

Received February 24, 2017; accepted March 11, 2017, date of publication March 16, 2017, date of current version August 8, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2683639

Outage Constrained Robust Hybrid Coordinated Beamforming for Massive MIMO Enabled Heterogeneous Cellular Networks

GUIXIAN XU¹, CHIA-HSIANG LIN², WEIGUO MA³, SHANZHI CHEN³, (Senior Member, IEEE), AND CHONG-YUNG CHI², (Senior Member, IEEE)

¹School of Communications and Information Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

²Department of Electrical Engineering, Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan

³State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology, Beijing 100083, China

Corresponding author: C.-Y. Chi (cychi@ee.nthu.edu.tw)

This work was supported by the Ministry of Science and Technology, R.O.C., under Grant MOST 105-2221-E-007-020-MY2. Part of this paper was presented at the IEEE ICC, Pairs, May 2017.

ABSTRACT Heterogeneous network (HetNet), employing massive multiple-input multiple-output (MIMO), has been recognized as a promising technique to enhance network capacity, and to improve energy efficiency for the fifth generation of wireless communications. However, most existing schemes for coordinated beamforming (CoBF) for a massive MIMO HetNet unrealistically assume the availability of perfect channel state information (CSI) on one hand, and cascade of each antenna with a distinct radio-frequency chain in massive MIMO is neither power nor cost-efficient on the other hand. In this paper, we consider a massive MIMO-enabled HetNet framework, consisting of one macrocell base station (MBS) equipped with an analog beamformer, followed by a digital beamformer, and one femtocell base station (FBS) equipped with a digital beamformer. In the presence of Gaussian CSI errors, we propose a robust hybrid CoBF (HyCoBF) design, including an analog beamforming design for MBS, and a digital CoBF design for both MBS and FBS. To this end, an outage probability-constrained robust HyCoBF problem is formulated by minimizing the total transmit power. The analog beamforming mechanism at MBS is a newly devised low-complexity beam selection scheme by selecting analog beams from a discrete Fourier transform matrix codebook. Then, a conservative approximate CoBF solution is obtained via semidefinite relaxation and an extended Bernstein-type inequality. Furthermore, a distributed implementation for the obtained CoBF solution using alternating direction method of multipliers is proposed. Finally, numerical simulations are provided to demonstrate the efficacy of the proposed robust HyCoBF algorithm.

INDEX TERMS Heterogeneous network (HetNet), massive MIMO, hybrid coordinated beamforming (HyCoBF), semidefinite relaxation (SDR), alternating direction method of multipliers (ADMM).

I. INTRODUCTION

One primary objective of fifth generation (5G) of wireless communications is to support the ever-increasing network services, including Mobile Internet, Internet of Things (IoT) and Vehicular Ad-hoc Network (VANET). To meet various demands in 5G wireless communication systems (e.g., high data rates, extremely low latency and high energy efficiency), a 5G wireless network may employ advanced deployments, such as massive multiple-input multiple-output (MIMO) and superdense heterogeneous deployment of cells [1], [2]. In a massive MIMO enabled 5G heterogeneous network (HetNet), the macrocell base station (MBS) can

be equipped with hundreds of antennas to simultaneously serve tens of user equipments (UEs), and the remaining degrees of freedom (DoF) of massive MIMO can be used to mitigate the inter-tier interference. Moreover, ultra dense femtocell base station (FBS) deployment can effectively increase network capacity by more than two order of magnitude and offload the wireless data from the MBS [3]. Besides, millimeter wave (mmWave) with short wavelength enables massive MIMO to pack more antennas into highly directional footprint and to readily adjust beamforming [4], making massive MIMO practically feasible.

Although massive MIMO enabled HetNet is a promising technique for meeting 5G system requirements, there are still quite some unresolved challenges [2]. Firstly, radio resource management (RRM) for HetNet plays a crucial role in achieving the desired system performance. In other words, all the available radio resources (e.g., bandwidth, transmission power, and antennas) ought to be maximally and efficiently exploited, and meanwhile the targeted quality of service (QoS) for active users must be guaranteed with a minimum amount of radio resources, by means of advanced inter-cell/tier and intra-cell interference management schemes. In addition, the deployment of reliable backhaul networks with the designed resource management schemes is also a very important issue. Secondly, it may not be realistic to assume the availability of instantaneous channel state information (CSI) of each user in massive MIMO enabled HetNet. Particularly, in a multicell setup, pilot contamination imposes a fundamental performance bottleneck for the massive MIMO systems [5]. Thirdly, 3GPP Long-Term Evolution (LTE) Release 13 has specified that each base station (BS) can have at most 64 antennas and at most 8 radio frequency (RF) chains [6], because deploying many RF chains in massive MIMO may not be a practical solution due to expensive hardware cost and low energy efficiency.

A judicious approach for reducing hardware cost and training overhead is hybrid beamforming (HyBF), constituted by analog beamforming in the RF domain, and digital beamforming (with a much smaller dimension than the former) in the baseband domain, having received enormous attention in recent years [7]. In fact, it has been shown that HyBF design, depending on available CSI (instantaneous versus statistical CSI), beamformer structure (full-connected, partial-connected, or switched), and carrier frequency (centimeter wave versus millimeter wave), can achieve the same performance as the fully digital (FD) design, when the number of RF chains is twice more than that of data streams [8], [9].

Assuming that full instantaneous CSI is available, some HyBF designs have been proposed, in the context of both full-connected and partial-connected structures for mmWave systems [9]–[13]. In [14] and [15], two-stage hybrid precoding for frequency-division duplex (FDD) massive MIMO systems, where the analog precoder is only adaptive to channel statistics and the digital precoder is designed by effective low-dimensional channel. Furthermore, to reduce the number of required RF chains, the so-called beamspace MIMO and the switched-beam selection have been recently proposed in mmWave/massive MIMO systems [16]–[20]. By employing the discrete lens array (DLA) (which induces negligible performance loss), a conventional spatial channel can be transformed to a beamspace channel so as to capture the channel sparsity at mmWave frequencies. Since each beam corresponds to a single RF chain in beamspace MIMO [17], [18], only a small number of beams can be selected according to the sparse beamspace channel for reduction of required RF chains. For switched-beam schemes (e.g., Butler method [21]), a fixed number of beams are

generated and pointed to different predetermined directions to cover the whole cell [19], [20].

In view of significant system throughput gains over the conventional single-cell design and a limited amount of information exchange (e.g., scheduling, power allocation) among the coordinated BSs, the multicell coordinated beamforming (CoBF) has been studied extensively both in academia and industry recently [22], [23]. Nevertheless, CSI is never perfectly known and the CSI uncertainty may vary from one BS to another, depending on the bandwidth of the network backhaul. To tackle CSI errors, two major types of robust transmit designs have been studied. One is the worst-case robust design, where CSI errors are constrained in a bounded set, and the other is the outage-constrained robust design, where CSI errors are modeled in a probabilistic fashion. The main bottleneck for the latter consists in intractable or even no explicit expression for the probabilistic constraint [24].

When an explicit expression exists for the probabilistic constraint, CoBF algorithms have been proposed such as using successive convex approximation (SCA) in [25] and using block successive upper bound minimization in [26]. For the case of no explicit expression for the probabilistic constraint, Monte Carlo sampling based approaches [27]–[29] have been proposed by generating channel realizations following the distribution of the channel uncertainty to obtain an approximate outage constraint. However, its computational complexity is high and its performance is sensitive to the accuracy of channel distribution information in some scenarios. The second approach is to find a restrictive (i.e., conservative) convex approximation to each outage constraint in terms of signal-to-interference-plus-noise ratio (SINR) [30], thereby obtaining an approximate CoBF solution with good performance and low complexity. Moreover, [31], [32] also studied the same problem for more specific scenarios. On the other hand, the distributed CoBF algorithm design has also been studied for practical applications [25], where multiple-input single-output (MISO) interference channel (IFC) is considered, but the corresponding design for interference broadcast channel (IBC) within the HetNet scenario still remains an open problem.

In this paper, we consider the hybrid coordinated beamforming (HyCoBF) within the HetNet scenario, as illustrated in Figure 1, where a large-scale antenna MBS equipped with an analog beamformer cascaded with a digital beamformer, and a conventional multiple-antenna FBS equipped with a digital beamformer, serve their respective users. The proposed HyCoBF design includes a beam selection algorithm for the analog beamformer at MBS and a digital CoBF design for both MBS and FBS. To the best of our knowledge, there is no existing work, that studies the HyCoBF in massive MIMO enabled HetNet such that the HyCoBF design is robust against CSI uncertainty under the preassigned outage probability constraints. We formulate such problem as a total power minimization problem, subject to preassigned users' outage probability constraints. However, this problem is almost intractable due to no closed-form

expressions for probabilistic outage constraints (induced by the CSI uncertainty), and due to a high-dimensional combinatorial optimization problem involved in the analog beam selection problem [24].

Because none of the existing beam search algorithms, such as greedy-pursuit algorithms [18], sum-rate-maximization based beam allocation algorithms (for a switched-beam based massive MIMO system) [19], and interference-aware (IA) beam selection methods (that consider multiuser interference) [17], can be applied due to complicated interference links involved in HetNet. Motivated by a well-known detection scheme for seismic events in geophysical signal processing, called the single most likely replacement (SMLR) detector [33], the proposed low-complexity analog beam selection algorithm tries to search for the best subset from the columns of a discrete Fourier transform (DFT) matrix codebook by enhancing the total channel power of macrocell user equipments (MUEs) and reducing the interference channel power of femtocell user equipments (FUEs) in the meantime.

With the designed analog beamformer applied, the power minimization problem (only for the digital CoBF design) is still hard to solve, due to no closed-form expressions for the probabilistic constraints. By virtue of semidefinite relaxation (SDR) technique, we reformulate the digital CoBF problem into a semidefinite program (SDP), where an extended Bernstein-type inequality is derived for finding conservative convex approximations to the original nonconvex probabilistic constraints. Furthermore, we develop a distributed implementation for the obtained digital CoBF by the alternating direction method of multipliers (ADMM) [34]. Finally, some numerical simulations are presented to demonstrate the efficacy of the proposed HyCoBF design.

The rest of this paper is organized as follows. In Section II, we present the system model and the problem formulation. The HyCoBF design (for both analog beam selection algorithm and robust digital CoBF) is proposed in Section III. In Section IV, we present a distributed implementation for the robust digital CoBF using ADMM. Simulation results are then provided in Section V to demonstrate the efficacy of the proposed HyCoBF design. Finally, some conclusions are drawn in Section VI.

Notation: \mathbb{R}^n , \mathbb{R}_+^n (\mathbb{R}_{++}^n , \mathbb{R}_-^n), \mathbb{S}^n (\mathbb{S}_+^n), \mathbb{C}^n , \mathbb{H}^n and $\mathbb{C}^{m \times n}$ stand for the sets of n -dimensional real vectors, non-negative (positive, non-positive) real n -vectors, real symmetric (positive semidefinite) matrices, complex n -vectors, $n \times n$ Hermitian matrices and $m \times n$ complex matrices, respectively. \mathbf{I}_n and $\mathbf{e}(i)$, respectively, denote the $n \times n$ identity matrix, and the i -th unit column vector of proper dimension. The superscripts T and H represent the matrix transpose and conjugate transpose, respectively. $\|\cdot\|$ and $\|\cdot\|_F$ denote the vector Euclidean norm and matrix Frobenius norm, respectively. The trace of matrix \mathbf{A} is denoted as $\text{Tr}(\mathbf{A})$. $\sup(C)$, \mathbf{A}^\dagger and $\lambda_{\max}(\mathbf{A})$ denote the supremum of a nonempty set C , the pseudoinverse and the maximum eigenvalue of matrix \mathbf{A} . $\mathbf{A} \geq \mathbf{0}$ means that \mathbf{A} is a positive semidefinite (PSD) matrix and $\mathbf{A}^{1/2} \geq \mathbf{0}$ is a square root of \mathbf{A} ; $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ represent

real and imaginary parts of the argument; $\mathcal{A} \setminus \mathcal{B}$ denotes the set by eliminating the elements of $\mathcal{A} \cap \mathcal{B}$ from \mathcal{A} ; $\mathcal{I}_K = \{1, \dots, K\}$ and $\{a_k\}$ ($\{\mathbf{w}_k\}$, $\{\mathbf{W}_k\}$) denotes the set of all scalars a_k (vectors \mathbf{w}_k , matrices \mathbf{W}_k) with the subscript k covering all the admissible integers that are defined in the context; $\mathcal{CN}(\cdot)$ denotes the complex Gaussian distribution; $\ln(\cdot)$, $\text{Pr}\{\cdot\}$ and $\mathbb{E}\{\cdot\}$ stand for the natural log function, probability function, and expectation operator, respectively.

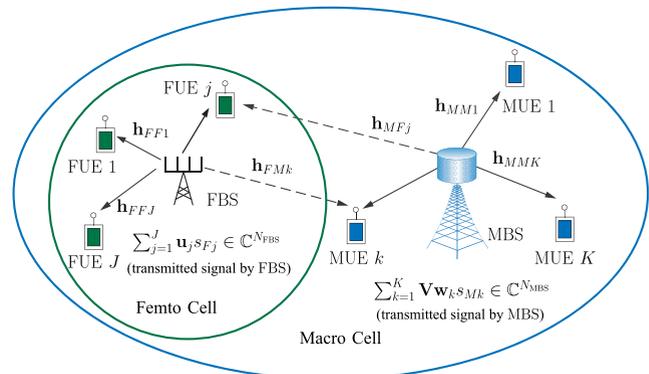


FIGURE 1. Illustration of a two-tier heterogeneous network, where the MBS employs large-scale antennas and the FBS employs conventional multiple antennas.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

We consider the time-division duplex (TDD) downlink multiuser transmission with full spectrum reuse in HetNet, which consists of an MBS equipped with large-scale N_{MBS} antennas, an FBS equipped with N_{FBS} antennas, K single-antenna MUEs and J single-antenna FUEs, as illustrated in Fig. 1. Instead of adopting FD beamforming, for which each antenna is connected with one distinct RF chain, we consider HyCoBF at MBS with N_{RF} RF chains, where $N_{\text{MBS}} \gg N_{\text{RF}} \geq K$. Specifically, the HyBF vector for MUE k at the MBS is given by $\mathbf{V}\mathbf{w}_k \in \mathbb{C}^{N_{\text{MBS}}}$, where $\mathbf{V} \in \mathbb{C}^{N_{\text{MBS}} \times N_{\text{RF}}}$ and $\mathbf{w}_k \in \mathbb{C}^{N_{\text{RF}}}$ denote the analog beamforming matrix and digital beamforming vector, respectively.

