IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 7, 2018, accepted April 3, 2018, date of publication April 10, 2018, date of current version May 2, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2825387

Secrecy Energy Efficiency in Wireless Powered Heterogeneous Networks: A Distributed ADMM Approach

XIN HU[1], BIN LI[2], KAIZHI HUANG[1], ZESONG FEI[2], (Senior Member, IEEE), AND KAI-KIT WONG[3], (Fellow, IEEE)

[1]National Digital Switching System Engineering and Technological R&D Center, Zhengzhou 450002, China
[2]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[3]Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, U.K.

Corresponding author: Kaizhi Huang (huangkaizhi@tsinghua.org.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61471396, Grant 61501516, Grant 61521003, Grant 61601514, and Grant 61701538.

ABSTRACT This paper investigates the physical layer security in heterogeneous networks (HetNets) supported by simultaneous wireless information and power transfer (SWIPT). We first consider a two-tier HetNet composed of a macrocell and several femtocells, where the macrocell base station (BS) serves multiple users in the presence of a malicious eavesdropper, while each femtocell BS serves a couple of Internet-of-Things (IoT) users. With regard to the energy constraint of IoT users, SWIPT is performed at the femtocell BSs, and IoT users accomplish the reception of information and energy in a time-switching manner, where information secrecy is to be protected. To enhance the secrecy performance, we inject artificial noise (AN) into the transmit beam at both macrocell and femtocell BSs, and for the sake of achieving green communications, we formulate the problem of maximizing secrecy energy efficiency while considering the fairness in a cross-tier multi-cell coordinated beamforming (MCBF) design. To handle this resulting nonconvex max–min fractional program problem, we propose an iterative algorithm by applying successive convex approximation method. Then, we further develop a decentralized solution based on alternative direction multiplier method (ADMM), which reduces the overhead of information exchange among coordinated BSs and achieves good approximation performance. Finally, simulation results demonstrate the performance of the proposed AN-aided cross-tier MCBF design and verify the validity of distributed ADMM-based approach.

INDEX TERMS Physical layer security, heterogeneous networks, secrecy energy efficiency, successive convex approximation, alternative direction multiplier method.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

With the tremendous growth of Internet-enabled smart devices, the next generation wireless communications have an ever-increasing demand for exceedingly high data rate, almost-ubiquitous coverage and reliable secrecy performance. In this regard, heterogeneous network (HetNet) [1], [2] has been emerged as a promising approach by increasing the cell density for higher spatial spectrum reuse. Unfortunately, the dense deployment of small cells (e.g., picocells, femtocells) also incurs great interferences that may deteriorate the quality of service (QoS) of users in macrocells, which is deemed as the main challenge in HetNets. Consequently, the transmit signals should be carefully designed to eliminate the mutual interferences.

Meanwhile, the energy consumption is bound to grow with the increase of data traffics, which may become unaffordable for low-power devices (e.g., nodes in wireless sensor networks or Internet of things (IoT)). Thus, more frequent recharging or replacement of the batteries is required to further prolong the lifetime of network, which accompanies the rising cost in maintenance. Recently, simultaneous wireless information and power transfer (SWIPT) [3], [4], acting as a novel research frontier of combining wireless power transfer and wireless communication, has been envisioned as a stable and continuous energy supply for the energy-constrained wireless systems. In SWIPT, devices can harvest energy from

radio-frequency (RF) signals ignoring the content carried, and this peculiarity contributes to the case of applying SWIPT into HetNets owning to the fact that all inter-tier and intra-tier interferences could contribute to the power supply for the harvesting devices.

The sharing of spectrum resources in addition to simultaneous transmission of power and information is achieved in virtue of the open nature of wireless propagation channels, which may however yield information leakage to the HetNets and SWIPT systems. To address this concern, innovative security technologies have been proposed. As an important complement to the traditional high-layer cryptographic methods, physical layer security (PLS) [5], [6] exploits the ''fingerprint'' of wireless channels to protect communications from being eavesdropped with low computational complexity, providing a novel approach to solving the issue of information leakage in HetNets and SWIPT systems, especially those including IoT devices that are poorly configured in hardwares. Thus, we focus on designing an efficient PLS scheme for SWIPT-enabled HetNets in this paper.

B. RELATED WORKS

Over the past few years, great achievements have been devoted to the issue of PLS in HetNets [7]–[11]. With the aid of stochastic geometry, Zhong *et al.* [7] proposed a tractable approach to analyze PLS in a K-tier HetNet, in which the secrecy coverage probability and the achievable secrecy rate for an arbitrary user were derived. Afterwards, Wang *et al.* [8] developed a user association policy based on an access threshold for PLS enhancement in the downlink of multi-tier HetNets. While Qi *et al.* [9] described a cognitive HetNet model under threshold-based protocols and derived the secrecy outage probability of cognitive users in the presence of non-colluding and colluding eavesdroppers, respectively. Moreover, the PLS performance in the downlink HetNets was promoted by using friendly jammers and full-duplex users in [10]. Previous research works studied the PLS in HetNets by focusing on the secrecy model and performance evaluations, but as mentioned before, the delicate design of transmit signals is also of great significance to the QoS of users in HetNets, which was rarely considered except in [11], where a joint resource allocation algorithm was proposed by simultaneously considering the PLS, cross-tier interferences in addition to the fairness requirement.

For securing SWIPT systems with PLS, there have also been remarkable accomplishments. To name a few, Liu *et al.* [12] considered the case where the BS aimed to transmit confidential information to an information receiver (IR) in the presence of multiple energy receivers (ERs).[1] They regarded the ERs as independent potential eavesdroppers and designed the transmit beam to maximize the secrecy rate of the IR subject to the energy constraint of ERs. On this basis, the robust beam design was carried out when the channel state information (CSI) was imperfectly known at the transmitter in [13]. While Chu *et al.* [14] dealt with the case where multiple ERs colluded with each other to overhear an IR jointly. In addition, Zhang *et al.* [15] expanded the number of available antennas at ERs, and the work in [16] considered a more complicated scenario, in which the IR and ERs were all equipped with multiple antennas. In addition, artificial noise (AN) was introduced to further safeguard SWIPT systems, and the joint design of AN vector and transmit beam was investigated detailedly in [17]–[19]. The above-mentioned works studied the PLS transmission in different scenarios, but were all limited to the single-cell coverage case, which ignored the deployment of SWIPT in multi-cell networks.

Recently, applying SWIPT to HetNets has attracted increasing interests by converting various interferences into power gain, and PLS was also introduced to the fusion case (i.e., SWIPT-enabled HetNets) in [20]–[22]. Specifically, Akbar *et al.* [20] established a tractable model for the joint uplink (UL) and downlink (DL) transmission in a K-tier HetNet with SWIPT, they derived the exact analytical expressions for the average received power and the outage probability in both DL and UL transmissions. Li *et al.* [21] considered a two-tier HetNet with SWIPT performed in the femtocells, and they accomplished the beam design to maximize the secrecy rate at IR in a femtocell, while satisfying the QoS requirements at each ER and macrocell user (MU). Similarly, the transmit beams and AN vectors at the macrocell BS and femtocell BSs were jointly optimized to maximize the secrecy rate of the eavesdropped MU under the QoS constraints of IRs and ERs in [22]. Nonetheless, [21] and [22] both only considered the security of one type of decoding users (i.e., either MU or IR), which ignored the worst case where both IRs and MUs are wiretapped simultaneously by malicious eavesdroppers (ERs included). Last but not least, the joint design was executed at a calculating center in the above works, and thus the exchange of CSI among BSs was inevitable. However, this will incur signaling overhead to the network and inestimable processing stress to the calculating center (especially when the number of coordinated BSs is large), which should be carefully handled.

C. OUR APPROACH AND CONTRIBUTIONS

In this paper, we consider a two-tier SWIPT-enabled HetNet, and focus on the transmit optimization while considering the aim of green communications as well as the security of various decoding users. Specifically, our main contributions can be summarized as follows:

- We establish the model of a two-tier HetNet, where SWIPT is performed at the femtocell BSs to power IoT users. We consider a more general case where both the IoT users and MUs are faced with the threat of secrecy leakage. To be specific, two kinds of malicious users are included, the MUs are wiretapped by a passive eavesdropper in the macrocell, and the harvesting IoT users

[1] IR and ER are both legitimated users in the network, where the IR has permission for information receiving and decoding, while the ER is only allowed to harvest energy from the ambient power of RF signals.

act as curious users that are interested in the information intended for decoding IoT users within the same femtocell.

- For the sake of security enhancement and power transfer, we deliberately inject the AN into the transmit beam at both the macrocell and femtocell BSs. To achieve green communications and promote the average secrecy performance, we formulate the joint design problem for secrecy energy efficiency maximization while considering the fairness among multiple cells.
- To solve the nonconvex max-min fractional program, we first apply the semi-definite relaxation (SDR) technique by holding the tightness of relaxations, and then propose an iterative algorithm based on successive convex approximation (SCA) with the complexity analyzed.
- To reduce the overhead of CSI exchange in the centralized multi-cell coordinated beamforming design (MCBF) and release the processing stress on the calculating center, we further introduce a decentralized solution based on the alternative direction multiplier method (ADMM), which can draw near the optimal solution while allowing each BS to handle its local CSI only.

The rest of this paper is organized as follows. Section II outlines the system model and formulates the optimization problem. Section III solves the nonconvex problem with SDR technique and proposes an SCA-based iterative algorithm. Section IV introduces a decentralized solution based on ADMM. Section V provides the simulation results and finally Section VI concludes this paper.

Notations: The expectation of a variable is denoted by $\mathbb{E}(\cdot)$. The magnitude of a complex number is denoted by $|\cdot|$. $\|\cdot\|_2$ represents the ℓ_2 norm. $\mathbb{C}^{p\times q}$ represents the space of complex matrices of dimension $p \times q$. The superscript $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and (Hermitian) conjugate transpose, respectively. Tr($\mathbf{A}$) and rank($\mathbf{A}$) represent the trace and the rank of matrix $\mathbf{A}$, respectively. $\{\mathbf{a}_m(a_m)\}_{m\in\mathcal{M}}$ denotes the vector composited by all the $\mathbf{a}_m(a_m)$ where $m \in \mathcal{M}$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. NETWORK STRUCTURE

We consider a two-tier HetNet with SWIPT, as illustrated in Fig. 1.[2] The network is composed of a macrocell BS (MBS) and N femtocell BSs (FBSs), which share the same spectrum resources to improve the spectral efficiency. The MBS is equipped with K_m antennas to serve M ($K_m > M$) MUs, and there exists an eavesdropper (Eve)[3] attempting to intercept

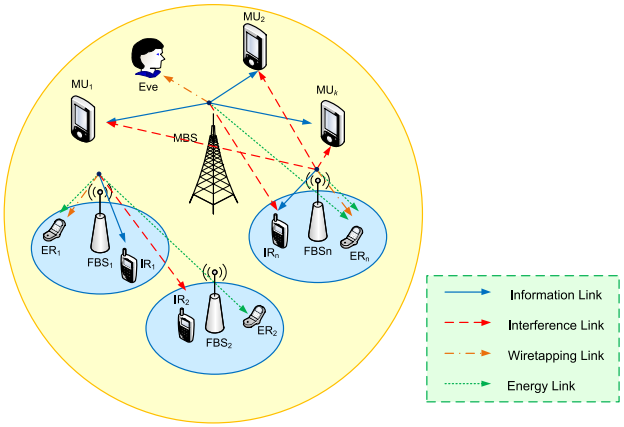

FIGURE 1. Illustration of a two-tier SWIPT-enabled HetNet.

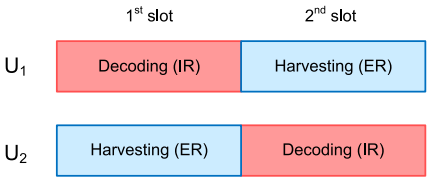

FIGURE 2. Parallel reception of information and power in the TS manner.

the information intended for multiple MUs; while each FBS is equipped with K_f ($K_f > 2$) antennas to serve a pair of femtocell users (FUs). All the users are equipped with a single antenna. We assume that the FUs are the devices deployed in IoT, which are severely constrained in energy, hence SWIPT is performed at the FBSs, and IoT users accomplish the reception of information and energy in a time-switching (TS) manner [23], which can be shown in Fig. 2. The DL transmission in each femtocell can be separated into two slots, which is a simple but representable case. In each slot, the FBS transmits confidential information to a certain FU (i.e., IR) and transfers energy to the other (i.e., ER), while the interconversion of roles takes place in the following slot. In this regard, there is one IR and ER in an arbitrary slot, which we extract for detailed studies.

We highlight that the ERs are deemed as legitimated users in the network but can only harvest energy from the ambient RF signals to accomplish the battery recharging. Unfortunately, the FBSs cannot strictly control the behaviors of IoT users, which indicates the ERs may turn curious to decode the private information intended for IRs [13], [17]. In this paper, we adopt this assumption and further suppose that each ER only attempts to overhear the IR in the same femtocell.[4]

For the sake of notational simplicity, we denote the m-th MU in the macrocell and the IR/ER in the n-th femtocell by MU_m ($m \le M$) and IR_n/ER_n ($n \le N$), respectively.

[2]In Fig. 1, the links between the macrocell and femtocell_n illustrate the inter-tier interferences, while the links between femtocell_1 and femtocell_2 demonstrate the intra-tier inferences. In other words, there indeed exist interference links between any two femtocells, but all are omitted except for those between femtocell_1 and femtocell_2 for the sake of clarity. Similarly, the interference links between the macrocell and femtocell_1 (femtocell_2) are represented by the links between the macrocell and femtocell_n.

[3]Note that we only consider one Eve in this paper for simplicity of presentation and calculation. In fact, our work can be extended to the scenario with multiple Eves in which the considered Eve is the one that achieves the maximal wiretapping capacity. The more general case with the existence of multiple Eves will be studied in our future work.

[4]When the femtocells belong to different parties, and there is no interest conflict among them but a competitive relationship between two IoT users in the same femtocell, it is a reasonable case where they are only interested in their competitors.

