Coordinated Beamforming for Multiuser MISO Interference Channel Under Rate Outage Constraints

Wei-Chiang Li, Tsung-Hui Chang, *Member, IEEE*, Che Lin, *Member, IEEE*, and Chong-Yung Chi, *Senior Member, IEEE*

Abstract—This paper studies the coordinated beamforming design problem for the multiple-input single-output (MISO) interference channel, assuming only channel distribution information (CDI) known to the transmitters. Under a given requirement on the rate outage probability for receivers, we aim to maximize the system utility (e.g., the weighted sum rate, weighted geometric mean rate, and the weighed harmonic mean rate) subject to the rate outage constraints and individual power constraints. The outage constraints, however, lead to a complicated, nonconvex structure for the considered beamforming design problem and render the optimization problem difficult to handle. Although this nonconvex optimization problem can be solved in an exhaustive search manner, this brute-force approach is only feasible when the number of transmitter-receiver pairs is small. For a system with a large number of transmitter-receiver pairs, computationally efficient alternatives are necessary. Hence, the focus of this paper is the design of such efficient approximation methods. In particular, by employing semidefinite relaxation (SDR) and first-order approximation techniques, we propose an efficient successive convex approximation (SCA) algorithm that provides high-quality approximate beamforming solutions via solving a sequence of convex approximation problems. The solution thus obtained is further shown to be a stationary point for the SDR of the original outage constrained beamforming design problem. Furthermore, we propose a distributed SCA algorithm where each transmitter optimizes its own beamformer using local CDI and information obtained from limited message exchange with the other transmitters. Our simulation results demonstrate that the proposed SCA algorithm and its distributed counterpart indeed converge, and promising performance can be achieved for all the considered system utilities.

Index Terms—Convex optimization, coordinated beamforming, interference channel, outage probability, semidefinite relaxation.

I. INTRODUCTION

I NTER-CELL interference is known to be one of the main bottlenecks that limit the system performance of a wireless cellular network where all transmitters share a universal frequency band. The performance degradation caused by such interference is severe especially for the users at the cell edge and can only be alleviated when some sort of cooperation is available between base stations (BSs) [2]. According to the level of cooperation, the coordinated transmission can be roughly divided into two classes: Network multiple-input multiple-output (MIMO) and interference coordination [3]. In network MIMO, all BSs work as a single virtual BS using all the available antennas for data transmission and reception. Each of the BSs requires to know all the channel state information (CSI) and data streams of users, demanding a large amount of message exchange between BSs [4]. Interference coordination, by contrast, only needs CSI sharing between BSs; based on the shared CSI, the BSs coordinate with each other in the design of transmission strategies, e.g., coordinated beamforming [5], [6] or power allocation [7]. Our interest in this paper lies in the coordinated beamforming design. To this end, we adopt the commonly used interference channel (IFC) model [8]–[10]. Under this model, a Pareto optimal transmission scheme is that the rate tuple of receivers resides on the boundary of the achievable rate region [11]. It is always desirable to have a Pareto optimal transmission scheme since, otherwise, the achievable rates of some of the receivers can be further improved.

Consider a multiple-input single-output (MISO) IFC, where the transmitters are equipped with multiple antennas while the receivers, i.e., mobile users, have only single antenna. We assume that the receivers employ single-user detection wherein the cross-link interference is treated as noise. Under such circumstance, analyses in [12]–[14] have shown that the Pareto optimal transmission strategy is transmit beamforming. While beamforming is a structurally simple transmission strategy, finding the optimal transmit beamformers for the MISO IFC is intrinsically difficult. More precisely, it has been proved [15] that finding the optimal beamformers that maximize system utilities, such as the weighted sum rate, the geometric mean rate, or the harmonic mean rate, is NP-hard in general. As a result, lots of efforts have focused on characterizing the optimal beamformer structures [12], [14], [16] in order to reduce the search dimension for finding the optimal beamforming vectors, or on investigating suboptimal but computationally efficient beamforming algorithms [15]–[17]. Another approach to studying these resource conflicts encountered in the IFC is to use Game theory; see [11], [18], [19] for related works.

