This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

Spectral and Energy Efficiencies in Full-Duplex Wireless Information and Power Transfer

Van-Dinh Nguyen, *Student Member, IEEE,* Trung Q. Duong, *Senior Member, IEEE,* Hoang Duong Tuan, Oh-Soon Shin, *Member, IEEE,* and H. Vincent Poor *Fellow, IEEE*

***Abstract*—A communication system is considered consisting of a full-duplex (FD) multiple-antenna base station (BS) and multiple single-antenna downlink users (DLUs) and single-antenna uplink users (ULUs), where the latter need to harvest energy for transmitting information to the BS. The communication is thus divided into two phases. In the first phase, the BS uses all available antennas for conveying information to DLUs and wireless energy to ULUs via information and energy beamforming, respectively. In the second phase, ULUs send their independent information to the BS using their harvested energy while the BS transmits the information to the DLUs. In the both phases, the communication is operated at the same time and over the same frequency band. The aim is to maximize the sum rate and energy efficiency under ULU achievable information throughput constraints by jointly designing beamformers and time allocation. The utility functions of interest are nonconcave and involved constraints are nonconvex, so these problems are computationally troublesome. To address them, path-following algorithms are proposed to arrive at least at local optima. The proposed algorithms iteratively improve the objectives with converge guaranteed. Simulation results demonstrate that they achieve fast convergence rate and outperform conventional solutions.**

***Index Terms*—Energy harvesting, full-duplex radios, full-duplex self-interference, transmit beamforming, wireless information and power transfer.**

I. Introduction

Radio-frequency (RF) energy harvesting (EH) communication has emerged as a promising cost-effective technology for supplying power to users [1], [2]. Enabling wireless devices to harvest energy from RF signals, RF-EH communication is expected to fundamentally reshape the landscape of power supply in Internet-of-Things (IoT) [3], [4]. Exploring RF EH communication allows one to transfer information and energy over the same RF channel [5], [6]. Various cooperative schemes with/without built-in batteries for energy storage in which multiple transceiver pairs communicate with each other via EH relays were studied in [7]–[10]. In this regard, transmit beamforming is used to focus information and/or RF energy at the desired users [11], [12].

An emerging trend is the development of wireless powered communication networks (WPCNs), which implement a downlink wireless energy transfer (DWET) followed by an uplink wireless information transmission (UWIT) [13]. A base station (BS) first transfers energy to the wireless powered users, who harvest the energy for transmitting their independent information to the BS. Time allocation for DWET and UWIT to optimize the sum information rate subject to per-user achievable information rate thresholds was considered in [14]. The joint downlink beamforming and uplink power allocation under fixed time durations to optimize the worst achievable information user rate was considered in [15] using alternating optimization. The optimal time allocation is then searched at grinding points. A similar optimization problem for the joint energy weight and power allocation with the BS of massive antenna array was analyzed asymptotically in [16], while [17] considered energy efficiency in the case of a single user. The joint time allocation for DWET, time separation and power allocation in UWIT for the users in optimizing the WPCN energy efficiency was proposed in [18]. Very recently, [19] considered a joint downlink (DL) and uplink (UL) transmission of K-tier heterogeneous cellular networks with downlink simultaneous wireless information and power transfer, where outage probability and ergodic capacity of both DL and UL have been derived.

Meanwhile, full duplex (FD) radio [20] offers enormous potential to significantly enhance the spectral efficiency compared to its half duplex (HD) counterpart. Recent studies (see, e.g., [21] and [22]) showed that FD radio may be deployable in next-generation networks because it can be implemented at reasonable cost and without complex radio hardware. The major challenge in FD radio is the residual self-interference (SI) from the transmit antennas to the receive antennas, which are co-located and function at the same time and over the same frequency band. A wide range of SI mitigation techniques were addressed in [20], [23] and [24]. More recently, FD multiple-input multiple-output (MIMO) precoding was studied in the context of multiuser MIMO (MU-MIMO) to improve the overall spectral efficiency of downlink and uplink channels [25]–[28]. A FD single-antenna architecture for energy-recycling was proposed in [29]. Beamformer design at the FD relay for harvesting energy and suppressing loop interference was considered in [30].

This work was supported in part by the Australian Research Councils Discovery Projects under Project DP130104617, in part by the U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415/14/22 and U.K. Engineering and Physical Sciences Research Council under Grant EP/P019374/1, and in part by the U.S. National Science Foundation under Grants CNS-1456793 and ECCS-1343210.

V.-D. Nguyen and O.-S. Shin are with the School of Electronic Engineering & Department of ICMC Convergence Technology, Soongsil University, Seoul 06978, Korea (e-mail: {nguyenvandinh, osshin}@ssu.ac.kr).

T. Q. Duong is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, United Kingdom (e-mail: trung.q.duong@qub.ac.uk).

H. D. Tuan is with the Faculty of Engineering and Information Technology, University of Technology Sydney, Broadway, NSW 2007, Australia (email: Tuan.Hoang@uts.edu.au).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

In this paper, we study the potential of FD radio in WPCNs to improve both spectral and energy efficiencies. There are both downlink users (DLUs) and uplink users (ULUs), where the ULUs need to harvest energy from the BS via DWET by the BS. The communication operates in two phases in the same time slot and over the same frequency band. In the first phase, the BS uses all available antennas to simultaneously transmit information to DLUs and transfer the energy to ULUs. In the second phase, the BS operates in FD mode for transmitting information to DLUs and receiving information from ULUs. Each ULU uses only the harvested energy in transmitting information to the BS. We jointly optimize the time allocation for the phases, the DL information and energy beamforming, and the UL transmit power allocation subject to the power budget at the BS and the individual ULU information rate thresholds. The ULU information rate threshold constraints are crucial to resolving the so called doubly near-far problem in WPCNs that discriminates the ULUs by favoring ones with better channel conditions for both energy transfer in phase I and information transmission in phase II. The residual SI and co-channel interference (CCI) from ULUs to DLUs are taken into account, which potentially offer the best performance make the optimizations more challenging. In fact, these problems involve optimization of nonconvex utility functions subject to nonconvex constraints, for which the optimal solutions are difficult computationally. Nevertheless, we propose path-following algorithms to address them. Our main contributions are summarized as follows:

- We propose a new model for WPCNs to optimize simultaneous uplink and downlink information transmission by exploring FD radio for the BS.
- Assuming perfect channel state information (CSI), we first develop a path-following algorithm of low complexity for the computational solution of sum rate maximization (SRM). The obtained solutions are at least local optima as they satisfy the Karush-Kuhn-Tucker (KKT) conditions. Numerical results show fast convergence of the proposed algorithm and greatly improve the system performance over the conventional approaches.
- The energy efficiency maximization (EEM) problem is a difficult nonlinear fraction program since the objective is not a ratio of a concave and convex function. The commonly-used Dinkelbach-type algorithms are not applicable. We develop a novel path-following algorithm that only invokes one simple convex quadratic program at each iteration, which again converges at least to a local optimum.

The rest of this paper is organized as follows. The system model and problem formulations of SRM and EEM are described in Section II. We devise the optimal solution to the SRM and EEM problems in Section III and IV, respectively. Numerical results are provided in Section V, and Section VI concludes the paper.

Notation: Bold lower and upper case letters represent vectors and matrices, respectively; $\mathbf{X}^H$, $\mathbf{X}^T$, $\mathbf{X}^*$, and $\mathrm{Trace}(\mathbf{X})$ are the Hermitian transpose, normal transpose, conjugate, and trace of a matrix $\mathbf{X}$, respectively. $\|\cdot\|$ and $|\cdot|$ denote the Euclidean norm of a matrix or vector and the absolute value of a complex scalar, respectively. $\mathbf{I}_N$ represents an $N \times N$ identity matrix. $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\eta}, \boldsymbol{Z})$ means that $\mathbf{x}$ is a random vector following a complex circular Gaussian distribution with mean $\boldsymbol{\eta}$ and covariance matrix $\boldsymbol{Z}$. $\mathbb{E}[\cdot]$ denotes the statistical expectation. The notation $\mathbf{X} \succeq \mathbf{0}$ ($\mathbf{X} \succ \mathbf{0}$, resp.) means the matrix $\mathbf{X}$ is positive semi-definite (definite, resp.). $\Re\{\cdot\}$ represents real part of the argument. The inner product $\langle \mathbf{X}, \mathbf{Y} \rangle$ is defined as $\mathrm{Trace}(\mathbf{X}^H\mathbf{Y})$. $\nabla_{\mathbf{x}} f(\mathbf{x})$ represents the gradient of $f(\cdot)$ with respect to vector $\mathbf{x}$.

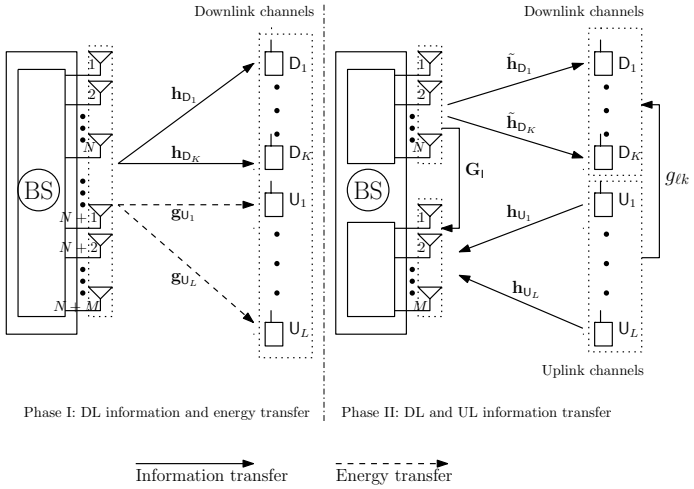


Fig. 1. A WPCN system with WET in the downlink channel and WIT in both uplink and downlink channels.

II. System Model and Optimization Problem Formulations

A. Signal Model

We consider a WPCN as illustrated in Fig. 1, which consists of a BS, K DLUs and L ULUs. The BS is equipped with M receive and N transmit antennas, while all the users are equipped with a single antenna. Denote by $\mathcal{D} \triangleq \{\mathsf{D}_1, \mathsf{D}_2, \cdots, \mathsf{D}_K\}$ and $\mathcal{U} \triangleq \{\mathsf{U}_1, \mathsf{U}_2, \cdots, \mathsf{U}_L\}$ the sets of DLUs and ULUs, respectively. All channels are assumed to follow independent quasi-static flat fading, i.e., remaining constant during a communication time block, denoted by T, but change independently from one block to another. Without loss of generality, the time block T is set as 1. Following [13]–[18], all ULUs $\mathsf{U}_\ell \in \mathcal{U}$ are assumed to harvest energy from the RF signal transmitted by the BS and then transmit information to BS as illustrated in Fig. 2. During the first fraction $0 < \alpha < 1$ of the time block, the BS simultaneously transmits information to all DLUs D_k and transfers energy to all ULU U_ℓ. In the remaining fraction $(1-\alpha)$, the BS operates in FD mode, i.e., it uses N antennas for transmitting information to the DLUs and M antennas for receiving information from ULUs.

With all $M+N$ antennas used in phase I, more degrees of freedom are added to the BS and all ULUs U_ℓ's are expected to harvest more energy from the RF signal. The complex baseband transmitted signal at the BS in phase I is then

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

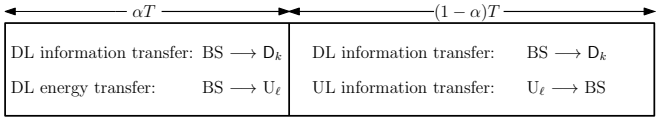


Fig. 2. The harvest-and-then-transmit protocol.