B. SIGNAL MODEL

Considering the existence of a malicious Eve in the macro-cell, we adopt the AN to enhance the security of transmission. The transmit signal at the MBS can be expressed as

$$\mathbf{x}_0(t) = \sum_{m=1}^{M} \mathbf{w}_m s_m(t) + \mathbf{z}_0(t), \tag{1}$$

where $\mathbf{w}_m \in \mathbb{C}^{K_m \times 1}$ is the beamforming vector aiming at MU_m; $s_m(t)$ with $\mathbb{E}\{|s_m|^2\} = 0$ represents the data symbol intended for MU_m; $\mathbf{z}_0 \in \mathbb{C}^{K_m \times 1}$ denotes the AN vector which follows Gaussian distribution, i.e., $\mathbf{z}_0 \sim \mathcal{CN}(0, \mathbf{\Sigma})$ and $\mathbf{\Sigma} \succeq \mathbf{0}$.

Similarly, the transmit signal at FBS_n, can be given by

$$\mathbf{x}_n(t) = \mathbf{v}_n c_n(t) + \mathbf{u}_n(t), \tag{2}$$

where $c_n(t)$ with $\mathbb{E}\{|c_n|^2\} = 0$ represents the data symbol intended for $\mathrm{IR_n}$ and $\mathbf{v}_n \in \mathbb{C}^{K_f \times 1}$ denotes the corresponding beamforming vector, while $\mathbf{u}_n \in \mathbb{C}^{K_f \times 1}$ is the energy vector utilized for transferring power to $\mathrm{ER_n}$ as well as confusing it in $c_n(t)$.

Then, the received signal at MU_m, Eve, IR_n and ER_n can be respectively expressed as

$$\mathbf{y}_m(t) = \underbrace{\mathbf{h}_m^H \mathbf{w}_m s_m(t)}_{\text{desired signal}} + \underbrace{\mathbf{h}_m^H \Big[\sum_{i=1, i\neq m}^{M} \mathbf{w}_i s_i(t) + \mathbf{z}_0(t)\Big]}_{\text{intra-tier interference}} + \underbrace{\sum_{j=1}^{N} \mathbf{h}_{jom}^H \Big[\mathbf{v}_j c_j(t) + \mathbf{u}_j(t)\Big]}_{\text{inter-tier interference}} + n_m(t), \tag{3}$$

$$\mathbf{y}_e^m(t) = \underbrace{\mathbf{h}_e^H \mathbf{w}_m s_m(t)}_{\text{desired signal}} + \underbrace{\mathbf{h}_e^H \Big[\sum_{i=1, i\neq m}^{M} \mathbf{w}_i s_i(t) + \mathbf{z}_0(t)\Big]}_{\text{intra-tier interference}} + \underbrace{\sum_{j=1}^{N} \mathbf{h}_{joe}^H \Big[\mathbf{v}_j c_j(t) + \mathbf{u}_j(t)\Big]}_{\text{inter-tier interference}} + n_e(t), \tag{4}$$

$$\mathbf{y}_n^{\mathrm{IR}}(t) = \underbrace{\mathbf{h}_{nn}^H \mathbf{v}_n c_n(t)}_{\text{desired signal}} + \underbrace{\mathbf{h}_{0n}^H \Big[\sum_{i=1}^{M} \mathbf{w}_i s_i(t) + \mathbf{z}_0(t)\Big]}_{\text{inter-tier interference}} + \underbrace{\mathbf{h}_{nn}^H \mathbf{u}_n(t) + \sum_{j=1, j\neq n}^{N} \mathbf{h}_{jn}^H \Big[\mathbf{v}_j c_j(t) + \mathbf{u}_j(t)\Big]}_{\text{intra-tier interference}} + n_{nir}(t), \tag{5}$$

and

$$\mathbf{y}_n^{\mathrm{ER}}(t) = \underbrace{\mathbf{g}_{nn}^H \mathbf{v}_n c_n(t)}_{\text{desired signal}} + \underbrace{\mathbf{g}_{0n}^H \Big[\sum_{i=1}^{N} \mathbf{w}_i s_i(t) + \mathbf{z}_0(t)\Big]}_{\text{inter-tier interference}} + \underbrace{\mathbf{g}_{nn}^H \mathbf{u}_n(t) + \sum_{j=1, j\neq n}^{N} \mathbf{g}_{jn}^H \Big[\mathbf{v}_j c_j(t) + \mathbf{u}_j(t)\Big]}_{\text{intra-tier interference}} + n_{ner}(t), \tag{6}$$

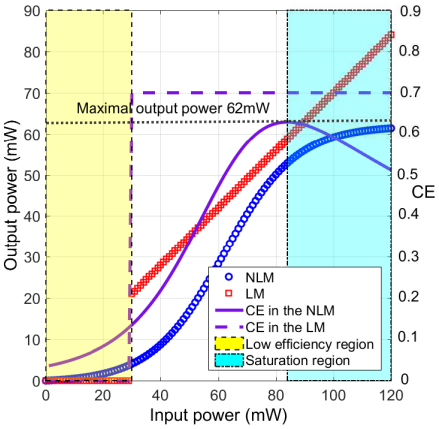

FIGURE 3. Comparison between non-linear model and linear model.

where $\mathbf{h}_m \in \mathbb{C}^{K_m \times 1}$, $\mathbf{h}_e \in \mathbb{C}^{K_m \times 1}$, $\mathbf{h}_{0n} \in \mathbb{C}^{K_m \times 1}$ and $\mathbf{g}_{0n} \in \mathbb{C}^{K_m \times 1}$ denote the channels from MBS to MU_m, Eve, IR_n and ER_n, respectively; likewise, $\mathbf{h}_{jom} \in \mathbb{C}^{K_f \times 1}$, $\mathbf{h}_{joe} \in \mathbb{C}^{K_f \times 1}$, $\mathbf{h}_{jn} \in \mathbb{C}^{K_f \times 1}$ and $\mathbf{g}_{jn} \in \mathbb{C}^{K_f \times 1}$ denote the channels from FBS_j to $\mathrm{MU_m}$, Eve, IR_n and ER_n, respectively; $n_m(t)$, $n_e(t)$, $n_{nir}(t)$ and $n_{ner}(t)$ denote the additive Gaussian white noise (AWGN) with variance σ_m^2, σ_e^2, σ_{nir}^2 and σ_{ner}^2 at MU_m, Eve, IR_n and ER_n, respectively.

Following (3), (4), (5) and (6), the received signal-to-interference-plus-noise ratio (SINR) at MU_m, Eve, IR_n and ER_n can be respectively given by (7), (8), (9) and (10), which are presented at the top of next page.

C. HARVESTING MODEL

Now we focus on the wireless power transfer in femto-cells and discuss the energy harvesting (EH) process of ERs. In most of the existing research works, EH process is modeled by the traditional linear model (LM), with which the functional relation between output direct current (DC) power P_{out} and input RF power P_{in} can be expressed as $P_{out} = \eta P_{in}$, where η denotes a constant conversion efficiency (CE). However, in practice, various non-linearities indicate that the CE dynamically varies with the input power, thus the simplified linear approximation of conversion-circuit can not accurately characterize the EH performance. To better reflect the dynamism of CE, we adopt a more practical non-linear model (NLM) for EH process in this paper, which was first put forward in [24] and [25].

With NLM, the harvested energy of ER_n can be expressed as

$$Q_{\mathrm{ER}}^n = \frac{\psi\left[1 + \exp(\mu\lambda)\right]}{\left[1 + \exp\left(-\mu\left(P_{in}^{ner} - \lambda\right)\right)\right]\exp(\mu\lambda)} - \frac{\psi}{\exp(\mu\lambda)}, \tag{11}$$

$$\mathrm{SINR}_m = \frac{\left|\mathbf{h}_m^H \mathbf{w}_m\right|^2}{\sum\limits_{i=1,i\neq m}^{M}\left|\mathbf{h}_m^H \mathbf{w}_i\right|^2 + \left|\mathbf{h}_m^H \mathbf{z}_0\right|^2 + \sum\limits_{j=1}^{N}\left(\left|\mathbf{h}_{jom}^H \mathbf{v}_n\right|^2 + \left|\mathbf{h}_{jom}^H \mathbf{u}_n\right|^2\right) + {\sigma_m}^2}, \tag{7}$$

$$\mathrm{SINR}_e^m = \frac{\left|\mathbf{h}_e^H \mathbf{w}_m\right|^2}{\sum\limits_{i=1,i\neq m}^{M}\left|\mathbf{h}_e^H \mathbf{w}_i\right|^2 + \left|\mathbf{h}_e^H \mathbf{z}_0\right|^2 + \sum\limits_{j=1}^{N}\left(\left|\mathbf{h}_{joe}^H \mathbf{v}_n\right|^2 + \left|\mathbf{h}_{joe}^H \mathbf{u}_n\right|^2\right) + {\sigma_e}^2}, \tag{8}$$

$$\mathrm{SINR}_{nir} = \frac{\left|\mathbf{h}_{nn}^H \mathbf{v}_n\right|^2}{\left|\mathbf{h}_{nn}^H \mathbf{u}_n\right|^2 + \sum\limits_{j=1,j\neq n}^{N}\left(\left|\mathbf{h}_{jn}^H \mathbf{v}_j\right|^2 + \left|\mathbf{h}_{jn}^H \mathbf{u}_j\right|^2\right) + \sum\limits_{i=1}^{M}\left(\left|\mathbf{h}_{0n}^H \mathbf{w}_i\right|^2\right) + \left|\mathbf{h}_{0n}^H \mathbf{z}_0\right|^2 + {\sigma_{nir}}^2}, \tag{9}$$

$$\mathrm{SINR}_{ner} = \frac{\left|\mathbf{g}_{nn}^H \mathbf{v}_n\right|^2}{\left|\mathbf{g}_{nn}^H \mathbf{u}_n\right|^2 + \sum\limits_{j=1,j\neq n}^{N}\left(\left|\mathbf{g}_{jn}^H \mathbf{v}_j\right|^2 + \left|\mathbf{g}_{jn}^H \mathbf{u}_j\right|^2\right) + \sum\limits_{i=1}^{N}\left(\left|\mathbf{g}_{0n}^H \mathbf{w}_i\right|^2\right) + \left|\mathbf{g}_{0n}^H \mathbf{z}_0\right|^2 + {\sigma_{ner}}^2}. \tag{10}$$

where ψ represents the maximum output power of RF-to-DC circuit, μ and λ are both constant parameters depending on the hardware components (e.g., resistance and capacitance), which have influence on the sensitivity of RF-to-DC circuit.

Fig. 3 provides an example of the NLM with $\psi = 62$mW, $\mu = 80$, $\lambda = 0.062$ and compares that to the LM with $\eta = 0.7$ to illustrate the promotion of rationality in two different aspects. For one thing, the CE varies consecutively near the sensitivity threshold (30mW) rather than presenting a step response from 0 to 0.7, which accords better with the practice; for another, as the input power rises up, the output power gradually tends to saturate rather than growing infinitely with the increment of input power, which coincides with the simple configuration in hardwares of IoT devices.

In account of the sensitivity threshold of RF-to-DC circuit, the AWGN power is ignorant for EH process, hence the input power P_{in} at ER_n is given by

$$P_{in}^{ner} = \sum_{j=1}^{N}\left(\left|\mathbf{g}_{jn}^H \mathbf{v}_n\right|^2 + \left|\mathbf{g}_{jn}^H \mathbf{u}_n\right|^2\right) + \sum_{i=1}^{M}\left(\left|\mathbf{g}_{0n}^H \mathbf{w}_i\right|^2\right) + \left|\mathbf{g}_{0n}^H \mathbf{z}_0\right|^2. \tag{12}$$

D. PROBLEM FORMULATION

Based on above analysis, we aim to secure the transmissions to legitimated decoding users (i.e., MUs and IRs) in the presence of malicious terminals (i.e., Eve) and curious users (i.e., ERs) by carefully designing the AN-aided MCBF. We pay attention to the secrecy energy efficiency (SEE) to achieve green communications. Further, we aim to improve the average secrecy performance by guaranteeing that all the cells can achieve a certain SEE level so that they could be equal to each other, this is what "fairness" implies. To this end, we maximize the minimal SEE among all the cells as

$$\max_{\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0} \min_{\aleph\in[0,N]} \frac{C_s^{\aleph}}{P_{\aleph}/\varpi + P_{\varepsilon}} \tag{13a}$$

$$\text{s.t. } Q_{\mathrm{ER}}^n(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0) \geq \theta, \quad \forall n \in [1,N], \tag{13b}$$

$$\frac{1}{\varpi}P_0(\mathbf{w},\mathbf{z}_0) + P_{\varepsilon} \leq \Delta, \tag{13c}$$

$$\frac{1}{\varpi}P_n(\mathbf{v},\mathbf{u}) + P_{\varepsilon} \leq \Gamma, \quad \forall n \in [1,N], \tag{13d}$$

where $\mathbf{w} = [\mathbf{w}_1, \ldots, \mathbf{w}_M] \in \mathbb{C}^{K_m \times M}$ denotes the matrix that stores the beamforming vectors at MBS, $\mathbf{v} = [\mathbf{v}_1, \ldots, \mathbf{v}_N] \in \mathbb{C}^{K_f \times N}$ and $\mathbf{u} = [\mathbf{u}_1, \ldots, \mathbf{u}_N] \in \mathbb{C}^{K_f \times N}$ denote the matrices that store the beamforming and AN vectors at FBSs, respectively; $C_s^n = \log_2(1 + \mathrm{SINR}_{nir}) - \log_2(1 + \mathrm{SINR}_{ner})$ represents the secrecy capacity of IR_n ($n \in [1,N]$), and $C_s^0 = \sum\limits_{m=1}^{M}\left[\log_2(1 + \mathrm{SINR}_m) - \log_2\left(1 + \mathrm{SINR}_e^m\right)\right]$ denotes the sum of secrecy capacity at MUs; P_0 and P_n denote the power consumption of MBS and FBS_n, respectively; $\varpi \in (0,1]$ is the power amplifier efficiency defined as the ratio between the total RF output power and the DC input power, and P_{ε} represents the power loss in hardwares; θ denotes the required power at each ER; Δ is the maximum output power of MBS, while Γ represents that of each FBS.

Obviously, problem (13) is a max-min fractional program which is nonconvex and thus cannot be solved directly and efficiently. Therefore, we will first introduce some slack variables and solve it with semi-definite program (SDP) method in the next section.

