Distributed Robust Multicell Coordinated Beamforming With Imperfect CSI: An ADMM Approach

Chao Shen, Tsung-Hui Chang, *Member, IEEE*, Kun-Yu Wang, Zhengding Qiu, and Chong-Yung Chi, *Senior Member, IEEE*

***Abstract*—Multicell coordinated beamforming (MCBF), where multiple base stations (BSs) collaborate with each other in the beamforming design for mitigating the intercell interference (ICI), has been a subject drawing great attention recently. Most MCBF designs assume perfect channel state information (CSI) of mobile stations (MSs); however CSI errors are inevitable at the BSs in practice. Assuming elliptically bounded CSI errors, this paper studies the robust MCBF design problem that minimizes the weighted sum power of BSs subject to worst-case signal-to-interference-plus-noise ratio (SINR) constraints on the MSs. Our goal is to devise a distributed optimization method to obtain the worst-case robust beamforming solutions in a decentralized fashion with only local CSI used at each BS and limited backhaul information exchange between BSs. However, the considered problem is difficult to handle even in the centralized form. We first propose an efficient approximation method for solving the nonconvex centralized problem, using semidefinite relaxation (SDR), an approximation technique based on convex optimization. Then a distributed robust MCBF algorithm is further proposed, using a distributed convex optimization technique known as *alternating direction method of multipliers* (ADMM). We analytically show the convergence of the proposed distributed robust MCBF algorithm to the optimal centralized solution. We also extend the worst-case robust beamforming design as well as its decentralized implementation method to a fully coordinated scenario. Simulation results are presented to examine the effectiveness of the proposed SDR method and the distributed robust MCBF algorithm.**

***Index Terms*—Alternating direction method of multipliers (ADMM), convex optimization, coordinated multipoint (CoMP), distributed optimization, multicell processing, robust beamforming, semidefinite relaxation (SDR).**

Manuscript received July 04, 2011; revised November 07, 2011 and February 01, 2012; accepted February 05, 2012. Date of publication February 22, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ignacio Santamaria. The work of C. Shen was supported by the National Basic Research Program of China (973 program) by Grant No. 2007CB307106 and by the Nufront Fellowship. The work of T.-H. Chang, K.-Y. Wang, and C.-Y. Chi was supported by the National Science Council, R.O.C., by Grant NSC-99-2221-E-007-052-MY3. The work of Z. Qiu was supported by the National Basic Research Program of China (973 program) by Grant No. 2007CB307106. A portion of this work was presented at the IEEE ICC, Kyoto, Japan, June 5–9, 2011.

C. Shen and Z. Qiu are with Institute of Information Science, Beijing Jiaotong University, Beijing, 100044, China (e-mail: wwellday@gmail.com; zdqiu@bjtu.edu.cn).

T.-H. Chang was with the National Tsing Hua University. He is now with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: tsunghui.chang@ieee.org).

K.-Y. Wang and C.-Y. Chi are with Institute of Communications Engineering & Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C. (e-mail: kunyuwang7@gmail.com; cychi@ee.nthu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2188719

I. Introduction

RECENTLY, multicell coordinated signal processing has drawn great attention because it can provide significant system throughput gains compared to the conventional single-cell designs [2], [3]. We consider the scenario where the base stations (BSs) are equipped with multiple antennas and the mobile stations (MSs) are equipped with single antenna. The BSs in different cells employ multicell coordinated beamforming (MCBF) [2], aiming at jointly designing the beam patterns of each BS in order to effectively mitigate the intercell interference (ICI). To this end, various MCBF designs have been proposed [4]–[8]. Most of the MCBF designs assume that the BSs are connected with a control center which knows all the MSs' channel state information (CSI) and computes the beamforming solution in a centralized manner. In practical multicell systems, however, obtaining the MCBF solutions in a decentralized fashion using only local CSI at each BS is of central importance, thereby having drawn extensive studies on distributed beamforming [6]–[13]. The reasons are that 1) the future wireless systems prefer a flat Internet Protocol (IP) architecture where all BSs are directly connected with the core network [14]; 2) if the control center is still employed, a distributed optimization method can be used to decouple the original problem into multiple parallel subproblems with smaller problem size, thus reducing the required computation power of the control center [15]. One of the most popular beamforming design criteria is to minimize the transmit power of BSs subject to quality-of-service (QoS) constraints of MSs [4]–[6], [16]. In [6], by leveraging the uplink-downlink duality [4], a distributed optimization method was proposed for the power-minimization-based MCBF design problem. Distributed optimization methods based on primal and dual decomposition techniques [17] were, respectively, reported in [13] and [10] for fully exploiting the inherently block-separable structure of the MCBF problem. Game theory-based distributed optimization methods were also proposed in [9] recently. Though being suboptimal in general, game theoretic approaches demand a much smaller amount of information exchange among BSs. In [12], the idea of uplink-downlink duality was used for distributed optimization of a different max-min-fair MCBF design problem [18].

The efficacy of beamforming designs relies on the assumption that the BSs have the perfect CSI of MSs. In practical scenarios, however, the BSs can never have perfect CSI, due to, e.g., imperfect channel estimation and finite rate feedback [19]. In the multicell scenario, it is even more difficult for the BSs to obtain reliable intercell CSI (i.e., the CSI of MSs in the neighboring cells). In the presence of CSI errors, the MSs' QoS requirements can no longer be guaranteed. In view of this, robust MCBF designs, which take into account the CSI errors, are of great importance. Robust beamforming designs have been studied for the single-cell scenarios; e.g., see [20]–[24]. Depending on how the CSI errors are modeled, there are two major design criteria, namely, the chance constrained robust design [22]–[24] and the worst-case constrained robust design [20], [21]. The former assumes that the CSI errors are random following certain statistical distribution (usually Gaussian), and the robustness is achieved in a probabilistic sense; while the latter assumes that the CSI errors lie in a bounded uncertainty region, e.g., quantization errors, and MSs' QoS is guaranteed for all possible errors within the region. These existing works provide useful insights and mathematical tools for the development of the robust MCBF design [25].

The focus of this paper is on the worst-case signal-to-interference-plus-noise ratio (SINR) constrained MCBF design problem [1], [26], where the weighted sum power of BSs is minimized subject to constraints that guarantee worst-case SINR requirements for the MSs. Our goal is to develop a distributed beamforming optimization algorithm for the worst-case robust formulation; however, the considered problem itself is difficult to handle even in the centralized form, due to the fact that each of the worst-case SINR constraints corresponds to infinitely many nonconvex constraints. In contrast with the single-cell robust beamforming design in [20] and [21], the robust MCBF formulation considered in this paper is more challenging because the CSI errors not only appear in the desired signal and intracell interference terms, but also in the ICI term. To handle this problem, a convex restrictive approximation formulation was proposed in [26]. Distributed optimization algorithms based on dual decomposition and alternating optimization were also presented in [26]. However, due to the reduced feasible set, the approximation method in [26] is less power efficient than the original problem.

In this paper, we propose a new convex approximation method for the worst-case SINR constrained robust MCBF design problem. Our approach is based on a convex approximation technique known as semidefinite relaxation (SDR) [27]. SDR has been a popular approach to dealing with various transmit beamforming designs [28], [29]. By SDR, we come up with an approximation formulation, which is a convex semidefinite program (SDP) and can be efficiently solved by interior-point methods [30]. We also identify several conditions under which the proposed SDR method can yield the global optimal solution.

We further develop a distributed optimization algorithm for solving the proposed SDR approximation formulation. While, as shown in [10], the dual decomposition method [17] can successfully handle the nonrobust MCBF design, direct application of the dual decomposition method to the proposed SDR-based robust MCBF formulation may result in unbounded subproblems due to lack of strict convexity, and the resulting algorithm is thus numerically unstable. To overcome this problem, we instead consider the so-called *alternating direction method of multipliers* (ADMM) [31], [32]. ADMM is an advanced dual decomposition method that combines the idea of dual decomposition and the augmented Lagrangian method [33], where the latter is often used for bringing numerical robustness to the dual ascent method [32] by adding strictly convex penalty terms. Hence, ADMM is in general more numerically stable and faster in convergence than the conventional dual decomposition method [32], [34]. Based on the principle of ADMM, we propose a distributed robust MCBF algorithm that is provably able to converge to the global optimum of the centralized problem.

In addition to the multicell coordinated scenario, we further consider a fully coordinated scenario where the BSs collaborate to serve the MSs near the cell edges, and extend the proposed worst-case robust design, as well as the distributed implementation method to this scenario. This fully coordinated scenario can enhance the QoS of the cell-edge MSs, but the BSs have to share with each other the data and CSI of the cell-edge MSs and synchronize their transmission phases, which is more complex and requires more backhaul signaling [2], [8].

The rest of this paper is organized as follows. Section II presents the multicell signal model and the worst-case SINR constrained robust MCBF design problem. In Section III, the proposed SDR approximation method and its optimality conditions are presented. Using ADMM, a distributed robust MCBF algorithm is proposed in Section IV. Section V extends the proposed method to a fully coordinated scenario. Simulation results are presented in Section VI. Finally, conclusions are drawn in Section VII.

Notations: $\mathbb{C}^n$, $\mathbb{R}^n$ and $\mathbb{H}^n$ stand for the sets of n-dimensional complex and real vectors and complex Hermitian matrices, respectively. $\mathbb{R}^n_+$ denotes the set of n-dimensional nonnegative orthant. Column vectors and matrices are written in boldfaced lowercase and uppercase letters, e.g., $\mathbf{a}$ and $\mathbf{A}$. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, and $\mathbf{0}$ denotes an all-zero vector (matrix) with appropriate dimension. The superscripts $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ represent the transpose, (Hermitian) conjugate transpose and pseudoinverse operations, respectively. $\mathrm{Rank}(\mathbf{A})$ and $\mathrm{Tr}(\mathbf{A})$ represent the rank and trace of matrix $\mathbf{A}$, respectively. $\mathbf{A} \succeq \mathbf{0}$ ($\succ \mathbf{0}$) means that matrix $\mathbf{A}$ is positive semidefinite (positive definite). For vector $\mathbf{a}$, $\|\mathbf{a}\|$ denotes the Euclidean norm. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. For a variable a_{nmk}, where $n \in \{1,\ldots,N\}$, $m \in \{1,\ldots,M\}$ and $k \in \{1,\ldots,K\}$, $\{a_{nmk}\}_k$ denotes the set containing $a_{nm1},\ldots,a_{nmK}$; while $\{a_{nmk}\}$ denotes the set containing all possible a_{nmk}, i.e., $a_{111},\ldots,a_{11K},a_{121},\ldots,a_{NMK}$.

II. Signal Model and Problem Statement

A. Signal Model

Consider a multicell downlink system that consists of N_c cells. Each cell is composed of one BS, which is equipped with N_t antennas, and K single-antenna MSs. The N_c BSs are assumed to operate over a common frequency band and each communicates with its K associated MSs using transmit beamforming. The scenario under consideration is that each MS is served by only one BS; extension to the scenario where one MS is served by multiple BSs will be given in Section V.

We use BS_n to denote the nth BS, and MS_{nk} the kth MS in the nth cell, for all $n \in \mathcal{N}_c \triangleq \{1,2,\ldots,N_c\}$ and $k \in \mathcal{K} \triangleq \{1,2,\ldots,K\}$. Let $s_{nk}(t) \in \mathbb{C}$ be the signal of interest for MS_{nk}, and $\mathbf{w}_{nk} \in \mathbb{C}^{N_t}$ be the associated beamforming vector. The transmit signal by BS_n is given by

$$\mathbf{x}_n(t) = \sum_{k=1}^{K} \mathbf{w}_{nk} s_{nk}(t) \tag{1}$$

for $n \in \mathcal{N}_c$. The received signal of MS_{nk} can be expressed as

$$\begin{aligned} y_{nk}(t) &= \sum_{m=1}^{N_c} \mathbf{h}_{mnk}^H \mathbf{x}_m(t) + z_{nk}(t) \\ &= \mathbf{h}_{nnk}^H \mathbf{w}_{nk} s_{nk}(t) + \sum_{i\neq k}^{K} \mathbf{h}_{nnk}^H \mathbf{w}_{ni} s_{ni}(t) \\ &\quad + \sum_{m\neq n}^{N_c} \sum_{i=1}^{K} \mathbf{h}_{mnk}^H \mathbf{w}_{mi} s_{mi}(t) + z_{nk}(t) \end{aligned} \tag{2}$$

where $\mathbf{h}_{mnk} \in \mathbb{C}^{N_t}$ denotes the channel vector from BS_m to MS_{nk}, and $z_{nk}(t) \in \mathbb{C}$ is the additive noise of MS_{nk}, which is assumed to have zero mean and variance $\sigma_{nk}^2 > 0$. The term $z_{nk}(t)$ may capture both the receiver noise and the interference from the noncoordinated BSs. In (2), the first term is the signal of interest for MS_{nk}, and the second and third terms are the intracell interference and ICI, respectively. Assume that $s_{nk}(t)$ are statistically independent, with zero mean and $\mathbb{E}\{|s_{nk}(t)|^2\} = 1$ for all $n \in \mathcal{N}_c$ and $k \in \mathcal{K}$, and that each MS employs single-user detection. By (2), the SINR of MS_{nk} is given by

$$\begin{aligned} &\mathrm{SINR}_{nk}\left(\{\mathbf{w}_{m1},\ldots,\mathbf{w}_{mK}\}_{m=1}^{N_c}, \{\mathbf{h}_{mnk}\}_{m=1}^{N_c}\right) \\ &= \frac{\left|\mathbf{h}_{nnk}^H \mathbf{w}_{nk}\right|^2}{\sum\limits_{i\neq k}^{K} \left|\mathbf{h}_{nnk}^H \mathbf{w}_{ni}\right|^2 + \sum\limits_{m\neq n}^{N_c} \sum\limits_{i=1}^{K} \left|\mathbf{h}_{mnk}^H \mathbf{w}_{mi}\right|^2 + \sigma_{nk}^2}. \end{aligned} \tag{3}$$

Single-cell beamforming designs [16], [29], [35] are developed mainly for handling the intracell interference. To effectively mitigate the ICI, the following MCBF design has been considered [36]

$$\min_{\substack{\mathbf{w}_{nk},k=1,\ldots,K\\ n=1,\ldots,N_c}} \sum_{n=1}^{N_c} \alpha_n \left(\sum_{k=1}^{K} \|\mathbf{w}_{nk}\|^2\right) \tag{4a}$$

$$\text{s.t. } \mathrm{SINR}_{nk}\left(\{\mathbf{w}_{m1},\ldots,\mathbf{w}_{mK}\}_{m=1}^{N_c}, \{\mathbf{h}_{mnk}\}_{m=1}^{N_c}\right) \geq \gamma_{nk} \quad \forall k \in \mathcal{K}, n \in \mathcal{N}_c \tag{4b}$$

where $\alpha_n > 0$ is the power priority weight for BS_n. As seen, the MCBF design jointly optimizes the beamforming vectors of all the BSs such that the weighted sum power is minimized, and meanwhile the MSs' SINR requirements $\gamma_{nk} > 0$ must be fulfilled. It has been shown that (4) can be reformulated as a convex second-order cone program (SOCP) [36], which can be efficiently solved via standard convex solvers, e.g., `SeDuMi` [37].