expressed as

$$\mathbf{x}_1 = \sum_{k=1}^{K} \mathbf{w}_{1,k} x_k + \mathbf{v}_e \tag{1}$$

where $\mathbf{w}_{1,k} \in \mathbb{C}^{(N+M)\times 1}$ denotes the k-th information beamforming vector, x_k with $\mathbb{E}\{|x_k|^2\} = 1$ is the message intended for DLU D_k. The energy beam vector $\mathbf{v}_e$ whose elements are zero-mean complex Gaussian random variables, is assumed $\mathbf{v}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{V}\mathbf{V}^H)$, where $\mathbf{V} \in \mathbb{C}^{(N+M)\times \tilde{L}}$ with $\tilde{L} \leq \min\big((N+M), L\big)$. The received signal at DLU D_k and ULU U_ℓ in phase I is, respectively, expressed as

$$y_{\mathsf{D}_k} = \mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k} x_k + \sum_{i=1, i\neq k}^{K} \mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,i} x_i + \mathbf{h}_{\mathsf{D}_k}^H \mathbf{v}_e + n_{\mathsf{D}_k}, \tag{2}$$

and

$$y_{\mathsf{U}_\ell} = \sum_{k=1}^{K} \mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{w}_{1,k} x_k + \mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{v}_e + n_{\mathsf{U}_\ell} \tag{3}$$

where $\mathbf{h}_{\mathsf{D}_k} \in \mathbb{C}^{(N+M)\times 1}$ and $\mathbf{g}_{\mathsf{U}_\ell} \in \mathbb{C}^{(N+M)\times 1}$ are the channel vectors from the BS to DLU D_k and ULU U_ℓ, respectively. They can be explicitly written as

$$\mathbf{h}_{\mathsf{D}_k} = [\underbrace{\hat{h}_{\mathsf{D}_k,1}, \cdots, \hat{h}_{\mathsf{D}_k,N}}_{\tilde{\mathbf{h}}_{\mathsf{D}_k}}, \hat{h}_{\mathsf{D}_k,(N+1)}, \cdots, \hat{h}_{\mathsf{D}_k,(N+M)}]^T, \tag{4}$$

$$\mathbf{g}_{\mathsf{U}_\ell} = [\hat{g}_{\mathsf{U}_\ell,1}, \cdots, \hat{g}_{\mathsf{U}_\ell,N}, \underbrace{\hat{g}_{\mathsf{U}_\ell,(N+1)}, \cdots, \hat{g}_{\mathsf{U}_\ell,(N+M)}}_{\tilde{\mathbf{g}}_{\mathsf{U}_\ell}}]^T \tag{5}$$

where $\hat{h}_{\mathsf{D}_k,i} \in \mathbb{C}$ and $\hat{g}_{\mathsf{U}_\ell,i} \in \mathbb{C}$, $\forall i = 1, \cdots, N+M$, denote the baseband channels from the i-th antenna at the BS to DLU D_k and ULU U_ℓ, respectively; $n_{\mathsf{D}_k} \sim \mathcal{CN}(0, \sigma_k^2)$ and $n_{\mathsf{U}_\ell} \sim \mathcal{CN}(0, \hat{\sigma}_\ell^2)$ represent the additive white Gaussian noise (AWGN) at DLU D_k and ULU U_ℓ, respectively. The harvested energy at ULU U_ℓ is defined by

$$\begin{aligned} E_{\mathsf{U}_\ell}(\mathbf{w}_1, \mathbf{V}, \alpha) &= \eta\alpha\mathbb{E}\left\{|y_{\mathsf{U}_\ell}|^2\right\} \\ &= \eta\alpha\Big(\sum_{k=1}^{K} |\mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{w}_{1,k}|^2 + \|\mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{V}\|^2\Big) \end{aligned} \tag{6}$$

where $\mathbf{w}_1 \triangleq [\mathbf{w}_{1,1}^T, \cdots, \mathbf{w}_{1,K}^T]^T \in \mathbb{C}^{(N+M)K}$, and η denotes the energy conversion efficiency at the receiver. In (6), the receive noise can be neglected since it will be negligible compared to energy transfer from the BS in practice. We can see that the harvested energy in (6) is also contributed by the DL information beams. We will show in Section V that the energy beam $\mathbf{v}_e$ is very beneficial when the ULUs are far from the DLUs.

From (2), the signal-to-interference-plus-noise ratio (SINR) at DLU D_k in phase I can be expressed as

$$\gamma_{1,k}(\mathbf{w}_1, \mathbf{V}) = \frac{|\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}|^2}{\sum_{i=1, i\neq k}^{K} |\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,i}|^2 + \|\mathbf{h}_{\mathsf{D}_k}^H \mathbf{V}\|^2 + \sigma_k^2}. \tag{7}$$

In phase II, the BS uses N antennas for transmitting information to the DLUs and M antennas for receiving information from the ULUs. The received signal at DLU D_k and the BS can be, respectively, written as

$$\begin{aligned} \tilde{y}_{\mathsf{D}_k} = \tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k} x'_k + \sum_{i=1, i\neq k}^{K} \tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,i} x'_i \\ + \sum_{\ell=1}^{L} p_\ell g_{\ell k} v_\ell + n_{\mathsf{D}_k}, \end{aligned} \tag{8}$$

and

$$\mathbf{y}_{\mathsf{U}} = \sum_{\ell=1}^{L} p_\ell \mathbf{h}_{\mathsf{U}_\ell} v_\ell + \sqrt{\rho} \sum_{k=1}^{K} \mathbf{G}_{\mathsf{I}}^H \mathbf{w}_{2,k} x'_k + \mathbf{n}_{\mathsf{U}} \tag{9}$$

where $\tilde{\mathbf{h}}_{\mathsf{D}_k} \in \mathbb{C}^{N\times 1}$, $\mathbf{w}_{2,k} \in \mathbb{C}^{N\times 1}$ and x'_k with $\mathbb{E}\{|x'_k|^2\} = 1$ are the transmit channel vector, information beam for DLU D_k and the message intended for DLU D_k in phase II. $p_\ell \in \mathbb{C}$, $\mathbf{h}_{\mathsf{U}_\ell} \in \mathbb{C}^{M\times 1}$ and v_ℓ with $\mathbb{E}\{|v_\ell|^2\} = 1$ are the transmit power, the receive channel vector, and the message of ULU U_ℓ, respectively. $\mathbf{n}_{\mathsf{U}} \sim \mathcal{CN}(\mathbf{0}, \tilde{\sigma}^2\mathbf{I})$ denotes the receive AWGN at the BS. Since the channel remains unchanged during a transmission block time, $\tilde{\mathbf{h}}_{\mathsf{D}_k}$ corresponds to first N elements of $\mathbf{h}_{\mathsf{D}_k}$ in (4). We also assume the reciprocity for UL and DL links, i.e., $\mathbf{h}_{\mathsf{U}_\ell} = \tilde{\mathbf{g}}_{\mathsf{U}_\ell}, \forall \ell$, where $\tilde{\mathbf{g}}_{\mathsf{U}_\ell}$ is defined in (5). Note that the term $\sqrt{\rho}\sum_{k=1}^{K} \mathbf{G}_{\mathsf{I}}^H \mathbf{w}_{2,k} x_k$ in (9) represents the FD SI left over from the so called analog-circuit domain SI cancellation [24], where $\mathbf{G}_{\mathsf{I}}$ is a fading loop channel and $0 \leq \rho \leq 1$ is used for modeling the degree of SI propagation [20]. The CCI from ULU U_ℓ to DLU D_k is denoted by $g_{\ell k}$.

From (8), the SINR at DLU D_k in phase II is

$$\gamma_{2,k}(\mathbf{w}_2, \mathbf{p}) = \frac{|\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}|^2}{\sum_{i=1, i\neq k}^{K} |\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,i}|^2 + \sum_{\ell=1}^{L} p_\ell^2 |g_{\ell k}|^2 + \sigma_k^2} \tag{10}$$

where $\mathbf{w}_2 \triangleq [\mathbf{w}_{2,1}^T, \mathbf{w}_{2,2}^T, \cdots, \mathbf{w}_{2,K}^T]^T \in \mathbb{C}^{NK}$ and $\mathbf{p} \triangleq [p_1, p_2, \cdots, p_L]^T$. The achieved SR of DL transmission is thus

$$\begin{aligned} R_{\mathsf{D}}\big(\mathbf{w}_1, \mathbf{w}_2, \mathbf{V}, \mathbf{p}, \alpha\big) = \alpha \sum_{k=1}^{K} \ln\Big(1 + \gamma_{1,k}\big(\mathbf{w}_1, \mathbf{V}\big)\Big) \\ + \big(1-\alpha\big) \sum_{k=1}^{K} \ln\Big(1 + \gamma_{2,k}\big(\mathbf{w}_2, \mathbf{p}\big)\Big). \end{aligned} \tag{11}$$

For simplicity, we adopt the minimum mean square error and successive interference cancellation (MMSE-SIC) receiver at the BS to maximize the received SINR of U_ℓ in (9). Assuming that the decoding order follows the ULU index order, the resultant SINR in decoding U_ℓ's information is [31]

$$\begin{aligned} \gamma_\ell(\mathbf{w}_2, \mathbf{p}) = & p_\ell^2 \mathbf{h}_{\mathsf{U}_\ell}^H \Bigg(\sum_{j>\ell}^{L} p_j^2 \mathbf{h}_{\mathsf{U}_j} \mathbf{h}_{\mathsf{U}_j}^H \\ & + \rho \sum_{k=1}^{K} \mathbf{G}_{\mathsf{I}}^H \mathbf{w}_{2,k} \mathbf{w}_{2,k}^H \mathbf{G}_{\mathsf{I}} + \tilde{\sigma}^2 \mathbf{I}\Bigg)^{-1} \mathbf{h}_{\mathsf{U}_\ell}. \end{aligned} \tag{12}$$

Then, the achieved SR of UL transmission is

$$R_{\mathsf{U}}(\mathbf{w}_2, \mathbf{p}, \alpha) = \big(1-\alpha\big) \sum_{\ell=1}^{L} \ln\Big(1 + \gamma_\ell\big(\mathbf{w}_2, \mathbf{p}\big)\Big). \tag{13}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

B. Optimization Problem Formulations

Our main goal is to maximize both the total SR and EE of the system by jointly deriving the time allocation (for two phases), the DL beamformers, and the UL transmit power allocation under the ULU rate thresholds.

1) SRM Problem Formulation: The SRM problem of jointly designing $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{V}$, $\mathbf{p}$, and α can be expressed as

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha}{\text{maximize}} \; R_{\mathsf{D}}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha)+R_{\mathsf{U}}(\mathbf{w}_2,\mathbf{p},\alpha) \tag{14a}$$

$$\text{s.t. } (1-\alpha)\ln\big(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\big)\geq \bar{r}_{\mathsf{u}},\ \forall \ell=1,\cdots,L, \tag{14b}$$

$$p_\ell^2 \leq p_{\mathsf{U}_\ell}^{\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha),\ \forall \ell=1,\cdots,L, \tag{14c}$$

$$p_\ell \geq 0,\ \forall \ell=1,\cdots,L, \tag{14d}$$

$$\alpha\Big(\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2\Big)+(1-\alpha)\|\mathbf{w}_2\|^2 \leq P_{\mathsf{BS}}, \tag{14e}$$

$$0<\alpha<1 \tag{14f}$$

where without loss of generality the same rate threshold $\bar{r}_{\mathsf{u}}$ for all ULUs is set, P_{BS} is the maximum transmit power at the BS, and

$$p_{\mathsf{U}_\ell}^{\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha)=\frac{E_{\mathsf{U}_\ell}(\mathbf{w}_1,\mathbf{V},\alpha)}{(1-\alpha)} =\frac{\eta\alpha}{1-\alpha}\Big(\sum_{k=1}^{K}|\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{w}_{1,k}|^2+\|\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{V}\|^2\Big). \tag{15}$$

The optimization problem in (14) is known as the spectral efficiency maximization problem. The constraints in (14b) impose a quality-of-service (QoS) requirement for ULU U_ℓ, i.e., the achievable information decoding rate should be not less than a given threshold $\bar{r}_{\mathsf{u}}$ to prevent the FD to maximize R_{D} only in maximizing the objective value in (14), leading to an extremely low QoS for ULUs. More importantly, they cognitively rule out the doubly near-far occurrence of favoring ULUs with better channel conditions. According to (6) and (11), the ULUs with better channel conditions are advantageous in both harvesting energy in phase I and transmitting information in phase II. The constraints (14c) merely mean that each ULU U_ℓ utilizes its energy harvested from phase I to transmit information to the BS, as illustrated in Fig. 2. The constraint in (14e) is the power constraint at the BS, which differs from the following one that was studied previously [15]:

$$\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2+\|\mathbf{w}_2\|^2\leq P_{\mathsf{BS}}. \tag{16}$$

However, in contrast to the left-hand side (LHS) of (14e), which is the total transmit power at the BS, the LHS of (16) is a sum of power rates so the constraint (16) is meaningless. Note that (16) is much stricter than (14e), i.e., by using (16), the BS does not use all allowable power and thus the corresponding performance is not optimal. This will be elaborated in the next sections.

2) EEM Problem Formulation: Another performance metric of interest is to maximize EE of the system. Energy consumption in green wireless networks has attracted much attention of both academia and industry recently [17], [32], [33]. In this paper, the power consumed by the BS and ULUs is taken into account.

