of transmit power of 7.44, 9.46, and 9.05 dBm at \( \gamma^{-1} = 0.02, 0.03, 0.04, \) respectively, compared to the suboptimal case.

D. A Common Energy Harvester

Finally, we validate the performance of the distributed estimation system with a common energy harvester. Specifically,
the power–distortion trade-off is ascertained by referring to Fig. 6, where the optimal MSE is depicted as a function of

the minimum transmit power \( P \) for distributed estimation when \( R_n = 10^{-2} I_{n_r}, \delta_0 = 1, n_s = n_r = 4, \) and \( n_s = n_r = 8. \) In this figure, the region above each trade-off curve is achievable. As \( P \) tends to infinity, the MSE converges to that of centralized estimations, i.e., \( [1^T R_n^{-1}]^{-1} \), plotted with the dotted curve. Moreover, as expected, the achievable region gets broader for a larger \((n_s, n_r)\) pair.

VI. CONCLUDING REMARKS

Using the SDR, we developed a new framework for solving the network lifetime problem of a WSN. To that end, we adopted the notion of RF-based WPT as well as the multiple-antenna technology so that both the life span and the estimation performance are substantially improved. In this paper, two optimization problems were formulated and iteratively solved by two proposed algorithms, which turned out to guarantee the convergence at least to a local optimum. We showed that power control is indeed required at the optimal solution. It was also shown that having multiple antennas at the FC provides a significant improvement in the estimation performance. Especially, it was shown that as the number of antennas grows large, the MSE of the distributed estimation with the BLUE approaches that of centralized estimations.

APPENDIX A

PROOF OF THEOREM 1

Proof: We start by proving the first property of Theorem 1. We exploit the strong duality and then examine the Karush-Kuhn-Tucker condition of (SDR1). Let \( \nu, \{ \lambda_k \}_{k=1}^{n_s}, \) and \( \beta \) be the dual variables of problem (SDR1). The Lagrangian of problem (SDR1) is defined as

\[
\mathcal{L}(Q, W, \eta, \nu, \lambda, \beta) = -\text{tr}(Q \Sigma) + \beta \left( \text{tr}(W) - \eta P \right) + \sum_{k=1}^{n_s} \lambda_k \text{tr}(D_k Q - G_k W) + \nu \left( \text{tr}(Q \Psi) + \eta \nu^T R_s v - 1 \right).
\]

Then, the dual function of problem (SDR1) is given by

\[
\min_{Q \succeq 0, W \succeq 0, \eta > 0} \mathcal{L}(Q, W, \eta, \nu, \lambda, \beta),
\]

which can be equivalently expressed as

\[
\min_{Q \succeq 0, W \succeq 0, \eta > 0} \text{tr}(Q Y^*) + \text{tr}(WB) + \eta \xi - \nu,
\]

where

\[
\xi = \nu^T R_s v - \beta P,
\]

\[
Y = -\Sigma + \sum_{k=1}^{n_s} \lambda_k D_k + \nu \Psi
\]

\[
Z = -\sum_{k=1}^{n_s} \lambda_k G_k + \beta I.
\]

When we let \( \nu^*, \{ \lambda_k^* \}_{k=1}^{n_s}, \) and \( \beta^* \) be the optimal dual solutions to problem (SDR1), we define

\[
Y^* = -\Sigma + \sum_{k=1}^{n_s} \lambda_k^* D_k + \nu^* \Psi
\]

\[
Z^* = -\sum_{k=1}^{n_s} \lambda_k^* G_k + \beta^* I.
\]

Then, the optimal \( \bar{Q}^* \) must be the solution to the following problem:

\[
\begin{align*}
\min_{Q \succeq 0} \quad & \text{tr}(Q Y^*) \\
\text{subject to} \quad & Y \succeq 0, Z \succeq 0, \xi \geq 0 \\
& \beta \geq 0, \lambda_k \geq 0, \forall k \in [n_s].
\end{align*}
\]

Since the duality gap between problem (SDR1) and (58) is zero, \( \nu^* \) is equal to the optimal value of problem (SDR1), which is positive. Thus, we conclude that \( \nu^* > 0. \) Next, we will show that \( \beta^* > 0. \) First, if there exists a \( \lambda_k > 0, \) then from the condition \( Z \succeq 0, \) it follows that \( \beta^* > 0. \) From the condition \( \xi^* = 0 \) and the facts that \( R_n \succeq 0 \) and \( \nu^* > 0, \) we also conclude that \( \beta^* > 0. \)