III. OPTIMIZATION IN THE SEE MAXIMIZATION DESIGN

We start by introducing an equivalent reformulation which exposes the hidden convexity of the proposed max-min SEE problem. As a consequence, problem (13) can be rewritten as

$$\max_{\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0,r,\zeta} \quad \zeta \tag{14a}$$

$$\text{s.t. } \frac{P_{\aleph}}{\varpi} + P_{\varepsilon} \leq r_{\aleph}, \quad \forall \aleph \in [0,N], \tag{14b}$$

$$C_s^{\aleph}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0) \geq r_{\aleph}\zeta, \quad \forall \aleph \in [0,N], \tag{14c}$$

$$\text{constraints (13b)}\sim\text{(13d)}, \tag{14d}$$

where ζ and $\boldsymbol{r} \triangleq [r_0, r_1, \ldots, r_N]$ are newly introduced slack variables. It is easy to check that the equalities in (14b) and (14c) hold at the optimal solution. Otherwise, the equalities are acquirable through decreasing ζ, and this operation will not change the optimal objective value.

On this basis, to deal with the non-convexity of (14c), we introduce several new auxiliary variables, i.e., $\boldsymbol{l} \triangleq [l_0, l_1, \ldots, l_N]$, $\boldsymbol{\tau}^a \triangleq \left[\tau_1^a, \ldots, \tau_N^a\right]$, $\boldsymbol{\tau}^b \triangleq \left[\tau_1^b, \ldots, \tau_N^b\right]$, $\boldsymbol{\tau}_0^a \triangleq \left[\tau_{01}^a, \ldots, \tau_{0M}^a\right]$, $\boldsymbol{\tau}_0^b \triangleq \left[\tau_{01}^b, \ldots, \tau_{0M}^b\right]$, and reformulate the problem mentioned above as

$$\max_{\substack{\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0,\boldsymbol{r},\zeta \\ \boldsymbol{l},\boldsymbol{\tau}^a,\boldsymbol{\tau}^b,\boldsymbol{\tau}_0^a,\boldsymbol{\tau}_0^b}} \quad \zeta \tag{15a}$$

$$\text{s.t. } \log_2\left(\tau_n^a\right) \geq l_n, \quad \forall n \in [1, N], \tag{15b}$$

$$l_n - \log_2\left(\frac{1}{\tau_n^b}\right) \geq r_n\zeta, \quad \forall n \in [1, N], \tag{15c}$$

$$\sum_{m=1}^{M} \log_2\left(\tau_{0m}^a\right) \geq l_0, \tag{15d}$$

$$l_0 - \sum_{m=1}^{M} \log_2\left(\frac{1}{\tau_{0m}^b}\right) \geq r_0\zeta, \tag{15e}$$

$$1 + \mathrm{SINR}_{nir} \geq \tau_n^a, \quad \forall n \in [1, N], \tag{15f}$$

$$1 + \mathrm{SINR}_{ner} \leq \frac{1}{\tau_n^b}, \quad \forall n \in [1, N], \tag{15g}$$

$$1 + \mathrm{SINR}_m \geq \tau_{0m}^a, \quad \forall m \in [1, M], \tag{15h}$$

$$1 + \mathrm{SINR}_e^m \leq \frac{1}{\tau_{0m}^b}, \quad \forall m \in [1, M], \tag{15i}$$

$$\text{constraint (14d)}, \tag{15j}$$

$$\tau_n^a \geq 1, \quad \tau_{0m}^a \geq 1, \ 1 \geq \tau_n^b \geq 0, \ 1 \geq \tau_{0m}^b \geq 0,$$
$$\forall m \in [1, M], \quad \forall n \in [1, N].$$

In the same spirit, we transform (15f)∼(15i) into convex ones with the aid of extra auxiliary variables, i.e., $\boldsymbol{\kappa}^1 \triangleq \left[\kappa_1^1, \ldots, \kappa_N^1\right]$, $\boldsymbol{\kappa}^2 \triangleq \left[\kappa_1^2, \ldots, \kappa_N^2\right]$, $\boldsymbol{\kappa}_0^1 \triangleq \left[\kappa_{01}^1, \ldots, \kappa_{0M}^1\right]$ and $\boldsymbol{\kappa}_0^2 \triangleq \left[\kappa_{01}^2, \ldots, \kappa_{0M}^2\right]$. By applying SDR technique and remove the rank-one constraints, problem (15) can be recast as

$$\max_{\substack{\mathbf{V},\mathbf{U},\mathbf{W},\mathbf{Z}_0,\boldsymbol{r},\boldsymbol{l},\zeta \\ \boldsymbol{\tau}^a,\boldsymbol{\tau}^b,\boldsymbol{\tau}_0^a,\boldsymbol{\tau}_0^b, \\ \boldsymbol{\kappa}^1,\boldsymbol{\kappa}^2,\boldsymbol{\kappa}_0^1,\boldsymbol{\kappa}_0^2}} \quad \zeta \tag{16a}$$

$$\text{s.t. } \mathrm{Tr}\left(\mathbf{H}_{nn}\mathbf{V}_n\right) \geq \left(\tau_n^a - 1\right)\kappa_n^1, \tag{16b}$$

$$\sum_{j=1, j\neq n}^{N} \left[\mathrm{Tr}\left(\mathbf{H}_{jn}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{H}_{jn}\mathbf{U}_j\right)\right] + \mathrm{Tr}\left(\mathbf{H}_{nn}\mathbf{U}_n\right)$$
$$+ \sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{H}_{0n}\mathbf{W}_i\right) + \mathrm{Tr}\left(\mathbf{H}_{0n}\mathbf{Z}_0\right) + \sigma_{nir}^2 \leq \kappa_n^1, \tag{16c}$$

$$\sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{W}_i\right) + \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{Z}_0\right) + \sum_{j=1, j\neq n}^{N} \left[\mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{U}_j\right)\right] + \mathrm{Tr}\left(\mathbf{G}_{nn}\mathbf{U}_n\right) + \sigma_{ner}^2 \geq \kappa_n^2\tau_n^b, \tag{16d}$$

$$\sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{W}_i\right) + \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{Z}_0\right) + \sum_{j=1}^{N} \left[\mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{U}_j\right)\right] + \sigma_{ner}^2 \leq \kappa_n^2, \tag{16e}$$

$$\mathrm{Tr}\left(\mathbf{H}_m\mathbf{W}_m\right) \geq \left(\tau_{0m}^a - 1\right)\kappa_{0m}^1, \tag{16f}$$

$$\sum_{j=1}^{N} \left[\mathrm{Tr}\left(\mathbf{H}_{jom}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{H}_{jom}\mathbf{U}_j\right)\right] + \mathrm{Tr}\left(\mathbf{H}_m\mathbf{Z}_0\right)$$
$$+ \sum_{i=1, i\neq m}^{M} \mathrm{Tr}\left(\mathbf{H}_m\mathbf{W}_i\right) + \sigma_m^2 \leq \kappa_{0m}^1, \tag{16g}$$

$$\sum_{j=1}^{N} \left[\mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{U}_j\right)\right] + \mathrm{Tr}\left(\mathbf{H}_e\mathbf{Z}_0\right)$$
$$+ \sum_{i=1, i\neq m}^{M} \mathrm{Tr}\left(\mathbf{H}_e\mathbf{W}_i\right) + \sigma_e^2 \geq \kappa_{0m}^2\tau_{0m}^b, \tag{16h}$$

$$\sum_{j=1}^{N} \left[\mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{U}_j\right)\right] + \mathrm{Tr}\left(\mathbf{H}_e\mathbf{Z}_0\right)$$
$$+ \sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{H}_e\mathbf{W}_i\right) + \sigma_e^2 \leq \kappa_{0m}^2, \tag{16i}$$

$$\sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{W}_i\right) + \mathrm{Tr}\left(\mathbf{Z}_0\right) \leq \Delta\varpi - P_\varepsilon, \tag{16j}$$

$$\mathrm{Tr}\left(\mathbf{V}_n\right) + \mathrm{Tr}\left(\mathbf{U}_n\right) \leq \Gamma\varpi - P_\varepsilon, \tag{16k}$$

$$\sum_{j=1}^{N} \left[\mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{U}_j\right)\right] + \sum_{i=1}^{M} \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{W}_i\right)$$
$$+ \mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{Z}_0\right) \geq \Lambda\left(\theta\right), \tag{16l}$$

$$\text{constraints (15b)} \sim \text{(15e)}, \tag{16m}$$

$$\mathbf{V}_n \succeq \mathbf{0}, \quad \mathbf{U}_n \succeq \mathbf{0}, \ \mathbf{W}_m \succeq \mathbf{0}, \ \mathbf{Z}_0 \succeq \mathbf{0},$$
$$\kappa_n^1 \geq 0, \quad kappa_{0m}^1 \geq 0, \ \kappa_n^2 \geq 0, \ \kappa_{0m}^2 \geq 0,$$
$$\forall n \in [1, N], \quad \forall m \in [1, M],$$

where

$$\Lambda\left(\theta\right) = \lambda - \ln\left[\frac{\psi\left[1 + \exp\left(\mu\lambda\right)\right]}{\left(\theta + \psi/\exp\left(\mu\lambda\right)\right)\exp\left(\mu\lambda\right)} - 1\right] \Big/ \mu \tag{17}$$

represents the inverse function of RF-to-DC conversion in NLM; $\mathbf{H}_\ulcorner = \mathbf{h}_\ulcorner\mathbf{h}_\ulcorner^H (\ulcorner \in [m, e, 0n, nn, jn, jom, joe])$, $\mathbf{G}_\urcorner = \mathbf{g}_\urcorner\mathbf{g}_\urcorner^H (\urcorner \in [0n, nn, jn])$, $\mathbf{W}_m = \mathbf{w}_m\mathbf{w}_m^H (m \in [1, M])$, $\mathbf{Z}_0 = \mathbf{z}_0\mathbf{z}_0^H$, $\mathbf{V}_n = \mathbf{v}_n\mathbf{v}_n^H$ and $\mathbf{U}_n = \mathbf{u}_n\mathbf{u}_n^H (n \in [1, N])$.

However, problem (16) is still nonconvex owning to the form of $f(x, y) = xy$ in (16b), (16d), (16f), (16h) and (15c), (15e) in (16m). Fortunately, we notice that the coupling of variables all occur in the right hand side of a constraint with $\geq$, which indicates that we can replace them with their upper bounds in convex forms without relaxation. In this regard, we apply conservatively convex-constrained

conditions based on SCA algorithm, the detailed introduction of which can be obtained in [26]. For any ξ, we define the following function

$$g_\xi(x,y) = \frac{\xi}{2}x^2 + \frac{1}{2\xi}y^2, \tag{18}$$

which serves as the upper bound of $f(x,y)$ according to the arithmetic-geometric mean inequality, and the equality holds if and only if $\xi = \frac{x}{y}$. Obviously, if $f(x,y)$ is replaced by $g_\xi(x,y)$, each of above nonconvex constraints can be conservatively approximated into several convex quadratic constraints, and the problem can be transformed into solving a series of convex problems iteratively. Therefore, in the q-th iterative approximation, (15c), (15e), (16b), (16d), (16f) and (16h) can be transformed into (19)∼(24), respectively, which can be expressed as

$$l_n + \log_2\left(\tau_n^b\right) \geq g_{\xi_{o1}}^{(q)}(r_n,\zeta), \tag{19}$$

$$l_0 + \sum_{m=1}^{M}\log_2\left(\tau_{0m}^b\right) \geq g_{\xi_{o2}}^{(q)}(r_0,\zeta), \tag{20}$$

$$\mathrm{Tr}\left(\mathbf{H}_{nn}\mathbf{V}_n\right) \geq g_{\xi_a}^{(q)}\left(\tau_n^a,\kappa_n^1\right) - \kappa_n^1, \tag{21}$$

$$\begin{aligned}\mathrm{Tr}\left(\mathbf{G}_{nn}\mathbf{U}_n\right) + \sum_{j=1,j\neq n}^{N}\left[\mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{G}_{jn}\mathbf{U}_j\right)\right] \\ + \sum_{i=1}^{M}\mathrm{Tr}\left(\mathbf{G}_{0n}\mathbf{W}_i\right) + \sigma_{ner}^2 \\ \geq g_{\xi_b}^{(q)}\left(\tau_n^b,\kappa_n^2\right),\end{aligned} \tag{22}$$

$$\mathrm{Tr}\left(\mathbf{H}_m\mathbf{W}_m\right) \geq g_{\xi_c}^{(q)}\left(\tau_{0m}^a,\kappa_{0m}^1\right) - \kappa_{0m}^1, \tag{23}$$

$$\begin{aligned}\sum_{j=1}^{N}\left[\mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{V}_j\right) + \mathrm{Tr}\left(\mathbf{H}_{joe}\mathbf{U}_j\right)\right] \\ + \sum_{i=1,i\neq m}^{M}\mathrm{Tr}(\mathbf{H}_e\mathbf{W}_i) + \mathrm{Tr}\left(\mathbf{H}_e\mathbf{Z}_0\right) + \sigma_e^2 \\ \geq g_{\xi_d}^{(q)}(\tau_{0m}^b,\kappa_{0m}^2),\end{aligned} \tag{24}$$

where $g_{\xi_\circ}^{(q)}(x,y)\left(\circ \in \{o1,o2,a,b,c,d\}\right)$ can be expressed as

$$g_{\xi_\circ}^{(q)}(x,y) = \frac{\xi_\circ^{(q-1)}}{2}x^2 + \frac{1}{2\xi_\circ^{(q-1)}}y^2, \tag{25}$$

$\xi_{o1}^{(q-1)}, \xi_{o2}^{(q-1)}, \xi_a^{(q-1)}, \xi_b^{(q-1)}, \xi_c^{(q-1)}$ and $\xi_d^{(q-1)}$ are the $(q-1)$-th iterative approximation of $\frac{r_n}{\xi}$, $\frac{r_0}{\xi}$, $\frac{\tau_n^a}{\kappa_n^1}$, $\frac{\tau_n^b}{\kappa_n^2}$, $\frac{\tau_{0m}^a}{\kappa_{0m}^1}$ and $\frac{\tau_{0m}^b}{\kappa_{0m}^2}$, respectively. So far, problem (16) can be reformulated as a generic convex program

$$\max_{\substack{\mathbf{V},\mathbf{U},\mathbf{W},\mathbf{Z}_0,\boldsymbol{r},\boldsymbol{l},\zeta,\\ \boldsymbol{\tau}^a,\boldsymbol{\tau}^b,\boldsymbol{\tau}_0^a,\boldsymbol{\tau}_0^b,\\ \boldsymbol{\kappa}^1,\boldsymbol{\kappa}^2,\boldsymbol{\kappa}_0^1,\boldsymbol{\kappa}_0^2}} \Big\{\zeta \,|(15b),(15d),(16c),(16e),(16g),\ (16i)\sim(16l),(19)\sim(24)\Big\}. \tag{26}$$