The aforementioned beamforming designs all assume that the transmitters have the complete knowledge of CSI. To provide the transmitters with complete CSI, the receivers need to periodically send the CSI (e.g., for frequency division duplexing systems) or training signals (e.g., for time division duplexing systems) back to the transmitters. In contrast to the CSI, channel distribution information (CDI) can remain

unchanged over a relatively long period of time and thus the amount of feedback information can be significantly reduced. With CDI at the transmitters, the ergodic rate region of the K-user MISO IFC has been analyzed and the structure of the Pareto optimal beamformers has been characterized in [20]. For a two-user case, an efficient algorithm for finding the Pareto boundary of the ergodic rate region was presented in [21]. Unlike the ergodic achievable rate where the packet delay is not taken into consideration, the outage constrained achievable rate is more suitable for delay-sensitive applications, such as those involving voice or video data communications. For such outage constrained achievable rate region, the authors of [22], [23] presented a numerical method for finding the Pareto boundary; however, the complexity of this algorithm increases exponentially with the number of transmitter-receiver pairs. Developing efficient beamforming design algorithms that can approach the outage constrained Pareto boundary is therefore important. While several efficient beamforming algorithms can be found in [24], [25], a different power-minimization design criterion was considered, instead of rate utility maximization.

In this paper, we investigate efficient coordinated beamforming design algorithms for maximizing the system utility under rate outage constraints and individual power constraints. Specifically, we assume that the MISO channel between each transmitter and receiver is composed of zero-mean circularly symmetric complex Gaussian fading coefficients where the corresponding covariance matrix is known to the transmitter. We formulate an outage constrained coordinated beamforming design problem, aiming at finding the Pareto optimal beamformers that maximize the system utility (e.g., the weighted sum rate) subject to a pre-assigned rate outage probability requirement and power constraints. Due to the complicated nonconvex outage constraints, we propose a successive convex approximation (SCA) algorithm, where the original problem is successively approximated by a convex problem and the beamforming solution is refined in an iterative manner. The convex approximation formulation is obtained by applying the convex optimization based semidefinite relaxation (SDR) technique [26], followed by a logarithmic change of variables and first-order approximation techniques. We analytically show that the proposed SCA algorithm can yield a beamforming solution that is a stationary point for the SDR of the original problem. We further propose a round-robin-fashioned (block-coordinate) distributed SCA algorithm where each transmitter optimizes only its beamformer using local CDI with limited communication overhead of message exchange with the other transmitters. It is shown by simulations that the two proposed algorithms yield near-optimal performance with lower complexity compared with those reported in [22], [23].

The remaining part of this paper is organized as follows. The system model and the outage constrained coordinated beamforming problem are presented in Section II. In Section III, we present the proposed SCA algorithm and analyze its convergence property. In Section IV, the distributed SCA algorithm is developed and analyzed. Simulation results that demonstrate the efficacy of the proposed algorithms are presented in Section V. Finally, the conclusions are drawn in Section VI.

Notation: The n-dimensional complex vectors and complex Hermitian matrices are denoted by $\mathbb{C}^n$ and $\mathbb{H}^n$, respectively. The $n \times n$ identity matrix is denoted by $\mathbf{I}_n$. The superscripts

'T' and 'H' represent the matrix transpose and conjugate transpose, respectively. We denote $\|\cdot\|$ as the vector Euclidean norm. $\mathbf{A} \succeq \mathbf{0}$ and $\mathbf{a} \succeq \mathbf{0}$ respectively mean that matrix $\mathbf{A}$ is positive semidefinite (PSD) and vector $\mathbf{a}$ is element-wise nonnegative. The trace and rank of matrix $\mathbf{A}$ are denoted as $\mathrm{Tr}(\mathbf{A})$ and $\mathrm{rank}(\mathbf{A})$, respectively. We use the expression $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{Q})$ if $\mathbf{x}$ is circularly symmetric complex Gaussian distributed with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{Q}$. We denote $\exp(\cdot)$ (or simply $e^{(\cdot)}$) as the exponential function, while $\ln(\cdot)$ and $\Pr\{\cdot\}$ represent the natural log function and the probability function, respectively. For a variable a_{ik}, where $i, k \in \{1, \dots, K\}$, $\{a_{ik}\}_k$ denotes the set $\{a_{i1}, \dots, a_{iK}\}$, $\{a_{ik}\}_{k \neq i}$ denotes the set $\{a_{ik}\}_k$ excluding a_{ii}, and $\{a_{ik}\}$ is defined as the set containing all possible a_{ik}, i.e., $\{a_{11}, \dots, a_{1K}, a_{21}, \dots, a_{KK}\}$.

