IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 23, 2019, accepted August 4, 2019, date of publication August 16, 2019, date of current version September 3, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2935798

Robust Secrecy Energy Efficiency Optimization for Wireless Powered Heterogeneous Networks Using Distributed ADMM Algorithm

BO ZHANG[1,2], BIN LI[3], KAIZHI HUANG[1,2], ZHOU ZHONG[1,2], LIJIAN ZHANG[4], AND ZESONG FEI[5], (Senior Member, IEEE)

[1]Information Engineering University, Zhengzhou 450001, China
[2]National Digital Switching System Engineering and Technological Research and Developing Center, Zhengzhou 450001, China
[3]School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China
[4]Institute of Systems Engineering, Beijing 100071, China
[5]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

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

This work was supported in part by the National Natural Science Foundation of China under Grant 61871404, Grant 61701538, and Grant 61871032, in part by the Innovation Group Project of the China National Natural Science Foundation under Grant 61521003, in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) Fund, and in part by the Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) Fund.

ABSTRACT Under the case of imperfect channel state information (CSI), this paper investigates the robust secrecy energy efficiency (SEE) optimization for the heterogeneous networks (HeNets) supported by simultaneous wireless information and power transfer. Specifically, we first consider a two-tier HeNet composed of a macrocell base station (MBS) and several femtocell base stations (FBSs), where the MBS serves multiple macrocell users (MUs) while each FBS serves an information receiver (IR) and an multiple-antenna energy receiver (ER). Meanwhile, a malicious multiple-antenna eavesdropper (Eve) attempts to wiretap the downlink information of MUs and the ER acts as a potential Eve to eavesdrop the confidential information for IR in the same femtocell. To enhance the secrecy performance, artificial noise (AN) is injected into the downlink information beams of the MBS and FBSs. By considering the user fairness, the problem of SEE maximization of the whole network is formulated via a cross-tier multi-cell coordinated beamforming design. The resulting problem contains infinite constraints caused by CSI error, which is nonconvex and cannot be solved directly. To this regard, we resort to the successive convex approximation, S-procedure and semi-definite relaxation techniques to obtain a solvable form of it. Furthermore, to reduce the overhead of information exchange among coordinated BSs, we develop a distributed solution based on alternative direction multiplier method (ADMM) that can achieve a good approximation performance. Finally, simulation results verify the validity of the proposed AN-aided cross-tier multi-cell coordinated beamforming design and distributed ADMM-based design.

INDEX TERMS Heterogeneous networks, simultaneous wireless information and power transfer, secrecy energy efficiency, imperfect CSI, convex optimization, alternative direction multiplier method.

I. INTRODUCTION

A. BACKGROUND

Since the dramatic increase of Internet-enabled smart devices (e.g., smart phones and electronic tablets) has spurred the explosive growth of high-rate multimedia wireless services, it becomes difficult for mobile operators to utilize conventional homogeneous networks to achieve higher capacity and coverage for next-generation 5G wireless communications. Increasing cell density for higher spatial spectrum reuse is one viable solution to do so. In this regard, heterogeneous networks (HeNets) have been proposed as a promising network densification architecture, where different types of overlaid small cells are deployed under the coverage of conventional macrocell and share the same spectrum resource [1], [2]. Nevertheless, the resulting increased cross-tier interferences also appear with the increase of spectral efficiency, which may deteriorate the quality of service (QoS) of

The associate editor coordinating the review of this article and approving it for publication was Yuan Gao.

legitimate users. Consequently, to reduce/eliminate the mutual interferences, the transmit signals should be designed elaborately.

Meanwhile, to satisfy the high energy consumption resulting from the ever-increasing traffic needs in 5G, simultaneous wireless information and power transfer (SWIPT) acting as a novel research frontier of combining wireless power transfer and wireless communication, has envisioned as a promising energy supply for the energy-constrained wireless systems [3]. In SWIPT, wireless devices can harvest energy from man-made radio frequency (RF) signals while ignoring the content carried, which contributes to the case of applying SWIPT into HetNets, i.e., HeNets with SWIPT, because the inter-tier and intra-tier interferences in HeNets bears potentials to enable more efficient wireless power transfer and usage [4]. However, the broadcast nature of wireless channels and the more open network architecture of HeNets with SWIPT make it vulnerable to resist spiteful wiretapping from eavesdroppers (Eves). Therefore, it is of much concern facing HeNets with SWIPT to secure wireless data transmission. To address this concern, physical layer security (PLS) provides a new method to secure the wireless information, which exploits the random characteristics of wireless channels and acts as an important complement to the traditional high-layer cryptographic methods. Most recently, with the assumption of perfect channel state information (CSI), some progresses on the PLS of HeNets with SWIPT have been achieved [5].

B. RELATED WORKS

Nevertheless, due to existence of quantization and estimation errors in practical networks, it is very hard for the transmitter to acquire the perfect CSI. Generally speaking, there are two CSI error models, i.e., bounded CSI error model and probabilistic CSI error model. The bounded CSI error model is described by a bounded set, which has a low implementation complexity but may under estimate the actual performance. On the other hand, a probabilistic model is used to model the probabilistic CSI errors, which is mainly applicable to the delay-sensitive wireless application scenarios [6], [7]. With the presence of CSI errors, the transmit signal design may become inaccurate, which can make more confidential information leak to Eves. As a result, studying the robust secure communication for HeNets with SWIPT is of paramount importance.

Over the past few years, some progresses have been achieved to the issue of robust PLS in SWIPT systems. With the existence of bounded CSI errors, robust power minimization (PM) and secrecy rate maximization (SRM) designs were investigated in MISO SWIPT systems [8], [9], and were later extended to MIMO SWIPT systems [10], relay networks [11], respectively. In addition, robust harvesting energy maximization (HEM) design and multi-objective optimization were also studied for a secure MISO SWIPT system [12] and cognitive radio network [13], respectively. On the other hand, with the existence of probabilistic CSI errors, [14] and [15] explored the robust SRM designs with outage constraints for MISO SWIPT systems, and utilized Bernstein-type inequalities (BTI), S-Procedure, and large deviation inequalities (LDI) based conservative approximations to transform the outage constraints into tractable forms. Afterwards, the robust PM, SRM and HEM designs were investigated for MIMO SWIPT systems in [16]–[18], respectively. Generally, these above-mentioned studies only explored the robust PLS transmission in single-cell coverage case, but ignored the deployment of SWIPT in HeNets. For robustly securing HeNets with PLS, only a few works explored this topic and focused on the secrecy model and performance evaluations in [19] and [20], which ignored the delicate design of transmit signals.

A common of previous related works is to address the robust secure issue separately in HeNets and SWIPT systems, respectively. Recently, with the advantage of converting various interferences into power gain, applying the SWIPT into HetNets, i.e., HetNets with SWIPT, has attracted increasing interests, and the robust PLS was also subsequently explored for this fusion case in [21] and [22], respectively. Specifically, the transmit beamforming (TBF) and artificial noise (AN) vectors of the macro base station (MBS) and femto base stations (FBSs) were jointly designed to maximize the secrecy rate of system while considering the imperfect CSIs of potential Eves with single antenna and the QoS constraints of legitimate users in [21]. Then, the worst-case based solution was obtained with the aid of semi-definite relaxation (SDR), successive convex approximation (SCA) and S-procedure techniques. Afterwards, [22] further explored the similar topic while considering both the imperfect CSIs of legitimate users and Eves, and reformulated the robust quadratic matrix inequality (QMI) constraints as the linear matrix inequality representations in light of S-procedure and first-order Taylor series expansion.

C. MOTIVATIONS AND MAIN CONTRIBUTIONS

To the best of our knowledge, the study for robust secure communications about HetNets with SWIPT is still largely missing, and three existing weaknesses in this field still need to be further strengthen: 1) all the mentioned-above works only consider the single-antenna Eves, but the more dangerous case where Eves are equipped with multiple antennas is ignored; 2) moreover, only one type of users, i.e., macrocell user (MU) or femtocell user (FU), was considered but the worse case that both MU and FU suffered from wiretapping was ignored; 3) last but not least, the joint design in all the above works is executed at a calculating center, which may lead to inestimable processing stress for the calculating center and much signaling overhead for the network especially when the number of coordinated BSs is large. As far as the known of authors, these weaknesses in the field have not been explored jointly yet, which motivates our work in this paper.

In this paper, we consider a two-tier HetNet with SWIPT with imperfect CSIs, and focus on the robust transmit optimization while considering the security of various decoding users towards green communications. To be specific, our main contributions can be summarized as follows:

1) We establish the robust PLS model of a two-tier HetNet with SWIPT under the imperfect CSI case while facing multiple-antenna wiretapping, where multiple FBSs are deployed under the coverage of one MBS. The MBS serves multiple MUs in the presence of a malicious multiple-antenna Eve while each FBS serves a pair of information receiver (IR) and energy receiver (ER), where the multiple-antenna ER acts as a potential Eve to wiretap the information of IR in the same femtocell. To promote the secrecy performance and power transfer, AN is aided into the downlink beam at both MBS and FBSs simultaneously.
2) To achieve the secure and green communications, we explore the joint TBF and AN design to maximize the secrecy energy efficiency (SEE) while considering the fairness among multiple cells under the imperfect CSI case. More specifically, the formulated problem containing infinite constraints is non-convex and hard to solve directly. To address this challenge, we first acquire an equivalent form of the original problem by employing the SCA and SDR techniques; then, following the worst-case robustness philosophy to deal with the infinite constraints caused by CSI errors, we reformulate the worst-case SEE maximization problem into a quadratic matrix inequality(QMI) problem in light of S-procedure, and rely on the linear matrix inequality (LMI) reformulation for the semi-infinite QMI constraints to construct solvable SOCP problems.
3) A distributed design based on the alternative direction multiplier method (ADMM) is further developed to reduce the overhead of CSI exchange and release the processing stress on the calculating center. Moreover, this distributed design can draw near the optimal solution while allowing each BS to handle its local CSIs only.

The remaining part of this paper is organized as follows. In Section II, we introduce the system model and formulate the optimization problem. Then, the robust TBF and AN optimization in centralized design and distributed design with ADMM are explored in Section III and Section IV, respectively. Finally, Section V presents the simulation results to validate the effectiveness of the proposed design, followed by our conclusions in Section VI.

Notations: Boldface lowercase and uppercase letters denote vectors and matrices, respectively. $C^{N\times 1}$ represents a $N \times 1$ complex column vector. The transpose, conjugate transpose and inverse of the matrix are denoted as $\mathbf{A}^T$, $\mathbf{A}^H$ and $\mathbf{A}^{-1}$, respectively, while $\mathbf{A} \succeq \mathbf{0}$ indicates that $\mathbf{A}$ is a Hermitian positive semidefinite matrix. Tr() is the trace operator, $\|\|_F$ represents the matrix Frobenius norm, Rank() stands for the rank of a matrix. $CN(\mathbf{u}, \mathbf{\Phi})$ denotes a complex Gaussian variable with mean $\mathbf{u}$ and covariance $\mathbf{\Phi}$, while $\Pr\{A\}$ and $\otimes$ are the probability of an event A and the Kronecker product, respectively. In addition, for the convenience of readers, the relevant symbols in the following sections of this paper are summarized in Table1 to refer.

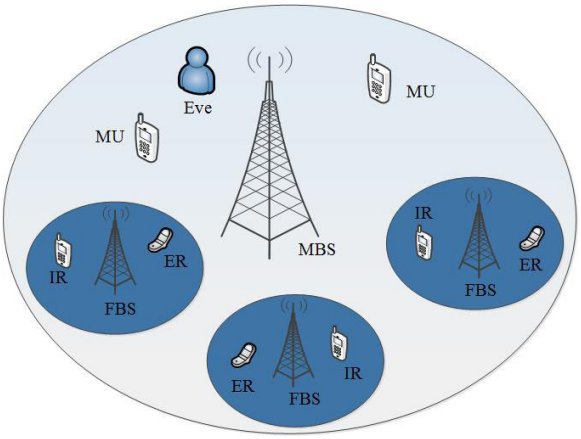

FIGURE 1. System model.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

We consider a two-tier HetNet with SWIPT, where N FBSs can be deployed within the coverage of an MBS, as depicted in Fig. 1. Each FBS serves a pair of femtocell users (FUs), one of which receives information (i.e., IR) while another one receives the energy (i.e., ER), respectively, and shares certain spectral resources as MBS serving M MUs to improve the spectrum efficiency. The MBS and each FBS are equipped with $N_M > M$ and $N_F > 2$ antennas, respectively, whereas each MU and IR are equipped with a single receive antenna while each ER has N_{ER} antennas. Meanwhile, a potential Eve with N_E antennas attempts to intercept the confidential information intended for MUs. Furthermore, ERs may be malicious for intercepting the information signal transmitted by the FBS to IR without any attacks. Thus, ERs are seen as potential Eves, and it is assumed that each ER only attempts to overhear the IR in the same femtocell.

For the sake of simplicity, let us denote the n-th FBS as FBS_n, the m-th MU as MU_m, the IR and ER of FBS_n as FU_n and ER_n, respectively. The channel vectors from the MBS to MU_m , Eve, FU_n and ER_n are denoted by $\mathbf{h}_m \in C^{N_M\times 1}$, $\mathbf{H}_E \in C^{N_M\times N_E}$, $\mathbf{h}_{MFn} \in C^{N_M\times 1}$ and $\mathbf{H}_{ERn} \in C^{N_F\times N_{ER}}$, respectively. Likewise, The channel vectors from the FBS_n to MU_m, Eve, FU_n, ER_n, $\mathrm{FU}_t(t \neq n)$ and $\mathrm{ER}_t(t \neq n)$ are denoted by $\mathbf{h}_{n,m} \in C^{N_F\times 1}$, $\mathbf{H}_{n,E} \in C^{N_F\times N_E}$, $\mathbf{h}_{Fn} \in C^{N_F\times 1}$, $\mathbf{H}_{ERn} \in C^{N_F\times N_{ER}}$, $\mathbf{h}_{nFt} \in C^{N_F\times 1}$ and $\mathbf{H}_{nERt} \in C^{N_F\times N_{ER}}$, respectively. These channel coefficients are independent, and the elements are independent and identically distributed (i.i.d.) complex Gaussian random variables. To support secure communication and facilitate energy harvesting at ERs, AN is aided into the downlink beam of the MBS and FBSs simultaneously. Then, the received signal at MU_m can be expressed as

$$\mathbf{y}_m = \mathbf{h}_m^H \mathbf{w}_m \mathbf{s}_m + \mathbf{h}_m^H \Big(\sum_{p\neq m}^{M} \mathbf{w}_p \mathbf{s}_p + \mathbf{z}_0\Big) + \sum_{n=1}^{N} \mathbf{h}_{n,m}^H (\mathbf{w}_{Fn}\mathbf{s}_{Fn} + \mathbf{z}_n) + n_m, \quad m \in [1, M], \quad (1)$$

where $\mathbf{w}_f \in C^{N_M \times 1}(f = m, p)$ and $\mathbf{w}_{Fn} \in C^{N_F \times 1}$ denote the beamforming vector from the MBS to MU_f and that from the FBS_n to FU_n, respectively; $\mathbf{s}_f(f = m, p)$ with $\mathrm{E}\{|\mathbf{s}_f|^2\} = 1$ and $\mathbf{s}_{Fn}$ with $\mathrm{E}\{|\mathbf{s}_{Fn}|^2\} = 1$ represent the message symbol from the MBS to MU_f and that from the FBS_n to FU_n, respectively; $\mathbf{z}_0 \in C^{N_M \times 1}$ and $\mathbf{z}_n \in C^{N_F \times 1}$ stand for the AN vectors at the MBS and FBS_n, and they both follow Gaussian distribution, i.e., $\mathbf{z}_0 \sim CN(0, \mathbf{Z}_0)$ and $\mathbf{z}_n \sim CN(0, \mathbf{Z}_n)$, $\mathbf{Z}_0$ and $\mathbf{Z}_n$ are the corresponding covariance matrix of $\mathbf{z}_0$ and $\mathbf{z}_n$; n_m indicates the additive white Gaussian noise (AWGN) at MU_m.

If the downlink information of MU_m is wiretapped by Eve, then the received signal at Eve can be expressed as

$$\mathbf{y}_{Em} = \mathbf{H}_E^H \mathbf{w}_m \mathbf{s}_m + \mathbf{H}_E^H \Big(\sum_{p \neq m}^{M} \mathbf{w}_p \mathbf{s}_p + \mathbf{z}_0\Big) + \sum_{n=1}^{N} \mathbf{H}_{n,E}^H (\mathbf{w}_{Fn} \mathbf{s}_{Fn} + \mathbf{z}_n) + n_E, \quad m \in [1, M], \tag{2}$$

where n_E indicates the AWGN at Eve.