Nevertheless, the convex constrains in logarithmic forms, i.e., (15b), (15d), (19) and (20), bring about high computational complexity, which may deteriorate the solving efficiency severely. To facilitate tractability, we convert them into second order cone (SOC) representable constraints while preserving the convexity. Take (15b) as an example, which is equivalent to

$$\tau_n^a\log_2\left(\tau_n^a\right) \geq \tau_n^a l_n, \forall n \in [1,N], \tag{27}$$

and the left hand side can be proved to be convex by its second order derivative, thus we have the following inequality

$$\begin{aligned}&\tau_n^a\log_2\left(\tau_n^a\right) \\ &\geq \tau_n^{a(q-1)}\log_2\left(\tau_n^{a(q-1)}\right) + \left(\tau_n^a - \tau_n^{a(q-1)}\right)\left[1+\log_2\left(\tau_n^{a(q-1)}\right)\right] \\ &= \tau_n^a\left[1+\log_2\left(\tau_n^{a(q-1)}\right)\right] - \tau_n^{a(q-1)},\end{aligned} \tag{28}$$

where $\tau_n^{a(q-1)}$ is the $(q-1)$-th approximation of τ_n^a. Hence we replace (15b) in the q-th iteration with

$$\tau_n^a\left[1+\log_2\left(\tau_n^{a(q-1)}\right)\right] - \tau_n^{a(q-1)} \geq \tau_n^a l_n, \tag{29}$$

which can be written as an SOC representable form as

$$\begin{aligned}&\tau_n^a - l_n + 1 + \log_2\left(\tau_n^{a(q-1)}\right) \\ &\geq \left\|\left[\tau_n^{a(q-1)} + l_n - 1 - \log_2\left(\tau_n^{a(q-1)}\right) \quad 2\sqrt{\tau_n^{a(q-1)}}\right]\right\|_2.\end{aligned} \tag{30}$$

Likewise, (19) can be transformed into

$$\begin{aligned}&\tau_n^b - g_{\xi_{o1}}^{(q)}(r_n,\zeta) + l_n + 1 + \log_2\left(\tau_n^{b(q-1)}\right) \\ &\geq \left\|\left[\tau_n^{b(q-1)} + g_{\xi_{o1}}^{(q)}(r_n,\zeta) - l_n \right.\right. \\ &\qquad \left.\left. - 1 - \log_2\left(\tau_n^{b(q-1)}\right) \quad 2\sqrt{\tau_n^{b(q-1)}}\right]\right\|_2.\end{aligned} \tag{31}$$

Afterwards, to deal with (15d) and (20), which are slightly different, we need to introduce several new slack variables $\boldsymbol{l}_{0\alpha} = \{l_{0\alpha}^m\}_{m\in[1,M]}, \boldsymbol{l}_{0\beta} = \{l_{0\beta}^m\}_{m\in[1,M]}$ and rewrite them as

$$\sum_{m=1}^{M} l_{0\alpha}^m \geq l_0, \tag{32}$$

$$\sum_{m=1}^{M} l_{0\beta}^m \geq g_{\xi_{o2}}^{(q)}(r_0,\zeta) - l_0, \tag{33}$$

$$\begin{aligned}&\tau_{0m}^a - l_{0\alpha}^m + 1 + \log_2\left(\tau_{0m}^{a(q-1)}\right) \\ &\geq \left\|\left[\tau_{0m}^{a\ (q-1)} + l_{0\alpha}^m - 1 - \log_2\left(\tau_{0m}^{a\ (q-1)}\right) \quad 2\sqrt{\tau_{0m}^{a\ (q-1)}}\right]\right\|_2,\end{aligned} \tag{34}$$

$$\begin{aligned}&\tau_{0m}^b - l_{0\beta}^m + 1 + \log_2\left(\tau_{0m}^{b(q-1)}\right) \\ &\geq \left\|\left[\tau_{0m}^{b\ (q-1)} + l_{0\beta}^m - 1 - \log_2\left(\tau_{0m}^{b\ (q-1)}\right) \quad 2\sqrt{\tau_{0m}^{b\ (q-1)}}\right]\right\|_2.\end{aligned} \tag{35}$$

TABLE 1. Computational complexity analysis.

Methods (processor)	Computation Complexity Order
Centralized (central calculating center)	$\sqrt{(M+1)K_m+2NK_f+8M+10N+3}\cdot n\cdot\Big\{n^3+[(M+1)K_m^2+2NK_f^2+4M+6N+3]n^2 +[(M+1)K_m^3+2NK_f^3+4M+6N+3]n\Big\}$, where $n=\mathcal{O}[(M+1)K_m^2+2NK_f^2+6M+6N+3]$
Decentralized (MBS)	$\sqrt{(M+1)K_m+8M+3}\cdot n\cdot\Big\{n^3+[(M+1)K_m^2+4M+3]n^2+[(M+1)K_m^3+4M+3]n\Big\}$, where $n=\mathcal{O}[(M+1)K_m^2+6M+3]$
Decentralized (FBS)	$\sqrt{2K_f+10}\cdot n\cdot[n^3+(4K_f+6)n^2+(2K_f^3+10)n]$, where $n=\mathcal{O}(2K_f^2+9)$

Finally, the problem can be expressed as

$$\max_{\substack{\mathbf{V},\mathbf{U},\mathbf{W},\mathbf{Z}_0,\boldsymbol{r},\boldsymbol{l},\zeta,\\ \boldsymbol{\tau}^a,\boldsymbol{\tau}^b,\boldsymbol{\tau}_0^a,\boldsymbol{\tau}_0^b,\\ \boldsymbol{\kappa}^1,\boldsymbol{\kappa}^2,\boldsymbol{\kappa}_0^1,\boldsymbol{\kappa}_0^2,\boldsymbol{l}_{0\alpha},\boldsymbol{l}_{0\beta}}} \Big\{\zeta\,|(16c),(16e),(16g),(16i)\sim(16l), (21)\sim(24),(30)\sim(35)\Big\}, \quad (36)$$

which can be solved efficiently via state-of-the-art conic solvers, such as CVX [27].

Notice that the rank-one constraints have been relaxed in problem (16), but generally, the optimal solution to the rank-relaxed problem may be not rank-one, hence the result of the rank-relaxed problem serves as a performance upper bound for the original problem. Only when the rank of the optimal solution to problem (36) is one, problem (36) and problem (13) have the same optimal result and solution. Fortunately, we have the following proposition.

Proposition 1: If problem (36) is feasible and $\Big\{\hat{\mathbf{W}}_1,\ldots,\hat{\mathbf{W}}_M,\hat{\mathbf{Z}}_0,\hat{\mathbf{V}}_1,\ldots,\hat{\mathbf{V}}_N,\hat{\mathbf{U}}_1,\ldots\hat{\mathbf{U}}_N\Big\}$ are the optimal solutions, they must be of rank-one.

Proof: Please see Appendix.

Complexity Analysis: According to [28], we know the complexity of solving an optimization problem through the method of interior point is $\mathcal{O}\big(\sqrt{\theta^*}L\ln\big(\frac{1}{\varepsilon}\big)\big)$, which is composed of two parts, namely, iteration complexity $\mathcal{O}\big(\sqrt{\theta^*}\ln\big(\frac{1}{\varepsilon}\big)\big)$ and per-iteration computational cost $\mathcal{O}(L)$. To be specific, θ^* is the barrier parameter determined by the number, the size of positive semi-definite definite (PSD) constraints and the number of SOC constraints, while $\ln\big(\frac{1}{\varepsilon}\big)$ is associated with the required precision of solutions; $\mathcal{O}(L)$ mainly stems from the number of optimization variables, the number of PSD constraints and their sizes, the SOC constraints and their sizes.

In our proposed iterative algorithm, the design problem has $(M+1)K_m^2+2NK_f^2$ design variables and $6N+6M+3$ slack variables, $M+1$ PSD constraints of size K_m, $2N$ PSD constraints of size K_f and $4M+6N+3$ PSD constraints of size 1, as well as $2N+2M$ SOC constraints of size 2. Ignoring the $\ln\big(\frac{1}{\varepsilon}\big)$ that is determined by the solution precision, we can obtain the computational complexity order of the iterative algorithm in Table 1.

IV. DECENTRALIZED DESIGN WITH ADMM

In the previous section, we obtain the optimal solution to the cross-tier MCBF design by applying SDR technique. However, the proposed method is executed in a centralized manner, which indicates the requirement of gathering all the CSI of coordinated BSs at a calculating center when joint designing the transmit beams and AN vectors, hence the centralized approach would doubtless incur heavy overhead to the network especially when the number of coordinated BSs is large. Furthermore, the scenario without an available central processing station should also be taken into account carefully.[5] Therefore, to find a decentralized solving method where each transmitter only needs to deal with its local CSI is of significant necessity.

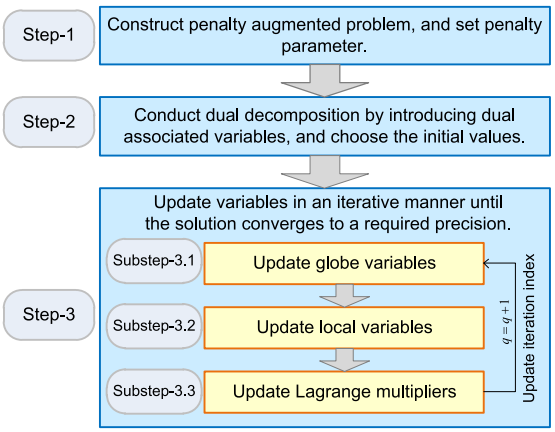

FIGURE 4. Key steps of the ADMM approach.

ADMM is a decentralized approach based on the combination of dual decomposition method and augmented Lagrangian method. In ADMM, there are mainly three key steps, as illustrated in Fig. 4. The first ingredient is to construct the penalty augmented problem while setting a penalty parameter to guarantee the convergence; then, the decomposition is conducted with dual variables introduced so that the problem can be solved in parallel at different BSs; finally,

[5]We provide two practical cases where no central processing station is available: a) there is a low trust degree among the BSs, hence they are not willing to send their local CSI to any other BS; b) the processing capacity of each BS is at a comparable level, while handing the CSI of all the BSs is unbearable for any one of them.

three kinds of variables are updated in an iterative manner until the solution converges to an expected precision.

To begin with, we first define some slack variables as

$$P_n = \mathrm{Tr}(\mathbf{V}_n) + \mathrm{Tr}(\mathbf{U}_n), \quad \forall n \in [1, N], \tag{37a}$$

$$P_0 = \sum_{m=1}^{M} \mathrm{Tr}(\mathbf{W}_m) + \mathrm{Tr}(\mathbf{Z}_0), \tag{37b}$$

$$X_{a,b}^{\mathrm{IR}} = \tilde{X}_{b,a}^{\mathrm{IR}} = \chi_{ab\mathrm{IR}}, \quad \forall a \in [0, N],\ \forall b \in [1, N],\ a \neq b, \tag{37c}$$

$$X_{a,b}^{\mathrm{ER}} = \tilde{X}_{b,a}^{\mathrm{ER}} = \chi_{ab\mathrm{ER}}, \quad \forall a \in [0, N],\ \forall b \in [1, N],\ a \neq b, \tag{37d}$$

$$X_{c,0}^{m} = \tilde{X}_{0,c}^{m} = \chi_{com}, \quad \forall c \in [1, N], \tag{37e}$$

$$X_{c,0}^{e} = \tilde{X}_{0,c}^{e} = \chi_{coe}, \quad \forall c \in [1, N], \tag{37f}$$

where all the variables in (37c)∼(37f) denote the inter-cell interference (ICI) brought by a BS to a user accessing another BS. $X_{a,b}^{\ltimes}$ / $\tilde{X}_{b,a}^{\ltimes}$ ($\ltimes \in \{\mathrm{IR}, \mathrm{ER}\}$) or $X_{c,0}^{\rtimes}$ / $\tilde{X}_{0,c}^{\rtimes}$ ($\rtimes \in \{m, e\}$) forms a pair of variables that express the same ICI, while the element without a hat (i.e., X) in the pair denotes the local variable stored at the interfering BS and the one with a hat (i.e., $\tilde{X}$) denotes that stored at the interfered BS, respectively. Specifically, $X_{a,b}^{\mathrm{IR}}$ and $X_{a,b}^{\mathrm{ER}}$ are the local variables of FBS_a ($a \geq 1$) or MBS ($a = 0$), while $\tilde{X}_{b,a}^{\mathrm{IR}}$ and $\tilde{X}_{b,a}^{\mathrm{ER}}$ are those of FBS_b. Moreover, $X_{c,0}^{m}$ and $X_{c,0}^{e}$ can only be stored at FBS_c ($c \geq 1$) while $\tilde{X}_{0,c}^{m}$ and $\tilde{X}_{0,c}^{e}$ are at the MBS. Notice that the slack variables $\chi_{ab\mathrm{IR}}$, $\chi_{ab\mathrm{ER}}$, χ_{com} and χ_{coe} can be deemed as the global copies of former ones, which are introduced to ensure the copies of the same ICI stored at different BSs are indeed equal to each other. Further, to facilitate tractability, we weaken the differences of heterogeneous BSs by denoting the MBS as BS_0 and FBS_n as BS_n, and rewrite problem (36) in a decomposable form as[6]

$$\min_{\substack{\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{z}_0,\boldsymbol{r},\boldsymbol{l},\boldsymbol{l}_{0\alpha},\boldsymbol{l}_{0\beta},\boldsymbol{\zeta}\\ \boldsymbol{\tau}^a,\boldsymbol{\tau}^b,\boldsymbol{\tau}_0^a,\boldsymbol{\tau}_0^b,\boldsymbol{\kappa}^1,\boldsymbol{\kappa}^2,\boldsymbol{\kappa}_0^1,\boldsymbol{\kappa}_0^2,\\ \{\chi_{nom}\},\{\chi_{noe}\},\{\chi_{\bar{n}i\mathrm{IR}}\},\{\chi_{\bar{n}i\mathrm{ER}}\}\\ \{X_{n,0}^m\},\{\tilde{X}_{0,n}^m\},\{X_{n,0}^e\},\{\tilde{X}_{0,n}^e\},\\ \{X_{n,i}^{\mathrm{IR}}\},\{\tilde{X}_{n,i}^{\mathrm{IR}}\},\{X_{n,i}^{\mathrm{ER}}\},\{\tilde{X}_{n,i}^{\mathrm{ER}}\}}} -\sum_{\bar{n}=0}^{N} \zeta_{\bar{n}} \tag{38a}$$