II. SIGNAL MODEL AND PROBLEM STATEMENT

We consider the K-user MISO IFC where each transmitter is equipped with N_t antennas and each receiver with a single antenna. It is assumed that transmitters employ transmit beamforming to communicate with their respective receivers. Let $s_i(t)$ denote the information signal sent from transmitter i, and let $\boldsymbol{w}_i \in \mathbb{C}^{N_t}$ be the corresponding beamforming vector. The received signal at receiver i is given by

$$x_i(t) = \boldsymbol{h}_{ii}^H \boldsymbol{w}_i s_i(t) + \sum_{k=1, k \neq i}^{K} \boldsymbol{h}_{ki}^H \boldsymbol{w}_k s_k(t) + n_i(t), \quad (1)$$

where $\boldsymbol{h}_{ki} \in \mathbb{C}^{N_t}$ denotes the channel vector from transmitter k to receiver i, and $n_i(t) \sim \mathcal{CN}(0, \sigma_i^2)$ is the additive white Gaussian noise at receiver i where $\sigma_i^2 > 0$ is the noise variance. As can be seen from (1), in addition to the noise, each receiver suffers from the cross-link interference $\sum_{k \neq i} \boldsymbol{h}_{ki}^H \boldsymbol{w}_k s_k(t)$. We assume that all receivers employ single-user detection where the cross-link interference is simply treated as background noise. Under Gaussian signaling, i.e., $s_i(t) \sim \mathcal{CN}(0, 1)$, the instantaneous achievable rate of the ith transmitter-receiver pair is known to be

$$r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_k\}) = \log_2\left(1 + \frac{\left|\boldsymbol{h}_{ii}^H \boldsymbol{w}_i\right|^2}{\sum_{k \neq i} \left|\boldsymbol{h}_{ki}^H \boldsymbol{w}_k\right|^2 + \sigma_i^2}\right).$$

In this paper, we assume that the channel coefficients $\boldsymbol{h}_{ki}$ are block-faded (i.e., quasi-static), and that the transmitters have only the statistical information of the channels, i.e., the CDI. In particular, it is assumed that $\boldsymbol{h}_{ki} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}_{ki})$ for all $k, i = 1, \dots, K$, where $\boldsymbol{Q}_{ki} \succeq \mathbf{0}$ denotes the channel covariance matrix and is known to all the transmitters. Since the transmission rate R_i cannot be adapted without CSI, the communication would be in outage whenever the transmission rate $R_i > 0$ is higher than the instantaneous capacity that the channel can support. For a given outage probability requirement $(\epsilon_1, \dots, \epsilon_K)$, the beamforming vectors $\{\boldsymbol{w}_i\}$ must satisfy $\Pr\{r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_i\}) < R_i\} \leq \epsilon_i$. Following [23], we define the corresponding ϵ_i-outage achievable rate region as follows.

Definition 1 [23]: *Let $P_i > 0$ denote the power constraint of transmitter i, for $i = 1, \dots, K$. The rate tuple $(R_1, \dots, R_K)$ is said to be achievable if $\Pr\{r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_k\}) < R_i\} \leq \epsilon_i$, $i = 1, \dots, K$, for some $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_K) \in \mathcal{W}_1 \times \cdots \times \mathcal{W}_K$ where $\epsilon_i \in (0, 1)$ is the maximum tolerable outage probability*

of receiver i and $\mathcal{W}_i \triangleq \{\boldsymbol{w} \in \mathbb{C}^{N_t} | \|\boldsymbol{w}\|^2 \leq P_i\}$. *The ϵ_i-outage achievable rate region is given by*

$$\mathcal{R} = \bigcup_{\substack{\boldsymbol{w}_i \in \mathcal{W}_i, \\ i=1,\ldots,K}} \left\{ (R_1,\ldots,R_K) \middle| \begin{array}{l} \Pr\{r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_k\}) < R_i\} \leq \epsilon_i \\ i = 1,\ldots,K \end{array} \right\}.$$

Given an outage requirement $(\epsilon_1,\ldots,\epsilon_K)$ and an individual power constraint $(P_1,\ldots,P_K)$, our goal is to optimize $\{\boldsymbol{w}_k\}$ such that the predefined system utility function $U(R_1,\ldots,R_K)$ is maximized. To this end, we consider the following outage constrained coordinated beamforming design problem

$$\max_{\substack{\boldsymbol{w}_i \in \mathbb{C}^{N_t}, R_i \geq 0, \\ i=1,\ldots,K}} U(R_1,\ldots,R_K) \tag{2a}$$

$$\text{s.t. } \Pr\{r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_k\}) < R_i\} \leq \epsilon_i, \tag{2b}$$

$$\|\boldsymbol{w}_i\|^2 \leq P_i, \quad i = 1,\ldots,K. \tag{2c}$$

Note that, as each user would prefer a higher transmission rate, a sensible system utility function should be strictly increasing with respect to the individual rate R_i for $i = 1,\ldots,K$, such that the optimal $(R_1,\ldots,R_K)$ of problem (2) would lie on the so-called Pareto boundary of $\mathcal{R}$ [11]. In this paper, we consider the following system utility function which captures a tradeoff between the system throughput and user fairness [27]