B. Worst-Case Robust MCBF Design

The MCBF design in (4) assumes that the BSs have perfect knowledge of the CSIs. As discussed in the introduction section, in practice, the BSs inevitably suffer from CSI errors. Let $\hat{\mathbf{h}}_{mnk} \in \mathbb{C}^{N_t}$, $n,m \in \mathcal{N}_c$, $k \in \mathcal{K}$ be the preassumed CSI at the BSs. Then the true CSI can be expressed as

$$\mathbf{h}_{mnk} = \hat{\mathbf{h}}_{mnk} + \mathbf{e}_{mnk} \ \forall m,n \in \mathcal{N}_c,\ k \in \mathcal{K} \tag{5}$$

where $\mathbf{e}_{mnk} \in \mathbb{C}^{N_t}$ denotes the CSI error vector satisfying the following elliptic model:

$$\mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1 \tag{6}$$

where $\mathbf{Q}_{mnk} \in \mathbb{H}^{N_t}$, $\mathbf{Q}_{mnk} \succ \mathbf{0}$ specifies the size and shape of the ellipsoid. When $\mathbf{Q}_{mnk} = \epsilon_{mnk}^{-2} \mathbf{I}_{N_t}$ where $\varepsilon_{mnk}^2 > 0$, (6) reduces to the popular spherical error model $\|\mathbf{e}_{mnk}\|^2 \leq \varepsilon_{mnk}^2$ [20]. The following simulation example motivates the need of robust designs for accounting for the CSI errors.

Example: Let us consider a two-cell system ($N_c = 2$), with two MSs in each cell ($K = 2$). Each of the BSs is equipped with four antennas ($N_t = 4$). A set of preassumed CSI $\{\hat{\mathbf{h}}_{mnk}\}$ is randomly generated following the independent and identically distributed (i.i.d.) complex Gaussian distribution with zero mean and unit variance (see Section VI-A for the detailed channel model used in the simulation). Using the preassumed CSI, an optimal beamforming solution is obtained by solving the MCBF problem (4) for $\gamma_{nk} = 20$ dB for $n,k = 1, 2$. To examine the impact of the CSI errors, we generated 10 000 sets of CSI errors satisfying $\|\mathbf{e}_{nnk}\|^2 \leq 0.01$, $\|\mathbf{e}_{mnk}\|^2 \leq 0.04$ for all $m \neq n$, $m,n \in \mathcal{N}_c$, $k \in \mathcal{K}$, and evaluated the achievable SINR values in (3). Fig. 1 displays the probability distribution of the achievable SINR values of MS_{11}. As seen, it is with very high chance that the achieved SINR is smaller than the target value, due to the CSI errors. The worst SINR value can be even less than 5 dB. ■

To guarantee that the SINR requirement γ_{nk} can be satisfied for all possible CSI errors, we consider the following worst-case robust MCBF design [1], [26]:

$$\min_{\{\mathbf{w}_{nk}\}_{n,k}} \sum_{n=1}^{N_c} \alpha_n \left(\sum_{k=1}^{K} \|\mathbf{w}_{nk}\|^2\right) \tag{7a}$$

$$\begin{aligned} \text{s.t. } &\mathrm{SINR}_{nk}\Big(\{\mathbf{w}_{m1},\ldots,\mathbf{w}_{mK}\}_{m=1}^{N_c}, \\ &\qquad \{\hat{\mathbf{h}}_{mnk} + \mathbf{e}_{mnk}\}_{m=1}^{N_c}\Big) \geq \gamma_{nk}, \\ &\forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1, m,n \in \mathcal{N}_c, k \in \mathcal{K}. \end{aligned} \tag{7b}$$

In comparison to the nonrobust design in (4), the above worst-case robust MCBF design can provide guaranteed QoS for the MSs, as illustrated in Fig. 1. Solving the robust design problem (7), however, is more challenging because, firstly, each of the SINR constraints is nonconvex, and second, there are infinitely many such nonconvex SINR constraints (due to the worst-case design criterion). In the next section, we propose to apply the

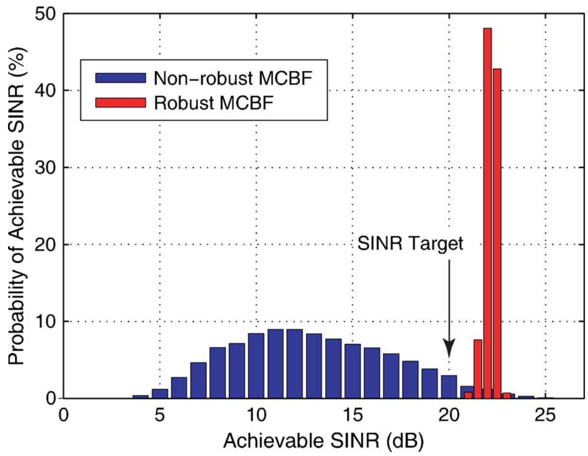


Fig. 1. Distribution of the achievable SINR values of MS_{11}, when the non-robust MCBF design (4) and the robust MCBF design (7) are, respectively, used. The target SINR value is 20 dB.

convex optimization-based SDR technique [27]. We will further present conditions under which SDR is optimal.

III. Proposed SDR–Based Method

A. Solving (7) by SDR and S-Lemma

Considering that each of the SINR constraints is non-convex, we first apply SDR to "linearize" the robust MCBF problem (7). Let us express the objective function of (7) as $\sum_{n=1}^{N_c}\alpha_n\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{w}_{nk}\mathbf{w}_{nk}^H)$, and express each of the worst-case SINR constraints in (7b) as

$$\left(\hat{\mathbf{h}}_{nnk}^H+\mathbf{e}_{nnk}^H\right)\left(\frac{1}{\gamma_{nk}}\mathbf{w}_{nk}\mathbf{w}_{nk}^H-\sum_{i\neq k}^{K}\mathbf{w}_{ni}\mathbf{w}_{ni}^H\right)(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})$$
$$\geq\sum_{m\neq n}^{N_c}\left(\hat{\mathbf{h}}_{mnk}^H+\mathbf{e}_{mnk}^H\right)\left(\sum_{i=1}^{K}\mathbf{w}_{mi}\mathbf{w}_{mi}^H\right)(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})$$
$$+\sigma_{nk}^2,\forall\mathbf{e}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{e}_{mnk}\leq 1, m\in\mathcal{N}_c. \tag{8}$$

The idea of SDR is to replace each rank-one matrix $\mathbf{w}_{nk}\mathbf{w}_{nk}^H$ by a general-rank positive semidefinite matrix $\mathbf{W}_{nk}\succeq\mathbf{0}$ [27]. After applying SDR to (7), we obtain the following:

$$\min_{\substack{\mathbf{W}_{nk}\succeq\mathbf{0},k=1,\ldots,K\\ n=1,\ldots,N_c}}\ \sum_{n=1}^{N_c}\alpha_n\left(\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})\right) \tag{9a}$$

$$\text{s.t. } \left(\hat{\mathbf{h}}_{nnk}^H+\mathbf{e}_{nnk}^H\right)\left(\frac{1}{\gamma_{nk}}\mathbf{W}_{nk}-\sum_{i\neq k}^{K}\mathbf{W}_{ni}\right)\left(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk}\right)$$
$$\geq\sum_{m\neq n}^{N_c}(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})^H\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)\left(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk}\right)$$
$$+\sigma_{nk}^2,\forall\mathbf{e}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{e}_{mnk}\leq 1, m\in\mathcal{N}_c, n\in\mathcal{N}_c, k\in\mathcal{K}. \tag{9b}$$

Note that the SDR problem (9) is convex, since the objective function and constraints are linear in $\mathbf{W}_{nk}$. However, (9) is still computationally intractable because it involves an infinite number of constraints. Fortunately, the infinitely many constraints can be recast as a finite number of convex constraints.

To show this, we first observe that the left- and right-hand side (LHS) (RHS) of the first inequality in (9b) involve independent CSI errors. Hence, (9b) for MS_{nk} can be alternatively expressed as follows:

$$\min_{\mathbf{e}_{nnk}^H\mathbf{Q}_{nnk}\mathbf{e}_{nnk}\leq 1}\left(\hat{\mathbf{h}}_{nnk}^H+\mathbf{e}_{nnk}^H\right)\left(\frac{1}{\gamma_{nk}}\mathbf{W}_{nk}-\sum_{i\neq k}^{K}\mathbf{W}_{ni}\right)$$
$$\times(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})$$
$$\geq\sum_{m\neq n}^{N_c}\left\{\max_{\mathbf{e}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{e}_{mnk}\leq 1}(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})^H\right.$$
$$\left.\times\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})\right\}+\sigma_{nk}^2. \tag{10}$$

By introducing the slack variable

$$t_{mnk}=\max_{\mathbf{e}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{e}_{mnk}\leq 1}(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})^H\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk}) \tag{11}$$

which is the worst-case ICI power from BS_m to MS_{nk}, for all $m\in\mathcal{N}_c\setminus\{n\}$, (10) can be written as

$$\min_{\mathbf{e}_{nnk}^H\mathbf{Q}_{nnk}\mathbf{e}_{nnk}\leq 1}\left\{\left(\hat{\mathbf{h}}_{nnk}^H+\mathbf{e}_{nnk}^H\right)\left(\frac{1}{\gamma_{nk}}\mathbf{W}_{nk}-\sum_{i\neq k}^{K}\mathbf{W}_{ni}\right)\right.$$
$$\left.\times(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})\right\}\geq\sum_{m\neq n}^{N_c}t_{mnk}+\sigma_{nk}^2. \tag{12}$$

By (11) and (12), the worst-case SINR constraint for MS_{nk} in (9b) can be decoupled into the following N_c worst-case constraints:

$$\left(\hat{\mathbf{h}}_{nnk}^H+\mathbf{e}_{nnk}^H\right)\left(\frac{1}{\gamma_{nk}}\mathbf{W}_{nk}-\sum_{i\neq k}^{K}\mathbf{W}_{ni}\right)(\hat{\mathbf{h}}_{nnk}+\mathbf{e}_{nnk})$$
$$\geq\sum_{m\neq n}^{N_c}t_{mnk}+\sigma_{nk}^2\ \forall\mathbf{e}_{nnk}^H\mathbf{Q}_{nnk}\mathbf{e}_{nnk}\leq 1 \tag{13}$$

$$\left(\hat{\mathbf{h}}_{mnk}^H+\mathbf{e}_{mnk}^H\right)\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)(\hat{\mathbf{h}}_{mnk}+\mathbf{e}_{mnk})\leq t_{mnk}$$
$$\forall\mathbf{e}_{mnk}^H\mathbf{Q}_{mnk}\mathbf{e}_{mnk}\leq 1, m\in\mathcal{N}_c\setminus\{n\}. \tag{14}$$

Since each of the constraints in (13) and (14) entails only one CSI error, we can use the same approach as in [21] and [38], to reformulate (13) and (14) into finite convex constraints. The ingredient is the S-lemma.

Lemma 1 (S-Lemma [30]): Let $\phi_i(\mathbf{x})\triangleq\mathbf{x}^H\mathbf{A}_i\mathbf{x}+\mathbf{b}_i^H\mathbf{x}+\mathbf{x}^H\mathbf{b}_i+c_i$, for $i=0,1$, where $\mathbf{A}_i\in\mathbb{H}^{N_t}$, $\mathbf{b}_i\in\mathbb{C}^{N_t}$ and $c_i\in\mathbb{R}$. Suppose that there exists an $\bar{\mathbf{x}}\in\mathbb{C}^{N_t}$ such that $\phi_1(\bar{\mathbf{x}})<0$. Then the two conditions are equivalent:

1) $\phi_0(\mathbf{x})\geq 0$ for all $\mathbf{x}$ satisfying $\phi_1(\mathbf{x})\leq 0$;

2) There exists a $\lambda \geq 0$ such that

$$\begin{bmatrix} \mathbf{A}_0 & \mathbf{b}_0 \\ \mathbf{b}_0^H & c_0 \end{bmatrix} + \lambda \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & c_1 \end{bmatrix} \succeq \mathbf{0}.$$

By applying S-lemma, we can recast (13) and (14) as the following linear matrix inequalities (LMIs) (15)–(16), shown at the bottom of the page, where $\mathbf{U}_{nk} \triangleq \frac{1}{\gamma_{nk}}\mathbf{W}_{nk} - \sum_{i\neq k}^{K}\mathbf{W}_{ni}$, and $\lambda_{mnk} \geq 0$ are slack variables. Consequently, one can reformulate (9) as the following:

$$\min_{\{\mathbf{W}_{nk}\},\{\lambda_{mnk}\},\{t_{mnk}\}} \sum_{n=1}^{N_c} \alpha_n \left(\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_{nk})\right) \tag{17a}$$

$$\text{s.t.} \quad \mathbf{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^{K}, \{t_{mnk}\}_m, \lambda_{nnk}\right) \succeq \mathbf{0}, \tag{17b}$$

$$\mathbf{\Psi}_{mnk}\left(\{\mathbf{W}_{mi}\}_{i=1}^{K}, t_{mnk}, \lambda_{mnk}\right) \succeq \mathbf{0} \; \forall m \neq n, \tag{17c}$$

$$\mathbf{W}_{nk} \succeq \mathbf{0}, \lambda_{mnk} \geq 0 \; \forall m, \; \forall n, k. \tag{17d}$$

Problem (17) is a convex SDP which can be efficiently solved by off-the-shelf convex solvers [37].

B. Optimality Conditions

An important aspect of SDR is whether the optimal solution $\{\mathbf{W}_{nk}^\star\}$ satisfies $\mathbf{W}_{nk}^\star = \mathbf{w}_{nk}^\star(\mathbf{w}_{nk}^\star)^H$ for some $\mathbf{w}_{nk}^\star \in \mathbb{C}^{N_t}$, for all n, k. If this is true, then $\{\mathbf{w}_{nk}^\star\}$ is an optimal solution of the original robust MCBF problem (7). It, therefore, is important to investigate the conditions under which the SDR problem (17) can yield rank-one $\{\mathbf{W}_{nk}^\star\}$. Some provable conditions are given in the following proposition.