Power consumption model: A linear power model [34] is adopted in this paper, where the total power consumption at the BS is modeled as

$$P_{\mathsf{DL}}=\frac{1}{\epsilon}\Big(\alpha(\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2)+(1-\alpha)\|\mathbf{w}_2\|^2\Big) +\alpha M P_{\mathsf{BS}}^{\mathsf{dyn}}+NP_{\mathsf{BS}}^{\mathsf{dyn}}+P_{\mathsf{BS}}^{\mathsf{sta}} \tag{17}$$

where $\epsilon\in(0,1]$ is the power amplifier efficiency, $P_{\mathsf{BS}}^{\mathsf{dyn}}$ is the dynamic power consumption associated to the power radiation of all circuit blocks in each active radio frequency chain, and $P_{\mathsf{BS}}^{\mathsf{sta}}$ is the static power consumed by cooling system, power supply, etc. Similarly, the total power consumption of all users in the UL channel is given by

$$P_{\mathsf{UL}}=(1-\alpha)\sum_{\ell=1}^{L}P_{\mathsf{U}_\ell}^{\mathsf{dyn}}+\sum_{\ell=1}^{L}P_{\mathsf{U}_\ell}^{\mathsf{sta}}. \tag{18}$$

Note that the formulation in (18) does not include a power consumed by sending data by ULUs as it is already incorporated in (17). For notational simplicity, we denote $P_0\triangleq NP_{\mathsf{BS}}^{\mathsf{dyn}}+P_{\mathsf{BS}}^{\mathsf{sta}}+\sum_{\ell=1}^{L}P_{\mathsf{U}_\ell}^{\mathsf{sta}}$ as the circuit power of the system, which is independent from the optimization variables. The EEM problem is thus

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha}{\text{maximize}} \; \frac{R_{\mathsf{D}}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha)+R_{\mathsf{U}}(\mathbf{w}_2,\mathbf{p},\alpha)}{\chi(\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\alpha)+P_0} \tag{19a}$$

$$\text{s.t. } (14\text{b}),(14\text{c}),(14\text{d}),(14\text{e}),(14\text{f}) \tag{19b}$$

where $\chi(\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\alpha)\triangleq\frac{1}{\epsilon}\Big(\alpha\big(\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2\big)+(1-\alpha)\|\mathbf{w}_2\|^2\Big)+\alpha MP_{\mathsf{BS}}^{\mathsf{dyn}}+(1-\alpha)\sum_{\ell=1}^{L}P_{\mathsf{U}_\ell}^{\mathsf{dyn}}$.

Naturally, the reciprocity for UL and DL links in time duplex division (TDD) mode is adopted for small cell systems as those considered in this paper, under which CSI is easily obtained by requesting all DLUs and ULUs to send their pilots to the BS and thus can be assumed perfectly available. The performance under perfect CSI also serves as a benchmark for the achievement of the FD systems. Moreover, our proposed algorithms can be further adjusted to robust optimization problems in dealing with the worst case of imperfect CSI.

III. Sum Rate Maximization

Finding an optimal solution to the SRM problem (14) is challenging due to the non-concavity of its objective function and nonconvexity of its feasible set. In this section, we propose a path-following computation procedure to obtain a local optimum.

The following inequalities, whose proofs are given in the Appendix A, will be frequently used in the paper:

$$\ln\Big(1+\frac{|x|^2}{y}\Big)\geq\ln\Big(1+\frac{|\bar{x}|^2}{\bar{y}}\Big)-\frac{|\bar{x}|^2}{\bar{y}}+2\frac{\Re\{\bar{x}^*x\}}{\bar{y}} -\frac{|\bar{x}|^2(|x|^2+y)}{\bar{y}(\bar{y}+|\bar{x}|^2)}, \tag{20}$$

$$\frac{|x|^2}{y}\geq 2\frac{\bar{x}^*x}{\bar{y}}-\frac{|\bar{x}|^2}{\bar{y}^2}y, \tag{21}$$

$$\forall x\in\mathbb{C},\bar{x}\in\mathbb{C},y>0,\bar{y}>0.$$

As the first step, we use the additional variables τ and β, which satisfy the convex constraints

$$\alpha\tau \geq 1 \text{ and } (1-\alpha)\beta \geq 1 \tag{22}$$

to equivalently express (14) as

$$\begin{aligned}
&\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta}{\text{maximize}} \sum_{k=1}^{K} \frac{\ln\big(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\big)}{\tau} \\
&+\sum_{k=1}^{K} \frac{\ln\big(1+\gamma_{2,k}(\mathbf{w}_2,\mathbf{p})\big)}{\beta} + \sum_{\ell=1}^{L} \frac{\ln\big(1+\gamma_{\ell}(\mathbf{w}_2,\mathbf{p})\big)}{\beta} \quad (23a)\\
&\text{s.t. } \ln\big(1+\gamma_{\ell}(\mathbf{w}_2,\mathbf{p})\big) \geq \bar{r}_{\mathsf{u}}/(1-\alpha),\ \forall \ell = 1,\cdots,L, \quad (23b)\\
&\frac{p_\ell^2}{\alpha} \leq \eta \frac{\phi_\ell^{\text{eh}}(\mathbf{w}_1,\mathbf{V})}{1-\alpha},\ \forall \ell = 1,\cdots,L, \quad (23c)\\
&\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2 + \|\mathbf{w}_2\|^2 - \frac{\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2}{\beta} \\
&\qquad - \frac{\|\mathbf{w}_2\|^2}{\tau} \leq P_{\mathsf{BS}} \quad (23d)\\
&(14\text{d}), (14\text{f}), (22), \quad (23e)
\end{aligned}$$

with

$$\phi_\ell^{\text{eh}}(\mathbf{w}_1,\mathbf{V}) = \sum_{k=1}^{K} |\mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{w}_{1,k}|^2 + \|\mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{V}\|^2.$$

Note that the objective (23a) is nonconcave and (23b)-(23d) are also nonconvex constraints. Let us treat the nonconcave objective (23a) first. As observed in [35], (7) can be equivalently replaced by

$$\gamma_{1,k}(\mathbf{w}_1,\mathbf{V}) = \frac{\big(\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\}\big)^2}{\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V})} \tag{24}$$

with the additional linear constraint

$$\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\} \geq 0 \tag{25}$$

where

$$\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) = \sum_{i=1,i\neq k}^{K} |\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,i}|^2 + ||\mathbf{h}_{\mathsf{D}_k}^H \mathbf{V}||^2 + \sigma_k^2.$$

At a feasible point $\big(\mathbf{w}_1^{(n)}, \mathbf{V}^{(n)}\big)$ it follows from the inequality (20) that

$$\begin{aligned}
&\ln\big(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\big) \geq \\
&\ln\big(1+\gamma_{1,k}(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)})\big) - \gamma_{1,k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big) \\
&+ 2\frac{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}}{\varphi_{\mathsf{D}_k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big)} - \\
&\frac{\Big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\Big)^2 \Big(\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) + \big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}\big)^2\Big)}{\varphi_{\mathsf{D}_k}(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)})\Big(\varphi_{\mathsf{D}_k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big) + \big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\big)^2\Big)}
\end{aligned} \tag{26}$$

Upon setting

$$\begin{aligned}
&a_{1,\mathsf{D}_k}^{(n)} = \ln\big(1+\gamma_{1,k}(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)})\big) - \gamma_{1,k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big) < 0,\\
&b_{1,\mathsf{D}_k}^{(n)} = 2\frac{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}}{\varphi_{\mathsf{D}_k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big)} > 0 \quad \big(\text{thanks to } (25)\big),\\
&c_{1,\mathsf{D}_k}^{(n)} = \\
&\frac{\big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\big)^2}{\varphi_{\mathsf{D}_k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big)\Big(\varphi_{\mathsf{D}_k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big) + \big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\big)^2\Big)} > 0
\end{aligned} \tag{27}$$

it follows from (26) and the inequality (21) that

$$\begin{aligned}
&\frac{\ln\big(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\big)}{\tau} \geq \frac{a_{1,\mathsf{D}_k}^{(n)}}{\tau} + b_{1,\mathsf{D}_k}^{(n)} \frac{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}}{\tau} \\
&- c_{1,\mathsf{D}_k}^{(n)} \frac{\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) + \big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}\big)^2}{\tau} \\
&\geq \frac{a_{1,\mathsf{D}_k}^{(n)}}{\tau} + b_{1,\mathsf{D}_k}^{(n)} \Bigg(2\frac{\sqrt{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}}\sqrt{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}}}{\tau^{(n)}} \\
&- \frac{\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\big\}\tau}{(\tau^{(n)})^2}\Bigg) - c_{1,\mathsf{D}_k}^{(n)} \frac{\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) + \big(\Re\big\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\big\}\big)^2}{\tau} \quad (28)\\
&:= f_{1,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_1,\mathbf{V},\tau\big). \quad (29)
\end{aligned}$$

Note that $f_{1,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_1,\mathbf{V},\tau\big)$ is concave and is a global lower bound of $\ln\big(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\big)/\tau$ with

$$f_{1,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)},\tau^{(n)}\big) = \frac{\ln\Big(1+\gamma_{1,k}\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)}\big)\Big)}{\tau^{(n)}}. \tag{30}$$

Analogously, (10) can be equivalently replaced by

$$\gamma_{2,k}(\mathbf{w}_2,\mathbf{p}) = \frac{\big(\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\}\big)^2}{\psi_{\mathsf{D}_k}(\mathbf{w}_2,\mathbf{p})} \tag{31}$$

with the additional linear constraint

$$\Re\big\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\big\} \geq 0 \tag{32}$$

where

$$\psi_{\mathsf{D}_k}(\mathbf{w}_2,\mathbf{p}) = \sum_{i=1,i\neq k}^{K} |\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,i}|^2 + \sum_{\ell=1}^{L} p_\ell^2 |g_{\ell k}|^2 + \sigma_k^2.$$

At a feasible point $\big(\mathbf{w}_2^{(n)}, \mathbf{p}^{(n)}, \beta^{(n)}\big)$ it can be shown in a similar manner that

$$\begin{aligned}
&\frac{\ln\big(1+\gamma_{2,k}(\mathbf{w}_2,\mathbf{p})\big)}{\beta} \geq \frac{a_{2,\mathsf{D}_k}^{(n)}}{\beta} \\
&+ b_{2,\mathsf{D}_k}^{(n)} \Bigg(2\frac{\sqrt{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}^{(n)}\}}\sqrt{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\}}}{\beta^{(n)}} - \frac{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}^{(n)}\}\beta}{(\beta^{(n)})^2}\Bigg) \\
&- c_{2,\mathsf{D}_k}^{(n)} \frac{\psi_{\mathsf{D}_k}(\mathbf{w}_2,\mathbf{p}) + \big(\Re\big\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\big\}\big)^2}{\beta} \quad (33)\\
&:= f_{2,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_2,\mathbf{p},\beta\big) \quad (34)
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

where

$$a_{2,\mathsf{D}_k}^{(n)} = \ln\big(1+\gamma_{2,k}(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)})\big) - \gamma_{2,k}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big) < 0,$$
$$b_{2,\mathsf{D}_k}^{(n)} = 2\frac{\Re\big\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H\mathbf{w}_{2,k}^{(n)}\big\}}{\psi_{\mathsf{D}_k}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)} > 0 \quad \big(\text{thanks to (32)}\big),$$
$$c_{2,\mathsf{D}_k}^{(n)} = \frac{\big(\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H\mathbf{w}_{2,k}^{(n)}\}\big)^2}{\psi_{\mathsf{D}_k}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)\Big(\psi_{\mathsf{D}_k}(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}) + \big(\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H\mathbf{w}_{2,k}^{(n)}\}\big)^2\Big)} > 0 \tag{35}$$

and $f_{2,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_2,\mathbf{p},\beta\big)$ is concave and is a global lower bound of $\ln\big(1+\gamma_{2,k}\big(\mathbf{w}_2,\mathbf{p}\big)\big)/\beta$ with

$$f_{2,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)},\beta^{(n)}\big) = \frac{\ln\Big(1+\gamma_{2,k}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)\Big)}{\beta^{(n)}}. \tag{36}$$

Next, at the feasible point $\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)$, based on the inequality (20) the left-hand side (LHS) of (23b) is lower bounded by

$$\ln\big(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\big) \geq \tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) \tag{37}$$

where $\tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p})$ is defined as

$$\tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) = a_\ell^{(n)} + b_\ell^{(n)}p_\ell - \varphi_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) \tag{38}$$

with

$$a_\ell^{(n)} = \ln\big(1+\gamma_\ell(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)})\big) - \gamma_\ell\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big) < 0,$$
$$b_\ell^{(n)} = 2p_\ell^{(n)}\mathbf{h}_{\mathsf{U}_\ell}^H\mathbf{\Phi}_\ell^{(n)}\mathbf{h}_{\mathsf{U}_\ell} > 0,$$
$$\varphi_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) = p_\ell^2\mathbf{h}_{\mathsf{U}_\ell}^H\mathbf{\Omega}_\ell^{(n)}\mathbf{h}_{\mathsf{U}_\ell} + \sum_{j>\ell}^L p_j^2\mathbf{h}_{\mathsf{U}_j}^H\mathbf{\Omega}_\ell^{(n)}\mathbf{h}_{\mathsf{U}_j} + \rho\sum_{k=1}^K \mathbf{w}_{2,k}^H\mathbf{G}_{\mathsf{I}}\mathbf{\Omega}_\ell^{(n)}\mathbf{G}_{\mathsf{I}}^H\mathbf{w}_{2,k} + \tilde{\sigma}^2\mathrm{Trace}\big(\mathbf{\Omega}_\ell^{(n)}\big),$$
$$\mathbf{\Phi}_\ell^{(n)} = \Big(\sum_{j>\ell}^L\big(p_j^{(n)}\big)^2\mathbf{h}_{\mathsf{U}_j}\mathbf{h}_{\mathsf{U}_j}^H + \rho\sum_{k=1}^K\mathbf{G}_{\mathsf{I}}^H\mathbf{w}_{2,k}^{(n)}\big(\mathbf{w}_{2,k}^{(n)}\big)^H\mathbf{G}_{\mathsf{I}} + \tilde{\sigma}_\ell^2\mathbf{I}\Big)^{-1} \succ \mathbf{0},$$
$$\mathbf{\Omega}_\ell^{(n)} = \mathbf{\Phi}_\ell^{(n)} - \Big(\big(p_\ell^{(n)}\big)^2\mathbf{h}_{\mathsf{U}_\ell}\mathbf{h}_{\mathsf{U}_\ell}^H + \sum_{j>\ell}^L\big(p_j^{(n)}\big)^2\mathbf{h}_{\mathsf{U}_j}\mathbf{h}_{\mathsf{U}_j}^H + \rho\sum_{k=1}^K\mathbf{G}_{\mathsf{I}}^H\mathbf{w}_{2,k}^{(n)}\big(\mathbf{w}_{2,k}^{(n)}\big)^H\mathbf{G}_{\mathsf{I}} + \tilde{\sigma}^2\mathbf{I}\Big)^{-1} \succeq \mathbf{0}. \tag{39}$$