To prove the second property of Theorem 1, we use the fact that for any two matrices of the same size \( A \) and \( B, \) rank \((A - B) \geq \text{rank}(A) - \text{rank}(B)) [43]. \) Since \( \beta^* > 0 \) and rank \( \sum_{k=1}^{n_s} \lambda_k^* G_k \leq n_s, \) it follows from (55) that rank \( \bar{Z}^* \geq [n_r - n_s] \geq n_r - n_s. \) Let Null \( (\bar{Z}^*) \) be the null space of \( \bar{Z}^*. \) Then from the condition \( \bar{W}^* \bar{Z}^* = 0, \) we must have \( \bar{W}^* \in \text{null}(\bar{Z}^*). \) Since rank \( \bar{Z}^* \geq n_r - n_s \) and rank \( \bar{W}^* \leq \dim(\text{null}(\bar{Z}^*)), \) it follows that rank \( \bar{W}^* \leq n_s. \) Using the fact that \( \bar{W}^* \) is an \( n_r \times n_r \) matrix, we conclude that rank \( \bar{W}^* \leq n_s. \)

Finally, we prove the property of the optimal solution \( \bar{Q}^*. \) Since \( \bar{W}^* \succeq 0, \) \( \sum_{k=1}^{n_s} \lambda_k^* D_k \succeq 0, \) and \( \nu^* > 0, \) we obtain

\[
\text{rank} \left( \nu^* \Psi + \sum_{k=1}^{n_s} \lambda_k^* D_k \right) = n_s.
\]

Hence, from the definition of \( Y^* \) in (54), it follows that

\[
\text{rank}(Y^*) \geq n_s - \text{rank}(\Sigma) = n_s - 1.
\]

From the condition (57), \( \bar{Q}^* \) must lie in the null space of \( Y^*. \) Therefore, rank \( \text{rank}(\bar{Q}^*) \leq \dim(\text{null}(Y^*)), \) which is upper-bounded by one due to (60). Now, assume that rank \( \text{rank}(Y^*) = n_s. \) Then from (57), it follows that \( \bar{Q}^* = 0, \) which cannot be the
optimal solution to problem (SDR1). In consequence, we must have rank \( (Y^*) = n_r - 1 \) and thus rank \( (Q^*) = 1 \), which completes the proof of Theorem 1.

### Appendix B

#### Proof of Theorem 3

Proof: We derive the result in (44) by showing that the optimal MSE in (P1–Sum) and the optimal power in (P2–Sum) are the inverse of each other. We start the proof by introducing the following lemma.

**Lemma 1:** For a given \( v \) in (12), let \((a^*_1, w^*_1)\) and \((a^*_2, w^*_2)\) be the objective functions to problems (P1–Sum) and (P2–Sum), respectively. We also let \( f_1 \cdot \) and \( f_2 \cdot \) be the objective functions in (P1–Sum) and (P2–Sum), respectively. Then, they obey the property: if \( \gamma = f_1 a^*_1 \cdot w^*_1 \), then \((a^*_2, w^*_2) := (f^*_2 a^*_1, w^*_1)\); and if \( P = f_2 a^*_2 \cdot w^*_2 \), then \((a^*_1, w^*_1) := (f^*_1 a^*_2, w^*_2)\).

**Proof:** First, it is worth noting that all the inequality constraints in (P1–Sum) and (P2–Sum) are satisfied with equality at the optimal solutions. For a given \( v \), we have \( ||w^*_1||^2 = P \). We will prove that \((a^*_1, w^*_1)\) is also a solution to (P2–Sum), i.e., \((a^*_2, w^*_2) := (f^*_2 a^*_1, w^*_1)\). We prove it by contradiction. Assume that \((a^*_1, w^*_1)\) is not a solution to (P2–Sum), that is, there exists a feasible solution \((a^*_1, w^*_1)\) to (P2–Sum) such that \( ||w^*_1|| < ||w^*_1|| \). In other words, we can find a constant \( c > 1 \) such that

\[
||w^*_1|| < c ||w^*_1|| \leq ||w^*_1||.
\]

(61)

Since the objective function in (P1–Sum) is monotonically increasing with the norm of \( a \), it follows that \( f_1 c a^*_1 \cdot w^*_1 > f_1 a^*_1 \cdot w^*_1 \). Since \((a^*_1, w^*_1)\) is feasible to (P1–Sum), we also have that \( f_1 c a^*_1 \cdot w^*_1 > f_1 a^*_1 \cdot w^*_1 \). Thus, we obtain

\[
f_1 c a^*_1 \cdot w^*_1 > f_1 a^*_1 \cdot w^*_1 \geq f_1 a^*_1 \cdot w^*_1.
\]

(62)

From (61) and the second constraint in (P1–Sum), the solution \((c a^*_1, w^*_1)\) is also feasible to (P1–Sum), and yields a higher objective value than the optimal \((a^*_1, w^*_1)\) does. This contradicts the assumption that \((a^*_1, w^*_1)\) is optimal to (P1–Sum), and thus \((a^*_1, w^*_1)\) must be a solution to (P2).

The proof for the second claim can be found using the similar steps to the proof for the first one, and hence is omitted here.

From (34), the optimal mse satisfies (44). Since the optimal \( P \) and mse are the inverse of each other, it follows that optimal \( P \) also satisfies (34). Due to the fact that the above property holds for any \( v \), it holds for the optimal \( v \) as well, which completes the proof of Theorem 3.
[32] O. Orhan, D. Guindüz, and E. Erkip, “Source-channel coding under energy, delay, and buffer constraints,” IEEE Trans. Wireless Commun., vol. 14, no. 7, pp. 3836–3849, Jul. 2015.