$$\text{s.t. } X_{\bar{n},i}^{\mathrm{IR}} \geq \mathrm{Tr}(\mathbf{H}_{\bar{n}i} P_{\bar{n}}), \quad \forall \bar{n} \in [0, N],\ \forall i \in [1, N],\ i \neq \bar{n}, \tag{38b}$$

$$X_{\bar{n},i}^{\mathrm{ER}} \leq \mathrm{Tr}(\mathbf{G}_{\bar{n}i} P_{\bar{n}}), \quad \forall \bar{n} \in [0, N],\ \forall i \in [1, N],\ i \neq \bar{n} \tag{38c}$$

$$X_{n,0}^{m} \geq \mathrm{Tr}(\mathbf{H}_{nom} P_n), \quad \forall n \in [1, N],\ \forall m \in [1, M], \tag{38d}$$

$$X_{n,0}^{e} \leq \mathrm{Tr}(\mathbf{H}_{noe} P_n), \quad \forall n \in [1, N], \tag{38e}$$

$$\mathrm{Tr}(\mathbf{H}_{nn}\mathbf{U}_n) + \sum_{i=0, i\neq n}^{N} \tilde{X}_{n,i}^{\mathrm{IR}} + \sigma_{nir}^2 \leq \kappa_n^1, \tag{38f}$$

$$\mathrm{Tr}(\mathbf{G}_{nn}\mathbf{U}_n) + \sum_{i=0, i\neq n}^{N} \tilde{X}_{n,i}^{\mathrm{ER}} + \sigma_{ner}^2 \geq g_{\xi_b}^{(q)}\left(\tau_n^b, \kappa_n^2\right), \tag{38g}$$

$$\mathrm{Tr}(\mathbf{G}_{nn}\mathbf{U}_n) + \mathrm{Tr}(\mathbf{G}_{nn}\mathbf{V}_n) + \sum_{i=0, i\neq n}^{N} \tilde{X}_{n,i}^{\mathrm{ER}} + \sigma_{ner}^2 \leq \kappa_n^2, \tag{38h}$$

$$\sum_{i=1, i\neq m}^{M} \mathrm{Tr}(\mathbf{H}_m\mathbf{W}_i) + \mathrm{Tr}(\mathbf{H}_m\mathbf{Z}_0) + \sum_{j=1}^{N} \tilde{X}_{0,j}^{m} + \sigma_m^2 \leq \kappa_{0m}^1, \tag{38i}$$

$$\sum_{i=1, i\neq m}^{M} \mathrm{Tr}(\mathbf{H}_e\mathbf{W}_i) + \mathrm{Tr}(\mathbf{H}_e\mathbf{Z}_0) + \sum_{j=1}^{N} \tilde{X}_{0,j}^{e} + \sigma_e^2 \geq g_{\xi_d}^{(q)}\left(\tau_{0m}^b, \kappa_{0m}^2\right), \tag{38j}$$

$$\sum_{i=1}^{M} \mathrm{Tr}(\mathbf{H}_e\mathbf{W}_i) + \mathrm{Tr}(\mathbf{H}_e\mathbf{Z}_0) + \sum_{j=1}^{N} \tilde{X}_{0,j}^{e} + \sigma_e^2 \leq \kappa_{0m}^2, \tag{38k}$$

$$\zeta_{\bar{n}} = \zeta, \quad \forall \bar{n} \in [0, N], \tag{38l}$$

$$\text{constraints}(16j),\ (16k),\ (21),\ (23),\ (30)\sim(35),\ (37). \tag{38m}$$

Next we include all the constraints that can be handled locally at $\mathrm{BS}_{\bar{n}}$ in the set $\mathbf{S}_{\bar{n}}^q$ which is defined as

$$\mathbf{S}_{\bar{n}}^q \triangleq \left\{\mathbf{s}_{\bar{n}} | (16j),\ (16k),\ (21),\ (23),\ (30)\sim(35),\ (38\mathrm{b})\sim(38\mathrm{k})\right\}, \tag{39}$$

where $\mathbf{s}_n \triangleq \left[\zeta_n, \mathbf{v}_n, \mathbf{u}_n, r_n, l_n, \kappa_n^1, \kappa_n^2, \tau_n^a, \tau_n^b, \mathbf{A}_n^T\right], n \in [1, N]$, simply stacks all the local variables that can be locally handled at BS_n, and specially for $\mathbf{s}_0 \triangleq \left[\zeta_0, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_M, \mathbf{z}_0, r_0, l_0, \boldsymbol{\kappa}_0^1, \boldsymbol{\kappa}_0^2, \boldsymbol{\tau}_0^a, \boldsymbol{\tau}_0^b, \boldsymbol{l}_{0\alpha}, \boldsymbol{l}_{0\beta}, \mathbf{A}_0^T\right]$. In what follows, we define

$$\begin{aligned}\mathbf{A}_n \triangleq \big[&X_{n,0}^1, X_{n,0}^2, \ldots, X_{n,0}^M, X_{n,0}^e, X_{n,1}^{\mathrm{IR}}, X_{n,1}^{\mathrm{ER}}, \ldots, X_{n,n-1}^{\mathrm{ER}},\\ &X_{n,n+1}^{\mathrm{IR}}, \ldots, X_{n,N}^{\mathrm{IR}}, X_{n,N}^{\mathrm{ER}}, \tilde{X}_{n,0}^{\mathrm{IR}}, \tilde{X}_{n,0}^{\mathrm{ER}}, \ldots, \tilde{X}_{n,n-1}^{\mathrm{ER}},\\ &\tilde{X}_{n,n+1}^{\mathrm{IR}}, \ldots, \tilde{X}_{n,N}^{\mathrm{IR}}, \tilde{X}_{n,N}^{\mathrm{ER}}\big]^T, \quad n \in [1, N],\end{aligned} \tag{40}$$

which only includes the ICIs relevant to BS_n, and in particular

$$\begin{aligned}\mathbf{A}_0 \triangleq \big[&X_{0,1}^{\mathrm{IR}}, X_{0,1}^{\mathrm{ER}}, \ldots, X_{0,N}^{\mathrm{IR}}, X_{0,N}^{\mathrm{ER}}, \tilde{X}_{0,1}^1, \ldots, \tilde{X}_{0,1}^M, \tilde{X}_{0,1}^e, \ldots,\\ &\tilde{X}_{0,N}^1, \ldots, \tilde{X}_{0,N}^M, \tilde{X}_{0,N}^e\big]^T.\end{aligned} \tag{41}$$

Similarly we denote

$$\begin{aligned}\mathbf{B}_n \triangleq \big[&\chi_{n01}, \chi_{n02}, \ldots, \chi_{n0M}, \chi_{n0e}, \chi_{n1\mathrm{IR}}, \chi_{n1\mathrm{ER}}, \ldots,\\ &\chi_{n(n-1)\mathrm{ER}}, \chi_{n(n+1)\mathrm{IR}}, \ldots, \chi_{nN\mathrm{IR}}, \chi_{nN\mathrm{ER}}, \chi_{0n\mathrm{IR}},\\ &\chi_{0n\mathrm{ER}}, \ldots, \chi_{(n-1)n\mathrm{IR}}, \chi_{(n-1)n\mathrm{ER}}, \ldots, \chi_{Nn\mathrm{IR}}, \chi_{Nn\mathrm{ER}}\big]^T,\\ &n \in [1, N]\end{aligned} \tag{42}$$

[6]Though we include the MBS and the FBSs in the same set, some associated variables are still in different forms. In this regard, we use two different subscripts (i.e., $_n$ and $_{\bar{n}}$) to maintain the difference. When using $_{\bar{n}}(0 \leq \bar{n} \leq N)$ we refer to the set including all the BSs, and when $_n(1 \leq n \leq N)$ we only refer to the FBSs with MBS (BS_0) excluded.

and

$$\mathbf{B}_0 \triangleq \left[\chi_{01\mathrm{IR}}, \chi_{01\mathrm{ER}}, \ldots, \chi_{0N\mathrm{IR}}, \chi_{0N\mathrm{ER}}, \chi_{1o1}, \chi_{1o2}, \ldots, \chi_{1o\mathrm{M}}, \chi_{1o\mathrm{e}}, \chi_{2o1}, \ldots, \chi_{No1}, \ldots, \chi_{No\mathrm{e}}\right]^T \quad (43)$$

by sorting the corresponding global versions of ICI in $\mathbf{A}_{\bar{n}}$.

From the definitions above, (38) can be equivalently rewritten in a more compact form as

$$\min_{\mathbf{S},\mathbf{\Omega}} -\sum_{\bar{n}=0}^{N} \zeta_{\bar{n}} \quad (44a)$$

$$\text{s.t. } \mathbf{s}_{\bar{n}} \in \mathbf{S}_{\bar{n}}^q, \quad \forall \bar{n} \in [0, N], \quad (44b)$$

$$\mathbf{A}_{\bar{n}} = \mathbf{B}_{\bar{n}}, \quad \forall \bar{n} \in [0, N], \quad (44c)$$

$$\zeta_{\bar{n}} = \zeta, \quad \forall \bar{n} \in [0, N], \quad (44d)$$

where $\mathbf{S} \triangleq [\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_N]$ and $\mathbf{\Omega} = \left[\zeta, \mathbf{B}_0^T, \mathbf{B}_1^T, \ldots, \mathbf{B}_N^T\right]$.

Now, we are in a position to describe the ADMM to solve (36) in a decentralized manner. First, we construct the penalty augmented problem of (44) as

$$\min_{\mathbf{S},\mathbf{\Omega}} \sum_{\bar{n}=0}^{N} \left[-\zeta_{\bar{n}} + \frac{c}{2}(\zeta_{\bar{n}} - \zeta)^2 + \frac{c}{2}\left\|\mathbf{A}_{\bar{n}} - \mathbf{B}_{\bar{n}}\right\|_2^2\right] \quad (45a)$$

$$\text{s.t. constraints (44b)}, \quad (45b)$$

where $c > 0$ is the penalty parameter. Note that the optimal result and solution are not changed, and the added terms $\frac{c}{2}(\zeta_n - \zeta)^2$ and $\frac{c}{2}\left\|\mathbf{A}_n - \mathbf{B}_n\right\|_2^2$ are required to ensure the convergence of ADMM algorithm.

Afterwards, we carry out the dual decomposition, which enables each BS to deal with its local CSI in a coordinated design, and the penalty Lagrangian function can be expressed as

$$\begin{aligned}\mathcal{L}^q(\mathbf{\Omega}, \mathbf{S}, \mathbf{\Theta}, \mathbf{\Upsilon}) = \sum_{\bar{n}=0}^{N} \Big[&-\zeta_{\bar{n}} + \frac{c}{2}(\zeta_{\bar{n}} - \zeta)^2 \\ &+ \frac{c}{2}\left\|\mathbf{A}_{\bar{n}} - \mathbf{B}_{\bar{n}}\right\|_2^2 + \Theta_{\bar{n}}(\zeta_{\bar{n}} - \zeta) \\ &+ \Upsilon_{\bar{n}}\left\|\mathbf{A}_{\bar{n}} - \mathbf{B}_{\bar{n}}\right\|_2^2\Big], \end{aligned} \quad (46)$$

where $\Theta_{\bar{n}}$ and $\Upsilon_{\bar{n}}$ are the associated Lagrange multipliers, and we define $\mathbf{\Theta} \triangleq [\Theta_0, \Theta_1, \ldots, \Theta_N]$, $\mathbf{\Upsilon} \triangleq [\Upsilon_0, \Upsilon_1, \ldots, \Upsilon_N]$.

We now come to the central part of ADMM, which is to update the global variables (i.e., $\mathbf{\Omega}$), the local variables (i.e., $\mathbf{S}$), and the Lagrange multipliers (i.e., $\mathbf{\Theta}$ and $\mathbf{\Upsilon}$) by using the Gauss-Seidel method. Succinctly speaking, at each iteration, all BSs agree a common knowledge of associated interference variables and the consensus SEE by updating global variables. Then, each BS independently solves its own sub-problem and updates its Lagrange multipliers to drive the corresponding variables into equality. The detailed updating procedures are shown as follows.