Similarly, the received signal at FU_n can be expressed as

$$\mathbf{y}_{Fn} = \mathbf{h}_{Fn}^H (\mathbf{w}_{Fn} \mathbf{s}_{Fn} + \mathbf{z}_n) + \sum_{t \neq n}^{K} \mathbf{h}_{tFn}^H (\mathbf{w}_{Ft} \mathbf{s}_{Ft} + \mathbf{z}_t) + \mathbf{h}_{MFn}^H \Big(\sum_{m=1}^{M} \mathbf{w}_m \mathbf{s}_m + \mathbf{z}_0\Big) + n_{Fn}, \quad n \in [1, N], \tag{3}$$

where n_{Fn} is the AWGN at FU_n.

The received signal at ER_n is denoted as

$$\mathbf{y}_{ERn} = \mathbf{H}_{ERn}^H (\mathbf{w}_{Fn} \mathbf{s}_{Fn} + \mathbf{z}_n) + \sum_{t \neq n}^{N} \mathbf{H}_{tERn}^H (\mathbf{w}_{Ft} \mathbf{s}_{Ft} + \mathbf{z}_t) + \mathbf{H}_{MERn}^H \Big(\sum_{m=1}^{M} \mathbf{w}_m \mathbf{s}_m + \mathbf{z}_0\Big) + n_{ERn}, \quad n \in [1, N], \tag{4}$$

where n_{ERn} represents the AWGN at ER_n. Without loss of generality, we assume that all the AWGN in (1)∼(4) follow the independent and identical Gaussian distribution, i.e., $CN(0, \sigma^2)$.

Based on the definition of secrecy rate, we let the sum secrecy rate of MUs as the secrecy performance of the macrocell. Defining $\mathbf{W}_m = \mathbf{w}_m \mathbf{w}_m^H$ and $\mathbf{W}_{Fn} = \mathbf{w}_{Fn} \mathbf{w}_{Fn}^H$, then it can be expressed as

$$C_S^0 = \sum_{m=1}^{M} \{\log_2(1 + \mathrm{SINR}_m) - \log_2 |\mathbf{R}_{Em}|\}^+, \tag{5}$$

where $\{a\}^+$ denotes $\max\{a, 0\}$; SINR_m represents the signal to interference plus noise ratio (SINR) of MU_m, and

$$\mathrm{SINR}_m = \frac{\mathrm{Tr}(\mathbf{H}_m \mathbf{W}_m)}{\sum_{n=1}^{N} \mathrm{Tr}(\mathbf{H}_{n,m} \mathbf{A}_n) + \mathrm{Tr}(\mathbf{H}_m \mathbf{C}_m) + \sigma^2}, \tag{6a}$$

$$\mathbf{R}_{Em} = \mathbf{I}_{N_E} + (\mathbf{H}_E^H \mathbf{W}_m \mathbf{H}_E) \times \{\mathbf{H}_E^H (\sum_{p \neq m}^{M} \mathbf{W}_p + \mathbf{Z}_0) \mathbf{H}_E + \sum_{n=1}^{N} \mathbf{H}_{n,E}^H (\mathbf{W}_{Fn} + \mathbf{Z}_n) \mathbf{H}_{n,E} + \sigma^2 \mathbf{I}_{N_E}\}^{-1}, \tag{6b}$$

where $\mathbf{H}_m = \mathbf{h}_m^H \mathbf{h}_m$, $\mathbf{H}_{n,m} = \mathbf{h}_{n,m}^H \mathbf{h}_{n,m}$, $\mathbf{A}_i = \mathbf{W}_{Fi} + \mathbf{Z}_i$ $(i = \{n, t\})$, $\mathbf{C}_m = \sum_{p \neq m}^{M} \mathbf{W}_p + \mathbf{Z}_0$.

Similarly, the secrecy performance of the n-th femtocell is given by

$$C_S^n = \{\log_2(1 + \mathrm{SINR}_{\mathrm{IR}n}) - \log_2 |\mathbf{R}_{\mathrm{EIR}n}|\}^+, \quad n \in [1, N], \tag{7}$$

where $\mathrm{SINR}_{\mathrm{IR}n}$ denotes the SINR of IR_n, and

$$\mathrm{SINR}_{\mathrm{IR}n} = \frac{\mathrm{Tr}(\mathbf{H}_{Fn} \mathbf{W}_{Fn})}{\mathrm{Tr}(\mathbf{H}_{Fn} \mathbf{Z}_n) + \mathrm{Tr}(\mathbf{H}_{tFn} \mathbf{A}_t) + \mathrm{Tr}(\mathbf{H}_{MFn} \mathbf{B}) + \sigma^2}, \tag{8a}$$

$$\mathbf{R}_{\mathrm{EIR}n} = \mathbf{I}_{N_{ER}} + (\mathbf{H}_{ERn}^H \mathbf{W}_{Fn} \mathbf{H}_{ERn}) \times \{\mathbf{H}_{ERn}^H \mathbf{A}_n \mathbf{H}_{ERn} + \mathbf{H}_{MERn}^H \mathbf{B} \mathbf{H}_{MERn} + \sigma^2 \mathbf{I}_{N_{ER}}\}^{-1}, \tag{8b}$$

where $\mathbf{H}_{Fn} = \mathbf{h}_{Fn}^H \mathbf{h}_{Fn}$, $\mathbf{H}_{tFn} = \mathbf{h}_{tFn}^H \mathbf{h}_{tFn}$, $\mathbf{H}_{MFn} = \mathbf{h}_{MFn}^H \mathbf{h}_{MFn}$, $\mathbf{B} = \sum_{m=1}^{M} \mathbf{W}_m + \mathbf{Z}_0$.

For the EH process, we consider non-linear model (NLM) [23] in this paper, then the harvested energy of ER_n is expressed as

$$\mathrm{E}_{\mathrm{h\_out}}(n) = \frac{A[1 + \exp(ab)]}{[1 + \exp(-a(\mathrm{E}_{\mathrm{h\_in}}(n)))] \exp(ab)} - \frac{A}{\exp(ab)}, \tag{9}$$

where A, a and b are the maximum output power of RF-to-DC circuit, constant parameters depending on the hardware components; $\mathrm{E}_{\mathrm{h\_in}}(n)$ represents the input RF signal power at ER_n and is expressed as

$$\mathrm{E}_{\mathrm{h\_in}}(n) = \mathrm{Tr}(\mathbf{H}_{ERn}^H \mathbf{H}_{ERn} \mathbf{A}_n) + \sum_{t \neq n}^{N} \mathrm{Tr}(\mathbf{H}_{tERn}^H \mathbf{H}_{tERn} \mathbf{A}_t) + \sum_{m=1}^{M} \mathrm{Tr}(\mathbf{H}_{MERn}^H \mathbf{H}_{MERn} \mathbf{B}), \quad n \in [1, N]. \tag{10}$$

B. PROBLEM FORMULATION

From (5), (7) and (10), we can see that the interference from the MBS invokes a negative impact on the quality of wireless transmission of both IR and ERs; meanwhile, it is beneficial for EH at ERs. Similarly, FBSs also have effects on the reception of MUs and Eves. As a result, it is not trivial to design beamforming of the MBS and FBSs for guaranteeing the secure transmissions of MUs and IRs in the presence of potential Eves while having a minimal effect on MUs and IRs. Meanwhile, it should be highlighted that the quality of CSIs have significant effects on the joint design process.

However, due to the existence of quantization and estimation errors in practical networks, it is very hard for the transmitter to acquire the perfect CSI. Thus, we consider the imperfect CSIs with the deterministic CSI errors in this paper. Then, the relation between actual and estimated CSIs can be respectively expressed as

$$\mathbf{h}_m = \hat{\mathbf{h}}_m + \mathbf{e}_m, \quad \|\mathbf{e}_m\| \le \varepsilon_m, \tag{11a}$$
$$\mathbf{h}_{n,m} = \hat{\mathbf{h}}_{n,m} + \mathbf{e}_{n,m}, \quad \|\mathbf{e}_{n,m}\| \le \varepsilon_{n,m}, \tag{11b}$$
$$\mathbf{h}_{Fn} = \hat{\mathbf{h}}_{Fn} + \mathbf{e}_{Fn}, \quad \|\mathbf{e}_{Fn}\| \le \varepsilon_{Fn}, \tag{11c}$$
$$\mathbf{h}_{tFn} = \hat{\mathbf{h}}_{tFn} + \mathbf{e}_{tFn}, \quad \|\mathbf{e}_{tFn}\| \le \varepsilon_{tFn}, \tag{11d}$$
$$\mathbf{h}_{MFn} = \hat{\mathbf{h}}_{MFn} + \mathbf{e}_{MFn}, \quad \|\mathbf{e}_{MFn}\| \le \varepsilon_{MFn}, \tag{11e}$$
$$\mathbf{H}_E = \hat{\mathbf{H}}_E + \mathbf{E}_E, \quad \|\mathbf{E}_E\| \le \varepsilon_E, \tag{11f}$$
$$\mathbf{H}_{n,E} = \hat{\mathbf{H}}_{n,E} + \mathbf{E}_{n,E}, \quad \|\mathbf{E}_{n,E}\| \le \varepsilon_{n,E}, \tag{11g}$$
$$\mathbf{H}_{ERn} = \hat{\mathbf{H}}_{ERn} + \mathbf{E}_{ERn}, \quad \|\mathbf{E}_{ERn}\| \le \varepsilon_{ERn}, \tag{11h}$$
$$\mathbf{H}_{tERn} = \hat{\mathbf{H}}_{tERn} + \mathbf{E}_{tERn}, \quad \|\mathbf{E}_{tERn}\| \le \varepsilon_{tERn}, \tag{11i}$$
$$\mathbf{H}_{MERn} = \hat{\mathbf{H}}_{MERn} + \mathbf{E}_{MERn}, \quad \|\mathbf{E}_{MERn}\| \le \varepsilon_{MERn}, \tag{11j}$$

where $\mathbf{h}_m$, $\hat{\mathbf{h}}_m$ and $\mathbf{e}_m$ are the actual channel vector, estimated channel vector and channel error vector of MU_m, respectively; ε_m denotes the corresponding channel error bound; the other channel vectors follow the same relation, so we do not repeat it for the sake of simplicity.

Under such a scenario and to improve the average SEE performance of the whole network, we focus on the joint design of beamforming vectors $\mathbf{w}_m$, $\mathbf{w}_{Fn}$ and AN matrixes $\mathbf{Z}_0$, $\mathbf{Z}_n$ to maximize the minimal SEE among all the cells. Hence, the joint design problem is formulated as

$$\max_{\substack{\{\mathbf{W}_m\},\mathbf{Z}_0,\\ \{\mathbf{W}_{Fn}\},\{\mathbf{Z}_n\}}} \ \{\min_{k\in[0,N]} \{S_{EE}(k)\}\} \tag{12a}$$
$$s.t. \quad \min_{\substack{\{\mathbf{E}_{ERn}\},\{\mathbf{E}_{tERn}\},\\ \{\mathbf{E}_{MERn}\}}} \{\mathrm{E_{h\_out}}(n)\} \ge \theta_n, \quad n \in [1,N], \tag{12b}$$
$$P_0 \le P_M, \tag{12c}$$
$$P_n \le P_F, \quad n \in [1,N], \tag{12d}$$
$$\mathbf{W}_m \succeq \mathbf{0}, \quad \mathbf{W}_{Fn} \succeq \mathbf{0}, \ \mathbf{Z}_0 \succeq \mathbf{0},$$
$$\mathbf{Z}_n \succeq \mathbf{0}, \quad m \in [1,M], \ n \in [1,N], \tag{12e}$$
$$\mathrm{Rank}(\mathbf{W}_m) = \mathrm{Rank}(\mathbf{W}_{Fn}) = 1,$$
$$m \in [1,M], \quad n \in [1,N], \tag{12f}$$

where $S_{EE}(k) = C_S^k/(P_k/\varpi + P_\varepsilon)$, $P_0 = \sum_{m=1}^{M} \mathrm{Tr}(\mathbf{W}_m) + \mathrm{Tr}(\mathbf{Z}_0)$, $P_n = \mathrm{Tr}(\mathbf{W}_{Fn}) + \mathrm{Tr}(\mathbf{Z}_n)$; $\varpi \in (0,1)$ is the ratio between the total RF output power and the DC input power while P_ε denotes the power loss in hardware; θ_n indicates the EH threshold of ER_n; P_M and P_F represent the maximum output power of the MBS and each FBS, respectively. Obviously, problem (12) is a non-convex problem and contains infinite constraints caused by CSI errors, which cannot be solved directly.

III. ROBUST TBF AND AN DESIGN IN CENTRALIZED DESIGN

In this section, we explore the joint design in a centralized approach. Specifically, FBSs first obtain their local CSIs including estimated CSIs and error information, and send them to the MBS. Then, the MBS acquires the global CSIs and solves the problem (12) to acquire the optimal TBF and AN matrixes. Finally, the MBS send the optimal design to the FBSs.

A. EQUIVALENT TRANSFORMATION OF ORIGINAL PROBLEM

In this section, we aim to transform the originally non-convex problem into a solvable form. To be more specific, with the aid of the SCA, first-order Taylor series expansion and SDR techniques, we first get rid of the non-convexity of some constraints; then, following the worst-case robustness philosophy to deal with the infinite constraints, we reformulate the worst-case SEE maximization problem into a QMI problem in light of S-procedure.

1) GETTING RID OF THE NON-CONVEXITY

To begin with, we aim to acquire the equivalent form of (12b). Thus, in combination with (9) and (10), it can be rewritten as

$$\min_{\substack{\{\mathbf{E}_{ERn}\},\{\mathbf{E}_{tERn}\},\\ \{\mathbf{E}_{MERn}\}}} \{\mathrm{Tr}(\mathbf{H}_{ERn}^H \mathbf{H}_{ERn} \mathbf{A}_n) + \sum_{t\ne n}^{N} \mathrm{Tr}(\mathbf{H}_{tERn}^H \mathbf{H}_{tERn} \mathbf{A}_t) + \mathrm{Tr}(\mathbf{H}_{MERn}^H \mathbf{H}_{MERn} \mathbf{B}) \} \ge B(\theta_n), \quad n \in [1,N], \tag{13}$$

where $B(\theta_n) = b - \ln[\frac{A[1+\exp(ab)]}{(\theta_n + A/\exp(ab))\exp(ab)} - 1]/a$. It should be noted that we donot consider noise in the equation (13) due to the small power of noise.

Further, to expose the hidden convexity of problem (12), we rewrite it as

$$\max_{\substack{\{\mathbf{W}_m\},\mathbf{Z}_0\\ \{\mathbf{W}_{Fn}\},\{\mathbf{Z}_n\},\{x,y_k\}}} \quad x \tag{14a}$$
$$s.t. \quad \min_{\substack{\{\mathbf{e}_m\}\{\mathbf{e}_{n,m}\},\{\mathbf{e}_{Fn}\},\\ \{\mathbf{e}_{tFn}\},\{\mathbf{e}_{MFn}\},\mathbf{E}_E,\{\mathbf{E}_{n,E}\}\\ \{\mathbf{E}_{ERn}\},\{\mathbf{E}_{tERn}\},\{\mathbf{E}_{MERn}\}}} \{\{C_S^k\} \ge x y_k\}, \quad k \in [0,N], \tag{14b}$$
$$\sum_{m=1}^{M} \mathrm{Tr}(\mathbf{W}_m) + \mathrm{Tr}(\mathbf{Z}_0) \le \varpi(y_0 - P_\varepsilon), \tag{14c}$$
$$\mathrm{Tr}(\mathbf{W}_{Fn}) + \mathrm{Tr}(\mathbf{Z}_n) \le \varpi(y_n - P_\varepsilon), \quad n \in [1,N], \tag{14d}$$
$$(12c)\ (12f), (13), \tag{14e}$$

where x and y_k are introduced variables. Intuitively, the equalities in (14b)$\sim$(14d) must hold at the optimal solution. Otherwise, the equalities can always be acquirable by increasing x and decreasing y_k, which will not change the optimal objective value.