Typically, the analog beamformer is implemented using a phase-shifter network with constant modulus for each entry of \mathbf{V} . In contrast to the full-connected HyBF structure, the HyBF with a selection structure is more economical and energy efficient [19] and hence is considered in this work. The analog beamformer can be implemented using a RF switch, followed by a fixed DFT beamformer using RF phase-shifter network [15]. Let $\mathbf{F} \in \mathbb{C}^{N_{\text{MBS}} \times N_{\text{MBS}}}$ be the N_{MBS} -point DFT matrix, and

$$\mathcal{B} \triangleq \{\mathbf{b} \in \{1, \dots, N_{\text{MBS}}\}^{N_{\text{RF}}} \mid b_i \neq b_j, \forall i \neq j\}, \quad (1)$$

where b_i denotes the i th component of \mathbf{b} , collects all permutations of any subset of $\{1, \dots, N_{\text{MBS}}\}$ with cardinality N_{RF} . The selected analog beamforming matrix \mathbf{V} and the RF switch matrix $\mathbf{A}(\mathbf{b})$ consisting of N_{RF} unit column vectors of

dimension N_{MBS} can then be conveniently expressed as

$$\begin{aligned} \mathbf{A}(\mathbf{b}) &= [\mathbf{e}(b_1), \dots, \mathbf{e}(b_{N_{\text{RF}}})], \mathbf{b} \in \mathcal{B}, \\ \mathbf{V} &= \mathbf{F}\mathbf{A}(\mathbf{b}) \in \mathbb{C}^{N_{\text{MBS}} \times N_{\text{RF}}}, \end{aligned} \quad (2)$$

which contains N_{RF} distinct columns of \mathbf{F} . Let s_{Mk} and s_{Fj} denote the transmit signals intended for MUE k and FUE j , respectively. Without loss of generality (w.l.o.g.), assume that $\mathbb{E}\{|s_{Fj}|^2\} = 1$ and $\mathbb{E}\{|s_{Mk}|^2\} = 1$. Then, the received signal of MUE k is given by

$$\begin{aligned} y_{Mk} &= \mathbf{h}_{MMk}^H \mathbf{V} \mathbf{w}_k s_{Mk} + \sum_{l=1, l \neq k}^K \mathbf{h}_{MMk}^H \mathbf{V} \mathbf{w}_l s_{Ml} \\ &+ \sum_{j=1}^J \mathbf{h}_{FMk}^H \mathbf{u}_j s_{Fj} + n_{Mk}, \end{aligned} \quad (3)$$

where $\mathbf{h}_{MMk} \in \mathbb{C}^{N_{\text{MBS}}}$ and $\mathbf{h}_{FMk} \in \mathbb{C}^{N_{\text{FBS}}}$, respectively, denote the channel vector from MBS and FBS to MUE k , $n_{Mk} \sim \mathcal{CN}(0, \sigma_{Mk}^2)$ is the additive white Gaussian noise (AWGN) at MUE k with noise variance $\sigma_{Mk}^2 > 0$, and $\mathbf{u}_j \in \mathbb{C}^{N_{\text{FBS}}}$ is the beamforming vector for FUE j at FBS. Similarly, the received signal of FUE j is given by

$$\begin{aligned} y_{Fj} &= \mathbf{h}_{FFj}^H \mathbf{u}_j s_{Fj} + \sum_{m=1, m \neq j}^J \mathbf{h}_{FFj}^H \mathbf{u}_m s_{Fm} \\ &+ \sum_{k=1}^K \mathbf{h}_{MFj}^H \mathbf{V} \mathbf{w}_k s_{Mk} + n_{Fj}, \end{aligned} \quad (4)$$

where $\mathbf{h}_{MFj} \in \mathbb{C}^{N_{\text{MBS}}}$ and $\mathbf{h}_{FFj} \in \mathbb{C}^{N_{\text{FBS}}}$, respectively, denote the channel vectors from MBS and FBS to FUE j , and $n_{Fj} \sim \mathcal{CN}(0, \sigma_{Fj}^2)$ is the AWGN at FUE j . Then, the SINRs of MUE k and FUE j can be expressed as

$$\text{SINR}_{Mk} = \frac{|\mathbf{h}_{MMk}^H \mathbf{V} \mathbf{w}_k|^2}{\sum_{l=1, l \neq k}^K |\mathbf{h}_{MMk}^H \mathbf{V} \mathbf{w}_l|^2 + \sum_{j=1}^J |\mathbf{h}_{FMk}^H \mathbf{u}_j|^2 + \sigma_{Mk}^2}, \quad (5a)$$

$$\text{SINR}_{Fj} = \frac{|\mathbf{h}_{FFj}^H \mathbf{u}_j|^2}{\sum_{m=1, m \neq j}^J |\mathbf{h}_{FFj}^H \mathbf{u}_m|^2 + \sum_{k=1}^K |\mathbf{h}_{MFj}^H \mathbf{V} \mathbf{w}_k|^2 + \sigma_{Fj}^2}. \quad (5b)$$

B. PROBLEM FORMULATION

Due to the channel reciprocity of TDD systems, the downlink CSI can be estimated¹ from the uplink pilot symbols transmitted from the users [35]. However, besides the CSI estimation error, some more CSI uncertainties are also inevitable due to delays in CSI acquisition, partial CSI acquisition and hardware impairment, etc. With regard to imperfect CSI, the actual channels of MUEs and FUEs can be modeled as [36]

$$\mathbf{h}_{MMk} = \hat{\mathbf{h}}_{MMk} + \mathbf{e}_{MMk}, \quad \mathbf{h}_{MFj} = \hat{\mathbf{h}}_{MFj} + \mathbf{e}_{MFj}, \quad (6a)$$

$$\mathbf{h}_{FFj} = \hat{\mathbf{h}}_{FFj} + \mathbf{e}_{FFj}, \quad \mathbf{h}_{FMk} = \hat{\mathbf{h}}_{FMk} + \mathbf{e}_{FMk}, \quad (6b)$$

where $\hat{\mathbf{h}}_{MMk}$, $\hat{\mathbf{h}}_{MFj} \in \mathbb{C}^{N_{\text{MBS}}}$, and $\hat{\mathbf{h}}_{FFj}$, $\hat{\mathbf{h}}_{FMk} \in \mathbb{C}^{N_{\text{FBS}}}$ are the given channel estimates and known to MBS and FBS, and

CSI errors \mathbf{e}_{MMk} , $\mathbf{e}_{MFj} \in \mathbb{C}^{N_{\text{MBS}}}$ and \mathbf{e}_{FFj} , $\mathbf{e}_{FMk} \in \mathbb{C}^{N_{\text{FBS}}}$ are modeled as

$$\mathbf{e}_{MMk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{MMk}), \quad \mathbf{e}_{MFj} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{MFj}), \quad (7a)$$

$$\mathbf{e}_{FFj} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{FFj}), \quad \mathbf{e}_{FMk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{FMk}), \quad (7b)$$

in which all \mathbf{C}_{MMk} and \mathbf{C}_{FFj} are positive definite. Thus, the outage constrained robust HyCoBF design problem can be formulated as

$$\min_{\mathbf{V}, \{\mathbf{w}_k\}, \{\mathbf{u}_j\}} \sum_{k \in \mathcal{I}_K} \|\mathbf{V} \mathbf{w}_k\|^2 + \sum_{j \in \mathcal{I}_J} \|\mathbf{u}_j\|^2 \quad (8a)$$

$$\text{s.t. } \Pr(\text{SINR}_{Mk} \geq \gamma_{Mk}) \geq 1 - \rho_{Mk}, \quad \forall k \in \mathcal{I}_K, \quad (8b)$$

$$\Pr(\text{SINR}_{Fj} \geq \gamma_{Fj}) \geq 1 - \rho_{Fj}, \quad \forall j \in \mathcal{I}_J, \quad (8c)$$

$$\mathbf{V} \in \{\mathbf{F}\mathbf{A}(\mathbf{b}) \mid \mathbf{b} \in \mathcal{B}\} \text{ (cf. (2))}, \quad (8d)$$

where γ_{Mk} and γ_{Fj} are target SINRs for MUEs and FUEs, respectively, ρ_{Mk} and ρ_{Fj} denote the associated outage probabilities, respectively.

III. PROPOSED OUTAGE CONSTRAINED HyCoBF DESIGN

Solving the HyCoBF design problem (8) is a daunting task, partly because the nonconvex probabilistic constraints (8b) and (8c) do not have closed-form expressions in general [24]; partly because the analog beam selection constraint (8d) makes the reformulation of (8) into a tractable convex problem almost formidable. We handle this problem by decoupling the design of \mathbf{V} (analog beamforming) and the joint design of \mathbf{w}_k and \mathbf{u}_j (digital CoBF), and they are presented in the following subsections, respectively.

A. ANALOG BEAMFORMING ALGORITHM

Motivated by the fact that SINRs of MUEs are larger for larger received signal power, and SINRs of FUEs are larger for smaller inter-cell interference power, we propose a new beam selection criterion named power ratio maximization (PRM), by maximizing the ratio of the total channel power of MUEs to the total inter-cell interference channel power from MBS to FUEs, as follows:

$$\mathbf{V}^* = \mathbf{F}\mathbf{A}(\mathbf{b}^*) \text{ (cf. (2))}, \quad (9)$$

$$\mathbf{b}^* = \arg \max_{\mathbf{b} \in \mathcal{B}} \left\{ \mathcal{J}(\mathbf{b}) \triangleq \frac{\|\hat{\mathbf{H}}_{MM}^H \mathbf{F}\mathbf{A}(\mathbf{b})\|_F^2}{\|\hat{\mathbf{H}}_{MF}^H \mathbf{F}\mathbf{A}(\mathbf{b})\|_F^2} \right\}, \quad (10)$$

where \mathcal{B} was defined in (1), $\hat{\mathbf{H}}_{MM} \triangleq [\hat{\mathbf{h}}_{MM1}, \dots, \hat{\mathbf{h}}_{MMK}]$ and $\hat{\mathbf{H}}_{MF} \triangleq [\hat{\mathbf{h}}_{MF1}, \dots, \hat{\mathbf{h}}_{MFJ}]$. Obviously, the aim of (10) is to select N_{RF} best distinct beams from a total of N_{MBS} beams, by maximizing $\mathcal{J}(\mathbf{b})$. However, obtaining the global optimum of (10) is computationally prohibitive. For instance, supposing that $N_{\text{MBS}} = 128$ and $N_{\text{RF}} = K = 16$, the number of possible beam selection combinations is of the order $\mathcal{O}(10^{20})$. In view of this, a low-complexity beam selection

¹Essentially, for a quasi-stationary fading channel, MBS can only estimate partial CSI in a hybrid-structured TDD massive MIMO system with the received pilot or training signal due to limited RF chains. However, it can repeatedly obtain other partial CSI within a coherence interval until the full CSI estimates are obtained.

Algorithm 1 Proposed Analog Beam Selection Algorithm (PRM)

-
- 1: Given $\hat{\mathbf{H}}_{MM}, \hat{\mathbf{H}}_{MF}$;
 - 2: Initialize $\mathbf{b} = (b_1, b_2, \dots, b_{N_{RF}}) \in \mathcal{B}$,
and let $\mathcal{S} = \mathcal{I}_{N_{MBS}} \setminus \{b_1, b_2, \dots, b_{N_{RF}}\}$;
 - 3: Compute $\mathcal{J}(\mathbf{b})$ by (10);
 - 4: **repeat**
 - 5: Obtain $j_\ell = \arg \max_{j \in \mathcal{S}} \mathcal{J}(\mathbf{b}_\ell(j))$, $\ell = 1, \dots, N_{RF}$,
where $\mathbf{b}_\ell(j) \triangleq (b_1, \dots, b_{\ell-1}, j, b_{\ell+1}, \dots, b_{N_{RF}})$;
 - 6: Obtain $l = \arg \max \{\mathcal{J}(\mathbf{b}_\ell(j_\ell))\}$, $\ell = 1, \dots, N_{RF}$;
 - 7: If $\mathcal{J}(\mathbf{b}_l(j_l)) > \mathcal{J}(\mathbf{b})$, update $\mathbf{b} := \mathbf{b}_l(j_l)$, $\mathcal{S} := \mathcal{I}_{N_{MBS}} \setminus \{b_1, \dots, b_{N_{RF}}\}$; and $\mathcal{J}(\mathbf{b}) := \mathcal{J}(\mathbf{b}_l(j_l))$;
 - 8: **until** $\mathcal{J}(\mathbf{b}_l(j_l)) \leq \mathcal{J}(\mathbf{b})$.
 - 9: Output selected analog beamforming matrix $\mathbf{V}^* = \mathbf{F}\mathbf{A}(\mathbf{b})$.
-

algorithm for solving (10) is proposed, by employing the idea of SMLR used in seismic deconvolution for detecting a Bernoulli-Gaussian signal with nonzero magnitudes [33]. The proposed algorithm is summarized in Algorithm 1, that updates only one component in \mathbf{b} by maximizing $\mathcal{J}(\mathbf{b})$ at each iteration. The associated objective value $\mathcal{J}(\mathbf{b})$ increases monotonically whenever \mathbf{b} is updated, so the convergence can be guaranteed. The computational cost mainly consists in calculating $\mathcal{J}(\mathbf{b})$ $N_{RF} \times (N_{MBS} - N_{RF})$ times (in step 5 of Algorithm 1) at each iteration. Surely, an initial beam switch vector \mathbf{b} is needed to initialize the Algorithm 1. A fast beam selection algorithm proposed in [17], that selects best N_{RF} distinct beam indices according to row magnitudes of $\hat{\mathbf{H}}_{MM}^H \mathbf{F}$, can be used to obtain the initial \mathbf{b} for Algorithm 1.

B. OUTAGE CONSTRAINED DIGITAL CoBF: CONSERVATIVE APPROXIMATION

Even when the analog beamformer \mathbf{V}^* is given, problem (8) is still intractable due to the probabilistic constraints ((8b) and (8c)) that do not have closed-form expressions in general. It is noticeable that problem (8) can be thought of as a dimension-reduced problem with effective channel $\hat{\mathbf{h}}_{MMk}^{(\text{eff})} = (\mathbf{V}^*)^H \hat{\mathbf{h}}_{MMk} \in \mathbb{C}^{N_{RF}}$ and $\hat{\mathbf{h}}_{MFj}^{(\text{eff})} = (\mathbf{V}^*)^H \hat{\mathbf{h}}_{MFj} \in \mathbb{C}^{N_{RF}}$ for MBS. Applying SDR (i.e., replacing $\mathbf{w}_k \mathbf{w}_k^H$ by $\mathbf{W}_k \geq \mathbf{0}$ and $\mathbf{u}_j \mathbf{u}_j^H$ by $\mathbf{U}_j \geq \mathbf{0}$) [37] to problem (8) yields

$$\min_{\{\mathbf{W}_k\}, \{\mathbf{U}_j\}} \sum_{k \in \mathcal{I}_K} \text{Tr}(\mathbf{V}^* \mathbf{W}_k (\mathbf{V}^*)^H) + \sum_{j \in \mathcal{I}_J} \text{Tr}(\mathbf{U}_j) \quad (11a)$$

$$\text{s.t. Pr} \left\{ \delta_1^H \mathbf{Q}_{MMk} \delta_1 + \delta_2^H \mathbf{Q}_{FMk} \delta_2 + 2\text{Re}\{\delta_1^H \mathbf{r}_{MMk}\} + 2\text{Re}\{\delta_2^H \mathbf{r}_{FMk}\} + c_{Mk} \geq 0 \right\} \geq 1 - \rho_{Mk}, \quad \forall k \in \mathcal{I}_K, \quad (11b)$$

$$\text{Pr} \left\{ \delta_1^H \mathbf{Q}_{MFj} \delta_1 + \delta_2^H \mathbf{Q}_{FFj} \delta_2 + 2\text{Re}\{\delta_1^H \mathbf{r}_{MFj}\} + 2\text{Re}\{\delta_2^H \mathbf{r}_{FFj}\} + c_{Fj} \geq 0 \right\} \geq 1 - \rho_{Fj}, \quad \forall j \in \mathcal{I}_J, \quad (11c)$$

$$\mathbf{W}_k \geq \mathbf{0}, \quad \forall k \in \mathcal{I}_K, \quad \mathbf{U}_j \geq \mathbf{0}, \quad \forall j \in \mathcal{I}_J, \quad (11d)$$

where \mathbf{V}^* is the analog beamformer obtained by Algorithm 1, $\delta_1 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{RF}})$, $\delta_2 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{FBS}})$, and the remaining parameters in (11) are defined in (12) at the top of next page.