A. UPDATE OF GLOBAL VARIABLES

It is the update of global variables that starts the iteration. When updating the global variables, the local variables are fixed, and the updated values can be acquired by solving the convex problem $\min_{\mathbf{\Omega}} \mathcal{L}^q\left(\mathbf{\Omega}, \mathbf{S}^t, \mathbf{\Theta}^t, \mathbf{\Upsilon}^t\right)$, where t denotes the ADMM iteration counter. Meanwhile, we notice that $\mathcal{L}^q\left(\mathbf{\Omega}, \mathbf{S}^t, \mathbf{\Theta}^t, \mathbf{\Upsilon}^t\right)$ is separable in ζ and $\{\mathbf{B}_n\}$, and thus the minimization of $\mathbf{\Omega}$ can be further decomposed into several independent subproblems, which can be expressed as

$$\zeta^{t+1} = \arg\min_{\zeta} \sum_{\bar{n}=0}^{N}\left[-\Theta_{\bar{n}}^t\zeta + \frac{c}{2}\left(\zeta_{\bar{n}}^t - \zeta\right)^2\right] \quad (47)$$

and

$$\mathbf{B}_{\bar{n}}^{t+1} = \arg\min_{\mathbf{B}_{\bar{n}}} -\left(\mathbf{\Upsilon}_{\bar{n}}^t\right)^T\mathbf{B}_{\bar{n}} + \frac{c}{2}\left\|\mathbf{A}_{\bar{n}}^t - \mathbf{B}_{\bar{n}}\right\|_2^2, \forall \bar{n} \in [0, N], \quad (48)$$

where $\zeta_{\bar{n}}^t$ and $\mathbf{A}_{\bar{n}}^t$ are obtained by solving (53) in the t-th iteration and exchanging corresponding variables among the coordinated BSs. According to [29], the closed-form solutions to (47) and (48) can be expressed as

$$\zeta^{t+1} = \frac{1}{N+1}\sum_{\bar{n}=0}^{N}\left(\zeta_{\bar{n}}^t + \frac{\Theta_{\bar{n}}^t}{c}\right) \quad (49)$$

and

$$\chi_{ni*}^{t+1} = \frac{1}{2}\left(X_{n,i}^{*\,t} + \tilde{X}_{i,n}^{*\,t}\right) + \frac{1}{2c}\left(\Upsilon_{n,i}^{*\,t} + \tilde{\Upsilon}_{i,n}^{*\,t}\right), \quad n \in [1, N], * \in \begin{cases}[1, M], & i = 0\\ \{\mathrm{IR}, \mathrm{ER}\}, & i \in [1, N],\ i \neq n,\end{cases} \quad (50a)$$

$$\chi_{0i*}^{t+1} = \frac{1}{2}\left(X_{0,i}^{*\,t} + \tilde{X}_{i,0}^{*\,t}\right) + \frac{1}{2c}\left(\Upsilon_{0,i}^{*\,t} + \tilde{\Upsilon}_{i,0}^{*\,t}\right), \quad i \in [1, N], \quad i \neq n,\ * \in \{\mathrm{IR}, \mathrm{ER}\}, \quad (50b)$$

or

$$\chi_{i\bar{n}*}^{t+1} = \frac{1}{2}\left(X_{i,\bar{n}}^{*\,t} + \tilde{X}_{\bar{n},i}^{*\,t}\right) + \frac{1}{2c}\left(\Upsilon_{i,\bar{n}}^{*\,t} + \tilde{\Upsilon}_{\bar{n},i}^{*\,t}\right), \quad i \in [0, N], * \in \begin{cases}[1, M], & \bar{n} = 0\\ \{\mathrm{IR}, \mathrm{ER}\}, & \bar{n} \in [1, N],\ i \neq \bar{n},\end{cases} \quad (51)$$

where $\left\{\Upsilon_{\bar{n},i}^*, \tilde{\Upsilon}_{\bar{n},i}^*\right\}$ and $\left\{\tilde{\Upsilon}_{i,\bar{n}}^*, \Upsilon_{i,\bar{n}}^*\right\}$ are the dual variables associated with the primal variables $\left\{X_{\bar{n},i}^*, \tilde{X}_{\bar{n},i}^*\right\}$ of BS_n and $\left\{\tilde{X}_{i,\bar{n}}^*, X_{i,\bar{n}}^*\right\}$ of $\mathrm{BS}_i (i \neq n)$, receptively. In other words, each BS can run an average consensus algorithm [30], [31] to acquire ζ^{t+1} in (49) and collect $\left\{\tilde{X}_{i,\bar{n}}^{*\,t}, X_{i,\bar{n}}^{*\,t}\right\}$, $\left\{\tilde{\Upsilon}_{i,\bar{n}}^*, \Upsilon_{i,\bar{n}}^*\right\}$ from other BSs to finally obtain $\chi_{\bar{n}i*}^{t+1}$ and $\chi_{i\bar{n}*}^{t+1}$.

B. UPDATE OF LOCAL VARIABLES

Afterwards, the following convex problem is solved to update the set of local variables $\mathbf{S}$

$$\mathbf{S}^{t+1} = \arg\min_{\mathbf{s}_{\bar{n}} \in \mathbf{S}_{\bar{n}}^q, \forall \bar{n} \in [0,N]} \mathcal{L}^q\left(\mathbf{\Omega}^{t+1}, \mathbf{S}, \mathbf{\Theta}^t, \mathbf{\Upsilon}^t\right). \quad (52)$$

Considering the key feature of the ADMM, which makes it a distributed approach, the augmented Lagrangian function

is decomposable in $\mathbf{s}_{\bar{n}}$, hence problem (52) can be solved in parallel at each coordinated BS. That is, $\mathrm{BS}_{\bar{n}}$ only needs to handle its local CSI to solve the following subproblem

$$\begin{aligned}\mathbf{s}_{\bar{n}}^{t+1} = & \underset{\mathbf{s}_{\bar{n}}\in\mathbf{S}_{\bar{n}}^{q},\forall\bar{n}\in[0,N]}{\arg\min} \quad -\zeta_{\bar{n}} + \Theta_{\bar{n}}^{t}\left(\zeta_{\bar{n}} - \zeta^{t+1}\right) \\ & + \left(\boldsymbol{\Upsilon}_{\bar{n}}^{t}\right)^{T}\left(\mathbf{A}_{\bar{n}} - \mathbf{B}_{\bar{n}}^{t+1}\right) + \frac{c}{2} \\ & \times\left[\left(\zeta_{\bar{n}} - \zeta^{t+1}\right)^{2} + \left\|\mathbf{A}_{\bar{n}} - \mathbf{B}_{\bar{n}}^{t+1}\right\|_{2}^{2}\right], \end{aligned} \tag{53}$$

where ζ^{t+1} and $\mathbf{B}_{\bar{n}}^{t+1}$ are acquired by updating global variables in the previous step. Since all the constraints in (39) are convex so that each subproblem can be solved with the aid of existing convex solvers.

C. UPDATE OF LAGRANGE MULTIPLIERS

Finally, the Lagrange multipliers are updated as

$$\Theta_{\bar{n}}^{t+1} = \Theta_{\bar{n}}^{t} + c\left(\zeta_{\bar{n}}^{t+1} - \zeta^{t+1}\right), \quad \forall\bar{n}\in[0,N], \tag{54}$$

and

$$\boldsymbol{\Upsilon}_{\bar{n}}^{t+1} = \boldsymbol{\Upsilon}_{\bar{n}}^{t} + c\left(\mathbf{A}_{\bar{n}}^{t+1} - \mathbf{B}_{\bar{n}}^{t+1}\right), \quad \forall\bar{n}\in[0,N]. \tag{55}$$

We recall that after the previous two steps, the local and global variables are all available to the BSs in this step, hence there is no more information exchange required for updating the Lagrange multipliers, which in turn incurs no extra signaling overhead [29]. Furthermore, we define the primal residual as $\Xi = \sqrt{\sum_{\bar{n}=0}^{N} \| [\zeta_{\bar{n}}^{t}, (\mathbf{A}_{\bar{n}}^{t})^{T}]^{T} - [\zeta^{t}, (\mathbf{B}_{\bar{n}}^{t})^{T}]^{T} \|_{2}^{2}}$, which is utilized for the ADMM-loop break control. When the primal residual is inferior to the preset precision, e.g., $\Xi \leq 10^{-4}$, the ADMM algorithm is deemed to converge. Then, the SCA parameters can be updated and it comes to the $(q+1)$-th SCA iteration. In general, the ADMM-based beam design can be summarized as Algorithm 1 (i.e., a dual-loop optimization algorithm), where the outer loop is the SCA-based iterative procedure while the inner one is the ADMM-based iteration.

1) INFORMATION EXCHANGE

As shown in Algorithm 1, the information exchange mainly occurs in the steps 4 and 5. To be more specific, in step 4, each BS needs to exchange ξ_{n}^{q} and Θ_{n}^{q} with other N BSs; and in step 5, the MBS needs to send out $2N$ variables including $\tilde{X}_{i,n}^{*\,q}$ and $X_{i,n}^{*\,q}$ to the FBSs; while each FBS BS_{n} $(n \geq 1)$ is bound to send out 2 variables to each of the other $(N-1)$ FBSs and M variables to the MBS. In the current LTE systems, the information exchange process among BSs can be accomplished via the X2 interface. We also note that the exchanged variables are all real-valued numbers that measure corresponding interference power, which brings about an obvious reduction in signaling overhead compared to the complex CSI exchanges [29].

2) PER BS COMPLEXITY ANALYSIS

We notice that the complexity of algorithm 1 is mainly determined by step 6, during which each BS solves the problem (53). Thus, the computational cost can be obtained in the similar way as in the previous section. To be more specific, the subproblem at the MBS has $(M+1)K_{m}^{2}$ variables and $6M+3$ slack variables, $M+1$ PSD constraints of size K_{m}, $4M+3$ PSD constraints of size 1, as well as $2M$ SOC constraints of size 2; at each FBS, the problem has $2K_{f}^{2}$ variables and 9 slack variables, 2 PSD constraints of size K_{m}, 6 PSD constraints of size 1, as well as 2 SOC constraints of size 2. Then, the computational complexity orders at different BSs are presented in Table 1, and obviously the processing stress is averaged on multiple BSs instead of being concentrated on the calculating center.

Algorithm 1 ADMM-Based Dual-Loop Optimization Algorithm for SEE Maximization in MCBF Design

Initialization: Set $q := 0$, $t := 0$ and choose initial feasible values for $(\mathbf{v}^{q}, \mathbf{u}^{q}, \mathbf{w}^{q}, \mathbf{z}^{q}, \boldsymbol{r}^{q}, \boldsymbol{l}^{q}, \boldsymbol{\kappa}^{1,q}, \boldsymbol{\kappa}^{2,q}, \boldsymbol{\tau}^{a,q}, \boldsymbol{\tau}^{b,q}, \boldsymbol{l}_{0\alpha}^{q}, \boldsymbol{l}_{0\beta}^{q})$ and choose initial values for $\left\{\zeta_{\bar{n}}^{0}\right\}, \left\{\mathbf{A}_{\bar{n}}^{0}\right\}, \boldsymbol{\Theta}^{0}, \boldsymbol{\Upsilon}^{0}$.
Output: optimal value of $(\mathbf{v}^{*}, \mathbf{u}^{*}, \mathbf{w}^{*}, \mathbf{z}^{*})$

1: **repeat** {outer loop-SCA}
2: **while** $\Xi \leq 10^{-4}$ **do** {inner loop-ADMM}
3: **for** $\bar{n} \in [0, N]$ **do**
4: $\mathrm{BS}_{\bar{n}}$ updates ζ^{t+1} through an average consensus algorithm [30].
5: $\mathrm{BS}_{\bar{n}}$ updates $\chi_{\bar{n}i*}^{t+1}$ and $\chi_{i\bar{n}*}^{t+1}$ by (50) and (51).
6: $\mathrm{BS}_{\bar{n}}$ updates $\mathbf{s}_{\bar{n}}^{t+1}$ by (53)
7: $\mathrm{BS}_{\bar{n}}$ updates Lagrange multipliers $\Theta_{\bar{n}}^{t+1}$ and $\boldsymbol{\Upsilon}_{\bar{n}}^{t+1}$ by (54) and (55), respectively.
8: **end for**
9: $t := t + 1$
10: **end while**
11: Obtain optimal value of $(\hat{\mathbf{v}}, \hat{\mathbf{u}}, \hat{\mathbf{w}}, \hat{\mathbf{z}})$ in outer q-th loop
12: Update SCA parameters $(\mathbf{v}^{q+1}, \mathbf{u}^{q+1}, \mathbf{w}^{q+1}, \mathbf{z}^{q+1}, \boldsymbol{r}^{q+1}, \boldsymbol{l}^{q+1}, \boldsymbol{\kappa}^{1,q+1}, \boldsymbol{\kappa}^{2,q+1}, \boldsymbol{\tau}^{a,q+1}, \boldsymbol{\tau}^{b,q+1}, \boldsymbol{l}_{0\alpha}^{q+1}, \boldsymbol{l}_{0\beta}^{q+1}) = (\hat{\mathbf{v}}, \hat{\mathbf{u}}, \hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{l}}, \hat{\boldsymbol{\kappa}}^{1}, \hat{\boldsymbol{\kappa}}^{2}, \hat{\boldsymbol{\tau}}^{a}, \hat{\boldsymbol{\tau}}^{b}, \hat{\boldsymbol{l}}_{0\alpha}, \hat{\boldsymbol{l}}_{0\beta})$ and $(\left\{\zeta_{\bar{n}}^{0}\right\}, \left\{\mathbf{A}_{\bar{n}}^{0}\right\}, \boldsymbol{\Theta}^{0}, \boldsymbol{\Upsilon}^{0}) = (\left\{\hat{\zeta}_{\bar{n}}^{0}\right\}, \left\{\hat{\mathbf{A}}_{\bar{n}}^{0}\right\}, \hat{\boldsymbol{\Theta}}^{0}, \hat{\boldsymbol{\Upsilon}}^{0})$
13: $q := q + 1$, $t := 0$.
14: **until** SCA coverages.
15: **return** $(\mathbf{v}^{*}, \mathbf{u}^{*}, \mathbf{w}^{*}, \mathbf{z}^{*})$

V. SIMULATION RESULTS

In this section, we provide simulation results and discussions. We consider the simulation scenario where two MUs and an Eve are randomly distributed in a macrocell with a radius of 500m, while a couple of IoT users are scattered randomly in each femtocell with a radius of 150m. With the MBS located at the origin of coordinate system, the coordinates of FBSs are randomly generated within the coverage of macrocell, as shown in Fig. 5. In the simulations, the small-scale fading coefficients all follow Gaussian distributed. The path loss and shadowing are modeled as $\beta(\mathrm{dB}) = 38\log_{10}(d) + 34.5 + \mathcal{CN}(0, 8)$, where d denotes the distance from transmitter to receiver in the corresponding channel.

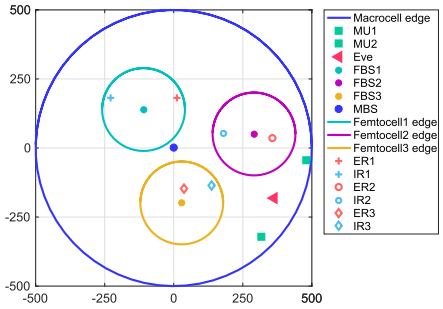

FIGURE 5. Illustration of simulation scenario.

TABLE 2. Simulation parameters.

Parameters	Values
Available antennas at MBS	8
Available antennas at FBS	5
Maximum transmit power at MBS	40dBm
Maximum transmit power at FBS	36dBm
Power loss in hardwares	33dBm
Power amplifier efficiency	0.7
Signal bandwidth	10kHz
Power spectral density of noise	-174dBm/Hz
Penalty parameter	0.05

Other general simulation parameters are listed in Table 2 and specific ones are given in the corresponding figures. All simulation results are averaged over 1000 randomly generated channel realizations.

