Wirelessly-powered Sensor Networks: Power Allocation for Channel Estimation and Energy Beamforming

Rong Du, Hossein Shokri Ghadikolaei, Carlo Fischione School of Electrical Engineering and Computer Science KTH Royal Institute of Technology, Stockholm, Sweden Email: {rongd, hshokri, carlofi}@kth.se

Abstract

Wirelessly-powered sensor networks (WPSNs) are becoming increasingly important in different monitoring applications. We consider a WPSN where a multiple-antenna base station, which is dedicated for energy transmission, sends pilot signals to estimate the channel state information and consequently shapes the energy beams toward the sensor nodes. Given a fixed energy budget at the base station, in this paper, we investigate the novel problem of optimally allocating the power for the channel estimation and for the energy transmission. We formulate this non-convex optimization problem for general channel estimation and beamforming schemes that satisfy some qualification conditions. We provide a new solution approach and a performance analysis in terms of optimality and complexity. We also present a closed-form solution for the case where the channels are estimated based on a least square channel estimation and a maximum ratio transmit beamforming scheme. The analysis and simulations indicate a significant gain in terms of the network sensing rate, compared to the fixed power allocation, and the importance of improving the channel estimation efficiency.

Index Terms

Wirelessly-powered sensor network, wireless energy transfer, power allocation, channel acquisition, non-linear energy harvesting

I. INTRODUCTION

Traditional battery-powered wireless sensor networks suffer a major problem of the limited energy budget of the nodes. Unless the battery of the nodes are replaced periodically, the network

The paper has been accepted in IEEE Transactions on Wireless Communications on Jan. 19th, 2020

lifetime is limited. Thus, for long-term monitoring applications, rechargeable sensor networks are more appealing than traditional battery-powered ones [1]. A promising technique to recharge the sensor nodes is called wireless energy transmission (WET) [2], [3], in which electromagnetic waves periodically recharge sensor nodes to extend their lifetime. Compared to the ambient energy harvesting, WET provides a better predictability, controllability, and reliability [4], leading to a more consistent performance of the network.

A. Related Works and Motivations

WET systems can be broadly divided into two categories, according to the transmission of the data: simultaneous wireless information and power transfer [5] and wirelessly-powered communication networks (WPCNs) [6]. In the first category, the transmitter simultaneously sends energy and data at the same time, and the receiver allocates some resources (e.g., time, power, or antennas) for energy harvesting and the rest for data communication [7]. In the second category, WPCN [6], [8], the process of energy and data transmissions is sequential, meaning that the energy receivers send their data using the energy harvested from the transmitters. Most of the existing studies in WPCN focus on maximizing the throughput of wireless devices via optimizing frequency or time schedules for energy transmissions and for data transmissions [8–11].

The severe propagation loss in wireless medium may result in a small received energy that may not be enough for the data transmission task [2]. To overcome this problem, instead of adopting an arbitrarily large transmission power, which is not possible due to the safety issue [12–14], we can substantially improve the WET efficiency by energy beamforming [15], [16]. More specifically, we can steer the energy toward the receivers, such that with the same transmission power the receivers can harvest substantially more energy compared to an omnidirectional energy transmission scheme. To this end, we need multiple antennas and channel state information (CSI). Fig. 1 illustrates a typical wirelessly-powered sensor networks (WPSNs) [17], which is a special case of WPCN. In WPSNs, some energy sources, hereafter called base stations (BSs), provide energy to the nodes using WET, and the nodes use the received energy to make measurements and send them to a sink, which could be an energy source. In this paper, we focus on the WPSNs with one BS.

There are some results on optimal beamforming design for energy transfer. With perfect CSI, the work in [18] shows that the optimal energy beamforming in terms of received energy of a point-to-point MIMO system can be achieved by the eigenvector corresponding to the largest

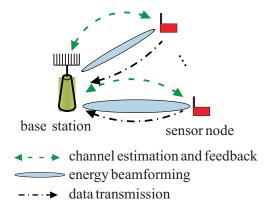


Fig. 1: The wirelessly-powered sensor network considered in this paper. The network consists of one base station and multiple sensor nodes. The base station uses energy beamforming to transmit energy to the nodes, and the nodes use the harvested energy for sensing and data transmission back to a sink node.

eigenvalue of the channel matrix. For a WPCN with multiple energy receivers, the authors of [9] study a joint time allocation and energy beamforming problem to maximize the network sumthroughput and provide a solution approach based on semi-definite relaxation. In a wireless sensor network, however, monitoring performance and lifetime are more important than the network sum-rate. References [19–22] study the monitoring performance and the lifetime of WPSNs, assuming that the energy sources (BSs) know perfect CSI a priori at no cost. However, in practice, the energy transmitters should always spend power and time to acquire CSI [6], and the WET efficiency greatly depends on the channel estimation quality.

The accuracy of the CSI estimation for WET has thus been investigated in [3], [16], [23–28]. In particular, the work in [23] investigates the interplay between the power allocation for channel estimation and the expected received energy at the receiver. However, in the context of wireless sensor networks, there is no study for the power allocation for channel estimation and energy transmission to maximize the network performance. The authors of [29] study a network with one multiple-antenna energy transmitter and one single-antenna energy receiver, and formulate an optimization problem on how much power and time to spend in channel estimation and channel feedback to maximize the rate in downlink data transmission. The objective function is non-convex and not analytically tractable. Thus, they maximize an upper bound of the downlink data rate instead. The work in [25] considers a similar setting, where the transmitter sends predefined energy beams to the receiver, and the receiver feeds back the strength of the received signal. The authors optimize the size of the training codebook for channel acquisition to balance the trade-off between channel estimation precision and energy transmission time. These results are extended to

multiple single-antenna energy receivers in [26]. Based on the feedback, the transmitter employs a maximum likelihood channel estimation method and clusters the receivers according to their channel phases during energy transmission. The work in [30] studies a network with one singleantenna energy transmitter and multiple single-antenna energy receivers, where the transmitter applies minimum mean square error (MMSE) to estimate the channels. The authors optimize the time allocation for channel estimation and energy transmission to maximize the energy efficiency of the network. In [27], the authors investigate energy harvesting (EH) with an one-bit feedback, where every receiver only reports whether the harvested energy in the current time interval is larger than that of the previous interval, and the transmitter estimates the channel using analytic center cutting plane method. With a similar training scheme, the authors of [28] consider the cases where the energy transmitter sends multi-sine energy signals over different frequencies, and model the EH by a non-linear function. Reference [24] proposes a pilot design approach that transmits pilots from the energy receiver to the energy transmitter for a point-topoint WET system and trades-off the channel estimation accuracy and the corresponding spent energy. Although this work provides the optimal solution for some special cases, the general cases remain unsolved due to the intractability of the expressions. Reference [16] extends this work to a scenario with multiple transmitters and single receiver.

Most of the above-mentioned studies use a linear EH model in the formulation. However, in reality, the EH is non-linear due to circuit sensitivity limitations, current leakage [31], and filters in the circuits [32], among others. A linear EH model sometimes can be considered as a special case of the non-linear harvesting models [32]. Therefore, the algorithms designed for the problem with linear EH models and the resulting insights may not hold for a general non-linear model. There are some works studying robust energy transmission with non-linear EH model [33], [34]. They model the imperfect CSI by a channel error with either a deterministic bound, or a random one with a deterministic and bounded variance. However, they do not study how to improve the EH performance by properly allocating resources in channel acquisition. Moreover, recent attempts on the design of algorithms for the non-linear EH cases are based on some restrictive assumptions [31], [34]. For example, the solution for [31] works for a specific EH model, and whether it is valid for other models is unsure. The approach in [34] can deal with the cases where the EH related constraint requires that the harvested energy is larger or smaller than a pre-defined threshold, which allows us to transform the EH related constraint into a linear constraint on the received energy. The solutions for a fairly general EH model are still open. In this paper, we

try to address this important research gap. We consider a wireless sensor network with multiple energy receivers, where the input-output model of the EH circuits of these energy receivers does not need to be a particular function, but rather whatever function as long as it is monotone and increasing. Hereafter, we refer to this EH model (formally defined in Section II) as a generic model to distinguish it from a particular function/model. We investigate the fundamental tradeoff between the channel estimation and energy transmission performance in the presence of an arbitrary approach for pilot transmission, channel estimation, and beamforming. We exemplify the use of our framework and show its application in benchmarking the performance of various approaches for pilot transmission, channel estimation, and beamforming. Our novelty is the joint consideration of the following: 1) controlling the CSI quality to benefit the energy transmission process, instead of having it as a given input; 2) generic EH model and channel estimation methods.

B. Contributions

In this paper, we consider a WPSN comprised of one BS and multiple energy receivers, and investigate a new problem of optimal power allocation for channel estimation and energy transmission that maintains a required monitoring performance throughout the network. We substantially extend our preliminary results [35] by optimizing the power allocation for a generic model that characterizes the performance of the channel estimation and energy beamforming schemes. We also extend our preliminary results in [35] to a class of non-linear EH models as long as the function that characterizes the input-output relation of the energy of the EH circuit is monotone and increasing. We develop a novel solution approach based on a bisection search and iterative feasibility checking. We exemplify the proposed solution approach for a specific case of least square (LS) channel estimation and maximum ratio transmission for energy beamforming. To summarize, the main contributions of the paper are as follows:

• We propose a novel problem of power allocation for channel estimation and energy transmission for multiple sensor nodes, to maximize the monitoring performance of a WPSN. We consider a generic EH model, which makes our optimization problem more challenging. However, due to the generality in the model, our results are general and we can directly apply them to any channel estimation and energy beamforming schemes that satisfy the technical conditions in Section II-B.

Symbols	Definition
E	Total energy to be allocated in a time block
E_i^{t}	Energy to be transmitted to v_i from the BS
$E_{\rm s}^*(w)$	Minimum total energy to satisfy the network sensing rate w
N	Number of sensor nodes
$N_{ m t}$	Number of antennas of each BS
$P^{\mathbf{p}}$	Power of the pilots for channel acquisition by the BS
c_i	Static power consumption of v_i
e_i	Power consumption of v_i to transmit a unit size data
$g_i(P^p)$	Expected pilot-estimation-beamforming gain of the BS to v_i
$oldsymbol{h}_i$	Channel from the BS to v_i
w_i	Sensing rate of v_i
w_{\min}	Minimum of the sensing rates of all nodes, i.e., $w_{\min} = \min_i \{w_i\}$
$\eta(\cdot)$	Function of the RF-DC conversion of the nodes

TABLE I: Main notations used throughout the paper.

• We show that the proposed power allocation problem is non-convex in general. Thus, we develop a novel solution method based on iteratively solving a convex optimization instance even when the RF energy conversion model is non-linear. We show that our proposed algorithm can achieve a solution that possesses the desired optimality. We provide a closed-form solution for the cases where we can simplify the EH model to be linear, and the BS has massive antennas and uses orthogonal pilot transmission combined with LS channel estimation and maximum ratio transmission.

C. Paper Structure

The rest of the paper is organized as follows. We describe our WPSN system model and formulate a novel power allocation problem in Section II. We study the problem and provide a solution approach and the corresponding analysis of the approach in Section III, followed by the numerical results in Section IV. We conclude this paper in Section V. To improve the readability of the paper, we present all the proofs in Appendix A.

In this paper, we use the notations as follows: For a vector x, x^T and x^H is its transpose and conjugate transpose, respectively. Notation $||x||^2 \triangleq x^H x$. Notation $\stackrel{\text{a.s.}}{\longrightarrow}$ means converge almost surely. Table I summarizes the main notations of the paper.

II. MODELLING AND PROBLEM FORMULATION

In this section, we introduce our WPSN setting. We then formulate a novel power allocation problem and analyze its complexity.

A. Network Model

We consider a WPSN as shown in Fig. 1. The network has one BS and N > 1 homogeneous sensor nodes, v_1, v_2, \ldots, v_N , demanding low energy consumption rates. We denote set $\mathcal{N} = \{1, 2, \ldots, N\}$. The BS has N_t antennas and uses energy beamforming to transmit RF energy to the sensor nodes. Accordingly, the sensor nodes use the harvested energy to sense and to transmit data. Here, we consider a star topology for the network where the sensor nodes transmit their measurements directly to the data sink, and we comment on how our framework easily applies to a general mesh topology in Section II-D. We assume that the design of energy transmit waveform and channel estimation have been pre-decided, and they are the input of the problem. Here, we optimize the power allocation of the BS in pilot signals (for channel estimation) and in the energy transmission.