A good approach, as proposed in [30], for handling (11b) and (11c), is to find conservative convex approximations to the constraints (11b) and (11c). Such approach can safely approximate the two probability inequalities by convex ones so that the resulting algorithm is computationally tractable. Specifically, the Bernstein-type inequality [38] for finding such conservative convex approximations has been applied to robust digital beamforming design under similar constraints as in problem (11) for the single-cell case [30]. However, the Bernstein-type inequality in [30] is not directly applicable to problem (11), mainly owing to different antenna deployments at MBS and FBS, resulting in channel vectors of different dimensions. In view of this, we need an extension form of the Bernstein-type inequality as derived in the following lemma.

Lemma 1: Let $\delta_1 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{RF}})$, $\delta_2 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{FBS}})$, $\mathbf{Q}_{MMk} \in \mathbb{H}^{N_{RF}}$, $\mathbf{Q}_{FMk} \in \mathbb{H}^{N_{FBS}}$, $\mathbf{r}_{MMk} \in \mathbb{C}^{N_{RF}}$, $\mathbf{r}_{FMk} \in \mathbb{C}^{N_{FBS}}$, and define

$$g_1(\delta_1, \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}) \triangleq \delta_1^H \mathbf{Q}_{MMk} \delta_1 + 2\text{Re}\{\mathbf{r}_{MMk}^H \delta_1\}, \quad (13)$$

$$g_2(\delta_2, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) \triangleq \delta_2^H \mathbf{Q}_{FMk} \delta_2 + 2\text{Re}\{\mathbf{r}_{FMk}^H \delta_2\}. \quad (14)$$

Then, the following inequality holds true, $\forall k \in \mathcal{I}_K$,

$$\begin{aligned} & \Pr \left\{ g_1(\delta_1, \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}) + g_2(\delta_2, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) \right. \\ & \quad \left. \geq \Upsilon(\ln(1/\rho_{Mk}) \mid \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) \right\} \\ & \geq 1 - \rho_{Mk}, \end{aligned} \quad (15)$$

where $\Upsilon: \mathbb{R}_{++} \rightarrow \mathbb{R}$ is defined as

$$\begin{aligned} \Upsilon(\ln(1/\rho_{Mk}) \mid \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) & \triangleq \text{Tr}(\mathbf{Q}_{MMk}) \\ & + \text{Tr}(\mathbf{Q}_{FMk}) - \ln(1/\rho_{Mk}) \cdot \lambda^+(\mathbf{Q}_{MMk}, \mathbf{Q}_{FMk}) - \alpha_{Mk} \\ & \cdot \sqrt{\|\mathbf{Q}_{MMk}\|_F^2 + 2\|\mathbf{r}_{MMk}\|^2 + \|\mathbf{Q}_{FMk}\|_F^2 + 2\|\mathbf{r}_{FMk}\|^2}, \end{aligned} \quad (16)$$

in which $\alpha_{Mk} = \sqrt{2 \ln(1/\rho_{Mk})}$ and $\lambda^+(\mathbf{Q}_{MMk}, \mathbf{Q}_{FMk}) \triangleq \max\{\lambda_{\max}(-\mathbf{Q}_{MMk}), \lambda_{\max}(-\mathbf{Q}_{FMk}), 0\}$.

The proof of Lemma 1 is relegated to Appendix A. By virtue of Lemma 1, problem (11) can be approximated as the following convex SDP and the detailed derivations are relegated to Appendix B:

$$\min_{\{\mathbf{W}_k\}, \{\mathbf{U}_j\}, P_M, P_F, \mathbf{t}_M, \mathbf{t}_F} P_M + P_F \quad (17a)$$

$$\text{s.t. } (\{\mathbf{W}_k\}, \{\mathbf{U}_j\}, \mathbf{t}_M, P_M) \in \mathcal{C}_M, \quad (17b)$$

$$(\{\mathbf{W}_k\}, \{\mathbf{U}_j\}, \mathbf{t}_F, P_F) \in \mathcal{C}_F, \quad (17c)$$

$$\mathbf{W}_k \geq \mathbf{0}, \quad \forall k \in \mathcal{I}_K, \quad \mathbf{U}_j \geq \mathbf{0}, \quad \forall j \in \mathcal{I}_J, \quad (17d)$$

where \mathcal{C}_M and \mathcal{C}_F defined in (41) and (43) in Appendix B denote the conservative convex approximations to the constraint sets associated with (11b) and (11c), respectively;

$$\mathbf{Q}_{MMk} \triangleq \tilde{\mathbf{C}}_{MMk}^{1/2} \mathbf{B}_k \tilde{\mathbf{C}}_{MMk}^{1/2}; \mathbf{Q}_{FMk} \triangleq \mathbf{C}_{FMk}^{1/2} \mathbf{D} \mathbf{C}_{FMk}^{1/2}; \mathbf{Q}_{FFj} \triangleq \mathbf{C}_{FFj}^{1/2} \mathbf{F}_j \mathbf{C}_{FFj}^{1/2}; \mathbf{Q}_{MFj} \triangleq \tilde{\mathbf{C}}_{MFj}^{1/2} \mathbf{G} \tilde{\mathbf{C}}_{MFj}^{1/2}; \quad (12a)$$

$$\mathbf{r}_{MMk} \triangleq \tilde{\mathbf{C}}_{MMk}^{1/2} \mathbf{B}_k \hat{\mathbf{h}}_{MMk}^{(\text{eff})}; \mathbf{r}_{FMk} \triangleq \mathbf{C}_{FMk}^{1/2} \mathbf{D} \hat{\mathbf{h}}_{FMk}; \mathbf{r}_{FFj} \triangleq \mathbf{C}_{FFj}^{1/2} \mathbf{F}_j \hat{\mathbf{h}}_{FFj}; \mathbf{r}_{MFj} \triangleq \tilde{\mathbf{C}}_{MFj}^{1/2} \mathbf{G} \hat{\mathbf{h}}_{MFj}^{(\text{eff})}; \quad (12b)$$

$$c_{Mk} \triangleq \underbrace{(\hat{\mathbf{h}}_{MMk}^{(\text{eff})})^H \mathbf{B}_k \hat{\mathbf{h}}_{MMk}^{(\text{eff})}}_{\triangleq c_{MMk}} + \underbrace{\hat{\mathbf{h}}_{FMk}^H \mathbf{D} \hat{\mathbf{h}}_{FMk}}_{\triangleq c_{FMk}} - \sigma_{Mk}^2; c_{Fj} \triangleq \underbrace{\hat{\mathbf{h}}_{FFj}^H \mathbf{F}_j \hat{\mathbf{h}}_{FFj}}_{\triangleq c_{FFj}} + \underbrace{(\hat{\mathbf{h}}_{MFj}^{(\text{eff})})^H \mathbf{G} \hat{\mathbf{h}}_{MFj}^{(\text{eff})}}_{\triangleq c_{MFj}} - \sigma_{Fj}^2; \quad (12c)$$

$$\mathbf{B}_k \triangleq \gamma_{Mk}^{-1} \mathbf{W}_k - \sum_{l \neq k}^K \mathbf{W}_l; \mathbf{D} \triangleq - \sum_{j=1}^J \mathbf{U}_j; \quad (12d)$$

$$\mathbf{F}_j \triangleq \gamma_{Fj}^{-1} \mathbf{U}_j - \sum_{l \neq j}^J \mathbf{U}_l; \mathbf{G} \triangleq - \sum_{k=1}^K \mathbf{W}_k; \quad (12e)$$

$$\tilde{\mathbf{C}}_{MMk} \triangleq \mathbb{E} \left[(\mathbf{V}^*)^H \mathbf{e}_{MMk} \mathbf{e}_{MMk}^H \mathbf{V}^* \right] = (\mathbf{V}^*)^H \mathbf{C}_{MMk} \mathbf{V}^* \text{ (cf. (7a))}; \quad (12f)$$

$$\tilde{\mathbf{C}}_{MFj} \triangleq \mathbb{E} \left[(\mathbf{V}^*)^H \mathbf{e}_{MFj} \mathbf{e}_{MFj}^H \mathbf{V}^* \right] = (\mathbf{V}^*)^H \mathbf{C}_{MFj} \mathbf{V}^* \text{ (cf. (7a))}. \quad (12g)$$

$\mathbf{t}_M \in \mathbb{R}^{3K}$, $\mathbf{t}_F \in \mathbb{R}^{3J}$, P_M (denoting transmit power of MBS), P_F (denoting transmit power of FBS) are auxiliary variables. Note that \mathbf{W}_k^* and \mathbf{U}_j^* (the solution of problem (17)) may not be rank-one matrices. If they are of rank one, i.e., $\mathbf{W}_k^* = \mathbf{w}_k^* (\mathbf{w}_k^*)^H$ and $\mathbf{U}_j^* = \mathbf{u}_j^* (\mathbf{u}_j^*)^H$. Then, given the reference point $\mathbf{V} = \mathbf{V}^*$ (obtained by Algorithm 1), the block minimizer of (8) with respect to the remaining block variables (i.e., $\{\mathbf{w}_k\}$ and $\{\mathbf{u}_j\}$) can be directly obtained as \mathbf{w}_k^* and \mathbf{u}_j^* ; otherwise, Gaussian randomization [37] can be employed to obtain a rank-one approximate solution.

Off-the-shelve convex solvers (e.g., CVX) can be used to obtain the centralized solution of problem (17). As previously mentioned, distributed solution to (17) is essential and indispensable to HetNet (due to frequent reconfiguration and scalability). Next, we present a distributed algorithm, for MBS and FBS to solve (17) with only local CSI and limited information exchange in a cooperative fashion.

IV. DISTRIBUTED ROBUST DIGITAL CoBF USING ADMM

Note that the two convex constraint sets \mathcal{C}_M and \mathcal{C}_F (cf. (41) and (43)) in problem (17) are coupled. We need to reformulate the feasible set of problem (17) into a pair of uncoupled constraint sets, denoted as $\tilde{\mathcal{C}}_M$ (cf. (19)) and $\tilde{\mathcal{C}}_F$ (cf. (20)), and some linear equality constraints that couple $\tilde{\mathcal{C}}_M$ and $\tilde{\mathcal{C}}_F$, so that ADMM can be employed for efficient distributed algorithm design, as detailed in Subsection IV-A. Some implementation issues and complexity analysis are then discussed in Subsection IV-B.

A. DISTRIBUTED CoBF ALGORITHM

Let us define the following vector variables:

$$\boldsymbol{\tau} \triangleq [\mathbf{a}^T, \mathbf{b}^T, \mathbf{y}^T]^T \in \mathbb{R}^{3(K+J)}, \quad (18a)$$

$$\mathbf{a} \triangleq [a_{MF1}, \dots, a_{MFJ}, a_{FM1}, \dots, a_{FMK}] \in \mathbb{R}_+^{(K+J)}, \quad (18b)$$

$$\mathbf{b} \triangleq [t_{MF1}, \dots, t_{MFJ}, t_{FM1}, \dots, t_{FMK}] \in \mathbb{R}_+^{(K+J)}, \quad (18c)$$

$$\mathbf{y} \triangleq [t_{M1}, \dots, t_{MK}, t_{F1}, \dots, t_{FJ}] \in \mathbb{R}_+^{(K+J)}. \quad (18d)$$

Then the feasible set of problem (17) can be equivalently re-expressed as follows:

$$\begin{aligned} \tilde{\mathcal{C}}_M \triangleq \{ & (P_M, \boldsymbol{\tau}_M, \{\mathbf{W}_k\}, \{a_{MMk}\}, \{t_{MMk}\}) \mid \\ & a_{MMk} + a_{FMk} + \ln(\rho_{Mk})t_{Mk} - \sigma_{Mk}^2 \\ & - \alpha_{Mk} \left\| [t_{MMk}, t_{FMk}]^T \right\| \geq 0, \forall k \in \mathcal{I}_K \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{MMk}) \\ \sqrt{2} \mathbf{r}_{MMk} \end{bmatrix} \right\| \leq t_{MMk}, \forall k \in \mathcal{I}_K \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{MFj}) \\ \sqrt{2} \mathbf{r}_{MFj} \end{bmatrix} \right\| \leq t_{MFj}, \forall j \in \mathcal{I}_J \\ & t_{Mk} \mathbf{I}_{N_{\text{RF}}} + \mathbf{Q}_{MMk} \geq \mathbf{0}, t_{Mk} \geq 0, \forall k \in \mathcal{I}_K \\ & t_{Fj} \mathbf{I}_{N_{\text{RF}}} + \mathbf{Q}_{MFj} \geq \mathbf{0}, t_{Fj} \geq 0, \forall j \in \mathcal{I}_J \\ & a_{MMk} \triangleq \text{Tr}(\mathbf{Q}_{MMk}) + c_{MMk}, \forall k \in \mathcal{I}_K \\ & a_{MFj} \triangleq \text{Tr}(\mathbf{Q}_{MFj}) + c_{MFj}, \forall j \in \mathcal{I}_J \\ & P_M \triangleq \sum_{k \in \mathcal{I}_K} \text{Tr}(\mathbf{V}^* \mathbf{W}_k (\mathbf{V}^*)^H), \\ & \boldsymbol{\tau}_M \triangleq \boldsymbol{\tau}, \mathbf{W}_k \geq \mathbf{0}, \forall k \in \mathcal{I}_K \}, \end{aligned} \quad (19)$$

$$\begin{aligned} \tilde{\mathcal{C}}_F \triangleq \{ & (P_F, \boldsymbol{\tau}_F, \{\mathbf{U}_j\}, \{a_{FFj}\}, \{t_{FFj}\}) \mid \\ & a_{FFj} + a_{MFj} + \ln(\rho_{Fj})t_{Fj} - \sigma_{Fj}^2 \\ & - \beta_{Fj} \left\| [t_{FFj}, t_{MFj}]^T \right\| \geq 0, \forall j \in \mathcal{I}_J \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{FFj}) \\ \sqrt{2} \mathbf{r}_{FFj} \end{bmatrix} \right\| \leq t_{FFj}, \forall j \in \mathcal{I}_J \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{FMk}) \\ \sqrt{2} \mathbf{r}_{FMk} \end{bmatrix} \right\| \leq t_{FMk}, \forall k \in \mathcal{I}_K \\ & t_{Fj} \mathbf{I}_{N_{\text{FBS}}} + \mathbf{Q}_{FFj} \geq \mathbf{0}, t_{Fj} \geq 0, \forall j \in \mathcal{I}_J \\ & t_{Mk} \mathbf{I}_{N_{\text{FBS}}} + \mathbf{Q}_{FMk} \geq \mathbf{0}, t_{Mk} \geq 0, \forall k \in \mathcal{I}_K \\ & a_{FFj} \triangleq \text{Tr}(\mathbf{Q}_{FFj}) + c_{FFj}, \forall j \in \mathcal{I}_J \\ & a_{FMk} \triangleq \text{Tr}(\mathbf{Q}_{FMk}) + c_{FMk}, \forall k \in \mathcal{I}_K \\ & P_F \triangleq \sum_{k \in \mathcal{I}_K} \text{Tr}(\mathbf{U}_j), \\ & \boldsymbol{\tau}_F \triangleq \boldsymbol{\tau}, \mathbf{U}_j \geq \mathbf{0}, \forall j \in \mathcal{I}_J \}. \end{aligned} \quad (20)$$

