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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



# Intelligent Reflecting Surface Aided MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer

Cunhua Pan, Hong Ren, Kezhi Wang, Maged Elkashlan, Arumugam Nallanathan, *Fellow, IEEE*, Jiangzhou Wang, *Fellow*, *IEEE* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—An intelligent reflecting surface (IRS) is invoked for enhancing the energy harvesting performance of a simultaneous wireless information and power transfer (SWIPT) aided system. Specifically, an IRS-assisted SWIPT system is considered, where a multi-antenna aided base station (BS) communicates with several multi-antenna assisted information receivers (IRs), while guaranteeing the energy harvesting requirement of the energy receivers (ERs). To maximize the weighted sum rate (WSR) of IRs, the transmit precoding (TPC) matrices of the BS and passive phase shift matrix of the IRS should be jointly optimized. To tackle this challenging optimization problem, we first adopt the classic block coordinate descent (BCD) algorithm for decoupling the original optimization problem into several subproblems and alternatively optimize the TPC matrices and the phase shift matrix. For each subproblem, we provide a low-complexity iterative algorithm, which is guaranteed to converge to the Karush-Kuhn-Tucker (KKT) point of each subproblem. The BCD algorithm is rigorously proved to converge to the KKT point of the original problem. We also conceive a feasibility checking method to study its feasibility. Our extensive simulation results confirm that employing IRSs in SWIPT beneficially enhances the system performance and the proposed BCD algorithm converges rapidly, which is appealing for practical applications.**

***Index Terms*—Intelligent Reflecting Surface (IRS), Large Intelligent Surface (LIS), SWIPT, Energy Harvesting, MIMO.**

C. Pan, H. Ren, M. Elkashlan and A. Nallanathan are with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (e-mail:{c.pan, h.ren, maged.elkashlan, a.nallanathan}@qmul.ac.uk). K. Wang is with Department of Computer and Information Sciences, Northumbria University, UK. (e-mail: kezhi.wang@northumbria.ac.uk). J. Wang is with the School of Engineering and Digital Arts, University of Kent, Canterbury, Kent, CT2 7NZ, U.K. (e-mail: J.Z.Wang@kent.ac.uk). L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk). A. Nallanathan would like to thank the U.K. Engineering and the Physical Sciences Research Council under Grant EP/N029666/1. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/Noo4558/1, EP/PO34284/1, COALESCE, of the Royal Society's Global Challenges Research Fund Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom. (*Corresponding author: Hong Ren.*)

## I. Introduction

Recently, intelligent reflecting surface (IRS)-assisted wireless communication has received considerable research attention, since it is capable of supporting cost-effective and energy-efficient high data rate communication for next-generation communication systems [1]–[3]. In simple tangible terms, an IRS is composed of a vast number of low-cost and passive reflective components, each of which is capable of imposing a phase change on the signals incident upon them. Thanks to the recent advances in meta-materials [4], it has become feasible to reconfigure the phase shifts in real time. As a result, the phase shifts of all reflective components can be collaboratively adjusted for ensuring that the signals reflected from the IRS can be added constructively or destructively at the receiver in order to beneficially steer the signal component arriving from the base station (BS) for enhancing the desired signal power or alternatively for suppressing the undesired signals, such as interference. In contrast to conventional physical layer techniques that are designed for accommodating the hostile time-varying wireless channels [5], [6], IRSs constitute a new paradigm capable of 'reprogramming' the wireless propagation environment into a more favorable transmission medium. Since the reflective components are passive, they impose a much lower power consumption than conventional relay-aided communication systems relying on active transmission devices. Additionally, no thermal noise is imposed by the IRS, since it directly reflects the incident signals without decoding or amplifying them, which is in contrast to conventional relays. Furthermore, the reflective phase arrays can be fabricated in small size and low weight, which enables them to be easily coated in the buildings' facade, ceilings, walls, etc. Furthermore, as IRS is a complementary device, it can be readily integrated into current wireless networks without modifying the physical layer standardization, making it transparent to the users. To fully exploit the benefits of IRS, the active beamforming at the BS and the passive beamforming at the IRS should be jointly designed. However, the optimization variables are coupled and the joint design leads to a complex

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



optimization problem that is difficult to solve.

Some innovative efforts have been devoted to the transceiver design when integrating IRS into various wireless communication systems, including the single-user scenarios of [7]–[11], the downlink multiple-user scenarios of [12]–[15], the physical layer security design of [16]–[21], the mobile edge computing (MEC) networks of [22], multigroup multicast networks of [23] and the multicell multiuser multiple-input multiple-output (MIMO) case in [24]. Concretely, Wu *et al.* proposed joint active and passive beamforming for a single-user scenario in [7], where semidefinite relaxation (SDR) was proposed for optimizing the phase shift matrix. However, its complexity is high since the number of optimization variables increases quadratically with the number of phase shifts. Additionally, the Gaussian random approximation employed leads to certain performance loss. To resolve this issue, Yu *et al.* [8] proposed a pair of efficient algorithms termed as fixed point iteration and manifold optimization techniques, which can guarantee locally optimal solutions. As a further advance, the authors of [9] considered realistic frequency-selective channels. The phase shift design was studied in [10] when only statistical channel state information (CSI) is available. A sophisticated phase shift model was derived in [11], by taking into account a realistic amplitude-phase relationship. For the multiuser case, the authors in [12] considered the total transmit power minimization problem, while guaranteeing the users' signal-to-interference-plus-noise ratio (SINR) constraints. The associated energy efficiency maximization problem was studied in [13] and zero-forcing beamforming was adopted by the BS for simplifying the optimization problem. By contrast, a weighted sum rate (WSR) maximization problem was considered in [14] and the fairness issues were studied in [15]. The authors of [16]–[18] studied the security issues of a single-user case, while the authors of [19]–[21] considered multiple-user scenarios. In [22], the IRS was shown to be beneficial in reducing the latency of MEC networks. In addition, the IRS can help enhance the WSR performance for the multigroup multicast network in [23]. Most recently, we considered an IRS-assisted multicell MIMO communications scenario [24], where we demonstrated that deploying an IRS at the cell edge is also capable of mitigating the adjacent-cell interference. Channel state information (CSI) is challenging to obtain in IRS-assisted communication system due to its passive feature. There are some initial efforts to handle this issue such as channel estimation and/or robust transmission design [25]–[28]. Specifically, Huang *et al.* [25] proposed a deep learning method for efficient online configuration of the phase shifts, where the phase values can be immediately obtained by inputting the user location into the trained deep neural network. A two-stage channel estimation method based on a sparse matrix factorization and a matrix completion was proposed in [26]. A pair of algorithms based on compressed sensing and deep learning were conceived by Taha *et al.* [27] for tackling the challenging channel estimation issues of IRS-assisted systems. Most recently, we first studied the robust beamforming design for IRS-assisted communication systems in [28], where the imperfect channel from an IRS to users was considered and the channel estimation error was assumed to be within a bounded elliptical region.

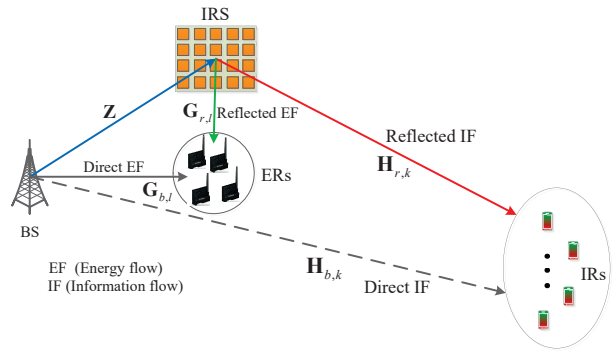


Fig. 1. An IRS-assisted SWIPT system.

On the other hand, information transmission enabled simultaneous wireless information and power transfer (SWIPT) is an appealing technique for future energy-hungry Internet-of-Things (IoTs) networks. Specifically, a base station (BS) with constant power supply will transmit wireless signals to a set of devices. Some devices intend to decode the information from the received signal, which are termed as information receivers (IRs), while the others will harvest the signal energy, which are called energy receiver (ER). In [29], Zhang *et al.* studied the trade-off between the information rate attained and the amount of harvested energy for a single-user MIMO system. In practice, a typical ER such as a humidity sensor requires much higher energy for its operation than that required by IRs. Due to severe channel attenuation, the power received by the ERs is weak, which limits the maximum link-distance of ERs. To mitigate this issue, we propose to deploy an IRS in the vicinity of ERs to provide additional transmission links to support the ERs for enhancing their harvested power as shown in Fig. 1, since there is a paucity of IRS-assisted SWIPT contributions in the literature [30]. Explicitly in [30], the weighted sum power maximization problem was studied by Wu and Zhang, who proved that no dedicated energy-carrying signals were required for an IRS-aided SWIPT system. The SDR method was adopted for solving the optimization problem, which exhibits a high computational complexity as well as imposing a performance degradation due to the associated rank-one extraction. However, this method is not applicable when each user is equipped with multiple antennas. Hence, in this paper we formulate a

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



weighted sum rate (WSR) maximization problem for the IRS-assisted SWIPT MIMO system of Fig. 1, in which an IRS is installed in the vicinity of ERs for compensating the associated power loss, while maximizing the WSR of distant IRs with the aid of passive beamforming.

Against this background, the main contributions of this paper are summarized as follows:

1) We formulate the WSR maximization problem by jointly optimizing the transmit precoding (TPC) matrices of the BS and those of the passive beamforming at the IRS for our IRS-assisted SWIPT MIMO system subject to a non-convex unit-modulus constraint imposed on the phase shifts, while simultaneously satisfying the energy harvesting requirement of the ERs. To the best of our knowledge, this is the first treatise considering the WSR maximization problem of IRS-assisted SWIPT MIMO systems, which is much more challenging than the weighted sum power minimization problem of [30] since the latter can be readily transformed into a convex optimization problem. In contrast to the multicell system of [24], an additional energy harvesting constraint is also imposed in our current study, which further complicates the analysis. Specifically, this constraint is non-convex and the optimization problem may become infeasible. The WSR maximization problem is challenging to solve, since the optimization variables are highly coupled and the data rate expressions of the IRs are complex. To deal with this issue, we first reformulate the original problem into an equivalent form by exploiting the equivalence between the data rate and the weighted minimum mean-square error (WMMSE). Then, an alternating optimization algorithm based on the popular block coordinate descent (BCD) algorithm is proposed for alternatively updating the active TPC matrices of the BS and the phase shift matrix of the IRS, which is rigorously proved to converge to the Karush-Kuhn-Tucker (KKT) point of the original optimization problem.
2) For a given phase shift matrix, we then proceed by developing an iterative algorithm based on the successive convex approximation (SCA) method and on the Lagrangian dual decomposition method to derive a nearly closed-form solution for the TPC matrices. A low-complexity bisection search method is proposed for finding the optimal dual variables. The solutions generated by our iterative algorithm are guaranteed to converge to the KKT point of the TPC optimization problem.
3) For the given TPC matrices, we formulate the phase shift optimization problem as a non-convex quadratically constrained quadratic program (QCQP) subject to an additional energy harvesting constraint by invoking some further matrix manipulations. Then, a novel iterative algorithm based on the majorization-minimization (MM) algorithm [31] and on the price-based method [32] is developed for solving the QCQP. We strictly prove that the final solution generated by the iterative algorithm is guaranteed to converge to the KKT point of the phase shift optimization problem.
4) The associated feasibility issue is also studied by formulating an alternative optimization problem and an iterative algorithm is proposed for solving this problem.
5) Extensive simulation results are provided for verifying the performance advantages of employing IRS in SWIPT in order to enhance the energy harvesting performance. It is shown that the operating range of the ERs can be dramatically expanded by placing IRSs in the ERs' vicinity. Furthermore, the BCD algorithm converges rapidly, and it is eminently suitable for practical applications. Our simulation results also show that as expected, the path loss exponent substantially affects the system's performance and thus the location of the IRS should be carefully chosen.

The remainder of this paper is organized as follows. In Section II, we introduce the IRS-assisted SWIPT system model and our problem formulation. The detailed algorithms used for solving the optimization problem are presented in Section III. The feasibility issues of the original problem are discussed in Section IV, followed by our extensive simulations and discussions in Section V. Finally, our conclusions are provided in Section VI.

*Notations*: For matrix $\mathbf{A}$, $\mathbf{A}^*$ and $\mathbf{A}^\star$ represent the conjugate operator and converged solution, respectively. $\mathrm{Re}\{a\}$ represents the real part of a complex value $a$. $\mathbb{C}^M$ denotes the set of $M \times 1$ complex vectors. $\mathbb{E}\{\cdot\}$ denotes the expectation operation. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \odot \mathbf{B}$ represents Hadamard product of $\mathbf{A}$ and $\mathbf{B}$. $\|\mathbf{A}\|_F$, $\mathrm{tr}\,(\mathbf{A})$ and $|\mathbf{A}|$ denote the Frobenius norm, trace operation and determinant of $\mathbf{A}$, respectively. $\nabla f_{\mathbf{x}}\,(\mathbf{x})$ denotes the gradient of the function $f$ with respect to (w.r.t.) the vector $\mathbf{x}$. $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ represents a random vector following the distribution of zero mean and unit variance matrix. $\arg\{\cdot\}$ means the extraction of phase information. $\mathrm{diag}(\cdot)$ denotes the diagonalization operation. $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote the conjugate, transpose and Hermitian operators, respectively. $\arg(\cdot)$ means the phase extraction operation.

## II. System Model and Problem Formulation

### A. System Model

Consider the IRS-aided multiuser MIMO downlink of a SWIPT system operating over the same frequency band both for data and energy transmission, as shown in Fig. 1. Let us assume that there are $K_I$ IRs and $K_E$ ERs,

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



respectively. It is also assumed that the BS is equipped with $N_B \geq 1$ antennas, while each IR and ER is equipped with $N_I \geq 1$ and $N_E \geq 1$ antennas, respectively. Let us denote the sets of IRs and ERs as $\mathcal{K}_I$ and $\mathcal{K}_E$, respectively. In general, low-power sensors require a certain amount of power (e.g., 0.1 mW) for their real-time operation. Due to the associated severe channel attenuation, the sensors should be deployed sufficiently close to the BS, which limits their practical implementation. To resolve this issue, we propose to employ an IRS, which has $M$ reflective elements in the ERs' vicinity for extending the operational range of sensors, as shown in Fig. 1. Firstly, the IRS increases the energy harvested by the ERs, and additionally it also assists in enhancing the signal strength for distant IRs through careful phase shift optimization.