It follows from (37) that $\tilde{f}_\ell^{(n)}\big(\mathbf{w}_2,\mathbf{p}\big)$ is a concave quadratic function with

$$\ln\Big(1+\gamma_\ell\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)\Big) = \tilde{f}_\ell^{(n)}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big). \tag{40}$$

As a result, the constraint (23b) is innerly approximated by the following convex quadratic constraint:

$$\tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) \geq \bar{r}_{\mathsf{u}}/(1-\alpha),\ \forall\ell=1,\cdots,L \tag{41}$$

while the last term of the objective (23a) is lower bounded by

$$\frac{\ln\big(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\big)}{\beta} \geq \frac{a_\ell^{(n)}}{\beta} + b_\ell^{(n)}\left(2\frac{\sqrt{p_\ell^{(n)}}\sqrt{p_\ell}}{\beta^{(n)}} - \frac{p_\ell^{(n)}\beta}{(\beta^{(n)})^2}\right) - \frac{\varphi_\ell^{(n)}\big(\mathbf{w}_2,\mathbf{p}\big)}{\beta} \tag{42}$$
$$:= f_\ell^{(n)}\big(\mathbf{w}_2,\mathbf{p},\beta\big) \tag{43}$$

with the concave function $f_\ell^{(n)}(\mathbf{w}_2,\mathbf{p},\beta)$ satisfying

$$\frac{\ln\Big(1+\gamma_\ell\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)}\big)\Big)}{\beta^{(n)}} = f_\ell^{(n)}\big(\mathbf{w}_2^{(n)},\mathbf{p}^{(n)},\beta^{(n)}\big). \tag{44}$$

For the constraint (23c) in (23), note that both of its sides are convex functions. Thus, at a feasible point $\big(\mathbf{w}_1^{(n)},\mathbf{V}^{(n)},\alpha^{(n)}\big)$, it follows from the inequality (21) that the function $\phi_\ell^{\mathrm{eh}}(\mathbf{w}_1,\mathbf{V})/(1-\alpha)$ in (23c) is lower bounded by

$$\frac{\phi_\ell^{\mathrm{eh}}(\mathbf{w}_1,\mathbf{V})}{1-\alpha} \geq \phi_\ell^{(n),\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha)$$
$$= \frac{\sum_{k=1}^K 2\Re\big\{(\mathbf{w}_{1,k}^{(n)})^H\mathbf{g}_{\mathsf{U}_\ell}\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{w}_{1,k}\big\}}{1-\alpha^{(n)}} - \frac{|\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{w}_{1,k}^{(n)}|^2(1-\alpha)}{\big(1-\alpha^{(n)}\big)^2}$$
$$+\frac{2\Re\big\{\mathrm{Trace}\big((\mathbf{V}^{(n)})^H\mathbf{g}_{\mathsf{U}_\ell}\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{V}\big)\big\}}{1-\alpha^{(n)}} - \frac{\|\mathbf{g}_{\mathsf{U}_\ell}^H\mathbf{V}^{(n)}\|^2(1-\alpha)}{(1-\alpha^{(n)})^2}. \tag{45}$$

Therefore, (23c) is innerly approximated by the convex constraint

$$\frac{p_\ell^2}{\alpha} \leq \eta\phi_\ell^{(n),\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha),\ \forall\ell=1,\cdots,L. \tag{46}$$

By using the inequality (21), the nonconvex constraint (23d) is innerly approximated by

$$\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2 + \|\mathbf{w}_2\|^2 + \frac{\big(\|\mathbf{w}_1^{(n)}\|^2 + \|\mathbf{V}^{(n)}\|^2\big)\beta}{\big(\beta^{(n)}\big)^2} - 2\frac{\Re\Big\{\big(\mathbf{w}_1^{(n)}\big)^H\mathbf{w}_1\Big\} + \Re\Big\{\mathrm{Trace}\big((\mathbf{V}^{(n)})^H\mathbf{V}\big)\Big\}}{\beta^{(n)}} - 2\frac{\Re\big\{(\mathbf{w}_2^{(n)})^H\mathbf{w}_2\big\}}{\tau^{(n)}} + \frac{\|\mathbf{w}_2^{(n)}\|^2\tau}{(\tau^{(n)})^2} \leq P_{\mathsf{BS}}. \tag{47}$$

From the above discussions, at $(n+1)$-th iteration, we solve the following convex problem:

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta}{\text{maximize}} \sum_{k=1}^K\Big(f_{1,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_1,\mathbf{V},\tau\big) + f_{2,\mathsf{D}_k}^{(n)}\big(\mathbf{w}_2,\mathbf{p},\beta\big)\Big) + \sum_{\ell=1}^L f_\ell^{(n)}\big(\mathbf{w}_2,\mathbf{p},\beta\big) \tag{48a}$$
$$\text{s.t.}\quad (14\text{d}),(14\text{f}),(22),(25),(32),(41),(46),(47). \tag{48b}$$

To find a feasible point of (14) to initialize the computational procedure, we consider the following nonconvex optimization subject to the convex constraints:

$$\max_{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta}\ \min_{\ell=1,\cdots,L}\ \min\Big\{\ln\big(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\big) - \frac{\bar{r}_{\mathsf{u}}}{1-\alpha}, \eta\frac{\phi_\ell^{\mathrm{eh}}(\mathbf{w}_1,\mathbf{V})}{1-\alpha} - \frac{p_\ell^2}{\alpha}\Big\} \tag{49a}$$
$$\text{s.t.}\quad (14\text{d}),(14\text{f}),(22),(47). \tag{49b}$$

Algorithm 1 Path-following algorithm for the SRM problem (14)

Initialization: Set $n := 0$ and solve (50) to generate an initial feasible point $\left(\mathbf{w}_1^{(n)}, \mathbf{w}_2^{(n)}, \mathbf{V}^{(n)}, \mathbf{p}^{(n)}, \alpha^{(n)}, \tau^{(n)}, \beta^{(n)}\right)$

1: **repeat**
2: Solve (48) to obtain the optimal solutions: $\left(\mathbf{w}_1^\star, \mathbf{w}_2^\star, \mathbf{V}, \mathbf{p}^\star, \alpha^\star, \tau^\star, \beta^\star\right)$.
3: Update: $\mathbf{w}_1^{(n+1)} := \mathbf{w}_1^\star$, $\mathbf{w}_2^{(n+1)} := \mathbf{w}_2^\star$, $\mathbf{V}^{(n+1)} := \mathbf{V}^\star$, $\mathbf{p}^{(n+1)} := \mathbf{p}^\star$, $\alpha^{(n+1)} := \alpha^\star$, $\tau^{(n+1)} := \tau^\star$ and $\beta^{(n+1)} := \beta^\star$.
4: Set $n := n + 1$.
5: **until** Convergence

Initialized by any feasible $\left(\mathbf{w}_1^{(0)}, \mathbf{w}_2^{(0)}, \mathbf{V}^{(0)}, \mathbf{p}^{(0)}, \alpha^{(0)}, \tau^{(0)}, \beta^{(0)}\right)$ to the convex constraints (14d), (14f), (22), and (47), we solve the following convex optimization at the n-th iteration:

$$\max_{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta} \min_{\ell=1,\ldots,L} \left\{ \tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) - \frac{\bar{r}_{\mathsf{u}}}{1-\alpha}, \eta\phi_\ell^{(n),\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha) - \frac{p_\ell^2}{\alpha} \right\} \quad (50a)$$

$$\text{s.t.} \quad (14d), (14f), (22), (47). \quad (50b)$$

and stops upon reaching

$$\min_{\ell=1,\ldots,L} \left\{ \tilde{f}_\ell^{(n)}\left(\mathbf{w}_2^{(n+1)}, \mathbf{p}^{(n+1)}\right) - \frac{\bar{r}_{\mathsf{u}}}{1-\alpha^{(n+1)}}, \eta\phi_\ell^{(n),\mathrm{eh}}\left(\mathbf{w}_1^{(n+1)}, \mathbf{V}^{(n+1)}, \alpha^{(n+1)}\right) - \frac{\left(p_\ell^{(n+1)}\right)^2}{\alpha^{(n+1)}} \right\} \geq 0. \quad (51)$$

In Algorithm 1, we summarize our proposed path-following optimization algorithm to solve the SRM problem (14). After solving (48), we update the involved variables for the next iteration until convergence, which is stated the following proposition.

Proposition 1: Algorithm 1 generates a sequence $\left\{\left(\mathbf{w}_1^{(n)}, \mathbf{w}_2^{(n)}, \mathbf{V}^{(n)}, \mathbf{p}^{(n)}, \alpha^{(n)}, \tau^{(n)}, \beta^{(n)}\right)\right\}$ of improved points of (23) and (14), which converges to a KKT point.

Proof: See Appendix B. ∎

Complexity Analysis: The convex optimization problem (48) involves $\tilde{n} = (3N + M)K + (N + M)\tilde{L} + 2L + 4$ scalar real variables and $\tilde{m} = 2K + 3L + 4$ quadratic and linear constraints, so the per-iteration computational complexity of solving (48) is $\mathcal{O}(\tilde{n}^2\tilde{m}^{2.5} + \tilde{m}^{3.5})$.

IV. Energy Efficiency Maximization

We now consider the EEM problem (19), which is also nonconvex and more computationally difficult than the SRM problem (14). Again, we use the additional variables τ and β which satisfy the convex constraint

$$\sqrt{\tau} \geq 1/\alpha \text{ and } \sqrt{\beta} \geq 1/(1-\alpha). \quad (52)$$

Also introduce the new variable

$$\lambda > 0 \quad (53)$$

to tackle the consumed power which satisfies the constraint

$$\frac{\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2 + \|\mathbf{w}_2\|^2}{\epsilon} - \frac{\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2}{\epsilon\sqrt{\beta}} - \frac{\|\mathbf{w}_2\|^2}{\epsilon\sqrt{\tau}} + MP_{\mathsf{BS}}^{\mathsf{dyn}}\alpha + \left(\sum_{\ell=1}^{L} P_{\mathsf{U}_\ell}^{\mathsf{dyn}}\right)(1-\alpha) + P_0 \leq \sqrt{\lambda}. \quad (54)$$

Accordingly, the constraint (14e) is written as

$$\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2 + \|\mathbf{w}_2\|^2 - \frac{\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2}{\sqrt{\beta}} - \frac{\|\mathbf{w}_2\|^2}{\sqrt{\tau}} \leq P_{\mathsf{BS}}. \quad (55)$$

With above settings, we rewrite the EEM problem (19) by

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta,\lambda}{\text{maximize}} \sum_{k=1}^{K} \frac{\ln\left(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\right)}{\sqrt{\tau\lambda}} + \sum_{k=1}^{K} \frac{\ln\left(1+\gamma_{2,k}(\mathbf{w}_2,\mathbf{p})\right)}{\sqrt{\beta\lambda}} + \sum_{\ell=1}^{L} \frac{\ln\left(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\right)}{\sqrt{\beta\lambda}} \quad (56a)$$

$$\text{s.t.} \quad (14d), (14f), (23b), (23c), (52), (53), (54), (55). \quad (56b)$$

Like (23), the problem (56) is highly nonconvex. However, we will show in the following that the convex inner approximation approach to SRM can be extended to EEM.