Proposition 1: *Suppose that the SDR problem (17) is feasible. Consider the following three conditions:*
C1) $K = 1$, i.e., there is only one MS in each cell;
C2) $\mathbf{Q}_{nnk} = \infty \mathbf{I}_{N_t}$ for all n, k, i.e., $\mathbf{e}_{nnk} = \mathbf{0}$ for all n, k, i.e., perfect intracell CSI $\{\mathbf{h}_{nnk}\}$;
C3) For the spherical model, i.e., $\|\mathbf{e}_{mnk}\|^2 \leq \varepsilon_{mnk}^2$ for all m, n, k, the CSI error bounds $\{\varepsilon_{mnk}\}$ satisfy

$$\varepsilon_{mnk} \leq \bar{\varepsilon}_{mnk} \quad \text{and} \quad \varepsilon_{nnk} < \sqrt{\frac{\sigma_{nk}^2 \alpha_n \gamma_{nk}}{f^\star}} \tag{18}$$

for all m, n, k, where $\{\bar{\varepsilon}_{mnk}\}$ are some CSI error bounds under which (17) is feasible, with $f^\star > 0$ denoting the associated optimal objective value.
If any one of the above three conditions is satisfied, then the SDR problem (17) must yield a rank-one solution, that is, $\{\mathbf{W}_{nk}^\star\}$ must satisfy $\mathbf{W}_{nk}^\star = \mathbf{w}_{nk}^\star(\mathbf{w}_{nk}^\star)^H$ for all n, k.

Proof: The proofs of C1) and C2) are presented in Appendix A. Case C3) is a generalization of the result in [38] where the tightness of SDR for the worst-case robust beamforming problem in the single-cell scenario ($N_c = 1$) is studied. Case C3) can be proved following exactly the same idea as in [38] and thus the details are omitted here. ■

Case C1) of Proposition 1 shows that the global optimum of the robust MCBF problem (7) can be attained by solving the SDR problem (17) whenever $K = 1$. While this result is interesting from the theoretical point of view, we should mention that, when $K = 1$, (7) can also be optimally solved by employing an SOCP reformulation idea presented in [26], [39]. It is known that solving an SOCP is in general computationally cheaper than solving an SDP [e.g., (17)]. For the general case of $K > 1$, Case C2) shows that if the BSs have channel uncertainty only for intercell CSI, i.e., $\{\mathbf{h}_{mnk}\}$ where $m \neq n$, and perfectly know the intracell CSI, i.e., $\{\mathbf{h}_{nnk}\}$, then rank-one solutions are guaranteed. If errors occur in both intracell and intercell CSI, Case C3) implies that rank-one solutions can also be obtained as long as that the CSI errors are sufficiently small.

For a general setup, it is not known yet whether the SDR problem (17) has a rank-one solution. If the obtained solution is not of rank one, then additional solution approximation procedure, such as the Gaussian randomization method [28], can be employed to obtain a rank-one approximate solution to (7). Quite surprisingly, we found in our simulation tests that (17) with spherical CSI errors always yields rank-one $\{\mathbf{W}_{nk}^\star\}$. Hence for these problem instances, we can simply perform rank-one decomposition of $\mathbf{W}_{nk}^\star = \mathbf{w}_{nk}^\star(\mathbf{w}_{nk}^\star)^H$. Investigating the reasons behind would be an interesting future research; see [38] and [40] for recent endeavors.

IV. Distributed Robust MCBF Algorithm Using ADMM

Solving the SDR problem (17) calls for a control center which requires all the CSI of MSs. As discussed in the Introduction, it is desirable to obtain the beamforming solutions in a decentralized fashion using local CSI at each BS, i.e., BS_n uses $\{\hat{\mathbf{h}}_{nmk}\}_{m,k}$ only for all $n \in \mathcal{N}_c$. A simple approach would be applying the dual decomposition method [17], similar to the approach in [10] and [26]. However, as will be explained later, the dual decomposition method is not suitable for (17) due to the fact that the decomposed problems lack strict convexity

$$\mathbf{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^{K}, \{t_{mnk}\}_m, \lambda_{nnk}\right) \triangleq \begin{bmatrix} \mathbf{U}_{nk} + \lambda_{nnk}\mathbf{Q}_{nnk} & \mathbf{U}_{nk}\hat{\mathbf{h}}_{nnk} \\ \hat{\mathbf{h}}_{nnk}^H\mathbf{U}_{nk} & \hat{\mathbf{h}}_{nnk}^H\mathbf{U}_{nk}\hat{\mathbf{h}}_{nnk} - \lambda_{nnk} - \sum_{m\neq n} t_{mnk} - \sigma_{nk}^2 \end{bmatrix} \succeq \mathbf{0}, \tag{15}$$

$$\mathbf{\Psi}_{mnk}\left(\{\mathbf{W}_{mi}\}_{i=1}^{K}, t_{mnk}, \lambda_{mnk}\right) \triangleq \begin{bmatrix} -\sum_{i=1}^{K}\mathbf{W}_{mi} + \lambda_{mnk}\mathbf{Q}_{mnk} & -\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)\hat{\mathbf{h}}_{mnk} \\ -\hat{\mathbf{h}}_{mnk}^H\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right) & -\hat{\mathbf{h}}_{mnk}^H\left(\sum_{i=1}^{K}\mathbf{W}_{mi}\right)\hat{\mathbf{h}}_{mnk} + t_{mnk} - \lambda_{mnk} \end{bmatrix} \succeq \mathbf{0}, m \in \mathcal{N}_c \setminus \{n\}. \tag{16}$$

and can be unbounded below. To fix this issue, we consider the use of ADMM [31], [32]. In the first subsection, we briefly review ADMM. In the second subsection, we present how a distributed robust MCBF algorithm can be developed following the principle of ADMM.

A. Review of ADMM

To illustrate the idea of ADMM, let us consider the following convex optimization problem [31], [32]:

$$\min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{z}\in\mathbb{R}^m} \quad F(\mathbf{x})+G(\mathbf{z}) \tag{19a}$$
$$\text{s.t.} \quad \mathbf{x}\in\mathcal{S}_1, \quad \mathbf{z}\in\mathcal{S}_2, \tag{19b}$$
$$\mathbf{z}=\mathbf{A}\mathbf{x} \tag{19c}$$

where $F:\mathbb{R}^n\mapsto\mathbb{R}$ and $G:\mathbb{R}^m\mapsto\mathbb{R}$ are convex functions, $\mathbf{A}$ is an $m\times n$ matrix, and $\mathcal{S}_1\subset\mathbb{R}^n$ and $\mathcal{S}_2\subset\mathbb{R}^m$ are nonempty convex sets. Assume that (19) is solvable and strong duality holds.

ADMM considers the following penalty augmented problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{z}\in\mathbb{R}^m} \quad F(\mathbf{x})+G(\mathbf{z})+\frac{c}{2}\|\mathbf{A}\mathbf{x}-\mathbf{z}\|^2 \tag{20a}$$
$$\text{s.t.} \quad \mathbf{x}\in\mathcal{S}_1, \quad \mathbf{z}\in\mathcal{S}_2, \tag{20b}$$
$$\mathbf{z}=\mathbf{A}\mathbf{x} \tag{20c}$$

where $c>0$ is the penalty parameter. It is easy to see that (20) is essentially equivalent to (19). The term $\frac{c}{2}\|\mathbf{A}\mathbf{x}-\mathbf{z}\|^2$ guarantees strict convexity of the objective function with respect to $\mathbf{x}$ and to $\mathbf{z}$, respectively.

The second ingredient of ADMM is dual decomposition [32] where the dual of (20) is concerned:

$$\max_{\boldsymbol{\xi}\in\mathbb{R}^m}\left\{\begin{array}{lll}\min_{\mathbf{x},\mathbf{z}} & F(\mathbf{x})+G(\mathbf{z})+\frac{c}{2}\|\mathbf{A}\mathbf{x}-\mathbf{z}\|^2+\boldsymbol{\xi}^T(\mathbf{A}\mathbf{x}-\mathbf{z}) \\ \text{s.t.} & \mathbf{x}\in\mathcal{S}_1, \quad \mathbf{z}\in\mathcal{S}_2\end{array}\right\} \tag{21}$$

in which $\boldsymbol{\xi}$ is the dual variable associated with the constraint (20c). Given a dual variable $\boldsymbol{\xi}$, the inner problem is a convex problem and can be efficiently solved. The outer variable $\boldsymbol{\xi}$ can be updated by the subgradient method [17]. In a standard dual optimization procedure, one usually updates the outer variable $\boldsymbol{\xi}$ when the associated inner problem has been solved for the global optimum. For example, one can use the nonlinear Gauss-Seidel method [31] to optimally solve the inner problem. Specifically, one iteratively solves the following two subproblems:

$$\mathbf{z}(q+1)=\arg\min_{\mathbf{z}\in\mathcal{S}_2} G(\mathbf{z})-\boldsymbol{\xi}(q)^T\mathbf{z}+\frac{c}{2}\|\mathbf{A}\mathbf{x}(q)-\mathbf{z}\|^2 \tag{22a}$$
$$\mathbf{x}(q+1)=\arg\min_{\mathbf{x}\in\mathcal{S}_1} F(\mathbf{x})+\boldsymbol{\xi}(q)^T\mathbf{A}\mathbf{x}+\frac{c}{2}\|\mathbf{A}\mathbf{x}-\mathbf{z}(q+1)\|^2 \tag{22b}$$

until convergence, where q is the iteration number. Instead, ADMM, as its name suggests, alternatively performs one iteration of the Gauss-Seidel step (22) and one step of the outer subgradient update for speeding up its convergence. The steps of ADMM are summarized in Algorithm 1.

ADMM can actually converge to the global optimum of (19) under relatively loose conditions by the following lemma.

Lemma 2 [31, Proposition 4.2]: Assume that $\mathcal{S}_1$ is bounded or that $\mathbf{A}^T\mathbf{A}$ is invertible. A sequence $\{\mathbf{x}(q),\mathbf{z}(q),\boldsymbol{\xi}(q)\}$ generated by Algorithm 1 is bounded, and every limit point of $\{\mathbf{x}(q),\mathbf{z}(q)\}$ is an optimal solution of (19).

Algorithm 1: ADMM

1: Set $q=0$, choose $c>0$,
2: Initialize $\boldsymbol{\xi}(q)$ and $\mathbf{x}(q)$;
3: **repeat**
4: Solve the two subproblems in (22).
5: $\boldsymbol{\xi}(q+1)=\boldsymbol{\xi}(q)+c(\mathbf{A}\mathbf{x}(q+1)-\mathbf{z}(q+1))$;
6: $q:=q+1$;
7: **until** the predefined stopping criterion is satisfied.

B. Applying ADMM to (17)

Our intention in this subsection is to reformulate (17) such that the corresponding ADMM steps in Algorithm 1 are decomposable and thus (17) can be solved in a distributed fashion. To this end, we first introduce the following two auxiliary variables:

$$p_n=\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk}), \quad T_{nk}=\sum_{m\neq n}^{N_c} t_{mnk} \tag{23}$$

for all $n\in\mathcal{N}_c$ and $k\in\mathcal{K}$, where p_n represents the transmission power of BS_n, and T_{nk} stands for the total worst-case ICI power from the neighboring BSs to MS_{nk}. Then (17) can be rewritten as

$$\min_{\substack{\{\mathbf{W}_{nk}\succeq\mathbf{0}\},\{\lambda_{mnk}\geq 0\},\\ \{t_{mnk}\}\{p_n\},\{T_{nk}\}}} \quad \sum_{n=1}^{N_c}\alpha_n p_n \tag{24a}$$
$$\text{s.t.} \quad \boldsymbol{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, T_{nk}, \lambda_{nnk}\right)\succeq\mathbf{0}, \tag{24b}$$
$$\boldsymbol{\Psi}_{mnk}\left(\{\mathbf{W}_{mi}\}_{i=1}^K, t_{mnk}, \lambda_{mnk}\right)\succeq\mathbf{0}, \tag{24c}$$
$$\sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})=p_n, \tag{24d}$$
$$\sum_{m\neq n}^{N_c} t_{mnk}=T_{nk}\geq 0, \quad \forall n,k,m\neq n. \tag{24e}$$

It is interesting to observe from (24b) that each MS_{nk} concerns only the total worst-case ICI power T_{nk} instead of the individual worst-case ICI powers $\{t_{mnk}\}$. Note that, with the problem unchanged, we can interchange the subindices m and n in (24c). Hence, the constraints in (24b) to (24d) can be decomposed into N_c independent convex sets

$$\begin{aligned}\mathcal{C}_n=\Big\{&\left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\{T_{nk}\}_k,\{t_{nmk}\}_{m,k},p_n\right)\Big| \\ &\boldsymbol{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K,T_{nk},\lambda_{nnk}\right)\succeq\mathbf{0} \quad \forall k \\ &\boldsymbol{\Psi}_{nmk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K,t_{nmk},\lambda_{nmk}\right)\succeq\mathbf{0} \quad \forall m\neq n,k \\ &\lambda_{nmk}\geq 0\ \forall m,k,\ \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_{nk})=p_n \\ &\mathbf{W}_{nk}\succeq\mathbf{0}\ \forall k, T_{nk}\geq 0 \quad \forall k\Big\}, \quad \forall n\in\mathcal{N}_c.\end{aligned} \tag{25}$$

Further define the following variables:

$$\mathbf{t}=\left[[t_{121},\ldots,t_{12K}],\ldots,\left[t_{N_c(N_c-1)1},\ldots,t_{N_c(N_c-1)K}\right]\right]^T \in\mathbb{R}^{N_c(N_c-1)K}, \tag{26a}$$

$$\mathbf{t}_n = [[T_{n1},\ldots,T_{nK}],[t_{n11},\ldots,t_{n1K}],\ldots,\\ [t_{nN_c1},\ldots,t_{nN_cK}]]^T \in \mathbb{R}_+^{N_cK}, n\in\mathcal{N}_c \quad (26b)$$

where $\mathbf{t}$ collects all the ICI variables, and $\mathbf{t}_n$ collects variables $\{T_{nk}\}_{k=1}^K$ and $\{t_{nmk}\}_{m,k}$ (where $m\neq n$) that are only relevant to BS_n. It is not difficult to check that there exists a linear mapping matrix $\mathbf{E}_n \in \{0,1\}^{N_cK\times N_c(N_c-1)K}$, such that $\mathbf{t}_n=\mathbf{E}_n\mathbf{t}$ for all $n\in\mathcal{N}_c$. By (25), (26), and $\mathbf{t}_n=\mathbf{E}_n\mathbf{t}$, we can rewrite (24) in a compact form as

$$\min_{\substack{\{\mathbf{W}_{nk}\},\{\lambda_{nmk}\},\\ \{\mathbf{t}_n\},\{p_n\},\mathbf{t}}} \sum_{n=1}^{N_c}\alpha_n p_n \quad (27a)$$
$$\text{s.t. } \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n \quad (27b)$$
$$\mathbf{t}_n=\mathbf{E}_n\mathbf{t},\quad \forall n\in\mathcal{N}_c. \quad (27c)$$

Note that in (26a) and (27), the elements of $\mathbf{t}$ are relaxed to the real space rather than explicitly required to be nonnegative so as to simplify the later optimization steps of ADMM. It will be seen that ADMM will converge asymptotically with a nonnegative iterate $\mathbf{t}$ in the sequel due to the constraints (26b) and (27c).