The number of data streams destined for each IR is assumed to be $d$, satisfying $1 \leq d \leq \min\{N_B, N_I\}$. The signal transmitted from the BS is given by

$$\mathbf{x} = \sum_{k=1}^{K_I} \mathbf{F}_k \mathbf{s}_k, \tag{1}$$

where $\mathbf{s}_k \in \mathbb{C}^{d\times 1}$ is the $(d \times 1)$-element data symbol vector designated for the $k$th IR satisfying $\mathbb{E}\left[\mathbf{s}_k \mathbf{s}_k^H\right] = \mathbf{I}_d$ and $\mathbb{E}\left[\mathbf{s}_i \mathbf{s}_j^H\right] = \mathbf{0}$, for $i \neq j$, while $\mathbf{F}_k \in \mathbb{C}^{N_B \times d}$ is the linear TPC matrix used by the BS for the $k$th IR. Assuming non-dispersive narrow-band transmission, the baseband equivalent channels spanning from the BS to the IRS, from the BS to the $k$th IR, from the BS to the $l$th ER, from the IRS to the $k$th IR, and finally from the IRS to the $l$th ER are modelled by the matrices $\mathbf{Z} \in \mathbb{C}^{M\times N_B}$, $\mathbf{H}_{b,k} \in \mathbb{C}^{N_I \times N_B}$, $\mathbf{G}_{b,l} \in \mathbb{C}^{N_E \times N_B}$, $\mathbf{H}_{r,k} \in \mathbb{C}^{N_I \times M}$, and $\mathbf{G}_{r,l} \in \mathbb{C}^{N_E \times M}$, respectively. Let us denote the diagonal reflection-coefficient matrix at the IRS by $\mathbf{\Phi} = \mathrm{diag}\left\{e^{j\theta_1}, \cdots, e^{j\theta_m}, \cdots, e^{j\theta_M}\right\}$ [1], where $\theta_m \in [0, 2\pi]$ is the phase shift of the $m$-th reflective element. Due to absorption and diffraction, the signal power that has been reflected multiple times is ignored. As a result, the signal received at the $k$th IR is given by

$$\mathbf{y}_{I,k} = \left(\mathbf{H}_{b,k} + \mathbf{H}_{r,k}\mathbf{\Phi}\mathbf{Z}\right)\mathbf{x} + \mathbf{n}_{I,k}, \tag{2}$$

where $\mathbf{n}_{I,k}$ is the $k$th IR's noise vector satisfying $\mathcal{CN}\left(\mathbf{0}, \sigma_I^2 \mathbf{I}_{N_I}\right)$. Similarly, the signal received at the $l$th ER is given by

$$\mathbf{y}_{E,l} = \left(\mathbf{G}_{b,l} + \mathbf{G}_{r,l}\mathbf{\Phi}\mathbf{Z}\right)\mathbf{x} + \mathbf{n}_{E,l}, \tag{3}$$

where $\mathbf{n}_{E,l}$ is the $l$th ER's noise vector obeying the distribution of $\mathcal{CN}\left(\mathbf{0}, \sigma_E^2 \mathbf{I}_{N_E}\right)$.

We assume that all the CSIs are perfectly known at the BS, and the BS is responsible for calculating the phase shifts of the IRS, which are then fed back by them to the IRS controller through dedicated feedback channels. Given this idealized and simplified assumption, the results obtained represent a performance upper bound of how much performance gain can be achieved by an IRS. Let us define the equivalent channel spanning from the BS to the $k$th IR by $\bar{\mathbf{H}}_k \triangleq \mathbf{H}_{b,k} + \mathbf{H}_{r,k}\mathbf{\Phi}\mathbf{Z}$. Upon substituting $\mathbf{x}$ into (2), $\mathbf{y}_{I,k}$ can be rewritten as

$$\mathbf{y}_{I,k} = \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{s}_k + \sum_{i=1, i\neq k}^{K_I} \bar{\mathbf{H}}_k \mathbf{F}_i \mathbf{s}_i + \mathbf{n}_{I,k}. \tag{4}$$

Then, the achievable data rate (nat/s/Hz) of the $k$th IR is given by [33]

$$R_k\left(\mathbf{F}, \mathbf{\Phi}\right) = \log\left|\mathbf{I} + \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{F}_k^{\mathrm{H}} \bar{\mathbf{H}}_k^{\mathrm{H}} \mathbf{J}_k^{-1}\right|, \tag{5}$$

where $\mathbf{F}$ denotes the collection of TPC matrices, while $\mathbf{J}_k$ is the interference-plus-noise covariance matrix given by $\mathbf{J}_k = \sum_{m=1, m\neq k}^{K_I} \bar{\mathbf{H}}_k \mathbf{F}_m \mathbf{F}_m^{\mathrm{H}} \bar{\mathbf{H}}_k^{\mathrm{H}} + \sigma_I^2 \mathbf{I}$.

On the other hand, due to the broadcast nature of wireless channels, the ERs can extract energy from the electromagnetic wave. In general, the harvested power is nonlinear over the received radio frequency (RF) power due to the nonlinear RF-to-DC conversion, which depends on the input RF power level. This nonlinear EH model has been characterized in [34], which is a complex function of the RF power. Based on this nonlinear EH model, various transmission designs have been proposed in [35] and [36]. However, there is still lack of a general model that can accurately characterize this nonlinear relationship by capturing all practical factors. Hence, for simplicity, we adopt the simple linear EH model as widely used in the existing literature [29], [37], [38]. By ignoring the noise power at the ERs, the total harvested power is proportional to the total received power. Let us define the equivalent channel spanning from the BS to the $l$th ER by $\bar{\mathbf{G}}_l \triangleq \mathbf{G}_{b,l} + \mathbf{G}_{r,l}\mathbf{\Phi}\mathbf{Z}$. Then, the total power harvested by the $l$th ER is

$$Q_i = \eta \mathrm{tr}\left(\sum_{k=1}^{K_I} \bar{\mathbf{G}}_l \mathbf{F}_k \mathbf{F}_k^{\mathrm{H}} \bar{\mathbf{G}}_l^{\mathrm{H}}\right), \tag{6}$$

where $0 < \eta \leq 1$ is the energy harvesting efficiency. In this paper, we consider the constraint that the weighted sum of the power harvested by all ERs should be higher than a predefined value, which is

$$Q = \sum_{l=1}^{K_E} \alpha_l Q_l = \mathrm{tr}\left(\sum_{k=1}^{K_I} \mathbf{F}_k^{\mathrm{H}} \mathbf{G} \mathbf{F}_k\right) \geq \bar{Q}, \tag{7}$$

where $\mathbf{G} = \sum_{l=1}^{K_E} \alpha_l \eta \bar{\mathbf{G}}_l^{\mathrm{H}} \bar{\mathbf{G}}_l$, $\alpha_l$ is the energy weighting factor of the $l$th ER, with a higher value of $\alpha_l$ representing a higher priority for the $l$th ER than for others. Finally, $\bar{Q}$ is the minimum harvested power threshold.

### B. Problem Formulation

Upon introducing the notations of $\phi_m = e^{j\theta_m}, \forall m$, we have $\mathbf{\Phi} = \mathrm{diag}\left\{\phi_1, \cdots, \phi_M\right\}$. Again, we aim for jointly

[1] $j$ is the imaginary unit.

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



optimizing the TPC matrices $\mathbf{F}$ and phase shift matrix $\mathbf{\Phi}$ with the goal of maximizing the WSR of all IRs subject to the total power budget, to the unit modulus of the phase shifters and to the harvested power requirement. Then, this problem can be formulated as follows:

$$\max_{\mathbf{F},\mathbf{\Phi}} \quad \sum_{k=1}^{K_I} \omega_k R_k\left(\mathbf{F},\mathbf{\Phi}\right) \tag{8a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K_I} \|\mathbf{F}_k\|_F^2 \leq P_T, \tag{8b}$$

$$\operatorname{tr}\left(\sum_{k=1}^{K_I} \mathbf{F}_k^{\mathrm{H}}\mathbf{G}\mathbf{F}_k\right) \geq \bar{Q}, \tag{8c}$$

$$|\phi_m| = 1, m = 1, \cdots, M, \tag{8d}$$

where $\omega_k$ is the weighting factor controlling the scheduling priority for each IR and $P_T$ is the power limit at the BS, while (8d) is the unit-norm constraint imposed on the phase shifters.

As the IRS is passive and both the ERs and IRs are energy constrained, we assume that this optimization problem is solved at the BS which posses the knowledge of the CSI of all related links and other related parameters such as $\bar{Q}$. After computing the phase shift values for the IRS, they are sent to the IRS controller through dedicated control channels. Problem (8) is difficult to solve, since the TPC matrices and the phase shifts are coupled. If we remove the energy harvesting (EH) constraint, the problem reduces to the WSR maximization problem recently studied in [24]. However, the additional EH constraint makes the optimization more challenging to solve and the algorithms developed in [24] cannot be directly applied for two reasons. Firstly, the EH constraint is non-convex. Secondly, this problem may be infeasible due to the conflicting constraints (8b) and (8c). In the following, we first conceive a low-complexity algorithm to solve this problem by assuming that it is feasible. Then, we study the feasibility of this problem.

## III. Low-Complexity Algorithm Development

In this section, we first transform Problem (8) into a more tractable one, which allows the decoupling of the TPC matrices and of the phase shifts. Then, the classic block coordinate descent (BCD) algorithm [33] is proposed for solving the transformed problem.

### A. Reformulation of the Original Problem

To deal with the complex objective function, we reformulate Problem (8) by employing the well-known WMMSE method [39]. The appealing feature of this method is that it can transform the original complex problem into an equivalent form, which facilitates the application of the BCD method.

Specifically, the linear decoding matrix $\mathbf{U}$ is applied to estimate the signal vector $\hat{\mathbf{s}}_k$ for each IR

$$\hat{\mathbf{s}}_k = \mathbf{U}_k^{\mathrm{H}}\mathbf{y}_{I,k}, \forall k, \tag{9}$$

where $\mathbf{U}_k \in \mathbb{C}^{N_I \times d}$ is the decoding matrix of the $k$th IR. Then, the MSE matrix of the $k$th IR is given by

$$\begin{aligned}\mathbf{E}_k =& \mathbb{E}_{\mathbf{s},\mathbf{n}}\left[\left(\hat{\mathbf{s}}_k - \mathbf{s}_k\right)\left(\hat{\mathbf{s}}_k - \mathbf{s}_k\right)^H\right] \\ =& \left(\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k - \mathbf{I}\right)\left(\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k - \mathbf{I}\right)^{\mathrm{H}} + \\ & \sum_{m=1, m\neq k}^{K_I} \mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_m\mathbf{F}_m^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k + \sigma^2\mathbf{U}_k^{\mathrm{H}}\mathbf{U}_k, \forall k \in \mathcal{K}_I,\end{aligned} \tag{10}$$

where $\mathbf{s}$ and $\mathbf{n}$ denote the collections of data symbols and noise vectors of all IRs, respectively.

By introducing a set of auxiliary matrices $\mathbf{W} = \{\mathbf{W}_k \succeq \mathbf{0}, \forall k \in \mathcal{K}_I\}$ and defining $\mathbf{U} = \{\mathbf{U}_k, \forall k \in \mathcal{K}_I\}$, Problem (8) can be reformulated as follows [33], [39]:

$$\max_{\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}} \quad \sum_{k=1}^{K_I} \omega_k h_k\left(\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}\right) \tag{11a}$$

$$\text{s.t.} \quad (8b), (8c), (8d), \tag{11b}$$

where $h_k\left(\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}\right)$ is given by

$$h_k\left(\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}\right) = \log|\mathbf{W}_k| - \operatorname{Tr}\left(\mathbf{W}_k\mathbf{E}_k\right) + d. \tag{12}$$

Although Problem (11) has more optimization variables than Problem (8), the objective function (OF) in Problem (11) is much easier to handle, which allows the BCD algorithm to solve this problem by iteratively obtaining one set of variables while keeping the others fixed. Note that the decoding matrices $\mathbf{U}$ and the auxiliary matrices $\mathbf{W}$ only appear in the function $h_k\left(\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}\right)$. Hence, the optimal solution of $\mathbf{U}$ and $\mathbf{W}$ can be obtained while keeping the other matrices fixed. Specifically, given $\mathbf{\Phi}$, $\mathbf{W}$, and $\mathbf{F}$, setting the first-order derivative of $h_k\left(\mathbf{W},\mathbf{U},\mathbf{F},\mathbf{\Phi}\right)$ with respect to $\mathbf{U}_k$ and $\mathbf{W}_k$ to zero, we can obtain the optimal solution of $\mathbf{U}_k$ and $\mathbf{W}_k$ respectively as follows

$$\mathbf{U}_k^\star = \left(\mathbf{J}_k + \bar{\mathbf{H}}_k\mathbf{F}_k\mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\right)^{-1}\bar{\mathbf{H}}_k\mathbf{F}_k, \mathbf{W}_k^\star = \mathbf{E}_k^{\star-1}, \tag{13}$$

where $\mathbf{E}_k^\star$ is obtained by inserting $\mathbf{U}_k^\star$ into the $k$th IR's MSE matrix in (10), yielding

$$\mathbf{E}_k^\star = \mathbf{I}_d - \mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\left(\sum_{m=1}^{K_I}\bar{\mathbf{H}}_k\mathbf{F}_m\mathbf{F}_m^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}} + \sigma_I^2\mathbf{I}\right)^{-1}\bar{\mathbf{H}}_k\mathbf{F}_k. \tag{14}$$

In the following, we focus our attention on the optimization of TPC matrices $\mathbf{F}$ and phase shifts $\mathbf{\Phi}$, when $\mathbf{U}$ and $\mathbf{W}$ are given.