As a result, problem (17) is equivalent to the following problem:

$$\min_{\mathcal{W}, \mathcal{U}, \delta} P_M + P_F \quad (21a)$$

$$\text{s.t. } \mathcal{W} \triangleq (P_M, \boldsymbol{\tau}_M, \{\mathbf{W}_k\}, \{a_{MMk}\}, \{t_{MMk}\}) \in \tilde{\mathcal{C}}_M, \quad (21b)$$

$$\mathcal{U} \triangleq (P_F, \boldsymbol{\tau}_F, \{\mathbf{U}_j\}, \{a_{FFj}\}, \{t_{FFj}\}) \in \tilde{\mathcal{C}}_F, \quad (21c)$$

$$\boldsymbol{\delta} = \boldsymbol{\tau}_M = \boldsymbol{\tau}_F. \quad (21d)$$

Note that the auxiliary variable $\boldsymbol{\delta}$ serves as a public variable for MBS and FBS to update their local variables independently in the distributed algorithm to be presented next.

According to ADMM, the proposed distributed robust CoBF algorithm tries to solve the following penalty terms augmented problem:

$$\min_{\substack{\mathcal{W}, \mathcal{U} \\ \mathcal{X}_M, \mathcal{X}_F}} P_M + P_F + \frac{c}{2} \{ \|\boldsymbol{\delta} - \boldsymbol{\tau}_M\|^2 + \|\boldsymbol{\delta} - \boldsymbol{\tau}_F\|^2 \\ + (p_M - P_M)^2 + (p_F - P_F)^2 \} \quad (22a)$$

$$\text{s.t. } \mathcal{W} \triangleq (P_M, \boldsymbol{\tau}_M, \{\mathbf{W}_k\}, \{a_{MMk}\}, \{t_{MMk}\}) \in \tilde{\mathcal{C}}_M, \quad (22b)$$

$$\mathcal{U} \triangleq (P_F, \boldsymbol{\tau}_F, \{\mathbf{U}_j\}, \{a_{FFj}\}, \{t_{FFj}\}) \in \tilde{\mathcal{C}}_F, \quad (22c)$$

$$\mathcal{X}_M \triangleq (\boldsymbol{\delta}, p_M) = (\boldsymbol{\tau}_M, P_M) \in \mathbb{R}^{3(K+J)+1}, \quad (22d)$$

$$\mathcal{X}_F \triangleq (\boldsymbol{\delta}, p_F) = (\boldsymbol{\tau}_F, P_F) \in \mathbb{R}^{3(K+J)+1}, \quad (22e)$$

where $c > 0$ is a preassigned penalty parameter, p_M and p_F are auxiliary variables. The corresponding ADMM for solving (22) actually solves the dual optimization problem of (22), which is also a max-min problem defined as

$$\max_{\substack{\mu_M, \mu_F \in \mathbb{R} \\ \mathbf{v}_M, \mathbf{v}_F \in \mathbb{R}^{3(K+J)}}} \left\{ \min_{\substack{\mathcal{W} \in \tilde{\mathcal{C}}_M, \mathcal{U} \in \tilde{\mathcal{C}}_F \\ \mathcal{X}_M, \mathcal{X}_F \in \mathbb{R}^{3(K+J)+1}}} (g_M(\mathcal{W}, \mathcal{X}_M, \mathbf{v}_M, \mu_M) \\ + g_F(\mathcal{U}, \mathcal{X}_F, \mathbf{v}_F, \mu_F)) \right\} \quad (23)$$

where

$$g_M(\mathcal{W}, \mathcal{X}_M, \mathbf{v}_M, \mu_M) \\ \triangleq P_M + \frac{c}{2} \{ \|\boldsymbol{\delta} - \boldsymbol{\tau}_M\|^2 + (p_M - P_M)^2 \} \\ + \mathbf{v}_M^T (\boldsymbol{\delta} - \boldsymbol{\tau}_M) + \mu_M (p_M - P_M), \quad (24a)$$

$$g_F(\mathcal{U}, \mathcal{X}_F, \mathbf{v}_F, \mu_F) \\ \triangleq P_F + \frac{c}{2} \{ \|\boldsymbol{\delta} - \boldsymbol{\tau}_F\|^2 + (p_F - P_F)^2 \} \\ + \mathbf{v}_F^T (\boldsymbol{\delta} - \boldsymbol{\tau}_F) + \mu_F (p_F - P_F), \quad (24b)$$

in which $\{\mathbf{v}_M, \mu_M\}$ and $\{\mathbf{v}_F, \mu_F\}$ are the dual variables associated with constraints (22d) and (22e), respectively. The resulting iterative distributed algorithm is summarized in Algorithm 2, that can yield a global optimal solution of (17) after convergence by the following theorem (whose proof is given in Appendix C):

Theorem 1: Assume that problem (17) is solvable and strictly feasible (i.e., strong duality holds). Every limit point $\mathbf{W}_k(q)$ and $\mathbf{U}_j(q)$ generated by Algorithm 2 is an optimal solution of problem (17).

In Algorithm 2, steps 4-6 update the primal variables \mathcal{W} , \mathcal{U} , \mathcal{X}_M , and \mathcal{X}_F by solving the inner minimization problem of (23). Specifically, the primal variables \mathcal{W} , \mathcal{U} are updated

Algorithm 2 Proposed Distributed Robust CoBF Algorithm

- 1: **Input** a set of the initial variables $\{\boldsymbol{\delta}(0), \mathbf{v}_M(0), \mathbf{v}_F(0), \mu_M(0), \mu_F(0), p_M(0), p_F(0)\}$ that are known to both MBS and FBS; choose a penalty parameter $c > 0$ and an over-relaxation parameter $\theta \in (1, 2)$.
- 2: Set $q = 0$.
- 3: **repeat**
- 4: MBS and FBS update primal variables $\mathcal{W}(q+1)$ and $\mathcal{U}(q+1)$ by (25a) and (25b), respectively, and then further update $\{\boldsymbol{\tau}_M(q+1), P_M(q+1)\}$ in $\mathcal{W}(q+1)$ and $\{\boldsymbol{\tau}_F(q+1), P_F(q+1)\}$ in $\mathcal{U}(q+1)$ by (26a) and (26b), respectively;
- 5: MBS and FBS exchange local iterates $\boldsymbol{\tau}_M(q+1)$ and $\boldsymbol{\tau}_F(q+1)$;
- 6: MBS and FBS update $\boldsymbol{\delta}(q+1)$ using (28a), and then update $p_M(q+1)$, and $p_F(q+1)$ by (28b) and (28c), respectively;
- 7: MBS updates $\{\mathbf{v}_M(q+1), \mu_M(q+1)\}$ by (29a) and (29b), and FBS updates $\{\mathbf{v}_F(q+1), \mu_F(q+1)\}$ by (29c) and (29d);
- 8: Set $q := q + 1$;
- 9: Set $c := \min\{qc, 1\}$;
- 10: **until** the predefined stopping criterion is met.
- 11: **Output** $\mathbf{W}_k^* = \mathbf{W}_k(q+1)$, $\mathbf{U}_j^* = \mathbf{U}_j(q+1)$, $\forall k \in \mathcal{I}_K, j \in \mathcal{I}_J$ (yielded in step 4) and the associated beamformers $\mathbf{w}_k^*, \mathbf{u}_j^*$.

in step 4 by solving the following convex subproblems:

$$\mathcal{W}(q+1) = \arg \min_{\mathcal{W} \in \tilde{\mathcal{C}}_M} g_M(\mathcal{W}, \mathcal{X}_M(q), \mathbf{v}_M(q), \mu_M(q)), \quad (25a)$$

$$\mathcal{U}(q+1) = \arg \min_{\mathcal{U} \in \tilde{\mathcal{C}}_F} g_F(\mathcal{U}, \mathcal{X}_F(q), \mathbf{v}_F(q), \mu_F(q)), \quad (25b)$$

where q denotes the iteration number. Moreover, let $\mathbf{z}_M = [\boldsymbol{\tau}_M^T(q+1), P_M(q+1)]^T$ and $\mathbf{z}_F = [\boldsymbol{\tau}_F^T(q+1), P_F(q+1)]^T$. Then, over-relaxation strategy [34] for faster convergence is performed by

$$[\boldsymbol{\tau}_M^T(q+1), P_M(q+1)]^T := \theta \mathbf{z}_M + (1-\theta)[\boldsymbol{\delta}^T(q), P_M(q)]^T, \quad (26a)$$

$$[\boldsymbol{\tau}_F^T(q+1), P_F(q+1)]^T := \theta \mathbf{z}_F + (1-\theta)[\boldsymbol{\delta}^T(q), P_F(q)]^T, \quad (26b)$$

where $\theta \in (1, 2)$ is the over-relaxation parameter. Step 5 is interchange of local iterates $\boldsymbol{\tau}_M(q+1)$ and $\boldsymbol{\tau}_F(q+1)$ between MBS and FBS. Step 6 solves the convex subproblems:

$$(\mathcal{X}_M(q+1), \mathcal{X}_F(q+1)) \\ = \arg \min_{\mathcal{X}_M, \mathcal{X}_F} \left\{ g_M(\mathcal{W}(q+1), \mathcal{X}_M, \mathbf{v}_M(q), \mu_M(q)) \\ + g_F(\mathcal{U}(q+1), \mathcal{X}_F, \mathbf{v}_F(q), \mu_F(q)) \right\}, \quad (27)$$

thereby yielding the closed-form solutions:

$$\boldsymbol{\delta}(q+1) = \boldsymbol{\Omega}^\dagger(\tilde{\boldsymbol{\tau}}(q+1) - \tilde{\boldsymbol{v}}(q)/c), \quad (28a)$$

$$p_M(q+1) = P_M(q+1) - \mu_M(q)/c, \quad (28b)$$

$$p_F(q+1) = P_F(q+1) - \mu_F(q)/c, \quad (28c)$$

where $\tilde{\tau}(q+1) = [\tau_M^T(q+1), \tau_F^T(q+1)]^T$ and $\tilde{\mathbf{v}}(q) = [\mathbf{v}_M^T(q), \mathbf{v}_F^T(q)]^T$, and Ω is defined in (45c) in Appendix C. Finally, in step 7, the dual variables $\{\mathbf{v}_M^T, \mu_M\}$ and $\{\mathbf{v}_F^T, \mu_F\}$ are updated using

$$\mathbf{v}_M(q+1) = \mathbf{v}_M(q) + c(\delta(q+1) - \tau_M(q+1)), \quad (29a)$$

$$\mu_M(q+1) = \mu_M(q) + c(p_M(q+1) - P_M(q+1)), \quad (29b)$$

$$\mathbf{v}_F(q+1) = \mathbf{v}_F(q) + c(\delta(q+1) - \tau_F(q+1)), \quad (29c)$$

$$\mu_F(q+1) = \mu_F(q) + c(p_F(q+1) - P_F(q+1)). \quad (29d)$$

B. COMPLEXITY AND INFORMATION EXCHANGE OVERHEAD ANALYSIS

The dominant computation complexity of Algorithm 2 consists in step 4 where the two convex subproblems (25a) and (25b) can be solved by using the primal-dual interior-point method (IPM), thereby involving linear matrix inequality (LMI), second-order cone (SOC) and equality constraints. The worst-case complexity analysis for solving a conic problem defined in (46) by the primal-dual IPM is summarized in Appendix D [39]. Because the complexity induced by equality constraints are negligible compared to that of the inequality constraints, we apply this analysis to (25a) and (25b) by ignoring the complexity induced by equality constraints. Next, we focus on the complexity analysis for (25a).

Before presenting the complexity analysis for subproblem (25a), we need to reformulate it into the same problem form as (46). By epigraph representation and Schur complement, problem (25a) can be written as the following SDP:

$$\min_{\mathcal{W}} P_M + \frac{c}{2}(t_1 + t_2) + \mathbf{v}_M^T(q)(\delta(q) - \tau_M) + \mu_M(q)(p_M(q) - P_M) \quad (30a)$$

$$\text{s.t. } \mathcal{W} \triangleq (P_M, \tau_M, \{\mathbf{W}_k\}, \{a_{MMk}\}, \{t_{MMk}\}) \in \tilde{\mathcal{C}}_M, \quad (30b)$$

$$\begin{bmatrix} \mathbf{I} & \delta(q) - \tau_M \\ (\delta(q) - \tau_M)^T & t_1 \end{bmatrix} \succeq \mathbf{0}, \quad (30c)$$

$$\begin{bmatrix} 1 & p_M(q) - P_M \\ p_M(q) - P_M & t_2 \end{bmatrix} \succeq \mathbf{0}. \quad (30d)$$

Secondly, we need to convert the complex SDP (30) into the corresponding real SDP, which maintains the same problem type except for doubled problem dimension. For simplicity, our complexity analysis for solving (30) assumes w.l.o.g. that it is a real SDP.