Before applying ADMM, let us first see why the conventional dual decomposition method [17] is not suitable for (27). Consider the dual problem of (27)

$$\max_{\substack{\boldsymbol{\nu}_n\in\mathbb{R}^{N_cK}\forall n,\\ \sum_{n=1}^{N_c}\mathbf{E}_n^T\boldsymbol{\nu}_n=\mathbf{0}}} \left\{\begin{array}{ll}\min & \sum\limits_{n=1}^{N_c}\alpha_n p_n-\sum_{n=1}^{N_c}\boldsymbol{\nu}_n^T\mathbf{t}_n\\ \text{s.t.} & \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n\forall n\end{array}\right\} \quad (28)$$

where $\boldsymbol{\nu}_n\in\mathbb{R}^{N_cK}$, $n\in\mathcal{N}_c$, are the dual variables associated with (27c). While the inner minimization problem of (28) is obviously decomposable, given $\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_{N_c}$, it is possible to obtain an inner solution of $\mathbf{t}_n$ such that $-\boldsymbol{\nu}_n^T\mathbf{t}_n\to-\infty$, i.e., the inner minimization problem is unbounded below, due to the unbounded feasible sets $\mathcal{C}_n$ [see (25), (15), and (16)]. In fact, by our numerical experience, this undesired situation happens very often, especially when $N_c>2$.

To overcome this issue, we apply the augmented Lagrangian method to (27) according to the principle of ADMM; this leads to the following problem:

$$\min_{\substack{\{\mathbf{W}_{nk}\},\{\lambda_{nmk}\},\{\mathbf{t}_n\},\\ \{p_n\},\{\rho_n\},\mathbf{t}}} \left\{\sum_{n=1}^{N_c}\alpha_n p_n+\frac{c}{2}\sum_{n=1}^{N_c}\|\mathbf{E}_n\mathbf{t}-\mathbf{t}_n\|^2 +\frac{c}{2}\sum_{n=1}^{N_c}(\rho_n-p_n)^2\right\} \quad (29a)$$
$$\text{s.t. } \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n\ \forall n\in\mathcal{N}_c \quad (29b)$$
$$\mathbf{t}_n=\mathbf{E}_n\mathbf{t}\quad\forall n\in\mathcal{N}_c \quad (29c)$$
$$p_n=\rho_n\quad\forall n\in\mathcal{N}_c \quad (29d)$$

where $\rho_n\geq 0$, $n=1,\ldots,N_c$, are slack variables, which are introduced in order to impose the penalty term $\frac{c}{2}\sum_{n=1}^{N_c}(\rho_n-p_n)^2$. Problem (29) is equivalent to (27), but the added penalty terms can resolve the numerically unbounded below issue; see [31].

Now we are ready to apply ADMM. Consider the following correspondences between (29) and (20):

$$\mathbf{x}\triangleq\left[\mathbf{t}^T,\rho_1,\ldots,\rho_{N_c}\right]^T,\ \mathbf{z}\triangleq\left[\mathbf{t}_1^T,\ldots,\mathbf{t}_{N_c}^T,p_1,\ldots,p_{N_c}\right]^T,$$
$$F(\mathbf{x})\triangleq 0,\ G(\mathbf{z})\triangleq\sum_{n=1}^{N_c}\alpha_n p_n,\ \mathbf{A}\triangleq\begin{bmatrix}\mathbf{E} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{N_c}\end{bmatrix},$$
$$\boldsymbol{\xi}\triangleq\left[\boldsymbol{\nu}_1^T,\ldots,\boldsymbol{\nu}_{N_c}^T,\mu_1,\ldots,\mu_{N_c}\right]^T,$$
$$\mathcal{S}_1\triangleq\mathbb{R}^{N_c(N_c-1)K+N_c}$$
$$\mathcal{S}_2\triangleq\left\{\left[\mathbf{t}_1^T,\ldots,\mathbf{t}_{N_c}^T,p_1,\ldots,p_{N_c}\right]^T \;\middle|\; \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n,n\in\mathcal{N}_c\right\} \quad (30)$$

where $\mathbf{E}\triangleq\left[\mathbf{E}_1^T,\cdots,\mathbf{E}_{N_c}^T\right]^T$ and $\boldsymbol{\nu}_n\in\mathbb{R}^{N_cK}$, $\mu_n\in\mathbb{R}$, $n=1,\ldots,N_c$, are the dual variables associated with constraints (29c) and (29d), respectively. According to Algorithm 1, the corresponding ADMM step 4 for (29) is to solve the following:

$$\min_{\substack{\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\\ \mathbf{t}_n,p_n,n=1,\ldots,N_c}} \sum_{n=1}^{N_c}\begin{pmatrix}\alpha_n p_n+\frac{c}{2}\|\mathbf{E}_n\mathbf{t}(q)-\mathbf{t}_n\|^2\\ +\frac{c}{2}\left(\rho_n(q)-p_n\right)^2\\ -\boldsymbol{\nu}_n^T(q)\mathbf{t}_n-\mu_n(q)p_n\end{pmatrix}$$
$$\text{s.t. } \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n,\ n=1,\ldots,N_c. \quad (31)$$

As one can see, (31) can be decomposed as the following N_c subproblems:

$$\{\mathbf{t}_n(q+1),p_n(q+1)\} =\arg\min\begin{pmatrix}\alpha_n p_n+\frac{c}{2}\|\mathbf{E}_n\mathbf{t}(q)-\mathbf{t}_n\|^2\\ +\frac{c}{2}\left(\rho_n(q)-p_n\right)^2\\ -\boldsymbol{\nu}_n^T(q)\mathbf{t}_n-\mu_n(q)p_n\end{pmatrix}$$
$$\text{s.t. } \left(\{\mathbf{W}_{nk}\}_k,\{\lambda_{nmk}\}_{m,k},\mathbf{t}_n,p_n\right)\in\mathcal{C}_n. \quad (32)$$

for $n=1,\ldots,N_c$. Since (32) is convex, it can be efficiently solved.

Second, the corresponding ADMM step 5 is given by solving the following two problems:

$$\mathbf{t}(q+1)=\arg\min_{\mathbf{t}\in\mathbb{R}^{N_c(N_c-1)K}}\frac{c}{2}\sum_{n=1}^{N_c}\|\mathbf{E}_n\mathbf{t}-\mathbf{t}_n(q+1)\|^2 +\sum_{n=1}^{N_c}\boldsymbol{\nu}_n^T(q)\mathbf{E}_n\mathbf{t}, \quad (33)$$

$$\{\rho_n(q+1)\}_{n=1}^{N_c}=\arg\min_{\substack{\rho_n\in\mathbb{R},\\ n=1,\ldots,N_c}}\frac{c}{2}\sum_{n=1}^{N_c}\left(\rho_n-p_n(q+1)\right)^2 +\sum_{n=1}^{N_c}\mu_n(q)\rho_n. \quad (34)$$

Because both (33) and (34) are convex quadratic problems, they have closed-form solutions given by

$$\mathbf{t}(q+1)=\mathbf{E}^\dagger\left(\tilde{\mathbf{t}}(q+1)-\frac{1}{c}\tilde{\boldsymbol{\nu}}(q)\right), \quad (35a)$$
$$\rho_n(q+1)=p_n(q+1)-\frac{1}{c}\mu_n(q),\quad n\in\mathcal{N}_c, \quad (35b)$$

where $\tilde{\mathbf{t}}(q+1) = [\mathbf{t}_1^T(q+1), \ldots, \mathbf{t}_{N_c}^T(q+1)]^T$ and $\tilde{\boldsymbol{\nu}}(q) = [\boldsymbol{\nu}_1^T(q), \ldots, \boldsymbol{\nu}_{N_c}^T(q)]^T$.

Finally, the corresponding ADMM step 6 is given by the following dual variable update

$$\boldsymbol{\nu}_n(q+1) = \boldsymbol{\nu}_n(q) + c\left(\mathbf{E}_n\mathbf{t}(q+1) - \mathbf{t}_n(q+1)\right), \quad (36a)$$
$$\mu_n(q+1) = \mu_n(q) + c\left(\rho_n(q+1) - p_n(q+1)\right), \quad (36b)$$

for all $n \in \mathcal{N}_c$. It is important to note that the ADMM steps (32), (35), and (36) can be implemented in a distributed fashion. Essentially, given the knowledge of local CSI $\{\hat{\mathbf{h}}_{nmk}\}_{m,k}$, the optimization problem (32) can be independently solved by BS_n, for all $n = 1, \ldots, N_c$. After that, each BS_n broadcasts its new $\mathbf{t}_n$ to the other BSs. With the knowledge of $\{\mathbf{t}_n\}$, each BS can compute the public ICI iterate $\mathbf{t}$ by (35a) and then use it to update the dual variables $\{\boldsymbol{\nu}_n(q+1)\}$ by (36a). Moreover, both $\rho_n(q+1)$ and $\mu_n(q+1)$ in (35b) and (36b) can be independently updated by each BS_n, for $n = 1, \ldots, N_c$. Summarizing the above steps, we thus obtain the distributed robust MCBF algorithm in Algorithm 2.

Algorithm 2: Proposed Distributed Robust MCBF Algorithm

1: **Input** a set of initial variables $\{\boldsymbol{\nu}_n(0), \mu_n(0), \mathbf{t}(0), \rho_n(0)\}_{n=1}^{N_c}$ that are known to all BSs; choose a penalty parameter $c > 0$.

2: Set $q = 0$.

3: **repeat**

4: Each BS_n solves the local beamforming design problem (32) to obtain the local ICI iterate $\mathbf{t}_n(q+1)$ and the local power $p_n(q+1)$.

5: Each BS_n informs the other BSs of its local ICI iterate $\mathbf{t}_n$.

6: Each BS_n updates the public ICI $\mathbf{t}$ and ρ_n by (35a) and (35b), respectively.

7: Each BS_n updates the dual variables $\{\boldsymbol{\nu}_n\}$ and μ_n by (36a) and (36b), respectively.

8: Set $q := q + 1$;

9: **until** the predefined stopping criterion is met.

Interestingly, Algorithm 2 can be interpreted as an adaptive ICI regularization scheme where the cooperative BSs gradually attain their own beamforming solutions until a consensus on the induced ICI powers among BSs is reached, i.e., $\mathbf{E}_n\mathbf{t}(q+1) = \mathbf{t}_n(q+1)$ for all n.

Algorithm 2 is guaranteed to converge to the global optimum of the SDR problem (17). Specifically, one can verify that the matrix $\mathbf{A}$ in (30) satisfies $\mathbf{A}^T\mathbf{A} \succ \mathbf{0}$. Therefore, by Lemma 2, we obtain the following result on the convergence of Algorithm 2.

Proposition 2: *For the proposed distributed robust MCBF algorithm in Algorithm 2, the iterates* $\mathbf{t}(q)$, $\{p_n(q), \mathbf{t}_n(q), \rho_n(q)\}_{n=1}^{N_c}$ *and* $\{\boldsymbol{\nu}_n(q), \mu_n(q)\}_{n=1}^{N_c}$ *will, respectively, converge to the optimal primal and dual solutions of (29) as* $q \to \infty$. *When the algorithm converges, the optimal* $\{\mathbf{W}_{n1}, \ldots, \mathbf{W}_{nK}\}_{n=1}^{N_c}$ *obtained in Step 4 is a global optimal solution of the SDR problem (17).*

Three remarks regarding the proposed distributed robust MCBF algorithm are in order.

Remark 1: Since ADMM operates in the dual domain without guarantee of the constraint $\mathbf{E}_n\mathbf{t}(q+1) = \mathbf{t}_n(q+1)$ to hold true before convergence, the obtained $\{\mathbf{W}_{nk}\}$ and $\{\lambda_{nmk}\}$ in Step 4 thus may not be feasible to the primal problem (17). However, each BS may perform one more optimization of

$$\begin{aligned}
\min_{\substack{\{\mathbf{W}_{nk}\succeq\mathbf{0}\}_k,\\ \{\lambda_{nmk}\geq 0\}_{m,k}}} \quad & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_{nk}) \\
\text{s.t.} \quad & \boldsymbol{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, \{t_{mnk}(q+1)\}_{m\neq n}, \lambda_{nnk}\right) \succeq \mathbf{0}\ \forall k \\
& \boldsymbol{\Psi}_{nmk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, t_{nmk}(q+1), \lambda_{nmk}\right) \succeq \mathbf{0}\ \forall m \neq n, k
\end{aligned} \quad (37)$$

using the tentatively consented ICI power vector $\mathbf{t}(q+1)$. It can be shown that the obtained $\{\mathbf{W}_{nk}\}$ and $\{\lambda_{nmk}\}$ are feasible to the SDR problem (17), provided that (37) is feasible for all BSs. If at least one of the BSs declares infeasibility of (37), then more iterations are needed for Algorithm 2 since it may stop too early to reach a reasonable consensus on the global ICI $\mathbf{t}(q+1)$.

Remark 2: In Algorithm 2, each BS_n is required to inform the other BSs of its local ICI iterate $\mathbf{t}_n$ at each iteration. In ad-hoc networks [41], this information exchange is commonly achieved through broadcasting [41]. However, for a cellular system where the BSs are physically connected via dedicated fiber links or by microwave radio links, BSs are preferred to communicate with each other in a point-to-point fashion. In that case, Algorithm 2 requires each BS to send out N_cK real values (i.e., $\mathbf{t}_n$) to the other $(N_c - 1)$ BSs in a one-by-one manner, resulting in a total signaling overhead of $(N_c - 1)N_cK$ real values. Interestingly, Algorithm 2 can actually be modified so that the excessive amount of signaling overhead due to point-to-point information exchange can be appreciably reduced for large N_c or K. Such an alternative distributed robust MCBF algorithm is detailed in Appendix B.

Remark 3: There actually exist several variants of ADMM in the literature of convex optimization, e.g., the fast iterative shrinkage-thresholding algorithm (FISTA) [42] and the Bregman distance-based primal-dual algorithm [34], that may potentially be applied to solving the SDR problem (17) for faster convergence and reduced backhaul signaling overhead. Further investigation of such possibilities will be an interesting direction for future research.