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



### B. Optimizing the Precoding Matrices $\mathbf{F}$

In this subsection, we aim to optimize the TPC matrices $\mathbf{F}$ with fixed $\mathbf{W}, \mathbf{U}$ and $\mathbf{\Phi}$. By inserting $\mathbf{E}_k$ in (10) into the OF of (11) and discarding the constant terms, the TPC matrices of our optimization problem can be transformed as follows

$$\min_{\mathbf{F}} \quad \sum_{k=1}^{K_I} \mathrm{tr}\left(\mathbf{F}_k^{\mathrm{H}}\mathbf{A}\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{Tr}\left(\mathbf{W}_k\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\right) \tag{15a}$$

$$\text{s.t.} \quad (8b), (8c), \tag{15b}$$

where $\mathbf{A} = \sum_{m=1}^{K_I} \omega_m \bar{\mathbf{H}}_m^{\mathrm{H}}\mathbf{U}_m\mathbf{W}_m\mathbf{U}_m^{\mathrm{H}}\bar{\mathbf{H}}_m$.

However, due to the non-convexity of the EH constraint, Problem (15) is still non-convex. To resolve this issue, we observe that it can be viewed as a difference of convex (d.c.) program, which can be efficiently solved by the successive convex approximation (SCA) method [40]. In particular, we can approximate it by its first-order Taylor expansion. By applying [41, Appendix B] and Jensen' inequality, we have

$$\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}\mathbf{G}\mathbf{F}_k\right) \geq -\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{(n)}\right) + 2\mathrm{Re}\left[\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k\right)\right], \tag{16}$$

where $\left\{\mathbf{F}_k^{(n)}, \forall k\right\}$ is the solution obtained from the previous iteration. Then, upon replacing the constraint (8c) by the following constraint:

$$2\mathrm{Re}\left[\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k\right)\right] \geq \tilde{Q}, \tag{17}$$

where $\tilde{Q} = \bar{Q} + \mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{(n)}\right)$, we may consider the following optimization problem:

$$\min_{\mathbf{F}} \quad \sum_{k=1}^{K_I} \mathrm{tr}\left(\mathbf{F}_k^{\mathrm{H}}\mathbf{A}\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\right) \tag{18a}$$

$$\text{s.t.} \quad (8b), (17). \tag{18b}$$

Since the OF is convex w.r.t. $\mathbf{F}$, and the constraints (8b) and (17) are convex, Problem (18) constitutes a convex optimization problem, which can be solved by standard convex solver packages, such as CVX [42]. However, the resultant computational complexity is high. In the following, we provide a low-complexity algorithm for obtaining a nearly optimal closed-form solution by resorting to the Lagrangian dual decomposition method [43]. Since Problem (18) is a convex problem and satisfies the slater's condition[2], the dual gap is zero and the optimal solution can be obtained by solving its dual problem instead of its original one. We first introduce the Lagrange multiplier $\lambda$ associated with the power constraint, and derive the partial Lagrangian function of Problem (18) as follows

$$\mathcal{L}(\mathbf{F}, \lambda) = \sum_{k=1}^{K_I} \mathrm{tr}\left(\mathbf{F}_k^{\mathrm{H}}\mathbf{A}\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\right) + \lambda\sum_{k=1}^{K_I}\mathrm{tr}\left(\mathbf{F}_k^{\mathrm{H}}\mathbf{F}_k\right) - \lambda P_T. \tag{19}$$

The dual function can be obtained by solving the following problem

$$g(\lambda) \triangleq \min_{\mathbf{F}} \mathcal{L}(\mathbf{F}, \lambda) \quad \text{s.t. } (17). \tag{20}$$

Then, the dual problem is given by

$$\max_{\lambda} \quad g(\lambda) \tag{21a}$$

$$\text{s.t.} \quad \lambda \geq 0. \tag{21b}$$

Before solving the dual problem (21), we have to derive the expression of the dual function $g(\lambda)$ by solving Problem (20) for a given $\lambda$. By introducing the dual variable $\mu \geq 0$ associated with the constraint (17), the Lagrangian function for Problem (20) is given by

$$\mathcal{L}(\mathbf{F}, \mu) = \sum_{k=1}^{K_I} \mathrm{tr}\left(\mathbf{F}_k^{\mathrm{H}}\left(\mathbf{A} + \lambda\mathbf{I}\right)\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{U}_k^{\mathrm{H}}\bar{\mathbf{H}}_k\mathbf{F}_k\right) - \sum_{k=1}^{K_I} \omega_k \mathrm{tr}\left(\mathbf{W}_k\mathbf{F}_k^{\mathrm{H}}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\right) \tag{22}$$

By setting the first-order derivative of $\mathcal{L}(\mathbf{F}, \mu)$ w.r.t. $\mathbf{F}_k^*$ to the zero matrix, we obtain the optimal solution of $\mathbf{F}_k$ as follows:

$$\mathbf{F}_k^{\star}(\mu) = \left(\mathbf{A} + \lambda\mathbf{I}\right)^{\dagger}\left(\omega_k\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\mathbf{W}_k + \mu\mathbf{G}\mathbf{F}_k^{(n)}\right), \tag{23}$$

where $(\cdot)^{\dagger}$ denotes the matrix pseudoinverse. The value of $\mu$ should be chosen for ensuring that the complementary slackness condition for constraint (17) is satisfied:

$$\mu\left(2\mathrm{Re}\left[\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{\star}(\mu)\right)\right] - \tilde{Q}\right) = 0. \tag{24}$$

[2]According to line 1 in Algorithm 2, the initial precoding matrix is initialized by the solution obtained from Section IV. Assume the original problem is feasible. Due to the randomness of channel matrices of $\mathbf{G}$ and $\mathbf{H}$, the precoding matrix obtained in Section IV must be strictly larger than the minimum EH requirement, i.e., $\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(0)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{(0)}\right) > \bar{Q}$. Then, based on [29], there must exist a strictly feasible solution, and thus the slater's condition holds.

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



Hence, if the following condition holds

$$2\mathrm{Re}\left[\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{\star}(0)\right)\right]\geq\tilde{Q}, \tag{25}$$

the optimal solution of Problem (20) is given by $\mathbf{F}_k^{\star}(0),\forall k\in\mathcal{K}_I$. Otherwise, the optimal $\mu$ is

$$\mu=\frac{\tilde{Q}-2\mathrm{Re}\left[\mathrm{tr}\left(\sum_{k=1}^{K_I}\omega_k\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}(\mathbf{A}+\lambda\mathbf{I})^{-1}\bar{\mathbf{H}}_k^{\mathrm{H}}\mathbf{U}_k\mathbf{W}_k\right)\right]}{2\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}(\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{G}\mathbf{F}_k^{(n)}\right)}. \tag{26}$$

With the aid of the dual function, we may now commence the solution of the dual problem (21) to find the optimal $\lambda$. Given $\lambda$, we denote the optimal solution of Problem (20) by $\mathbf{F}_k(\lambda)$. The value of $\lambda$ should be chosen for ensuring that the complementary slackness condition for the power constraint is satisfied:

$$\lambda\left(\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}(\lambda)\mathbf{F}_k(\lambda)\right)-P_T\right)=0. \tag{27}$$

If the following condition holds:

$$\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}(0)\mathbf{F}_k(0)\right)\leq P_T, \tag{28}$$

then the optimal solution is given by $\mathbf{F}_k(0)$. Otherwise, we have to find $\lambda$ for ensuring that the following equation holds:

$$P(\lambda)\triangleq\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}(\lambda)\mathbf{F}_k(\lambda)\right)=P_T. \tag{29}$$

Unfortunately, due to the complex expression of $\mu$ in (26), we are unable to prove its monotonic nature by using the explicit expression of $P(\lambda)$ as in [24]. In the following lemma, we prove that $P(\lambda)$ is a monotonically decreasing function of $\lambda$, which enables the bisection search method to find $\lambda$.

***Lemma 1:*** The total power $P(\lambda)$ is a monotonically decreasing function of $\lambda$.

***Proof:*** Please refer to Appendix A. □

Based on Lemma 1, the bisection search method can be used for finding the solution of equation (29). In Algorithm 1, we provide the detailed steps of solving Problem (18) for the case of $\lambda>0$. In each iteration of Algorithm 1, we have to calculate $\mathbf{F}_k^{\star}(\mu)$ in (23), which involves the calculation of $(\mathbf{A}+\lambda\mathbf{I})^{-1}$ at a complexity order of $\mathcal{O}(N_B^3)$. If the total number of iterations is $T$, then the total complexity of calculating $(\mathbf{A}+\lambda\mathbf{I})^{-1}$ is $\mathcal{O}(TN_B^3)$, which may be excessive. Here, we provide a method for reducing the computational complexity. Specifically, as $\mathbf{A}$ is a non-negative definite matrix, it can be decomposed as $\mathbf{A}=\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}}$ by using the singular value decomposition (SVD), where $\mathbf{Q}\mathbf{Q}^{\mathrm{H}}=\mathbf{Q}^{\mathrm{H}}\mathbf{Q}=\mathbf{I}_{N_T}$ and $\mathbf{\Lambda}$ is a diagonal matrix with non-negative diagonal elements. Then, we have $(\mathbf{A}+\lambda\mathbf{I})^{-1}=\mathbf{Q}(\lambda\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{Q}^{\mathrm{H}}$. Hence, in each iteration, we only have to calculate the product of two matrices, which has much lower complexity than calculating the inverse of the matrix having the same dimension.

**Algorithm 1** Bisection Search Method to Solve Problem (18)

1: Initialize the accuracy $\varepsilon$, the bounds $\lambda_l$ and $\lambda_u$;
2: Calculate $\lambda=(\lambda_l+\lambda_u)/2$;
3: If condition (25) is satisfied, $\mu$ is equal to zero. Otherwise, update $\mu$ in (26);
4: Calculate $\{\mathbf{F}_k(\lambda),\forall k\}$ according to (23);
5: If $P(\lambda)\geq P_T$, set $\lambda_l=\lambda$. Otherwise, set $\lambda_u=\lambda$;
6: If $|\lambda_l-\lambda_u|\leq\varepsilon$, terminate. Otherwise, go to step 2.

Based on the above discussions, in Algorithm 2 we provide the detailed steps of the SCA algorithm conceived for solving Problem (15).

**Algorithm 2** SCA Algorithm to Solve Problem (15)

1: Initialize the accuracy $\varepsilon$, the precoding matrices $\mathbf{F}^{(0)}$ from Section 2, the iteration index $n=0$, the maximum number of iterations $n_{\max}$, calculate the OF value of Problem (15) as $z(\mathbf{F}^{(0)})$;
2: Calculate $\tilde{Q}^{(n)}=\bar{Q}+\mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{(n)\mathrm{H}}\mathbf{G}\mathbf{F}_k^{(n)}\right)$;
3: With $\tilde{Q}^{(n)}$, calculate $\{\mathbf{F}_k^{(n+1)},\forall k\}$ by solving Problem (18) using Algorithm 1;
4: If $n\geq n_{\max}$ or $\left|z(\mathbf{F}^{(n+1)})-z(\mathbf{F}^{(n)})\right|/\left|z(\mathbf{F}^{(n+1)})\right|<\varepsilon$, terminate. Otherwise, set $n\leftarrow n+1$ and go to step 2.

In the following, we show that Algorithm 2 converges to the KKT point of Problem (15).

***Theorem 1:*** The sequences of $\{\mathbf{F}^{(n)},n=1,2,\cdots\}$ generated by Algorithm 2 converge to the KKT optimum point of Problem (15).

***Proof:*** The proof is similar to that of [44] and hence it is omitted for simplicity. □

Next, we briefly analyze the complexity of Algorithm 2. We assume that $N_B\geq N_I\geq d$. In each iteration of Algorithm 2, the main complexity contribution is the calculation of $\{\mathbf{F}_k^{(n+1)},\forall k\}$ by using the bisection search method in Algorithm 1. In each iteration of Algorithm 1, the main complexity lies in calculating $\mathbf{F}$ in (23), which is on the order of $\mathcal{O}(K_IN_B^3)$. The number of iterations required for Algorithm 1 to converge is given by $\log_2\left(\frac{\lambda_u-\lambda_l}{\varepsilon}\right)$. Hence, the total complexity of Algorithm 1 is $\mathcal{O}(\log_2\left(\frac{\lambda_u-\lambda_l}{\varepsilon}\right)K_IN_B^3)$. Then, the total complexity of Algorithm 2 is given by $\mathcal{O}(n_{\max}\log_2\left(\frac{\lambda_u-\lambda_l}{\varepsilon}\right)K_IN_B^3)$.

### *C. Optimizing the Phase Shift Matrix* $\mathbf{\Phi}$

In this subsection, we focus our attention on optimizing the phase shift matrix $\mathbf{\Phi}$, while fixing the other

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



parameters. Upon substituting $\mathbf{E}_k$ in (10) into (12) and removing the terms that are independent of $\mathbf{\Phi}$, the phase shift optimization problem is formulated as:

$$\min_{\mathbf{\Phi}} \sum_{k=1}^{K_I} \mathrm{tr}\left(\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \tilde{\mathbf{F}} \bar{\mathbf{H}}_k^{\mathrm{H}} \mathbf{U}_k\right) - \sum_{k=1}^{K_I} \mathrm{tr}\left(\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \mathbf{F}_k\right) - \sum_{k=1}^{K_I} \mathrm{tr}\left(\omega_k \mathbf{W}_k \mathbf{F}_k^{\mathrm{H}} \bar{\mathbf{H}}_k^{\mathrm{H}} \mathbf{U}_k\right) \tag{30a}$$

$$\text{s.t.} \quad (8c), (8d), \tag{30b}$$

where $\tilde{\mathbf{F}} = \sum_{m=1}^{K_I} \mathbf{F}_m \mathbf{F}_m^{\mathrm{H}}$.