To further deal with the non-convexity of (14b), we introduce some new auxiliary variables, i.e., $\{\alpha_{0m}\}_{m=1}^{M}$, β_0, $\{\chi_{0m}\}_{m=1}^{M}$, $\{\alpha_n\}_{n=1}^{N}$, $\{\beta_n\}_{n=1}^{N}$, $\{\chi_n\}_{n=1}^{N}$, and reformulate it as

$$\beta_0 - \sum_{m=1}^{M} \log_2(\frac{1}{\chi_{0m}}) \geq xy_0, \tag{15a}$$

$$\sum_{m=1}^{M} \log_2(\alpha_{0m}) \geq \beta_0, \tag{15b}$$

$$1 + \mathrm{SINR}_m \geq \alpha_{0m}, \quad m \in [1, M], \tag{15c}$$

$$\max_{\mathbf{E}_E, \{\mathbf{E}_{n,E}\}} \{|\mathbf{R}_{Em}|\} \leq \frac{1}{\chi_{0m}}, \quad m \in [1, M], \tag{15d}$$

$$\beta_n - \log_2(\frac{1}{\chi_n}) \geq xy_n, \quad n \in [1, N] \tag{15e}$$

$$\log_2(\alpha_n) \geq \beta_n, \quad n \in [1, N], \tag{15f}$$

$$1 + \mathrm{SINR}_{Fn} \geq \alpha_n, \quad n \in [1, N], \tag{15g}$$

$$\min_{\{\mathbf{E}_{ERn}\}, \{\mathbf{E}_{tERn}\}, \{\mathbf{E}_{MERn}\}} \{|\mathbf{R}_{\mathrm{EIR}n}|\} \leq \frac{1}{\chi_n}, \quad n \in [1, N]. \tag{15h}$$

After the above procedures, (14b) is equivalently transformed into many constraints in (15a)∼(15f). Fortunately, we notice that (15a)∼(15d) and (15e)∼(15h) have a similar structure so they can be processed in a similar way in the following.

Taking (15c) as an example, with the aid of slack variables $\{\phi_m\}_{m=1}^{M}$, we can equivalently transform it into

$$\min_{\{\mathbf{e}_m\}} \{\mathrm{Tr}(\mathbf{H}_m \mathbf{W}_m)\} \geq (\alpha_{0m} - 1)\phi_m, \tag{16a}$$

$$\max_{\{\mathbf{e}_m, \mathbf{e}_{n,m}\}} \{\mathrm{Tr}(\mathbf{H}_m \mathbf{C}_m) + \sum_{n=1}^{N} \mathrm{Tr}(\mathbf{H}_{n,m} \mathbf{A}_n) + \sigma^2\} \leq \phi_m, \tag{16b}$$

Similarly, (15g) can be rewritten as

$$\min_{\{\mathbf{e}_{Fn}\}} \{\mathrm{Tr}(\mathbf{H}_{Fn} \mathbf{W}_{Fn})\} \geq (\alpha_n - 1)\phi_{Fn}, \quad n \in [1, N], \tag{17a}$$

$$\min_{\{\mathbf{e}_{Fn}, \mathbf{e}_{tFn}, \mathbf{e}_{MFn}\}} \{\mathrm{Tr}(\mathbf{H}_{Fn} \mathbf{Z}_n) + \sum_{t \neq n}^{K} \mathrm{Tr}(\mathbf{H}_{tFn} \mathbf{A}_t) + \mathrm{Tr}(\mathbf{H}_{MFn} \mathbf{B}) + \sigma^2\} \leq \phi_{Fn}, \quad n \in [1, N], \tag{17b}$$

where $\{\phi_{Fn}\}_{n=1}^{N}$ are newly introduced variables. In what follows, we aim to acquire an equivalent form of (15d). Based on this consideration, we introduce the following Lemma 1.

Lemma 1 [24]: For any positive semi-definite matrix $\mathbf{A}$, the following inequality holds

$$|\mathbf{I} + \mathbf{A}| \geq 1 + \mathrm{Tr}(\mathbf{A}), \tag{18}$$

and the equality holds if and only if $\mathrm{Rank}(\mathbf{A}) \leq 1$.

Resorting to the Lemma 1, (15d) can be reformulated as

$$\max_{\mathbf{E}_E, \{\mathbf{E}_{n,E}\}} \{\mathrm{Tr}\{(\mathbf{H}_E^H \mathbf{C}_m \mathbf{H}_E + \sum_{n=1}^{N} \mathbf{H}_{n,E}^H \mathbf{A}_n \mathbf{H}_{n,E} + \sigma^2 \mathbf{I}_{N_E})^{-1} \times (\mathbf{H}_E^H \mathbf{W}_m \mathbf{H}_E)\} + 1\} \leq \frac{1}{\chi_{0m}}. \tag{19}$$

With the introduced auxiliary variables u_{0m}, v_{0m}, t_{0m}, x_{0m}, y_{0m} and z_{0m}, (19) can be further rewritten as

$$\chi_{0m} + u_{0m}\chi_{0m} \leq 1, \tag{20a}$$

$$\max_{\mathbf{E}_E, \{\mathbf{E}_{n,E}\}} \{\mathrm{Tr}(\mathbf{H}_E^H \mathbf{W}_m \mathbf{H}_E)\} \leq v_{0m}, \tag{20b}$$

$$v_{0m} \leq e^{x_{0m}}, \tag{20c}$$

$$u_{0m} \geq e^{y_{0m}}, \quad t_{0m} \geq e^{z_{0m}}, \tag{20d}$$

$$x_{0m} \leq y_{0m} + z_{0m}, \tag{20e}$$

$$t_{0m} \leq \min_{\mathbf{E}_E, \{\mathbf{E}_{n,E}\}} \{\mathrm{Tr}(\mathbf{H}_E^H \mathbf{C}_m \mathbf{H}_E + \sum_{n=1}^{N} \mathbf{H}_{n,E}^H \mathbf{A}_n \mathbf{H}_{n,E}) + \sigma^2\}. \tag{20f}$$

Similarly, (15h) can be reformulated as

$$\chi_n + u_{ERn}\chi_n \leq 1, \tag{21a}$$

$$\max_{\mathbf{E}_E, \{\mathbf{E}_{n,E}\}} \{\mathrm{Tr}(\mathbf{H}_{ERn}^H \mathbf{W}_{Fn} \mathbf{H}_{ERn})\} \leq v_{ERn}, \tag{21b}$$

$$v_{ERn} \leq e^{x_{ERn}}, \tag{21c}$$

$$u_{ERn} \geq e^{y_{ERn}}, \quad t_{ERn} \geq e^{z_{ERn}}, \tag{21d}$$

$$x_{ERn} \leq y_{ERn} + z_{ERn}, \tag{21e}$$

$$t_{ERn} \leq \min_{\substack{\{\mathbf{E}_{ERn}\}, \{\mathbf{E}_{tERn}\} \\ \{\mathbf{E}_{MERn}\}}} \{\Upsilon + \Gamma + \Lambda + \sigma^2\}, \tag{21f}$$

where u_{ERn}, v_{ERn}, t_{ERn}, x_{ERn}, y_{ERn} and z_{ERn} are introduced auxiliary variables.

At this moment, (14b) has been transformed into (15a), (15b), (16a),(16b), (20a)∼(20f), (15e), (15f), (17a), (17b), (21a)∼(21f). Meanwhile, (15a), (15e), (16a), (17a), (20a) and (21a) have a same form of $f(x, y) = xy$, where the coupling of variables all lie in the right hand side of a constraint with $\geq$. Thus, with the aid of the SCA algorithm, new reformulations of these six constraints can be acquired. Specifically, the upper bound of $f(x, y) = xy$ can be defined as $g_\zeta(x, y) = {}^{\zeta}\!/_{2}\, x^2 + {}^{1}\!/_{2\zeta}\, y^2$ for any ζ because $g_\zeta(x, y) \geq f(x, y)$ always holds. Moreover, if $\zeta = {}^{y}\!/_{x}$, $g_\zeta(x, y)$ is equivalent to $f(x, y)$. Therefore, in the q-th SCA iterative approximation, these six constraints can be respectively transformed into

$$\beta_0 + \sum_{m=1}^{M} \log_2(\chi_{0m}) \geq g_{\zeta_0}^{(q)}(x, y_0), \tag{22a}$$

$$\beta_n + \log_2(\chi_n) \geq g_{\zeta_n}^{(q)}(x, y_n), \quad n \in [1, N], \tag{22b}$$

$$\min_{\{\mathbf{e}_m\}} \{\mathrm{Tr}(\mathbf{H}_m \mathbf{W}_m)\} \geq g_{\zeta_{0m}}^{(q)}(\alpha_{0m}, \phi_m) - \phi_m, \quad m \in [1, M], \tag{22c}$$

$$\min_{\{\mathbf{e}_{Fn}\}} \{\mathrm{Tr}(\mathbf{H}_{Fn} \mathbf{W}_{Fn})\} \geq g_{\zeta_{Fn}}^{(q)}(\alpha_n, \phi_{Fn}) - \phi_{Fn}, \quad n \in [1, N], \tag{22d}$$

$$1 \geq \chi_{0m} + g_{\zeta_{0Mm}}^{(q)}(u_{0m}, \chi_{0m}), \quad m \in [1, M], \tag{22e}$$

$$1 \geq \chi_n + g_{\zeta_{ERn}}^{(q)}(u_{ERn}, \chi_n), \quad n \in [1, N], \tag{22f}$$

where $g^{(q)}_{\zeta_0}(x,y_0) = \frac{\zeta_0^{(q-1)}}{2}x^2 + \frac{1}{2\zeta_0^{(q-1)}}y_0^2$, $g^{(q)}_{\zeta_n}(x,y_n) = \frac{\zeta_n^{(q-1)}}{2}x^2 + \frac{1}{2\zeta_n^{(q-1)}}y_n^2$, $g^{(q)}_{\zeta_{0m}}(\alpha_{0m},\phi_m) = \frac{\zeta_{0m}^{(q-1)}}{2}\alpha_{0m}^2 + \frac{1}{2\zeta_{0m}^{(q-1)}}\phi_m^2$, $g^{(q)}_{\zeta_{Fn}}(\alpha_n,\phi_{Fn}) = \frac{\zeta_{Fn}^{(q-1)}}{2}\alpha_n^2 + \frac{1}{2\zeta_{Fn}^{(q-1)}}\phi_{Fn}^2$, $g^{(q)}_{\zeta_{0Mm}}(u_{0m},\chi_{0m}) = \frac{\zeta_{0Mm}^{(q-1)}}{2}u_{0m}^2 + \frac{1}{2\zeta_{0Mm}^{(q-1)}}\chi_{0m}^2$, $g^{(q)}_{\zeta_{ERn}}(u_{ERn},\chi_n) = \frac{\zeta_{ERn}^{(q-1)}}{2}u_{ERn}^2 + \frac{1}{2\zeta_{ERn}^{(q-1)}}\chi_n^2$; $\zeta_0^{(q-1)}$, $\zeta_n^{(q-1)}$, $\zeta_{0m}^{(q-1)}$, $\zeta_{Fn}^{(q-1)}$, $\zeta_{0Mm}^{(q-1)}$ and $\zeta_{ERn}^{(q-1)}$ are the $(q-1)$-th iterative approximation of y_0/x, y_n/x, ϕ_m/α_{0m}, ϕ_{Fn}/α_n, u_{0m}/χ_{0m} and u_{ERn}/χ_n, respectively.

Meanwhile, it is easily seen that the constraints in (20d) and (21d) are convex now. However, the constraints in (20c) and (21c) are still nonconvex due to the fact that if a constraint is such that a convex function is smaller than or equal to a concave function, then the constraint is a convex constraint. Fortunately, the functions $e^{x_{0m}}$ and $e^{x_{ERn}}$ on the right-hand sides of (20c) and (21c) are convex. To this end, we define $x_{0m}^{(q)}$ and $x_{ERn}^{(q)}$ as the $(q-1)$ th iteration of the variables x_{0m} and x_{ERn} for an iterative algorithm given below. By applying first-order Taylor series expansions on $e^{x_{0m}}$ and $e^{x_{ERn}}$, i.e., $e^{x_{0m}^{(q)}}(x_{0m} - x_{0m}^{(q)} + 1) \le e^{x_{0m}}$ and $e^{x_{ERn}^{(q)}}(x_{ERn} - x_{ERn}^{(q)} + 1) \le e^{x_{ERn}}$, the linearization of nonconvex constraints in (20c) and (21c) are given by [25]

$$v_{0m} \le e^{x_{0m}^{(q)}}(x_{0m} - x_{0m}^{(q)} + 1), \quad m \in [1,M], \tag{23a}$$

$$v_{ERn} \le e^{x_{ERn}^{(q)}}(x_{ERn} - x_{ERn}^{(q)} + 1), \quad n \in [1,N]. \tag{23b}$$

After the above procedures, we further find that (15b), (15f), (20d), (21d), (22a) and (22b) have logarithmic or exponential forms, which results to high computational complexity and the decrease of solving efficiency. Furthermore, (20d) and (21d) can be further reformulated as

$$\ln(u_{0m}) \ge y_{0m}, \quad \ln(t_{0m}) \ge z_{0m}, \ m \in [1,M], \tag{24a}$$

$$\ln(u_{ERn}) \ge y_{ERn}, \quad \ln(t_{ERn}) \ge z_{ERn}, \ n \in [1,N]. \tag{24b}$$

Therefore, to facilitate tractability, we convert these logarithmic constraints into convex second order cone (SOC) representable constraints in the following.

Transformation of (15f) and (22b): (15f) is equivalent to

$$\alpha_n \log_2(\alpha_n) \ge \alpha_n \beta_n, n \in [1,N], \tag{25}$$

The left hand side of $\ge$ is convex, so we have the following inequality

$$\begin{aligned}\alpha_n \log_2(\alpha_n) \ge \{&\alpha_n^{(q-1)}\log_2(\alpha_n^{(q-1)}) + (\alpha_n - \alpha_n^{(q-1)}) \\ &\times (1 + \log_2(\alpha_n^{(q-1)})) \\ &= \alpha_n(1 + \log_2(\alpha_n^{(q-1)})) - \alpha_n^{(q-1)}\},\end{aligned} \tag{26}$$

where $\alpha_n^{(q-1)}$ is the $(q-1)$-th approximation of α_n. Therefore, (15f) in the q-th iteration can be replaced with

$$\alpha_n(1 + \log_2(\alpha_n^{(q-1)})) - \alpha_n^{(q-1)} \ge \alpha_n \beta_n, \quad n \in [1,N]. \tag{27}$$

Further, (27) can be written as the following SOC representable form to improve efficiency

$$\begin{aligned}&\left\| [2\sqrt{\alpha_n^{(q-1)}}, (\alpha_n + \beta_n - 1 - \log_2(\alpha_n^{(q-1)}))] \right\|_2 \\ &\quad \le \alpha_n + 1 + \log_2(\alpha_n^{(q-1)}) - \beta_n, \quad n \in [1,N].\end{aligned} \tag{28}$$

Similarly, (22b) can be rewritten as

$$\begin{aligned}&\left\| [2\sqrt{\chi_n^{(q-1)}}, (\chi_n + g^q_{\zeta_n}(x,y_n) - \beta_n - 1 - \log_2(\chi_n^{(q-1)}))] \right\|_2 \\ &\le \chi_n + 1 + \log_2(\chi_n^{(q-1)}) - g^q_{\zeta_n}(x,y_n) + \beta_n, \quad n \in [1,N].\end{aligned} \tag{29}$$

Transformation of (24a) and (24b): Following the similar procedures, these two constraints can be transformed into

$$\begin{aligned}&\left\| [2\sqrt{u_{0m}^{(q-1)}}, (u_{0m} + y_{0m} - 1 - \ln(u_{0m}^{(q-1)}))] \right\|_2 \\ &\quad \le u_{0m} + 1 + \ln(u_{0m}^{(q-1)}) - y_{0m}, \quad m \in [1,M],\end{aligned} \tag{30a}$$

$$\begin{aligned}&\left\| [2\sqrt{t_{0m}^{(q-1)}}, (t_{0m} + z_{0m} - 1 - \ln(t_{0m}^{(q-1)}))] \right\|_2 \\ &\quad \le t_{0m} + 1 + \ln(t_{0m}^{(q-1)}) - z_{0m}, \quad m \in [1,M],\end{aligned} \tag{30b}$$

$$\begin{aligned}&\left\| [2\sqrt{u_{ERn}^{(q-1)}}, (u_{ERn} + y_{ERn} - 1 - \ln(u_{ERn}^{(q-1)}))] \right\|_2 \\ &\quad \le u_{ERn} + 1 + \ln(u_{ERn}^{(q-1)}) - y_{ERn}, \quad n \in [1,N],\end{aligned} \tag{31a}$$

$$\begin{aligned}&\left\| [2\sqrt{t_{ERn}^{(q-1)}}, (t_{ERn} + z_{ERn} - 1 - \ln(t_{ERn}^{(q-1)}))] \right\|_2 \\ &\quad \le t_{ERn} + 1 + \ln(t_{ERn}^{(q-1)}) - z_{ERn}, \quad n \in [1,N].\end{aligned} \tag{31b}$$