V. Extension to Fully Coordinated BSs

In this section, we extend the robust beamforming design to the scenario where some of the MSs are near the cell boundary and thus desire to receive the signal of interest sent from multiple BSs for the guaranteed QoS. To simultaneously serve these MSs, the BSs have to be fully coordinated, with shared data streams and CSI of these cell-edge MSs [2]. Assume that there are L cell-edges MSs, in addition to the K intracell MSs in each cell. The transmit signal of BS_n is given by

$$\tilde{\mathbf{x}}_n(t) = \mathbf{x}_n(t) + \sum_{\ell=1}^{L} \mathbf{f}_{n\ell} d_\ell(t) \quad (38)$$

where $\mathbf{x}_n(t)$ has been defined in (1) which is intended for the K intracell MSs, $d_\ell(t)$ is the data stream for the ℓth cell-edge MS, and $\mathbf{f}_{n\ell} \in \mathbb{C}^{N_t}$ is the beamforming vector of BS_n for sending $d_\ell(t)$. The received signals of intracell MS_{nk} and cell-edge MS_ℓ are, respectively, given by

$$y_{nk}^{(\mathrm{Intra})}(t) = \sum_{m=1}^{N_c} \mathbf{h}_{mnk}^H \mathbf{x}_m \left(t - \tau_{nk}^{(m)}\right) + \sum_{m=1}^{N_c} \sum_{j=1}^{L} \mathbf{h}_{mnk}^H \mathbf{f}_{mj} d_j \left(t - \tau_{nk}^{(m)}\right) + z_{nk}(t) \quad (39)$$

$$y_{\ell}^{(\mathrm{Edge})}(t) = \sum_{m=1}^{N_c} \mathbf{g}_{m\ell}^H \mathbf{x}_m \left(t - \tau_{\ell}^{(m)}\right) + \sum_{m=1}^{N_c} \sum_{j=1}^{L} \mathbf{g}_{m\ell}^H \mathbf{f}_{mj} d_j \left(t - \tau_{\ell}^{(m)}\right) + z_{\ell}(t) \quad (40)$$

for $n \in \mathcal{N}_c, k \in \mathcal{K}$, and $\ell \in \mathcal{L} \triangleq \{1, \ldots, L\}$, where $\mathbf{g}_{m\ell} \in \mathbb{C}^{N_t}$ is the channel vector from BS_m to cell-edge MS_ℓ, and $z_\ell(t)$ is the additive noise at cell-edge MS_ℓ, which is assumed to have zero mean and variance $\sigma_\ell^2 > 0$. Note from (39) and (40) that we have taken into account the inevitable time delays $\tau_{nk}^{(m)}, \tau_\ell^{(m)} > 0$ between the BSs and MSs [8], [43]. Assume that $\tau_\ell^{(m)} \neq \tau_\ell^{(n)}$ for all $m \neq n$, and that each $d_\ell(t)$ is temporally uncorrelated with zero mean and unit variance. The receiver SINRs corresponding to (39) and (40) are given by (41) and (42) [8], shown at the bottom of the page.

Our goal here is, again, to find the beamforming vectors that are robust against the possible CSI errors. As the channel error model for intracell MSs, we model the cell-edge MSs' channel as $\mathbf{g}_{m\ell} = \hat{\mathbf{g}}_{m\ell} + \mathbf{v}_{m\ell}\ \forall m \in \mathcal{N}_c, \ell \in \mathcal{L}$, where $\hat{\mathbf{g}}_{m\ell} \in \mathbb{C}^{N_t}$ is the preassumed CSI, and $\mathbf{v}_{m\ell} \in \mathbb{C}^{N_t}$ is the CSI error satisfying $\mathbf{v}_{m\ell}^H \tilde{\mathbf{Q}}_{m\ell} \mathbf{v}_{m\ell} \leq 1$ in which $\tilde{\mathbf{Q}}_{m\ell} \succ \mathbf{0}$. We consider the following worst-case robust formulation:

$$\min_{\{\mathbf{w}_{nk}\},\{\mathbf{f}_{n\ell}\}} \sum_{n=1}^{N_c} \alpha_n \left(\sum_{k=1}^{K} \|\mathbf{w}_{nk}\|^2 + \sum_{\ell=1}^{L} \|\mathbf{f}_{n\ell}\|^2 \right) \quad (43a)$$

$$\text{s.t. } \mathrm{SINR}_{nk}^{(\mathrm{Intra})}\left(\{\mathbf{w}_{mi}\}, \{\mathbf{f}_{mj}\}, \{\hat{\mathbf{h}}_{mnk} + \mathbf{e}_{mnk}\}_{m=1}^{N_c}\right) \geq \gamma_{nk},$$
$$\forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1, m, n \in \mathcal{N}_c, k \in \mathcal{K}, \quad (43b)$$
$$\mathrm{SINR}_{\ell}^{(\mathrm{Edge})}\left(\{\mathbf{w}_{mi}\}, \{\mathbf{f}_{mj}\}, \{\hat{\mathbf{g}}_{m\ell} + \mathbf{v}_{m\ell}\}_{m=1}^{N_c}\right) \geq \gamma_{\ell},$$
$$\forall \mathbf{v}_{m\ell}^H \tilde{\mathbf{Q}}_{m\ell} \mathbf{v}_{m\ell} \leq 1, m \in \mathcal{N}_c, \ell \in \mathcal{L}. \quad (43c)$$

The proposed method based on SDR and S-lemma in Section III-A can be used to handle (43) as well. First, replace each $\mathbf{w}_{nk}\mathbf{w}_{nk}^H$ and each $\mathbf{f}_{n\ell}\mathbf{f}_{n\ell}^H$ by general-rank $\mathbf{W}_{nk} \succeq \mathbf{0}$ and $\mathbf{F}_{n\ell} \succeq \mathbf{0}$, respectively. Second, follow the steps in (8)–(16) to decouple and transform the constraints in (43b) and (43c) into a finite number of LMIs. The resulting SDR problem can be shown to be

$$\min_{\substack{\{\mathbf{W}_{nk}\},\{\mathbf{F}_{nk}\},\\ \{\lambda_{mnk}\},\{t_{mnk}\}\\ \{\tilde{\lambda}_{m\ell}\},\{\eta_{m\ell}\}}} \sum_{n=1}^{N_c} \alpha_n \left(\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_{nk}) + \sum_{\ell=1}^{L} \mathrm{Tr}(\mathbf{F}_{n\ell}) \right)$$
$$\text{s.t. } \boldsymbol{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, \{t_{mnk}\}_m, \lambda_{nnk}\right) - \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}_{nnk}^H \end{bmatrix} \left(\sum_{j=1}^{L} \mathbf{F}_{nj} \right) \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}_{nnk}^H \end{bmatrix}^H \succeq \mathbf{0},\ \forall n, k,$$
$$\boldsymbol{\Psi}_{mnk}\left(\{\mathbf{W}_{mi}\}_{i=1}^K, t_{mnk}, \lambda_{mnk}\right) - \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}_{mnk}^H \end{bmatrix} \left(\sum_{j=1}^{L} \mathbf{F}_{mj} \right) \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}_{mnk}^H \end{bmatrix}^H \succeq \mathbf{0},\ \forall n, m \neq n, k,$$
$$\begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{g}}_{m\ell}^H \end{bmatrix} \left(\frac{1}{\gamma_\ell} \mathbf{F}_{m\ell} - \sum_{i=1}^{K} \mathbf{W}_{mi} - \sum_{j\neq\ell}^{L} \mathbf{F}_{mj} \right) \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{g}}_{m\ell}^H \end{bmatrix}^H + \begin{bmatrix} \tilde{\lambda}_{m\ell} \tilde{\mathbf{Q}}_{m\ell} & \mathbf{0} \\ \mathbf{0} & -\tilde{\lambda}_{m\ell} - \eta_{m\ell} \end{bmatrix} \succeq \mathbf{0},\ \forall m, \ell,$$
$$\mathbf{W}_{nk} \succeq \mathbf{0}, \mathbf{F}_{m\ell} \succeq \mathbf{0}, \lambda_{mnk} \geq 0, \tilde{\lambda}_{m\ell} \geq 0,$$
$$\sum_{m=1}^{N_c} \eta_{m\ell} - \sigma_\ell^2 \geq 0,\ \forall m, n, k, \ell \quad (44)$$

where $\{\tilde{\lambda}_{m\ell}\}, \{\eta_{m\ell}\}$ are slack variables. For the spherical error model, a sufficient condition for the tightness of SDR, which is similar to Proposition 1, can be shown to be

$$\varepsilon_{nmk} \leq \bar{\varepsilon}_{nmk},\ \varepsilon_{n\ell} \leq \bar{\varepsilon}_{n\ell},\ \varepsilon_{nnk} < \sqrt{\alpha_n \gamma_{nk} \sigma_{nk}^2 / g^\star}, \quad \text{and}$$
$$\varepsilon_{n\ell} < \sqrt{\alpha_n \gamma_\ell \sigma_\ell^2 / g^\star} \quad \forall n, m, k, \ell \quad (45)$$

where $\{\bar{\varepsilon}_{nmk}\}, \{\bar{\varepsilon}_{n\ell}\}$ are the CSI error bounds for which (44) is feasible, and $g^\star > 0$ is the associated optimal objective value. The condition in (45) implies that (44) can attain the global optimum of (43) if the CSI errors are sufficiently small. A distributed optimization algorithm for (44) can also be developed

$$\mathrm{SINR}_{nk}^{(\mathrm{Intra})}\left(\{\mathbf{w}_{mi}\}, \{\mathbf{f}_{mj}\}, \{\mathbf{h}_{mnk}\}_{m=1}^{N_c}\right) = \frac{\left|\mathbf{h}_{nnk}^H \mathbf{w}_{nk}\right|^2}{\sum_{i\neq k}^{K} \left|\mathbf{h}_{nnk}^H \mathbf{w}_{ni}\right|^2 + \sum_{m\neq n}^{N_c} \sum_{i=1}^{K} \left|\mathbf{h}_{mnk}^H \mathbf{w}_{mi}\right|^2 + \sum_{m=1}^{N_c} \sum_{j=1}^{L} \left|\mathbf{h}_{mnk}^H \mathbf{f}_{mj}\right|^2 + \sigma_{nk}^2}, \quad (41)$$

$$\mathrm{SINR}_{\ell}^{(\mathrm{Edge})}\left(\{\mathbf{w}_{mi}\}, \{\mathbf{f}_{mj}\}, \{\mathbf{g}_{m\ell}\}_{m=1}^{N_c}\right) = \frac{\sum_{m=1}^{N_c} \left|\mathbf{g}_{m\ell}^H \mathbf{f}_{m\ell}\right|^2}{\sum_{m=1}^{N_c} \sum_{i=1}^{K} \left|\mathbf{g}_{m\ell}^H \mathbf{w}_{mi}\right|^2 + \sum_{j\neq\ell}^{L} \sum_{m=1}^{N_c} \left|\mathbf{g}_{m\ell}^H \mathbf{f}_{mj}\right|^2 + \sigma_\ell^2}. \quad (42)$$

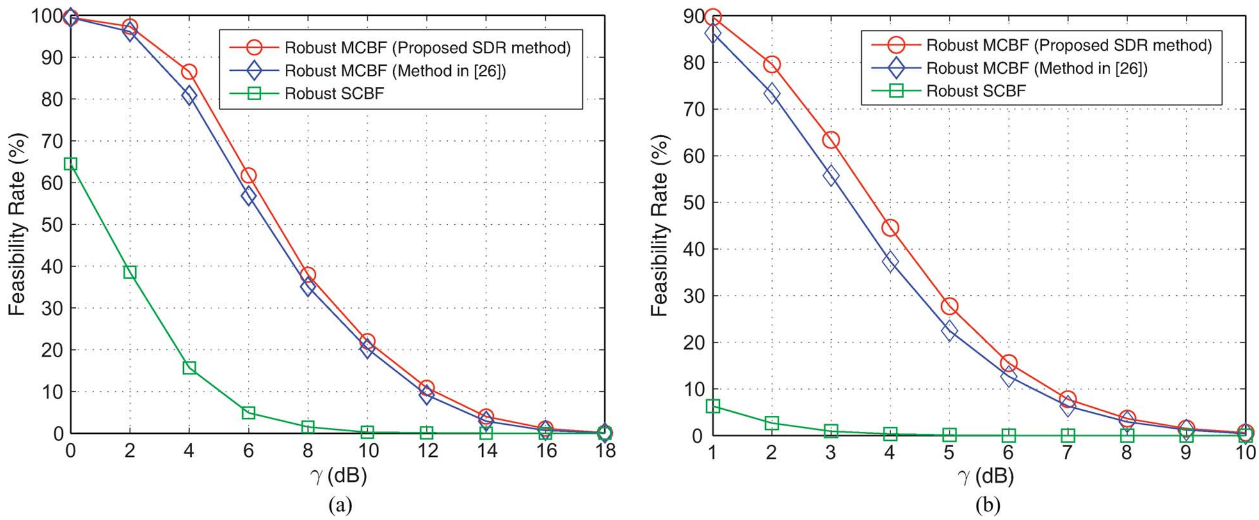


Fig. 2. Feasibility rate (%) versus SINR requirement γ for $K = 4, N_t = 6, \varepsilon = 0.1$. (a) $N_c = 2$. (b) $N_c = 3$.

by ADMM, using the same ideas as presented in Section IV-B for (17). We will provide a simulation example in Section VI to demonstrate the efficacy of (44) in providing guaranteed QoS for the cell-edge MSs.

VI. Simulation Results

A. Simulation Setting

In the simulations, we not only consider the small scale channel fading but also the large scale fading effects such as shadowing and path loss, in order to simulate the multicell scenario. Specifically, we follow the channel model [6], [44]

$$\mathbf{h}_{mnk} = 10^{-(128.1+37.6\log_{10}(d_{mnk}))/20} \cdot \psi_{mnk} \cdot \varphi_{mnk} \cdot (\hat{\mathbf{h}}_{mnk} + \mathbf{e}_{mnk}) \quad (46)$$

where the exponential term is due to the path loss depending on the distance between the mth BS and MS_{nk} (denoted by d_{mnk} in kilometers), ψ_{mnk} reflects the shadowing effect, and φ_{mnk} represents the transmit-receive antenna gain. The term inside the parentheses in (46) denotes the small scale fading which consists of the preassumed CSI $\hat{\mathbf{h}}_{mnk}$ and the CSI error $\mathbf{e}_{mnk}$. As seen from (46), it is assumed that the BSs can accurately track the large scale fading, and suffers only from the small scale CSI errors.