By substituting $\bar{\mathbf{H}}_k = \mathbf{H}_{b,k} + \mathbf{H}_{r,k}\mathbf{\Phi}\mathbf{Z}$ into (30a), we have

$$\begin{aligned} &\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \tilde{\mathbf{F}} \bar{\mathbf{H}}_k^{\mathrm{H}} \mathbf{U}_k \\ =&\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k} \mathbf{\Phi} \mathbf{Z} \tilde{\mathbf{F}} \mathbf{Z}^{\mathrm{H}} \mathbf{\Phi}^{\mathrm{H}} \mathbf{H}_{r,k}^{\mathrm{H}} \mathbf{U}_k + \\ &\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{b,k} \tilde{\mathbf{F}} \mathbf{Z}^{\mathrm{H}} \mathbf{\Phi}^{\mathrm{H}} \mathbf{H}_{r,k}^{\mathrm{H}} \mathbf{U}_k + \\ &\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k} \mathbf{\Phi} \mathbf{Z} \tilde{\mathbf{F}} \mathbf{H}_{b,k}^{\mathrm{H}} \mathbf{U}_k + \omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{b,k} \tilde{\mathbf{F}} \mathbf{H}_{b,k}^{\mathrm{H}} \mathbf{U}_k, \end{aligned} \tag{31}$$

and

$$\begin{aligned} &\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \mathbf{F}_k \\ =&\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k} \mathbf{\Phi} \mathbf{Z} \mathbf{F}_k + \omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{b,k} \mathbf{F}_k. \end{aligned} \tag{32}$$

Let us define $\mathbf{B}_k \triangleq \omega_k \mathbf{H}_{r,k}^{\mathrm{H}} \mathbf{U}_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k}$, $\mathbf{C} \triangleq \mathbf{Z}\tilde{\mathbf{F}}\mathbf{Z}^{\mathrm{H}}$ and $\mathbf{D}_k \triangleq \omega_k \mathbf{Z}\tilde{\mathbf{F}}^{\mathrm{H}} \mathbf{H}_{b,k}^{\mathrm{H}} \mathbf{U}_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k}$. By using (31), we arrive at:

$$\begin{aligned} &\mathrm{tr}\left(\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \tilde{\mathbf{F}} \bar{\mathbf{H}}_k^{\mathrm{H}} \mathbf{U}_k\right) \\ =&\mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{B}_k \mathbf{\Phi} \mathbf{C}\right) + \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{D}_k^{\mathrm{H}}\right) + \mathrm{tr}\left(\mathbf{\Phi} \mathbf{D}_k\right) + \mathrm{const}_1, \end{aligned} \tag{33}$$

where $\mathrm{const}_1$ is a constant term that is independent of $\mathbf{\Phi}$.

Similarly, by defining $\mathbf{T}_k \triangleq \omega_k \mathbf{Z} \mathbf{F}_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \mathbf{H}_{r,k}$, from (32) we have

$$\mathrm{tr}\left(\omega_k \mathbf{W}_k \mathbf{U}_k^{\mathrm{H}} \bar{\mathbf{H}}_k \mathbf{F}_k\right) = \mathrm{tr}\left(\mathbf{\Phi} \mathbf{T}_k\right) + \mathrm{const}_2, \tag{34}$$

where $\mathrm{const}_2$ is a constant term that is independent of $\mathbf{\Phi}$.

By defining $\mathbf{G}_b \triangleq \sum_{l=1}^{K_E} \alpha_l \eta \mathbf{G}_{b,l}^{\mathrm{H}} \mathbf{G}_{b,l}$, $\mathbf{G}_r \triangleq \sum_{l=1}^{K_E} \alpha_l \eta \mathbf{G}_{r,l}^{\mathrm{H}} \mathbf{G}_{r,l}$, and $\mathbf{G}_{br} \triangleq \mathbf{Z}\tilde{\mathbf{F}} \sum_{l=1}^{K_E} \alpha_l \eta \mathbf{G}_{b,l}^{\mathrm{H}} \mathbf{G}_{r,l}$, the EH constraint in (8c) can be recast as follows:

$$\begin{aligned} &\mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{G}_r \mathbf{\Phi} \mathbf{C}\right) + \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{G}_{b,r}^{\mathrm{H}}\right) + \mathrm{tr}\left(\mathbf{\Phi} \mathbf{G}_{br}\right) \\ &+ \mathrm{tr}\left(\mathbf{G}_b \tilde{\mathbf{F}}\right) \geq \bar{Q}. \end{aligned} \tag{35}$$

By inserting (33) and (34) into the OF of Problem (30) and removing the constant terms, we have

$$\min_{\mathbf{\Phi}} \quad \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{B} \mathbf{\Phi} \mathbf{C}\right) + \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{V}^{\mathrm{H}}\right) + \mathrm{tr}\left(\mathbf{\Phi} \mathbf{V}\right) \tag{36a}$$

$$\text{s.t.} \quad (8d), (35), \tag{36b}$$

where $\mathbf{B}$ and $\mathbf{V}$ are given by $\mathbf{B} = \sum_{k=1}^{K_I} \mathbf{B}_k$ and $\mathbf{V} = \sum_{k=1}^{K_I} \mathbf{D}_k - \sum_{k=1}^{K_I} \mathbf{T}_k$, respectively.

Upon denoting the collection of diagonal elements of $\mathbf{\Phi}$ by $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_M]^{\mathrm{T}}$ and adopting the matrix identity of [45, Eq. (1.10.6)], it follows that

$$\begin{aligned} \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{B} \mathbf{\Phi} \mathbf{C}\right) &= \boldsymbol{\phi}^{\mathrm{H}} \left(\mathbf{B} \odot \mathbf{C}^{\mathrm{T}}\right) \boldsymbol{\phi}, \\ \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{G}_r \mathbf{\Phi} \mathbf{C}\right) &= \boldsymbol{\phi}^{\mathrm{H}} \left(\mathbf{G}_r \odot \mathbf{C}^{\mathrm{T}}\right) \boldsymbol{\phi}. \end{aligned} \tag{37}$$

Upon denoting the collections of diagonal elements of $\mathbf{V}$ and $\mathbf{G}_{br}$ by $\mathbf{v} = \left[[\mathbf{V}]_{1,1}, \cdots, [\mathbf{V}]_{M,M}\right]^{\mathrm{T}}$ and $\mathbf{g} = \left[[\mathbf{G}_{br}]_{1,1}, \cdots, [\mathbf{G}_{br}]_{M,M}\right]^{\mathrm{T}}$, we arrive at

$$\begin{aligned} \mathrm{tr}\left(\mathbf{\Phi} \mathbf{V}\right) = \mathbf{v}^{\mathrm{T}} \boldsymbol{\phi}, \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{V}^{\mathrm{H}}\right) = \boldsymbol{\phi}^{\mathrm{H}} \mathbf{v}^*, \\ \mathrm{tr}\left(\mathbf{\Phi} \mathbf{G}_{br}\right) = \mathbf{g}^{\mathrm{T}} \boldsymbol{\phi}, \mathrm{tr}\left(\mathbf{\Phi}^{\mathrm{H}} \mathbf{G}_{br}^{\mathrm{H}}\right) = \boldsymbol{\phi}^{\mathrm{H}} \mathbf{g}^*. \end{aligned} \tag{38}$$

Moreover, the constraint (35) can be rewritten as

$$\boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi} + 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}} \mathbf{g}^*\right\} \geq \widehat{Q}, \tag{39}$$

where we have $\widehat{Q} = \bar{Q} - \mathrm{Tr}\left(\mathbf{G}_b \tilde{\mathbf{F}}\right)$ and $\mathbf{\Upsilon} = \mathbf{G}_r \odot \mathbf{C}^{\mathrm{T}}$. It can be verified that $\mathbf{G}_r$ and $\mathbf{C}^{\mathrm{T}}$ are non-negative semidefinite matrices. Then, according to [45], the Hadamard product $\mathbf{G}_r \odot \mathbf{C}^{\mathrm{T}}$ (or equivalently $\mathbf{\Upsilon}$) is also a semidefinite matrix.

Thus, Problem (36) can be transformed as

$$\min_{\boldsymbol{\phi}} \quad \boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Xi} \boldsymbol{\phi} + 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}} \mathbf{v}^*\right\} \tag{40a}$$

$$\text{s.t.} \quad (8d), (39), \tag{40b}$$

where we have $\mathbf{\Xi} = \mathbf{B} \odot \mathbf{C}^{\mathrm{T}}$. Again, $\mathbf{B}$ can be verified to be a non-negative semidefinite matrix, and thus $\mathbf{\Xi}$ is a non-negative semidefinite matrix.

Due to the non-convex constraint (39), Problem (40) is difficult to solve. To deal with this constraint, we again employ the SCA method [40]. Specifically, since $\boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi}$ is convex w.r.t. $\boldsymbol{\phi}$, its lower bound can be obtained as follows:

$$\boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi} \geq -\boldsymbol{\phi}^{(n)\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi}^{(n)} + 2\mathrm{Re}\left[\boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi}^{(n)}\right], \tag{41}$$

where $\boldsymbol{\phi}^{(n)}$ is obtained in the previous iteration. Then, constraint (39) is replaced by the following constraint

$$2\mathrm{Re}\left[\boldsymbol{\phi}^{\mathrm{H}} \left(\mathbf{g}^* + \mathbf{\Upsilon} \boldsymbol{\phi}^{(n)}\right)\right] \geq \widehat{Q} + \boldsymbol{\phi}^{(n)\mathrm{H}} \mathbf{\Upsilon} \boldsymbol{\phi}^{(n)} \triangleq \hat{Q}, \tag{42}$$

which is a linear constraint. Then, Problem (40) then becomes

$$\min_{\boldsymbol{\phi}} \quad \boldsymbol{\phi}^{\mathrm{H}} \mathbf{\Xi} \boldsymbol{\phi} + 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}} \mathbf{v}^*\right\} \tag{43a}$$

$$\text{s.t.} \quad (8d), (42). \tag{43b}$$

In the following, we conceive the Majorization-Minimization (MM) algorithm [31] for solving Problem (43). The key idea is to solve a challenging problem by introducing a series of more tractable subproblems. Upon

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



denoting the objective function of Problem (43) by $f(\boldsymbol{\phi})$, in the $(n+1)$th iteration we have to find the upper bound of the OF, denoted as $g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$, which should satisfy the following three conditions:

$$\begin{aligned}&1) g(\boldsymbol{\phi}^{(n)}|\boldsymbol{\phi}^{(n)}) = f(\boldsymbol{\phi}^{(n)});\\&2) \nabla_{\boldsymbol{\phi}^*} g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})\Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}^{(n)}} = \nabla_{\boldsymbol{\phi}^*} f(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^{(n)}};\\&3) g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) \geq f(\boldsymbol{\phi}).\end{aligned} \tag{44}$$

Then, we solve the approximate subproblem defined by a more tractable new OF $g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$. To find $g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$, we introduce the following lemma [46].

***Lemma 2:*** For any given $\boldsymbol{\phi}^{(n)}$, the following inequality holds for any feasible $\boldsymbol{\phi}$:

$$\begin{aligned}\boldsymbol{\phi}^{\mathrm{H}}\boldsymbol{\Xi}\boldsymbol{\phi} \leq& \boldsymbol{\phi}^{\mathrm{H}}\mathbf{X}\boldsymbol{\phi} - 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}}\left(\mathbf{X}\boldsymbol{\Xi}\right)\boldsymbol{\phi}^{(n)}\right\}\\&+\left(\boldsymbol{\phi}^{(n)}\right)^{\mathrm{H}}\left(\mathbf{X}-\boldsymbol{\Xi}\right)\boldsymbol{\phi}^{(n)} \triangleq y(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}),\end{aligned} \tag{45}$$

where $\mathbf{X} = \lambda_{\max}\mathbf{I}_M$ and $\lambda_{\max}$ is the maximum eigenvalue of $\boldsymbol{\Xi}$. □

Then, the function $g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ can be constructed as follows:

$$g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) = y(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) + 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}}\mathbf{v}^*\right\}, \tag{46}$$

where $y(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ is defined in (45). The new OF $g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ is more tractable than the original OF $f(\boldsymbol{\phi})$. The subproblem to be solved is given by

$$\min_{\boldsymbol{\phi}} \quad g(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) \tag{47a}$$

$$\text{s.t.} \quad (8d), (42). \tag{47b}$$

Since $\boldsymbol{\phi}^{\mathrm{H}}\boldsymbol{\phi} = M$, we have $\boldsymbol{\phi}^{\mathrm{H}}\mathbf{X}\boldsymbol{\phi} = M\lambda_{\max}$, which is a constant. By removing the other constants, Problem (47) can be rewritten as follows:

$$\max_{\boldsymbol{\phi}} \quad 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}}\mathbf{q}^{(n)}\right\} \tag{48a}$$

$$\text{s.t.} \quad (8d), (42), \tag{48b}$$

where $\mathbf{q}^{(n)} = (\lambda_{\max}\mathbf{I}_M - \boldsymbol{\Xi})\boldsymbol{\phi}^{(n)} - \mathbf{v}^*$. Due to the additional constraint (42), the optimal solution of Problem (48) cannot be obtained as in [24]. Furthermore, due to the non-convex unit-modulus constraint (8d), Problem (48) is a non-convex optimization problem. As a result, the Lagrangian dual decomposition method developed for the convex problem (18) is not applicable here, since the dual gap is not zero.

In the following, we propose a price mechanism for solving Problem (48) that can obtain the globally optimal solution. Specifically, we consider the following problem by introducing a non-negative price $p$ on the left hand side of constraint (42):

$$\max_{\boldsymbol{\phi}} \; 2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}}\mathbf{q}^{(n)}\right\} + 2p\mathrm{Re}\left[\boldsymbol{\phi}^{\mathrm{H}}\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^{(n)}\right)\right] \tag{49a}$$

$$\text{s.t.} \; (8d). \tag{49b}$$

For a given $p$, the globally optimal solution is given by

$$\boldsymbol{\phi}(p) = e^{j\arg\left(\mathbf{q}^{(n)} + p\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^{(n)}\right)\right)}. \tag{50}$$

Our objective is to find a $p$ value for ensuring that the complementary slackness condition for constraint (42) is satisfied:

$$p\left(J(p) - \hat{Q}\right) = 0, \tag{51}$$

where $J(p) = 2\mathrm{Re}\left[\boldsymbol{\phi}(p)^{\mathrm{H}}\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^{(n)}\right)\right]$. To solve this equation, we consider two cases: 1) $p = 0$; 2) $p > 0$.

***Case I:*** In this case, $\boldsymbol{\phi}(0) = e^{j\arg\left(\mathbf{q}^{(n)}\right)}$ has to satisfy constraint (42). Otherwise, $p > 0$.

***Case II:*** Since $p > 0$, equation (51) holds only when $J(p) = \hat{Q}$. To solve this equation, we first provide the following lemma.

***Lemma 3:*** Function $J(p)$ is a monotonically increasing function of $p$.

***Proof:*** The proof is similar to Lemma 1 and thus omitted. □

Based on Lemma 3, the bisection search method can be adopted for finding the solution of $J(p) = \hat{Q}$. Based on the above discussions, we provide the algorithm to solve Problem (48) in Algorithm 3. Although Problem (48) is a non-convex problem, in the following theorem we prove that Algorithm 3 is capable of finding the globally optimal solution.