B. Channel Estimation and Beamforming

We consider a block-fading wireless channel where the channels from the BS to the sensor nodes remain constant during a coherence interval, hereafter called time block [18]. The energy is carried by a single-carrier signal. We normalize the length of a time block to be 1. In each time block of interest, the BS first sends pilot signals in the initial t^p time with power P^p , receives feedbacks from the sensor nodes, and gets the estimation of the channel. Such an approach is similar to the forward-link training and the power probing scheme in [3]. Our framework allows for both 1) the sensor nodes estimate the channel and feed the estimation back, and 2) the sensor nodes feed the received signal back and the BS estimates the channel. For both cases, the quantization and noise at the feedback process can be considered as additional white noises, which do not affect the results of the paper. In the remaining $1-t^p$ time, the BS transmits energy E_i^t to each sensor nodes. More specifically, within a time block of interest, the BS first transmits energy to v_1 , then to v_2 , and so on. Although such a time-splitting scheme is suboptimal, it has lower complexity and the performance is close to the optimum [36].

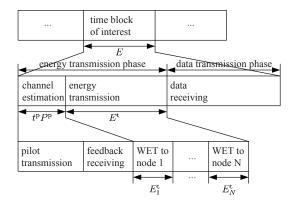


Fig. 2: In each time block of interest, the base station allocates its power for estimating the channels of the sensor nodes, and the energy to transmit to each sensor node.

Let E be the total energy in the time block for channel estimation and energy transmission. Then, we have $E = t^{p}P^{p} + \sum_{j} E_{i}^{t}$. Based on the estimation of the channels, the BS can form energy beams using the existing approaches [3], [14], [36]. We consider the energy transmission as two consecutive processes. The first process denotes how much RF energy is received at the antenna of each node i, E_{i}^{r} , as a function of how much energy is transmitted from the BS, E_{i}^{t} . According to the study in [37], we have that $E_{i}^{r} = g_{TX-RX,i}E_{i}^{t}$,¹ which is a proportional relation and the gain $g_{TX-RX,i}$ depends on the exact channel gain, the accuracy of the channel estimation, and the transmitted RF signals that carry the energy. The output of the first process will be the input of the second process, which denotes how much energy is harvested by node i, E_{i}^{h} , as a function of E_{i}^{r} . We represent it by $E_{i}^{h} = \eta(E_{i}^{r})$, where $\eta(\cdot)$ is the RF-DC conversion function that depends on the rectenna model (here we assume that the recetenna circuits of all sensor nodes are the same). We describe the details of the two processes in the following.

In the first process, we call a combination of the approaches for pilot transmission, channel estimation, and energy beamforming as a pilot-estimation-beamforming (PEB) scheme (here, we do not limit to any particular PEB scheme, but we formulate a general approach that can be applied to different cases of practical interest as long as they satisfy the qualification conditions that will be described in Assumption 1 and 2). Recall that $g_{TX-RX,i}$ depends on the accuracy of the channel estimation, which relates to the power of transmitted pilots P^{p} . Thus, we model

¹We should note that this model is also valid for modulated single-carrier signals, and also the multisine carrier signals with uniform power allocation on each antenna when the channel is frequency-flat, since the performance of this uniform power allocation is very closes to the optimum in frequency-flat channels [38].

 $g_{\text{TX-RX},i} = g_i(P^p)$, and call it PEB gain.² We give an example of g_i for the cases of single continuous wave energy beamforming as follows:

Consider a block-fading channel h_i from the BS to v_i with an additive white Gaussian noise. The covariance of the noise is $\sigma_n^2 I$. In a PEB scheme, the BS transmits pilots with power P^p . After receiving N_t pilots, v_i has its received signal and then transmits it back. Based on the feedback, the BS makes an estimation of the channel $\hat{h}_i(P^p)$, which is a function of P^p . Then, during the energy transmission, if the BS transmits energy E_i^t to v_i with beamforming $\boldsymbol{b}_i(\hat{\boldsymbol{h}}_i(P^p))$ (we simplify it as $\boldsymbol{b}_i(P^p)$ in the following for the notation convenience), the received energy of v_i would be $E_i^r = E_i^t \|\boldsymbol{b}_i^H(P^p)\boldsymbol{h}_i\|^2 / \|\boldsymbol{b}_i(P^p)\|^2$. Thus, the PEB gain in this example is $g_i(P^p) = \|\boldsymbol{b}_i^H(P^p)\boldsymbol{h}_i\|^2 / \|\boldsymbol{b}_i(P^p)\|^2$.

As we can see, the PEB gain $g_i(\cdot)$ abstracts the combined effects of the PEB scheme. Any set of schemes (including the scheme of power probing [3]) that satisfy the following conditions are compatible with our framework.

Assumption 1 (Beamforming qualification condition). We assume that for the BS and any sensor node *i*, and any pilot transmission power $P^{p} > 0$, the PEB gain $g_{i}(P^{p})$ satisfies the following conditions: *i*) $g_{i}(\cdot)$ is an increasing non-negative and bounded function w.r.t. P^{p} , and *ii*) it is smooth and concave w.r.t. P^{p} , thus $g'_{i}(\cdot) > 0$, and $g''_{i}(\cdot) \le 0$.

Remark 1. The example we provided is for unmodulated continuous wave beamforming. Regarding modulated energy-carrier signals and multisine signals, we are not sure whether the beamforming of these signals would satisfy Assumption 1, due to the lack of research on how the accuracy of the channel would affect the received energy under such beamformings. In fact, we do not limit the algorithm to a specific channel estimation method or beamforming scheme, in which case the developed algorithms may not be applicable for other channel estimation methods. Instead, our results are valid for any PEB scheme that satisfies the beamforming qualification condition of Assumption 1. We provide examples in Section III-C where we show its validity for orthogonal pilot transmission combined with LS channel estimation or mimimum mean square error (MMSE) estimation [39], and maximum ratio transmission beamforming [40]. For any other PEB schemes that pass this condition and the EH condition that will be describe next,

²The gain also depends on other factors, such as modulation of the signals and channel estimation methods. However, such factors are considered as input rather than decision variables in our paper. Thus, they are abstracted by the function $g_i(\cdot)$.

one can find the optimal power allocation for channel estimation and beamforming based on the method provided here. For those that do not pass this beamforming qualification condition, we will provide some discussions in Remark 6.

From the monotonicity of g_i , it is straightforward to know that, the PEB gain g_i that the BS can have is bounded by a lower bound 0, and an upper bound $g_i(E/t^p)$. Such a lower bound and upper bound will be used in the development of our solution algorithm, as described in Section III.

Next, we consider the RF-DC conversion process, and we assume that the EH model $\eta(\cdot)$ satisfies the following condition:

Assumption 2 (Energy harvesting condition). The mapping from the received RF energy to the harvested energy, i.e., $E^{\rm h} = \eta(E^{\rm r})$ is a non-negative and monotone increasing function, w.r.t. the received energy $E^{\rm r}$.

Remark 2. This model says that, for a given input signal/waveform and a realization of the wireless channels, the RF-DC circuit of a node can harvest more RF energy with more received RF energy at the antenna. Many non-linear RF EH models [31], [32], [37] (for both singlesine or multisine carriers with some given waveform design schemes, e.g. the uniform power, adaptive single sinewave, and adaptive matched filter in [38], as discussed in Appendix B) and linear models [9], [23], [36] (i.e., $E_i^{\rm h} = \eta E_i^{\rm r}$) can satisfy this assumption. However, it should be noticed that not all beamforming schemes and RF-DC circuits follow the assumption, e.g., when the transmitter could use adaptive waveform design that suboptimally allocates power to different waveforms according to the channel state. The theoretical results of the paper are valid whenever Assumptions 1 and 2 are satisfied. When they are not satisfied, the solution approach may not work and it will be an open question for these cases. In addition, when the transmit waveform strategy of the BS is not predefined and static, the assumption might not hold if there exists a candidate transmit waveform strategy in the strategy set that violates the assumption.

C. Energy Consumption Model

Each sensor node *i* uses a predefined fixed power to transmit data in a predetermined data rate. We denote the energy consumption to sense, to process, and to transmit a unit data to the sink node by $e_i > 0$. Besides, its static energy consumption is c_i , which accounts for circuits

consumptions and also the power of sending channel feedback to the BS. Denote the sensing rate of node v_i by w_i . Then, we have that the total energy consumption of v_i is $e_i w_i + c_i$.³ We require that the average consumed energy of each node is no larger than the average harvested energy, i.e., $e_i w_i + c_i \leq E_i^h$, $\forall i$. Under this requirement, we will optimize the monitoring performance of the WPSN, as we describe next.

D. Power Allocation Problem

The monitoring performance of the WPSN considered here depends on how many measurements are received at the sink node. Naturally, we hope that the nodes make as many measurements as possible. Besides, we do not want to have some nodes to make little measurements whilst some other nodes make too many measurements. Thus, we use the minimum of the sensing rate of all sensor nodes, $w_{\min} \triangleq \min_i \{w_i\}$, as the monitoring performance metric of the WPSN. We also call it network sensing rate. Denote $\boldsymbol{w} = [w_1, w_2, \dots, w_N]^T$ and $\boldsymbol{E}^t = [E_1^t, \dots, E_N^t]^T$. Then, we are ready to formulate the power allocation problem as follows:

$$\max_{w_{\min}, \boldsymbol{w}, \boldsymbol{E}^{\mathrm{t}}, P^{\mathrm{p}}} w_{\min} \tag{1a}$$

s.t.
$$w_i \ge w_{\min}, \forall i \in \mathcal{N},$$
 (1b)

$$e_i w_i + c_i \le \eta \left(E_i^{\mathrm{t}} g_i(P^{\mathrm{p}}) \right), \forall i \in \mathcal{N},$$
 (1c)

$$t^{\mathrm{p}}P^{\mathrm{p}} + \sum_{i} E_{i}^{\mathrm{t}} \le E ,$$
 (1d)

$$w_i, E_i^{\mathrm{t}}, P^{\mathrm{p}} \ge 0, \forall i \in \mathcal{N},$$
 (1e)

where the objective is to maximize the network sensing rate, w_{\min} ; Constraint (1c) is the energy causality, i.e., the consumed energy of a node must be no larger than the energy it harvests; Constraint (1d) is the power limit of the BS; and Constraint (1e) is the non-negative constraint of the decision variables. The problem is to allocate the power of channel estimation and the energy to be transmitted to each sensor node, such that the network sensing rate is maximized.

Remark 3. The optimal solution of Problem (1) should hold that $w_i = w_{\min}, \forall i \in \mathcal{N}$. This can proved by contradiction. Briefly speaking, if there is a node with a sensing rate larger than

³This model is widely used for WSNs [21], [41], [42] because the power that can be used for data transmission by the sensor nodes is very limited, compared to other wireless devices such as mobile phones. However, if one uses the model based on Shannon capacity, our approach is still valid with proper modifications.

 w_{\min} , the BS can transmit less energy to this node, and increase the transmission energy to other nodes to increase their sensing rates, and thus the network sensing rate.

Remark 4 (Extension to mesh networks). For a mesh sensor network (i.e., the sensor nodes will relay other nodes measurements) with a fixed routing, we only need to modify parameter e_i as the summation of the power consumption of sensing, processing, and transmitting its measurement to its destination, and the power consumption of receiving and relaying measurement from each of its child node to its destination. Then, the results of the paper still hold.

E. Complexity Analysis

From Constraint (1c), even when the RF-DC conversion function $\eta(\cdot)$ is linear, simple algebra shows that the Hessian matrix is not necessarily positive semidefinite. Consequently, Problem (1) is not a convex optimization and the solution approach is non-trivial. Notice also that when N = 1, such a problem can be simplified to a convex optimization. Therefore, the difficulty of Problem (1) mainly comes from the power allocation for multiple sensor nodes, addressing which is one of the major technical novelty of this paper. In addition, the non-linear behaviour of $\eta(\cdot)$ makes the problem even more challenging. Despite the non-convexity of the resulting optimization problem, we propose an efficient algorithm that finds the optimal solution of Problem (1), as will be presented in the next section.