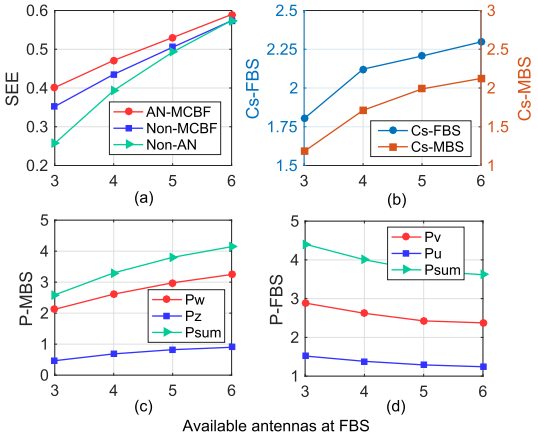

FIGURE 6. Secrecy performance versus the number of antennas configured at FBS.

Fig. 6 presents the secrecy performance versus the number of available antennas at each FBS. The SEE curves, the achievable secrecy rate of BSs, as well as the average power consumption at the MBS and FBSs are illustrated in sub-figures (a)∼(d), respectively. As shown in Fig. 6(a), the proposed method of AN-aided MCBF design outperforms the one without AN (hereinafter referred to as non-AN) and the one without joint design among neighbor BSs (hereinafter referred to as non-MCBF). We notice that as K_m increases, the SEE levels obtained via three methods are all promoted, whereas the performance gaps between AN-aided MCBF design and non-AN, non-MCBF methods tend to shrink, which can be contributed to the spatial degree of freedom brought by multi-antennas technique. When more antennas are available at each FBS, the resolution ratio of achievable space is concurrently promoted, and the diversity between legitimated channels (i.e., from FBSs to IRs) and wiretapping channels (i.e., from FBSs to ERs) can be guaranteed more easily with less AN injected. As a consequence, higher secrecy rate can be achieved with less power (of both beam and AN) consumed, as shown in Fig. 6(b) and Fig. 6(d), and the effect of AN tends to decrease.

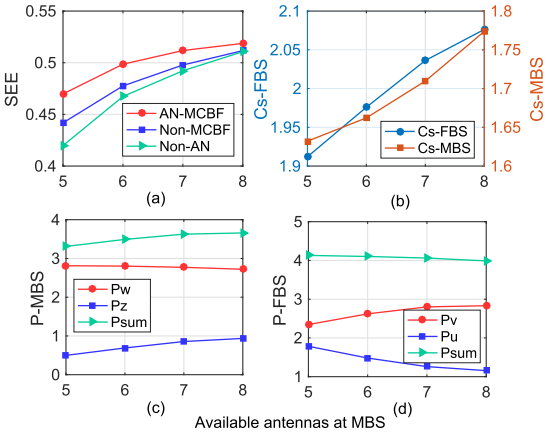

FIGURE 7. Secrecy performance versus the number of antennas configured at MBS.

Fig. 7 shows the impact of the number of available antennas at MBS on secrecy performance. With more antennas configured, the MBS could save power while providing comparative secrecy capacity, and the SEE in macrocell will be surely promoted. In order to improve the average SEE level, the MBS utilizes a portion of power saved by configuring more antennas to confuse as well as transfer energy to the ERs in femtocells. Therefore, each FBS can allocate more power for beam part and improve the secrecy rate as shown in Fig. 7(b) and Fig. 7(d), thus the average SEE is further promoted. However, by comparing Fig. 7(a) to Fig. 6(a), we discover that the enhancement brought by increasing K_m is not as obvious as increasing K_f. The reason lies that the power saved at the MBS needs to be distributed to three ERs, and the secrecy gain obtained at each femtocell is rather limited.

Fig. 8 depicts the secrecy performance versus the harvesting requirements of ERs. If more energy is required at each ER, more power needs to be allocated to AN and thus the SEE tends to decrease in femtocells. As mentioned in section II, the non-linear harvesting model is adopted, where the required RF power is bound to present exponential growth as the harvesting requirement increases linearly, hence the varying rates of secrecy rate and transmit power both rise up with the increment of harvesting requirement, as depicted in Fig. 8(b) and Fig. 8(d). Then we focus on the varying trend of power consumption and secrecy rate in the macrocell. Though SWIPT is not performed at the MBS and the

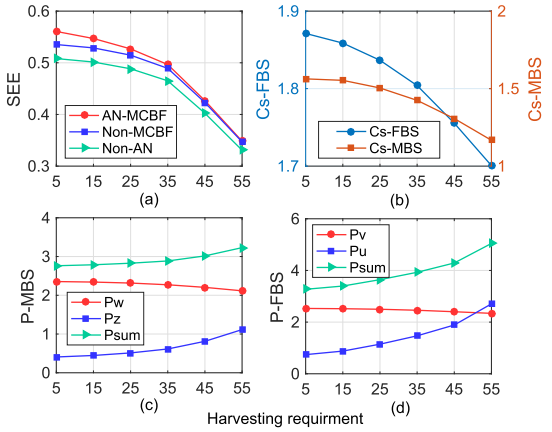

FIGURE 8. Secrecy performance versus the EH requirement.

increasing of harvesting requirement should not make any change, the SEE in the macrocell tends to decrease owning to the fairness requirement. In other words, when the average SEE in femtocells decreases, the average performance is further deteriorated, and there is no need for the macrocell to stay at a high SEE level. On the contrary, the MBS shares part of the task of power transfer, as illustrated in Fig. 8(c), which helps retard the decreasing of globe SEE.

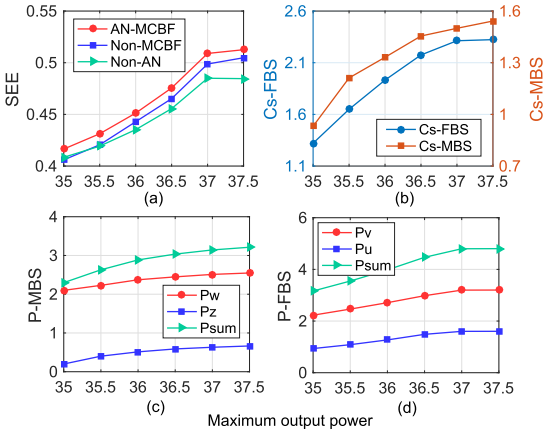

FIGURE 9. Secrecy performance versus the maximum transmit power at FBS.

Next, we pay attention to the effect of maximum output power at BSs on the SEE performance. Considering the power supply is usually sufficient at the MBS, we mainly focus on the FBSs in this part, as presented in Fig. 9. The validity of AN-aided MCBF design is firstly verified in Fig. 9(a), and the varying trend can be analyzed as follows. Owning to the fast fading of power along with distance, the transmit power for satisfying the harvesting requirement of ER is relatively high. If the maximum output power is poorly guaranteed at the FBS, the SEE will be limited to a low level. As the maximum transmit power grows, more available power can be allocated for improving secrecy rate, which can surely bring about substantial performance gain. In addition, we also notice that the curve of SEE tends to level off when the transmit power is relatively sufficient. This is because increasing transmit power linearly results in a logarithmic growth in secrecy rate, and when the secrecy rate gain can not catch up with the increment of transmit power, the SEE will not be promoted by increasing transmit power anymore, which we call the "power saturation."

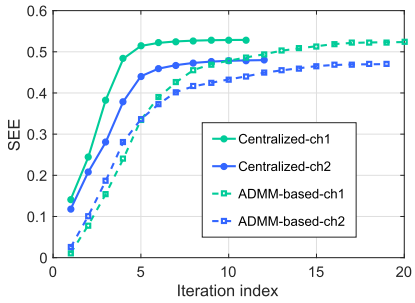

FIGURE 10. Optimization results versus the iteration index.

Fig. 10 illustrates the convergence performance of the two algorithms proposed in this paper in terms of SEE versus the iteration index. As can be observed, with the centralized SCA-based algorithm, the SEE converges to a stabilized value after 8 iterations. With the expected precision set to 10^{-4}, the stable SEE can be achieved within 12 iterations under different channel realizations (i.e., ch1 and ch2). As for ADMM, we should point out that the converging rate is much slower than the centralized algorithm, which is unsurprising and can be contributed to the distributed solving manner. Moreover, awaiting ADMM-based inner loop to converge in each SCA iteration (outer loop) also slows down the converging rate severely. But actually in first several outer loops, the results obtained by ADMM procedure are still rough estimations of the optimal solutions, hence there is no need to arrive at the results that completely converge, instead, the ADMM-loop can be forced to break by the inner loop counter (e.g., $t \geq 5$), which works well in improving the converging rate.

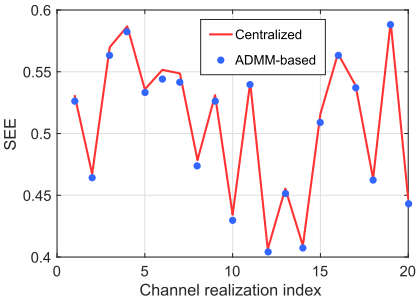

FIGURE 11. Optimization results on random channel realizations.

Fig. 11 provides the performance comparison between the SCA-based centralized approach and the ADMM-based decentralized algorithm in terms of the SEE over 20 random channel realizations. As illustrated, under most of the channel realizations, the ADMM-based algorithm can almost

converge to the optimal solution obtained by the SCA-based centralized design, which further implies the validity of distributed ADMM-based algorithm.

VI. CONCLUSION

This paper investigated the PLS in HetNets, where the femtocells performed SWIPT to power energy-constrained IoT users. To guard both the MUs and the IoT users, we considered an AN-aided cross-tier MCBF design. To secure the transmissions while achieving green communications, the SEE was deemed as the optimization target. On this basis, we formulated the problem of maximizing the minimum SEE to promote the average performance of the entire network. To handle the formulated nonconvex problem, we first applied SDR technique by maintaining the tightness and then introduced an SCA-based algorithm for complexity reduction. To further release the processing stress on the calculating center and the signaling overhead of the network, we proposed an ADMM-based decentralized algorithm that allows each BS to deal with its local CSI. Simulation results verified the performance of AN-aided MCBF design and the validity of decentralized algorithm.

APPENDIX

The Lagrange function of problem (16) can be written as

$$\begin{aligned}
&\mathcal{L}(\mathbf{V},\mathbf{U},\mathbf{W},\mathbf{Z}_0,\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\lambda_1,\lambda_2,\lambda_3)\\
&= -\sum_{n=1}^{N}\left\{\log_2\left(1+\frac{\mathbf{H}_{nn}\mathbf{V}_n}{\hat{I}_n^{IR}}\right)-\log_2\left(1+\frac{\mathbf{G}_{nn}\mathbf{V}_n}{\hat{I}_n^{ER}}\right)\right.\\
&\quad\left.-\xi\left\{\left[\mathrm{Tr}\left(\mathbf{V}_n\right)+\mathrm{Tr}\left(\mathbf{U}_n\right)\right]\frac{1}{\varpi}+P_\varepsilon\right\}\right\}\\
&\quad-\left\{\sum_{m=1}^{M}\left[\log_2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)-\log_2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\right]\right.\\
&\quad\left.-\xi\left\{\left[\sum_{m=1}^{M}\mathrm{Tr}\left(\mathbf{W}_m\right)+\mathrm{Tr}\left(\mathbf{Z}_0\right)\right]\frac{1}{\varpi}+P_\varepsilon\right\}\right\}\\
&\quad-\sum_{n=1}^{N}\mathbf{V}_n\mathbf{A}_n-\sum_{n=1}^{N}\mathbf{U}_n\mathbf{B}_n-\sum_{m=1}^{M}\mathbf{W}_m\mathbf{C}_m-\mathbf{ZD}\\
&\quad-\lambda_1\sum_{n=1}^{N}\left\{\sum_{i=1}^{N}\left[\mathrm{Tr}\left(\mathbf{G}_{in}\mathbf{V}_i\right)+\mathrm{Tr}\left(\mathbf{G}_{in}\mathbf{U}_i\right)\right]\right.\\
&\quad\left.+\sum_{j=1}^{M}\mathrm{Tr}\left(\mathbf{G}_{on}\mathbf{W}_j\right)+\mathbf{G}_{on}Z-\Lambda\left(\theta\right)\right\}\\
&\quad+\lambda_2\sum_{n=1}^{N}\left\{\left[\mathrm{Tr}\left(\mathbf{V}_n\right)+\mathrm{Tr}\left(\mathbf{U}_n\right)\right]\frac{1}{\varpi}+P_\varepsilon-\Gamma\right\}\\
&\quad+\lambda_3\left\{\left[\sum_{m=1}^{M}\mathrm{Tr}\left(\mathbf{W}_m\right)+\mathrm{Tr}\left(\mathbf{Z}\right)\right]\frac{1}{\varpi}+P_\varepsilon-\Delta\right\}, \qquad (56)
\end{aligned}$$

where $\mathbf{A}\in\mathbb{H}_+^{N_f}$, $\mathbf{B}\in\mathbb{H}_+^{N_f}$, $\mathbf{C}\in\mathbb{H}_+^{N_m}$, $\mathbf{D}\in\mathbb{H}_+^{N_m}$, $\lambda_1\in\mathbb{R}_+$, $\lambda_2\in\mathbb{R}_+$, $\lambda_3\in\mathbb{R}_+$ are the Lagrange multipliers associated with problem (16). We accomplish the proof by taking an example of $\mathbf{V}_\gamma(\gamma\in[1,N])$, hence we derive the corresponding Karush-Kuhn-Tucker conditions as

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\mathbf{V}_\gamma} &= -\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}+\frac{\mathbf{G}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{G}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{ER}}\right)\hat{I}_\gamma^{ER}}\\
&\quad-\lambda_1\sum_{n=1}^{N}\mathbf{G}_{\gamma n}-\mathbf{A}_\gamma+\left(\xi+\lambda_2\right)\frac{1}{\varpi}I\\
&\quad-\sum_{i=1,1\neq\gamma}^{N}\left[\frac{\mathbf{V}_i\mathbf{G}_{ii}\mathbf{G}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{G}_{ii}\mathbf{V}_i}{\hat{I}_i^{ER}}\right)\left(\hat{I}_i^{ER}\right)^2}\right.\\
&\quad\left.-\frac{\mathbf{V}_i\mathbf{H}_{ii}\mathbf{H}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{H}_{ii}\mathbf{V}_i}{\hat{I}_i^{IR}}\right)\left(\hat{I}_i^{IR}\right)^2}\right]\\
&\quad-\sum_{m=1}^{M}\left[\frac{\mathbf{H}_m\mathbf{W}_m\mathbf{H}_{\gamma om}}{\ln 2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)\left(\hat{I}_m\right)^2}\right.\\
&\quad\left.-\frac{\mathbf{H}_e\mathbf{W}_m\mathbf{H}_{\gamma oe}}{\ln 2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\left(\hat{I}_e\right)^2}\right]=0, \qquad (57)\\
&\mathbf{V}_n\mathbf{A}_n=0,\quad \mathbf{A}_n\succeq\mathbf{0},\ \forall n\in[1,N]. \qquad (58)
\end{aligned}$$