***Theorem 2:*** Algorithm 3 is capable of finding the globally optimal solution of Problem (48) and thus also of Problem (47).

***Proof:*** Please refer to Appendix B. □

**Algorithm 3** Bisection Search Method to Solve Problem (48)

1: Calculate $J(0)$. If $J(0) \leq \hat{Q}$, terminate. Otherwise, go to step 2.
2: Initialize the accuracy $\varepsilon$, bounds $p_l$ and $p_u$;
3: Calculate $p = (p_l + p_u)/2$;
4: Update $\boldsymbol{\phi}(p)$ in (50) and calculate $J(p)$;
5: If $J(p) \geq \hat{Q}$, set $p_u = p$; Otherwise, set $p_l = p$;
6: If $|p_l - p_u| \leq \varepsilon$, terminate; Otherwise, go to step 3.

Based on the above, we now provide the details of solving Problem (30) in Algorithm 4.

In the following theorem, we prove that the sequence of $\{\boldsymbol{\phi}^{(n)}, n = 1, 2, \cdots\}$ generated by Algorithm 4 converges to the KKT-optimal point of Problem (30).

***Theorem 3:*** The sequences of the OF value produced by Algorithm 4 are guaranteed to converge, and the final solution satisfies the KKT point of Problem (30).

***Proof:*** Please refer to Appendix C. □

Let us now analyze the complexity of Algorithm 4. The complexity is dominated by calculating $\boldsymbol{\phi}^{(n+1)}$ in step 4 using Algorithm 3. The complexity mainly depends on calculating the maximum eigenvalue of $\boldsymbol{\Xi}$. Its

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



**Algorithm 4** MM Combined with SCA Algorithm to Solve Problem (30)

1: Initialize the accuracy $\varepsilon$, the phase shifts $\boldsymbol{\phi}^{(0)}$, the iteration index to $n=0$, the maximum number of iterations to $n_{\max}$, calculate the OF value of Problem (43) as $f(\boldsymbol{\phi}^{(0)})$;
2: Calculate $\hat{Q}^{(n)}=\widehat{Q}+\boldsymbol{\phi}^{(n)\mathrm{H}}\boldsymbol{\Upsilon}\boldsymbol{\phi}^{(n)}$;
3: Calculate $\mathbf{q}^{(n)}=(\lambda_{\max}\mathbf{I}_M-\boldsymbol{\Xi})\,\boldsymbol{\phi}^{(n)}-\mathbf{v}^*$;
4: Update $\boldsymbol{\phi}^{(n+1)}$ by solving Problem (48) using Algorithm 3;
5: If $n\geq n_{\max}$ or $\left|f(\boldsymbol{\phi}^{(n+1)})-f(\boldsymbol{\phi}^{(n)})\right|/f(\boldsymbol{\phi}^{(n+1)})\leq\varepsilon$ holds, terminate; Otherwise, set $n\leftarrow n+1$ and go to step 2.

complexity is on the order of $\mathcal{O}(M^3)$. The number of iterations required for Algorithm 3 is $\log_2\left(\frac{p_u-p_l}{\varepsilon}\right)$. Then, the total complexity of step 3 is $\mathcal{O}(\log_2\left(\frac{p_u-p_l}{\varepsilon}\right)M^3)$. Hence, the total complexity of Algorithm 4 is given by $\mathcal{O}(n_{\max}\log_2\left(\frac{p_u-p_l}{\varepsilon}\right)M^3)$.

### *D. Overall Algorithm to Solve Problem (8)*

Based on the above analysis, we provide the detailed steps of the BCD algorithm to solve Problem (8) in Algorithm 5, where $R(\mathbf{F}^{(n)},\boldsymbol{\phi}^{(n)})$ denotes the OF value of Problem (8) in the $n$th iteration.

**Algorithm 5** Block Coordinate Descent Algorithm

1: Initialize iterative number $n=1$, maximum number of iterations $n_{\max}$, feasible $\mathbf{F}^{(1)}$, $\boldsymbol{\phi}^{(1)}$, error tolerance $\varepsilon$, calculate $R(\mathbf{F}^{(1)},\boldsymbol{\phi}^{(1)})$, calculate the optimal decoding matrices $\mathbf{U}^{(1)}$ and auxiliary matrices $\mathbf{W}^{(1)}$ based on (13);
2: Given $\mathbf{U}^{(n)}$, $\mathbf{W}^{(n)}$ and $\boldsymbol{\phi}^{(n)}$, calculate the optimal precoding matrices $\mathbf{F}^{(n+1)}$ by solving Problem (15) using Algorithm 2;
3: Given $\mathbf{U}^{(n)}$, $\mathbf{W}^{(n)}$ and $\mathbf{F}^{(n+1)}$, calculate the optimal $\boldsymbol{\phi}^{(n+1)}$ by solving Problem (30) using Algorithm 4;
4: Given $\mathbf{F}^{(n+1)}$ and $\boldsymbol{\phi}^{(n+1)}$, calculate the optimal decoding matrices $\mathbf{U}^{(n+1)}$ in (13);
5: Given $\mathbf{F}^{(n+1)}$, $\mathbf{U}^{(n+1)}$ and $\boldsymbol{\phi}^{(n+1)}$, calculate the optimal auxiliary matrices $\mathbf{W}^{(n+1)}$ in (13);
6: If $n\geq n_{\max}$ or $\left|R(\mathbf{F}^{(n+1)},\boldsymbol{\phi}^{(n+1)})-R(\mathbf{F}^{(n)},\boldsymbol{\phi}^{(n)})\right|/R(\mathbf{F}^{(n+1)},\boldsymbol{\phi}^{(n+1)})<\varepsilon$, terminate. Otherwise, set $n\leftarrow n+1$ and go to step 2.

The following theorem shows the convergence and solution properties of Algorithm 5.

***Theorem 4:*** The OF value sequence $\{R(\mathbf{F}^{(n)},\boldsymbol{\phi}^{(n)}),n=1,2,\cdots\}$ generated by Algorithm 5 is guaranteed to converge, and the final solution satisfies the KKT conditions of Problem (8).

***Proof:*** Please refer to Appendix D. □

The complexity of Algorithm 5 mainly depends on that of Step 2 and Step 3, the complexity of which has been analyzed in the above subsections. In specific, the total complexity of step 2 and step 3 are respectively given by $\mathcal{O}(n_1^{\max}\log_2\left(\frac{\lambda_u-\lambda_l}{\varepsilon}\right)K_IN_B^3)$ and $\mathcal{O}(n_2^{\max}\log_2\left(\frac{p_u-p_l}{\varepsilon}\right)M^3)$, where $n_1^{\max}$ and $n_2^{\max}$ denote the number of iterations for Algorithm 2 and Algorithm 4 to converge. Denote the total number of iterations of Algorithm 5 as $N_{\max}$. Then, the overall complexity of Algorithm 5 is given by $O\big(N_{\max}\big(n_1^{\max}\log_2\big(\frac{\lambda_u-\lambda_l}{\varepsilon}\big)K_IN_B^3+n_2^{\max}\log_2\big(\frac{p_u-p_l}{\varepsilon}\big)M^3\big)\big)$. Additionally, the simulation results show that Algorithm 5 converges rapidly, which demonstrates the low complexity of this algorithm.

## IV. Feasibility Check for Problem (8)

Due to the conflicting EH and limited transmit power constraints, Problem (8) may be infeasible. Hence, we have to first check whether Problem (8) is feasible or not. To this end, we construct the following optimization problem:

$$\max_{\mathbf{F},\boldsymbol{\Phi}}\quad \mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}\mathbf{G}\mathbf{F}_k\right) \tag{52a}$$

$$\text{s.t.}\quad (8b),(8d). \tag{52b}$$

If the optimal OF value is larger than $\bar{Q}$, Problem (8) is feasible. Otherwise, it is infeasible. As the TPC matrices and phase shift matrix are coupled, the globally optimal solution is difficult to obtain. In the following, we can obtain a suboptimal solution by alternatively optimizing the TPC matrices and phase shifts.

For a given phase shift matrix, the TPC matrix optimization problem is given by

$$\max_{\mathbf{F}}\quad \mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}\mathbf{G}\mathbf{F}_k\right) \tag{53a}$$

$$\text{s.t.}\quad (8b). \tag{53b}$$

Upon denoting the maximum eigenvalue and the corresponding eigenvector of $\mathbf{G}$ by $\chi$ and $\mathbf{b}$ respectively, the optimal solution can be readily obtained as $\mathbf{F}_k=\left[\sqrt{p_k}\mathbf{b},\mathbf{0}_{N_B\times(d-1)}\right],\forall k=1,\cdots,K_I$, where $\sum_{k=1}^{K_I}p_k=P_T$ and $p_k\geq 0,\forall k=1,\cdots,K_I$. Without loss of generality, we can set $p_i=P_T/K_I,\forall i\in\mathcal{K}_I$. The OF value is given by $\chi P_T$. In this case, the optimal TPC matrix represents the optimal energy beamforming, which is the same as that for the single-antenna IR case of [38].

For a given TPC matrix $\mathbf{F}$, the phase shift optimization problem is formulated as:

$$\max_{\boldsymbol{\phi}}\quad \boldsymbol{\phi}^{\mathrm{H}}\boldsymbol{\Upsilon}\boldsymbol{\phi}+2\mathrm{Re}\left\{\boldsymbol{\phi}^{\mathrm{H}}\mathbf{g}^*\right\} \tag{54a}$$

$$\text{s.t.}\quad (8d), \tag{54b}$$

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



where $\boldsymbol{\Upsilon}$ and $\mathbf{g}$ are defined in the above section. The OF is convex w.r.t. $\boldsymbol{\phi}$, and maximizing a convex function is a d.c program. Hence, it can be solved by using the SCA method by approximating $\boldsymbol{\phi}^{\mathrm{H}}\boldsymbol{\Upsilon}\boldsymbol{\phi}$ as its first-order Taylor expansion, details of which are omitted.

Finally, alternatively solve Problem (53) and (54) until the OF is larger than $\bar{Q}$.

## V. Simulation Results

In this section, we provide simulation results for demonstrating the benefits of applying IRS to SWIPT systems, as seen in Fig. 2, where there are four ERs and two IRs. The ERs and IRs are uniformly and randomly scattered in a circle centered at $(x_{\mathrm{ER}}, 0)$ and $(x_{\mathrm{IR}}, 0)$ with radius 1 m and 4 m, respectively. The IRS is located at $(x_{\mathrm{IRS}}, 2)$. In the simulations, we assume that the IRS is just above the ERs and thus we set $x_{\mathrm{ER}} = x_{\mathrm{IRS}}$. The large-scale path loss is modeled in dB as

$$PL = PL_0 \left(\frac{D}{D_0}\right)^{-\alpha}, \tag{55}$$

where $\mathrm{PL}_0$ is the path loss at the reference distance $D_0$, $D$ is the link length in meters, and $\alpha$ is the path loss exponent. Here, we set $D_0 = 1$ and $\mathrm{PL}_0 = -30\mathrm{dB}$. The path loss exponents of the BS-IRS, IRS-ER, IRS-IR, BS-IR and BS-ER links are respectively set as $\alpha_{\mathrm{BSIRS}} = 2.2$, $\alpha_{\mathrm{IRSER}} = 2.2$, $\alpha_{\mathrm{IRSIR}} = 2.4$, $\alpha_{\mathrm{BSIR}} = 3.6$ and $\alpha_{\mathrm{BSER}} = 3.6$. Unless otherwise stated, the other parameters are set as follows: Channel bandwidth of 1 MHz, noise power density of $-160$ dBm/Hz, $N_B = 4$, $N_I = N_E = 2$, $d = 2$, $\bar{Q} = 2 \times 10^{-4}$ W, $\eta = 0.5$, $M = 50$, $P_T = 10\ W$, weight factors $\omega_k = 1, \forall k \in \mathcal{K}_I, \alpha_l = 1, \forall l \in \mathcal{K}_E$, $x_{\mathrm{ER}} = 5$ m, and $x_{\mathrm{IR}} = 400$ m. The following results are obtained by averaging over 100 random locations and channel generations. Due to the severe blockage and long distance, the channels from the BS and the IRS to the IRs are assumed to be Rayleigh fading. However, as the BS, the ERs and the IRS are close to each other, the small-scale channels are assumed to be Rician fading. In particular, the small-scale channels from the IRS to the ERs are denoted as:

$$\tilde{\mathbf{G}}_{r,l} = \sqrt{\frac{\beta_{\mathrm{irser}}}{\beta_{\mathrm{irser}}+1}}\tilde{\mathbf{G}}_{r,l}^{\mathrm{LoS}} + \sqrt{\frac{1}{\beta_{\mathrm{irser}}+1}}\tilde{\mathbf{G}}_{r,l}^{\mathrm{NLoS}}, \quad l = 1, \cdots, K_E, \tag{56}$$

where $\beta_{\mathrm{irser}}$ is the Rician factor, $\tilde{\mathbf{G}}_{r,l}^{\mathrm{LoS}}$ is the deterministic line of sight (LoS), and $\tilde{\mathbf{G}}_{r,l}^{\mathrm{NLoS}}$ is the non-LoS (NLoS) component that is Rayleigh fading. The LoS component $\tilde{\mathbf{G}}_{r,l}^{\mathrm{LoS}}$ can be modeled as $\tilde{\mathbf{G}}_{r,l}^{\mathrm{LoS}} = \mathbf{a}_{N_E}\left(\vartheta_{\mathrm{irser},l}^{AoA}\right)\mathbf{a}_M^H\left(\vartheta_{\mathrm{irser},l}^{AoD}\right)$, where $\mathbf{a}_{N_E}\left(\vartheta_{\mathrm{irser},l}^{AoA}\right)$ is defined as

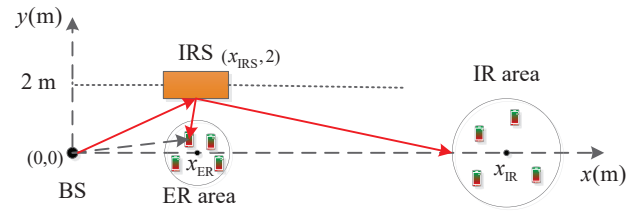


Fig. 2. The simulated IRS-aided SWIPT MIMO communication scenario.