III. SOLUTION METHOD

In this section, we investigate a solution approach to solve Problem (1). Then, we show the convergence properties and the computational complexity of the algorithm. We then provide some illustrative examples, in which we can find closed-form solutions of (1). Last, we will discuss some special cases, such as the problem with linear EH models.

A. Algorithm Development

To develop the solution algorithm, we first study a sub-problem on checking the feasibility of the total energy E, based on the assumption that w_{\min} is given. Assume that w_{\min} is given. The following sub-problem finds the minimum total energy to satisfy w_{\min} :

$$\min_{E_{\rm s}, \boldsymbol{E}^{\rm t}, P^p} \quad E_{\rm s} \tag{2a}$$

s.t.
$$e_i w_{\min} + c_i \le \eta \left(E_i^{\mathsf{t}} g_i(P^{\mathsf{p}}) \right), \forall i \in \mathcal{N},$$
 (2b)

$$t^{\mathbf{p}}P^{\mathbf{p}} + \sum_{i \in \mathcal{N}} E_i^{\mathbf{t}} \le E_{\mathbf{s}} \,, \tag{2c}$$

$$P^{\mathbf{p}}, E_i^{\mathbf{t}} \ge 0, \forall i \in \mathcal{N}.$$
 (2d)

Since $\eta(\cdot)$ is non-negative and monotone increasing, we have that $\eta(\cdot)$ has an inverse function, denoted by $\eta^{-1}(\cdot)$, and $\eta^{-1}(\cdot)$ is also non-negative and monotone increasing. Therefore, Constraint (2b) gives us that $E_i^t \ge \eta^{-1}(e_i w_{\min} + c_i)/g_i(P^p)$. To make it more concise, we define $f_i(P^p; w_{\min}) \triangleq \eta^{-1}(e_i w_{\min} + c_i)/(g_i(P^p))$. Then, Problem (2) is equivalent to the following one:

$$\min_{0 \le P^{\mathbf{p}} \le E/t^{\mathbf{p}}} \quad E_{\mathbf{s}}(P^{\mathbf{p}}; w_{\min}) \triangleq t^{\mathbf{p}}P^{\mathbf{p}} + \sum_{i \in \mathcal{N}} f_i(P^{\mathbf{p}}; w_{\min}),$$
(3)

where $E_{\rm s}(P^{\rm p}; w_{\rm min})$ is the total energy to satisfy the required sensing rate $w_{\rm min}$, when the BS uses power $P^{\rm p}$ for the channel estimation. We denote by $E_{\rm s}^*(w_{\rm min})$ the optimum of Problem (3), then we have the following proposition for Problem (3), proved in Appendix:

Proposition 1. Problem (3) is a single variable convex optimization problem, and the optimal solution is either at $P^{p} = 0$ or at the point where its derivative $E'_{s}(P^{p}; w_{\min})$ is 0.

Remark 5. (The proof of) Proposition 1 implies that $E'_{s}(P^{p}; w_{\min})$ is monotone increasing with P^{p} . Thus, we have that

- If E'_s(0; w_{min}) = t^p + ∑_{i∈N} f'_i(0; w_{min}) ≥ 0 for all P^p ≥ 0, then the optimal solution of Problem (3) is P^p = 0.
- If E'_s(E/t^p; w_{min}) ≤ 0, then due to the monotonicity of E'_s(P^p; w_{min}), we have that the P^p that satisfies Equation (11) is larger than E/t^p. This means that the given w_{min} is not achievable with the total energy constraint E.
- Otherwise, i.e., $E'_s(0; w_{\min}) < 0 < E'_s(E/t^p; w_{\min})$, we can achieve the unique solution of Equation (11) numerically using a bisection search algorithm in the region $P^p \in [0, E/t^p]$.

Let $E_{s}^{*}(w)$ be the optimum of Problem (3) given w, and w_{\min}^{*} be the optimum of Problem (1). If $E_{s}^{*}(w) < E$, we have that $w < w_{\min}^{*}$; otherwise $w \ge w_{\min}^{*}$. This gives us the solution algorithm for Problem (1) based on bisection searching. The idea is as follows: We first find a lower bound and an upper bound of w_{\min} , which is denoted by w_{\min}^{l} and w_{\min}^{u} respectively. For the lower bound, we can easily choose $w_{\min}^{l} = 0$. For the upper bound, it corresponds to the case that the BS can achieve the upper bound of the PEB gain, i.e., $g_i(E/t^p)$, without spending any power in pilot transmission. Thus, one can choose w_{\min}^{u} be the optimal solution of the following linear optimization problem:

$$\max_{w_{\min}, \boldsymbol{E}^{t}} \quad w_{\min} \tag{4a}$$

s.t.
$$e_i w_{\min} + c_i \le \eta_{\max} E_i^{t} g_i \left(\frac{E}{t^{p}}\right), \forall i \in \mathcal{N},$$
 (4b)

$$\sum_{i\in\mathcal{N}} E_i^{\mathsf{t}} \le E \,, \tag{4c}$$

$$w_{\min}, P_i^{t} \ge 0, \forall i \in \mathcal{N},$$
(4d)

where $\eta_{\max} \triangleq \max_{x \ge 0} \eta(x)/x$, such that $\eta(x) \le \eta_{\max} x$ for $x \ge 0$. This gives us the proposition of the upper bound rate as follows:

Lemma 1. Consider a feasible Problem (1). The upper bound of w_{\min} is given by

$$w_{\min}^{u} = \frac{\eta_{\max}E - \sum_{i \in \mathcal{N}} \frac{c_i}{g_i(\frac{E}{t^p})}}{\sum_{i \in \mathcal{N}} \frac{e_i}{g_i(\frac{E}{t^p})}}.$$
(5)

Once we have known the upper bound and the lower bound of w_{\min} , we can check the feasibility of $w_{\min} = 0.5(w_{\min}^{l} + w_{\min}^{u})$ for Problem (3). If w_{\min} is feasible, we update the new lower bound by w_{\min} ; otherwise, we update the new upper bound by w_{\min} . This proceeds iteratively until the lower bound and the upper bound converge. Algorithm 1 summarizes this procedure.

Remark 6. For the cases where the PEB gain $g_i(P^p)$ is not concave w.r.t. the pilot power, it is not sufficient to achieve that Problem (3) is convex. However, since Problem (3) is a single variable optimization in a bounded region with no other constraints, one can use numerical approaches, such as bisection search and Newton's method, to find a solution that is close to the optimum. Thus, we can still use Algorithm 1 to find the solution of Problem (1) by properly modifying its Line 11. **Input:** $e_i, c_i, g_i(\cdot), \forall i \in \mathcal{N}, E, t^p, \varepsilon, \eta(\cdot)$ **Output:** $E_i^t, \forall i \in \mathcal{N}, P^p, w_{\min}$ 1: Set $w_{\min}^{l} = 0$ 2: if $\sum_{i\in\mathcal{N}} c_i/(g_i(E/t^p)) > \eta_{\max}E$ then The problem is infeasible and return w = 0. 3: 4: else 5: Find w_{\min}^{u} according to Equation (5) while $w_{\min}^{\mathrm{u}} - w_{\min}^{\mathrm{l}} \ge \varepsilon$ do 6: Set $w_{\min} = 0.5(w_{\min}^{u} + w_{\min}^{l})$ if $t^{p} + \sum_{i \in \mathcal{N}} f'_{i}(0; w_{\min}) \ge 0$ then $P^{p} \leftarrow 0, E^{t}_{i} \leftarrow f_{i}(0; w_{\min}), \forall i \in \mathcal{N}, E^{*}_{s}(w_{\min}) = \sum_{i \in \mathcal{N}} E^{t}_{i}$ 7: 8: 9: 10: else Find $P^{\rm p}$ that satisfies Equation (11). $E_i^{\rm t} \leftarrow f_i(P^{\rm p}; w_{\min}), \forall i, E_{\rm s}^*(w_{\min}) = t^{\rm p}P^{\rm p} + \sum_{i \in \mathcal{N}} E_i^{\rm t}$ 11: 12: end if if $E_{s}^{*}(w_{\min}) - E > 0$ then 13: Update $w_{\min}^{u} = w_{\min}$ 14: 15: else 16: Update $w_{\min}^{l} = w_{\min}$ end if 17: 18: end while Set $w_{\min} = w_{\min}^{l}$, and set P^{p} as the optimal solution of convex optimization Problem (3) 19: for i = 1 to N do 20: Set $E_i^t = f_i(P^p; w_{\min})$. 21: end for 22: return $E_i^t, P^p, \forall i \in \mathcal{N}, w_{\min}.$ 23: 24: end if

B. Performance Analysis

Now, we are ready to analyze the performance of Algorithm 1 in solving Problem (1), in terms of the optimality of the final solution and its computational complexity. The near optimality of the algorithm is given by the following proposition:

Proposition 2. Let Problem (1) be feasible and let its optimum be $w_{\min}^{o} > 0$. Given any arbitrary small gap ε , Algorithm 1 finds a feasible solution ($w_{\min}, w, E^{t}, P^{p}$) that satisfies $w_{\min}^{o} - w_{\min} < \varepsilon$.

Regarding the complexity of Algorithm 1, we have the following proposition:

Proposition 3. Let Problem (1) be feasible and let its optimum be w > 0. Given the arbitrary optimality gap ε , the time complexity of Algorithm 1 is at most $\mathcal{O}(N \log(E) \log(E/(N\varepsilon)))$, where recall that N is the number of sensor nodes, and E is the total energy of BS for each time block.

Consequently, we conclude that Algorithm 1 is an efficient approach (sublinear in N) to find

a feasible and near optimal solution (arbitrarily close to the optimal w_{\min}). However, for the cases where Assumptions 1 and 2 are not satisfied, the optimality gap and time complexity of Algorithm 1 will depend on the the objective function and the solution approach of Problem (3). For example, if the objective function of Problem (3) is unimodal, then we can still use bisection method in Line 11 of Algorithm 1 and observe the same time complexity as in Proposition 3. Next, we provide an illustrative example where the BS uses a simple LS algorithm to estimate the CSI and then uses maximum ratio beamforming for the energy transmission.

C. Illustrative Example

In this subsection, we will give a simple example that the BS uses orthogonal pilot transmission combined with LS channel estimation and maximum ratio transmission, to show this PEB scheme follows qualification conditions of Assumption 1. In this special case, we can further simplify the solutions of our iterative optimization approach and derive closed-form expressions.