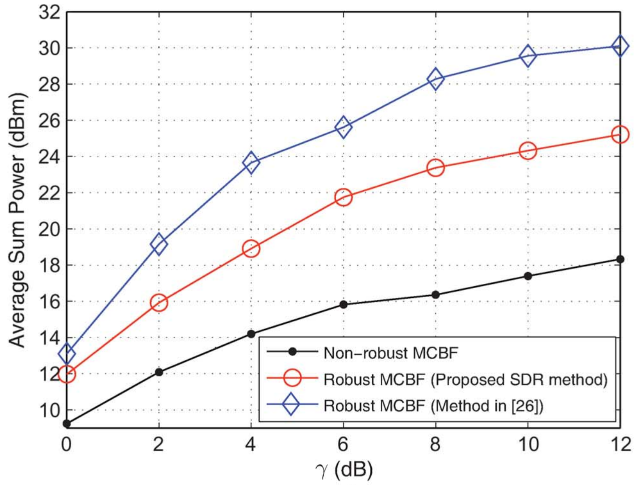
The inter-BS distance is 500 m, and the locations of the MSs in each cell are randomly determined with the distance to the serving BS at least 35 m, i.e., $d_{nnk} \geq 0.035$ for all n, k. The shadowing coefficient ψ_{mnk} follows the log-normal distribution with zero mean and standard deviation equal to 8. The elements of the preassumed CSI $\{\hat{\mathbf{h}}_{mnk}\}$ are i.i.d. complex Gaussian random variables with zero mean and unit variance. We also assume that all MSs have the same noise power spectral density equal to -162 dBm/Hz (-92 dBm over a 10–MHz bandwidth), and each BS has a maximum power limit 46 dBm [44]. The SINR requirements of MSs are the same, i.e., $\gamma_{mnk} \triangleq \gamma$, and each link has the same antenna gain $\varphi_{mnk} = 15$ dBi. The power weight α_n for BS_n is set to one for all n (i.e., sum power). For the CSI errors, the spherical error model is considered, i.e., $\mathbf{Q}_{mnk} = \epsilon_{mnk}^{-2}\mathbf{I}_{N_t}$ for all m, n and k. If not mentioned specifically, all error radii ε_{mnk} are the same and equal to ε.

B. Performance Comparison With Existing Methods

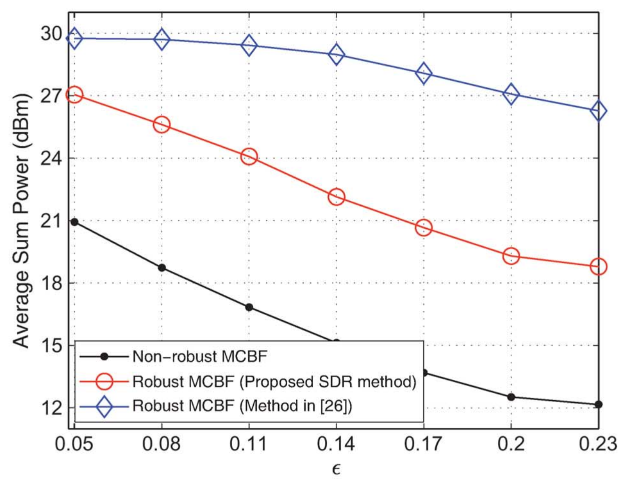
For the robust MCBF design (7), we compare the proposed SDR method with the convex restrictive approximation method in [26] which is the only existing method as far as we know. The single-cell beamforming (SCBF) design with independent ICI constraints [1], [45] is also compared. All the design formulations are solved by `SeDuMi` [37].

We first present the feasibility rates of the three beamforming designs. In the simulation, (7) is considered feasible for a channel realization if it can yield an optimal solution with each BS's power no greater than 46 dBm. Fig. 2 presents the simulation results obtained by testing over seven thousand channel realizations. One can observe from this figure that both robust MCBF designs exhibit much higher feasibility rates than the robust SCBF design, showing the improved capability of coordinated beamforming by exploiting the degrees of freedom provided by multiple BSs. Second, one can see that the proposed SDR method exhibits a slightly higher feasibility rate than the method in [26]. We should emphasize that in the simulation tests, the SDR problem (17) all yields rank-one solutions. Hence, the feasibility rate shown here is in fact that of the original problem (7).

Next, let us examine the average sum powers of the two robust MCBF designs. As a performance benchmark, we also present the average sum power of the nonrobust MCBF design (4). Fig. 3(a) and (b) shows the results of average sum power (dBm) versus the SINR requirement γ and versus the CSI error radius ε, respectively. Note that, over the seven thousand channel realizations, each of the results in Fig. 3 for one value of γ or ε is obtained by averaging over the channel realizations for which the three methods under test are all feasible. One can observe from both figures that, as a price for worst-case performance guarantee, the robust MCBF designs require higher average transmission powers than the nonrobust design, but the proposed SDR method is more power efficient than the method in [26]. For example, for $\gamma = 10$ dB in Fig. 3(a), the proposed



(a)



(b)

Fig. 3. Average transmission sum power (dBm) of various methods for $N_c = 2$, $K = 4$ and $N_t = 6$. (a) $\varepsilon = 0.1$. (b) $\gamma = 10$ dB.

SDR method consumes around 24 dBm while the method in [26] requires 29 dBm. On the other hand, it is noticeable from Fig. 3(b) that the average powers of all methods decrease with ε, which seems counter-intuitive. The reason for this is that the set and the total number of feasible channel realizations used for evaluating the average powers significantly vary with ε, and thus the obtained average powers turn out not necessarily to increase with ε.

C. Performance of Proposed Distributed Robust MCBF Algorithm

Now, let us examine the performance of the proposed distributed robust MCBF algorithm (Algorithm 2). In the simulations, the initial input values $\{\boldsymbol{\nu}_n(0), \mu_n(0), \mathbf{t}(0), \rho_n(0)\}_{n=1}^{N_c}$ are all set to zero. The augmented penalty parameter c is not a constant, but given by

$$c(q+1) := \begin{cases} qc(q), & \text{if } c(q) < 1, \\ 1, & \text{otherwise,} \end{cases} \tag{47}$$

where $c(0) = 10^{-6}$. As one can verify that $c(q)$ will reach the value of one after nine iterations, Algorithm 2 following (47) will still converge to the global optimum for a sufficiently large q, according to Proposition 2. First, we compare the optimal sum power of the centralized problem (7) with that obtained by Algorithm 2. The simulation results by testing over 50 randomly generated channel realizations are presented in Fig. 4, under various simulation settings. From Fig. 4(a) and Fig. 4(b), where $N_c = 2$ and $K = 2$ and $K = 4$, respectively, we can observe that Algorithm 2 can yield near-optimal solutions within 50 iterations. As observed, for most of the cases, 10 and 20 iterations are quite sufficient for the scenarios in Fig. 4(a) and Fig. 4(b), respectively. When the number of cells N_c increases to three $(N_c = 3)$, as shown in Fig. 4(c), 25 iterations are sufficient to obtain a near-optimal solution. General speaking, as the number of cells and that of MSs increase, the number of iterations needed to achieve a near-optimal performance also increases. As seen from Fig. 4(d), where $N_c = 8$, at least 100 iterations are required.

To further look into the convergence behavior of Algorithm 2, we show in Fig. 5(a) the typical convergence curves of Algorithm 2 in the scenarios considered in Fig. 4(a) to Fig. 4(c). In Fig. 5, the normalized power accuracy is defined as

$$\text{Normalized power accuracy} = \frac{|P^\star(q) - P^\star|}{P^\star} \tag{48}$$

where $P^\star(q) = \sum_{n=1}^{N_c} p_n(q)$ is the sum power at iteration q, and $P^\star$ denotes the centralized solution of (17). We can see from Fig. 5(a) that Algorithm 2 can yield a solution with the normalized power accuracy smaller than 0.1 within 50 iterations. Fig. 5(b) shows the convergence curves of Algorithm 2 for $N_c = 2$, $K = 2$, $N_t = 4$, and various CSI error radii and SINR requirements. It can be seen from this figure that the convergence speed can be slowed down as the CSI error radius or SINR requirement increases. Nevertheless, for the scenarios considered in Fig. 5(b), less than 40 iterations are needed for achieving 0.01 normalized power accuracy. The simulation results shown in Fig. 4 and Fig. 5 well demonstrate the convergence of Algorithm 2, as stated in Proposition 2.

D. Performance of Robust Fully Coordinated BF

In this subsection, we examine the effectiveness of the robust fully coordinated BF design (43) in serving the cell-edge MSs. To this end, let us consider a three-cell system $(N_c = 3)$ with two MSs in each cell $(K = 2)$. As illustrated in Fig. 6, we divide each cell into two parts, namely, the intracell region (circular disks) and the cell-edge region. In particular, the inter-BS distance is set to 500 meters and the radius for the intracell region is 235 meters. In each cell, the position of one of the MSs is randomly generated within the intracell region; while the other MS is randomly located in the cell-edge region within the equilateral triangle formed by the three BSs (see Fig. 6). As the robust fully coordinated BF design (43) is applied, the three MSs in the cell-edge regions will be served simultaneously by the three BSs, i.e., $K = 1$ and $L = 3$ in (43); while when the robust MCBF design (7) is applied, each BS will serve the two MSs located inside its cell region (the hexagon), i.e., $K = 2$ in (7). Fig. 7 shows the performance comparison results of the robust

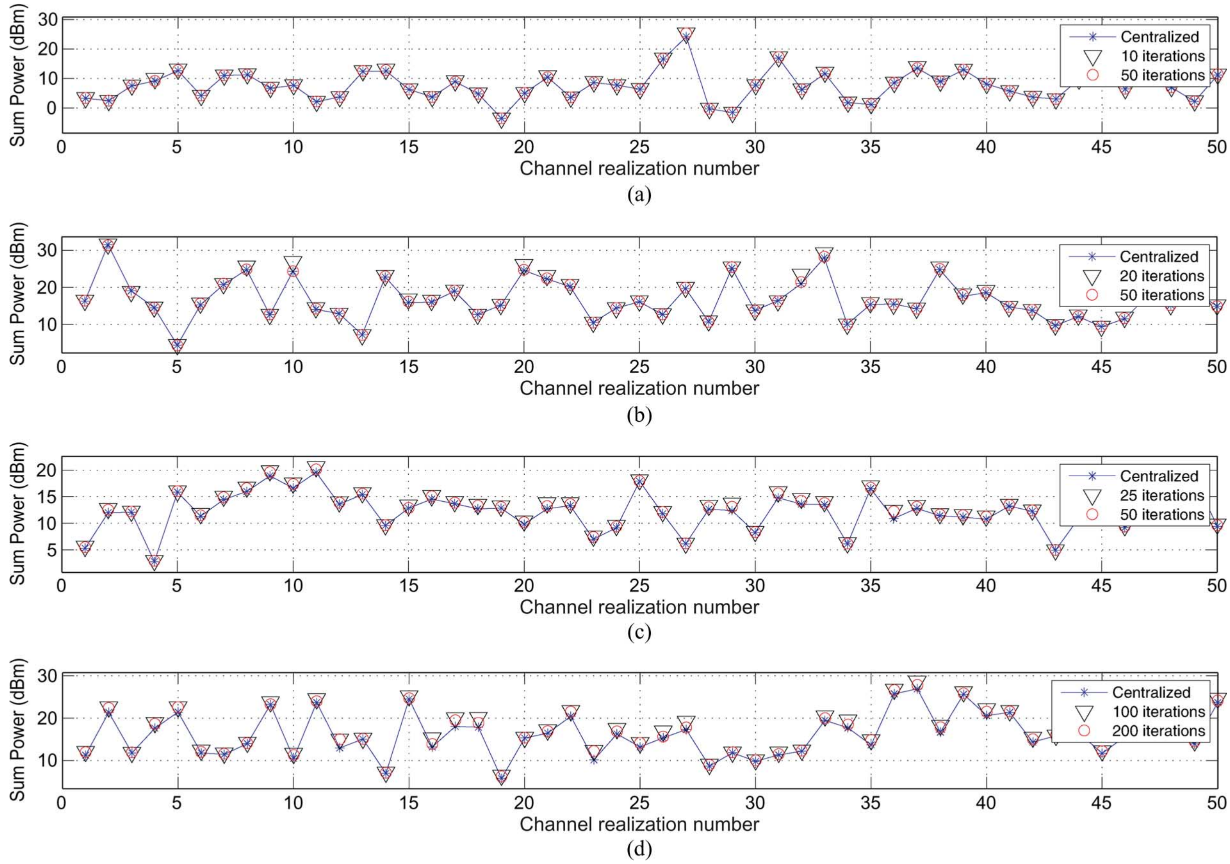


Fig. 4. Sum power comparison between the centralized robust MCBF solution and the distributed robust MCBF solution (by Algorithm 2) over 50 randomly generated channel realizations, for $\gamma = 10$ dB and $\varepsilon = 0.05$. (a) $N_c = 2, K = 2, N_t = 8$; (b) $N_c = 2, K = 4, N_t = 8$; (c) $N_c = 3, K = 3, N_t = 8$; (d) $N_c = 8, K = 1, N_t = 4$.

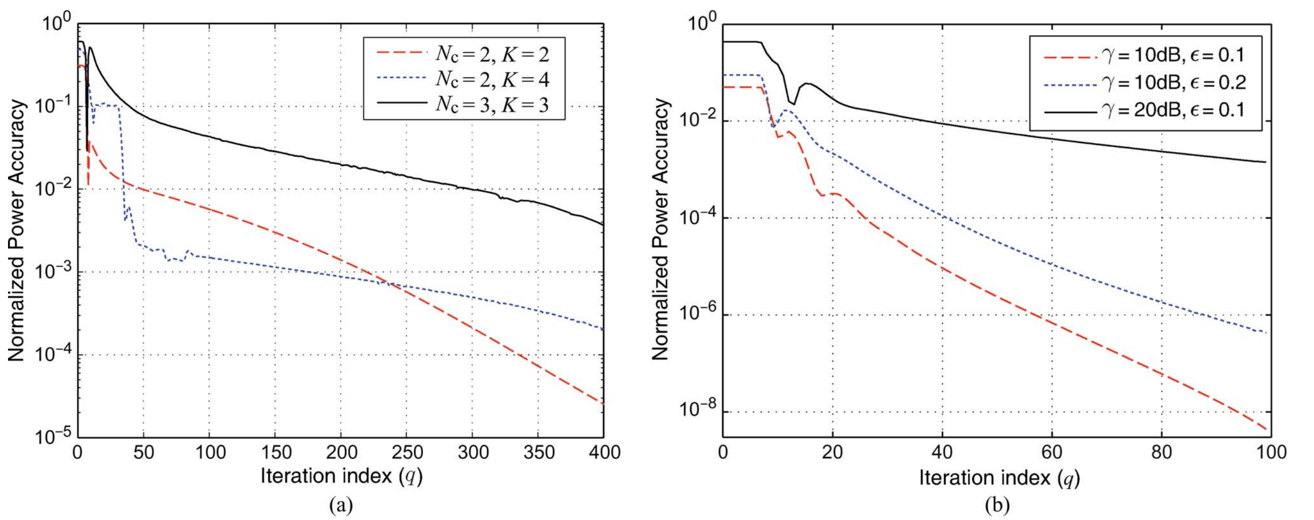


Fig. 5. Typical convergence curves of Algorithm 2 under various simulation settings. (a) $N_t = 8, \gamma = 10$ dB, $\varepsilon = 0.05$. (b) $N_c = 2, K = 2, N_t = 4$.

fully coordinated BF design (43) and the robust MCBF design (7) by testing over 17 000 channel realizations. The SDR formulation (44) is used as an approximation to (43). It is found in the simulation that SDR formulation (44) always yields rank-one solutions; hence the obtained solution is exactly the optimal solution of (43) for the tested problem instances. From Fig. 7, it can be observed that the robust fully coordinated (with cell-edge users simultaneously served by all the coordinated BSs) BF design is more feasible and more power efficient (by about 3 dB) than the robust MCBF design.