By (19), the inequality constraints in (30b), (30c), and (30d), consist of $K + J$ LMI constraints of size 1; $2K + J$ LMI constraints of size N_{RF} ; two LMI constraints of size $3(K + J) + 1$ and 2, respectively, (i.e., $p = 3K + 2J + 2$, cf. (46b)); $K + J$ SOC constraints of size $N_{RF}^2 + N_{RF} + 1$ and K SOC constraints of size 6 (i.e., $m = p + 2K + J$; cf. (46c)). Moreover, the number of decision variables of problem (30) is on the order of $(4K + 2J)N_{RF}^2 \triangleq \tilde{n}$. Hence, the complexity order of the primal-dual IPM for solving (25a) is given by

$$C_{(25a)} = \sqrt{\tilde{\varphi}(\mathcal{K})} \cdot (\tilde{\mathcal{C}}_{\text{form}} + \tilde{\mathcal{C}}_{\text{fact}}) \cdot \ln(1/\epsilon), \quad (\text{by (49)}) \quad (31)$$

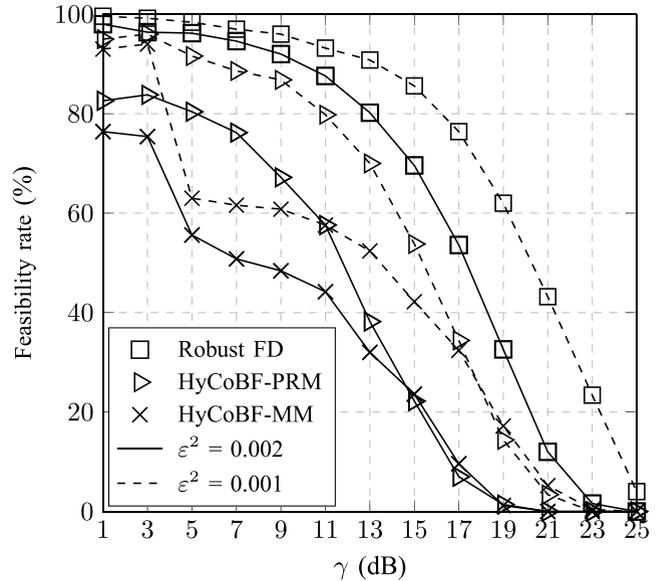


FIGURE 2. Feasibility rates of the various methods versus target SINR γ , for $N_{MBS} = 16, N_{RF} = 4, K = 4; N_{FBS} = 2, J = 2; \rho = 0.1, \epsilon^2 \in \{0.001, 0.002\}$.

where ϵ denotes the solution accuracy, and

$$\tilde{\varphi}(\mathcal{K}) = (2K + J)N_{RF} + 8K + 6J + 3, \quad (\text{by (47)})$$

$$\begin{aligned} \tilde{\mathcal{C}}_{\text{form}} = & \tilde{n}(2K + J)(N_{RF}^3 + \tilde{n}N_{RF}^2) \\ & + \tilde{n}(3(K + J) + 1)^3 + \tilde{n}^2(3(K + J) + 1)^2 \\ & + \tilde{n}(K + J)(\tilde{n} + 1) + \tilde{n}(8 + 4\tilde{n}) \\ & + \tilde{n}(K + J)(N_{RF}^2 + N_{RF} + 1)^2 \\ & + 36\tilde{n}K, \quad (\text{by (48a)}) \end{aligned}$$

$$\tilde{\mathcal{C}}_{\text{fact}} = \tilde{n}^3. \quad (\text{by (48b)})$$

Similarly, one can obtain the complexity order of the primal-dual IPM for solving (25b), $C_{(25b)}$, by interchanging K with J , and replacing N_{RF} with N_{FBS} in (31). Thus, the total complexity order of Algorithm 2 is approximately $I_{\text{iter}} \times (C_{(25a)} + C_{(25b)})$, where I_{iter} is the number of outer iterations spent before convergence of Algorithm 2.

On the other hand, the information exchange overhead of Algorithm 2 can be easily seen to be the exchange of local iterates τ_M of MBS and τ_F of FBS in step 5 at each iteration. In LTE-Advanced based cellular systems, this information exchange of $6(K + J)$ real values can be achieved by X2 interface (wired fiber connection).

V. SIMULATION RESULTS

In this section, we present some simulation results to demonstrate the effectiveness of the proposed HyCoBF algorithm, constituted by Algorithm 1 for analog beam selection, and Algorithm 2 for digital CoBF, where problem (25) (for the distributed solution) is solved by using the off-the-shelf convex optimization parser software, CVX [40]. The simulation results for the centralized solution (obtained by solving (17) using CVX) are also provided for justifying the performance

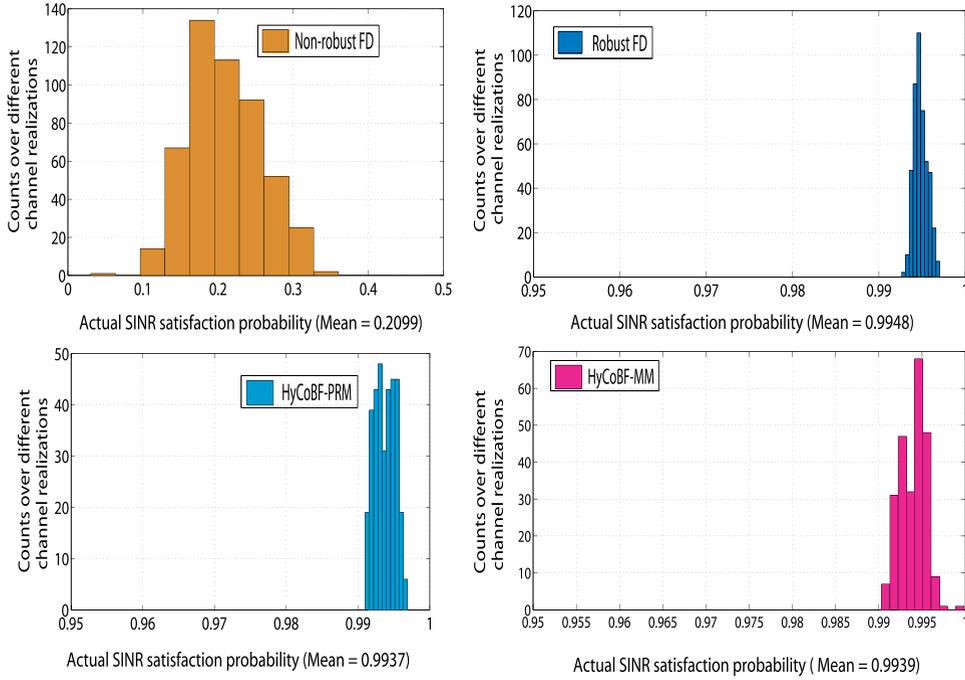


FIGURE 3. Histograms of the actual SINR satisfaction probabilities of various methods at target SINR $\gamma = 9$ dB, for $N_{\text{MBS}} = 16$, $N_{\text{RF}} = 4$, $K = 4$; $N_{\text{FBS}} = 2$, $J = 2$; $\rho = 0.1$, $\varepsilon^2 = 0.002$.

of the proposed HyCoBF design. The simulation settings are as follows. Users' noise powers are identical, i.e., $\sigma_k^2 = \sigma_j^2 = \sigma^2$, $\forall k \in \mathcal{I}_K$, $\forall j \in \mathcal{I}_J$; target SINRs for all MUEs and FUEs are identical, i.e., $\gamma_{Mk} = \gamma_{Fj} = \gamma$, $\forall k \in \mathcal{I}_K$, $\forall j \in \mathcal{I}_J$; SINR outage probabilities are also identically set to $\rho_{Mk} = \rho_{Fj} = \rho = 0.1$ for all k and j , i.e., SINR satisfaction probabilities are higher than 90%. Both MBS and FBS are assumed to be able to track the large-scale fading while the small-scale components for CSI errors are complex Gaussian distributed with zero mean and identical covariance matrices $\mathbf{C}_{MMk} = \mathbf{C}_{MFj} = \varepsilon^2 \mathbf{I}_{N_{\text{MBS}}}$ and $\mathbf{C}_{FFj} = \mathbf{C}_{FMk} = \varepsilon^2 \mathbf{I}_{N_{\text{FBS}}}$, where $\varepsilon^2 > 0$ denotes the variance of each component of channel error vectors. Channel estimates of $\hat{\mathbf{h}}_{MMk}$, $\hat{\mathbf{h}}_{FMk}$, $\hat{\mathbf{h}}_{FFj}$ and $\hat{\mathbf{h}}_{MFj}$ are randomly generated according to the standard circularly symmetric complex Gaussian distribution with variances determined by large scale fading due to path loss and shadowing effect. The simulation results are obtained by averaging over all 500 channel realizations for which all the methods under test yield feasible solutions.

A. PERFORMANCE OF THE PROPOSED HyCoBF DESIGN

We first investigate the feasibility rate of the proposed HyCoBF design. As Algorithm 1 is used for analog beam selection, the obtained design is denoted as HyCoBF-PRM; as magnitude maximization based beam selection is used instead [20], the corresponding design is denoted as HyCoBF-MM. The results of feasibility rate obtained by calculating the percentage proportions (i.e., feasibility rates) of feasible solutions over the generated 500 channel realizations, are illustrated in Fig. 2, where the corresponding

results for the robust FD design (for which $N_{\text{MBS}} = N_{\text{RF}}$, denoted as ‘‘Robust FD’’) are also provided. One can see, from this figure that, the proposed designs for $\varepsilon^2 = 0.001$ (dashed lines) yield much higher feasibility rates than for $\varepsilon^2 = 0.002$ (solid lines). It can be observed that the robust FD outperforms the proposed HyCoBF design for $N_{\text{RF}} = 4$ RF chains equipped at the MBS, simply because the latter has less DoF at the MBS, thereby leading to lower feasibility rates. Moreover, the HyCoBF-PRM performs better than the HyCoBF-MM for SINR from 1 dB to 15 dB, while for SINR higher than 15 dB, their feasibility rates are comparable.

The conservatism in terms of satisfaction probability for the various methods is evaluated by histograms (counted over all feasible channel realizations associated with the results shown in Fig. 2) versus actual SINR satisfaction probabilities. The actual SINR satisfaction probability, defined as the minimum of actual SINR satisfaction probabilities of all the MUEs and FUEs by applying the designed beamformers (i.e., \mathbf{V}^* , \mathbf{w}_k^* at MBS and \mathbf{u}_j^* at FBS) to probability function in (8b) and (8c), was obtained by calculating the relative frequency over 10,000 randomly generated CSI errors (cf. (7)). Fig. 3 shows the obtained histograms for $\gamma = 9$ dB and $\varepsilon^2 = 0.002$. From this figure, one can see that the non-robust FD design (that treats all the given channel estimates as perfect channels) does not achieve the target SINR satisfaction probability due to its actual SINR satisfaction probabilities below 40% (average satisfaction probability is only around 21%) for all the channel realizations. This justifies that the non-robust FD design is quite sensitive to CSI errors. It can be seen that HyCoBF-PRM, HyCoBF-MM and robust FD designs indeed

TABLE 1. For some results in Fig.2, counts of rank-one solutions and all the feasible solutions for each simulation case for $\epsilon^2 \in \{0.01, 0.002\}$ and $\gamma \in \{1, 5, 9, 13\}$ dB are shown here.

ϵ^2	0.01				0.002				
	γ (dB)	1	5	9	13	1	5	9	13
Robust FD	(447, 447)	(396, 396)	(286, 286)	(102, 102)	(490, 490)	(481, 481)	(460, 460)	(401, 401)	
HyCoBF-PRM	(418, 418)	(250, 250)	(63, 63)	(35, 35)	(413, 413)	(402, 402)	(336, 336)	(191, 191)	

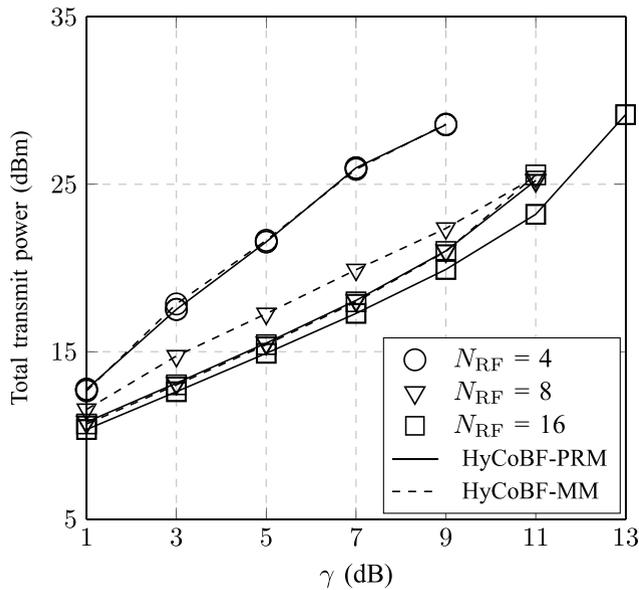


FIGURE 4. Total transmit power of the proposed HyCoBF versus target SINR γ , for $N_{MBS} = 16, N_{RF} = \{4, 8, 16\}, K = 4; N_{FBS} = 2, J = 2; \rho = 0.1, \epsilon^2 = 0.002$.

have higher than 90% SINR satisfaction probability over all feasible channel realizations. In spite of both of HyCoBF-PRM design and HyCoBF-MM design over meet the target satisfaction probability (as shown in Figure 3), the former (for which the PRM beam selection scheme is initialized by the MM beam scheme) also yields higher feasibility rate (as shown in Fig. 2) than the latter, implying that the former is also more power efficient than the latter as shown in Fig. 4.