First, let us derive the expected harvested energy when the BS uses the LS estimation. During the channel acquisition phase, the sensor nodes use a switch circuit to connect their antenna element with their communication module [3]. Recall that the BS has N_t antennas, and each sensor node has one antenna. Thus, the channel from the BS to node *i* is h_i of size $N_t \times 1$, and we assume that h_i is independent to h_k for $i \neq k$. To estimate the channels toward all the nodes, the BS broadcasts N_t pilots to the nodes. For simplicity, we assume that the pilots are the column vectors of the identical matrix I_{N_t} . If the power of the pilots are P^p , then for each node *i*, it receives the signal as $y_i = \sqrt{P^p/N_t}Ih_i + n_i$, where n_i is an additive white Gaussian noise at the node *i* with covariance $\sigma_n^2 I$. Consider the feedback with quantization, then the feedback of *y* is $\hat{y}_i = y + e_{q,i}$, where $e_{q,i}$ is the zero-mean quantization error. We assume that the quantization is i.i.d, and is independent of the channel and the noise at the receiver. The LS estimation of *h*, based on \hat{y}_i is

$$\hat{\boldsymbol{h}}_{i}^{\text{LS}} = \left(\frac{P^{\text{p}}}{N_{\text{t}}}\boldsymbol{I}^{H}\boldsymbol{I}\right)^{-1}\sqrt{\frac{P^{\text{p}}}{N_{\text{t}}}}\boldsymbol{I}\hat{\boldsymbol{y}}_{i} = \sqrt{\frac{N_{\text{t}}}{P^{\text{p}}}}\hat{\boldsymbol{y}}_{i} = \boldsymbol{h}_{i} + \sqrt{\frac{N_{\text{t}}}{P^{\text{p}}}}(\boldsymbol{n}_{i} + \boldsymbol{e}_{\text{q},i})$$

Define $\tilde{n}_i = n_i + e_{q,i}$, and the covariance of \tilde{n}_i be $\sigma_{\tilde{n}}^2 I$. By setting $b_i = \hat{h}_i^{LS}$ as the beamforming vector, the expected received power of node *i* becomes

$$E_i^{\mathrm{r}}(P^{\mathrm{p}}) = E_i^{\mathrm{t}} \mathbb{E} \left[\frac{\|\boldsymbol{h}_i^H \boldsymbol{h}_i + \sqrt{\frac{N_{\mathrm{t}}}{P^{\mathrm{p}}}} \boldsymbol{h}_i^H \tilde{\boldsymbol{n}}_i\|^2}{\|\boldsymbol{h}_i + \sqrt{\frac{N_{\mathrm{t}}}{P^{\mathrm{p}}}} \tilde{\boldsymbol{n}}_i\|^2} \right]$$

Therefore, we have that

$$g_i(P^{\mathbf{p}}) = \mathbb{E}\left[\frac{\|\sqrt{P^{\mathbf{p}}}\boldsymbol{h}_i^H\boldsymbol{h}_i + \sqrt{N_{\mathbf{t}}}\boldsymbol{h}_i^H\tilde{\boldsymbol{n}}_i\|^2}{\|\sqrt{P^{\mathbf{p}}}\boldsymbol{h}_i + \sqrt{N_{\mathbf{t}}}\tilde{\boldsymbol{n}}_i\|^2}\right].$$
(6)

where the expectation is taken over the distribution of \tilde{n}_i . This is because, when the BS models the function $g_i(\cdot)$ and formulates the optimization problem, it has not sent out the pilots and the noise is not realized. We show in Appendix C that $g_i(P^p)$ is concave when $P^p \ge (2\sqrt{3} - 1)(N_t^2\sigma_n^2)/||\mathbf{h}_i||^2$. For the cases when P^p is smaller than such a threshold, it is not sufficient to determine whether it is concave or not. To have a better understanding of its concavity, we run a simulation scenario as follows. We place the BS at (0,0), and we put a sensor node at a random location near the BS. The distance of the node to the BS turns out to be 11.69 meters. The BS has $N_t = 100$ antennas, transmitting energy pilot at frequency 915MHz. The wireless channel is modelled as Rician with factor 10, and we generate 1000 instances of the channels \mathbf{h}_i . The noise at the sensor node is $\sigma_n^2 = -90$ dBm. We first consider the perfect feedback case, i.e., no quantization error. Then, we vary P^p from 0.1 mW to 0.1W, and calculate the averaged $g_i(P^p)$ for each P^p . The result is shown in Fig 3(a). It can be observed that the curve of $g(P^p)$ is an increasing and concave function of P^p .

To simplify the simulations, we want to find an approximation of $g_i(\cdot)$ in (6). Therefore, we approximate $g_i(P^p)$ by

$$g_i(P^{\mathrm{p}}) \approx \hat{g}_i(P^{\mathrm{p}}) = \sigma_{h_i}^2 \frac{P^{\mathrm{p}} \sigma_{h_i}^2 + N_{\mathrm{t}} \sigma_{\tilde{\mathrm{n}}}^2}{P^{\mathrm{p}} \sigma_{h_i}^2 + N_{\mathrm{t}}^2 \sigma_{\tilde{\mathrm{n}}}^2},\tag{7}$$

where $\sigma_{h_i}^2 = \mathbb{E}[\boldsymbol{h}_i^H \boldsymbol{h}_i].$

To compare $g_i(P^p)$ and its approximation $\hat{g}_i(P^p)$, we plot $\hat{g}_i(P^p)$ in Fig 3(a) by the red dots. It can be observed that the two functions are close, i.e., the difference $|\hat{g}(P^p) - g(P^p)|/g(P^p)$ is small. Thus, we consider $\hat{g}_i(P^p)$ as a good approximation of $g_i(P^p)$, and we will use \hat{g}_i instead of g_i in the following and in the simulations. One can examine that, $\hat{g}_i(P^p)$ is a concave and monotone increasing function by simple algebra. We also check the effect of feedback quantization in Fig 3(a). The green dotted line and the blue solid line in Fig. 3(a) correspond to quantization with 4-bit mantissa and no quantization (perfect feedback), respectively. Observing up to 4% relative difference between these curves indicates the marginal effect of quantization compare to that of pilot power optimization. Thus, in the following, we would neglect the quantization error, i.e., $\tilde{n} = n$, for the sake of simplicity.

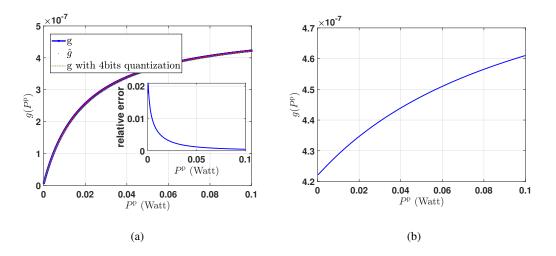


Fig. 3: (a) The PEB gain and its approximation at different pilot power P^{p} of the least square estimator cases; (b) The PEB gain at different pilot power P^{p} of the MMSE estimator cases.

We start from Remark 5 and compute $E'(P^{p}; w_{\min})$. We replace function g_i in Equation (11) (in Appendix A) by \hat{g}_i , and have that

$$E'(P^{\mathbf{p}}; w_{\min}) = t^{\mathbf{p}} + \sum_{i \in \mathcal{N}} f'_{i}(P^{\mathbf{p}}|w_{\min}) = t^{\mathbf{p}} - \sum_{i \in \mathcal{N}} \eta^{-1} \left(e_{i}w_{\min} + c_{i} \right) \frac{N_{\mathrm{t}}\sigma_{\mathrm{n}}^{2}(N_{\mathrm{t}} - 1)}{(P^{\mathbf{p}}\sigma_{h_{i}}^{2} + N_{\mathrm{t}}\sigma_{\mathrm{n}}^{2})^{2}}.$$
 (8)

We can see from Equation (8) that $E'(P^{\rm p}; w_{\min})$ is strictly increasing with $P^{\rm p}$. Therefore, if Equation (8) is feasible, the $P^{\rm p}$ that satisfies $E'(P^{\rm p}; w_{\min}) = 0$ is unique, which can be achieved by the bisection approach. For a special case where $\sigma_{h_i}^2 = \sigma_{h_k}^2 = \sigma_{h_k}^2 = \sigma_{h_k}^2 = \delta_{h_k}^2 = \sigma_{h_k}^2 = \delta_{h_k}^2 = \delta_{h_k}$

$$P^{\mathbf{p}} = \sqrt{\sum_{i \in \mathcal{N}} \frac{\eta^{-1}(e_i w_{\min} + c_i) N_{\mathbf{t}} \sigma_{\mathbf{n}}^2 (N_{\mathbf{t}} - 1)}{t^{\mathbf{p}} \sigma_{\mathbf{h}}^4}} - \frac{N_{\mathbf{t}} \sigma_{\mathbf{n}}^2}{\sigma_{\mathbf{h}}^2}.$$

If we use such a result in Line 11 of Algorithm 1, its computational complexity is then reduced from $\mathcal{O}(N\log(E)\log(E/(N\varepsilon)))$ to $\mathcal{O}(N\log(E/(N\varepsilon)))$. We can see that, given a certain channel estimation approach, and some other special conditions, we can revise Algorithm 1 to a simpler and faster one. Also, since $\sigma_{h_i}^2 = \sigma_{h_k}^2 = \sigma_{h}^2$, $\forall k \neq i$, we have that $g_i(x) = g_k(x)$. Thus, under such a special case, the upper bound of w_{\min} corresponding to Lemma 1 can be simplified to

$$w_{\min}^{u} = \frac{\eta_{\max}g\left(\frac{E}{t^{p}}\right)P - \sum_{i \in \mathcal{N}}c_{i}}{\sum_{i \in \mathcal{N}}e_{i}}.$$
(9)

We will use this result in the simulation in Section IV to check the performance of Algorithm 1 in the special case of this section.

Besides the LS estimator, we also study the cases where the BS uses the minimum mean square error (MMSE) [39], [43] to estimate the channels, i.e., the estimated channel is

$$\hat{\boldsymbol{h}}_{i}^{\mathrm{MMSE}} = \boldsymbol{R}_{h_{i}} \sqrt{\frac{P^{\mathrm{p}}}{N_{\mathrm{t}}}} \left(\frac{P^{\mathrm{p}}}{N_{\mathrm{t}}} \boldsymbol{R}_{h_{i}} + \sigma_{\mathrm{n}}^{2} \boldsymbol{I}\right)^{-1} \boldsymbol{y}_{i},$$

where \mathbf{R}_{h_i} is the channel covariance matrix. Then, similar to what we have done for the LS estimator case, we run the same simulation by setting $\mathbf{b}_i = \hat{\mathbf{h}}_i^{\text{MMSE}}$. We present the result in Fig. 3(b), which shows the concavity of the PEB gain $g_i(P^p)$.

D. Solution for Linear Energy Harvesting Model

Previously, we have provided the solution algorithm for Problem (1). In this subsection, we consider a special case where the EH of the nodes are modelled by a linear function. That is, the EH model is $\eta(x) = \alpha x$, where α is a constant representing the EH rate. In this case, we can still use Algorithm 1 to solve the problem. The analytical results that we have provided in the previous subsections are also valid for the case of linear EH modelling. It means that, given a non-linear EH model, we can find a linear approximation of the model and compare the solutions of Algorithms 1 for the non-linear EH model case with its linear approximation case. Such comparison enables the sensitivity analysis of the solution to the linear approximation.

E. Asymptotic Case

In this subsection, we consider the problem with linear EH model we discussed in the previous subsection. We study the behaviour of LS channel estimation and maximum ratio transmit beamforming with $N_t \rightarrow +\infty$. We have the following results:

Remark 7. Consider a BS with sufficiently large number of antennas, constant transmit power per antenna, orthogonal pilot transmission combined with least-square channel estimation and maximum ratio transmission scheme. Then

$$g_i(P^{\rm p}) \xrightarrow{a.s.} \frac{N_{\rm t}\sigma_i^4 P^{\rm p}}{\sigma_i^2 P^{\rm p} + N_{\rm t}\sigma_{\rm p}^2},$$
(10)

where $\sigma_i^2 = \mathbb{E}_k[||h_{ik}||^2]$, and h_{ik} is the k-th element of vector h_i . That is, $g_i(P^p)$ converges almost surely to a monotone increasing and concave function.

Notice that the total pilot transmit power can grow arbitrarily large in Remark 7. We have added that assumption for the sake of mathematical tractability in the proof. In many wireless

systems, however, we can replace N_t by the number of multi-path components of the channel, which is a constant number, independent of N_t [44]. This number could be very small, e.g., less than 5 in millimeter-wave networks [45].