VII. Conclusion

We have presented an efficient approximation method based on SDR, formulated as (17), for the worst-case SINR constrained robust MCBF design problem [in (7)]. We have shown that when there is only one MS in each cell or when the CSI

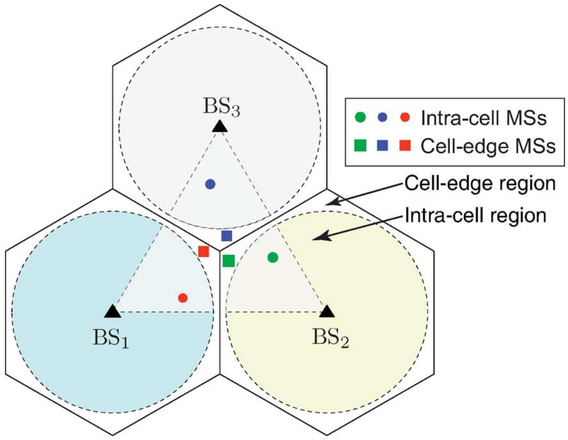


Fig. 6. Illustration of the simulation scenario for the robust fully coordinated BF design (43).

errors are sufficiently small, the proposed SDR method can yield the global optimal solution to the original problem, i.e., (7) (Proposition 1). Moreover, by using ADMM, we have presented a distributed robust MCBF algorithm (Algorithm 2). The proposed distributed algorithm is appealing because it is proven to converge to the global optimum of the centralized problem (Proposition 2). Extension of the proposed SDR method to the fully coordinated scenario has been presented as well. The presented simulation results have shown that the proposed SDR method is more power efficient than the existing method reported in [26], and that the proposed distributed optimization algorithm can obtain beamforming solutions with the normalized power accuracy smaller than 0.1 in several tens of iterations for the typical scenario of $N_c \leq 3$. As a future research, it is also possible to apply other variations of ADMM to distributed MCBF design for faster convergence and smaller backhaul signaling overhead.

Appendix A
Proof of Proposition 1

Proof of Case C1): Let us rewrite (17) for $K=1$ as follows:

$$\min_{\{\mathbf{W}_n\},\{\lambda_{mn}\},\{t_{mn}\}_{m\neq n}} \sum_{n=1}^{N_c} \alpha_n \mathrm{Tr}(\mathbf{W}_n) \tag{A.1a}$$
$$\text{s.t.}\quad \mathbf{\Phi}_n\left(\mathbf{W}_n, \{t_{mn}\}_{m\neq n}, \lambda_{nn}\right) \succeq \mathbf{0} \quad \forall n, \tag{A.1b}$$
$$\mathbf{\Psi}_{mn}(\mathbf{W}_m, t_{mn}, \lambda_{mn}) \succeq \mathbf{0} \quad \forall n, m\neq n, \tag{A.1c}$$
$$\mathbf{W}_n \succeq \mathbf{0} \quad \forall n, \tag{A.1d}$$
$$\lambda_{mn} \geq 0 \quad \forall m,n, \tag{A.1e}$$

where the subindices k and i in (17) have been removed for notational simplicity. Proposition 1 can be proved by investigating the KKT conditions of (A.1). Specifically, let $\{\mathbf{Y}^\star_{nn} \succeq \mathbf{0}\}$, $\{\mathbf{Y}^\star_{mn} \succeq \mathbf{0}\}$ and $\{\mathbf{Z}^\star_n \succeq \mathbf{0}\}$ be the optimal dual variables associated with (A.1b), (A.1c), and (A.1d), respectively. Moreover, let

$$\mathbf{Y}^\star_{nn} \triangleq \begin{bmatrix} \mathbf{A}_n & \mathbf{b}_n \\ \mathbf{b}_n^H & c_n \end{bmatrix} \succeq \mathbf{0}. \tag{A.2}$$

According to the KKT conditions of (A.1), one can verify that the optimal $\{\mathbf{W}^\star_n\}$, $\{\lambda^\star_{mn}\}$, $\{t^\star_{mn}\}$ satisfy the following conditions:

$$\mathbf{Z}^\star_n \mathbf{W}^\star_n = \mathbf{0}, \quad \mathbf{W}^\star_n \neq \mathbf{0}, \tag{A.3}$$
$$\mathbf{Z}^\star_n = \mathbf{I}_{N_t} + \sum_{m\neq n}^{N_c} [\mathbf{I}_{N_t}\ \hat{\mathbf{h}}_{nm}] \mathbf{Y}^\star_{nm} \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}^H_{nm} \end{bmatrix} - \frac{1}{\gamma_n} [\mathbf{I}_{N_t}\ \hat{\mathbf{h}}_{nn}] \mathbf{Y}^\star_{nn} \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}^H_{nn} \end{bmatrix} \succeq \mathbf{0}, \tag{A.4}$$
$$\mathbf{\Phi}_n\left(\mathbf{W}^\star_n, \{t^\star_{mn}\}_{m\neq n}, \lambda^\star_{nn}\right) \mathbf{Y}^\star_{nn} = \mathbf{0}, \mathbf{Y}^\star_{nn} \neq \mathbf{0}, \tag{A.5}$$
$$\mathrm{Tr}(\mathbf{Q}_{nn}\mathbf{A}_n) \leq c_n, \tag{A.6}$$
$$t^\star_{mn} > 0 \quad \forall m \neq n, \tag{A.7}$$
$$\lambda^\star_{nn} > 0, \tag{A.8}$$

for $n \in \mathcal{N}_c$. Specifically, to show (A.8), we note that if $\lambda^\star_{nn} = 0$, then it follows from (15) that:

$$\begin{bmatrix} -\hat{\mathbf{h}}^H_{nn} & 1 \end{bmatrix} \mathbf{\Phi}_n\left(\mathbf{W}^\star_n, \{t^\star_{mn}\}_{m\neq n}, \lambda^\star_{nn}\right) \begin{bmatrix} -\hat{\mathbf{h}}_{nn} \\ 1 \end{bmatrix} = -\sum_{m\neq n} t^\star_{mn} - \sigma_n^2 < 0, \tag{A.9}$$

which contradicts with (A.1b).

The first step of the proof is to show that $\mathbf{W}^\star_n$ has rank one whenever $\mathbf{Y}^\star_{nn}$ has rank one. Suppose that $\mathbf{Y}^\star_{nn}$ is of rank one and that $\mathbf{Y}^\star_{nn} = \mathbf{y}\mathbf{y}^H$ where $\mathbf{y} \in \mathbb{C}^{N_t+1}$. Let $\mathbf{X}_n \triangleq \mathbf{X}_n^{\frac{1}{2}}\mathbf{X}_n^{\frac{1}{2}} = \mathbf{I}_{N_t} + \sum_{m\neq n}^{N_c} [\mathbf{I}_{N_t}\ \hat{\mathbf{h}}_{nm}] \mathbf{Y}^\star_{nm} \left[\begin{smallmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}^H_{nm} \end{smallmatrix}\right] \succ \mathbf{0}$. Then by (A.4)

$$\begin{aligned} &\mathrm{Rank}\left(\mathbf{Z}^\star_n\right) \\ &= \mathrm{Rank}\left(\mathbf{X}_n^{1/2}\mathbf{X}_n^{1/2} - \gamma_n^{-1}\left[\mathbf{I}_{N_t}\ \hat{\mathbf{h}}_{nn}\right] \mathbf{y}\mathbf{y}^H \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}^H_{nn} \end{bmatrix}\right) \\ &= \mathrm{Rank}\left(\mathbf{I}_{N_t} - \gamma_n^{-1}\mathbf{X}_n^{-1/2}\left[\mathbf{I}_{N_t}\ \hat{\mathbf{h}}_{nn}\right] \mathbf{y}\mathbf{y}^H \begin{bmatrix} \mathbf{I}_{N_t} \\ \hat{\mathbf{h}}^H_{nn} \end{bmatrix} \mathbf{X}_n^{-1/2}\right) \\ &\geq N_t - 1. \end{aligned} \tag{A.10}$$

It follows from (A.3) and (A.10) that $\mathbf{W}^\star_n$ must be of rank one due to

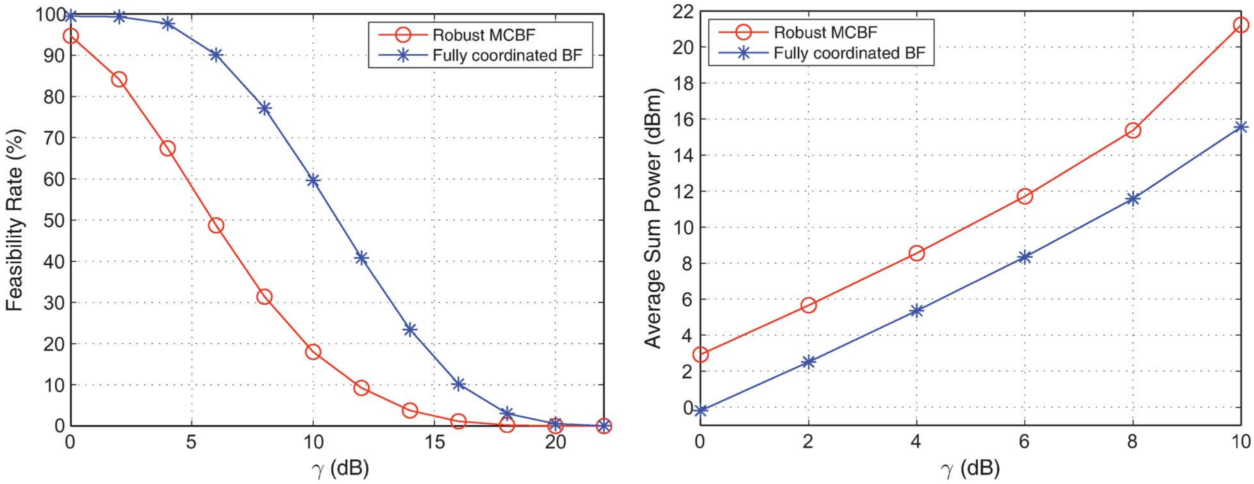
$$0 < \mathrm{Rank}\left(\mathbf{W}^\star_n\right) \leq \mathrm{Nullity}\left(\mathbf{Z}^\star_n\right) = N_t - \mathrm{Rank}\left(\mathbf{Z}^\star_n\right) \leq 1. \tag{A.11}$$

What remains is to show that $\mathbf{Y}^\star_{nn}$ is indeed of rank one. First, one can show that $c_n > 0$ since, if not, by (A.2), (A.6), and the fact of $\mathbf{Q}_{nn} \succ \mathbf{0}$, we must have $\mathbf{Y}^\star_{nn} = \mathbf{0}$, which, however, leads to $\mathbf{Z}^\star_n \succ \mathbf{0}$ by (A.4) and thus $\mathbf{W}^\star_n = \mathbf{0}$ by (A.11), and consequently contradicts with (A.3). By substituting (A.2) into (A.5), where $\mathbf{\Phi}_n$ is given by (15), it is not difficult to verify that the following two equalities hold:

$$\left(\gamma_n^{-1}\mathbf{W}^\star_n + \lambda^\star_{nn}\mathbf{Q}_{nn}\right)\mathbf{A}_n + \gamma_n^{-1}\mathbf{W}^\star_n\hat{\mathbf{h}}_{nn}\mathbf{b}_n^H = \mathbf{0}, \tag{A.12}$$
$$\left(\gamma_n^{-1}\mathbf{W}^\star_n + \lambda^\star_{nn}\mathbf{Q}_{nn}\right)\mathbf{b}_n + \gamma_n^{-1}\mathbf{W}^\star_n\hat{\mathbf{h}}_{nn}c_n = \mathbf{0}. \tag{A.13}$$

By postmultiplying (A.13) with $-\frac{\mathbf{b}_n^H}{c_n}$, and adding the resultant equality to (A.12), one can obtain

$$\left(\gamma_n^{-1}\mathbf{W}^\star_n + \lambda^\star_{nn}\mathbf{Q}_{nn}\right)\left(\mathbf{A}_n - \mathbf{b}_n\mathbf{b}_n^H/c_n\right) = \mathbf{0}. \tag{A.14}$$



(a) (b)

Fig. 7. Performance comparison results of the robust MCBF design (7) and the fully coordinated BF design (43), for $N_c = 3$, $K = 2$, $N_t = 4$, and $\varepsilon = 0.1$. (a) Feasibility rate versus SINR requirement. (b) Average sum power versus SINR requirement.

Since $\left(\frac{1}{\gamma_n}\mathbf{W}_n^\star + \lambda_{nn}^\star \mathbf{Q}_{nn}\right) \succ \mathbf{0}$ due to both $\lambda_{nn}^\star > 0$ and $\mathbf{Q}_{nn} \succ \mathbf{0}$, (A.14) implies that $\mathbf{A}_n = \frac{\mathbf{b}_n \mathbf{b}_n^H}{c_n}$, and thus (A.2) reduces to

$$\mathbf{Y}_{nn}^\star = \begin{bmatrix} \mathbf{b}_n \mathbf{b}_n^H / c_n & \mathbf{b}_n \\ \mathbf{b}_n^H & c_n \end{bmatrix} = \begin{bmatrix} \mathbf{b}_n/\sqrt{c_n} \\ \sqrt{c_n} \end{bmatrix} \left[\mathbf{b}_n^H/\sqrt{c_n} \quad \sqrt{c_n} \right] \tag{A.15}$$

which is a rank-one matrix. Case C1) is thus proved.