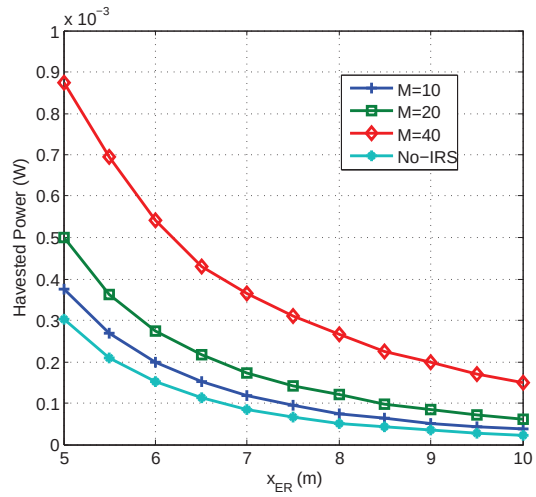


Fig. 3. Maximum harvested power achieved by various schemes.

$$\mathbf{a}_{N_E}\left(\vartheta_{\mathrm{irser},l}^{AoA}\right) = \left[1, e^{j\frac{2\pi d}{\lambda}\sin\vartheta_{\mathrm{irser},l}^{AoA}}, \cdots, e^{j\frac{2\pi d}{\lambda}(N_E-1)\sin\vartheta_{\mathrm{irser},l}^{AoA}}\right]^T \tag{57}$$

and

$$\mathbf{a}_M\left(\vartheta_{\mathrm{irser},l}^{AoD}\right) = \left[1, e^{j\frac{2\pi d}{\lambda}\sin\vartheta_{\mathrm{irser},l}^{AoD}}, \cdots, e^{j\frac{2\pi d}{\lambda}(M-1)\sin\vartheta_{\mathrm{irser},l}^{AoD}}\right]^T. \tag{58}$$

In (57) and (58), $d$ is the antenna separation distance, $\lambda$ is the wavelength, $\vartheta_{\mathrm{irser},l}^{AoD}$ is the angle of departure and $\vartheta_{\mathrm{irser},l}^{AoA}$ is the angle of arrival. It is assumed that $\vartheta_{\mathrm{irser},l}^{AoD}$ and $\vartheta_{\mathrm{irser},l}^{AoA}$ are randomly distributed within $[0, 2\pi]$. For simplicity, we set $d/\lambda = 1/2$. The small-scale channels from the BS to the ERs and the IRS are similarly defined. For simplicity, the Rician factors for all Rician fading channels are assumed to be the same as $\beta = 3$.

We first compare the maximum power harvested by various schemes in Fig. 3. Specifically, we solve the EH maximization problem (52) by using the feasibility check method in Section IV. Additionally, we also present the results without using IRS. Fig. 3 shows the maximum EH power versus the ER circle center location $x_{\mathrm{EH}}$. As expected, the EH power gleaned by all schemes decreases, when the ERs are far away from the BS. As expected,

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



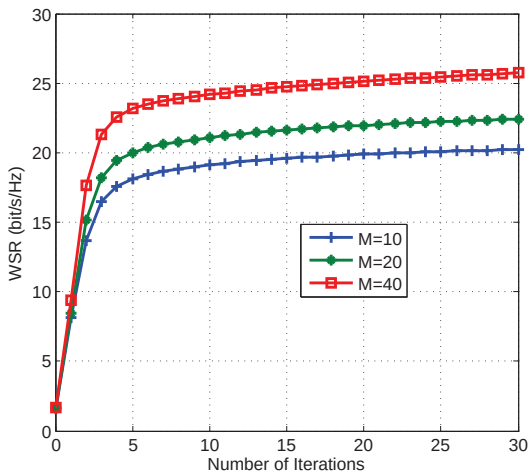


Fig. 4. Convergence behaviour of the BCD algorithm.

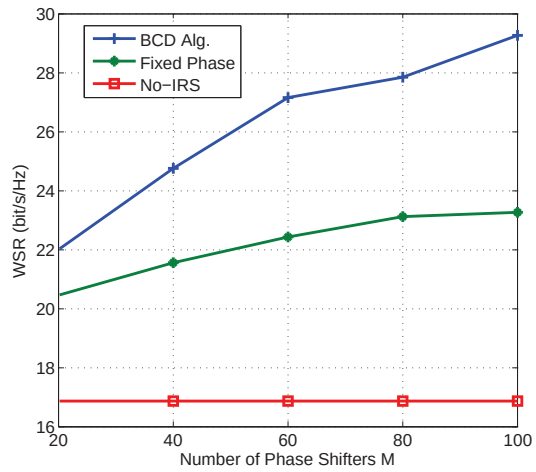


Fig. 5. WSR versus the number of phase shifters.

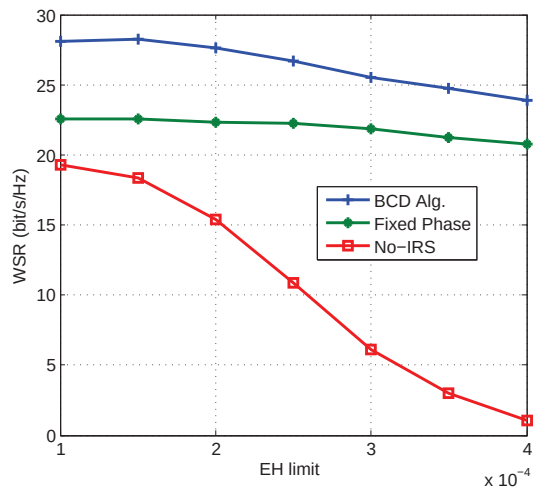


Fig. 6. WSR versus the harvested power requirement $\bar{Q}$.

more power can be harvested with the aid of IRS than that without IRS, especially when the number of phase shifters $M$ is large. This is mainly due to the fact that an additional strong link is reflected by the IRS, which can be harvested by the ERs. This figure also shows that the IRS is effective in expanding the operational range of ERs. For example, when the harvested power limit is $\bar{Q} = 2 \times 10^{-4}$ W, the maximum operational range of the system without IRS is only 5.5 m, while the system having $M = 40$ phase shifters can operate for distances up to 9 m.

In Fig. 4, we study the convergence behaviour of the BCD algorithm for different numbers of phase shifters $M$. It is observed from Fig. 4 that the WSR achieved for various $M$ values increases monotonically with the number of iterations, which verifies Theorem 4. Additionally, the BCD algorithm converges rapidly and in general a few iterations are sufficient for the BCD algorithm to achieve a large portion of the converged WSR. This reflects the low complexity of the BCD algorithm, which is appealing for practical applications.

In the following, we compare our proposed BCD algorithms to a pair of benchmark schemes: 1)'No-IRS': In this scheme, there is no IRS to assist the transmission as in conventional systems; 2) 'Fixed Phase': In this method, the phase shifts are fixed at the solutions obtained by solving the harvested power maximization problem (52), while they are not optimized, when using the BCD algorithm by removing Step 3 of the BCD algorithm. When any of the methods fails to obtain a feasible solution, its achievable WSR is set to zero.

In Fig. 5, we first study the impact of the number of phase shifters $M$ on the performance of various algorithms. As expected, the WSR achieved by all the algorithms - except for the No-IRS method - increases with $M$, since a higher degree of freedom can be exploited for optimizing the system performance. By carefully optimizing the phase shifts for maximizing the WSR, the BCD algorithm significantly outperforms the fixed-phase scheme. Additionally, the performance gain increases with $M$, which emphasizes the importance of optimizing the phase shifts. By employing the IRS in our SWIPT system, the WSR obtained by the BCD algorithm becomes drastically higher than that of No-IRS. For example, when $M = 60$, the WSR performance gain is up to 10 bit/s/Hz. These results demonstrate that introducing the IRS into our SWIPT system is a promising technique of enhancing the system performance.

In Fig. 6, the impact of harvested power requirement $\bar{Q}$ is investigated. It is seen from this figure that the WSR achieved by all the algorithms decreases upon increasing $\bar{Q}$, because the probability of infeasibility increases, which in turn reduces the average WSR value. We also find that the WSR obtained by the No-IRS scheme decreases more

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



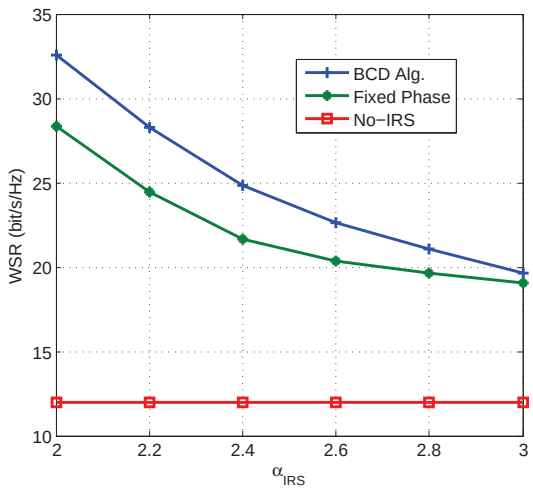


Fig. 7. WSR versus the IRS-related path loss exponent $\alpha_{\mathrm{IRS}}$.

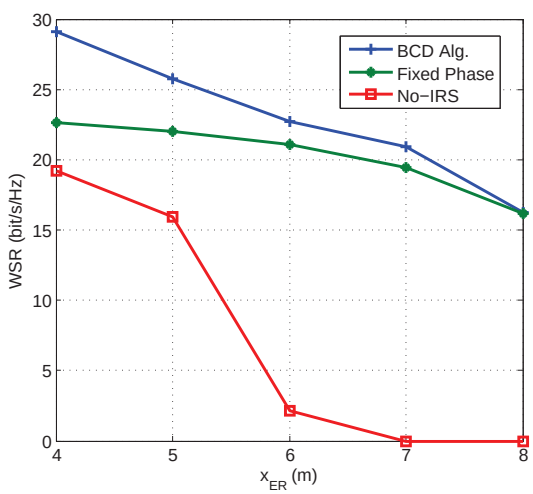


Fig. 8. WSR versus the location of ER circle center $x_{\mathrm{ER}}$.

rapidly than that of the other two IRS-aided transmission schemes. The WSR of the No-IRS is approaching zero when $\bar{Q} = 4 \times 10^{-4}$ W, while those relying on IRSs achieve a WSR gain in excess of 20 bit/s/Hz. It is observed again that the BCD algorithm performs better than the fixed-phase scheme, but the gap narrows with the increase of $\bar{Q}$. This can be explained as follows. With the increase of $\bar{Q}$, both the TPC matrices and the phase shifts should be designed for maximizing the power harvested at the ERs, and thus the final solutions of the fixed-phase and BCD method will become the same.

The above results are obtained for $\alpha_{\mathrm{BSIRS}} = 2.2$, $\alpha_{\mathrm{IRSER}} = 2.2$, $\alpha_{\mathrm{IRSIR}} = 2.4$ based on the assumption that the IRS relies on an obstacle-free scenario. In practice, this ideal scenario is seldom encountered. Hence, it is imperative to investigate the impact of $\alpha_{\mathrm{IRS}} \triangleq \alpha_{\mathrm{BSIRS}} = \alpha_{\mathrm{IRSER}} = \alpha_{\mathrm{IRSIR}}$ on the system performance, which is shown in Fig. 7. Observe from this figure that the WSR

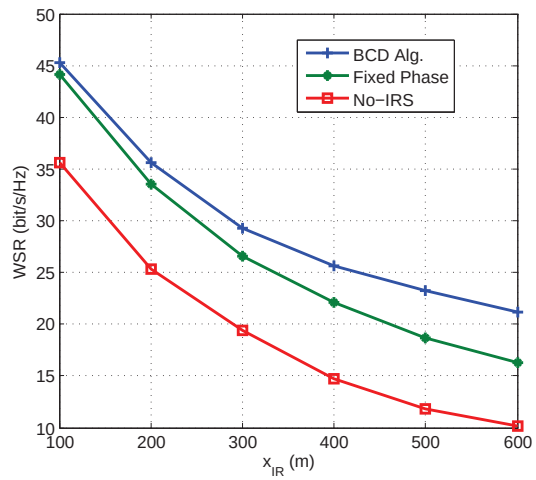


Fig. 9. WSR versus the location of IR circle center $x_{\mathrm{IR}}$.

achieved by the algorithms using IRS decreases drastically with $\alpha_{\mathrm{IRS}}$. When $\alpha_{\mathrm{IRS}} = 3$, the WSR-performance gain of our algorithm over the No-IRS scenario is only 7 bit/s/Hz, because upon increasing $\alpha_{\mathrm{IRS}}$, the signal power reflected from the IRS becomes weaker. Hence, the benefits of the IRS can be eroded. This provides an important engineering design insight: the location of IRS should be carefully considered for finding an obstacle-free scenario associated with a low $\alpha_{\mathrm{IRS}}$.

In Fig. 8, we study the impact of ER locations on the system performance. As expected, the WSR achieved by all the schemes decreases with $x_{\mathrm{IRS}}$, since the ERs become more distant from the BS and the signals gleaned from both the BS and IRS become weaker. The WSR achieved by the No-IRS approaches zero when $x_{\mathrm{IRS}} = 8$ m, hence this method cannot reach the energy transmission target of the ERs. The proposed algorithm is again observed to significantly outperform the other two algorithms, especially when the ERs are close to the BS.

Finally, the impact of IR locations is investigated in Fig. 9. It is observed that the WSR achieved by all the algorithms decreases with $x_{\mathrm{IR}}$ since the IRs become farther away from the BS when increasing $x_{\mathrm{IR}}$. The proposed algorithm is shown to achieve nearly the WSR gain of 10 bit/s/Hz over the No-IRS when $x_{\mathrm{IR}} = 100\ m$, and the WSR gain slightly increases with $x_{\mathrm{IR}}$. This means that the IRS is more advantageous when the IRs are far away from the BS, and the IRS can provide one additional favorable link.

## VI. Conclusions

In this paper, we have invoked an IRS in a SWIPT MIMO system for enhancing the performance of both the ERs and IRs. By carefully adjusting the phase shifts at the IRS, the signal reflected by the IRS can be added constructively at both the ERs and IRs. We considered

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



the WSR maximization problem of IRs, while guaranteeing the energy harvesting requirements of the ERs and the associated non-convex unit-modulus constraints. We conceived a BCD algorithm for alternatively optimizing the TPC matrices at the BS and the phase shift matrix at the IRSs. For each subproblem, a low-complexity iterative algorithm was proposed, which guarantees to be at worst locally optimal. Our simulation results demonstrated that the IRS enhances the performance of the SWIPT system and that the proposed algorithm converges rapidly, hence it is eminently suitable for practical implementations.