From Remark 7, when N_t is fixed and large enough (e.g., massive antenna regime), the PEB gain $g_i(P^p)$ is monotone increasing and concave with P^p , and therefore it satisfies the qualification conditions of Assumption 1. Thus, our proposed algorithm is applicable to such a case. However, in the asymptotic regime, described by Equation (10), we can have a simpler solution. Recall that in Algorithm 1, we need to solve sub-problem (3) in each iteration. By substituting (10) into (3), the sub-problem turns to be

$$\min_{P^{\mathrm{p}} \ge 0} \quad t^{\mathrm{p}} P^{\mathrm{p}} + \sum_{i \in \mathcal{N}} \frac{\sigma_{\mathrm{n}}^{2}(e_{i}w_{\mathrm{min}} + c_{i})}{\alpha \sigma_{i}^{4} P^{\mathrm{p}}} + \sum_{i \in \mathcal{N}} \frac{e_{i}w_{\mathrm{min}} + c_{i}}{\alpha N_{\mathrm{t}}\sigma_{i}^{2}}.$$

The optimal solution can be easily achieved as $P^{\rm p} = \sqrt{\sum_{i \in \mathcal{N}} \sigma_{\rm n}^2 (e_i w_{\min} + c_i) (\alpha \sigma_i^4 t^{\rm p})^{-1}}$, and $E_{\rm s}^*(w_{\min}) = 2\sqrt{\sum_{i \in \mathcal{N}} \sigma_{\rm n}^2 (e_i w_{\min} + c_i) t^{\rm p} (\alpha \sigma_i^4)^{-1}} + \sum_{i \in \mathcal{N}} (e_i w_{\min} + c_i) (\alpha N_{\rm t} \sigma_i^2)^{-1}$. For notation simplicity, we define here $A = \sum_{i \in \mathcal{N}} \sigma_{\rm n}^2 e_i t^{\rm p} (\alpha \sigma_i^4)^{-1}$, $B = \sum_{i \in \mathcal{N}} \sigma_{\rm n}^2 c_i t^{\rm p} (\alpha \sigma_i^4)^{-1}$, $C = \sum_{i \in \mathcal{N}} e_i (\alpha N_{\rm t} \sigma_i^2)^{-1}$, and $D = \sum_{i \in \mathcal{N}} c_i (\alpha N_{\rm t} \sigma_i^2)^{-1}$. Then, $E_{\rm s}^*(w_{\min}) = 2\sqrt{Aw_{\min} + B} + Cw_{\min} + D$. The termination of Algorithm 1 gives us that the optimal w_{\min} should satisfy that $E = E^*(w_{\min}) = 2\sqrt{Aw_{\min} + B} + Cw_{\min} + D$. This gives us the closed-form solution of w_{\min} as $w_{\min}^* = \frac{(E - D)C + 2A - \sqrt{[(E - D)C + 2A]^2 - C^2[(E - D)^2 - 4B]}}{C^2}$.

We can now derive the closed-form solution of P^{p} and E_{i}^{t} . We observe that, in such an asymptotic case, the complexity of the approach is as low as $\mathcal{O}(N)$, which corresponds to the calculation of auxiliary variables A, B, C, and D.

IV. NUMERICAL RESULTS

In this section, we numerically evaluate the performance of Algorithm 1 to solve Problem (1) with non-linear and linear EH models. We use Matlab for performing numerical simulations. We first describe the set-ups of the simulations. Then, we test the convergence of the algorithm. Finally, we evaluate the average network sensing rates achieved by the algorithm with different network parameters.

A. Simulation Set-ups

The default set-ups of the simulations are given as follows. We deploy N = 20 sensor nodes randomly in a disk region with radius R meters. One BS is located at the centre of the region

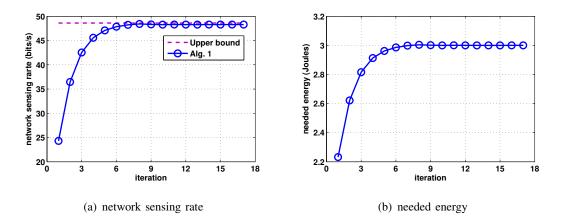


Fig. 4: Convergence of Algorithm 1 (non-linear energy harvesting case).

to transmit energy and to collect data. The BS has $N_{\rm t} = 32$ antennas, and it has 3 Joules energy available in a time block of 1 second length. The frequency of the RF energy carrier is 915 MHz. The path-loss depends on the distance between the BS and the node and is calculated according to the Friis equation. In the non-linear harvesting model, the RF-DC conversion rate function is $\eta(x) = P_{\rm max}(1 - \exp(-\eta_{\rm max}x/P_{\rm max}))$, where $P_{\rm max} = 20$ mW corresponds to the saturation of the RF-DC conversion power and $\eta_{\rm max} = 0.3$ is the maximum RF-DC conversion rate. This non-linear model captures the saturation behaviour of the EH. For the linear model, we set the RF-DC conversion rate to be $\alpha = 0.3$ by default.

Each sensor node transmits data at the standard 2.4 GHz frequency. In each time block, the energy consumption to transmit a unit size data to the BS is $10^{-7}d^2$ Joules, where d is the distance of the node to the BS. The static energy consumption of a node is $c_j = 3 \times 10^{-6}$ Joules. The time duration of channel estimation, $t^{\rm p}$, is 10% of the time block. We use a simple least-square estimator to obtain CSI, with a noise level, $\sigma_{\rm n}^2$, at -90 dBm, and then apply maximum ratio transmit beamforming to send energy. The reason we use this PEB scheme is that we have a good closed-form approximation to the PEB gain $g(\cdot)$. However, it should be noticed that, the proposed Algorithm 1 can be applied to different PEB schemes, as long as the PEB gain meets the qualification condition of Assumption 1.

B. Convergence Tests

1) Non-linear energy harvesting model: First, we will show the convergence of Algorithm 1 for the non-linear EH model. The nodes are deployed in an annulus region with inner radius 25 meters and outer radius 50 meters. The termination parameter ε in Line 6 of the algorithm

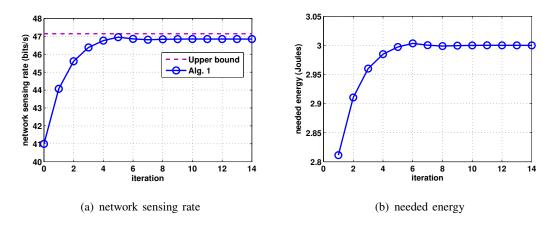


Fig. 5: Convergence of Algorithm 1 (linear energy harvesting case).

is set to be 0.001 bit/s. Fig. 4 shows w_{\min} and the corresponding needed energy $E_s^*(w_{\min})$ achieved by Algorithm 1 in each iteration step. Recall that the optimal solution should satisfy that $E_s^*(w_{\min}) = E$. We have that E_s should converge to E = 3 Joules. Initially, the sensing rate is w = 24.3 bits/s, and the corresponding energy is 2.23 Joules. Thus, in the second iteration, the threshold sensing rate increases to 36.46 bits/s, and it keeps increasing until the 8th iteration, where the needed energy is slightly above 3 Joules. Then, the threshold sensing rate starts decreasing. The algorithm terminates at the 17th iteration, where the resulting sensing rate is 48.29 bits/s. The resulting energy is slightly less than 3 Joules, which indicates the sensing rate is near optimal and feasible. The resulting sensing rate is close to the upper bound achieved according to Lemma 1, as shown by the purple dashed line. We can also see that the total number of iterations is not too large. Thus, the algorithm is efficient to achieve a near optimal solution.

2) Linear energy harvesting model: Next, we show the convergence of Algorithm 1 in terms of network sensing rate for the problem with linear EH model. The convergence of the sensing rate and the required energy in a simulation case is shown in Fig. 5. The result is similar to the non-linear EH case in terms of convergence. More specifically, the initial sensing rate is w = 41 bits/s, and the corresponding energy is 2.811 Joules. The sensing rate increases accordingly to 44.07 bits/s until the 6th iteration, where the needed energy is slightly above 3 Joules. Then, the threshold sensing rate starts to decrease. The algorithm terminates at the 15th iteration, where the resulting sensing rate is 46.85 bits/s and it is close to the upper bound indicated by the purple dashed line.

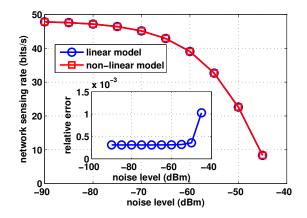


Fig. 6: The network sensing rates achieved by Algorithm 1, and the relative difference between the non-linear and linear model.

C. Comparing Non-linear and Linear Models

In this subsection, we will compare the results of Algorithm 1 for the problem with non-linear EH model and the one with linear EH model, to see whether the non-linear model we use here can be simplified to a linear model. We run simulation for 1000 times for different noise levels in channel estimation. In each simulation instance, we set the same network parameters, including the location of the nodes, the wireless channels, and the energy consumption model for the linear model and the non-linear model cases. The only difference is the EH model. We check the relative error of the results of using Algorithm 1 to solve the problem with the non-linear EH model and the ones with the linear model, i.e., $|w_{\min,NL} - w_{\min,L}|/w_{\min,NL}$, where $w_{\min,NL}$ and $w_{\min,L}$ are the results for the non-linear and linear model respectively. The result is shown in Fig. 6. The blue line with circles and the red line with square marks represents the network sensing rate of the non-linear and the linear EH model, respectively, under different noise level in channel estimation. The two curves are close to each other, thus in the mini figure we show the relative error of the network sensing rate between the non-linear and linear model. We observe that, the relative error is less than 0.15%, which means that the solution for the linear case is close to the solution for the non-linear case. The main reason comes from the fact that, in our simulation, the optimal pilot power for the non-linear EH model cases lies in a region where the EH model has a good linearity, and far from saturation. If we use other non-linear EH models or different network set-ups, the difference may be large. However, investigating which EH models can be well approximated by a linear model is beyond the scope of this work.

D. Performance Tests

To evaluate the performance of Algorithm 1, we make simulations with different parameters, such as network size and noise level of channel estimation. For each combination of parameter, we simulate 1000 times with different deployments of the sensor nodes, and take the average.

The performance is compared to the upper bound achieved by Lemma 1, a sensing rate achieved by a random based power allocation, a rate achieved by a fixed power allocation [10], [15], and a rate achieved by energy broadcasting, where there is no energy beamforming. The upper bound corresponds to the solution of Problem (4), which is the case where the BS has perfect CSI a priori. For the random based power allocation, the idea is that the BS first allocates the power for channel estimation P^{p} randomly, and then finds the solution of Problem (1) with P^{p} fixed. The fixed power allocation is similar to the random based power allocation, where the difference is that the BS always uses a fixed ratio of the total power (in the simulation we use 10%) for channel estimation. Regarding the energy broadcasting case, the base station spends no power in channel estimation and just broadcasts energy with a fixed power of 3 Watts. As default, we use purple dashed lines to represent the upper bound of the sensing rate, blue lines with circle marks for the sensing rate achieved by Algorithm 1, red lines with square marks for the rate achieved by the fixed power allocation, green lines with crosses for the rate achieved by the random power allocation, and yellow lines with diamond marks for the rate achieved by energy broadcasting.

1) Identical Expected Channel Estimation Gain: Since the linear model is a good approximation of the non-linear model, we consider the problem with linear model here for simplicity. Recall in Section III-C, we studied a special case where the BS uses LS Estimator for channel estimation. For the case where the expected channel estimation gains are identical, i.e., $\sigma_{h_i}^2 = \sigma_{h_k}^2, \forall i \neq k$, we have a closed-form solution for Equation (8), and it reduces the complexity of Algorithm 1 to $\mathcal{O}(N \log(E/(N\varepsilon)))$. Therefore, we first study the performance of Algorithm 1 in such an special case with different radius R, and different numbers of nodes N.

For the case with different R, we vary it from 10 to 50. N is fixed to be 20, and the distance of each node to the BS is R. The result is shown in Fig. 7(a), where the X axis stands for the radius R, the Y axis stands for the network sensing rate, and the upper bound of the rate is achieved by Equation (9). In general, with larger distance from BS to the nodes, the energy received by each node drops significantly. Thus, the sensing rate decreases sharply. We observe that, the rate

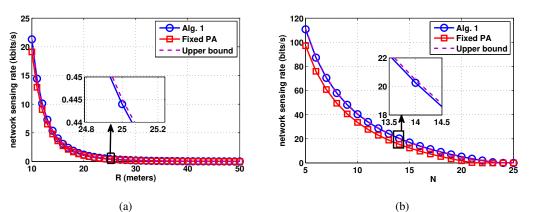


Fig. 7: (a) Comparison of Algorithm 1 to other approaches with different radius R; (b) Comparison of Algorithm 1 to other approaches with different numbers of nodes N.

achieved by energy broadcasting is 0 even when R is 10 meters. This means that the energy received by the nodes is smaller than the constant energy consumption c_i . However, if we use energy beamforming, the nodes can harvest more energy then the constant energy consumption. When R is 10 meters, the rate achieved by Algorithm 1 is approximately 21 kbits/s. Even when R is 50 meters, the rate achieved by Algorithm 1 is 5 bits/s, which means that the power is still enough for the applications that require low sensing rate, such as agricultural monitoring, etc. Besides, the results show that, the rate achieved by Algorithm 1 is very close to the theoretical upper bound, as can be seen in the mini figure, and the relative difference is approximately 0.2%, which is small enough. Compared to the fixed power allocation, the rate achieved by Algorithm 1 when R = 14 is approximately 10% higher, and it is approximately 30% higher when R = 40.