Proof of Case C2): Case C2) can be proved following similar derivations in (A.10) and (A.11), but using the KKT conditions of (17) with $\mathbf{Q}_{nnk} = \infty \mathbf{I}_{N_t}$ (i.e., $\mathbf{e}_{nnk} = \mathbf{0}$) for all n, k. ∎

Appendix B
An Alternative Distributed Robust MCBF Algorithm

In the appendix, we present an alternative distributed robust MCBF algorithm which has a reduced amount of point-to-point information exchange between BSs. Note that the SDR problem (16) can be expressed as

$$\min_{\substack{\{\mathbf{W}_{nk} \succeq \mathbf{0}\}, \{p_n\}, \\ \{\lambda_{mnk} \geq 0\}, \{t_{mnk}\}}} \sum_{n=1}^{N_c} \alpha_n p_n \tag{B.1a}$$

$$\text{s.t.} \quad \mathbf{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, \{t_{mnk}\}_m, \lambda_{nnk}\right) \succeq \mathbf{0}, \quad \forall n \in \mathcal{N}_c, k \in \mathcal{K}, \tag{B.1b}$$

$$\mathbf{\Psi}_{nmk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, t_{nmk}, \lambda_{nmk}\right) \succeq \mathbf{0}, \quad \forall n \in \mathcal{N}_c, m \in \mathcal{N}_c \setminus \{n\}, k \in \mathcal{K}, \tag{B.1c}$$

$$\sum_{k=1}^K \mathrm{Tr}(\mathbf{W}_{nk}) = p_n \quad \forall n \in \mathcal{N}_c. \tag{B.1d}$$

Different from (26b), we follow the idea as in [10] to define the local ICI variable vector $\mathbf{t}_n$ as follows:

$$\mathbf{t}_n = \left[\left[t_{1n1}^{(n)}, \ldots, t_{1nK}^{(n)}\right], \ldots, \left[t_{N_c n1}^{(n)}, \ldots, t_{N_c nK}^{(n)}\right], \left[t_{n11}^{(n)}, \ldots, t_{n1K}^{(n)}\right], \left[t_{n21}^{(n)}, \ldots, t_{nN_cK}^{(n)}\right]\right]^T \in \mathbb{R}^{2(N_c-1)K} \tag{B.2}$$

where the superscript (n) indicates that $\{t_{mnk}^{(n)}\}_{mk}$ and $\{t_{nmk}^{(n)}\}_{mk}$ are the local incoming and outgoing worst-case ICI variables maintained by BS_n, for $n = 1, \ldots, N_c$. Then there exists a linear mapping matrix $\tilde{\mathbf{E}}_n \in \{0,1\}^{2(N_c-1)K \times N_c(N_c-1)K}$ such that

$$\mathbf{t}_n = \tilde{\mathbf{E}}_n \mathbf{t} \tag{B.3}$$

for all $n \in \mathcal{N}_c$, where $\mathbf{t}$ is as defined in (26a). Equation (B.3) ensures $t_{nmk}^{(n)} = t_{nmk}^{(m)} = t_{nmk}$ for all n, m, k. We therefore can rewrite (B.1) in the following compact form

$$\min_{\{\mathbf{W}_{nk}\}, \{\lambda_{nmk}\}, \{\mathbf{t}_n\}, \{p_n\}, \mathbf{t}} \sum_{n=1}^{N_c} \alpha_n p_n \tag{B.4a}$$

$$\text{s.t.} \quad \left(\{\mathbf{W}_{nk}\}_k, \{\lambda_{nmk}\}_{m,k}, \mathbf{t}_n, p_n\right) \in \mathcal{C}_n, \tag{B.4b}$$

$$\mathbf{t}_n = \tilde{\mathbf{E}}_n \mathbf{t}, \quad \forall n \in \mathcal{N}_c \tag{B.4c}$$

where

$$\begin{aligned} \mathcal{C}_n = \Bigg\{ & \left(\{\mathbf{W}_{nk}\}_k, \{\lambda_{nmk}\}_{m,k}, \mathbf{t}_n, p_n\right) \mid \\ & \mathbf{\Phi}_{nk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, \left\{t_{mnk}^{(n)}\right\}_{m \neq n}, \lambda_{nnk}\right) \succeq \mathbf{0} \ \forall k \in \mathcal{K}, \\ & \mathbf{\Psi}_{nmk}\left(\{\mathbf{W}_{ni}\}_{i=1}^K, t_{nmk}^{(n)}, \lambda_{nmk}\right) \succeq \mathbf{0} \ \forall m \in \mathcal{N}_c \setminus \{n\}, \\ & k \in \mathcal{K}, \lambda_{nmk} \geq 0 \ \forall m \in \mathcal{N}_c, k \in \mathcal{K}, \\ & \sum_{k=1}^K \mathrm{Tr}(\mathbf{W}_{nk}) = p_n, \mathbf{W}_{nk} \succeq \mathbf{0} \ \forall k \in \mathcal{K}, \\ & t_{mnk}^{(n)} \geq 0, t_{nmk}^{(n)} \geq 0 \ \forall m \in \mathcal{N}_c \setminus \{n\}, k \in \mathcal{K} \Bigg\}. \end{aligned}$$

Note that (B.4) has exactly the same form as (27). Therefore, by applying ADMM to (B.4), one can obtain the same ADMM steps as in (32)to (36) for (B.4), except $\mathbf{t}_n$ now defined in (B.2) and $\mathbf{E}_n$ replaced by $\tilde{\mathbf{E}}_n$ instead. We show next that the two differences can result in a quite different algorithm which has a reduced amount of signaling overhead for point-to-point information exchange between BSs.

One should note that for updating (35a), each BS_n in fact does not need to compute the whole $\mathbf{t}(q+1)$; instead, BS_n only needs $\tilde{\mathbf{E}}_n\mathbf{t}(q+1)$, which is composed of $\{t_{mnk}(q+1)\}_{mk}$ and $\{t_{nmk}(q+1)\}_{mk}$. Similarly, to compute (36a), each BS_n also only needs $\tilde{\mathbf{E}}_n\mathbf{t}(q)$. The interesting part is that, by exploiting the structure of $\tilde{\mathbf{E}}_n$, one can show that BS_n can actually learn $\{t_{mnk}(q+1)\}_k$ and $\{t_{nmk}(q+1)\}_k$ by communicating only with BS_m. Specifically, one can show that (35a) has an explicit structure of

$$t_{nmk}(q+1)=\frac{1}{2}\left(t_{nmk}^{(n)}(q+1)-\frac{1}{c}\nu_{nmk}^{(n)}(q)\right)+\frac{1}{2}\left(t_{nmk}^{(m)}(q+1)-\nu_{nmk}^{(m)}(q)\right), \quad \text{(B.5a)}$$

$$t_{mnk}(q+1)=\frac{1}{2}\left(t_{mnk}^{(n)}(q+1)-\frac{1}{c}\nu_{mnk}^{(n)}(q)\right)+\frac{1}{2}\left(t_{mnk}^{(m)}(q+1)-\nu_{mnk}^{(m)}(q)\right) \quad \text{(B.5b)}$$

for all n,m,k. From (B.5), we can see that BS_m only needs to send the $2K$ real values of $\{t_{nmk}^{(m)}(q+1)-\nu_{nmk}^{(m)}(q)\}_k$ and $\{t_{mnk}^{(m)}(q+1)-\nu_{mnk}^{(m)}(q)\}_k$ to BS_n. In summary, each BS totally needs to send out $2(N_c-1)K$ real values in each iteration, which is more backhaul efficient than Algorithm 2 ($(N_c-1)N_cK$ real values as stated in Remark 2) when $N_c>2$.

REFERENCES

[1] C. Shen, K.-Y. Wang, T.-H. Chang, Z. Qiu, and C.-Y. Chi, "Worst-case SINR constrained robust coordinated beamforming for multicell wireless systems," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 5–9, 2011, pp. 1–5.

[2] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multicell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[3] R. Irmer, V. H. Droste, D. Telekom, P. Marsch, M. Grieger, and G. Fettweis, "Coordinate multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[4] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1999.

[5] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*. Boca Raton, FL: CRC, 2002, ch. 18.

[6] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[7] L. Venturino, N. Prasad, and X.-D. Wang, "Coordinated linear beamforming in downlink multicell wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451–1461, Apr. 2010.

[8] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimality properties and low-complexity solutions to coordinated multicell transmission," in *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 6–10, 2010, pp. 1–6.

[9] D. H. N. Nguyen and T. Le-Ngoc, "Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 3326–3338, Jul. 2011.

[10] A. Tölli, H. Pennanen, and P. Komulanen, "Decentralized minimum power multicell beamforming with limited backhaul signaling," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 570–580, Feb. 2011.

[11] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.

[12] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, "Distributed multicell beamforming with limited intercell coordination," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 728–738, Feb. 2011.

[13] H. Pennanen, A. Tölli, and M. Latva-aho, "Decentralized coordinated downlink beamforming via primal decomposition," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 647–650, Nov. 2011.

[14] S. Sim, S.-J. Han, J.-S. Park, and S.-C. Lee, "Seamless IP mobility support for flat architecture mobile WiMAX networks," *IEEE Commun. Mag.*, vol. 47, no. 6, pp. 142–148, Jun. 2009.

[15] J. Qiu, R. Zhang, Z.-Q. Luo, and S. Cui, "Optimal distributed beamforming for MISO interference channels," in *Proc. ASILOMAR*, Pacific Grove, CA, Nov. 7–10, 2010, pp. 277–281.

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

[17] S. Boyd, L. Xiao, A. Mutapcic, and J. Mattingley, Notes on Decomposition Methods [Online]. Available: http://see.stanford.edu/materials/lsocoee364b/08-decomposition_notes.pdf

[18] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[19] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[20] M. Shenouda and T. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 714–724, Dec. 2007.

[21] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–15, June 2008, Article ID 609028.

[22] M. B. Shenouda and T. N. Davidson, "Outage-based designs for multi-user transceivers," in *Proc. IEEE ICASSP 2009*, Taipei, Taiwan, April 19–24, 2009, pp. 2389–2392.

[23] K.-Y. Wang, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "A semidefinite relaxation based conservative approach to robust transmit beamforming with probabilistic SINR constraints," in *Proc. EUSIPCO*, Aalborg, Denmark, Aug. 23–27, 2010, pp. 407–411.

[24] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A. M.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: A Bernstein-type inequality based conservative approach," in *Proc. IEEE ICASSP*, Prague, Czech, May 22–27, 2011, pp. 3080–3083.

[25] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.* vol. 60, no. 5, pp. 2508–2523, May 2012.

[26] A. Tajer, N. Prasad, and X.-D. Wang, "Robust linear precoder design for multicell downlink transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[27] 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.

[28] Z.-Q. Luo and T.-H. Chang, "SDP relaxation of homogeneous quadratic optimization: Approximation bounds and applications," in *Convex Optimization in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ., 2010, ch. 4.

[29] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[31] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ: Prentice-Hall, 1989.

[32] 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 in Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2008.

[33] D. P. Bertsekas and J. N. Tsitsiklis, *Nonlinear Programming: 2nd Edition*. Belmont, MA: Athena Scientific, 1999.

[34] X. Zhang, M. Burger, and S. Osher, "A unified primal-dual algorithm framework based on Bregman iteration," *J. Sci. Comput.*, vol. 46, pp. 20–46, 2011.

[35] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Ann. Allerton Conf. Commun., Contr., Comput.*, Sep. 1999, pp. 987–996.

[36] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[37] J. Sturm, SeDuMi: Version 1.1 [Online]. Available: http://sedumi.ie.lehigh.edu/, Oct. 2004

[38] E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?," in *Proc. IEEE ICASSP*, Progue, Czech, May 22–27, 2011, pp. 3096–2099.

[39] S. Vorobyov, A. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[40] T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Worst-case robust multiuser transmit beamforming using semidefinite relaxation: Duality and implications," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Glove, CA, Nov. 6–9, 2011, pp. 1–5.

[41] B. Williams and T. Camp, "Comparison of broadcasting techniques for mobile ad hoc networks," in *Proc. 3rd ACM Int. Symp. on Mobile Ad Hoc Netw. Comput.*, Lausanne, Switzerland, Jun. 9–11, 2002, pp. 1–16.

[42] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Image Sci.*, vol. 2, pp. 183–202, Mar. 2009.

[43] H. Zhang, N. B. Mehta, A. F. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 155–165, Jan. 2008.

[44] Technical Specification Group Radio Access Network; Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA), 3GPP TR 25.814, v.7.1.0, Sep. 2006 [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25814.htm

[45] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MISO transmitter optimization for intercell interference mitigation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4272–4285, Aug. 2010.

Chao Shen received the B.S. degree in communication engineering and the M.S. degree in digital signal processing from the Beijing Jiaotong University, China, in 2003 and 2006, respectively.

He is currently working toward the Ph.D. degree in the field of the future wireless communication at the same university. His research interests include wireless communication and digital signal processing.

Tsung-Hui Chang (S'07–M'08) 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 2003 and 2008, respectively.

He was an exchange Ph.D. student of University of Minnesota, Minneapolis, MN, a visiting scholar of The Chinese University of Hong Kong, Hong Kong, and a Postdoctoral Researcher with the Institute of Communications Engineering, NTHU. He currently works as a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, University of California, Davis. His research interests are widely in wireless communications, digital signal processing, and convex optimization and its applications.

Kun-Yu Wang received the B.S. degree in electrical engineering from the Yuan Ze University, Taoyuan, Taiwan, in 2008.

Currently, he is pursuing the Ph.D. degree in communications engineering at the National Tsing Hua University, Hsinchu, Taiwan. His research interests are in wireless communications and signal processing.

Zhengding Qiu received the B.Sc. degree in communication engineering and the M.Sc. degree in digital signal processing from Northern Jiaotong University, Beijing, China, in 1967 and 1981, respectively.

Currently, he is a Professor with the Institute of Information Science at Beijing Jiaotong University, Beijing, China. His main research includes digital signal processing, biometrics, digital watermarking, multimedia communication, and Internet technology.

Prof. Qiu is a Fellow with the Chinese Institute of Communication and the Chinese Institute of Electronics, respectively.

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, 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 during 2002–2005), National Tsing Hua University, Hsinchu, Taiwan. He has published more than 190 technical papers, including more than 65 journal papers (mostly in IEEE TRANSACTIONS ON SIGNAL PROCESSING), two book chapters, and more than 120 peer-reviewed conference papers, as well as a graduate-level textbook, *Blind Equalization and System Identification* (New York: Springer-Verlag, 2006). His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, biomedical, and hyperspectral image analysis.

Dr. Chi has been a Technical Program Committee member for many IEEE sponsored and cosponsored workshops, symposiums, and conferences on signal processing and wireless communications, including Co-Organizer and General Co-chairman of 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), and Co-Chair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008, and Lead Co-Chair of SPC Symposium, ChinaCOM 2009. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (May 2001–April 2006), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (January 2006–December 2007), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (January 2008–December 2009), Associate Editor of the IEEE SIGNAL PROCESSING LETTERS (June 2006–May 2010), and a member of the Editorial Board of the EURASIP *Signal Processing Journal* (June 2005–May 2008), and an editor (July 2003–Decemer 2005) as well as a Guest Editor (2006) of the EURASIP *Journal on Applied Signal Processing*. He was a member of IEEE Signal Processing Committee on Signal Processing Theory and Methods (2005–2010). Currently, he is a member of IEEE Signal Processing Committee on Signal Processing for Communications and Networking.