[33] M. Gastpar and M. Vetterli, Source-Channel Communication in Sensor Networks, (Lecture Notes in Computer Science). New York, NY, USA: Springer, 2006, vol. 2634.

[34] H. Ju and R. Zhang, “Throughput maximization in wireless powered communication networks,” IEEE Trans. Signal Process., vol. 13, no. 1, pp. 418–428, Jan. 2014.

[35] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[36] A. Charnes and W. W. Cooper, “Programming with linear fractional functionals,” Naval Res. Logist. Quart., vol. 9, pp. 181–186, 1962.

[37] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” Mar. 2004. [Online]. Available: http://cvxr.com/cvx

[38] B.-T. Aharon and A. Nemirovski, Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, (MOS-SIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2001.

[39] R. Hunger, “Floating point operations in matrix-vector calculus,” Technische Universität München, Munich, Germany, Tech. Rep. TUM-LNS-TR-05-05, 2007.

[40] K. Gomadam, V. R. Cadambe, and S. A. Jafar, “A distributed numerical approach to interference alignment and applications to wireless interference networks,” IEEE Trans. Inf. Theory, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[41] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, “An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel,” IEEE Trans. Signal Process., vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[42] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[43] R. Horn and C. R. Johnson, Matrix Analysis. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[44] S. Y. Seidel and T. S. Rappaport, “914 MHz path loss prediction models for indoor wireless communications in multifloored buildings,” IEEE Trans. Antennas Propag., vol. 40, no. 2, pp. 207–217, Feb. 1992.

[45] J. A. Gutierrez, M. Naeve, E. Callaway, M. Bourgeois, V. Mitter, and B. Heile, “IEEE 802.15.4: A developing standard for low-power, low-cost wireless personal area networks,” IEEE Netw., vol. 15, no. 5, pp. 12–19, Sep. 2001.

Vien V. Mai received the B.E. degree in electronic and telecommunication from Posts & Telecommunications Institute of Technology, Ho Chi Minh, Vietnam, in 2013, the M.S. degree in electronics and radio engineering from Kyung Hee University, Yongin, South Korea, in 2015. He is currently working toward the Ph.D. degree KTH Royal Institute of Technology, Stockholm, Sweden. From September 2015 to July 2016, he was with the Communications and Networking Laboratory, Dankook University, Yongin, as a Researcher. Since August 2016, he has been with KTH Royal Institute of Technology. His research interests include communication theory, distributed detection and estimation, and design and complexity analysis of efficient randomized algorithms suitable for scalable convex optimization.

Won-Yong Shin (S’02–M’08–SM’16) received the B.S. degree in electrical engineering from Yonsei University, Seoul, South Korea, in 2002, and the M.S. and Ph.D. degrees in electrical engineering and computer science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2004 and 2008, respectively. From February 2008 to April 2008, he was a Visiting Scholar in the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. From September 2008 to April 2009, he was with the Brain Korea Institute and CHIPS at KAIST, as a Postdoctoral Fellow. From August 2009 to April 2009, he was with the Lumicomm, Inc., Daejeon, as a Visiting Researcher. In May 2009, he joined Harvard University as a Postdoctoral Fellow and was promoted to a Research Associate in October 2011. Since March 2012, he has been with the Department of Computer Science and Engineering, Dankook University, Yongin, South Korea, where he is currently a Tenured Associate Professor. His research interests include the areas of information theory, communications, signal processing, mobile computing, big data analytics, and online social networks analysis.

Dr. Shin has served as an Associate Editor of the IEEE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, the IEICE Transactions on Smart Processing & Computing, and the Journal of Korea Information and Communications Society. He served as a Guest Editor of The Scientific World Journal (Special Issue on Challenges towards 5G Mobile and Wireless Communications) and the International Journal of Distributed Sensor Networks (Special Issue on Cloud Computing and Communication Protocols for IoT Applications). He also served as an Organizing Committee for the 2015 IEEE Information Theory Workshop. He received the Bronze Prize of the Samsung Humantech Paper Contest (2008) and the KICS Haedong Young Scholar Award (2016).

Koji Ishibashi (S’01–M’07) received the B.E. and M.E. degrees in engineering from the University of Electro-Communications, Tokyo, Japan, in 2002 and 2004, respectively, and the Ph.D. degree in engineering from Yokohama National University, Yokohama, Japan, in 2007. From 2007 to 2012, he was an Assistant Professor in the Department of Electrical and Electronic Engineering, Shizuoka University, Hamamatsu, Japan. From 2010 to 2012, he was a Visiting Scholar in the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Since April 2012, he has been with the Advanced Wireless & Communication Research Center, The University of Electro-Communications, Tokyo, Japan, where he is currently an Associate Professor. His current research interests are energy harvesting communications, wireless power transfer, channel codes, signal processing, and information theory.