To handle the objective of (56a), in the same manner to (29), we derive its lower bound as

$$\begin{aligned}
\frac{\ln\left(1+\gamma_{1,k}(\mathbf{w}_1,\mathbf{V})\right)}{\sqrt{\tau\lambda}} &\geq \frac{a_{1,\mathsf{D}_k}^{(n)}}{\sqrt{\tau\lambda}} + b_{1,\mathsf{D}_k}^{(n)} \frac{\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\}}{\sqrt{\tau\lambda}} - c_{1,\mathsf{D}_k}^{(n)} \frac{\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) + \left(\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\}\right)^2}{\sqrt{\tau\lambda}} \\
&\geq \frac{a_{1,\mathsf{D}_k}^{(n)}}{\sqrt{\tau\lambda}} + b_{1,\mathsf{D}_k}^{(n)} \left(2\frac{\sqrt{\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\}}\sqrt{\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\}}}{\sqrt{\tau^{(n)}\lambda^{(n)}}} - \frac{\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\}\tau}{2(\tau^{(n)})^{3/2}\sqrt{\lambda^{(n)}}} - \frac{\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}^{(n)}\}\lambda}{2(\lambda^{(n)})^{3/2}\sqrt{\tau^{(n)}}} \right) \\
&\quad - c_{1,\mathsf{D}_k}^{(n)} \frac{\varphi_{\mathsf{D}_k}(\mathbf{w}_1,\mathbf{V}) + \left(\Re\{\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}\}\right)^2}{\sqrt{\tau\lambda}} \quad (57) \\
&:= \hat{f}_{1,\mathsf{D}_k}^{(n)}(\mathbf{w}_1,\mathbf{V},\tau,\lambda) \quad (58)
\end{aligned}$$

where $a_{1,\mathsf{D}_k}^{(n)}, b_{1,\mathsf{D}_k}^{(n)}$, and $c_{1,\mathsf{D}_k}^{(n)}$ are defined in (27). Similarly to (34)

$$\begin{aligned}
\frac{\ln\left(1+\gamma_{2,k}(\mathbf{w}_2,\mathbf{p})\right)}{\sqrt{\beta\lambda}} &\geq \\
&\frac{a_{2,\mathsf{D}_k}^{(n)}}{\sqrt{\beta\lambda}} + b_{2,\mathsf{D}_k}^{(n)} \left(2\frac{\sqrt{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}^{(n)}\}}\sqrt{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\}}}{\sqrt{\beta^{(n)}\lambda^{(n)}}} - \frac{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}^{(n)}\}\beta}{2(\beta^{(n)})^{3/2}\sqrt{\lambda^{(n)}}} - \frac{\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}^{(n)}\}\lambda}{2(\lambda^{(n)})^{3/2}\sqrt{\beta^{(n)}}} \right) \\
&\quad - c_{2,\mathsf{D}_k}^{(n)} \frac{\psi_{\mathsf{D}_k}(\mathbf{w}_2,\mathbf{p}) + \left(\Re\{\tilde{\mathbf{h}}_{\mathsf{D}_k}^H \mathbf{w}_{2,k}\}\right)^2}{\sqrt{\beta\lambda}} \quad (59) \\
&:= \hat{f}_{2,\mathsf{D}_k}^{(n)}(\mathbf{w}_2,\mathbf{p},\beta,\lambda) \quad (60)
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

where $a_{2,\mathsf{D}_k}^{(n)}, b_{2,\mathsf{D}_k}^{(n)}$, and $c_{2,\mathsf{D}_k}^{(n)}$ are defined in (35). Next, similarly to (42), we have

$$\frac{\ln\big(1+\gamma_\ell(\mathbf{w}_2,\mathbf{p})\big)}{\sqrt{\beta\lambda}} \geq \hat{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p},\beta,\lambda) \tag{61}$$

where

$$\hat{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p},\beta,\lambda) = \frac{a_\ell^{(n)}}{\sqrt{\beta\lambda}} + b_\ell^{(n)}\left(2\frac{\sqrt{p_\ell^{(n)}}\sqrt{p_\ell}}{\sqrt{\beta^{(n)}\lambda^{(n)}}} - \frac{p_\ell^{(n)}\beta}{2(\beta^{(n)})^{3/2}\sqrt{\lambda^{(n)}}} - \frac{p_\ell^{(n)}\lambda}{2(\lambda^{(n)})^{3/2}\sqrt{\beta^{(n)}}}\right) - \frac{\varphi_\ell^{(n)}(\mathbf{w}_2,\mathbf{p})}{\sqrt{\beta\lambda}} \tag{62}$$

with $a_\ell^{(n)}, b_\ell^{(n)}$, and $\varphi_\ell^{(n)}(\mathbf{w}_2,\mathbf{p})$ are defined in (39).

Turning attention to the constraints in (56b), we see that the nonconvex constraints (23b) and (23c) were innerly approximated by convex constraints (41) and (45), respectively. Then, we need to deal with the two last nonconvex constraints (54) and (55), which are innerly approximated in the same manner to (47) by

$$\begin{aligned}&\frac{\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2+\|\mathbf{w}_2\|^2}{\epsilon} + \frac{\|\mathbf{w}_1^{(n)}\|^2+\|\mathbf{V}^{(n)}\|^2}{2\epsilon\sqrt{\beta^{(n)}}}\\ &+\frac{\|\mathbf{w}_2^{(n)}\|^2}{2\epsilon\sqrt{\tau^{(n)}}} - 2\frac{\Re\big\{(\mathbf{w}_1^{(n)})^H\mathbf{w}_1\big\}+\Re\big\{\mathrm{Trace}\big((\mathbf{V}^{(n)})^H\mathbf{V}\big)\big\}}{\epsilon\sqrt{\beta^{(n)}}}\\ &+\frac{\big(\|\mathbf{w}_1^{(n)}\|^2+\|\mathbf{V}^{(n)}\|^2\big)\beta}{2\epsilon(\beta^{(n)})^{3/2}} - 2\frac{\Re\big\{(\mathbf{w}_2^{(n)})^H\mathbf{w}_2\big\}}{\epsilon\sqrt{\tau^{(n)}}} + \frac{\|\mathbf{w}_2^{(n)}\|^2\tau}{2\epsilon(\tau^{(n)})^{3/2}}\\ &+MP_{\mathsf{BS}}^{\mathsf{dyn}}\alpha + \Big(\sum_{\ell=1}^{L}P_{\mathsf{U}_\ell}^{\mathsf{dyn}}\Big)(1-\alpha)+P_0 \leq \sqrt{\lambda}\end{aligned} \tag{63}$$

and

$$\begin{aligned}&\|\mathbf{w}_1\|^2+\|\mathbf{V}\|^2+\|\mathbf{w}_2\|^2 + \frac{\|\mathbf{w}_1^{(n)}\|^2+\|\mathbf{V}^{(n)}\|^2}{2\sqrt{\beta^{(n)}}} + \frac{\|\mathbf{w}_2^{(n)}\|^2}{2\sqrt{\tau^{(n)}}}\\ &-2\frac{\Re\big\{(\mathbf{w}_1^{(n)})^H\mathbf{w}_1\big\}+\Re\big\{\mathrm{Trace}\big((\mathbf{V}^{(n)})^H\mathbf{V}\big)\big\}}{\sqrt{\beta^{(n)}}}\\ &+\frac{\big(\|\mathbf{w}_1^{(n)}\|^2+\|\mathbf{V}^{(n)}\|^2\big)\beta}{2(\beta^{(n)})^{3/2}} - 2\frac{\Re\big\{(\mathbf{w}_2^{(n)})^H\mathbf{w}_2\big\}}{\sqrt{\tau^{(n)}}} + \frac{\|\mathbf{w}_2^{(n)}\|^2\tau}{2(\tau^{(n)})^{3/2}}\\ &\leq P_{\mathsf{BS}}.\end{aligned} \tag{64}$$

In summary, at the $(n+1)$-th iteration we solve the following innerly approximated problem for (56):

$$\begin{aligned}\underset{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta,\lambda}{\text{maximize}} \;&\sum_{k=1}^{K}\Big(\hat{f}_{1,\mathsf{D}_k}^{(n)}(\mathbf{w}_1,\mathbf{V},\tau,\lambda)+\hat{f}_{2,\mathsf{D}_k}^{(n)}(\mathbf{w}_2,\mathbf{p},\beta,\lambda)\Big)\\ &+\sum_{\ell=1}^{L}\hat{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p},\beta,\lambda)\end{aligned} \tag{65a}$$

$$\text{s.t. } (14\text{d}),(14\text{f}),(25),(32),(41),(45),(52),(53),(63),(64). \tag{65b}$$

To find an initial feasible point to (19), we solve the following convex optimization problem:

$$\begin{aligned}\max_{\mathbf{w}_1,\mathbf{w}_2,\mathbf{V},\mathbf{p},\alpha,\tau,\beta}\; \min_{\ell=1,\cdots,L}\; &\Big\{\tilde{f}_\ell^{(n)}(\mathbf{w}_2,\mathbf{p}) - \frac{\bar{r}_{\mathsf{u}}}{1-\alpha},\\ &\eta\phi_\ell^{(n),\mathrm{eh}}(\mathbf{w}_1,\mathbf{V},\alpha) - \frac{p_\ell^2}{\alpha}\Big\}\end{aligned} \tag{66a}$$

$$\text{s.t.} \quad (14\text{d}),(14\text{f}),(52),(64). \tag{66b}$$

Algorithm 2 Path-following algorithm for EEM (19)

Initinalization: Set $n := 0$ and solve (66) to generate an initial feasible point $\big(\mathbf{w}_1^{(n)},\mathbf{w}_2^{(n)},\mathbf{V}^{(n)},\mathbf{p}^{(n)},\tau^{(n)},\beta^{(n)},\lambda^{(n)}\big)$.

1: **repeat**
2: Solve (65) to obtain the optimal solutions: $\big(\mathbf{w}_1^\star,\mathbf{w}_2^\star,\mathbf{V},\mathbf{p}^\star,\tau^\star,\beta^\star,\lambda^\star\big)$.
3: Update: $\mathbf{w}_1^{(n+1)} := \mathbf{w}_1^\star$, $\mathbf{w}_2^{(n+1)} := \mathbf{w}_2^\star, \mathbf{V}^{(n+1)} := \mathbf{V}^\star, \mathbf{p}^{(n+1)} := \mathbf{p}^\star$, $\alpha^{(n+1)} := \alpha^\star$, $\tau^{(n+1)} := \tau^\star$, $\beta^{(n+1)} := \beta^\star$, and $\lambda^{(n+1)} := \lambda^\star$.
4: Set $n := n+1$.
5: **until** Convergence

and stop upon reaching

$$\begin{aligned}\min_{\ell=1,\ldots,L}\Bigg\{&\tilde{f}_\ell^{(n)}\Big(\mathbf{w}_2^{(n+1)},\mathbf{p}^{(n+1)}\Big) - \frac{\bar{r}_{\mathsf{u}}}{1-\alpha^{(n+1)}},\\ &\eta\phi_\ell^{(n),\mathrm{eh}}\Big(\mathbf{w}_1^{(n+1)},\mathbf{V}^{(n+1)},\alpha^{(n+1)}\Big) - \frac{\big(p_\ell^{(n+1)}\big)^2}{\alpha^{(n+1)}}\Bigg\} \geq 0.\end{aligned} \tag{67}$$

We outline the proposed iterative method in Algorithm 2. Analogously to Proposition 1, it can be shown that Algorithm 2 generates a sequence $\Big\{\big(\mathbf{w}_1^{(n)},\mathbf{w}_2^{(n)},\mathbf{V}^{(n)},\mathbf{p}^{(n)},\alpha^{(n)},\tau^{(n)},\beta^{(n)},\lambda^{(n)}\big)\Big\}$ of improved points of (19), which converges to a KKT point. The optimization problem (65) involves $\tilde{n}=(3N+M)K+(N+M)\tilde{L}+2L+5$ scalar real variables and $\tilde{m}=2K+3L+6$ quadratic and linear constraints, so that the per-iteration cost of solving (65) is $\mathcal{O}(\tilde{n}^2\tilde{m}^{2.5}+\tilde{m}^{3.5})$.

Remark 1: We have observed in our numerical experiments that solving (50) and (66) requires no more than three iterations to generate a feasible initial point of (14) and (19) in all cases.

V. Numerical Results

This section evaluates the numerical performance of the proposed algorithms. The used convex solver is MOSEK [36] in the MATLAB environment. The simulation results are derived by averaging over 10000 runs. The channel vectors from the BS to a user and from ULU to DLU are assumed to undergo the path loss model for line of sight (LOS) and non-line-of-sight (NLOS), respectively [26], [28]. Namely, the channel vector from the BS to DLU D_k is modeled by $\mathbf{h}_{\mathsf{D}_k} = \sqrt{\hbar_{\mathsf{D}_k}}\bar{\mathbf{h}}_{\mathsf{D}_k}$, where the entries of $\bar{\mathbf{h}}_{\mathsf{D}_k}$ are independent circularly symmetric complex Gaussian (CSCG) random variables with zero means and unit variances, and $\hbar_{\mathsf{D}_k} = 10^{(-\sigma_{\mathsf{LOS}}/10)}$ represents the path loss. The channels $\mathbf{g}_{\mathsf{U}_\ell}$ and $g_{\ell k}$ are generated similarly as $\mathbf{g}_{\mathsf{U}_\ell} = \sqrt{\hbar_{\mathsf{U}_\ell}}\bar{\mathbf{g}}_{\mathsf{U}_\ell}$ and $g_{\ell k} = \sqrt{\hbar_{\ell k}}\bar{g}_{\ell k}$, where the entries of $\bar{\mathbf{g}}_{\mathsf{U}_\ell}$ and $\bar{g}_{\ell k}$ are independent CSCG random variables with distribution $\mathcal{CN}(0,1)$, and their path losses are $\hbar_{\mathsf{U}_\ell} = 10^{(-\sigma_{\mathsf{LOS}}/10)}$ and $\hbar_{\ell k} = 10^{(-\sigma_{\mathsf{NLOS}}/10)}$. The entries of the fading loop channel $\mathbf{G}_{\mathsf{I}}$ are independently drawn from the CSCG distribution $\mathcal{CN}(0,1)$ [20]. Unless stated otherwise, the parameters given in Table I follow those studied in [11], [12], [15], [32]. The bandwidth of the system was omitted in the previous sections to simplify the notation without affecting the

solution because it is a constant. In Table I, d ($\tilde{d}$, resp.) is the distance between the BS and a user (between ULU and DLU, resp.). We simulate small-cell scenarios, so $5 \leq d \leq 30$ m. To emphasize the effectiveness of using all available antennas in simultaneous information transmission and energy transfer in phase I, we also compare the performance of our proposed FD system with other two options:

- The BS uses N antennas for transmitting information to DLUs and M antennas for transferring energy to ULUs in phase I. We call this option "conventional FD".
- The BS uses all available antennas for transferring energy to ULUs in phase I only, i.e., $\mathbf{w}_1 = 0$ is set. We call this option "no downlink WIT (DWIT) in phase I".