As discussed in the Section III-B, when the designed \mathbf{W}_k^* and \mathbf{U}_j^* are not of rank one, one may need to find approximate rank-one solutions from them by Gaussian randomization. Therefore, we examine the proportion of rank-one solutions out of those feasible solutions associated with the results shown in Fig. 2. Numerically, \mathbf{W}_k^* and \mathbf{U}_j^* are regarded as rank-one matrices if the following conditions hold:

$$\frac{\lambda_{\max}(\mathbf{W}_k^*)}{\text{Tr}(\mathbf{W}_k^*)} \geq 0.9999, \quad \frac{\lambda_{\max}(\mathbf{U}_j^*)}{\text{Tr}(\mathbf{U}_j^*)} \geq 0.9999, \quad k \in \mathcal{I}_K, \quad j \in \mathcal{I}_J. \quad (32)$$

As shown in Table I, each entry is a pair (p, q) in which p (q) denotes the number of realizations for which the obtained solutions are feasible (feasible and rank one) for each simulation case for $\gamma \in \{1, 5, 9, 13\}$ (dB) and $\epsilon^2 \in \{0.01, 0.02\}$. Due to $p = q$ for all the entries in Table I, all the yielded

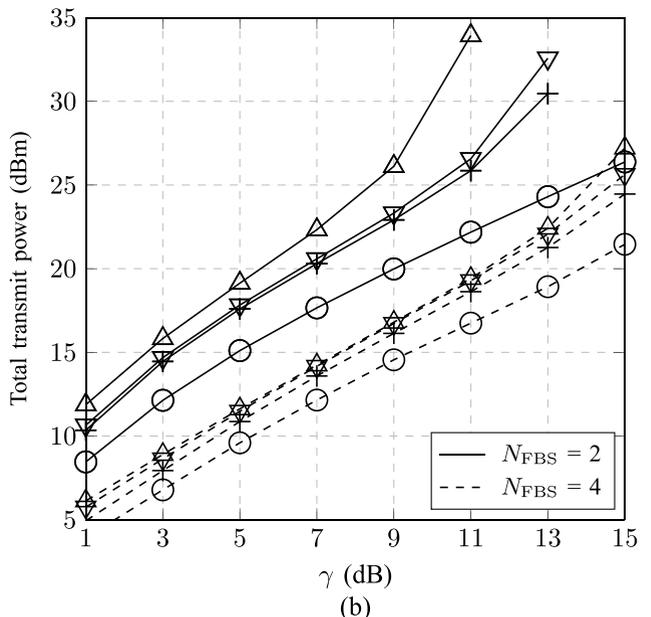
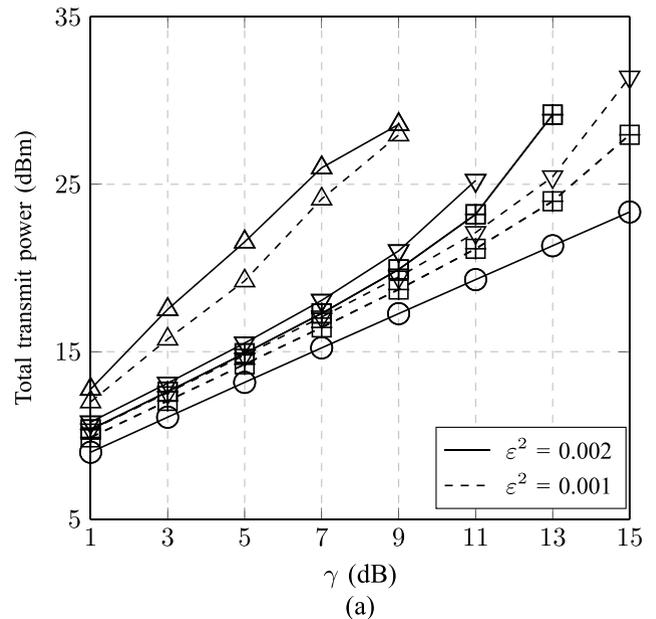


FIGURE 5. Total transmit power of the HyCoBF-PRM (“ Δ ” for $N_{RF} = 4$, “ ∇ ” for $N_{RF} = 8$, and “ $+$ ” for $N_{RF} = 16$); non-robust FD design (denoted as “ \circ ”) and robust FD design (denoted as “ \square ”), for $K = 4$ and $J = 2; \rho = 0.1$, and (a) $N_{MBS} = 16, N_{FBS} = 2, \epsilon^2 \in \{0.001, 0.002\}$, (b) $N_{MBS} = 64, N_{FBS} \in \{2, 4\}, \epsilon^2 = 0.002$.

feasible solutions are of rank one, indicating that the desired HyCoBF strategy is also a single stream transmission scheme, thus suitable for practical deployment.

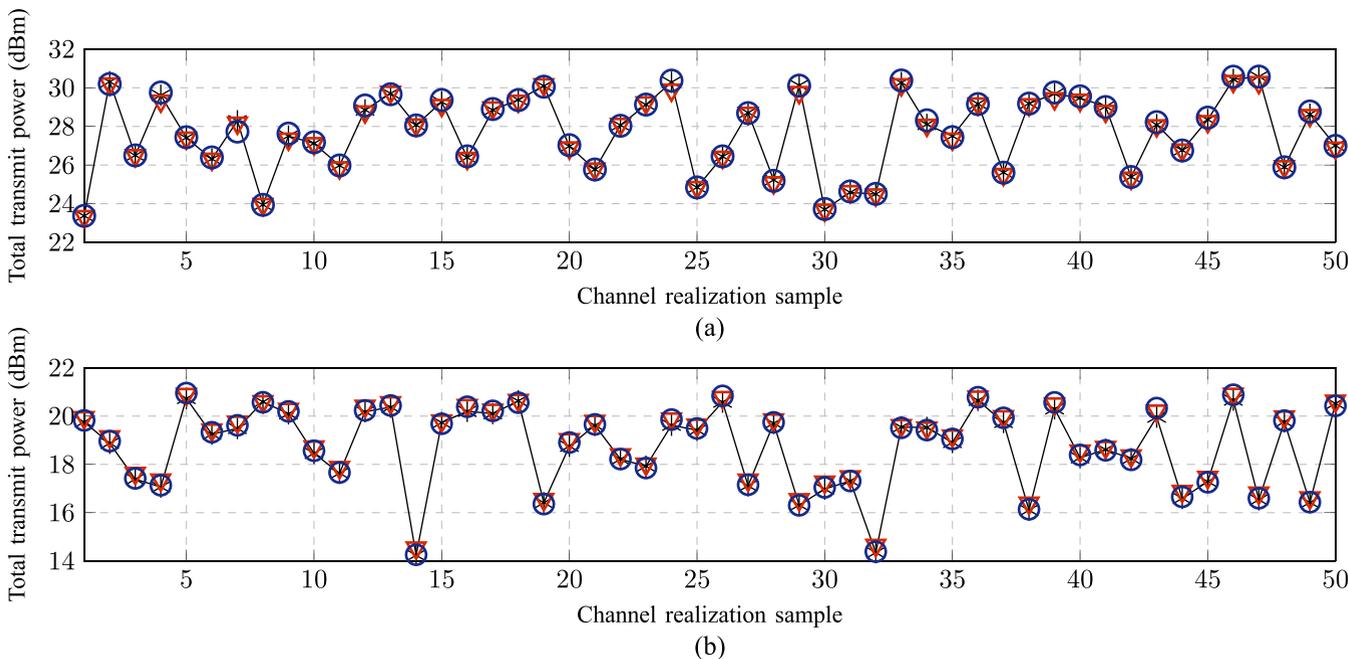


FIGURE 6. Total transmit power of the proposed HyCoBF-PRM (“*” for centralized solutions, and “▽” and “○” for distributed solutions obtained using Algorithm 2 after 15 and 25 iterations, respectively) versus 50 randomly generated channel realizations, for $N_{MBS} = 16$, $N_{RF} = 4$, $K = 4$; $\gamma = 9$ dB, $\rho = 0.1$, $\varepsilon^2 = 0.002$, and (a) $N_{FBS} = J = 2$, (b) $N_{FBS} = J = 4$.

Next, let us examine the transmit power performances of the proposed HyCoBF-PRM by the averaged transmit powers over the channel realizations for which all the designs yield feasible solutions. Fig. 5(a) shows the power performances of the proposed HyCoBF-PRM design, and the corresponding results for the robust FD design as well as the non-robust FD method. It can be observed that the power performances of all the designs under test are better for smaller γ , and that the transmit powers of the non-robust method are the least. As expected, the proposed HyCoBF-PRM design performs better for larger N_{RF} and smaller ε^2 ; and its performance for $N_{RF} = N_{MBS} = 16$ is the same as that of robust FD design. Let us emphasize that the SINR outage probabilities (i.e., (8b) and (8c)) for the non-robust FD method are never satisfied over all feasible channel realizations (cf. Fig. 3). Compared with the robust FD design, the extra power consumption of the proposed HyCoBF-PRM design for $N_{RF} = 8$ is quite small for both $\varepsilon^2 = 0.001$ and $\varepsilon^2 = 0.002$, demonstrating its promising performance for this case.

The corresponding simulation results (not including the robust FD design) for $N_{MBS} = 64$, $N_{RF} \in \{4, 8, 16\}$, $N_{FBS} \in \{2, 4\}$ and $\varepsilon^2 = 0.002$ are shown in Fig. 5(b). It can be seen from this figure that the relative performances among all the designs under test remain the same as shown in Fig. 5(a). Moreover, the total transmit powers for the proposed HyCoBF-PRM for $N_{RF} = 8$ and $N_{RF} = 16$ are comparable, indicating that the MBS has almost achieved the “best” performance for this case of $N_{MBS} = 64$, whereas increasing N_{FBS} (i.e., more spatial DoF) can further reduce the total transmit power by more than 5 dB for this case. These results demonstrate the efficacy of the proposed

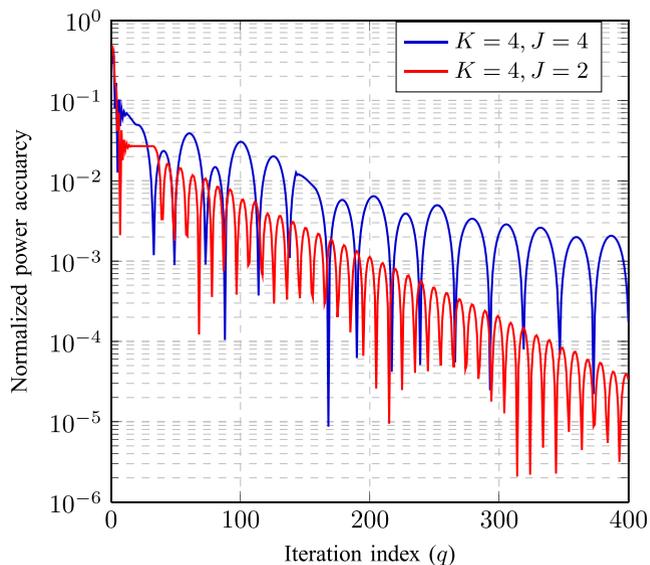


FIGURE 7. Typical convergence curves of Algorithm 2 under $N_{MBS} = 16$, $N_{RF} = 4$, $K = 4$; $N_{FBS} = J \in \{2, 4\}$; $\gamma = 9$ dB, $\rho = 0.1$, $\varepsilon^2 = 0.002$.

HyCoBF-PRM design, and the performance dependence and perspective over different values of N_{MBS} , N_{RF} and N_{FBS} .

B. PERFORMANCE OF THE PROPOSED DISTRIBUTED DIGITAL CoBF DESIGN

To examine the performance of the proposed distributed CoBF design (Algorithm 2), we set the initial input values $\{\delta(0), \mathbf{v}_M(0), \mathbf{v}_F(0), \mu_M(0), \mu_F(0), p_M(0), p_F(0)\}$ all to zero, the over-relaxation parameter $\theta = 1.8$, and for less dependence on the initial values [34], the augmented penalty

parameter $c(q)$ is iteratively updated by rules described in steps 8 and 9 of Algorithm 2 (with $c(0) = 10^{-6}$) [41]. According to Theorem 1, Algorithm 2 will converge to the global optimum of (17) as q increases. From Fig. 6, it can be observed that Algorithm 2 can yield near-optimal solutions within 25 iterations for different simulation settings over different channel realizations.

To further look into the convergence behavior of Algorithm 2, the normalized power accuracy defined as

$$\text{Normalized power accuracy} = \frac{|P^*(q) - P^*|}{P^*}, \quad (33)$$

where P^* is the centralized solution (obtained by solving (17)); $P^*(q) = P_M^*(q) + P_F^*(q)$ is the total power at iteration q , where $P_M^*(q)$ and $P_F^*(q)$ are obtained in step 4 of Algorithm 2. A typical convergence curve for each simulation case shown in Fig. 6 is shown in Fig. 7. As can be seen from this figure that Algorithm 2 yields a solution with the normalized power accuracy smaller than 0.01 within 30 iterations. Moreover, only in a few tens of iterations, Algorithm 2 can yield a solution with the normalized power accuracy less than 10^{-3} . One can also see that higher normalized power accuracy (e.g., 10^{-5}) needs more iterations due to the fluctuating convergence behavior, and that the convergence speed is slower for larger network scale. Nevertheless, these simulation results well demonstrate the convergence of Algorithm 2, as proved in Theorem 1.

VI. CONCLUSION

Within the umbrella of the massive MIMO enabled HetNet, we have presented an outage constrained robust HyCoBF design for the 5G wireless communications, in the presence of Gaussian CSI errors. The proposed HyCoBF design is a cascade of a low-complexity analog beam selection mechanism (Algorithm 1) followed by a robust digital CoBF design. The former obtains the analog beamformer by maximizing the ratio of the total channel power of all the MUEs to the total interference channel power impinging on FUEs. With the designed analog beamformer by Algorithm 1, the later obtains the digital beamformers at MBS and FBS by solving a conservative convex approximation problem, by the use of SDR and an extended form of Bernstein-type inequality. Furthermore, by using ADMM, we have presented a distributed robust digital CoBF algorithm (Algorithm 2). The convergence property of the proposed distributed algorithm, which yields the same solution as the centralized version, has been demonstrated theoretically (Theorem 1) and numerically. Our simulation results have demonstrated that the proposed conservative convex approximation can yield promising performance and, most importantly, it can achieve acceptable performance comparable to the FD beamforming scheme with much smaller number of RF chains. Although the proposed algorithm can solve the SDR based robust CoBF problem within polynomial time, the computational cost of the IPM may still be too expensive as the problem size increases, such as in the scenarios with large number

of antennas and/or dense wireless networks. As a future research, it is worth studying other variations of ADMM for distributed CoBF design for faster convergence.

APPENDIX A PROOF OF LEMMA 1

To prove Lemma 1, we need the following Lemma:

Lemma 2: [38, Lemma 0.1] Let $\mathbf{a} = [a_1, \dots, a_{p+q}]^T \in \mathbb{R}^{p+q}$ and $\mathbf{b} = [b_1, \dots, b_{p+q}]^T \in \mathbb{R}^{p+q}$ be real vectors, $\mathbf{z} = [z_1, \dots, z_{p+q}]^T \in \mathbb{R}^{p+q}$ be a real random vector, where z_1, \dots, z_{p+q} are independent identically distributed (i.i.d.) real random variables following the standard normal distribution $\mathcal{N}(0, 1)$, and define

$$g = \sum_{\ell=1}^{p+q} a_\ell z_\ell^2 + 2b_\ell z_\ell. \quad (34)$$

Then, given $\eta > 0$, the following concentration result holds true:

$$\Pr \left\{ g \leq \sum_{\ell=1}^{p+q} a_\ell - 2\sqrt{\eta} \sqrt{\sum_{\ell=1}^{p+q} a_\ell^2 + 2b_\ell^2} - 2\eta a^- \right\} \leq e^{-\eta}, \quad (35)$$

where $a^- \triangleq \sup\{\sup\{-a_\ell \mid 1 \leq \ell \leq p+q\}, 0\}$.