Transformation of (15b) and (22a): The procedures to reformulate (15b) and (22a) are slightly different from the above steps. Specifically, some slack variables $\{a_m\}_{m=1}^M$ and $\{b_m\}_{m=1}^M$ need to be introduced, then they can be finally reformulated as

$$\sum_{m=1}^{M} a_m \ge \beta_0, \tag{32a}$$

$$\begin{aligned}&\left\| [2\sqrt{\alpha_{0m}^{(q-1)}}, (\alpha_{0m} + b_m - 1 - \log_2(\alpha_{0m}^{(q-1)}))] \right\|_2 \\ &\quad \le \alpha_{0m} + 1 + \log_2(\alpha_{0m}^{(q-1)}) - a_m, \quad m \in [1,M],\end{aligned} \tag{32b}$$

$$\sum_{m=1}^{M} b_m \ge g^{(q)}_{\zeta_0}(x,y_0) - \beta_0, \tag{33a}$$

$$\begin{aligned}&\left\| [2\sqrt{\chi_{0m}^{(q-1)}}, (\chi_{0m} + c_m - 1 - \log_2(\chi_{0m}^{(q-1)}))] \right\|_2 \\ &\quad \le \chi_{0m} + 1 + \log_2(\chi_{0m}^{(q-1)}) - b_m, \quad m \in [1,M].\end{aligned} \tag{33b}$$

Resorting to the SDR technique to abandon the rank-one constraint, then in the q-th iterative approximation, original

problem (12) can be expressed as

$$\max_{\substack{\{\mathbf{W}_m\},\mathbf{Z}_0,\\ \{\mathbf{W}_{Fn}\},\{\mathbf{Z}_n\},x}} \quad x \tag{34a}$$

$$\text{s.t.} \quad \left\{\begin{array}{c}(12c)\sim(12e),\ (13),\ (14c),\ (14d),\{(33a),(33b)\},\ (32a),\\ (32b),(16b),(22c),(22e),(20b),(23a),(20e),(20f),(30a),\\ (30b),(29),(28),\ (17b),(22d),(22f),(21b),(23b),(31a),\\ (31b),(21e),(21f)\end{array}\right\} \tag{34b}$$

where (33a), (33b), (32a), (32b), (16b), (22c), (22e), (20b), (23), (20e), (20f), (30a), (30b), (29), (28), (17b), (22d) and (22f), (21b), (23b), (31a), (31b), (21e), (21f) are the reformulations of (15a) (15h), respectively. If the CSI errors are not considered in this problem (34), then it is convex and can be solved efficiently via state-of-the-art conic solvers, such as CVX. Nevertheless, due to the existence of CSI errors, there are infinite constraints so problem (34) still cannot be solved directly. Therefore, we continue to process the infinite constraints of this problem in the following section.

2) REFORMULATION OF INFINITE CONSTRAINTS CAUSED BY CSI ERRORS

Obviously, (13), (16b), (22c), (20b), (20f), (17b), (22d) and (21b), (21f) contain infinite constraints caused by CSI uncertainty. To address this issue, we first introduce some auxiliary variables $\{\{\xi_{tERn}\}_{t\neq n}^N\}_{n=1}^N$, $\{\xi_{MERn}\}_{n=1}^N$, $\{\{\xi_{n,m}\}_{n=1}^N\}_{m=1}^M$, $\{\xi_{n,E}\}_{n=1}^N$, $\{\{\xi_{tFn}\}_{t\neq n}^N\}_{n=1}^N$ and $\{\xi_{MFn}\}_{n=1}^N$, then reformulate them with the aid of equivalent transformation and S-procedure.

Reformulation of (13), (21b) and (21f): For these three constraints, we notice that they contain some common items, so we can first reformulate (13) as (21f) and the following expressions

$$\min_{\mathbf{E}_{ERn}} \{\mathrm{Tr}(\mathbf{H}_{ERn}^H\mathbf{W}_{Fn}\mathbf{H}_{ERn})\} \geq B(\theta_n) - t_{ERn} + \sigma^2, \tag{35a}$$

$$\min_{\mathbf{E}_{ERn}} \{\mathrm{Tr}(\mathbf{H}_{ERn}^H\mathbf{H}_{ERn}\mathbf{Z}_n)\} \geq t_{ERn} - \sum_{t\neq n}^N \xi_{tERn} - \xi_{MERn} - \sigma^2, \tag{35b}$$

$$\min_{\mathbf{E}_{tERn}} \{\mathrm{Tr}(\mathbf{H}_{tERn}^H\mathbf{H}_{tERn}\mathbf{A}_t)\} \geq \xi_{tERn}, \tag{36a}$$

$$\min_{\mathbf{E}_{MERn}} \{\mathrm{Tr}(\mathbf{H}_{MERn}^H\mathbf{B}\mathbf{H}_{MERn})\} \geq \xi_{MERn}. \tag{36b}$$

To further deal with these constraints, we rewrite (36a) as

$$\mathrm{Tr}(\hat{\mathbf{H}}_{ERn}^H\mathbf{Z}_n\hat{\mathbf{H}}_{ERn} + \mathbf{E}_{ERn}^H\mathbf{Z}_n\mathbf{E}_{ERn} + 2\,\mathrm{Re}\{\mathbf{E}_{ERn}^H\mathbf{Z}_n\hat{\mathbf{H}}_{ERn}\}) \geq t_{ERn} - \sum_{t\neq n}^N \xi_{tERn} - \xi_{MERn} - \sigma^2 \tag{37}$$

Applying the following matrix identities

$$\begin{aligned}\mathrm{vec}(\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3) &= (\mathbf{M}_3^T\otimes\mathbf{M}_1)\mathrm{vec}(\mathbf{M}_2),\\ \mathrm{Tr}(\mathbf{M}_1^H\mathbf{M}_2) &= \mathrm{vec}(\mathbf{M}_1)^H\mathrm{vec}(\mathbf{M}_2)\end{aligned} \tag{38}$$

to reformulate (37) as

$$\hat{\mathbf{h}}_{ERn}^H\mathbf{Z}_n\hat{\mathbf{h}}_{ERn} + 2\,\mathrm{Re}\{\mathbf{e}_{ERn}^H\mathbf{Z}_n\hat{\mathbf{h}}_{ERn}\} + \mathbf{e}_{ERn}^H\mathbf{Z}_n\mathbf{e}_{ERn} \geq t_{ERn} - \sum_{t\neq n}^N \xi_{tERn} - \xi_{MERn} - \sigma^2. \tag{39}$$

where $\hat{\mathbf{h}}_{ERn} = \mathrm{vec}(\hat{\mathbf{H}}_{ERn})$, $\mathbf{e}_{ERn} = \mathrm{vec}(\mathbf{E}_{ERn})$, $\mathbf{D}_n = \mathbf{I}_{N_{ER}}\otimes\mathbf{A}_n$. To further process (39), we introduce the following Lemma 2.

Lemma 2 (S-Procedure) [10]: Let $g_k(x) = x^HF_kx + 2\,\mathrm{Re}\{g_k^Hx\} + q_k$, $k = 1, 2$, where $F_k \in H^L$, $g_k \in C^L$, $q_k \in R$. When $g_1(\widetilde{x}) < 0$, then $g_1(\widetilde{x}) \leq 0 \Rightarrow g_2(\widetilde{x}) \leq 0$ holds if and only if there is $\lambda \geq 0$ and

$$\lambda\begin{bmatrix}F_1 & g_1\\ g_1^H & q_1\end{bmatrix} - \begin{bmatrix}F_2 & g_2\\ g_2^H & q_2\end{bmatrix} \geq 0. \tag{40}$$

Resorting the Lemma 2, (36a) can be finally reformulated as

$$\begin{bmatrix}\lambda_{ERn}\mathbf{I}_{N_FN_{ER}} + \mathbf{Z}_n & \mathbf{Z}_n\hat{\mathbf{h}}_{ERn}\\ \hat{\mathbf{h}}_{ERn}^H\mathbf{Z}_n & \hat{\mathbf{h}}_{ERn}^H\mathbf{Z}_n\hat{\mathbf{h}}_{ERn} + e_{ERn}\end{bmatrix} \geq 0, \tag{41}$$

where $\lambda_{ERn} \geq 0$ is the auxiliary variable, $e_{ERn} = \sum_{t\neq n}^N \xi_{tERn} + \xi_{MERn} + \sigma^2 - t_{ERn} - \lambda_{ERn}\varepsilon_{ERn}^2$. Similarly, (36b), (36c) and (35) are respectively reformulated as

$$\begin{bmatrix}\lambda_{tERn}\mathbf{I}_{N_FN_{ER}} + \mathbf{D}_t & \mathbf{D}_t\hat{\mathbf{h}}_{tERn}\\ \hat{\mathbf{h}}_{tERn}^H\mathbf{D}_t & \hat{\mathbf{h}}_{tERn}^H\mathbf{D}_t\hat{\mathbf{h}}_{tERn} - e_{tERn}\end{bmatrix} \geq 0, \tag{42a}$$

$$\begin{bmatrix}\lambda_{MERn}\mathbf{I}_{N_MN_{ER}} + \mathbf{F} & \mathbf{F}\hat{\mathbf{h}}_{MERn}\\ \hat{\mathbf{h}}_{MERn}^H\mathbf{F} & \hat{\mathbf{h}}_{MERn}^H\mathbf{F}\hat{\mathbf{h}}_{MERn} - e_{MERn}\end{bmatrix} \geq 0, \tag{42b}$$

$$\begin{bmatrix}\lambda_{EERn}\mathbf{I}_{N_FN_{ER}} + \mathbf{W}_{Fn} & \mathbf{W}_{Fn}\hat{\mathbf{h}}_{ERn}\\ \hat{\mathbf{h}}_{ERn}^H\mathbf{W}_{Fn} & \hat{\mathbf{h}}_{ERn}^H\mathbf{W}_{Fn}\hat{\mathbf{h}}_{ERn} + e_{EERn}\end{bmatrix} \geq 0, \tag{42c}$$

where $\lambda_{tERn} \geq 0$, $\lambda_{MERn} \geq 0$ and $\lambda_{EERn} \geq 0$ are auxiliary variables; $\hat{\mathbf{h}}_{tERn} = \mathrm{vec}(\hat{\mathbf{H}}_{tERn})$, $\hat{\mathbf{h}}_{MERn} = \mathrm{vec}(\hat{\mathbf{H}}_{MERn})$, $\mathbf{D}_t = \mathbf{I}_{N_{ER}}\otimes\mathbf{A}_t$, $\mathbf{F} = \mathbf{I}_{N_{ER}}\otimes\mathbf{B}$, $e_{tERn} = \xi_{tERn} + \lambda_{tERn}\varepsilon_{tERn}^2$, $e_{MERn} = \xi_{MERn} + \lambda_{MERn}\varepsilon_{MERn}^2$, $e_{EERn} = t_{ERn} - B(\theta_n) - \sigma^2 - \lambda_{EERn}\varepsilon_{ERn}^2$.

Reformulation of (16b) and (22c): Applying the Lemma 2, we can respectively transform them into

$$\begin{bmatrix}\lambda_m\mathbf{I}_{N_M} - \mathbf{C}_m & -\mathbf{C}_m\hat{\mathbf{h}}_m\\ -\hat{\mathbf{h}}_m^H\mathbf{C}_m & e_m - \hat{\mathbf{h}}_m^H\mathbf{C}_m\hat{\mathbf{h}}_m\end{bmatrix} \geq 0, \tag{43a}$$

$$\begin{bmatrix}\lambda_{n,m}\mathbf{I}_{N_F} - \mathbf{A}_n & -\mathbf{A}_n\hat{\mathbf{h}}_{n,m}\\ -\hat{\mathbf{h}}_{n,m}^H\mathbf{A}_n & e_{n,m} - \hat{\mathbf{h}}_{n,m}^H\mathbf{A}_n\hat{\mathbf{h}}_{n,m}\end{bmatrix} \geq 0, \tag{43b}$$

$$\begin{bmatrix}\lambda_{Mm}\mathbf{I}_{N_M} + \mathbf{W}_m & \mathbf{W}_m\hat{\mathbf{h}}_m\\ \hat{\mathbf{h}}_m^H\mathbf{W}_m & \hat{\mathbf{h}}_m^H\mathbf{W}_m\hat{\mathbf{h}}_m + e_{Mm}\end{bmatrix} \geq 0, \tag{43c}$$

where $\lambda_m \geq 0$, $\lambda_{n,m} \geq 0$ and $\lambda_{Mm} \geq 0$; $e_m = \phi_m - \sigma^2 - \sum_{n=1}^{N} \xi_{n,m} - \lambda_m \varepsilon_m^2$, $e_{n,m} = \xi_{n,m} - \lambda_{n,m}\varepsilon_{n,m}^2$, $e_{Mm} = \phi_m - \lambda_{Mm}\varepsilon_m^2 - g_{\zeta_{0m}}^{(q)}(\alpha_{0m}, \phi_m)$.

Reformulation of (20b) and (20f): Following the similar procedures as (36a), they can be finally formulated as

$$\begin{bmatrix} \lambda_{Em}\mathbf{I}_{N_M N_E} - (\mathbf{I}_{N_E} \otimes \mathbf{W}_m) & -(\mathbf{I}_{N_E} \otimes \mathbf{W}_m)\hat{\mathbf{h}}_E \\ -\hat{\mathbf{h}}_E^H(\mathbf{I}_{N_E} \otimes \mathbf{W}_m) & e_{Em} - \hat{\mathbf{h}}_E^H(\mathbf{I}_{N_E} \otimes \mathbf{W}_m)\hat{\mathbf{h}}_E \end{bmatrix} \geq 0, \tag{44a}$$

$$\begin{bmatrix} \lambda_{n,E}\mathbf{I}_{N_F N_{ER}} + \mathbf{D}_n & \mathbf{D}_n\hat{\mathbf{h}}_{n,E} \\ \hat{\mathbf{h}}_{n,E}^H\mathbf{D}_n & e_{n,E} + \hat{\mathbf{h}}_{n,E}^H\mathbf{D}_n\hat{\mathbf{h}}_{n,E} \end{bmatrix} \geq 0, \tag{44b}$$

$$\begin{bmatrix} \lambda_{EEm}\mathbf{I}_{N_M N_E} + (\mathbf{I}_{N_E} \otimes \mathbf{C}_m) & (\mathbf{I}_{N_E} \otimes \mathbf{C}_m)\hat{\mathbf{h}}_E \\ \hat{\mathbf{h}}_E^H(\mathbf{I}_{N_E} \otimes \mathbf{C}_m) & e_{EEm} + \hat{\mathbf{h}}_E^H(\mathbf{I}_{N_E} \otimes \mathbf{C}_m)\hat{\mathbf{h}}_E \end{bmatrix} \geq 0, \tag{44c}$$

where $\hat{\mathbf{h}}_E = \mathrm{vec}(\hat{\mathbf{H}}_E)$, $\hat{\mathbf{h}}_{n,E} = \mathrm{vec}(\hat{\mathbf{H}}_{n,E})$; $\lambda_{Em} \geq 0$, $\lambda_{n,E} \geq 0$ and $\lambda_{EEm} \geq 0$; $e_{Em} = v_{0m} - \lambda_{Em}\varepsilon_E^2$, $e_{n,E} = -\xi_{n,E} - \lambda_{n,E}\varepsilon_{n,E}^2$, $e_{EEm} = \sum_{n=1}^{N} \xi_{n,E} + \sigma^2 - \lambda_{EEm}\varepsilon_E^2 - t_{0m}$.