This paper assumes perfect CSI at the BS, which is challenging to obtain. For the future work, we will consider the robust transmission design for the IRS-aided SWIPT system, where the CSI is assumed to be imperfectly known. In addition, how to design the discrete phase shifts will be left for future work.

## Appendix A
## Proof of Lemma 1

We consider a pair of variables $\lambda$ and $\lambda'$, where $\lambda > \lambda'$. Let $\mathbf{F}(\lambda)$ and $\mathbf{F}(\lambda')$ be the optimal solutions of Problem (20) with $\lambda$ and $\lambda'$, respectively. Since $\mathbf{F}(\lambda)$ is the optimal solution of Problem (20) with $\lambda$, we have

$$\mathcal{L}[\mathbf{F}(\lambda), \lambda] \leq \mathcal{L}[\mathbf{F}(\lambda'), \lambda]. \tag{A.1}$$

Similarly, we have

$$\mathcal{L}[\mathbf{F}(\lambda'), \lambda'] \leq \mathcal{L}[\mathbf{F}(\lambda), \lambda']. \tag{A.2}$$

By adding these two inequalities and simplifying them, we have $(\lambda - \lambda')\, P(\lambda) \leq (\lambda - \lambda')\, P(\lambda')$. Since $\lambda > \lambda'$, we have $P(\lambda) \leq P(\lambda')$, which completes the proof.

## Appendix B
## Proof of Theorem 2

Denote the globally optimal solution of Problem (48) by $\boldsymbol{\phi}^\star$. According to [43], for a non-convex optimization problem, all its locally optimal solutions (including the globally optimal solution) should satisfy the Karush-Kuhn-Tucker (KKT) optimality conditions, one of which is the complementary slackness condition for constraint (42):

$$\lambda^\star \left( 2\mathrm{Re}\left[ \boldsymbol{\phi}^{\star \mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right] - \hat{Q} \right) = 0, \tag{B.1}$$

where $\lambda^\star$ is the corresponding optimal dual variable. We consider two cases: 1) $\lambda^\star = 0$; 2) $\lambda^\star > 0$.

The first case means that constraint (42) is not tight in the optimum. Then, the optimal solution can be obtained as $\boldsymbol{\phi}^\star = e^{j \arg\left(\mathbf{q}^{(n)}\right)}$, which is equal to $\boldsymbol{\phi}(0)$. Hence, Algorithm 3 achieves the optimal solution of Problem (48).

For the second case, the following equality should hold:

$$2\mathrm{Re}\left[ \boldsymbol{\phi}^{\star \mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right] = \hat{Q}. \tag{B.2}$$

We prove the second case by using the method of contradiction. Denote the optimal $p$ obtained by Algorithm 3 as $p^\star$, and the corresponding $\boldsymbol{\phi}$ as $\boldsymbol{\phi}(p^\star)$. Then, we have

$$2\mathrm{Re}\left[ \boldsymbol{\phi}(p^\star)^{\mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right] = \hat{Q}. \tag{B.3}$$

Let us assume that $\boldsymbol{\phi}(p^\star)$ is not the globally optimal solution of Problem (48). Then, we have

$$2\mathrm{Re}\left\{ \boldsymbol{\phi}(p^\star)^{\mathrm{H}} \mathbf{q}^{(n)} \right\} < 2\mathrm{Re}\left\{ \boldsymbol{\phi}^{\star \mathrm{H}} \mathbf{q}^{(n)} \right\}. \tag{B.4}$$

Since $\boldsymbol{\phi}(p^\star)$ is the globally optimal solution of Problem (49) when $p = p^\star$, we have

$$\begin{aligned} &2\mathrm{Re}\left\{ \boldsymbol{\phi}(p^\star)^{\mathrm{H}} \mathbf{q}^{(n)} \right\} + 2p^\star \mathrm{Re}\left[ \boldsymbol{\phi}(p^\star)^{\mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right] \\ \geq &2\mathrm{Re}\left\{ \boldsymbol{\phi}^{\star \mathrm{H}} \mathbf{q}^{(n)} \right\} + 2p^\star \mathrm{Re}\left[ \boldsymbol{\phi}^{\star \mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right]. \end{aligned} \tag{B.5}$$

By substituting (B.2) and (B.3) into (B.5), we have

$$2\mathrm{Re}\left\{ \boldsymbol{\phi}(p^\star)^{\mathrm{H}} \mathbf{q}^{(n)} \right\} \geq 2\mathrm{Re}\left\{ \boldsymbol{\phi}^{\star \mathrm{H}} \mathbf{q}^{(n)} \right\}, \tag{B.6}$$

which contradicts (B.4). Hence, the solution obtained by Algorithm 3 is the globally optimal solution of Problem (48). Since Problem (47) is equivalent to Problem (48), the proof is complete.

## Appendix C
## Proof of Theorem 3

Let us define the following functions:

$$T(\boldsymbol{\phi}) \triangleq \boldsymbol{\phi}^{\mathrm{H}} \boldsymbol{\Upsilon} \boldsymbol{\phi} + 2\mathrm{Re}\left\{ \boldsymbol{\phi}^{\mathrm{H}} \mathbf{g}^* \right\} + \mathrm{tr}\left( \mathbf{G}_b \tilde{\mathbf{F}} \right), \tag{C.1}$$

$$\begin{aligned} \bar{T}(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) \triangleq &-\boldsymbol{\phi}^{(n)\mathrm{H}} \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} + 2\mathrm{Re}\left[ \boldsymbol{\phi}^{\mathrm{H}} \left( \mathbf{g}^* + \boldsymbol{\Upsilon} \boldsymbol{\phi}^{(n)} \right) \right] \\ &+ \mathrm{tr}\left( \mathbf{G}_b \tilde{\mathbf{F}} \right). \end{aligned} \tag{C.2}$$

It can be verified that $T(\boldsymbol{\phi}^{(n)}) = \bar{T}(\boldsymbol{\phi}^{(n)}|\boldsymbol{\phi}^{(n)})$.

We first show that the solution sequence $\{\boldsymbol{\phi}^{(n)}, n = 1, 2, \cdots\}$ is feasible for Problem (30). The unit-modulus constraint is guaranteed in (50). We only have to check the EH constraint in (8c). Note that $\boldsymbol{\phi}^{(n+1)}$ is a feasible solution of Problem (48), and thus satisfies constraint (41). Hence, we have $\bar{T}(\boldsymbol{\phi}^{(n+1)}|\boldsymbol{\phi}^{(n)}) \geq \bar{Q}$. By using inequality (41), we have $T(\boldsymbol{\phi}^{(n+1)}) \geq \bar{T}(\boldsymbol{\phi}^{(n+1)}|\boldsymbol{\phi}^{(n)})$. Then, $T(\boldsymbol{\phi}^{(n+1)}) \geq \bar{Q}$ holds, which means that the sequence of $\boldsymbol{\phi}^{(n+1)}$ satisfies the EH constraint in (8c).

Now, we show that the OF value sequence $\{f(\boldsymbol{\phi}^{(n)}), n = 1, 2, \cdots\}$ is monotonically decreasing. Based on Theorem 2, the globally optimal solution $\boldsymbol{\Phi}$ to Problem (47) can be obtained. Then, we have $g(\boldsymbol{\phi}^{(n+1)}|\boldsymbol{\phi}^{(n)}) \leq g(\boldsymbol{\phi}^{(n)}|\boldsymbol{\phi}^{(n)})$. According to the first condition in (44), we have $g(\boldsymbol{\phi}^{(n)}|\boldsymbol{\phi}^{(n)}) = f(\boldsymbol{\phi}^{(n)})$. Hence, we have $g(\boldsymbol{\phi}^{(n+1)}|\boldsymbol{\phi}^{(n)}) \leq f(\boldsymbol{\phi}^{(n)})$. By using the third condition of (44), we have $g(\boldsymbol{\phi}^{(n+1)}|\boldsymbol{\phi}^{(n)}) \geq f(\boldsymbol{\phi}^{(n+1)})$. As a result, we have $f(\boldsymbol{\phi}^{(n)}) \geq f(\boldsymbol{\phi}^{(n+1)})$. Additionally, the OF must have a lower bound due to the unit-modulus constraint. Hence,

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



the OF value sequence $\{f(\boldsymbol{\phi}^{(n)}), n = 1, 2, \cdots\}$ is guaranteed to converge.

Now, we prove that the converged solution satisfies the KKT conditions of Problem (30). Let us denote the converged solution by $\{\boldsymbol{\phi}^\star\}$. Since $\boldsymbol{\phi}^\star$ is the globally optimal solution of Problem (47), it must satisfy the KKT conditions of Problem (47). Specifically, the Lagrange function of Problem (47) is given by

$$\mathcal{L}(\boldsymbol{\phi}, \nu, \boldsymbol{\tau}) = g(\boldsymbol{\phi}|\boldsymbol{\phi}^\star) + \nu\left(\hat{Q} - 2\mathrm{Re}\left[\boldsymbol{\phi}^{\mathrm{H}}\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^\star\right)\right]\right) + \sum_{m=1}^{M} \tau_m\left(|\phi_m| - 1\right), \tag{C.3}$$

where $\nu$ and $\boldsymbol{\tau} = \{\tau_1, \cdots, \tau_M\}$ are the corresponding dual variables. Then, there must exist a $\nu^\star$ and $\boldsymbol{\tau}^\star = \{\tau_1^\star, \cdots, \tau_M^\star\}$ for ensuring that the following conditions are satisfied:

$$\nabla_{\boldsymbol{\phi}^*}\mathcal{L}(\boldsymbol{\phi}, \nu, \boldsymbol{\tau})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*} g(\boldsymbol{\phi}|\boldsymbol{\phi}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} - \nu^\star\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^\star\right) + \sum_{m=1}^{M}\tau_m^\star(\nabla_{\boldsymbol{\phi}^*}|\phi_m|)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \mathbf{0}, \tag{C.4}$$

$$\nu^\star\left(\hat{Q} - 2\mathrm{Re}\left[\boldsymbol{\phi}^{\star\mathrm{H}}\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^\star\right)\right]\right) = 0, \tag{C.5}$$

$$\tau_m^\star\left(|\phi_m^\star| - 1\right) = 0, \forall m. \tag{C.6}$$

According to the second condition of (44), we have

$$\nabla_{\boldsymbol{\phi}^*} g(\boldsymbol{\phi}|\boldsymbol{\phi}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*} f(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}. \tag{C.7}$$

Upon denoting the OF of Problem (30) as $\varphi(\boldsymbol{\phi})$, which is the same as $f(\boldsymbol{\phi})$ except that $\varphi(\boldsymbol{\phi})$ has more constants, we have $\nabla_{\boldsymbol{\phi}^*} f(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*}\varphi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}$. Combining with (C.7), we have $\nabla_{\boldsymbol{\phi}^*} g(\boldsymbol{\phi}|\boldsymbol{\phi}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*}\varphi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}$. By substituting it into (C.4), we arrive at

$$\nabla_{\boldsymbol{\phi}^*}\varphi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} - \nu^\star\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^\star\right) + \sum_{m=1}^{M}\tau_m^\star\left(\nabla_{\boldsymbol{\phi}^*}|\phi_m|\right)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \mathbf{0}. \tag{C.8}$$

It can be checked that the set of equations (C.5), (C.6) and (C.8) constitutes exactly the KKT conditions of Problem (30). Hence, the proof is complete.

## Appendix D
## Proof of Theorem 4

Let us define the OF of Problem (11) as

$$h\left(\mathbf{W}, \mathbf{U}, \mathbf{F}, \boldsymbol{\Phi}\right) \triangleq \sum_{k=1}^{K_I}\omega_k h_k\left(\mathbf{W}, \mathbf{U}, \mathbf{F}, \boldsymbol{\Phi}\right). \tag{D.1}$$

It can be readily verified that the sequence of solutions $\{\mathbf{F}^{(n)}, \boldsymbol{\phi}^{(n)}\}$ generated by Algorithm 5 is always feasible for Problem (8). The monotonic property of Algorithm 5 can be similarly proved by using the method of [33].

In the following, we prove that the converged solution satisfies the KKT conditions of Problem (8). Let us denote the converged solution as $\{\mathbf{W}^\star, \mathbf{U}^\star, \mathbf{F}^\star, \boldsymbol{\Phi}^\star\}$.