To employ the real-valued Bernstein-type inequality (35) in proving the complex-valued inequality (15), we define the following real-form counterparts:

$$\bar{\mathbf{Q}}_1 = \frac{1}{2} \begin{bmatrix} \text{Re}\{\mathbf{Q}_{MMk}\} & -\text{Im}\{\mathbf{Q}_{MMk}\} \\ \text{Im}\{\mathbf{Q}_{MMk}\} & \text{Re}\{\mathbf{Q}_{MMk}\} \end{bmatrix} \in \mathbb{S}^p,$$

$$\bar{\mathbf{Q}}_2 = \frac{1}{2} \begin{bmatrix} \text{Re}\{\mathbf{Q}_{FMk}\} & -\text{Im}\{\mathbf{Q}_{FMk}\} \\ \text{Im}\{\mathbf{Q}_{FMk}\} & \text{Re}\{\mathbf{Q}_{FMk}\} \end{bmatrix} \in \mathbb{S}^q,$$

$$\bar{\mathbf{r}}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \text{Re}\{\mathbf{r}_{MMk}\} \\ \text{Im}\{\mathbf{r}_{MMk}\} \end{bmatrix} \in \mathbb{R}^p,$$

$$\bar{\mathbf{r}}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} \text{Re}\{\mathbf{r}_{FMk}\} \\ \text{Im}\{\mathbf{r}_{FMk}\} \end{bmatrix} \in \mathbb{R}^q,$$

$$\bar{\mathbf{z}}_1 = \sqrt{2} \begin{bmatrix} \text{Re}\{\delta_1\} \\ \text{Im}\{\delta_1\} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p),$$

$$\bar{\mathbf{z}}_2 = \sqrt{2} \begin{bmatrix} \text{Re}\{\delta_2\} \\ \text{Im}\{\delta_2\} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_q),$$

where $p \triangleq 2N_{\text{RF}}$ and $q \triangleq 2N_{\text{FBS}}$. Moreover, by the eigenvalue decomposition of the real symmetric matrices $\bar{\mathbf{Q}}_1 = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^T$ and $\bar{\mathbf{Q}}_2 = \mathbf{U}_2 \mathbf{\Lambda}_2 \mathbf{U}_2^T$ (where \mathbf{U}_1 and \mathbf{U}_2 are orthogonal matrices, and $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are diagonal matrices), we can re-express g_1 (cf. (13)) and g_2 (cf. (14)) in the same form as (34):

$$\begin{aligned} g_1(\delta_1, \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}) &= \bar{\mathbf{z}}_1^T \bar{\mathbf{Q}}_1 \bar{\mathbf{z}}_1 + 2\bar{\mathbf{r}}_1^T \bar{\mathbf{z}}_1 \\ &= \bar{\mathbf{z}}_1^T \mathbf{\Lambda}_1 \bar{\mathbf{z}}_1 + 2\bar{\mathbf{r}}_1^T \bar{\mathbf{z}}_1 = \sum_{k=1}^p \mathbf{h}_{1k} \tilde{z}_{1k}^2 + 2\tilde{\mathbf{r}}_{1k} \tilde{z}_{1k}, \end{aligned} \quad (36)$$

$$\begin{aligned} g_2(\delta_2, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) &= \bar{\mathbf{z}}_2^T \bar{\mathbf{Q}}_2 \bar{\mathbf{z}}_2 + 2\bar{\mathbf{r}}_2^T \bar{\mathbf{z}}_2 \\ &= \bar{\mathbf{z}}_2^T \mathbf{\Lambda}_2 \bar{\mathbf{z}}_2 + 2\bar{\mathbf{r}}_2^T \bar{\mathbf{z}}_2 = \sum_{j=1}^q \mathbf{h}_{2j} \tilde{z}_{2j}^2 + 2\tilde{\mathbf{r}}_{2j} \tilde{z}_{2j}, \end{aligned} \quad (37)$$

where h_{1k} is the k th diagonal element of $\mathbf{\Lambda}_1$, h_{2j} is the j th diagonal element of $\mathbf{\Lambda}_2$, $\tilde{\mathbf{z}}_1 = [\tilde{z}_{11}, \dots, \tilde{z}_{1p}]^T \triangleq \mathbf{U}_1^T \bar{\mathbf{z}}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p)$, $\tilde{\mathbf{z}}_2 = [\tilde{z}_{21}, \dots, \tilde{z}_{2q}]^T \triangleq \mathbf{U}_2^T \bar{\mathbf{z}}_2 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_q)$, $\tilde{\mathbf{r}}_1 = [\tilde{r}_{11}, \dots, \tilde{r}_{1p}]^T \triangleq \mathbf{U}_1^T \bar{\mathbf{r}}_1$, and $\tilde{\mathbf{r}}_2 = [\tilde{r}_{21}, \dots, \tilde{r}_{2q}]^T \triangleq \mathbf{U}_2^T \bar{\mathbf{r}}_2$.

By substituting $\mathbf{a} \triangleq [h_{11}, \dots, h_{1p}, h_{21}, \dots, h_{2q}]^T \in \mathbb{R}^{p+q}$, $\mathbf{b} \triangleq [\tilde{\mathbf{r}}_1^T, \tilde{\mathbf{r}}_2^T]^T \in \mathbb{R}^{p+q}$, and $\mathbf{z} = [\tilde{\mathbf{z}}_1^T, \tilde{\mathbf{z}}_2^T]^T$ into (35), we have the following inequality, $\forall \eta > 0$:

$$\begin{aligned} & \Pr \left\{ \sum_{k=1}^p h_{1k} \tilde{z}_{1k}^2 + 2\tilde{r}_{1k} \tilde{z}_{1k} + \sum_{j=1}^q h_{2j} \tilde{z}_{2j}^2 + 2\tilde{r}_{2j} \tilde{z}_{2j} \leq \sum_{k=1}^p h_{1k} \right. \\ & \quad \left. + \sum_{j=1}^q h_{2j} - 2\eta \sup \left\{ \sup_{k=1, \dots, p} \{-h_{1k}\}, \sup_{j=1, \dots, q} \{-h_{2j}\}, 0 \right\} \right. \\ & \quad \left. - 2\sqrt{\eta} \sqrt{\sum_{k=1}^p (h_{1k}^2 + 2\tilde{r}_{1k}^2) + \sum_{j=1}^q (h_{2j}^2 + 2\tilde{r}_{2j}^2)} \right\} \leq e^{-\eta}. \end{aligned} \quad (38)$$

Furthermore, one can easily verify the equalities that $\tilde{\mathbf{z}}_1^T \bar{\mathbf{Q}}_1 \tilde{\mathbf{z}}_1 + 2\tilde{\mathbf{r}}_1^T \tilde{\mathbf{z}}_1 = \delta_1^H \mathbf{Q}_{MMk} \delta_1 + 2\text{Re}\{\mathbf{r}_{MMk}^H \delta_1\}$, $\tilde{\mathbf{z}}_2^T \bar{\mathbf{Q}}_2 \tilde{\mathbf{z}}_2 + 2\tilde{\mathbf{r}}_2^T \tilde{\mathbf{z}}_2 = \delta_2^H \mathbf{Q}_{FMk} \delta_2 + 2\text{Re}\{\mathbf{r}_{FMk}^H \delta_2\}$, $\|\mathbf{r}_{MMk}\|^2 = 2\|\tilde{\mathbf{r}}_1\|^2 = 2\|\tilde{\mathbf{r}}_1\|^2$, $\|\mathbf{r}_{FMk}\|^2 = 2\|\tilde{\mathbf{r}}_2\|^2 = 2\|\tilde{\mathbf{r}}_2\|^2$, $\text{Tr}(\mathbf{Q}_{MMk}) = \sum_{k=1}^p h_{1k} = \text{Tr}(\bar{\mathbf{Q}}_1)$, $\text{Tr}(\mathbf{Q}_{FMk}) = \sum_{j=1}^q h_{2j} = \text{Tr}(\bar{\mathbf{Q}}_2)$, $\|\mathbf{Q}_{MMk}\|_F^2 = 2\sum_{k=1}^p h_{1k}^2 = 2\|\bar{\mathbf{Q}}_1\|_F^2$, $\|\mathbf{Q}_{FMk}\|_F^2 = 2\sum_{j=1}^q h_{2j}^2 = 2\|\bar{\mathbf{Q}}_2\|_F^2$, and $\lambda^+(\mathbf{Q}_{MMk}, \mathbf{Q}_{FMk}) = 2\lambda^+(\bar{\mathbf{Q}}_1, \bar{\mathbf{Q}}_2) = 2\sup\{\sup_{k=1, \dots, p} \{-h_{1k}\}, \sup_{j=1, \dots, q} \{-h_{2j}\}, 0\}$; these, together with (38) and $\eta = \ln(1/\rho_{Mk}) > 0$ (since $0 < \rho_{Mk} \leq 1$), directly imply that the probability inequality (15) holds true. ■

APPENDIX B CONSERVATIVE CONVEX APPROXIMATION PROBLEM (17)

Since Υ (cf. (16)) is monotonically decreasing, its inverse mapping $\Upsilon^{-1}: \mathbb{R} \rightarrow \mathbb{R}_{++}$ is well defined. Then, by applying Lemma 1 and the fact that $e^{-\Upsilon^{-1}(-c_{Mk})} > 0$ (where c_{Mk} is given in (12c)), we have the following inequality:

$$\begin{aligned} & \Pr \left\{ g_1(\delta_1, \mathbf{Q}_{MMk}, \mathbf{r}_{MMk}) + g_2(\delta_2, \mathbf{Q}_{FMk}, \mathbf{r}_{FMk}) \right. \\ & \quad \left. + c_{Mk} \geq 0 \right\} \geq 1 - e^{-\Upsilon^{-1}(-c_{Mk})}, \end{aligned} \quad (39)$$

implying that $e^{-\Upsilon^{-1}(-c_{Mk})} \leq \rho_{Mk}$ is a conservative approximation to (11b), which can be equivalently expressed as

$$\begin{aligned} & \text{Tr}(\mathbf{Q}_{MMk}) + \text{Tr}(\mathbf{Q}_{FMk}) + \ln(\rho_{Mk}) \cdot \lambda^+(\mathbf{Q}_{MMk}, \mathbf{Q}_{FMk}) \\ & \quad - \alpha_{Mk} \sqrt{\|\mathbf{Q}_{MMk}\|_F^2 + 2\|\mathbf{r}_{MMk}\|^2 + \|\mathbf{Q}_{FMk}\|_F^2 + 2\|\mathbf{r}_{FMk}\|^2} \\ & \quad + c_{Mk} \geq 0 \end{aligned} \quad (40)$$

Following a similar reformulation procedure for the single-cell case as reported in [30] and [42], the convex constraint (40) can be expressed in the following form:

$$\begin{aligned} C_M \triangleq & \left\{ (\{\mathbf{W}_k\}, \{\mathbf{U}_j\}, \mathbf{t}_M, \{a_{MMk}\}, \{a_{FMk}\}, P_M) \right\} \\ & a_{MMk} + a_{FMk} + \ln(\rho_{Mk})t_{Mk} - \sigma_{Mk}^2 \end{aligned}$$

$$\begin{aligned} & -\alpha_{Mk} \left\| [t_{MMk}, t_{FMk}]^T \right\| \geq 0, \forall k \in \mathcal{I}_K \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{MMk}) \\ \sqrt{2}\mathbf{r}_{MMk} \end{bmatrix} \right\| \leq t_{MMk}, \forall k \in \mathcal{I}_K \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{FMk}) \\ \sqrt{2}\mathbf{r}_{FMk} \end{bmatrix} \right\| \leq t_{FMk}, \forall k \in \mathcal{I}_K \\ & t_{Mk} \mathbf{I}_{N_{\text{RF}}} + \mathbf{Q}_{MMk} \geq \mathbf{0}, \forall k \in \mathcal{I}_K \\ & t_{Mk} \mathbf{I}_{N_{\text{FBS}}} + \mathbf{Q}_{FMk} \geq \mathbf{0}, t_{Mk} \geq 0, \forall k \in \mathcal{I}_K \\ & a_{MMk} \triangleq \text{Tr}(\mathbf{Q}_{MMk}) + c_{MMk}, \forall k \in \mathcal{I}_K \\ & a_{FMk} \triangleq \text{Tr}(\mathbf{Q}_{FMk}) + c_{FMk}, \forall k \in \mathcal{I}_K \\ & P_M \triangleq \sum_{k \in \mathcal{I}_K} \text{Tr}(\mathbf{V}^* \mathbf{W}_k (\mathbf{V}^*)^H) \end{aligned} \quad (41)$$

where $\mathbf{t}_M \triangleq [t_{MM1}, \dots, t_{MMK}, t_{FM1}, \dots, t_{FMK}, t_{M1}, \dots, t_{MK}]^T \in \mathbb{R}^{3K}$ (auxiliary variables), and c_{MMk} and c_{FMk} are given in (12c). Similarly, we can show that the following inequality is a conservative approximation to (11c):

$$\begin{aligned} & \text{Tr}(\mathbf{Q}_{FFj}) + \text{Tr}(\mathbf{Q}_{MFj}) + \ln(\rho_{Fj}) \cdot \lambda^+(\mathbf{Q}_{FFj}, \mathbf{Q}_{MFj}) \\ & \quad - \beta_{Fj} \sqrt{\|\mathbf{Q}_{FFj}\|_F^2 + 2\|\mathbf{r}_{FFj}\|^2 + \|\mathbf{Q}_{MFj}\|_F^2 + 2\|\mathbf{r}_{MFj}\|^2} \\ & \quad + c_{Fj} \geq 0 \end{aligned} \quad (42)$$

where $\beta_{Fj} \triangleq \sqrt{2 \ln(1/\rho_{Fj})}$, which can be represented as the following form:

$$\begin{aligned} C_F \triangleq & \left\{ (\{\mathbf{W}_k\}, \{\mathbf{U}_j\}, \mathbf{t}_F, \{a_{FFj}\}, \{a_{MFj}\}, P_F) \right\} \\ & a_{FFj} + a_{MFj} + \ln(\rho_{Fj})t_{Fj} - \sigma_{Fj}^2 \\ & - \beta_{Fj} \left\| [t_{FFj}, t_{MFj}]^T \right\| \geq 0, \forall j \in \mathcal{I}_J \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{FFj}) \\ \sqrt{2}\mathbf{r}_{FFj} \end{bmatrix} \right\| \leq t_{FFj}, \\ & \left\| \begin{bmatrix} \text{vec}(\mathbf{Q}_{MFj}) \\ \sqrt{2}\mathbf{r}_{MFj} \end{bmatrix} \right\| \leq t_{MFj}, \forall j \in \mathcal{I}_J \\ & t_{Fj} \mathbf{I}_{N_{\text{FBS}}} + \mathbf{Q}_{FFj} \geq \mathbf{0}, \forall j \in \mathcal{I}_J \\ & t_{Fj} \mathbf{I}_{N_{\text{RF}}} + \mathbf{Q}_{MFj} \geq \mathbf{0}, t_{Fj} \geq 0, \forall j \in \mathcal{I}_J \\ & a_{FFj} \triangleq \text{Tr}(\mathbf{Q}_{FFj}) + c_{FFj}, \forall j \in \mathcal{I}_J \\ & a_{MFj} \triangleq \text{Tr}(\mathbf{Q}_{MFj}) + c_{MFj}, \forall j \in \mathcal{I}_J \\ & P_F \triangleq \sum_{j \in \mathcal{I}_J} \text{Tr}(\mathbf{U}_j) \end{aligned} \quad (43)$$

where c_{FFj} and c_{MFj} are given in (12c), and $\mathbf{t}_F \triangleq [t_{FF1}, \dots, t_{FFJ}, t_{MF1}, \dots, t_{MFJ}, t_{F1}, \dots, t_{FJ}]^T \in \mathbb{R}^{3J}$ collects all the auxiliary variables in the derivation of (43). ■