For the case with different N, we vary it from 5 to 25. The radius of the filed is R = 50 meters. We compare the performance of Algorithm 1 to those achieved by the fixed power allocation, and the upper bound. The simulation results is shown in Fig. 7(b). In general, the energy transmitted to each node, E_i^t is inversely proportional to N. Thus, when N increases, each node receives less energy, and thus the sensing rate reduces. The rate achieves by Algorithm 1 is close to the theoretical upper bound, and it is approximately 10% larger than the one achieved by the fix power allocation. When N is larger than 24, the harvested energy of the nodes is smaller than c_i , and the sensing rate becomes 0. We can conclude that, the proposed Algorithm 1 achieves a near optimal solution, and the upper bound is tight when the noise level for channel estimation is low. Also, with the power allocation achieved from Algorithm 1, the WPSN with even one BS that provides energy can monitor a region of size of several hundred meters square with sensing rate at approximately several bits per second.

2) General Cases: Now we will consider the general cases where the channel estimation gain is not necessarily identical for all the nodes. The nodes are randomly deployed in a disk region with R = 50 meters.

To begin with, we test Algorithm 1 with different noise levels. For the same node deployment and channel state, we vary the noise level from -90 dBm to -45 dBm, and compare the network sensing rate achieved by different algorithms. With lower noise level, the channel estimation is more power efficient. The simulation result is shown in Fig. 8(a). In general, with higher noise level, the channel estimation error becomes larger. Thus, the BS tends to spend more energy in channel estimation, resulting in less energy that is transmitted to the nodes. Therefore, the sensing rate that is achieved by Algorithm 1 decreases from 47 bits/s to 7 bits/s when the noise level increases from -90 dBm to -45 dBm. By comparing the performance to the upper bound, we can see that the gap between the sensing rate achieved by Algorithm 1 and the upper bound becomes larger when the noise level is high. Therefore, the upper bound is loose at high noise level. For the performance of fixed power allocation, it can be observed that, the sensing rate is close to the one achieved by Algorithm 1 when the noise level is around -55 to -50 dBm. It means that, in these cases the pilot power of the fixed power scheme, 0.3 Watts, happens to be close to the optimal one. However, for other noise level, it is much worse than Algorithm 1. The rate drops to approximately 3 bits/s when the noise level is at -45 dBm. Besides, the sensing rate achieved by random power allocation is even worse than the fixed power allocation. It is approximately only one fourth to the rate achieved by Algorithm 1.

Next, we also compare the sensing rate achieved by different algorithms with different static power consumption c, as shown in Fig. 8(b). The noise level is -90 dBm. It can be seen that, the network sensing rate decreases linearly with c, until when it reaches 0. When c = 0, the rate achieved by energy broadcasting is approximately 1.2 bits/s. However, when c is slightly larger than 0, the harvested energy from energy broadcasting can not support the static power consumption, and thus the nodes that are far away from the BS have no residual power to transmit the data. With energy beamforming, the BS can better allocate the power to different nodes. Therefore, it may transmit more energy to the nodes that are further away than the nearby nodes, and the network sensing rate is much larger than the case by energy broadcasting. We observe that, under different c, the network sensing rate is always close to the upper bound, and it is basically linearly decreasing with c.

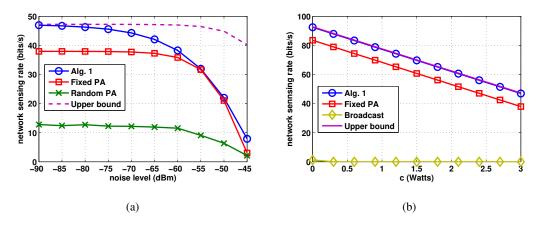


Fig. 8: (a) Comparison of Algorithm 1 to other approaches with different noise levels; (b) Comparison of Algorithm 1 to other approaches with different static power consumptions c.

To summarize, the simulation results show the convergence of Algorithm 1. Its performance is close to the upper bound if we have large power efficiency in channel estimation. Also, Algorithm 1 outperforms other power allocations scheme in terms of the sensing rate.

V. CONCLUSIONS AND FUTURE WORKS

We studied a wirelessly-powered sensor network where a BS with multiple antennas needs to optimally allocate its power in pilot transmission, and in energy beamforming to a set of sensor nodes. We considered a non-linear EH of the nodes, and represented the model by a generic class of non-linear functions. We showed that this optimal power allocation is a non-convex optimization problem. We proposed an algorithm to achieve a near optimal solution of the problem. We studied the convergence properties of the proposed algorithm and showed in the simulation that the solutions achieved by the proposed algorithm are close to the theoretical upper-bound. Moreover, the solutions improve the network sensing rate by around 10% compared to the existing fixed power allocation approach.

In the future, we will study the case where the channel states are time-correlated, and thus the base station can update the power allocation adaptively. Another interesting topic is to jointly consider the data routing decision in the problem. We are also interested in the cases with multiple BSs, where the BSs may perform cooperative channel estimation and beamforming.

APPENDIX A

PROOFS

A. Proposition 1

Recall that function g_i is monotone increasing and positive, and its first derivative is monotone decreasing. Recall that $f_i(P^p; w_{\min}) \triangleq \eta^{-1}(e_i w_{\min} + c_i)/(g_i(P^p))$. Thus, we have that $f_i > 0$,

$$\frac{df_i(P^{\rm p}; w_{\rm min})}{dP^{\rm p}} = -\eta^{-1} (e_i w_{\rm min} + c_i) \frac{g_i'(P^{\rm p})}{g_i^2(P^{\rm p})} \le 0,$$

and

$$\frac{d^2 f_i(P^{\rm p}; w_{\rm min})}{d(P^{\rm p})^2} = \eta^{-1} (e_i w_{\rm min} + c_i) \frac{2 \left(g_i'(P^{\rm p})\right)^2 - g_i''(P^{\rm p})g_i^2(P^{\rm p})}{g_i^4(P^{\rm p})} \ge 0,$$

i.e., f_i is monotone decreasing, positive, and its first derivative is monotone increasing w.r.t. P^p . Thus, $E_s(P^p|w_{\min})$ is convex, and the problem is a convex optimization problem. Thus, we have that

$$\frac{dE_{\rm s}(P^{\rm p}|w_{\rm min})}{dP^{\rm p}} = t^{\rm p} + \sum_{i\in\mathcal{N}} f_i'(P^{\rm p}|w_{\rm min}) \,.$$

Then, we have that the first derivative of $E_s(P^p|w_{\min})$ satisfies one of the following cases: 1) it is always positive; 2) it has a root; 3) it is always negative.

The first case gives us that the optimal solution is $P^{p} = 0$. The second case gives us that the optimal solution P^{p} satisfies

$$\sum_{i \in \mathcal{N}} f'_i(P^{\mathbf{p}} | w_{\min}) = -t^{\mathbf{p}} \,. \tag{11}$$

The third case means the optimal $P^{p} \to +\infty$, which means that the given w_{\min} is not achievable and thus we discard this case. To conclude, given w_{\min} , the optimal training power P^{p} such that E is minimized, is either $P^{p} = 0$, or it satisfies Equation (11). This completes the proof.

B. Lemma 1

We start the proof with comparing Problem (1) and Problem (4). Recall that $\eta_{\max} \triangleq \max_{x\geq 0} \eta(x)/x$. We have that the right hand side of Constraint (4b), $\eta_{\max} E_i^{t} g_i (E/t^{p})$ is larger than or equal to $\eta(E_i^{t} g_i (E/t^{p}))$. Together with Constraint (1d), we have that $\eta_{\max} E_i^{t} g_i (E/t^{p})$ is larger than or equal to $\eta(E_i^{t} g_i (P^{p}))$, which is the right hand side of Constraint (1c). Therefore, Constraint (4b) is a relaxation of Constraint (1c). Additionally, we have that Constraint (4c) is a relaxation of Constraint (1d). Therefore, we have that Problem (4) is a relaxation of Problem (1). Since the problems are maximization problems, we have that the optimum of Problem (1) is upper bounded by the optimum of Problem (4).

Next, we will find an upper bound of the optimum of Problem (4). Briefly speaking, Constraint (4b) gives us that the optimal w_{\min} should satisfy $(e_i w_{\min} + c_i)/(\eta_{\max} g_i(E/t^p)), \forall i \in \mathcal{N}$. Together with Constraint (4c), we have that

$$\sum_{i \in \mathcal{N}} \frac{e_i}{g_i(\frac{E}{t^{\mathrm{p}}})} w_{\min} + \sum_{i \in \mathcal{N}} \frac{c_i}{g_i(\frac{E}{t^{\mathrm{p}}})} = \eta_{\max} E \,,$$

whose solution gives us the right hand side of Equation (5). Thus, we have that such an upper bound of the optimum of Problem (4) is also an upper bound of that of Problem (1), which completes the proof.

C. Proposition 2

Denote w_{\min}^{o} , $E^{t,o}$, $P^{p,o}$ the optimal solution of Problem (1). In the initialization of Algorithm 1, if Problem (1) is feasible, then we have that $w_{\min}^{l} = 0 \le w_{\min}^{o} < w_{\min}^{u}$. We will prove that, after each iteration of Algorithm 1, w_{\min}^{o} always lies within the range $[w_{\min}^{l}, w_{\min}^{u}]$. This is equivalent to prove that, for a w_{\min} , if $E_{s}^{*}(w_{\min}) - E > 0$, then $w_{\min} > w_{\min}^{o}$; otherwise, $w_{\min} \le w_{\min}^{o}$.

We prove that if $E_s^*(w_{\min}) - E > 0$, then $w_{\min} > w_{\min}^o$, by contradiction. Assume that in this case $w_{\min} \le w_{\min}^o$. Since w_{\min}^o is feasible. We have that $E \ge E_s^*(w_{\min}^o) = t^p P^p(w_{\min}^o) + \sum_i f_i(P^p(w_{\min}^o); w_{\min}^o)$, where $P^p(w_{\min}^o)$ is the optimal P^p for Problem (2) when problem input w_{\min} is w_{\min}^o . Recall that function $f_i(P^p; w)$ is monotone increasing with w. Thus, $f_i(P^p(w_{\min}^o); w_{\min}^o) \ge$ $f_i(P^p(w_{\min}^o); w_{\min})$. Therefore, $E_s^*(w_{\min}^o) = t^p P^p(w_{\min}^o) + \sum_i f_i(P^p(w_{\min}^o); w_{\min}^o) \ge t^p P^p(w_{\min}^o) + \sum_i f_i(P^p(w_{\min}^o); w_{\min}) \ge P^p(w_{\min}) + \sum_i f_i(P^p(w_{\min}^o); w_{\min}) \ge E$, which contradicts to that $E^*(w_{\min}^o) \le E$.

Similarly, we can prove that if $E_{\rm s}^*(w_{\rm min}) - E \leq 0$, then $w_{\rm min} \leq w_{\rm min}^{\rm o}$. Thus, we can conclude that after each iteration of Algorithm 1, $w_{\rm min}^{\rm o}$ always lies within the range $[w_{\rm min}^{\rm l}, w_{\rm min}^{\rm u})$. Furthermore, since $w_{\rm min}^{\rm l}$ is always feasible in each iteration, and we set the $w_{\rm min}$ to be $w_{\rm min}^{\rm l}$, we have that the output of the algorithm is feasible. Moreover, since the algorithm terminates when $w_{\rm min}^{\rm u} - w_{\rm min}^{\rm l} \leq \varepsilon$, we have that $\varepsilon \geq w_{\rm min}^{\rm o} - w_{\rm min}^{\rm l} = w_{\rm min}^{\rm o} - w_{\rm min}$, where the equality comes from Line 19 of the algorithm. This completes the proof.