Reformulation of (17b) and (22d): Similar to the steps as (16b) and (22c), they can be finally transformed into

$$\begin{bmatrix} \lambda_{Fn}\mathbf{I}_{N_F} - \mathbf{Z}_n & -\mathbf{Z}_n\hat{\mathbf{h}}_{Fn} \\ -\hat{\mathbf{h}}_{Fn}^H\mathbf{Z}_n & e_{Fn} - \hat{\mathbf{h}}_{Fn}^H\mathbf{Z}_n\hat{\mathbf{h}}_{Fn} \end{bmatrix} \geq 0, \tag{45a}$$

$$\begin{bmatrix} \lambda_{tFn}\mathbf{I}_{N_F} - \mathbf{A}_t & -\mathbf{A}_t\hat{\mathbf{h}}_{tFn} \\ -\hat{\mathbf{h}}_{tFn}^H\mathbf{A}_t & e_{tFn} - \hat{\mathbf{h}}_{tFn}^H\mathbf{A}_t\hat{\mathbf{h}}_{tFn} \end{bmatrix} \geq 0, \tag{45b}$$

$$\begin{bmatrix} \lambda_{MFn}\mathbf{I}_{N_M} - \mathbf{B} & -\mathbf{B}\hat{\mathbf{h}}_{MFn} \\ -\hat{\mathbf{h}}_{MFn}^H\mathbf{B} & e_{MFn} - \hat{\mathbf{h}}_{MFn}^H\mathbf{B}\hat{\mathbf{h}}_{MFn} \end{bmatrix} \geq 0, \tag{45c}$$

$$\begin{bmatrix} \lambda_{FFn}\mathbf{I}_{N_F} + \mathbf{W}_{Fn} & \mathbf{W}_{Fn}\hat{\mathbf{h}}_{Fn} \\ \hat{\mathbf{h}}_{Fn}^H\mathbf{W}_{Fn} & \hat{\mathbf{h}}_{Fn}^H\mathbf{W}_{Fn}\hat{\mathbf{h}}_{Fn} + e_{FFn} \end{bmatrix} \geq 0, \tag{45d}$$

where $\lambda_{Fn} \geq 0$, $\lambda_{tFn} \geq 0$, $\lambda_{MFn} \geq 0$ and $\lambda_{FFn} \geq 0$; $e_{Fn} = \phi_{Fn} - \sigma^2 - \sum_{t\neq n}^{N} \xi_{tFn} - \xi_{MFn} - \lambda_{Fn}\varepsilon_{Fn}^2$, $e_{tFn} = \xi_{tFn} - \lambda_{tFn}\varepsilon_{tFn}^2$, $e_{MFn} = \xi_{MFn} - \lambda_{MFn}\varepsilon_{MFn}^2$,$e_{FFn} = \phi_{Fn} - \lambda_{FFn}\varepsilon_{Fn}^2 - g_{\zeta_{Fn}}^{(q)}(\alpha_n, \phi_{Fn})$.

So far, in the q-th iterative approximation, problem (12) can be finally expressed as

$$\max_{\substack{\{\mathbf{W}_m\},\mathbf{Z}_0,\\ \{\mathbf{W}_{Fn}\},\{\mathbf{Z}_n\},x}} \quad x \tag{46a}$$

$$\text{s.t.} \left\{ \begin{array}{c} \text{(12c),(12d),(12e), (14c), (14d), \{(33a),(33b)\},} \\ \text{\{(32a),(32b)\},(43a),(43b),(43c),(22e),(23a),(20e),} \\ \text{(30a),(30b),(44a),(44b), (44c),(29), (28), (45a),} \\ \text{(45b),(45c),(45d),(22f),(23b),(31a),(31b),(21e), (41),} \\ \text{(42a),(42b),(42c)} \end{array} \right\} \tag{46b}$$

which can be solved efficiently via state-of-the-art conic solvers, such as CVX [26]. However, it should be noticed that the rank-one constraint has been abandoned by the SDR technique, which means that the result of the rank-relaxed problem serves as a performance upper bound for the original problem if the rank of the optimal beamforming solution is not one. Fortunately, we have the following proposition 1 to prove that applying the SDR technique is tight, where the rank of one beamforming vectors can always be acquired.

Proposition 1: If problem (46) is feasible, and $\{\mathbf{W}_m^*\}$, $\{\mathbf{W}_{Fn}^*\}$ are the optimal beamforming solutions, there must be $\mathrm{Rank}(\mathbf{W}_m^*) = \mathrm{Rank}(\mathbf{W}_{Fn}^*) = 1, \forall m \in [1, M], n \in [1, N]$.

Proof: Please see the Appendix.

B. THE CENTRALIZED ALGORITHM WITH SCA ITERATION

In this section, Algorithm 1 concludes the steps of proposed centralized algorithm with SCA iteration for solving problem (12); subsequently, the convergence, computational complexity and information change between coordinated BSs in the algorithm are analyzed, respectively. For the sake of simplicity, let us use $\{\vartheta_j\}_{j=1}^{5M+6N+1}$ to represent some auxiliary variables, i.e., x, $\{y_k\}_{k=0}^{N}$, $\{\alpha_{0m}, \phi_m\}_{m=1}^{M}$, $\{\alpha_n, \phi_{Fn}\}_{n=1}^{N}$, $\{\chi_{0m}, u_{0m}, t_{0m}\}_{m=1}^{M}$ and $\{\chi_n, u_{ERn}, t_{ERn}\}_{n=1}^{N}$. These variables should be given initial value to start the SCA iteration and updated during the iteration process.

Algorithm 1 The Centralized Algorithm With SCA Iteration

Initialization: set iteration index $q = 0$, maximum iteration number $Q_{\max}$ and feasible points $\{\vartheta_j^{(0)}\}_{j=1}^{5M+6N+1}$;

Repeat:

solve the problem (46) and obtain their optimal solutions $\{\vartheta_j^*\}_{j=1}^{5M+6N+1}$;

update the feasible points $\{\vartheta_j^{(q+1)}\}_{j=1}^{5M+6N+1} = \{\vartheta_j^*\}_{j=1}^{5M+6N+1}$;

$q = q + 1$;

Until: $q = Q_{\max}$ or algorithm converges;

Output: Optimal beamforming and AN matrixes $\{\mathbf{W}_m^*\}_{m=1}^{M}$, $\{\mathbf{W}_{Fn}^*\}_{n=1}^{N}$, $\{\mathbf{Z}_k^*\}_{k=0}^{N}$, and optimal SEE objective value x^*.

1) ALGORITHM CONVERGENCE

For the $(q+1)$-th iteration, its initial feasible points can be acquired by $\{\vartheta_j^{(q+1)}\}_{j=1}^{5M+6N+1} = \{\vartheta_j^*\}_{j=1}^{5M+6N+1}$, where $\{\vartheta_j^*\}_{j=1}^{5M+6N+1}$ are the corresponding optimal variables obtained in the q-th iteration. Intuitively, $\{\vartheta_j^{(q+1)}\}_{j=1}^{5M+6N+1}$ must be the feasible points to start the $(q+1)$-th SCA iteration because they satisfy corresponding constraints in problem (46). Moreover, with the maximization objective, the corresponding objective value is no less than that obtained in the q-th iteration. Meanwhile, due to the power constraint of the MBS and FBSs, the SEE must has an upper bound. Therefore, the objective value generated by Algorithm 1 must converge to a stable and highest value with the increase of q.

TABLE 1. Computational complexity analysis.

Methods	Variables		PSD constraints PSD constraints	SOC constraints SOC constraints
	Design variables	Slack variables		
Algorithm 1	a_1	a_2	$(a_3, 2M)$, (a_4, N) (a_5, N), (a_6, a_7) (a_8, a_9), (a_{10}, a_{11}) $(1, a_{12})$,$(N_M, 2N)$ $(N_F, M+1)$	$4M+4N$
Algorithm 2 (MBS)	a_{13}	a_{14}	$(a_3, 2M)$,(a_4, N) (a_8, a_9),$(N_M, 2N)$, $(1, a_{12})$	$4M$
Algorithm 2 (FBS)	$2N_F^2$	$4N+M+13$	$(a_5, 1)$,$(a_6, N+2)$ $(a_{10}, a_{11}/N)$, $(N_F, M+1)$,$(1, a_{12})$	4

2) COMPUTATIONAL COMPLEXITY

The computational complexity for solving problem (46) based on Algorithm 1 is briefly discussed here. Since the proposed Algorithm 1 is based on SDP, the computational complexity is determined significantly by the number and size of variables (i.e., design variables and slack variables) and constraints (i.e., PSD constraints and slack constraints) [27]. Especially, the major computational complexity for problem (46) is summarized in Table 1.

3) INFORMATION CHANGE

To make the MBS acquire the global CSIs, each FBS should send the CSIs and CSI error bounds of their local users to the MBS. In other words, each FBS should send $NN_F(1+N_F)$ complex values and 2 real values to the MBS. After obtaining the optimal design, the MBS will send optimal beamforming and AN vectors, i.e., $2N_F$ complex values to each FBS, respectively.

In the Table 1, some parameters are respectively expressed as: $a_1 = (M+1)N_M^2+2NN_F^2$, $a_2 = 4N^2+NM+13M+15N+3$, $a_3 = N_M N_E+1$, $a_4 = N_M N_{ER}+1$, $a_5 = N_F N_E+1$, $a_6 = N_F N_{ER}+1$, $a_7 = N^2+2N$, $a_8 = N_M+1$, $a_9 = 2M+N^2-N$, $a_{10} = N_F+1$, $a_{11} = N^2+N+NM$, $a_{12} = 2N^2+NM+11M+12N+4$, $a_{13} = a_1-2NN_F^2$, $a_{14} = a_2-4N^2-NM-13N$.

IV. ROBUST TBF AND AN DESIGN IN DISTRIBUTED DESIGN WITH ADMM

In previous section, the MBS should gather all the CSIs of coordinated BSs to perform the centralized algorithm. Nevertheless, in some scenes, it may be difficult to find a centralized processing BS, such as: (1) The BSs have low trust relation with each other so they are not willing to exchange their local CSIs to other BSs; (2) The difference between BS's processing capability is minor so the centralized approach may pose heavy overhead to the centralized processing BS. Therefore, a distributed approach is developed where each BS only needs to deal with its local CSIs is of significant necessity. In this section, we first explore the robust distributed design with ADMM. Subsequently, the ADMM-based distributed approach is concluded.

A. ROBUST DISTRIBUTED DESIGN WITH ADMM

In ADMM, three mainly key steps should be executed: (1) To guarantee the convergence, the penalty augmented problem is constructed while a penalty parameter is set; (2) The decomposition is conducted with dual variables introduced to make the problem be solved in parallel at different BSs; (3) Three kinds of variables are updated in an iterative manner until the solution converges to an expected precision [27].

We first introduce the following slack variables as

$$\xi_{tERn} = \tilde{\xi}_{tERn} = X_{tERn}, \quad t \in [1,N],\ n \in [1,N],\ t \neq n, \tag{47a}$$

$$\xi_{MERn} = \tilde{\xi}_{MERn} = X_{MERn}, \quad n \in [1,N], \tag{47b}$$

$$\xi_{n,m} = \tilde{\xi}_{n,m} = X_{n,m}, \quad n \in [1,N],\ m \in [1,M], \tag{47c}$$

$$\xi_{n,E} = \tilde{\xi}_{n,E} = X_{n,E}, \quad n \in [1,N], \tag{47d}$$

$$\xi_{tFn} = \tilde{\xi}_{tFn} = X_{tFn}, \quad t \in [1,N],\ n \in [1,N],\ t \neq n, \tag{47e}$$

$$\xi_{MFn} = \tilde{\xi}_{MFn} = X_{MFn}, \quad n \in [1,N], \tag{47f}$$

where all the variables above represent the inter-cell interference (ICI) imposed by a BS on a user of another BS. More specifically, $\{\xi_{tERn}, \tilde{\xi}_{tERn}\}$, $\{\xi_{MERn}, \tilde{\xi}_{MERn}\}$, $\{\xi_{n,m}, \tilde{\xi}_{n,m}\}$, $\{\xi_{n,E}, \tilde{\xi}_{n,E}\}$, $\{\xi_{tFn}, \tilde{\xi}_{tFn}\}$ and $\{\xi_{MFn}, \tilde{\xi}_{MFn}\}$ represent a pair of variables expressing the same ICI, respectively. Meanwhile, the element with a hat and that without a hat in each pair of variables are the local variable stored at the BS of interfered user and that stored at the interfering BS, respectively. For example, ξ_{tERn} stored at the FBS_t denotes the interference from FBS_t to ER_n while $\tilde{\xi}_{tERn}$ stored at the FBS_n represents the same ICI. Moreover, the slack variables X_{tERn}, X_{MERn}, $X_{n,m}$, $X_{n,E}$, X_{tFn} and X_{MFn} are introduced to ensure that the copies of the same ICI stored at different BSs equal to each other.

Based on above analysis and to facilitate tractability, we rewrite problem (46) in a decomposable form as (48), shown at the top of the next page, where $\lambda_m \geq 0$, $\lambda_{EEm} \geq 0$,

$$\max_{\substack{\{\mathbf{W}_m\},\mathbf{Z}_0,\\ \{\mathbf{W}_{Fn}\},\{\mathbf{Z}_n\},x}} \quad x \tag{48a}$$

$$\text{s.t.} \left\{ \begin{array}{c} (12c),(12d),(12e),\ (14c),\ (14d),\ \{(33a),(33b)\},\ \{(32a),(32b)\}, \\ (43b),(43c),(22e),(23a),(20e),(30a),(30b),(44a),(44b),(29), \\ (28),\ (45b),(45c),(45d),(22f),(23b),(31a),(31b),(21e),(42a), \\ (42b),(42c) \end{array} \right\}, \tag{48b}$$

$$\left[\begin{array}{cc} \lambda_m \mathbf{I}_{N_M} - \mathbf{C}_m & -\mathbf{C}_m \hat{\mathbf{h}}_m \\ -\hat{\mathbf{h}}_m^H \mathbf{C}_m & -\hat{\mathbf{h}}_m^H \mathbf{C}_m \hat{\mathbf{h}}_m - \tilde{e}_m \end{array} \right] \succeq 0, \tag{48c}$$

$$\left[\begin{array}{cc} \lambda_{EEm} \mathbf{I}_{N_M N_E} + (\mathbf{I}_{N_E} \otimes \mathbf{C}_m) & (\mathbf{I}_{N_E} \otimes \mathbf{C}_m)\hat{\mathbf{h}}_E \\ \hat{\mathbf{h}}_E^H (\mathbf{I}_{N_E} \otimes \mathbf{C}_m) & \tilde{e}_{EEm} + \hat{\mathbf{h}}_E^H (\mathbf{I}_{N_E} \otimes \mathbf{C}_m)\hat{\mathbf{h}}_E \end{array} \right] \succeq 0, \tag{48d}$$

$$\left[\begin{array}{cc} \lambda_{Fn} \mathbf{I}_{N_F} - \mathbf{Z}_n & -\mathbf{Z}_n \hat{\mathbf{h}}_{Fn} \\ -\hat{\mathbf{h}}_{Fn}^H \mathbf{Z}_n & \tilde{e}_{Fn} - \hat{\mathbf{h}}_{Fn}^H \mathbf{Z}_n \hat{\mathbf{h}}_{Fn} \end{array} \right] \succeq 0, \tag{48e}$$

$$\left[\begin{array}{cc} \lambda_{ERn} \mathbf{I}_{N_F N_{ER}} + \mathbf{Z}_n & \mathbf{Z}_n \hat{\mathbf{h}}_{ERn} \\ \hat{\mathbf{h}}_{ERn}^H \mathbf{Z}_n & \hat{\mathbf{h}}_{ERn}^H \mathbf{Z}_n \hat{\mathbf{h}}_{ERn} + \tilde{e}_{ERn} \end{array} \right] \succeq 0, \tag{48f}$$

$\lambda_{Fn} \geq 0$, $\lambda_{ERn} \geq 0$; $\tilde{e}_m = \phi_m - \sigma^2 - \sum_{n=1}^{N} \tilde{\xi}_{n,m} - \lambda_m \varepsilon_m^2$, $\tilde{e}_{EEm} = \sum_{n=1}^{N} \tilde{\xi}_{n,E} + \sigma^2 - \lambda_{EEm}\varepsilon_E^2 - t_{0m}$, $\tilde{e}_{Fn} = \phi_{Fn} - \sigma^2 - \sum_{t\neq n}^{N} \tilde{\xi}_{tFn} - \tilde{\xi}_{MFn} - \lambda_{Fn}\varepsilon_{Fn}^2$, $\tilde{e}_{ERn} = \sum_{t\neq n}^{N} \tilde{\xi}_{tERn} + \tilde{\xi}_{MERn} + \sigma^2 - t_{ERn} - \lambda_{ERn}\varepsilon_{ERn}^2$.