Additionally, we also compare the performance of our proposed FD system with that of an HD system in which the BS uses all the antennas, i.e., $M + N$, for communication, and it uses half time and half power for transmitting information to DLUs and the remaining half time is for transferring the energy by half power and then receiving information from ULUs. Accordingly, there are two separate DL and UL sum rate optimization problems:

$$\max_{\mathbf{w}_1} \tilde{R}_{\mathsf{DL}}(\mathbf{w}_1) \triangleq \frac{1}{2}\sum_{k=1}^{K} \ln\Big(1 + \frac{|\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,k}|^2}{\sum_{i=1,i\neq k}^{K} |\mathbf{h}_{\mathsf{D}_k}^H \mathbf{w}_{1,i}|^2 + \sigma_k^2}\Big) \quad (68a)$$

$$\text{s.t. } \frac{1}{2}\|\mathbf{w}_1\|^2 \leq P_{\mathsf{BS}}/2 \quad (68b)$$

for communication with DLUs, and

$$\max_{\mathbf{V},\mathbf{p},\alpha} \tilde{R}_{\mathsf{UL}}(\mathbf{V},\mathbf{p},\alpha) \triangleq (1/2-\alpha) \times \sum_{\ell=1}^{L} \ln\Big(1 + p_\ell^2 \mathbf{h}_{\mathsf{U}_\ell}^H \big(\sum_{j>\ell}^{L} p_j^2 \mathbf{h}_{\mathsf{U}_j}\mathbf{h}_{\mathsf{U}_j}^H + \tilde{\sigma}^2\mathbf{I}\big)^{-1}\mathbf{h}_{\mathsf{U}_\ell}\Big) \quad (69a)$$

$$\text{s.t. } (1/2-\alpha)\ln\Big(1 + p_\ell^2 \mathbf{h}_{\mathsf{U}_\ell}^H \big(\sum_{j>\ell}^{L} p_j^2 \mathbf{h}_{\mathsf{U}_j}\mathbf{h}_{\mathsf{U}_j}^H + \tilde{\sigma}^2\mathbf{I}\big)^{-1}\mathbf{h}_{\mathsf{U}_\ell}\Big) \geq \bar{r}_{\mathsf{u}}, \forall \ell = 1,\cdots,L, \quad (69b)$$

$$p_\ell^2 \leq \frac{\eta\alpha}{1/2-\alpha}\|\mathbf{g}_{\mathsf{U}_\ell}^H \mathbf{V}\|^2, \ \forall \ell = 1,\cdots,L, \quad (69c)$$

$$\alpha\|\mathbf{V}\|^2 \leq P_{\mathsf{BS}}/2, \quad (69d)$$

$$0 < \alpha < 1/2, p_\ell \geq 0, \ \forall \ell = 1,\cdots,L \quad (69e)$$

for communication with ULUs. Their optimal values are then summed for the HD SR performance. Next, the EEM problem for HD system is thus

$$\underset{\mathbf{w}_1,\mathbf{V},\mathbf{p},\alpha}{\text{maximize}} \ \frac{\tilde{R}_{\mathsf{DL}}(\mathbf{w}_1) + \tilde{R}_{\mathsf{UL}}(\mathbf{V},\mathbf{p},\alpha)}{\chi_{\mathsf{HD}}(\mathbf{w}_1,\mathbf{V},\alpha) + P_{\mathsf{BS}}^{\mathsf{sta}} + \sum_{\ell=1}^{L} P_{\mathsf{U}_\ell}^{\mathsf{sta}}} \quad (70a)$$

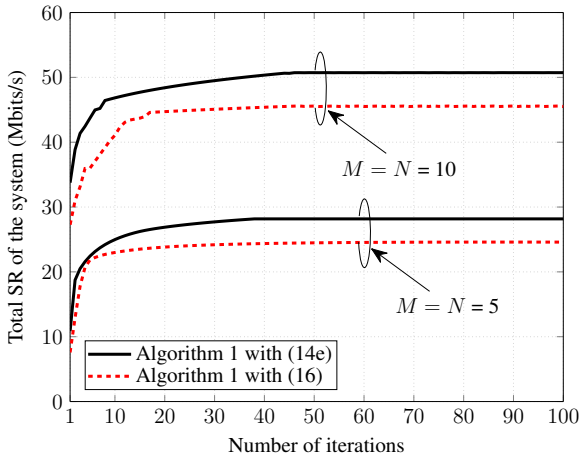
$$\text{s.t.} \quad (68b), (69b), (69c), (69d), (69e) \quad (70b)$$

where $\chi_{\mathsf{HD}}(\mathbf{w}_1,\mathbf{V},\alpha) \triangleq \frac{1}{\epsilon}\big(\frac{1}{2}\|\mathbf{w}_1\|^2 + \alpha\|\mathbf{V}\|^2\big) + (1/2+\alpha)(M+N)P_{\mathsf{BS}}^{\mathsf{dyn}} + (1/2-\alpha)\sum_{\ell=1}^{L} P_{\mathsf{U}_\ell}^{\mathsf{dyn}}$.

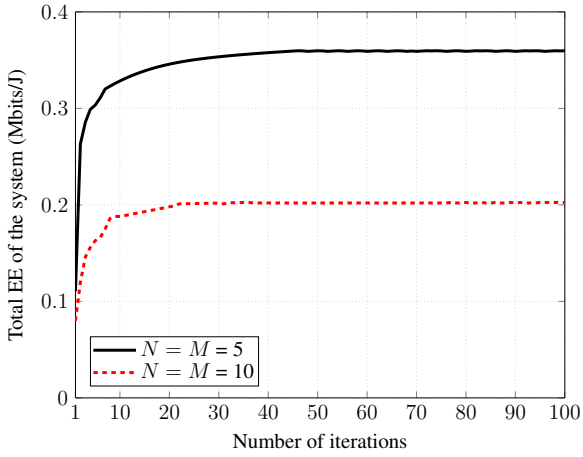
It is obvious that problems (68) and (69) can be solved by Algorithm 1 while problem (70) can be solved by Algorithm 2.

TABLE I
SIMULATION PARAMETERS

Parameters	Value
System bandwidth	1 [MHz]
Noise variances $(\sigma_k^2, \tilde{\sigma}^2)$	-80 [dBm]
Dynamic power consumptions: $P_{\mathsf{BS}}^{\mathsf{dyn}}$ and $P_{\mathsf{U}_\ell}^{\mathsf{dyn}}$	10 and 7 [dBm]
Static power consumptions: $P_{\mathsf{BS}}^{\mathsf{sta}}$ and $P_{\mathsf{U}_\ell}^{\mathsf{sta}}$	15 and 5 [dBm]
Power amplifier efficiency (ϵ)	0.69
Energy conversion efficiency (η)	1/2
Path loss from the BS to an user (σ_{LOS})	30.18 + 26log(d) [dB]
Path loss from the U_ℓ to the D_k $(\tilde{\sigma}_{\mathsf{NLOS}})$	145.4 + 37.5log$(\tilde{d})$ [dB]
The power budge at the BS, P_{BS}	30 [dBm]



(a) Convergence of Algorithm 1 for different numbers of antennas at the BS.



(b) Convergence of Algorithm 2 for different numbers of antennas at the BS.

Fig. 3. Convergence of Algorithms 1 and 2 for different numbers of antennas at the BS with $K = L = 3, \rho = -40$ dB, $\bar{r}_{\mathsf{u}} = 1$ Mbits/s.

A. Convergence Results

Figs. 3(a) and 3(b) depict typical convergence behavior of the proposed Algorithm 1 and Algorithm 2 in the scenario described in their caption. Both algorithms converge within tens of iterations. The convergence rate is not sensitive to the problem size which is proportional to N and M. The objectives are iteratively improved as expected. Fig. 3 also

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

TABLE II
AVERAGE RATIO (IN %) BETWEEN THE POWER USED AND POWER BUDGET FOR DIFFERENT NUMBERS OF ANTENNAS AT THE BS.

$N = M$	2	4	6	8	10	12
Algorithm 1 with (16)	43.22	50.85	61.13	69.40	59.72	57.56
Algorithm 1 with (14e)	100.00	100.00	100.00	100.00	100.00	100.00

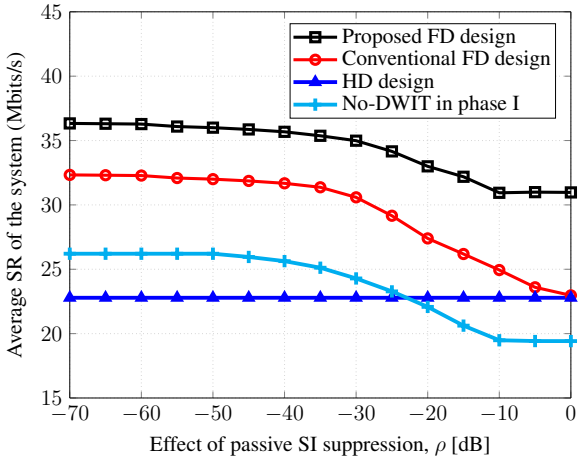
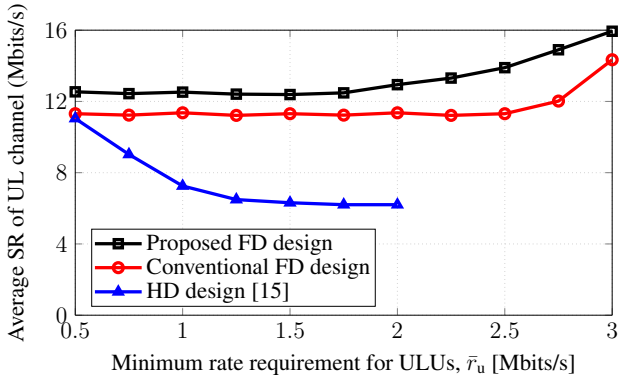


Fig. 4. SR versus ρ with $N = M = 4, K = L = 3, \bar{r}_{\rm u} = 1$ Mbits/s.

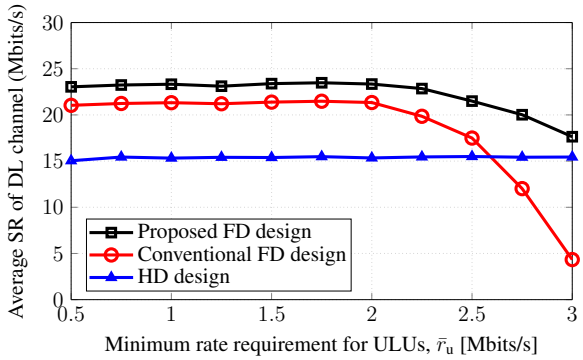
shows that at least 95% of the SR and EE values are reached within 30 iterations. Of course, Algorithm 1 using the real power constraint (14e) offers better performance compared to that using the relaxed power constraint (16) as Fig. 3(a) particularly shows because the BS can use all allowable power with (14e). In particular, Table II shows that using (16) can exploit only 69.40% of the allowable power with $N = M = 8$ while that of (14e) always uses 100% in all cases. This also supports the statement made in Section II-B, on the use of (14e) instead of (16).

B. Sum Rate Maximization

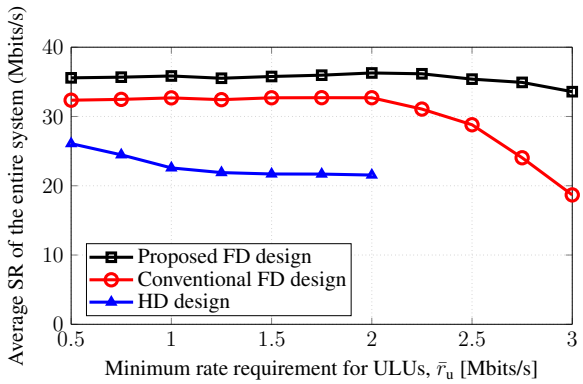
Fig. 4 illustrates the effect of FD residual SI. When the SI degree ρ is relatively small ($\rho \leq -40$ dB), the SR of FD systems is better than that of the HD system by about 59.48% and 41.92%. The SR of FD is degraded as the residual SI becomes larger ($\rho > -40$ dB). The SR of no DWIT in phase I is seen to yield slightly better SR performance than HD for $\rho \leq -23$ dB, but exhibits performance degradation, which worsens as $\rho > -23$ dB. The high transmit power in the DL transmission in phase II causes severe interference to the received signals of the ULUs at the BS. Thus, the BS needs to scale down the transmit power in phase II to meet QoS for ULUs, compromising the SR of FD. Another interesting observation is that the gap between the proposed FD and conventional FD designs increases in ρ. The main reason is that for such a case, the SR of the DLUs is mostly contributed by the DL channel in phase I with no effect from ρ. Therefore, the proposed FD is much more robust against the residual SI effect. The simulation results in Fig. 4 further confirm that incorporating transmit and receive antennas at the BS in phase I is a powerful means to combat the degree-of-freedom (DoF)



(a) SR of UL channel.