According to Theorem 1, $\mathbf{F}^\star$ is the KKT-optimum point of Problem (15). Upon denoting the OF of Problem (15) as $z(\mathbf{F}, \boldsymbol{\Phi}^\star)$, the Lagrange function of Problem (15) is given by

$$\mathcal{L}(\mathbf{F}, \lambda, \mu) = z(\mathbf{F}, \boldsymbol{\Phi}^\star) + \lambda\left(\sum_{k=1}^{K_I}\|\mathbf{F}_k\|_F^2 - P_T\right) + \mu\left(\bar{Q} - \mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\mathrm{H}}\mathbf{G}\mathbf{F}_k\right)\right), \tag{D.2}$$

where $\lambda$ and $\mu$ are the corresponding dual variables. Then, there must exist a $\lambda^\star$ and $\mu^\star$ for ensuring that the following conditions are satisfied [3]:

$$\nabla_{\mathbf{F}_k^*}\mathcal{L}(\mathbf{F}, \lambda, \mu)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star} = \nabla_{\mathbf{F}_k^*} z(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star} + \lambda^\star\mathbf{F}_k^\star - \mu^\star\mathbf{G}\mathbf{F}_k^\star = \mathbf{0}, \forall k \in \mathcal{K}_I, \tag{D.3}$$

$$\lambda^\star\left(\sum_{k=1}^{K_I}\|\mathbf{F}_k^\star\|_F^2 - P_T\right) = 0, \tag{D.4}$$

$$\mu^\star\left(\bar{Q} - \mathrm{tr}\left(\sum_{k=1}^{K_I}\mathbf{F}_k^{\star\mathrm{H}}\mathbf{G}\mathbf{F}_k^\star\right)\right) = 0. \tag{D.5}$$

Furthermore, it can be readily checked that

$$\nabla_{\mathbf{F}_k^*} h\left(\mathbf{W}^\star, \mathbf{U}^\star, \mathbf{F}, \boldsymbol{\Phi}^\star\right)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}, = \nabla_{\mathbf{F}_k^*} z(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}\forall k \in \mathcal{K}_I. \tag{D.6}$$

To expound a little further, we have the following chain of inequalities:

$$\nabla_{\mathbf{F}_k^*} h_k\left(\mathbf{W}^\star, \mathbf{U}^\star, \mathbf{F}, \boldsymbol{\Phi}^\star\right)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star} \tag{D.7}$$

$$= -\mathrm{tr}\left(\mathbf{W}_k^\star\left(\nabla_{\mathbf{F}_k^*}\mathbf{E}_k\left(\mathbf{U}^\star, \mathbf{F}, \boldsymbol{\Phi}^\star\right)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}\right)\right) \tag{D.8}$$

$$= -\mathrm{tr}\left(\left(\mathbf{E}_k\left(\mathbf{U}^\star, \mathbf{F}^\star, \boldsymbol{\Phi}^\star\right)\right)^{-1}\left(\nabla_{\mathbf{F}_k^*}\mathbf{E}_k\left(\mathbf{U}^\star, \mathbf{F}, \Phi^\star\right)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}\right)\right) \tag{D.9}$$

$$= \left(\nabla_{\mathbf{F}_k^*}\log\left|\left(\mathbf{E}_k\left(\mathbf{U}^\star, \mathbf{F}, \boldsymbol{\Phi}^\star\right)\right)^{-1}\right|\right)\Big|_{\mathbf{F}_k=\mathbf{F}_k^\star} \tag{D.10}$$

$$= \nabla_{\mathbf{F}_k^*} R_k(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}, \tag{D.11}$$

where (D.8) follows from the chain rule, and the final equality follows from applying the Woodbury matrix identity to (14). Combining (D.11) with (D.6), we have

$$\nabla_{\mathbf{F}_k^*} z(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star} = \nabla_{\mathbf{F}_k^*} R_k(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star}. \tag{D.12}$$

By substituting (D.12) into (D.3), we arrive at

$$\nabla_{\mathbf{F}_k^*} R_k(\mathbf{F}, \boldsymbol{\Phi}^\star)\big|_{\mathbf{F}_k=\mathbf{F}_k^\star} + \lambda^\star\mathbf{F}_k^\star - \mu^\star\mathbf{G}\mathbf{F}_k^\star = \mathbf{0}, \forall k \in \mathcal{K}_I. \tag{D.13}$$

[3]For simplicity, the prime constraints are omitted.

According to Theorem 3, $\boldsymbol{\phi}^\star$ satisfies the KKT conditions of Problem (30), and thus the set of equations (C.5), (C.6) and (C.8) hold.

Furthermore, it can be readily verified that

$$\nabla_{\boldsymbol{\phi}^*} h\left(\mathbf{W}^\star, \mathbf{U}^\star, \mathbf{F}^\star, \boldsymbol{\Phi}\right)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*} \varphi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}. \tag{D.14}$$

By using similar derivations as in (D.7)-(D.11), we can prove that

$$\nabla_{\boldsymbol{\phi}^*} h\left(\mathbf{W}^\star, \mathbf{U}^\star, \mathbf{F}^\star, \boldsymbol{\Phi}\right)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*} R_k(\boldsymbol{\phi}, \mathbf{F}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}. \tag{D.15}$$

Hence, we have

$$\nabla_{\boldsymbol{\phi}^*} \varphi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \nabla_{\boldsymbol{\phi}^*} R_k(\boldsymbol{\phi}, \mathbf{F}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star}. \tag{D.16}$$

By substituting (D.16) into (C.8), we arrive at:

$$\nabla_{\boldsymbol{\phi}^*} R_k(\boldsymbol{\phi}, \mathbf{F}^\star)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} - \nu^\star\left(\mathbf{g}^* + \boldsymbol{\Upsilon}\boldsymbol{\phi}^\star\right) + \sum_{m=1}^{M} \tau_m^\star (\nabla_{\boldsymbol{\phi}^*} |\phi_m|)|_{\boldsymbol{\phi}=\boldsymbol{\phi}^\star} = \mathbf{0}. \tag{D.17}$$

Then, the set of equations (D.13), (D.4), (D.5), (D.17), (C.5), and (C.6) constitute exactly the KKT conditions of Problem (8).

## REFERENCES


[1] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come," *EURASIP J. Wireless Commun. Networking*, vol. 2019, no. 1, p. 129, 2019.

[2] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network." [Online]. Available: https://arxiv.org/abs/1905.00152

[3] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. Love, "Multiple antenna technologies for beyond 5G." [Online]. Available: https://arxiv.org/abs/1910.00092

[4] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light: Science & Applications*, vol. 3, no. 10, p. e218, 2014.

[5] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.

[6] J. Zhang, L. Dai, Z. He, B. Ai, and O. A. Dobre, "Mixed-ADC/DAC multipair massive MIMO relaying systems: Performance analysis and power optimization," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 140–153, 2018.

[7] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 1–6.

[8] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces." [Online]. Available: https://arxiv.org/abs/1904.12199

[9] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization." [Online]. Available: https://arxiv.org/abs/1906.09956

[10] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, 2019.

[11] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization." [Online]. Available: https://arxiv.org/abs/1907.06002

[12] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[13] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[14] H. Guo, Y. Liang, J. Chen, and E. Larsson, "Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks." [Online]. Available: https://arxiv.org/abs/1905.07920

[15] Q. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. Alouini, "Large intelligent surface assisted MIMO communications." [Online]. Available: https://arxiv.org/abs/1903.08127

[16] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces." [Online]. Available: https://arxiv.org/abs/1904.09573

[17] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, 2019.

[18] H. Shen, W. Xu, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, pp. 1–1, 2019.

[19] J. Chen, Y. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security." [Online]. Available: https://arxiv.org/abs/1905.03689

[20] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for secure IRS-assisted multiuser MISO systems." [Online]. Available: https://arxiv.org/abs/1907.03085

[21] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication via joint beamforming and jamming." [Online]. Available: https://arxiv.org/abs/1907.12839

[22] T. Bai, C. Pan, Y. Deng, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Latency minimization for intelligent reflecting surface aided mobile edge computing." [Online]. Available: https://arxiv.org/abs/1910.07990

[23] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems." [Online]. Available: https://arxiv.org/abs/1909.04606

[24] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Intelligent reflecting surface for multicell MIMO communications." [Online]. Available: https://arxiv.org/abs/1907.10864

[25] C. Huang, G. C. Alexandropoulos, C. Yuan, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces." [Online]. Available: https://arxiv.org/abs/1905.07726

[26] Z. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO." [Online]. Available: https://arxiv.org/abs/1905.07948

[27] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning." [Online]. Available: https://arxiv.org/abs/1904.10136

[28] G. Zhou, C. Pan, H. Ren, K. Wang, M. D. Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems." [Online]. Available: https://arxiv.org/abs/1911.06237

[29] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[30] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT." [Online]. Available: https://arxiv.org/abs/1907.05558

[31] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[32] C. Pan, W. Xu, J. Wang, H. Ren, W. Zhang, N. Huang, and M. Chen, "Pricing-based distributed energy-efficient beamforming for MISO interference channels," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 710–722, Apr. 2016.

[33] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.

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

[35] D. Mishra, G. C. Alexandropoulos, and S. De, "Energy sustainable IoT with individual QoS constraints through MISO SWIPT multicasting," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2856–2867, Aug. 2018.

[36] K. Xiong, B. Wang, and K. R. Liu, "Rate-energy region of SWIPT for MIMO broadcasting under nonlinear energy harvesting model," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5147–5161, Aug. 2017.

[37] F. Wang, J. Xu, X. Wang, and S. Cui, "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1784–1797, Mar. 2018.

[38] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4798–4810, Sep. 2014.

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

[40] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, and L. Hanzo, "The non-coherent ultra-dense C-RAN is capable of outperforming its coherent counterpart at a limited fronthaul capacity," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2549–2560, Nov. 2018.

[41] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Robust beamforming design for ultra-dense user-centric C-RAN in the face of realistic pilot contamination and limited feedback." [Online]. Available: https://arxiv.org/abs/1804.03990

[42] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.

[43] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[44] C. Pan, W. Xu, W. Zhang, J. Wang, H. Ren, and M. Chen, "Weighted sum energy efficiency maximization in ad hoc networks," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 233–236, Jun. 2015.

[45] X.-D. Zhang, *Matrix analysis and applications*. Cambridge University Press, 2017.

[46] J. Song, P. Babu, and D. P. Palomar, "Sequence design to minimize the weighted integrated and peak sidelobe levels," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2051–2064, Apr. 2016.


**Cunhua Pan** received the B.S. and Ph.D. degrees from the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2010 and 2015, respectively. From 2015 to 2016, he was a Research Associate at the University of Kent, U.K. He held a post-doctoral position at Queen Mary University of London, U.K., from 2016 and 2019, where he is currently a Lecturer (Assistant Professor) .

His research interests mainly include ultra-dense C-RAN, machine learning, UAV, Internet of Things, and mobile edge computing. He serves as a TPC member for numerous conferences, such as ICC and GLOBECOM, and the Student Travel Grant Chair for ICC 2019. He also serves as an Editor of IEEE ACCESS.

**Hong Ren** received the B.S. degree from Southwest Jiaotong University, in Electrical Engineering in 2011, Chengdu, China, and the M.S. and Ph.D. degrees from Southeast University, Nanjing, China, in Electrical Engineering in 2014 and 2018, respectively. From October 2016 to Jan 2018, she was a visiting student in the School of Electronics and Computer Science, University of Southampton, UK. She is currently a postdoctoral scholar with the School of Electronic Engineering and Computer Science, Queen Mary University of London, UK. Her research interests lie in the areas of communication and signal processing, including cooperative transmission, Internet of Things and ultra reliability and low latency communications.

**Kezhi Wang** received his B.E. and M.E. degrees in School of Automation from Chongqing University, China, in 2008 and 2011, respectively. He received his Ph.D. degree in Engineering from the University of Warwick, U.K. in 2015. He was a Senior Research Officer in University of Essex, U.K. Currently he is a Senior Lecturer in Department of Computer and Information Sciences at Northumbria University, U.K. His research interests include mobile edge computing, UAV communication and machine learning.

**Maged Elkashlan** received the Ph.D. degree in Electrical Engineering from the University of British Columbia, Canada, 2006. From 2007 to 2011, he was with the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. During this time, he held visiting appointments at University of New South Wales and University of Technology Sydney. In 2011, he joined the School of Electronic Engineering and Computer Science at Queen Mary University of London, UK. His research interests fall into the broad areas of communication theory and statistical signal processing. Dr. Elkashlan currently serves as Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He received the Best Paper Awards at the IEEE International Conference on Communications (ICC) in 2016 and 2014, the International Conference on Communications and Networking in China (CHINACOM) in 2014, and the IEEE Vehicular Technology Conference (VTC-Spring) in 2013.

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/JSAC.2020.3000802, IEEE Journal on Selected Areas in Communications



**Jiangzhou Wang** (F17) is currently a Professor and the former Head of the School of Engineering and Digital Arts at the University of Kent, U.K. He has published over 300 papers in international journals and conferences and 4 books in the areas of wireless mobile communications. Professor Wang is a Fellow of the Royal Academy of Engineering, U.K., and a Fellow of the IEEE. He was a recipient of the Best Paper Award from the IEEE GLOBECOM2012. He was an IEEE Distinguished Lecturer from 2013 to 2014. He was the Technical Program Chair of the 2019 IEEE International Conference on Communications (ICC2019), Shanghai, the Executive Chair of the IEEE ICC2015, London, and the Technical Program Chair of the IEEE WCNC2013. He has served as an Editor for a number of international journals, including IEEE Transactions on Communications from 1998 to 2013.

**Arumugam Nallanathan** (S97-M00-SM05-F17) is Professor of Wireless Communications and Head of the Communication Systems Research (CSR) group in the School of Electronic Engineering and Computer Science at Queen Mary University of London since September 2017. He was with the Department of Informatics at Kings College London from December 2007 to August 2017, where he was Professor of Wireless Communications from April 2013 to August 2017 and a Visiting Professor from September 2017. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include Artificial Intelligence for Wireless Systems, Beyond 5G Wireless Networks, Internet of Things (IoT) and Molecular Communications. He published nearly 500 technical papers in scientific journals and international conferences. He is a co-recipient of the Best Paper Awards presented at the IEEE International Conference on Communications 2016 (ICC'2016) , IEEE Global Communications Conference 2017 (GLOBECOM'2017) and IEEE Vehicular Technology Conference 2018 (VTC'2018). He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science Highly Cited Researcher in 2016.

He is an Editor for IEEE Transactions on Communications. He was an Editor for IEEE Transactions on Wireless Communications (2006-2011), IEEE Transactions on Vehicular Technology (2006-2017), IEEE Wireless Communications Letters and IEEE Signal Processing Letters. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society and Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.

**Lajos Hanzo** (M91SM92F04) received his 5-year Master degree in Electronics in 1976 and his doctoral degree in 1983. In 2004 the University of Southampton bestowed the Doctor of Science (DSc) higher research degree upon him. During his 44-year career in Telecommunications, he has held various research and academic positions in Hungary, Germany. and U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., where he holds the Chair of Telecommunications. He has successfully supervised 119 Ph.D. students, coauthored 18 John Wiley/IEEE Press books on mobile radio communications, totaling in excess of 10 000 pages, published 1329 research papers at IEEE Xplore, acted as TPC and the General Chair of IEEE conferences, presented keynote lectures, and has been awarded a number of distinctions. He is currently directing an academic research team, working on a range of research projects in the field of wireless multimedia communications, sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC), U.K., the European Research Councils Advanced Fellow Grant, and the Royal Society, UK. He is an enthusiastic supporter of industrial and academic liaison, and he offers a range of industrial courses. He is a Fellow of the Royal Academy of Evgineering (FREng), Fellow of the Institute of Engineering and Technology (FIET), Fellow of the European Signal Processing Association (EURASIP). In 2009, he received an Honorary Doctorate from the Technical University of Budapest, and again in 2015 by from University of Edinburgh. From 2008 to 2012, he was the Editor-in-Chief of the IEEE Press and a Chaired Professor at Tsinghua University, Beijing. His research is funded by the European Research Council's Advanced Fellow Grant. He is a Governor of the IEEE VTS.