In the next, for the sake of simplicity, we denote all the BSs in the network as BS_k, $k \in [0, N]$, where BS_0 and $\text{BS}_n (n \in [1, N])$ are the MBS and FBS_n, respectively. To make it that the constraints can be handled locally at BS_k, all the constraints are put in the set $\mathbf{S}_k^{(q)}$, which is defined as

$$\mathbf{S}_k^{(q)} \triangleq \left\{ \mathbf{s}_k \middle| \left\{ \begin{array}{c} (12c),\ (12b),\ (12e),\ (14c),\ (14d),\ (34a), \\ (34b),\ (33a),\ (33b),\ (44b),\ (44c),\ (49c), \\ (22e),\ (23),\ (20e),\ (31a),\ (31b),\ (45a), \\ (45b),\ (49d),\ (30),\ (29),\ (46b),\ (46c), \\ (46d),\ (49e),\ (22f),\ (24),\ (32a),\ (32b), \\ (21e),\ (43a),\ (43b),\ (43c),\ (49f) \end{array} \right\} \right\}, \tag{49}$$

where

$$\mathbf{s}_n \triangleq [\left\{ \begin{array}{c} x_n, \mathbf{w}_{Fn}, \mathbf{z}_n, y_n, \alpha_n, \beta_n, \chi_n, \phi_{Fn}, u_{ERn}, \\ v_{ERn}, t_{ERn}, x_{ERn}, y_{ERn}, z_{ERn}, \{\lambda_i\}_{i=1}^{4N+M+1}, \mathbf{J}_n^T \end{array} \right\}]$$

includes all the local variables handled at FBS_n, and that for the MBS can be similarly denoted as

$$\mathbf{s}_0 \triangleq [\left\{ \begin{array}{c} x_0, \{\mathbf{w}_m\}_{m=1}^M, \mathbf{z}_0, y_0, \beta_0, \\ \{\phi_m, u_{0m}, v_{0m}, t_{0m}, x_{0m}, y_{0m}\}_{m=1}^M, \\ \{z_{0m}, \alpha_{0m}, \chi_{0m}\}_{m=1}^M, \{\lambda_j\}_{i=1}^{2N+4M}, \mathbf{J}_0^T \end{array} \right\}],$$

where

$$\{\lambda_i\}_{i=1}^{2N+M+4} = \left\{ \begin{array}{c} \lambda_{ERn}, \{\lambda_{tERn}\}_{t\neq n}^N, \lambda_{EERn}, \lambda_{Fn}, \\ \{\lambda_{tFn}\}_{t\neq n}^N, \lambda_{FFn}, \lambda_{n,E}, \{\lambda_{n,m}\}_{m=1}^M \end{array} \right\}, \tag{50a}$$

$$\mathbf{J}_n \triangleq \left[\left\{ \begin{array}{c} \{\xi_{n,m}\}_{m=1}^M, \xi_{n,E}, \xi_{nER1}, \xi_{nF1}, \cdots, \\ \xi_{nER(n-1)}, \xi_{nF(n-1)}, \xi_{nER(n+1)}, \xi_{nF(n+1)}, \\ \cdots, \xi_{nERN}, \xi_{nFN}, \tilde{\xi}_{MERn}, \tilde{\xi}_{MFn}, \tilde{\xi}_{nER1}, \\ \tilde{\xi}_{nF1}, \cdots, \tilde{\xi}_{nER(n-1)}, \tilde{\xi}_{nF(n-1)}, \tilde{\xi}_{nER(n+1)}, \\ \tilde{\xi}_{nF(n+1)}, \cdots, \tilde{\xi}_{nERN}, \tilde{\xi}_{nFN} \end{array} \right\} \right]^T, \tag{50b}$$

$$\{\lambda_j\}_{j=1}^{2N+4M} = \left\{ \begin{array}{c} \{\lambda_{MERn}\}_{n=1}^N, \{\lambda_m\}_{m=1}^M, \{\lambda_{Mm}\}_{m=1}^M, \\ \{\lambda_{Em}\}_{m=1}^M, \{\lambda_{EEm}\}_{m=1}^M, \{\lambda_{MFn}\}_{n=1}^N \end{array} \right\}, \tag{50c}$$

$$\mathbf{J}_0 \triangleq \left[\begin{array}{c} \{\xi_{MFn}\}_{n=1}^N, \{\xi_{MERn}\}_{n=1}^N, \{\tilde{\xi}_{1,m}\}_{m=1}^M, \\ \tilde{\xi}_{1,E}, \cdots, \{\tilde{\xi}_{N,m}\}_{m=1}^M, \tilde{\xi}_{N,E} \end{array} \right]^T, \tag{50d}$$

where the elements in $\mathbf{J}_n$ and $\mathbf{J}_0$ are the ICIs relevant to the FBS_n and MBS, respectively. Similarly, $\mathbf{K}_n$ and $\mathbf{K}_0$ contain the global versions of corresponding ICIs, respectively, which are written as

$$\mathbf{K}_n \triangleq \left[\left\{ \begin{array}{c} \{X_{n,m}\}_{m=1}^M, X_{n,E}, X_{nER1}, X_{nF1}, \cdots, \\ X_{nER(n-1)}, X_{nF(n-1)}, X_{nER(n+1)}, X_{nF(n+1)}, \\ \cdots, X_{nERN}, X_{nFN}, X_{MERn}, X_{MFn}, X_{nER1}, \\ X_{nF1}, \cdots, X_{nER(n-1)}, X_{nF(n-1)}, X_{nER(n+1)}, \\ X_{nF(n+1)}, \cdots, X_{nERN}, X_{nFN} \end{array} \right\} \right]^T, \tag{51a}$$

$$\mathbf{K}_0 \triangleq \left[\left\{ \begin{array}{c} \{X_{MFn}\}_{n=1}^N, \{X_{MERn}\}_{n=1}^N, \{X_{1,m}\}_{m=1}^M, \\ X_{1,E}, \cdots, \{X_{N,m}\}_{m=1}^M, X_{N,E} \end{array} \right\} \right]^T. \tag{51b}$$

Thus, we can rewrite problem (46) in a more compact form as

$$\min_{\mathbf{S},\boldsymbol{\Omega}} \ -\sum_{k=0}^{N} x_k \tag{52a}$$

$$s.t.\ \mathbf{s}_k \in \mathbf{S}_k^{(q)}, \quad k \in [0, N], \tag{52b}$$
$$x_k = x, \quad k \in [0, N], \tag{52c}$$
$$\mathbf{J}_k = \mathbf{K}_k, \quad k \in [0, N], \tag{52d}$$

where $\mathbf{S} = [\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_N]$, $\mathbf{\Omega} = [x, \mathbf{K}_0^T, \mathbf{K}_1^T, \cdots, \mathbf{K}_N^T]$. In what follows, it is in a position to describe the ADMM for solving problem (46) in a distributed manner.

Firstly, the augmented Lagrangian function of problem (52) is written as

$$\mathcal{L}(\mathbf{S}, \mathbf{\Omega}, \mathbf{\Theta}, \mathbf{\Psi}) = \sum_{k=0}^{N} \{-x_k + \frac{f}{2}((x_k - x)^2 + \|\mathbf{J}_k - \mathbf{K}_k\|_2^2) + \Theta_k(x_k - x) + \mathbf{\Psi}_k^T(\mathbf{J}_k - \mathbf{K}_k)\}, \tag{53}$$

where $\mathbf{\Theta} \triangleq [\Theta_0, \Theta_1, \cdots, \Theta_N]$, $\mathbf{\Psi} \triangleq [\mathbf{\Psi}_0^T, \mathbf{\Psi}_1^T, \cdots, \mathbf{\Psi}_N^T]$; $\frac{f}{2}((x_k - x)^2 + \|\mathbf{J}_k - \mathbf{K}_k\|^2)$ denotes the quadratic penalty term while $f > 0$ is the penalty parameter, which is to provides strict convexity w.r.t. $\mathbf{S}$ and $\mathbf{\Omega}$ for (52); Θ_k and $\mathbf{\Psi}_k$ are the Lagrange multipliers associated with (52c) and (52d), respectively. Therefore, problem (53) is always solvable.

Then, we are in a position to execute the central parts of the ADMM, which are 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 $\mathbf{\Psi}$) by the Gauss-Seidel method, respectively. Specifically, by the update of global variables, all BSs agree a common knowledge of corresponding ICI variables and the consensus SEE value at each iteration of the ADMM. Then, each BS utilizes its local CSIs to solves its own subproblem independently, and the involved variables are derived into equality by updating its Lagrange multipliers. The specific steps of the updated procedures are shown in following sections.

1) UPDATE OF GLOBAL VARIABLES

With the local variables fixed, an iteration begins with the update of global variables through solving the convex problem $\min_{\mathbf{\Omega}} \mathcal{L}^{(q)}\left(\mathbf{\Omega}, \mathbf{S}^{(u)}, \mathbf{\Theta}^{(u)}, \mathbf{\Psi}^{(u)}\right)$, where u denotes the iteration number of ADMM part. Meanwhile, due to the fact that $\mathcal{L}^{(q)}\left(\mathbf{\Omega}, \mathbf{S}^{(u)}, \mathbf{\Theta}^{(u)}, \mathbf{\Psi}^{(u)}\right)$ is separable in x and $\{\mathbf{K}_k\}$, the minimization problem can be divided into two subproblems, which are respectively written as

$$x^{(u+1)} = \arg\min_{x} \sum_{k=0}^{N} (\frac{f}{2}(x_k^{(u)} - x)^2 - \Theta_k^{(u)} x), \tag{54a}$$

$$\mathbf{K}_k^{(u+1)} = \arg\min_{\mathbf{K}_k} (-(\mathbf{\Psi}_k^{(u)})^T \mathbf{K}_k + \frac{f}{2} \left\| \mathbf{J}_k^{(u)} - \mathbf{K}_k \right\|_2^2), \tag{54b}$$

where $x_k^{(u)}$ and $\mathbf{J}_k^{(u)}$ can be obtained by solving (54a) and (54b) in the u-th ADMM iteration and are exchanged between all coordinated BSs, respectively. Furthermore, according to [28], the closed-form solutions to (54a) and (54b) are respectively written as

$$x^{(u+1)} = \frac{1}{N+1} \sum_{k=0}^{N} (x_k^{(u)} + \frac{\Theta_k^{(u)}}{f}), \tag{55a}$$

$$X_{n,m}^{(u+1)} = \frac{1}{2}(\xi_{n,m}^{(u)} + \tilde{\xi}_{n,m}^{(u)}) + \frac{1}{2f}(\Psi_{n,m}^{(u)} + \tilde{\Psi}_{n,m}^{(u)}), \tag{55b}$$

$$X_{n,E}^{(u+1)} = \frac{1}{2}(\xi_{n,E}^{(u)} + \tilde{\xi}_{n,E}^{(u)}) + \frac{1}{2f}(\Psi_{n,E}^{(u)} + \tilde{\Psi}_{n,E}^{(u)}), \tag{55c}$$

$$X_{nERt}^{(u+1)} = \frac{1}{2}(\xi_{nERt}^{(u)} + \tilde{\xi}_{nERt}^{(u)}) + \frac{1}{2f}(\Psi_{nERt}^{(u)} + \tilde{\Psi}_{nERt}^{(u)}), \tag{55d}$$

$$_{nFt}^{(u+1)} = \frac{1}{2}(\xi_{nFt}^{(u)} + \tilde{\xi}_{nFt}^{(u)}) + \frac{1}{2f}(\Psi_{nFt}^{(u)} + \tilde{\Psi}_{nFt}^{(u)}), \tag{55e}$$

$$X_{MERn}^{(u+1)} = \frac{1}{2}(\xi_{MERn}^{(u)} + \tilde{\xi}_{MERn}^{(u)}) + \frac{1}{2}(\Psi_{MERn}^{(u)} + \tilde{\Psi}_{MERn}^{(u)}), \tag{55f}$$

$$X_{MFn}^{(u+1)} = \frac{1}{2f}(\xi_{MFn}^{(u)} + \tilde{\xi}_{MFn}^{(u)}) + \frac{1}{2f}(\Psi_{MFn}^{(u)} + \tilde{\Psi}_{MFn}^{(u)}), \tag{55g}$$

where $\{\Psi_{n,m}^{u}, \tilde{\Psi}_{n,m}^{u}\}$ are the dual variables associated with the primal variables $\{\xi_{n,m}^{u}, \tilde{\xi}_{n,m}^{u}\}$of the u-th ADMM iteration in (55b), respectively. Meanwhile, the other variables in (55c) (55g) follow the similar relation, and we do not repeat it for the sake of simplicity. Moreover, it should be noticed that (55a),(55f),(55g) and (55a) (55e) are independently carried out at the MBS and FBS_n, respectively. Therefore, $\{x^{(u+1)}, X_{MERn}^{(u+1)}, X_{MFn}^{(u+1)}\}$ and $\{x^{(u+1)}, X_{n,m}^{(u+1)}, X_{n,E}^{(u+1)}, X_{nERt}^{(u+1)}, X_{nFt}^{(u+1)}\}$ can be obtained at the MBS and FBS_n, respectively. In other words, each BS can compute x^{u+1} by running an average consensus algorithm [29] in (55a). Meanwhile, $\{X_{n,m}^{(u+1)}, X_{n,E}^{(u+1)}, X_{MERn}^{(u+1)}, X_{MFn}^{(u+1)}\}$ and $\{X_{nERt}^{(u+1)}, X_{nFt}^{(u+1)}, X_{tERn}^{(u+1)}, X_{tFn}^{(u+1)}\}$ are acquired at the FBS_n after gathering $\frac{1}{2}\xi_{n,m}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{n,m}^{(u)}$, $\frac{1}{2}\xi_{n,E}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{n,E}^{(u)}$, $\frac{1}{2}\xi_{MERn}^{(u)} + \frac{1}{2f}\Psi_{MERn}^{(u)}$, $X_{MFn}^{(u+1)} = \frac{1}{2}\xi_{MFn}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{MFn}^{(u)}$ from the MBS and $\frac{1}{2}\xi_{nERt}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{nERt}^{(u)}$, $\frac{1}{2}\xi_{nFt}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{nFt}^{(u)}$, $\frac{1}{2}\xi_{tERn}^{(u)} + \frac{1}{2f}\Psi_{tERn}^{(u)}$ and $\frac{1}{2}\xi_{tFn}^{(u)} + \frac{1}{2f}\Psi_{tFn}^{(u)}$ from the FBS_t. In the similar way, $\{X_{MERn}^{(u+1)}, X_{MFn}^{(u+1)}, X_{n,m}^{(u+1)}, X_{n,E}^{(u+1)}\}$ are obtained at the MBS after collecting $\frac{1}{2}\xi_{MERn}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{MERn}^{(u)}$, $\frac{1}{2}\xi_{MFn}^{(u)} + \frac{1}{2f}\tilde{\Psi}_{MFn}^{(u)}$, $\frac{1}{2}\xi_{n,m}^{(u)} + \frac{1}{2f}\Psi_{n,m}^{(u)}$ and $\frac{1}{2}\xi_{n,E}^{(u)} + \frac{1}{2f}\Psi_{n,E}^{(u)}$ from the FBS_n. Notice that these information exchange between coordinated BSs are all real-valued so the amount of exchanged signaling overhead is significantly lower than the complex-value CSIs sharing.

2) UPDATE OF LOCAL VARIABLES

The update of local variables $\mathbf{S}$ is performed by solving the following convex problem

$$\mathbf{S}^{(u+1)} = \arg\min_{\mathbf{s}_k \in \mathbf{S}_k^{(q)}, k \in [0,N]} \mathcal{L}^{(q)}\left(\mathbf{\Omega}^{(u+1)}, \mathbf{S}, \mathbf{\Theta}^{(u)}, \mathbf{\Psi}^{(u)}\right) \tag{56}$$

According to the distributed feature of ADMM, the augmented Lagrangian function can be decomposable in $\mathbf{s}_k$. Succinctly speaking, by decomposing problem (56) independently, it can be solved in parallel at each BS. Thus, with its

local CSIs, BS_k only needs to solve the following subproblem

$$\begin{aligned}\mathbf{s}_k^{(u+1)} = & \underset{\mathbf{s}_k\in\mathbf{S}_k^{(q)},k\in[0,N]}{\arg\min} \{-x_k+\frac{f}{2}((x_k-x^{(u+1)})^2+\left\|\mathbf{J}_k-\mathbf{K}_k^{(u+1)}\right\|^2) \\ & +\Theta_k^{(u)}(x_k-x^{(u+1)})+(\mathbf{\Psi}_k^{(u)})^T(\mathbf{J}_k-\mathbf{K}_k^{(u+1)})\}, \\ & k\in[0,N], \end{aligned} \tag{57}$$

where $x^{(u+1)}$ and $\mathbf{K}_k^{(u+1)}$ denote the global variables, which have been updated in the $(u+1)$-th ADMM iteration. Subproblem (57) is convex so it can be solved directly by existing convex solvers.