(b) SR of DL channel.

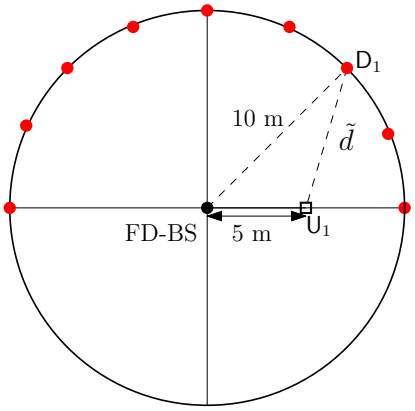


(c) SR of the entire system.

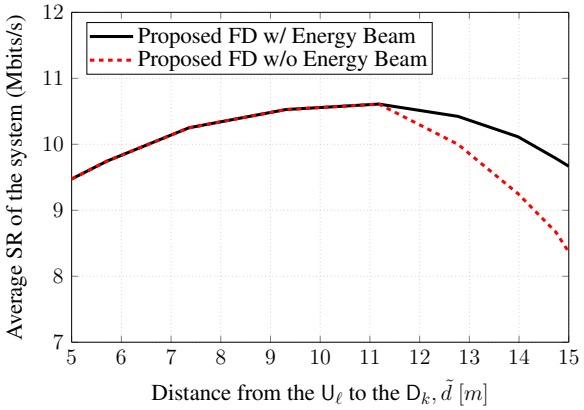
Fig. 5. SR versus the minimum rate requirement for ULUs, with $N = M = 4, K = L = 3, \rho = -40$ dB.

bottleneck for leveraging multiuser diversity.

The SR versus QoS threshold $\bar{r}_{\rm u} \in [0.5,\ 3]$ Mbits/s is shown by Fig. 5. Fig. 5(a) shows that (69) is infeasible for $\bar{r}_{\rm u} > 2$ Mbits/s, i.e., the HD system cannot offer such high QoS to ULUs. In contrast, the SR of the FD system monotonically increases even for $\bar{r}_{\rm u} > 2$ Mbits/s. As QoS $\bar{r}_{\rm u} \leq 2$ Mbits/s is easily met, FD-BS pays less attention to the ULUs. For higher QoS $\bar{r}_{\rm u} > 2$ Mbits/s, the FD-BS must pay more attention to serving ULUs by transferring more power to them in phase I and scaling down the transmitted power to DLUs in phase II. This results in a dramatic degradation of the SRs of the FD systems for the DL channel in Fig. 5(b). Although the conventional FD system is worse than the HD one when $\bar{r}_{\rm u} > 2.6$ Mbits/s in terms of the SR, the entire SRs of the former



(a) Simulation setup for Fig. 6(b) in the 2D plane.



(b) Average SR of the system.

Fig. 6. SR versus the distance from the ULU to the DLU with $N = M = 4, K = L = 1, \rho = -40$ dB, $\bar{r}_{\mathsf{u}} = 2$ Mbits/s.

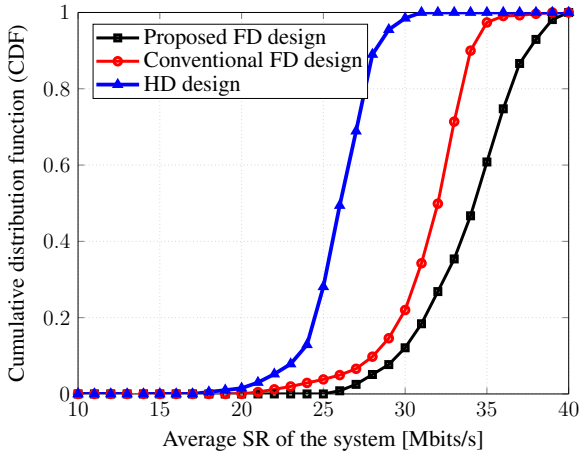


Fig. 7. CDF of SR with $N = M = 4, K = L = 3, \rho = -40$ dB, $\bar{r}_{\mathsf{u}} = 1$ Mbits/s.

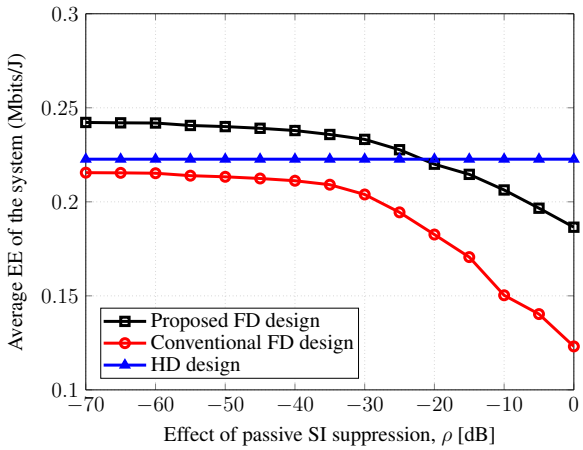


Fig. 8. EE versus ρ with $N = M = 4, K = L = 3, \bar{r}_{\mathsf{u}} = 1$ Mbits/s.

still outperform that of the latter, as shown in Fig. 5(c).

As mentioned earlier, the harvested energy of the ULUs is also contributed by the signals intended for DLUs, and it is of interest to investigate how the energy beam $\mathbf{v}_e$ affects the system performance. We set a scenario as in Fig. 6(a) with one ULU, which is fixed and one DLU, which moves on a circle with a radius of 10 m. The SRs of the proposed FD design with and without using an energy beam versus the distance between the ULU and DLU is shown in Fig. 6(b). Both of them first increase as the DLU moves further from the ULU and reach a peak at $\tilde{d} = 11.18$ m. Increasing $\tilde{d}$ leads to a decrease in the CCI, which is then enhanced for better SR of the DL channel. The two curves coincide for $\tilde{d} \leq 11.18$ m, where the ULU is able to harvest sufficient energy from the DL signal and the energy beam is not really useful. However, they decrease for $\tilde{d} > 11.18$ m because the ULU could harvest only a small amount of energy from the DL signal while the influence of CCI is negligible, making ULU information throughput at the BS extremely low. Certainly, the proposed FD system by using the energy beam achieves better SR than that without it and the gap between the two is even deeper. Without using the energy beam, the BS transmits not only the information for the DLU but also transfers energy to the ULU, which then requires the beamforming vectors to spread over a wide geographical range causing an SR loss of the FD system. This means that using the energy beam is necessary when $\tilde{d}$ is relatively large.

Fig. 7 plots the cumulative distribution function (CDF) of the SR of the FD and HD systems. As expected, the proposed FD design outperforms the conventional FD and HD systems. Specifically, the proposed FD design reaches 3 Mbits/s and 9.5 Mbits/s of SR higher than the conventional FD system and HD system, respectively, in about 80% of the simulated trials.

C. Energy Efficiency Maximization

Fig. 8 shows the comparison of the EE performance between the FD systems and HD one at different SI levels ρ. The proposed FD design achieves better EE performance than the HD design for $\rho \leq -22$ dB. However, the former performs worse than the latter for $\rho > -22$ dB, where the FD system's strong SI hurts the energy efficiency. For whatever SI level, the EE of the proposed FD system still outperforms the conventional FD one.

Next, we study EE versus the dynamic power consumption $P_{\mathsf{BS}}^{\mathsf{dyn}}$ and $P_{\mathsf{U}_\ell}^{\mathsf{dyn}}$ for different designs in Fig. 9 under SI level $\rho = -40$ dB. The proposed FD design offers better gain over the conventional FD and HD designs, especially when

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

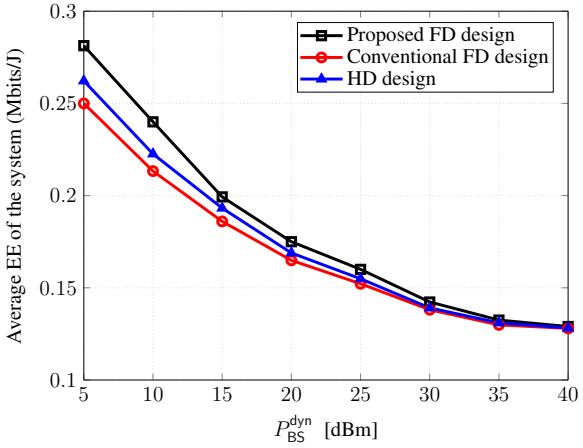


Fig. 9. EE versus $P_{\mathsf{BS}}^{\mathsf{dyn}}$ where we set $P_{\mathsf{U}_\ell}^{\mathsf{dyn}} = \xi P_{\mathsf{BS}}^{\mathsf{dyn}}, \forall \ell$ and $\xi = 0.7$, with $N = M = 4, K = L = 3, \rho = -40$ dB, $\bar{r}_{\mathsf{u}} = 1$ Mbits/s.

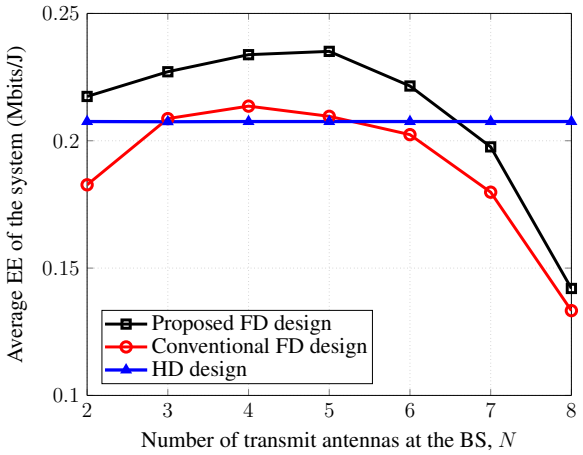


Fig. 10. EE versus N for $N + M = 10$, with $K = L = 3, \rho = -40$ dB, $\bar{r}_{\mathsf{u}} = 1$ Mbits/s.

the dynamic power consumption is relatively small. For small $P_{\mathsf{BS}}^{\mathsf{dyn}}$ and $P_{\mathsf{U}_\ell}^{\mathsf{dyn}}$, the total power consumption in (17) and (18) is mostly determined by the beamformer power, i.e., $\alpha(\|\mathbf{w}_1\|^2 + \|\mathbf{V}\|^2) + (1-\alpha)\|\mathbf{w}_2\|^2$, so it is energy-efficient to use a low transmit power since the BS has more antennas to transmit the signals in phase I as discussed in Section II. When the dynamic power consumption becomes large, all the system designs attain almost equal energy efficiency. The circuit power consumption suppresses the power consumed by beamfomers and thus, it significantly impacts the EE of all the systems.

Finally, EE versus the number of transmit antennas at the BS, N, for a fixed $N + M = 10$ is shown in Fig. 10. The EE of the two FD designs first increases to a certain value of N and then decreases beyond this value. Intuitively, the optimal number N of transmit antennas is 5 for the proposed FD design and 4 for the conventional FD design for this setting. Of course they may be different in other settings. When N becomes large, the EE of both FD systems is decreased because employing more antennas for phase I improves SR but consumes more the total power consumption, i.e., the dynamic power consumption. In addition, the gap between two FD designs is reduced for large N (also a small M) since the conventional FD design is nearly the same as the cooperative FD one. Obviously, the EE of the HD system is unchanged with varying N since its BS always uses all available antennas (i.e., $N + M$, which is fixed) for communication.

Remark 2: Fig. 10 particularly shows that increasing the number of transmit antennas may degrade the energy efficiency of the FD systems as it leads to increasing the total power consumption at the BS. Joint beamforming design and antenna selection for maximizing the energy efficiency is certainly interesting but beyond the scope of this paper.

VI. Conclusion

In this paper, a MIMO FD BS in WPCNs has been studied in which the users in the UL channel are designed to harvest energy from the BS before transmitting their information. To combat the degree-of-freedom bottleneck, a cooperative transmit strategy between the DL and UL transmission has been proposed. We have developed path-following algorithms for jointly designing the energy harvesting time and beamforming to maximize the sum rate and energy efficiency of the FD system. The proposed algorithms are guaranteed to converge monotonically to at least local optima of the nonconvex design problems. Numerical results have been presented to show the fast convergence rate and demonstrate the advantages of our proposed algorithms. These observations further confirm the efficacy of the proposed cooperative transmission in enhancing system performance, as well as the necessity of optimizing energy beamforming in order to provide significant performance improvement compared to the existing solutions.