D. Proposition 3

Recall that the algorithm is based on bisection search. Therefore, the number of iterations depends on the initial range, which is $[0, w_{\min}^{u}]$, and the termination condition ε . Recall that the initial w_{\min}^{u} is the optimal solution of Problem (4). We can achieve an upper bound of w_{\min}^{u} as $E\eta_{\max}/(N\min\{e_i/g_i\})$, which is $\mathcal{O}(E/N)$. Therefore, the number of iteration is $\mathcal{O}(\log(E/(N\varepsilon)))$. In each iteration, the algorithm achieves $E_{s}^{*}(w_{\min})$ according to Proposition 1, whose complexity is $\mathcal{O}(N\log(E))$. Thus, the total complexity is $\mathcal{O}(N\log(E)\log(E/(N\varepsilon)))$.

E. Remark 7

We denote $\alpha = P^{p}/N_{t}$. Then, we have that when N_{t} is fixed, the monotonicity and concavity of a function of P^{p} is the same to that of α . Here, we will analyze how $g_{i}(P^{p})$ behaves when N_{t} is sufficiently large. From Equation (6) we have that

$$\lim_{N_{t} \to +\infty} g_{i}(P^{p}) = \lim_{N_{t} \to +\infty} g_{i}(\alpha N_{t})$$
(12a)

$$= \lim_{N_{t} \to +\infty} \mathbb{E} \left[\frac{\alpha N_{t} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i} + 2\sqrt{\alpha} N_{t} \boldsymbol{n}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i} + N_{t} \boldsymbol{n}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{n}_{i}}{\alpha N_{t} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i} + 2\sqrt{\alpha} N_{t} \boldsymbol{h}_{i}^{H} \boldsymbol{n}_{i} + N_{t} \boldsymbol{n}_{i}^{H} \boldsymbol{n}_{i}} \right]$$
(12b)

$$= \mathbb{E}\left[\lim_{N_{t}\to+\infty} \frac{\alpha \frac{\boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i}}{N_{t}^{2}} + 2\sqrt{\alpha} \frac{\boldsymbol{n}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i}}{N_{t}^{2}} + \frac{\boldsymbol{n}_{i}^{H} \boldsymbol{h}_{i} \boldsymbol{h}_{i}^{H} \boldsymbol{n}_{i}}{N_{t}^{2}} N_{t}}{\frac{\alpha \boldsymbol{h}_{i}^{H} \boldsymbol{h}_{i}}{N_{t}} + 2\sqrt{\alpha} \frac{\boldsymbol{h}_{i}^{H} \boldsymbol{n}_{i}}{N_{t}} + \frac{\boldsymbol{n}_{i}^{H} \boldsymbol{h}_{i}}{N_{t}}} N_{t}}\right].$$
 (12c)

Recall that $\mathbf{h}_i = [h_{i1}, \dots, h_{iN_t}]^T$, and the element of \mathbf{h}_i are i.i.d.. Thus, according to the random matrix theory [46], we have that $\lim_{N_t \to +\infty} \mathbf{h}_i^H \mathbf{h}_i / N_t \xrightarrow{\mathbf{a.s.}} \sigma_i^2$. Similarly, we have that $\lim_{N_t \to +\infty} \mathbf{n}_i^H \mathbf{n}_i / N_t \xrightarrow{\mathbf{a.s.}} \sigma_n^2$. Since \mathbf{h}_i and \mathbf{n}_i are mutually independent, we have that $\lim_{N_t \to +\infty} \mathbf{h}_i^H \mathbf{n}_i / N_t \xrightarrow{\mathbf{a.s.}} 0$. Thus, we can continue to simplify Equation (12a) as follows:

$$\lim_{N_{t}\to+\infty} g_{i}(P^{p}) \xrightarrow{\mathbf{a.s.}} \mathbb{E}\left[\lim_{N_{t}\to+\infty} N_{t} \frac{\alpha \sigma_{i}^{4}}{\alpha \sigma_{i}^{2} + \sigma_{n}^{2}}\right] = \frac{\alpha \sigma_{i}^{4}}{\alpha \sigma_{i}^{2} + \sigma_{n}^{2}} \lim_{N_{t}\to+\infty} N_{t}$$

By replacing α by $P^{\rm p}/N_{\rm t}$, we have that $g_i(P^{\rm p})$ converges almost surely to $N_{\rm t}\sigma_i^4 P^{\rm p}(\sigma_i^2 P^{\rm p} + N_{\rm t}\sigma_{\rm n}^2)^{-1}$. One can easily check that the function is monotone increasing and concave w.r.t. $P^{\rm p}$ by checking its derivatives. This completes the proof.

APPENDIX B

DISCUSSION ON THE FEASIBILITY OF ASSUMPTION 2 FOR ADAPTIVE WAVEFORM DESIGN

When the BS can transmit energy with different sinewaves, it could allocate power to different sinewaves according to the CSI. This scheme is called adaptive waveform deign, and is shown to

be better than non-adaptive ones in terms of harvested energy in [38], especially under the nonlinear RF-DC model. Here, we will show that, using some adaptive waveform design schemes, the energy transmission using multisine carriers and the non-linear RF-DC model will satisfy our Assumption 2.

To illustrate the idea, we reuse the toy-example in [38, Section IV.B]: There are two sinewaves, with $s_0 > 0$, $s_1 > 0$ be the amplitude of the sinewaves, and $A_0 > 0$, $A_1 > 0$ the channels amplitude of the two sinewaves. Reference [38] shows that, to maximizing the output DC is (approximately) equivalent to maximizing

$$z_{DC}(s_0, s_1) = \tilde{k}_2(s_0^2 A_0^2 + s_1^2 A_1^2) + \tilde{k}_4[(s_0^2 A_0^2 + s_1^2 A_1^2)^2 + 2s_0^2 s_1^2 A_0^2 A_1^2]$$

where \tilde{k}_2 and \tilde{k}_4 are two positive coefficients that are related to the RF-DC circuit.

Our Assumption 2 requires that, if the received RF power at the antenna, $P_{ave,rf}$, increases to $P'_{ave,rf} > P_{ave,rf}$, the harvested power increases, and so as the corresponding $z_{DC}(s_0, s_1)$, where $P_{ave,rf} = 0.5(s_0^2 A_0^2 + S_1^2 A_1^2)$. This may not necessarily be true in general, though there exist some waveform design schemes for which this assumption holds:

- Allocates power to the strongest frequency component with adaptive single sinewave (ASS) strategy in [38]: Without loss of generality, assume that A₀ > A₁, i.e., frequency 0 has stronger amplitude than frequency 1. Then, P'_{ave,rf} > P_{ave,rf} indicates that s'₀ > s₀ and s₁ = s'₁ = 0. This yields z_{DC}(s₀, s₁) < z_{DC}(s'₀, s'₁), i.e., the harvested power increases.
- 2) Allocating same power ratio to two frequency components, i.e., s'₀/s'₁ = s₀/s₁: If the ratio depends on the channel strength, then it is the adaptive matched filter in [38]. Otherwise, it is a non-adaptive scheme (e.g. it is the uniform power scheme in [38] if the ratio is 1). In this case, P'_ave,rf > P_ave,rf indicates that s'₀ > s₀ and s₁ > s'₁ (since s'₀/s'₁ = s₀/s₁). This again yields z_{DC}(s₀, s₁) < z_{DC}(s'₀, s'₁).
- 3) Optimal scheme that maximizes z_{DC} : In this case, $(s_0, s_1) = \arg \max_{s_0, s_1} \{z_{DC}(s_0, s_1) | s_0^2 A_0^2 + S_1^2 A_1^2 = 2P_{ave,rf}\}$, and $(s'_0, s'_1) = \arg \max_{s_0, s_1} \{z_{DC}(s_0, s_1) | s_0^2 A_0^2 + S_1^2 A_1^2 = 2P'_{ave,rf}\}$. Since $P'_{ave,rf} > P_{ave,rf}$, we can construct a feasible point (s''_0, s_1) by setting

$$s_0'' = \sqrt{\frac{2(P'_{ave,rf} - P_{ave,rf})}{A_0^2}} + s_0^2,$$

such that $(s_0'')^2 A_0^2 + s_1^2 A_1^2 = 2P'_{ave,rf}$ and $s_0'' > s_0$. We have that $z_{DC}(s_0'', s_1) > z_{DC}(s_0, s_1)$. Meanwhile, since (s_0'', s_1) is a feasible point of $\arg \max_{s_0,s_1} \{z_{DC}(s_0, s_1) | s_0^2 A_0^2 + S_1^2 A_1^2 = 2P'_{ave,rf}\}$, we have that $z_{DC}(s_0'', s_1) \le z_{DC}(s_0', s_1')$. Thus again, $z_{DC}(s_0, s_1) < z_{DC}(s_0', s_1')$. These examples show that, the non-linear RF-DC model with some adaptive but suboptimal waveform design schemes could still satisfy Assumption 2. It also works for some non-adaptive schemes. However, there are some other waveform design schemes that may fail, such as randomly assigning the power to different frequencies (which is non-adaptive and suboptimal), and some adaptive but suboptimal schemes (for which we could not prove that $z_{DC}(s_0, s_1) < z_{DC}(s'_0, s'_1)$ when $P_{ave,rf} < P'_{ave,rf}$).

APPENDIX C

CONCAVITY OF PEB GAIN FOR LS CASES

We study the concavity of the PEB gain $g_i(P^p)$ in (6) and achieve the following result:

Proposition 4. $g_i(P^p)$ in Equation (6) is concave almost for sure in the region $P^p \ge (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)(N_t\sigma_n^2/\mathbf{h}_i^H\mathbf{h}_i)$, where $2Q^{-1}(N_t, 0, 0.99)$ is the inverse of the generalized regularized incomplete gamma function that represents the 99-th percentile from the lower cumulative distribution of the Chi-squared distribution with degree of freedom $2N_t$ (corresponding to the real parts and the imaginary parts of the noises), i.e., the probability that $\|\mathbf{n}\|^2 \le Q^{-1}(N_t, 0, 0.99)\sigma_n^2$ is 99%.

Proof: Recall that $g_i(P^p) = \mathbb{E}\left[\|\sqrt{P^p} \boldsymbol{h}_i^H \boldsymbol{h}_i + \sqrt{N_t} \boldsymbol{h}_i^H \boldsymbol{n}_i \|^2 / \|\sqrt{P^p} \boldsymbol{h}_i + \sqrt{N_t} \boldsymbol{n}_i \|^2 \right]$. Since the noise \boldsymbol{n} is zero mean and Gaussian, we have that $\mathbb{E}\left[\|\sqrt{P^p} \boldsymbol{h}_i^H \boldsymbol{h}_i + \sqrt{N_t} \boldsymbol{h}_i^H \boldsymbol{n}_i \|^2 / \|\sqrt{P^p} \boldsymbol{h}_i + \sqrt{N_t} \boldsymbol{n}_i \|^2 \right] = \mathbb{E}\left[\|\sqrt{P^p} \boldsymbol{h}_i^H \boldsymbol{h}_i - \sqrt{N_t} \boldsymbol{h}_i^H \boldsymbol{n}_i \|^2 / \|\sqrt{P^p} \boldsymbol{h}_i - \sqrt{N_t} \boldsymbol{n}_i \|^2 \right]$. Thus,

$$g_i(P^{\mathrm{p}}) = \frac{1}{2} \int_{\boldsymbol{n}} \left[\underbrace{\frac{\|\sqrt{P^{\mathrm{p}}} \boldsymbol{h}_i^H \boldsymbol{h}_i + \sqrt{N_{\mathrm{t}}} \boldsymbol{h}_i^H \boldsymbol{n}_i\|^2}{\|\sqrt{P^{\mathrm{p}}} \boldsymbol{h}_i + \sqrt{N_{\mathrm{t}}} \boldsymbol{n}_i\|^2}}_{\triangleq G_i(P^{\mathrm{p}})} + \frac{\|\sqrt{P^{\mathrm{p}}} \boldsymbol{h}_i^H \boldsymbol{h}_i - \sqrt{N_{\mathrm{t}}} \boldsymbol{h}_i^H \boldsymbol{n}_i\|^2}{\|\sqrt{P^{\mathrm{p}}} \boldsymbol{h}_i - \sqrt{N_{\mathrm{t}}} \boldsymbol{n}_i\|^2}} \right] p_r(\boldsymbol{n}) \mathrm{d}\boldsymbol{n}$$