3) UPDATE OF LAGRANGE MULTIPLIERS

Finally, the Lagrange multipliers are updated by

$$\Theta_k^{(u+1)} = \Theta_k^{(u)} + f(x_k^{(u+1)} - x^{(u+1)}), \quad k\in[0,N], \tag{58a}$$

$$\mathbf{\Psi}_k^{(u+1)} = \mathbf{\Psi}_k^{(u)} + f(\mathbf{J}_k^{(u+1)} - \mathbf{K}_k^{(u+1)}), \quad k\in[0,N]. \tag{58b}$$

BSs have acquired all the global and local variables after the previous two steps A and B. Therefore, no extra signaling overhead of BSs is needed in this step. Moreover, if the primal residual $\Xi^{(u)}$ is below a threshold Ξ_{th}, i.e., $\Xi^{(u)} < \Xi_{\mathrm{th}}$, the ADMM iteration is thought as convergence, where $\Xi^{(u)}$ is defined as $\sqrt{\sum_{k=0}^{N}\left\|[x_k^{(u)},(\mathbf{J}_k^{(u)})^T]^T-[x^{(u)},(\mathbf{K}_k^{(u)})^T]^T\right\|_2^2}$ to indicate the convergence in the u-th ADMM iteration. Then, it comes to the $(q+1)$-th SCA iteration.

B. THE DISTRIBUTED ALGORITHM WITH ADMM ITERATION

The steps of distributed algorithm with ADMM iteration are summarized in Algorithm 2, where the outer and the inner loop are the SCA-based and ADMM-based iterative procedure, respectively. Subsequently, the computational complexity and the information change between BSs are analyzed, respectively.

1) PER BS COMPLEXITY ANALYSIS

The computational complexity of each iteration in Algorithm 2 is mainly determined by solving the subproblem (58) at each BS. Therefore, the computational overhead can be acquired with the similar way as that in Algorithm 2. Note that the processing stress of whole network is averaged on $N+1$ BSs so the computational overhead is affordable for each BS.

2) INFORMATION EXCHANGE

The information exchange mainly occurs in the step 5 and 6, just as shown in Algorithm 2. To be more specific, $x_k^{(u)}$ and $\Theta_k^{(u)}$ are broadcasted by each BS in step 5; then, in step 6, $M+3$ real values are sent to each FBS by the MBS while 4 and $M+3$ real values are sent to the other $N-1$ FBSs and the MBS by each FBS, respectively. Meanwhile, notice that these information exchange between coordinated BSs are all real-valued so the amount of exchanged signaling overhead is significantly lower than the complex-value CSIs sharing.

Algorithm 2 Distributed Algorithm With ADMM Iteration

1: **Initialization:** set iteration index $q=0$, $u=0$ and chose feasible initial points $\{\vartheta_j\}_{j=1}^{5M+6N+1}$ of SCA outer loop, and feasible initial points of ADMM inner loop $\{x_k^{(0)}\}$, $\{\mathbf{J}_k^{(0)}\}$, $\mathbf{\Theta}^{(0)}$, $\mathbf{\Psi}^{(0)}$;
2: **Repeat :** (SCA outer loop)
3: **while** $\Xi \geq \Xi_{\mathrm{th}}$ do :(ADMM inner loop)
4: **for** $k\in[0,N]$ do:
5: BS_k updates $x_k^{(u+1)}$ through an average consensus algorithm;
6: BS_k updates $\mathbf{K}_k^{(u+1)}$ by (55b) (55g);
7: BS_k updates $\mathbf{s}_k^{(u+1)}$ by (57);
8: BS_k updates the Lagrange multipliers $\Theta_k^{(u+1)}$, $\mathbf{\Psi}_k^{(u+1)}$ by (58a) and (58b), respectively;
9: **end for**
10: $u=u+1$;
11: **end while**
12: obtain the optimal solutions $\{\vartheta_j^*\}_{j=1}^{5M+6N+1}$;
13: update $\{\{x_k^{(0)}\},\{\mathbf{J}_k^{(0)}\},\mathbf{\Theta}^{(0)},\mathbf{\Psi}^{(0)}\}=\{\{x_k^*\},\{\mathbf{J}_k^*\},\mathbf{\Theta}^*,\mathbf{\Psi}^*\}$ and the SCA iteration points $\{\vartheta_j^{(q+1)}\}_{j=1}^{5M+6N+1}=\{\vartheta_j^*\}_{j=1}^{5M+6N+1}$;
14: $q=q+1$, $u=0$;
15: **Until:** algorithm converges;
16: **Output:** optimal SEE objective value, beamforming and AN matrixes $\{\mathbf{W}_m^*\}_{m=1}^M$, $\{\mathbf{W}_{Fn}^*\}_{n=1}^N$, $\{\mathbf{Z}_k^*\}_{k=0}^N$.

V. SIMULATION RESULTS AND ANALYSIS

In this section, simulation results are provided to evaluate the performance of proposed robust design. We consider that two coordinated FBSs are deployed under the coverage of the MBS, where two MUs and an Eve are randomly distributed in a macrocell with a radius of 300m while an IR and an ER are scattered randomly in each femtocell with a radius of 100m. Meanwhile, the path loss and shadowing are modeled as $\beta(\mathrm{dB}) = 38\log_{10}(d)+34.5+CN(0,8)$ [28], where d represents the distance from transmitter to receiver in the corresponding channel while the small-scale fading coefficients all follow Gaussian distributed. Moreover, unless otherwise specified, the specific antenna settings are $N_M=8$, $N_F=5$ and $N_E=N_{ER}=N_{Eve}=2$, respectively; the total transmit power of the MBS and each FBS are $P_M=48\mathrm{dBm}$, $P_F=38\mathrm{dBm}$, power loss and power amplifier efficiency are $P_\varepsilon=33\mathrm{dBm}$ and $\varpi=0.8$, respectively; EH threshold of ERs is $\theta=\theta_n=-20\mathrm{dBm}$, $\forall n\in[1,N]$; all the CSI errors have the same bound $\varepsilon=0.01$; the penalty parameter $f=0.05$. All simulation results are averaged over 1000 randomly generated channel realizations.

In our simulations, two compared designs, i.e., zero-force AN design and non-AN design, are presented simultaneously

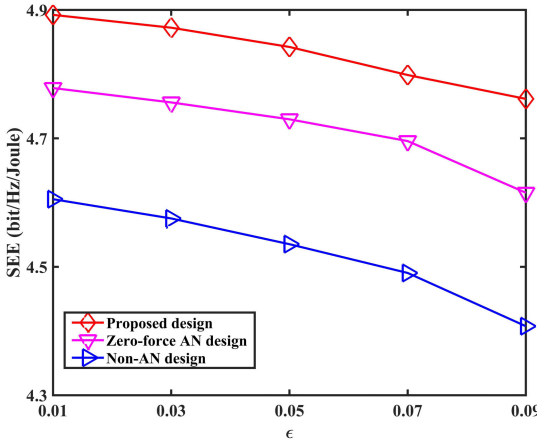

FIGURE 2. SEE performance versus the CSI error bound.

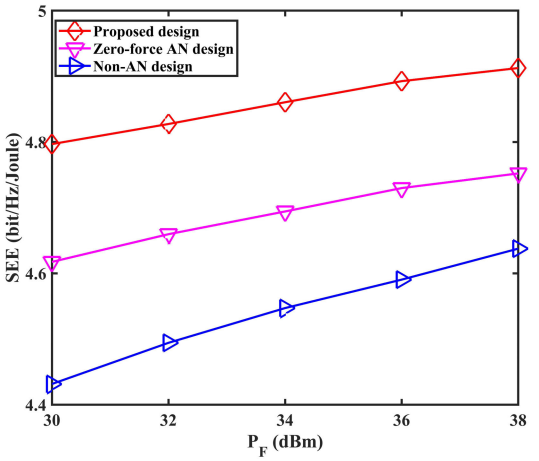

FIGURE 3. SEE performance versus the transmit power at each FBS.

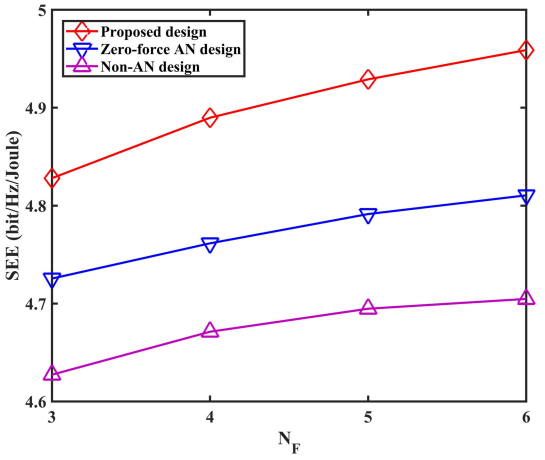

FIGURE 4. SEE performance versus the antennas number at each FBS.

to analyze the performance of proposed method. More specifically, non-AN design means that no AN is injected into the transmit beam of the MBS and FBSs while zero-force AN design denotes that AN is injected into the zero space of channel of legitimate users. Firstly, the comparison of these three methods with respect to the impact of the CSI error on SEE performance is displayed in Fig. 2. With the increase of ε, the SEE performance of them all tends to decrease. This can be explained that the deviation of downlink beamforming and AN at the MBS and FBSs occurs due to the existence of CSI error, which will surely bring about the waste of transmit power and the decrease of secrecy performance. Meanwhile, under the same parameters setting, we find that our proposed design always has the most SEE performance while zero-force AN design is second to it. This phenomenon indicates that the aided AN can better promote the SEE performance because AN carries the RF power for ERs while interfering with potential Eves, which utilizes the transmit power more efficiently; moreover, the AN with unfixed structure has more superior performance than AN with fixed structure, which validates the effectiveness of our method.

Next, we pay attention to the effect of maximum output power at BSs on the SEE performance. Considering that the power supply is usually sufficient at the MBS, we mainly focus on the FBSs in this part, as presented in Fig. 3. As it describes, the SEE performance of three methods all increase with the growth of P_F. The reason is that as P_F grows, some extra transmit power will be saved when achieving the comparative SEE performance. Then, these extra power can be allocated for improving the SEE performance of macrocell and femtocells.

Similarly, considering that the MBS is deployed in schedule, which means that its antennas setting is usually sufficient, we mainly focus on the impact of the number of antennas at each FBS on the SEE performance, just as presented in Fig. 4. Specifically, as N_F grows, the SEE performance also increases simultaneously because the resolution ratio of achievable space can be improved by more antennas available at each FBS, and then the diversity between legitimated channels (i.e., from FBSs to IRs) and wiretapping channels (i.e., from FBSs to ERs) is guaranteed more easily with less AN injected. Furthermore, our proposed design always has the most superior performance under the same parameters settings, which further validates the effectiveness of our design.

As shown in Fig. 5, the SEE performance tends to decrease as the increase of θ, and our proposed method always has the best performance among three methods. This phenomenon can be explained that the increase of θ leads to more power consumption utilized for power transfer, which will lead to the decrease of secrecy performance of femtocells and the average SEE performance of whole network because of the SEE fairness requirement between multiple cells and the transmit power limit. Meanwhile, we notice that if the potential Eve is equipped with more antennas, the SEE performance of our proposed design becomes lower, which results from the increased wiretapping performance of potential Eves and leads to the decrease of secrecy performance.

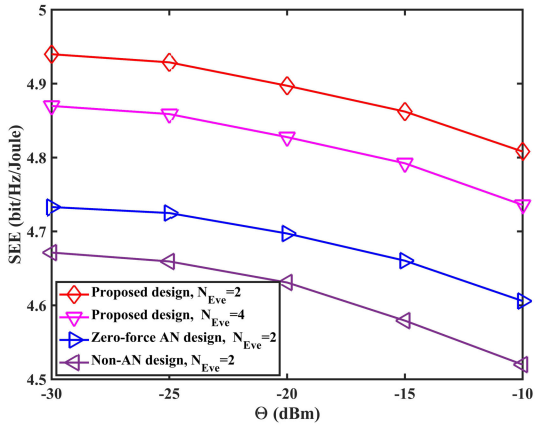

FIGURE 5. SEE performance versus the EH threshold of ERs.

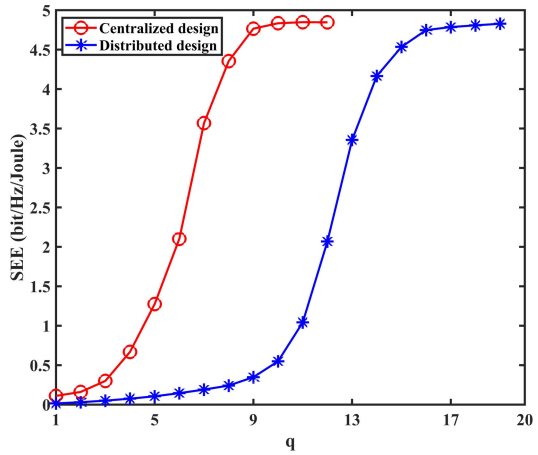

FIGURE 6. Convergence of the centralized and distributed design.

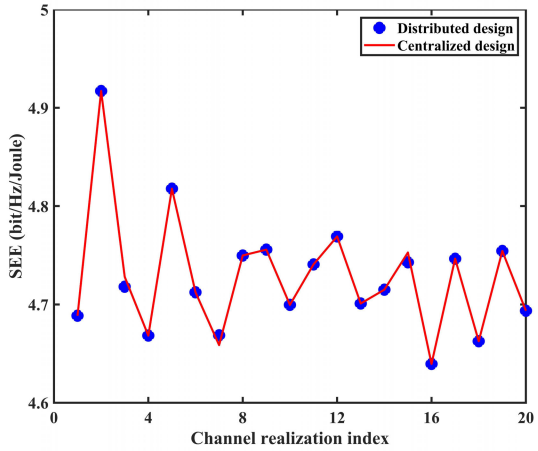

FIGURE 7. SEE performance on random channel realizations.

Fig. 6 illustrates the convergence performance versus the iteration number in the centralized and distributed design under the same channel. As can be observed, the SEE curve of centralized design converges to a stabilized value within about 10 iterations. As for the distributed design, the converging rate is slower than the centralized design, which is unsurprising for the distributed solving manner. Moreover, the ADMM-based inner loop also slows down the converging rate.

The performance comparison between the centralized and distributed design is provided in Fig. 7 in terms of the SEE over 20 random channel realizations. As it shows, the distributed design can almost converge to the optimal solution obtained by the centralized design under most of the channel realizations. Moreover, only small disparity exists between the optimal solution obtained by the centralized and distributed design in the 3, 7 and 15 channel realization, which further implies the validity of distributed design.

VI. CONCLUSION

In this paper, with the existence of CSI errors, we investigated the robust TBF and AN design for secure SWIPT in a two-tier HetNet, where both the MUs and IRs are in faced with potential secrecy leakage. To achieve green communications and promote the average secrecy performance robustly, we explored the joint design problem for SEE maximization while considering the fairness among multiple cells. Specifically, with the aid of the SCA, SDR techniques and S-procedure, we successfully transform it into a series of convex forms. Moreover, to further release the processing stress on the calculating center and the signaling overhead of the network, we proposed a distributed approach based on ADMM that allows each BS to deal with its local CSI. Simulation results verified the performance of robust TBF and AN design and the validity of the proposed distributed approach.

APPENDIX

To prove the Proposition 1, we denote the Lagrange function of (46) as $\mathcal{L}$. Due to the fact that there are too many constraints and some of them are not relevant to the proof process, the complete expression of $\mathcal{L}$ must be too long and not necessary. Therefore, we only present part of $\mathcal{L}$, which are relevant to the proof process. In the following, we will first prove $\text{Rank}(\mathbf{W}_m^*) = 1, \forall m \in [1, M]$ in the subsection A, then we can follow the similar process to prove $\text{Rank}(\mathbf{W}_{Fn}^*) = 1, \forall n \in [1, N]$ in the subsection B.