Then, the following equation holds

$$\begin{aligned}
&\left\{-\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}+\frac{\mathbf{G}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{G}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{ER}}\right)\hat{I}_\gamma^{ER}}\right.\\
&\quad-\lambda_1\sum_{n=1}^{N}\mathbf{G}_{\gamma n}-\mathbf{A}_\gamma+\left(\xi+\lambda_2\right)\frac{1}{\varpi}I\\
&\quad-\sum_{i=1,1\neq\gamma}^{N}\left[\frac{\mathbf{V}_i\mathbf{G}_{ii}\mathbf{G}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{G}_{ii}\mathbf{V}_i}{\hat{I}_i^{ER}}\right)\left(\hat{I}_i^{ER}\right)^2}\right.\\
&\quad\left.-\frac{\mathbf{V}_i\mathbf{H}_{ii}\mathbf{H}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{H}_{ii}\mathbf{V}_i}{\hat{I}_i^{IR}}\right)\left(\hat{I}_i^{IR}\right)^2}\right]\\
&\quad-\sum_{m=1}^{M}\left[\frac{\mathbf{H}_m\mathbf{W}_m\mathbf{H}_{\gamma om}}{\ln 2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)\left(\hat{I}_m\right)^2}\right.\\
&\quad\left.\left.-\frac{\mathbf{H}_e\mathbf{W}_m\mathbf{H}_{\gamma oe}}{\ln 2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\left(\hat{I}_e\right)^2}\right]\right\}\mathbf{V}_\gamma=0, \qquad (59)\\
&\Rightarrow\left\{\frac{\mathbf{G}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{G}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{ER}}\right)\hat{I}_\gamma^{ER}}-\lambda_1\sum_{n=1}^{N}\mathbf{G}_{\gamma n}+\left(\xi+\lambda_2\right)\frac{1}{\varpi}I\right.\\
&\quad-\sum_{i=1,1\neq\gamma}^{N}\left[\frac{\mathbf{V}_i\mathbf{G}_{ii}\mathbf{G}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{G}_{ii}\mathbf{V}_i}{\hat{I}_i^{ER}}\right)\left(\hat{I}_i^{ER}\right)^2}\right.\\
&\quad\left.\left.-\frac{\mathbf{V}_i\mathbf{H}_{ii}\mathbf{H}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{H}_{ii}\mathbf{V}_i}{\hat{I}_i^{IR}}\right)\left(\hat{I}_i^{IR}\right)^2}\right]\right.
\end{aligned}$$

$$-\sum_{m=1}^{M}\left[\frac{\mathbf{H}_m\mathbf{W}_m\mathbf{H}_{\gamma om}}{\ln 2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)\left(\hat{I}_m\right)^2}-\frac{\mathbf{H}_e\mathbf{W}_m\mathbf{H}_{\gamma oe}}{\ln 2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\left(\hat{I}_e\right)^2}\right]\Bigg\}\mathbf{V}_\gamma=\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}V_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}\mathbf{V}_\gamma, \tag{60}$$

$$\begin{aligned}\Rightarrow V_\gamma=&\left\{\frac{\mathbf{G}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{G}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{ER}}\right)\hat{I}_\gamma^{ER}}-\lambda_1\sum_{n=1}^{N}\mathbf{G}_{\gamma n}\right.\\&+(\xi+\lambda_2)\frac{1}{\varpi}I-\sum_{i=1,1\neq\gamma}^{N}\left[\frac{\mathbf{V}_i\mathbf{G}_{ii}\mathbf{G}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{G}_{ii}\mathbf{V}_i}{\hat{I}_i^{ER}}\right)\left(\hat{I}_i^{ER}\right)^2}\right.\\&\left.-\frac{\mathbf{V}_i\mathbf{H}_{ii}\mathbf{H}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{H}_{ii}\mathbf{V}_i}{\hat{I}_i^{IR}}\right)\left(\hat{I}_i^{IR}\right)^2}\right]\\&-\sum_{m=1}^{M}\left[\frac{\mathbf{H}_m\mathbf{W}_m\mathbf{H}_{\gamma om}}{\ln 2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)\left(\hat{I}_m\right)^2}\right.\\&\left.\left.-\frac{\mathbf{H}_e\mathbf{W}_m\mathbf{H}_{\gamma oe}}{\ln 2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\left(\hat{I}_e\right)^2}\right]\right\}^{-1}\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}\mathbf{V}_\gamma.\end{aligned} \tag{61}$$

Finally, the following rank relation holds

$$\begin{aligned}&\mathrm{rank}(\mathbf{V}_\gamma)\\&=\mathrm{rank}\left\{\left\{\frac{\mathbf{G}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{G}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{ER}}\right)\hat{I}_\gamma^{ER}}-\lambda_1\sum_{n=1}^{N}\mathbf{G}_{\gamma n}\right.\right.\\&+(\xi+\lambda_2)\frac{1}{\varpi}I-\sum_{i=1,1\neq\gamma}^{N}\left[\frac{\mathbf{V}_i\mathbf{G}_{ii}\mathbf{G}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{G}_{ii}\mathbf{V}_i}{\hat{I}_i^{ER}}\right)\left(\hat{I}_i^{ER}\right)^2}\right.\\&\left.-\frac{\mathbf{V}_i\mathbf{H}_{ii}\mathbf{H}_{\gamma i}}{\ln 2\left(1+\frac{\mathbf{H}_{ii}\mathbf{V}_i}{\hat{I}_i^{IR}}\right)\left(\hat{I}_i^{IR}\right)^2}\right]\\&-\sum_{m=1}^{M}\left[\frac{\mathbf{H}_m\mathbf{W}_m\mathbf{H}_{\gamma om}}{\ln 2\left(1+\frac{\mathbf{H}_m\mathbf{W}_m}{\hat{I}_m}\right)\left(\hat{I}_m\right)^2}\right.\\&\left.\left.\left.-\frac{\mathbf{H}_e\mathbf{W}_m\mathbf{H}_{\gamma oe}}{\ln 2\left(1+\frac{\mathbf{H}_e\mathbf{W}_m}{\hat{I}_e}\right)\left(\hat{I}_e\right)^2}\right]\right\}^{-1}\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}V_\gamma\right\}\\&\leq\mathrm{rank}\left[\frac{\mathbf{H}_{\gamma\gamma}}{\ln 2\left(1+\frac{\mathbf{H}_{\gamma\gamma}\mathbf{V}_\gamma}{\hat{I}_\gamma^{IR}}\right)\hat{I}_\gamma^{IR}}\right]\leq 1,\end{aligned} \tag{62}$$

which completes the proof that $\mathbf{V}_n$ is of rank-one. Likewise, the ranks of $\mathbf{U}_n$, $\mathbf{W}_m$, $\mathbf{Z}_0$ can all be proved to be one, and we omit the detailed process for the sake of briefness.

REFERENCES

[1] A. Damnjanovic *et al.*, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[2] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[3] 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.

[4] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, Nov. 2014.

[5] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: A tutorial," *IEEE Wireless Commun.*, vol. 18, no. 2, pp. 66–74, Apr. 2011.

[6] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 3rd Quart., 2014.

[7] Z. Zhong, P. Janhua, L. Wenyu, and H. Kaizhi, "A tractable approach to analyzing the physical-layer security in k-tier heterogeneous cellular networks," *China Commun.*, vol. 12, no. S1, pp. 166–173, Dec. 2015.

[8] H.-M. Wang, T.-X. Zheng, J. Yuan, D. Towsley, and M. H. Lee, "Physical layer security in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1204–1219, Mar. 2016.

[9] X. Qi, K. Huang, B. Li, L. Jin, and X. Ji, "Physical layer security in multi-antenna cognitive heterogeneous cellular networks: A unified secrecy performance analysis," *Sci. China Inf. Sci.*, vol. 61, no. 2, pp. 22–31, Feb. 2018.

[10] W. Tang, S. Feng, Y. Ding, and Y. Liu, "Physical layer security in heterogeneous networks with jammer selection and full-duplex users," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7982–7995, Dec. 2017.

[11] S. Gong, C. Xing, Z. Fei, and J. Kuang, "Resource allocation for physical layer security in heterogeneous network with hidden eavesdropper," *China Commun.*, vol. 13, no. 3, pp. 82–95, Mar. 2016.

[12] L. Liu, R. Zhang, and K.-C. Chua, "Secrecy wireless information and power transfer with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1850–1863, Apr. 2014.

[13] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure beamforming in MISO full-duplex two-way secure communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 408–414, Jan. 2016.

[14] Z. Chu, Z. Zhu, M. Johnston, and S. Y. Le Goff, "Simultaneous wireless information power transfer for MISO secrecy channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 6913–6925, Sep. 2015.

[15] H. Zhang, Y. Huang, C. Li, and L. Yang, "Secure transmission scheme for SWIPT in MISO broadcast channel with confidential messages and external eavesdroppers," in *Proc. 81st Veh. Technol. Conf.*, Lasgow, U.K., May 2015, pp. 1–5.

[16] M. R. A. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 40–54, Jan. 2015.

[17] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, May 2015.

[18] M. Tian, X. Huang, Q. Zhang, and J. Qin, "Robust AN-aided secure transmission scheme in MISO channels with simultaneous wireless information and power transfer," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 723–727, Jun. 2015.

[19] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Joint optimization of AN-aided beamforming and power splitting designs for MISO secrecy channel with SWIPT," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[20] S. Akbar, Y. Deng, A. Nallanathan, and M. Elkashlan, "Downlink and uplink transmission in K-tier heterogeneous cellular network with simultaneous wireless information and power transfer," in *Proc. IEEE GlobeCom*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[21] B. Li, Z. Fei, Z. Chu, and Y. Zhang, "Secure transmission for heterogeneous cellular networks with wireless information and power transfer," *IEEE Syst. J.*, to be published.

[22] Y. Ren, T. Lv, H. Gao, and Y. Li, "Secure wireless information and power transfer in heterogeneous networks," *IEEE Access*, vol. 5, pp. 4967–4979, May 2017.

[23] Y. Dong, M. J. Hossain, and J. Cheng, "Joint power control and time switching for SWIPT systems with heterogeneous QoS requirements," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 328–331, Feb. 2016.
[24] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2082–2085, Dec. 2015.
[25] E. Boshkovska, N. Zlatanov, L. Dai, D. W. K. Ng, and R. Schober, "Secure SWIPT networks based on a non-linear energy harvesting model," in *Proc. IEEE WCNC Workshop*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.
[26] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[27] K.-Y. Wang, A. Man-Cho So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
[28] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications* (MPSSIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2001.
[29] K.-G. Nguyen, Q.-D. Vu, M. Juntti, and L.-N. Tran, "Distributed solutions for energy efficiency fairness in multicell MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6232–6247, Sep. 2017.
[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[31] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.

XIN HU received the B.E. degree in technique and instrumentation of measurements from Xi'an Jiaotong University, Xi'an, China, in 2015. He is currently pursuing the M.S. degree with the National Digital Switching System Engineering and Technological R&D Center, Henan, China. His major research interests include physical layer security, simultaneous wireless information and power transfer and Internet of Things.

BIN LI received the M.S. degree in communication and information systems from the Guilin University of Electronic Technology, Guilin, China, in 2013. He is currently pursuing the Ph.D. degree with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. From 2013 to 2014, he was a Research Assistant with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. Since 2017, he has been a Visiting Student with the Department of Informatics, University of Oslo, Norway. His research interests include wireless cooperative networks, physical layer security, energy harvesting, and MIMO techniques.

KAIZHI HUANG received the Ph.D. degree in communication and information system from Tsinghua University, Beijing, China. She is currently a Professor and a Supervisor of postgraduate student at the National Digital Switching System Engineering and Technological R&D Center, Zhengzhou, China. She is also serving as the Leader of the Wireless Mobile Communication Innovation Technology Team, Henan. Her major research interests include wireless mobile communication network and information theory.

ZESONG FEI (M'07–SM'16) received the Ph.D. degree in electronic engineering from the Beijing Institute of Technology (BIT), in 2004. He is currently a Professor with the Research Institute of Communication Technology, BIT, where he is involved in the design of the next generation high-speed wireless communication. His research interests include wireless communications and multimedia signal processing. He is the Chief Investigator of the National Natural Science Foundation of China. He is a Senior Member of the Chinese Institute of Electronics and the China Institute of Communications.

KAI-KIT WONG (M'01–SM'08–F'16) received the B.Eng., M.Phil., and Ph.D. degrees in Electrical and Electronic Engineering from The Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively. After graduation, he took up academic and research positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, the Smart Antennas Research Group, Stanford University, and the University of Hull, U.K. He is the Chair of wireless communications with the Department of Electronic and Electrical Engineering, University College London, U.K.

His current research interests include 5G and beyond mobile communications, including topics such as massive MIMO, full-duplex communications, millimeter-wave communications, edge caching and fog networking, physical layer security, wireless power transfer and mobile computing, V2X communications, cognitive radios, fluid antenna communications systems, and remote ECG detection. He was a co-recipient of the 2013 IEEE Signal Processing Letters Best Paper Award and the 2000 IEEE VTS Japan Chapter Award at the IEEE Vehicular Technology Conference in Japan in 2000, and a few other international best paper awards.

Dr. Wong is a fellow of the IET. He is on the editorial board of several international journals. He served as an Associate Editor for the IEEE Signal Processing Letters from 2009 to 2012 and an Editor for the IEEE Transactions on Wireless Communications from 2005 to 2011. He was also a Guest Editor for the IEEE JSAC SI on virtual MIMO in 2013. He has been serving as a Senior Editor for the IEEE Communications Letters since 2012 and the IEEE Wireless Communications Letters since 2016. He is currently a Guest Editor for the IEEE JSAC SI on physical layer security for 5G.