Appendix A
Proof for inequalities (20) and (21)

The function $f(t) = -\ln(1-t)$ is obviously convex and increasing in the domain $0 \le t < 1$, while the function $g(x,z) = |x|^2/z$ is convex. Therefore, the composite function $f\big(g(x,z)\big) = -\ln\big(1 - |x|^2/z\big)$ is convex in the domain $z > |x|^2$ [37], for which

$$\begin{aligned} -\ln\Big(1-\frac{|x|^2}{z}\Big) &\ge -\ln\Big(1-\frac{|\bar{x}|^2}{\bar{z}}\Big) \\ &\quad + \big\langle \nabla f\big(g(\bar{x},\bar{z})\big), (x,z) - (\bar{x},\bar{z})\big\rangle \\ &= -\ln\Big(1-\frac{|\bar{x}|^2}{\bar{z}}\Big) - \frac{|\bar{x}|^2}{\bar{z}-|\bar{x}|^2} + 2\frac{\Re\{\bar{x}^* x\}}{\bar{z}-|\bar{x}|^2} \\ &\quad - \frac{|\bar{x}|^2 z}{(\bar{z}-|\bar{x}|^2)\bar{z}}. \end{aligned} \tag{71}$$

By noting

$$\ln\Big(1+\frac{|x|^2}{y}\Big) = -\ln\Big(1-\frac{|x|^2}{y+|x|^2}\Big) \tag{72}$$

(20) is obtained by applying (71) for $z = y + |x|^2$ and $\bar{z} = \bar{y} + |\bar{x}|^2$.

Furthermore, as $g(x,y) = |x|^2/y$ is convex in $x \in \mathbb{C}$ and $y > 0$, it is true that [37]

$$\begin{aligned} \frac{|x|^2}{y} &\ge \frac{|\bar{x}|^2}{\bar{y}} + \langle \nabla g(\bar{x},\bar{y}), (x,y) - (\bar{x},\bar{y})\rangle \\ &= 2\frac{\bar{x}^* x}{\bar{y}} - \frac{|\bar{x}|^2}{\bar{y}^2} y \end{aligned} \tag{73}$$

showing (21), and the proof is completed.

APPENDIX B
PROOF OF PROPOSITION 1

For compact representation we use the notation $\mathbf{\Psi}$ and $\mathbf{\Psi}^{(n)}$ to refer to $\big(\mathbf{w}_1, \mathbf{w}_2, \mathbf{V}, \mathbf{p}, \alpha, \tau, \beta\big)$ and $\big(\mathbf{w}_1^{(n)}, \mathbf{w}_2^{(n)}, \mathbf{V}^{(n)}, \mathbf{p}^{(n)}, \alpha^{(n)}, \tau^{(n)}, \beta^{(n)}\big)$, respectively. Denote by $\mathcal{F}(\mathbf{\Psi})$ and $\mathcal{F}^{(n)}(\mathbf{\Psi})$ the objectives of (23) and (48), respectively. We have

$$\mathcal{F}\big(\mathbf{\Psi}\big) \geq \mathcal{F}^{(n)}\big(\mathbf{\Psi}\big),\ \forall\, \mathbf{\Psi} \tag{74}$$

and

$$\mathcal{F}\big(\mathbf{\Psi}^{(n)}\big) = \mathcal{F}^{(n)}\big(\mathbf{\Psi}^{(n)}\big). \tag{75}$$

Therefore,

$$\begin{aligned}\mathcal{F}\big(\mathbf{\Psi}^{(n+1)}\big) &\geq \mathcal{F}^{(n)}\big(\mathbf{\Psi}^{(n+1)}\big)\\ &\geq \mathcal{F}^{(n)}\big(\mathbf{\Psi}^{(n)}\big)\\ &= \mathcal{F}\big(\mathbf{\Psi}^{(n)}\big)\end{aligned} \tag{76}$$

where the second inequality follows from the fact that $\mathbf{\Psi}^{(n+1)}$ and $\mathbf{\Psi}^{(n)}$ are the optimal solution and feasible point of (48), respectively. This result shows that $\mathbf{\Psi}^{(n+1)}$ is a better point for (23) than $\mathbf{\Psi}^{(n)}$. Since the sequence $\big\{\mathbf{\Psi}^{(n)}\big\}$ is bounded, by Cauchy's theorem, there is a convergent subsequence $\big\{\mathbf{\Psi}^{(n_\nu)}\big\}$ with a limit point $\bar{\mathbf{\Psi}}$, i.e.,

$$\lim_{\nu\to+\infty}\Big[\mathcal{F}(\mathbf{\Psi}^{(n_\nu)}) - \mathcal{F}(\bar{\mathbf{\Psi}})\Big] = 0. \tag{77}$$

For every n there is ν such that $n_\nu \leq n \leq n_{\nu+1}$, so

$$\begin{aligned}0 &= \lim_{\nu\to+\infty}\Big[\mathcal{F}(\mathbf{\Psi}^{(n_\nu)}) - \mathcal{F}(\bar{\mathbf{\Psi}})\Big]\\ &\leq \lim_{n\to+\infty}\Big[\mathcal{F}(\mathbf{\Psi}^{(n)}) - \mathcal{F}(\bar{\mathbf{\Psi}})\Big]\\ &\leq \lim_{\nu\to+\infty}\Big[\mathcal{F}(\mathbf{\Psi}^{(n_{\nu+1})}) - \mathcal{F}(\bar{\mathbf{\Psi}})\Big]\\ &= 0\end{aligned} \tag{78}$$

showing that $\lim_{n\to+\infty}\mathcal{F}(\mathbf{\Psi}^{(n)}) = \mathcal{F}(\bar{\mathbf{\Psi}})$. Then, each accumulation point $\bar{\mathbf{\Psi}}$ of the sequence $\big\{\mathbf{\Psi}^{(n)}\big\}$ is a KKT-point according to [38, Theorem 1]. Proposition 1 is thus proved.

REFERENCES

[1] H. J. Visser and R. J. M. Vullers, "RF energy harvesting and transport for wireless sensor network applications: Principles and requirements," *Proc. IEEE*, vol. 101, no. 6, pp. 1410–1423, June 2013.

[2] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2015.

[3] H. Nishimoto, Y. Kawahara, and T. Asami, "Prototype implementation of ambient RF energy harvesting wireless sensor networks," in *Proc. IEEE Sensors*, Nov. 2010, pp. 1282–1287.

[4] X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Wireless charger networking for mobile devices: Fundamentals, standards, and applications," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 126–135, Apr. 2015.

[5] A. Minasian, S. Shahbazpanahi, and R. Adve, "Energy harvesting cooperative communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6118–6131, Nov. 2014.

[6] Z. Ding, C. Zhong, D. W. K. Ng, M. Peng, H. A. Suraweera, R. Schober, and H. V. Poor, "Application of smart antenna technologies in simultaneous wireless information and power transfer," *IEEE Commun. Magazine*, vol. 53, no. 4, pp. 86–93, 2015.

[7] Z. Ding, S. M. Perlaza, I. Esnaola, and H. V. Poor, "Power allocation strategies in energy harvesting wireless cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 846–860, Feb. 2014.

[8] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, July 2013.

[9] Z. Ding, I. Krikidis, B. Sharif, and H. V. Poor, "Wireless information and power transfer in cooperative networks with spatially random relays," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4440–4453, Aug. 2014.

[10] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Select. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[11] S. Timotheou, I. Krikidis, G. Zheng, and B. Ottersten, "Beamforming for MISO interference channels with QoS and RF energy transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2646–2658, May 2014.

[12] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, June 2014.

[13] S. Bi, Y. Zeng, and R. Zhang, "Wireless powered communication networks: An overview," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 10–18, Apr. 2016.

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

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

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

[17] X. Chen, X. Wang, and X. Chen, "Energy efficient optimization for wireless information and power transfer in large MIMO employing energy beamforming," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 667–670, Dec. 2013.

[18] Q. Wu, M. Tao, D. W. K. Ng, W. Chen, and R. Schober, "Energy-efficient resource allocation for wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2312–2327, Mar. 2016.

[19] S. Akbar, Y. Deng, A. Nallanathan, M. Elkashlan, and A.-H. Aghvami, "Simultaneous wireless information and power transfer in K-tier heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5804–5817, Aug. 2016.

[20] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[21] "System scenarios and technical requirements for full-duplex concept," *DUPLO project, Deliverable D1.1*. [Online]. Available: at http://www.fp7-duplo.eu/index.php/deliverables.

[22] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[23] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[24] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Select. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Feb. 2014.

[25] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.

[26] ——, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, Sept. 2014.

[27] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, Apr. 2015.

[28] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, June 2016.

[29] M. Maso, C.-F. Liu, C.-H. Lee, T. Q. S. Quek, and L. S. Cardoso, "Energy-recyclying full-duplex radios for next-generation networks,"

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2665488, IEEE Transactions on Communications

IEEE J. Select. Areas Commun., vol. 33, no. 12, pp. 2948–2962, Dec. 2015.

[30] M. Mohammadi, B. K. Chalise, H. A. Suraweera, C. Zhong, G. Zheng, and I. Krikidis, "Throughput analysis and optimization of wireless-powered multiple antenna full-duplex relay systems," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1769–1785, Apr. 2016.

[31] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge Univ. Press, UK, 2005.

[32] Q.-D. Vu, L.-N. Tran, R. Farrell, and E.-K. Hong, "An efficiency maximization design for SWIPT," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2189–2193, Dec. 2015.

[33] S. He, Y. Huang, W. Chen, S. Jin, H. Wang, and L. Yang, "Energy efficient coordinated precoding design for a multicell system with RF energy harvesting," *EURASIP J. Wireless Commun. Netw.*, Mar. 2015.

[34] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. 19th Future Network & Mobile Summit (ICT Summit'10)*, Florence, Italy, June 2010, pp. 1–8.

[35] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[36] "I. MOSEK aps," 2014. [Online]. Available: at http://www.mosek.com

[37] H. Tuy, *Convex Analysis and Global Optimization (second edition).* Springer, 2016.

[38] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operation Research*, vol. 26, no. 4, pp. 681–683, 1978.

Van-Dinh Nguyen (S'14) received his B.S. degree in Telecommunications from Bach Khoa University, Vietnam, in 2012, and his M.S. degree in Electronic Engineering from Soongsil University, Seoul, Korea, in 2015. He was a visiting student at Queen's University Belfast, UK (June-July 2015 and August 2016). Currently, he is pursuing his Ph.D. degree in Electronic Engineering at Soongsil University, Seoul, Korea. His research interests include wireless communications, information theory, massive MIMO, stochastic geometry, physical layer security, energy harvesting, cooperative communications, and cognitive radio.

Trung Q. Duong (S'05–M'12–SM'13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Since 2013, he has joined Queen's University Belfast, UK as a Lecturer (Assistant Professor). His current research interests include small-cell networks, physical layer security, energy-harvesting communications, cognitive relay networks. He is the author or co-author of 240 technical papers published in scientific journals (125 articles) and presented at international conferences (115 papers).

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IET COMMUNICATIONS, and a Senior Editor for IEEE COMMUNICATIONS LETTERS. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014, and IEEE Global Communications Conference (GLOBECOM) 2016. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2016-2021).

Hoang Duong Tuan received the Diploma (Hons.) and Ph.D. degrees in applied mathematics from Odessa State University, Ukraine, in 1987 and 1991, respectively. He spent nine academic years in Japan as an Assistant Professor in the Department of Electronic-Mechanical Engineering, Nagoya University, from 1994 to 1999, and then as an Associate Professor in the Department of Electrical and Computer Engineering, Toyota Technological Institute, Nagoya, from 1999 to 2003. He was a Professor with the School of Electrical Engineering and Telecommunications, University of New South Wales, from 2003 to 2011. He is currently a Professor with the Centre for Health Technologies, University of Technology Sydney. He has been involved in research with the areas of optimization, control, signal processing, wireless communication, and biomedical engineering for more than 20 years.

Oh-Soon Shin (S'00-M'10) received his B.S., M.S., and Ph.D. degrees in Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea, in 1998, 2000, and 2004, respectively. From 2004 to 2005, he was with the Division of Engineering and Applied Sciences, Harvard University, MA, USA, as a Postdoctoral Fellow. From 2006 to 2007, he was a Senior Engineer at Samsung Electronics, Suwon, Korea. In September 2007, he joined the School of Electronic Engineering, Soongsil University, Seoul, Korea, where he is currently an Associate Professor. His research interests include communication theory, wireless communication systems, and signal processing for communication.

H. Vincent Poor (S'72, M'77, SM'82, F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. During 2006 to 2016, he served as Dean of Princetons School of Engineering and Applied Science. His research interests are in the areas of information theory, statistical signal processing and stochastic analysis, and their applications in wireless networks and related fields such as smart grid and social networks. Among his publications in these areas is the book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is a member of the National Academy of Engineering, the National Academy of Sciences, and is a foreign member of the Royal Society. He is also a fellow of the American Academy of Arts and Sciences, the National Academy of Inventors, and other national and international academies. He received the Marconi and Armstrong Awards of the IEEE Communications Society in 2007 and 2009, respectively. Recent recognition of his work includes the 2016 John Fritz Medal, the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships at Peking University and Tsinghua University, both conferred in 2016, and a D.Sc. *honoris causa* from Syracuse University awarded in 2017.