A. THE PROOF OF $\text{Rank}(\mathbf{W}_m^*) = 1$

From (46) and the Karush-Kuhn-Tucker conditions, we have

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}_m^*} = \mathbf{A}_m^* - \mathbf{X}_m^* + [\mathbf{I}_{N_M} \hat{\mathbf{h}}_m] \mathbf{T}_{Mm}^* [\mathbf{I}_{N_M} \hat{\mathbf{h}}_m]^H = 0$$

$$\Rightarrow \mathbf{A}_m^* = \mathbf{X}_m^* - [\mathbf{I}_{N_M} \hat{\mathbf{h}}_m] \mathbf{T}_{Mm}^* [\mathbf{I}_{N_M} \hat{\mathbf{h}}_m]^H, \quad \text{(A1)}$$

where

$$\mathbf{X}_m^* = (f^* + i^*)\mathbf{I}_{N_M} + \sum_{n=1}^{N} [\mathbf{I}_{N_M} \hat{\mathbf{h}}_{MFn}] \mathbf{T}_{MFn}^* [\mathbf{I}_{N_M} \hat{\mathbf{h}}_{MFn}]^H + \sum_{j=1}^{N_M} \mathbf{R}_{Em}^{*(j,j)} - \sum_{n=1}^{N} \sum_{j=1}^{N_M} \mathbf{R}_{MERn}^{*(j,j)},$$

$\mathbf{A}_m^* \geq \mathbf{0}$, $f^* \geq 0$, $i^* \geq 0$, $\mathbf{T}_{MFn}^* \geq \mathbf{0}$, $\mathbf{T}_{Mm}^* \geq 0$ indicate the optimal Lagrange multipliers associated with (12e), (12c), (14c), (45c) and (43c), respectively; $\mathbf{R}_{Em}^{*(j,j)} \in \mathbb{H}_+^{N_M}$, $\mathbf{R}_{MERn}^{*(j,j)} \in \mathbb{H}_+^{N_M}$ are the diagonal block sub-matrices of $[\mathbf{I}_{N_M}\hat{\mathbf{h}}_E]\mathbf{\Omega}_{Em}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_E]^H$ and $[\mathbf{I}_{N_M}\hat{\mathbf{h}}_{MERn}]\mathbf{\Omega}_{MERn}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_{MERn}]^H$, respectively. In particular,

$$[\mathbf{I}_{N_M}\hat{\mathbf{h}}_E]\mathbf{\Omega}_{Em}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_E]^H = \begin{bmatrix} \mathbf{R}_{Em}^{*(1,1)} \cdots \mathbf{R}_{Em}^{*(1,N_M)} \\ \cdots\cdots\cdots \\ \mathbf{R}_{Em}^{*(N_M,1)} \cdots \mathbf{R}_{Em}^{*(N_M,N_M)} \end{bmatrix}, \tag{A2}$$

$$[\mathbf{I}_{N_M}\hat{\mathbf{h}}_{MERn}]\mathbf{\Omega}_{MERn}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_{MERn}]^H = \begin{bmatrix} \mathbf{R}_{MERn}^{*(1,1)} \cdots \mathbf{R}_{MERn}^{*(1,N_M)} \\ \cdots\cdots\cdots \\ \mathbf{R}_{MERn}^{*(N_M,1)} \cdots \mathbf{R}_{MERn}^{*(N_M,N_M)} \end{bmatrix}, \tag{A3}$$

In what follows, we first prove the rank of $[\mathbf{I} \ \hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I} \ \hat{\mathbf{h}}_m]^H$ is no more than one, i.e., $\text{Rank}([\mathbf{I} \ \hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I} \ \hat{\mathbf{h}}_m]^H) \leq 1$. In this regard, we have the following definition expressed as

$$\begin{aligned} &[\mathbf{I}_{N_M}0][\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H = \mathbf{I}_{N_M} \\ &[\mathbf{I}_{N_M}0]\begin{bmatrix} \lambda_{Mm}^*\mathbf{I}_{N_M}0 \\ 0e_{Mm} \end{bmatrix} = \lambda_{Mm}^*[[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m] - [0\hat{\mathbf{h}}_m]], \end{aligned} \tag{A4}$$

At the optimal solution, according to the KKT condition, we have the following expression

$$\begin{bmatrix} \lambda_{Mm}^*\mathbf{I}_{N_M}0 \\ 0e_{Mm}^* \end{bmatrix}\mathbf{T}_{Mm}^* + \begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}^H\mathbf{W}_m^*\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^* = \mathbf{0} \tag{A5}$$

Let $[\mathbf{I}_{N_M}0]$ and $\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}^H$ multiply the left and right side of (A5), respectively, and combining with (A4), then we can acquire

$$\begin{aligned} &[\mathbf{I}_{N_M}0]\begin{bmatrix} \lambda_{Mm}^*\mathbf{I}_{N_M}0 \\ 0e_{Mm}^* \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H \\ &\quad + [\mathbf{I}_{N_M}0]\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}^H\mathbf{W}_m^*\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H = 0 \\ &\Rightarrow \lambda_{Mm}^*([\mathbf{I}_{N_M}\hat{\mathbf{h}}_m] - [0\hat{\mathbf{h}}_m])\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H \\ &\quad + \mathbf{W}_m^*\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H = 0 \\ &\Rightarrow \lambda_{Mm}^*[0\hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H \\ &= (\lambda_{Mm}^*\mathbf{I}_{N_M} + \mathbf{W}_m^*)\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H, \end{aligned} \tag{A6}$$

Due to $\lambda_{Mm}^* \geq 0$ and $\mathbf{W}_m^* \geq 0$, we see $(\lambda_{Mm}^*\mathbf{I}_{N_M} + \mathbf{W}_m^*) \geq 0$. Furthermore, at the optimal point, $\mathbf{W}_m^* = 0$ is not consistent with the optimization objective. Therefore, there must be $\mathbf{W}_m^* > 0 \Rightarrow (\lambda_{Mm}^*\mathbf{I}_{N_M} + \mathbf{W}_m^*) > 0$. Then, combining with (A6), we can get

$$\begin{aligned} &\text{Rank}(\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H) \\ &= \text{Rank}((\lambda_{Mm}^*\mathbf{I}_{N_M} + \mathbf{W}_m^*)\begin{bmatrix} \mathbf{I}_{N_M}\hat{\mathbf{h}}_m \end{bmatrix}\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M}\hat{\mathbf{h}}_m]^H) \\ &= \text{Rank}(\lambda_{Mm}^*[0\hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I}_{N_M} \ \hat{\mathbf{h}}_m]^H) \\ &\leq \text{Rank}([0\hat{\mathbf{h}}_m]) \leq 1, \end{aligned} \tag{A7}$$

Therefore, we can get $\text{Rank}([\mathbf{I}\hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I}\hat{\mathbf{h}}_m]^H) \leq 1$.

In the next, we aim to prove that the rank of $\mathbf{A}_m^*$ is $\text{Rank}(\mathbf{A}_m^*) = N_M - 1$. To begin with, combining (A1) and (A7), we have

$$\begin{aligned} \text{Rank}(\mathbf{A}_m^*) &\geq \text{Rank}(\mathbf{X}_m^*) - \text{Rank}([\mathbf{I}\hat{\mathbf{h}}_m]\mathbf{T}_{Mm}^*[\mathbf{I}\hat{\mathbf{h}}_m]^H) \\ &= \text{Rank}(\mathbf{X}_m^*) - 1, \end{aligned} \tag{A8}$$

and then we continue to analyze the following two cases:

1) If $\text{Rank}(\mathbf{X}_m^*) = N_M$, there is $\text{Rank}(\mathbf{A}_m^*) \geq N_M - 1$. However, if $\text{Rank}(\mathbf{A}_m^*) = N_M$, i.e., $\mathbf{A}_m^*$ is of full-rank, then it means $\mathbf{W}_m^* = \mathbf{0}$, which is not consistent with the optimization objective and can not be the optimal solution. Therefore, we can get $\text{Rank}(\mathbf{A}_m^*) = N_M - 1$ and $\text{Rank}(\mathbf{W}_m^*) = 1$ under this case;

2) If $\text{Rank}(\mathbf{X}_m^*) \leq N_M - 1$, we denote $\mathbf{\Pi}_m = [\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_{N_M - R_{\mathbf{X}_m^*}}]$ with $\mathbf{\Pi}_m^H\mathbf{\Pi}_m = \mathbf{I}_{N_M - R_{\mathbf{X}_m^*}}$ as the orthogonal basis for the null space of $\mathbf{X}_m^*$, i.e., $\mathbf{X}_m^*\mathbf{\Pi}_m = \mathbf{0}$. Then, we can acquire

$$\begin{aligned} \boldsymbol{\pi}_i^H\mathbf{A}_m^*\boldsymbol{\pi}_i &= \boldsymbol{\pi}_i^H(\mathbf{X}_m^* - r_m^*\mathbf{H}_m)\boldsymbol{\pi}_i \\ &= -r_m^*\left\|\mathbf{h}_m^H\boldsymbol{\pi}_i\right\|^2 \leq 0, \quad i \in [1, N_M - R_{\mathbf{X}_m^*}], \end{aligned} \tag{A9}$$

where $r_m^* = 0$ occurs if and only if

$$\text{Tr}(\mathbf{H}_m\mathbf{W}_m) \leq g_{\zeta_{0m}}^q(\alpha_{0m}, \phi_m) - \phi_m, \quad m \in [1, M], \tag{A10}$$

which means that the secrecy rate of MUs can not be guaranteed, which contracts with the formed equivalently constraint (13b) and the optimal design objective. Therefore, there must be $r_m^* > 0$. Since $\mathbf{A}_m^* \geq 0$ and $r_m^* > 0$, it follows from (A9) that $\left\|\mathbf{h}_m^H\boldsymbol{\pi}_i\right\|^2 = 0$. Considering $\mathbf{X}_m^*\mathbf{\Pi}_m = \mathbf{0}$, we have

$$\mathbf{h}_m\mathbf{h}_m^H\mathbf{\Pi}_m = 0 \Rightarrow \mathbf{A}_m^*\mathbf{\Pi}_m = 0, \tag{A11}$$

which indicates that all $\boldsymbol{\pi}_i$'s lie in the null space of $\mathbf{h}_m\mathbf{h}_m^H$ and no information will be transferred to MU_m [30]. Hence positive secrecy rate can not be satisfied, which contradicts with the objective of optimization SEE under this case

In summary, according to the two cases analyzed above, $\text{Rank}(\mathbf{A}_m^*) = N_M - 1$ must hold. According to $\mathbf{A}_m^*\mathbf{W}_m^* = 0$, we have

$$\text{Rank}(\mathbf{W}_m^*) \leq 1. \tag{A12}$$

Based on above analysis, $\text{Rank}(\mathbf{W}_m^*) = 0$ is not consistent with the design objective. Therefore, there must be $\text{Rank}(\mathbf{W}_m^*) = 1$.

B. THE PROOF OF Rank($\mathbf{W}_{Fn}^{*}$) = 1

Following the similar procedures, we can also prove Rank($\mathbf{W}_{Fn}^{*}$) = 1. To avoid redundancy, we do not present the specific proof process in this section.

REFERENCES

[1] G. Zhao, S. Chen, L. Zhao, and L. Hanzo, "Energy-spectral-efficiency analysis and optimization of heterogeneous cellular networks: A large-scale user-behavior perspective," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4098–4112, May 2018.

[2] Y. Wang, M. Haenggi, and Z. Tan, "SIR meta distribution of K-tier downlink heterogeneous cellular networks with cell range expansion," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 3069–3081, Apr. 2019.

[3] S. Jang, H. Lee, S. Kang, T. Oh, and I. Lee, "Energy efficient SWIPT systems in multi-cell MISO networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8180–8194, Dec. 2018.

[4] H. Zhang, J. Du, J. Cheng, K. Long, and V. C. M. Leung, "Incomplete CSI based resource optimization in SWIPT enabled heterogeneous networks: A non-cooperative game theoretic approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1882–1892, Mar. 2018.

[5] X. Hu, B. Li, K. Huang, and K. Wong, "Secrecy energy efficiency in wireless powered heterogeneous networks: A distributed ADMM approach," *IEEE Access*, vol. 6, pp. 20609–20624, 2018.

[6] B. Liu, F. Zhou, G. Lu, and R. Q. Hu, "Energy efficient and robust beamforming for MISO cognitive small cell networks," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5002–5014, Dec. 2018.

[7] B. Li, Z. Fei, Z. Chu, F. Zhou, K. K. Wong, and P. Xiao, "Robust chance-constrained secure transmission for cognitive satellite–terrestrial networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4208–4219, May 2018.

[8] H. Zhang, Y. Huang, C. Li, and L. Yang, "Secure beamforming design for SWIPT in MISO broadcast channel with confidential messages and external eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7807–7819, Nov. 2016.

[9] H. Niu, B. Zhang, D. Guo, and Y. Huang, "Joint robust design for secure AF relay networks with SWIPT," *IEEE Access*, vol. 5, pp. 9369–9377, Apr. 2017.

[10] S.-H. Wang and B.-Y. Wang, "Robust secure transmit design in MIMO channels with simultaneous wireless information and power transfer," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 2147–2151, Nov. 2015.

[11] Y. Feng, Z. Yang, W.-P. Zhu, Q. Li, and B. Lv, "Robust cooperative secure beamforming for simultaneous wireless information and power transfer in amplify-and-forward relay networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2354–2366, Mar. 2017.

[12] M. R. A. Khandaker and K.-K. Wong, "Robust secrecy beamforming with energy-harvesting eavesdroppers," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 10–13, Feb. 2015.

[13] D. Lo and R. Schober, "Multiobjective resource allocation for secure communication in cognitive radio networks with wireless information and power transfer," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3166–3184, Aug. 2016.

[14] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Outage constrained robust beamforming for secure broadcasting systems with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7610–7620, Nov. 2016.

[15] M. R. A. Khandaker, K.-K. Wong, Y. Zhang, and Z. Zheng, "Probabilistically robust SWIPT for secrecy MISOME systems," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 211–226, Jan. 2017.

[16] T. A. Le, Q.-T. Vien, H. X. Nguyen, D. W. K. Ng, and R. Schober, "Robust chance-constrained optimization for power-efficient and secure SWIPT systems," *IEEE Trans. Green Commun. Netw.*, vol. 1, no. 3, pp. 333–346, Sep. 2017.

[17] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. L. Goff, "Robust outage secrecy rate optimizations for a MIMO secrecy channel," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 86–89, Feb. 2015.

[18] H. Niu, D. Guo, Y. Huang, B. Zhang, and B. Gao, "Outage constrained robust energy harvesting maximization for secure MIMO SWIPT systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 614–617, Oct. 2017.

[19] Y. Xu, K. Huang, and X. Hu, "CoMP transmission for safeguarding dense heterogeneous networks with imperfect CSI," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 1, pp. 110–132, Jan. 2019.

[20] Y. Xu, K. Huang, and Y. Zou, "Dual-threshold based secure on-off transmission scheme for dense HCNs with imperfect CSI," *China Commun.*, vol. 16, no. 3, pp. 132–142, Jan. 2019.

[21] 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, 2017.

[22] B. Li, Z. Fei, Z. Chu, and Y. Zhang, "Secure transmission for heterogeneous cellular networks with wireless information and power transfer," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3755–3766, Dec. 2018.

[23] 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 Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

[24] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[25] 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. 2016.

[26] 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.

[27] 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.

[28] K. Nguyen, Q. Vu, and M. Juntti, "Distributed solutions for energy efficiency fairness in multicell MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 1–7, Sep. 2017.

[29] 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.

[30] 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.

BO ZHANG received the M.S. degree from the National Digital Switching System Engineering and Technological Research and Developing Center, Zhengzhou, China, in 2017, where he is currently pursuing the Ph.D. degree. His major research interests include physical layer security and heterogeneous networks.

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. From 2017 to 2018, he was 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 with the National Digital Switching System Engineering and Technological Research and Developing Center, Zhengzhou, China. She is also the Leader of the Wireless Mobile Communication Innovation Technology Team, Henan. Her major research interests include wireless mobile communication networks and information theory.

ZHOU ZHONG received the Ph.D. degree from the National Digital Switching System Engineering and Technological Research and Developing Center, Zhengzhou, China, in 2013, where he is currently a Lecturer. His major research interest includes wireless communication security.

LIJIAN ZHANG received the Ph.D. degree from the National Digital Switching System Engineering and Technological Research and Developing Center, Zhengzhou, China, in 2015. He is currently a Lecturer with the Institute of Systems Engineering, Beijing, China. His major research interest includes wireless communication security.

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. He is also the Chief Investigator of the National Natural Science Foundation of China. His research interests include wireless communications and multimedia signal processing. He is also a Senior Member of the Chinese Institute of Electronics and the China Institute of Communications.