To prove the concavity of $g_i(P^p)$, we can check the concavity of $G_i(P^p)$. For the sake of simplicity, we denote $A = \|\mathbf{h}\|^2$, $B = \sqrt{N_t} \operatorname{Re}\{\mathbf{h}^H \mathbf{n}\}$, $C = N_t \|\mathbf{n}\|^2$ (note that A, B, and C are different to the ones of Section III-E). We also discard the subscript i, and further simplify it to

$$G(x) = \frac{\sqrt{x}A^2 + 2\sqrt{x}AB + N_{\rm t} \|\boldsymbol{h}^H \boldsymbol{n}\|^2}{Ax + 2\sqrt{x}B + C} + \frac{\sqrt{x}A^2 - 2\sqrt{x}AB + N_{\rm t} \|\boldsymbol{h}^H \boldsymbol{n}\|^2}{Ax - 2\sqrt{x}B + C}.$$

Then, the second derivative of G(x) is

$$G''(x) = \frac{d^2 G(x)}{dx^2} = \frac{4(N_{\rm t} \| \boldsymbol{h}^H \boldsymbol{n} \|^2 - AC)}{(\|\sqrt{x} \boldsymbol{h}_i + \sqrt{N_{\rm t}} \boldsymbol{n}_i \|^2 \| \sqrt{x} \boldsymbol{h}_i - \sqrt{N_{\rm t}} \boldsymbol{n}_i \|^2)^3} G_2(x)$$

where $G_2(x) = A^5 x^3 + 3A^4 C x^2 + 3A^3 C^2 x - 12A^2 B^2 C x + A^2 C^3 - 12AB^2 C^2 + 16B^4 C$. The denominator of G''(x) is positive. Also, according to the Cauchy–Schwarz inequality, we have that $N_t \| \mathbf{h}^H \mathbf{n} \|^2 - AC \leq 0$, and thus $B^2 \leq AC$. Then, we know that G(x) is concave when $G_2(x)$ is positive. With $B^2 \leq AC$, we have that $G_2(x) \geq A^5 x^3 + 3A^4 C x^2 - 9A^3 C^2 x - 11A^2 C^3 + 16B^4 C$. Let $k \triangleq Ax/C$. We have that $G_2(x) \geq A^2 C^3 (k^3 + 3k^2 - 9k - 11) + 16B^4 C = A^2 C^3 (k+1)(k+$ $1 + 2\sqrt{3})(k - (2\sqrt{3} - 1)) + 16B^4 C$. Recall that $A \geq 0$, $C \geq 0$. We have that $G_2(x) \geq 0$ when $k \geq 2\sqrt{3} - 1$. Recall $C = N_t \| \mathbf{n} \|^2$, where $\| \mathbf{n} \|^2$ follows the Chi-squared distribution with degree of freedom $2N_t$, and the 99-th percentile of its lower cumulative distribution is $Q^{-1}(N_t, 0, 0.99)\sigma_n^2$. Then, we have that when $P^p \geq (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)N_t\sigma_n^2/A$, it is almost for sure that $G_i(P^p)$ is concave for all realizations of the noises. Thus, when $P^p \geq (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)N_t\sigma_n^2/A$, it is almost for sure that $G_i(P^p)$ is concave for all realizations of the noises. Thus, when $P^p \geq (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)N_t\sigma_n^2/A$, it is almost for sure that $G_i(P^p)$ is concave for all realizations of the noises. Thus, when $P^p \geq (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)N_t\sigma_n^2/A$, it is almost for sure that $G_i(P^p)$ is concave for all realizations of the noises. Thus, when $P^p \geq (2\sqrt{3} - 1)Q^{-1}(N_t, 0, 0.99)N_t\sigma_n^2/A$, it is almost for sure that $G_i(P^p)$.

Notice that, the threshold is sufficiently small and is practical for real-world use. For example, let σ^2 be -120 dBm, and $N_t = 16$. We have that $Q^{-1}(N_t, 0, 0.99) \approx 26.74$. Then, if the power of the received N_t pilots is larger than approximately -90 dBm, then $g(\cdot)$ is concave, i.e., the required transmission power of the pilots is also very low (P^p should be larger than -20 dBm if the pathloss is 70 dB). Therefore, the requirement is easy to fulfil.

REFERENCES

- S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Communications Surveys* & *Tutorials*, vol. 13, no. 3, pp. 443–461, 2011.
- [2] L. Xie, Y. Shi, Y. T. Hou, and A. Lou, "Wireless power transfer and applications to sensor networks," *IEEE Wireless Communications*, vol. 20, no. 4, pp. 140–145, 2013.
- [3] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and signals design for wireless power transmission," *IEEE Trans. Communications*, vol. 65, no. 5, pp. 2264–2290, 2017.
- [4] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 757–789, 2015.
- [5] C. Zhong, X. Chen, Z. Zhang, and G. K. Karagiannidis, "Wireless-powered communications: Performance analysis and optimization," *IEEE Trans. Communications*, vol. 63, no. 12, pp. 5178–5190, 2015.
- [6] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 117–125, 2015.
- [7] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 104–110, 2014.
- [8] S. Bi, Y. Zeng, and R. Zhang, "Wireless powered communication networks: An overview," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 10–18, 2016.

- [9] Q. Sun, G. Zhu, C. Shen, X. Li, and Z. Zhong, "Joint beamforming design and time allocation for wireless powered communication networks," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1783–1786, 2014.
- [10] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan, "Throughput optimization for massive MIMO systems powered by wireless energy transfer," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 8, pp. 1640–1650, 2015.
- [11] H. Lee, K.-J. Lee, H.-B. Kong, and I. Lee, "Sum-rate maximization for multiuser MIMO wireless powered communication networks," *IEEE Trans. Vehicular Technology*, vol. 65, no. 11, pp. 9420–9424, 2016.
- [12] M. M. Sohul, M. Yao, T. Yang, and J. H. Reed, "Spectrum access system for the citizen broadband radio service," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 18–25, 2015.
- [13] O. Galinina, H. Tabassum, K. Mikhaylov, S. Andreev, E. Hossain, and Y. Koucheryavy, "On feasibility of 5G-grade dedicated rf charging technology for wireless-powered wearables," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 28–37, 2016.
- [14] Y. Alsaba, S. K. A. Rahim, and C. Y. Leow, "Beamforming in wireless energy harvesting communications systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1329–1360, 2018.
- [15] L. Liu, R. Zhang, and K.-C. Chua, "Multi-antenna wireless powered communication with energy beamforming," *IEEE Trans. Communications*, vol. 62, no. 12, pp. 4349–4361, 2014.
- [16] S. Lee and R. Zhang, "Distributed wireless power transfer with energy feedback," *IEEE Trans. Signal Processing*, vol. 65, no. 7, pp. 1685–1699, 2017.
- [17] K. W. Choi, L. Ginting, P. A. Rosyady, A. A. Aziz, and D. I. Kim, "Wireless-powered sensor networks: How to realize," *IEEE Trans. Wireless Communications*, vol. 16, no. 1, pp. 221–234, 2017.
- [18] J. Xu and R. Zhang, "A general design framework for MIMO wireless energy transfer with limited feedback." *IEEE Trans. Signal Processing*, vol. 64, no. 10, pp. 2475–2488, 2016.
- [19] R. Du, A. Ozcelikkale, C. Fischione, and M. Xiao, "Optimal energy beamforming and data routing for immortal wireless sensor networks," in *Proc. IEEE International Conference on Communications*, 2017, pp. 242–247.
- [20] L. Xie, Y. Shi, Y. T. Hou, and H. D. Sherali, "Making sensor networks immortal: An energy-renewal approach with wireless power transfer," *IEEE/ACM Trans. Networking*, vol. 20, no. 6, pp. 1748–1761, 2012.
- [21] L. Xie, Y. Shi, Y. Hou, W. Lou, H. Sherali, H. Zhou, and S. Midkiff, "A mobile platform for wireless charging and data collection in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 8, pp. 1521–1533, Aug 2015.
- [22] R. Du, C. Fischione, and M. Xiao, "Lifetime maximization for sensor networks with wireless energy transfer," in *Proc. IEEE International Conference on Communications*, 2016, pp. 20–25.
- [23] S. Kashyap, E. Björnson, and E. G. Larsson, "On the feasibility of wireless energy transfer using massive antenna arrays," *IEEE Trans. Wireless Communications*, vol. 15, no. 5, pp. 3466–3480, 2016.
- [24] Y. Zeng and R. Zhang, "Optimized training design for wireless energy transfer," *IEEE Trans. Communications*, vol. 63, no. 2, pp. 536–550, 2015.
- [25] S. Abeywickrama, T. Samarasinghe, C. K. Ho, and C. Yuen, "Wireless energy beamforming using received signal strength indicator feedback," *IEEE Trans. Signal Processing*, vol. 66, no. 1, pp. 224–235, 2018.
- [26] S. Abeywickrama, T. Samarasinghe, C. Yuen, and R. Zhang, "Cluster-based wireless energy transfer for low complex energy receivers," in *Proc. International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, 2018, pp. 1–7.
- [27] J. Xu and R. Zhang, "Energy beamforming with one-bit feedback," *IEEE Trans. Signal Processing*, vol. 62, no. 20, pp. 5370–5381, 2014.

- [28] Y. Huang and B. Clerckx, "Waveform design for wireless power transfer with limited feedback," *IEEE Trans. Wireless Communications*, vol. 17, no. 1, pp. 415–429, 2018.
- [29] T.-D. Le, G. Kaddoum, and O.-S. Shin, "Joint channel resources allocation and beamforming in energy harvesting systems," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 884–887, 2018.
- [30] K. Lee and J.-P. Hong, "Energy-efficient resource allocation for simultaneous information and energy transfer with imperfect channel estimation," *IEEE Trans. Vehicular Technology*, vol. 65, no. 4, pp. 2775–2780, 2016.
- [31] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2082–2085, 2015.
- [32] B. Clerckx, "Wireless information and power transfer: Nonlinearity, waveform design and rate-energy tradeoff," *IEEE Trans. Signal Processing*, vol. 66, no. 4, pp. 847–862, 2018.
- [33] 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. Communications*, vol. 65, no. 5, pp. 1984–1999, 2017.
- [34] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and K. B. Letaief, "Coordinated beamforming with artificial noise for secure SWIPT under non-linear EH model: Centralized and distributed designs," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 7, pp. 1544–1563, 2018.
- [35] R. Du and C. Fischione, "Power allocation for channel estimation and energy beamforming in wirelessly powered sensor networks," in *Proc. IEEE International Conference on Communicationn Workshops*, 2018, pp. 1–6.
- [36] R. Du, A. Özçelikkale, C. Fischione, and M. Xiao, "Towards immortal wireless sensor networks by optimal energy beamforming and data routing," *IEEE Trans. Wireless Communications*, vol. 17, no. 8, pp. 5338–5352, 2018.
- [37] B. Clerckx and J. Kim, "On the beneficial roles of fading and transmit diversity in wireless power transfer with nonlinear energy harvesting," *IEEE Trans. Wireless Communications*, vol. 17, no. 11, pp. 7731–7743, 2018.
- [38] B. Clerckx and E. Bayguzina, "Waveform design for wireless power transfer," *IEEE Trans. Signal Processing*, vol. 64, no. 23, pp. 6313–6328, 2016.
- [39] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.
- [40] T. K. Lo, "Maximum ratio transmission," in Proc. IEEE International Conference on Communications, 1999, pp. 1310– 1314.
- [41] J.-H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Trans. on Networking*, vol. 12, no. 4, pp. 609–619, 2004.
- [42] I. Koutsopoulos and M. Halkidi, "Distributed energy-efficient estimation in spatially correlated wireless sensor networks," *Computer Communications*, vol. 45, pp. 47–58, 2014.
- [43] T. E. Bogale, L. B. Le, and X. Wang, "Hybrid analog-digital channel estimation and beamforming: Training-throughput tradeoff," *IEEE Trans. Communications*, vol. 63, no. 12, pp. 5235–5249, 2015.
- [44] N. N. Moghadam, H. S. Gadikolaei, G. Fodor, M. Bengtsson, and C. Fischione, "Pilot precoding and combining in multiuser MIMO networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1632–1648, 2017.
- [45] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [46] A. M. Tulino, S. Verdú *et al.*, "Random matrix theory and wireless communications," *Foundations and Trends* (*R) in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.