APPENDIX C PROOF OF THEOREM 1

Let us consider the following structured convex optimization problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m} F(\mathbf{x}) + G(\mathbf{z}) \quad (44a)$$

$$\text{s.t. } \mathbf{x} \in \mathcal{S}_1, \mathbf{z} \in \mathcal{S}_2, \quad (44b)$$

$$\mathbf{A}\mathbf{x} = \mathbf{z}, \quad (44c)$$

where \mathbf{A} is an $m \times n$ matrix, and $F : \mathbb{R}^n \mapsto \mathbb{R}$ and $G : \mathbb{R}^m \mapsto \mathbb{R}$ are convex functions; and $S_1 \subset \mathbb{R}^n$ and $S_2 \subset \mathbb{R}^m$ are nonempty convex sets. Assume that (44) is solvable and strong duality holds. Problem (21) (which is equivalent to problem (17)) can be re-expressed in the same form as (44) by the following correspondences:

$$\mathbf{x} = [\delta^T, p_M, p_F]^T, \mathbf{z} = [\tau_M^T, \tau_F^T, P_M, P_F]^T, \quad (45a)$$

$$F(\mathbf{x}) = 0, G(\mathbf{z}) = P_M + P_F, \quad (45b)$$

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}, \boldsymbol{\Omega} \triangleq [\mathbf{I}_{3(K+J)}, \mathbf{I}_{3(K+J)}]^T, \quad (45c)$$

$$S_1 \triangleq \mathbb{R}^{6(K+J)+2}, \quad (45d)$$

$$S_2 \triangleq \{\mathbf{z} \mid (P_M, \tau_M, \{\mathbf{W}_k\}, \{a_{MMk}\}, \{t_{MMk}\}) \in \tilde{\mathcal{C}}_M, \\ (P_F, \tau_F, \{\mathbf{U}_j\}, \{a_{FFj}\}, \{t_{FFj}\}) \in \tilde{\mathcal{C}}_F\}. \quad (45e)$$

Since $\mathbf{A}^T \mathbf{A} = \mathbf{I}_{6(K+J)+2}$ is invertible, and both $\tilde{\mathcal{C}}_M$ and $\tilde{\mathcal{C}}_F$ are convex sets, the ADMM based distributed algorithm (Algorithm 2) can be guaranteed to converge and the yielded solution is globally optimal to problem (17) by [41, Lemma 2]. ■

APPENDIX D SUMMARY OF WORST-CASE COMPLEXITY FOR A CONIC PROGRAM BY PRIMAL-DUAL IPM

Consider the following conic problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^T \mathbf{x} \quad (46a)$$

$$\text{s.t. } \sum_{i=1}^n x_i \mathbf{A}_i^j + \mathbf{B}^j \in \mathbb{S}_+^{k_j} \text{ for } j = 1, \dots, p, \quad (46b)$$

$$\mathbf{T}^j \mathbf{x} - \mathbf{b}^j \in \mathbb{L}^{k_j} \text{ for } j = p + 1, \dots, m. \quad (46c)$$

where $\mathbf{c} \in \mathbb{R}^n$; $\mathbf{A}_i^j, \mathbf{B}^j \in \mathbb{S}^{k_j}$ for $i = 1, \dots, n$ and $j = 1, \dots, p$; $\mathbf{T}^j \in \mathbb{R}^{k_j \times n}$ and $\mathbf{b}^j \in \mathbb{R}^{k_j}$; \mathbb{L}^{k_j} is the SOC of dimension $k_j \geq 1$; i.e., $\mathbb{L}^{k_j} = \{\mathbf{x} \in \mathbb{R}^{k_j} \mid x_{k_j} \geq \sqrt{x_1^2 + \dots + x_{k_j-1}^2}\}$. Note that linear constraint $\mathbf{a}^T \mathbf{x} - \mathbf{b} \geq 0$ is equivalent to the LMI constraint $\mathbf{a}^T \mathbf{x} - \mathbf{b} \in \mathbb{R}_+ = \mathbb{S}_+^1$ for $k = 1$.

According to [39], the worst-case complexity of a generic IPM for solving (46) consists of two parts:

1) *Iteration Complexity*: For obtaining an ϵ -suboptimal solution of (46), the number of required iterations is on the order of $\sqrt{\varphi(\mathcal{K})} \cdot \ln(1/\epsilon)$, where $\mathcal{K} = \prod_{j=1}^p \mathbb{S}_+^{k_j} \times \prod_{j=p+1}^m \mathbb{L}^{k_j}$ is a cone measuring the geometric complexity of the conic constraints with respect to (46b) and (46c), and

$$\varphi(\mathcal{K}) = \sum_{j=1}^p k_j + 2(m - p). \quad (47)$$

2) *Per-iteration Complexity*: In each iteration of IPM, a search direction is found by solving a system of \tilde{n} linear equations in \tilde{n} unknowns, where \tilde{n} is the total number of primal and dual variables. The computational cost is dominated by (i) the formation of the corresponding $\tilde{n} \times \tilde{n}$ coefficient matrix \mathcal{H} of \tilde{n} linear equations, and (ii) the factorization of the coefficient matrix \mathcal{H} . The cost of forming and that of

factorizing the coefficient matrix \mathcal{H} for (46) are, respectively, on the order of

$$C_{\text{form}} = \underbrace{\tilde{n} \sum_{j=1}^p k_j^3}_{\text{due to (46b)}} + \underbrace{\tilde{n}^2 \sum_{j=1}^p k_j^2}_{\text{due to (46c)}} + \underbrace{\tilde{n} \sum_{j=p+1}^m k_j^2}_{\text{due to (46c)}}, \quad (48a)$$

$$C_{\text{fact}} = \tilde{n}^3. \quad (48b)$$

Therefore, the worst-case complexity order for solving (46) is given by

$$\sqrt{\varphi(\mathcal{K})} \cdot (C_{\text{form}} + C_{\text{fact}}) \cdot \ln(1/\epsilon). \quad (49)$$

ACKNOWLEDGMENT

The authors would like to thank Tung-Chi Yeh, Yao-Rong Syu for their helps during the preparation of this manuscript. ■

REFERENCES

- [1] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, Mar. 2016.
- [2] D. Muirhead, M. A. Imran, and K. Arshad, "A survey of the challenges, opportunities and use of multiple antennas in current and future 5G small cell base stations," *IEEE Access*, vol. 4, pp. 2952–2964, Jul. 2016.
- [3] N. Bhushan et al., "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.
- [4] W. Roh et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.
- [5] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [6] *Study on Elevation Beamforming/Full-Dimension (FD) MIMO for LTE*, document TR 36.897, 3rd Generation Partnership Project (3GPP), 2014. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2580>
- [7] A. F. Molisch et al. (Sep. 2016). "Hybrid beamforming for massive MIMO—A survey." [Online]. Available: <http://arxiv.org/pdf/1609.05078v1.pdf>
- [8] S. Payami, M. Ghorashi, and M. Dianati, "Hybrid beamforming for large antenna arrays with phase shifter selection," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7258–7271, Nov. 2016.
- [9] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems," in *Proc. IEEE ICASSP*, Brisbane, QLD, Australia, Apr. 2015, pp. 2929–2933.
- [10] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [11] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.
- [12] X. Xu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
- [13] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, Jr., "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
- [14] A. Liu and V. K. N. Lau, "Hierarchical interference mitigation for massive MIMO cellular networks," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4786–4797, Sep. 2014.
- [15] A. Liu and V. K. N. Lau, "Impact of CSI knowledge on the codebook-based hybrid beamforming in massive MIMO," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6545–6556, Dec. 2016.

- [16] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.
- [17] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmWave massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1054–1057, May 2016.
- [18] P. V. Pierluigi and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212–2223, Jun. 2015.
- [19] J. Wang and H. Zhu, "Beam allocation and performance evaluation in switched-beam based massive MIMO systems," in *Proc. IEEE ICC*, London, U.K., Jun. 2015, pp. 2387–2392.
- [20] J. Wang, H. Zhu, L. Dai, N. J. Gomes, and J. Wang, "Low-complexity beam allocation for switched-beam based multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8236–8248, Dec. 2016.
- [21] J. Butler and R. Lowe, "Beamforming matrix simplifies design of electrically scanned antennas," *Electron. Design*, vol. 9, pp. 170–173, Apr. 1962.
- [22] S. Chen, S. Sun, Q. Gao, and X. Su, "Adaptive beamforming in TDD-based mobile communication systems: State of the art and 5G research directions," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 81–87, Dec. 2016.
- [23] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
- [24] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM J. Optim.*, vol. 17, no. 4, pp. 969–996, 2006.
- [25] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi, "Coordinated beamforming for multiuser MISO interference channel under rate outage constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1087–1103, Mar. 2013.
- [26] W.-C. Li, T.-H. Chang, and C.-Y. Chi, "Multicell coordinated beamforming with rate outage constraint—Part II: Efficient approximation algorithms," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2763–2778, Jun. 2015.
- [27] L. J. Hong, Y. Yang, and L. Zhang, "Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach," *Oper. Res.*, vol. 59, no. 3, pp. 617–630, May/Jun. 2011.
- [28] A. M.-C. So and Y. J. Zhang, "Distributionally robust slow adaptive OFDMA with soft QoS via linear programming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 947–958, May 2013.
- [29] Y. Shi, J. Zhang, and K. B. Letaief, "Optimal stochastic coordinated beamforming for wireless cooperative networks with CSI uncertainty," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 960–973, Feb. 2015.
- [30] K.-Y. Wang, A. M.-C. 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.
- [31] F. Sahrabi and T. N. Davidson, "Coordinate update algorithms for robust power loading for the MU-MISO downlink with outage constraints," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2761–2773, Jun. 2016.
- [32] M. Medra, Y. Huang, W.-K. Ma, and T. N. Davidson, "Low-complexity robust MISO downlink precoder design under imperfect CSI," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3237–3249, Jun. 2016.
- [33] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*. London, U.K.: Academic, 1983.
- [34] 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.
- [35] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.
- [36] K.-Y. Wang, N. Jacklin, Z. Ding, and C.-Y. Chi, "Robust MISO transmit optimization under outage-based QoS constraints in two-tier heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [37] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [38] I. Bechar. (Apr. 2009). "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables." [Online]. Available: <http://arxiv.org/abs/0909.3595>
- [39] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications* (MOS-SIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2001.
- [40] M. Grant and S. Boyd. (Apr. 2011). *CVX: MATLAB Software for Disciplined Convex Programming, Version 1.21*. [Online]. Available: <http://cvxr.com/cvx>
- [41] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2988–3003, Jun. 2012.
- [42] C.-Y. Chi, W.-C. Li, and C.-H. Lin, *Convex Optimization for Signal Processing and Communications: From Fundamentals to Applications*. Boca Raton, FL, USA: CRC Press, 2017.



GUIXIAN XU received the M.S. degree in signal processing from the Communication University of China, Beijing, China, in 2011. He is currently pursuing the Ph.D. degree in communications engineering with the Beijing University of Posts and Telecommunications, Beijing. Since 2011, he has been a TDLTE Hardware Engineer with Datang Mobile, Beijing. He has been an Intern with CBC/XEV of Ericsson China and the State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology, Beijing. He was a Visiting Ph.D. Student with the National Tsing Hua University, Hsinchu, Taiwan, from 2015 to 2016. His research interests are in massive MIMO, 5G, and convex optimization in wireless communications and signal processing.



CHIA-HSIANG LIN received the B.S. degree in electrical engineering and the Ph.D. degree in communications engineering from the National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2010 and 2016, respectively. He was also a Visiting Scholar with The Chinese University of Hong Kong, Hong Kong, in 2014 and 2017, respectively, and the Virginia Polytechnic Institute and State University, Arlington, VA, USA, from 2015 to 2016. He is currently a Post-Doctoral Researcher with NTHU. His research interests are in network science, game theory, convex geometry and optimization, and blind source separation. He received the 2016 Outstanding Doctoral Dissertation Award from the Chinese Image Processing and Pattern Recognition Society, and the 2016 Best Doctoral Dissertation Award from the IEEE Geoscience and Remote Sensing Society, Taipei Chapter.



WEIGUO MA was born in 1971. He received the B.Sc. degree in electrical engineering and the Ph.D. degree in signal and information system from the Beijing Institute of Technology in 1993 and 1998, respectively. He is currently the Professor of the State Key Laboratory of Wireless Telecommunications. He has 32 granted invention patents mainly in mobile communication area. His current research is focusing on 5G and LTE-V. He chaired several research projects from the National Hi-Tech Research and Development 863 Program of China and from the industry, mainly on TD-SCDMA NodeB and UE chipset IP cores and TD-LTE eNB.



SHANZHI CHEN (SM'04) received the bachelor's degree from Xidian University in 1991 and the Ph.D. degree from the Beijing University of Posts and Telecommunications, China, in 1997. He joined the Datang Telecom Technology & Industry Group in 1994, and has been served as EVP Engineering and CTO since 2008. He is currently the Director of the State Key Laboratory of Wireless Mobile Communications, and the Board Member of Semiconductor Manufacturing International Corporation.

He has contributed to the research and development of TD-LTE 4G. His current research interests include 5G mobile communications, network architectures, vehicular communication networks, and Internet of Things. He was a member of the steering expert group on information technology of the 863 Hi-Tech Research and Development Program of China from 1999 to 2011. His achievements have received many awards and honors, he received 2001, 2012, and 2016 National Awards for Science and Technology Progress, China, the 2015 National Award for Technological Invention, China, and the 2014 Distinguished Young Scholar Award of the National Natural Science Foundation, China.



CHONG-YUNG CHI (S'83–M'83–SM'89) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, in 1983. From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, CA. He has been a Professor with the Department of Electrical Engineering since 1989, and the Institute of Communications Engineering (ICE) since 1999 (also the Chairman of ICE from 2002 to 2005), National Tsing Hua University, Hsinchu, Taiwan.

He has authored over 210 technical papers, including over 80 journal papers (mostly in the *IEEE TRANSACTIONS ON SIGNAL PROCESSING*), over 130 peer-reviewed conference papers, four book chapters, and two books, including a new textbook (432 pages), *Convex Optimization for Signal Processing and Communications*, (CRC Press, 2017) (which has been popularly used in an invited 2-week intensive short course 12 times in eight major universities in China, since 2010, before its publication). His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, biomedical and hyperspectral image analysis. He has been a Technical Program Committee Member for many IEEE sponsored and co-sponsored workshops, symposiums, and conferences on signal processing and wireless communications, including a Co-Organizer and the General Co-Chairman of the 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications, and a Co-Chair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008 and a Lead Co-Chair of SPC Symposium, ChinaCOM 2009. He was a member of Signal Processing Theory and Methods Technical Committee from 2005 to 2010, and the IEEE Signal Processing Society, and a member of Signal Processing for Communications and Networking Technical Committee from 2011 to 2016. He is currently a member of Sensor Array and Multichannel Technical Committee, and the IEEE Signal Processing Society. He was an Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 2001 to 2006 and from 2012 to 2015, the *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II* from 2006 to 2007, the *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I* from 2008 to 2009, and the *IEEE SIGNAL PROCESSING LETTERS* from 2006 to 2010, and a member of Editorial Board of *Signal Processing* from 2005 to 2008, and an Editor from 2003 to 2005 and a Guest Editor in 2006 of the *EURASIP Journal on Applied Signal Processing*.

•••