$$U_\beta(\{R_i\}) = \begin{cases} \sum_{i=1}^K \frac{\alpha_i R_i^{1-\beta}}{1-\beta}, & 0 \leq \beta < \infty, \beta \neq 1, \\ \sum_{i=1}^K \alpha_i \ln(R_i), & \beta = 1, \end{cases} \tag{3}$$

where the coefficients $\alpha_i \in [0,1]$ for $i = 1,\ldots,K$ with $\sum_{i=1}^K \alpha_i = 1$ represent the user priority, and the parameter $\beta \in [0,\infty)$ reflects the user fairness. For example, for β being 0, 1, 2 and infinity, $U_\beta(\{R_i\})$ corresponds to the weighted sum rate, weighted geometric mean rate, weighted harmonic mean rate and the weighted minimal rate, respectively. Hence, for β being 0, 1, 2 and infinity, maximizing $U_\beta(\{R_i\})$ is respectively equivalent to achieving the maximal throughput, proportional fairness, minimal potential delay, and the max–min fairness of users [28]. It can be verified that $U_\beta(\cdot)$ is concave in $\{R_i\}$ for all β. However, since the outage constraints (2b) have a complicated structure as will be seen later, solving problem (2) is still challenging.

One possible approach to solving such a nonconvex problem is via exhaustive search [22]. In [22], each of the cross-link interference is discretized into M levels, and given a set of cross-link interference levels, the maximum achievable rate for each receiver can be computed [22]. Since there are a total of $K(K-1)$ cross-user links, one has to exhaustively search over $M^{K(K-1)}$ rate tuples. The complexity of this method thus increases exponentially with $K(K-1)$, making this approach only viable when K is small. For a simple example of $K = 3$ and $M = 10$, this method requires searching over 10^6 rate tuples, which is computationally prohibitive in practice.

III. PROPOSED CONVEX APPROXIMATION METHOD

Our goal in this section is to develop an efficient approximation algorithm that obtains near-optimal solutions of problem

(2) for any number of transmitter-receiver pairs, K. To begin with, we note from [24, Appendix I] that the outage probability function in (2b) can actually be expressed in closed form as

$$\Pr\{r_i(\{\boldsymbol{h}_{ki}\}_k, \{\boldsymbol{w}_k\}) < R_i\} = 1 - \left(\exp\left(\frac{-(2^{R_i}-1)\sigma_i^2}{\boldsymbol{w}_i^H \boldsymbol{Q}_{ii} \boldsymbol{w}_i} \right) \right.$$

$$\left. \times \prod_{k \neq i} \frac{\boldsymbol{w}_i^H \boldsymbol{Q}_{ii} \boldsymbol{w}_i}{\boldsymbol{w}_i^H \boldsymbol{Q}_{ii} \boldsymbol{w}_i + (2^{R_i}-1)\boldsymbol{w}_k^H \boldsymbol{Q}_{ki} \boldsymbol{w}_k} \right). \tag{4}$$

So, problem (2) can be rewritten as

$$\max_{\substack{\boldsymbol{w}_i \in \mathbb{C}^{N_t}, R_i \geq 0, \\ i=1,\ldots,K}} U(R_1,\ldots,R_K) \tag{5a}$$

$$\text{s.t. } \rho_i \exp\left(\frac{(2^{R_i}-1)\sigma_i^2}{\boldsymbol{w}_i^H \boldsymbol{Q}_{ii} \boldsymbol{w}_i} \right)$$

$$\times \prod_{k \neq i} \left(1 + \frac{(2^{R_i}-1)\boldsymbol{w}_k^H \boldsymbol{Q}_{ki} \boldsymbol{w}_k}{\boldsymbol{w}_i^H \boldsymbol{Q}_{ii} \boldsymbol{w}_i} \right) \leq 1, \tag{5b}$$

$$\|\boldsymbol{w}_i\|^2 \leq P_i, \quad i = 1,\ldots,K, \tag{5c}$$

where $\rho_i \triangleq 1 - \epsilon_i$. Although the outage probability can now be expressed in closed form, problem (5) is still difficult to solve, since the constraints in (5b) are still nonconvex and complicated. In the ensuing subsections, we present a convex approximation method to handle problem (5) efficiently.

A. Convex Approximation Formulation

The proposed convex approximation method starts with applying semidefinite relaxation (SDR), a convex optimization based approximation technique [26]. Specifically, through SDR, we approximate the quadratic terms $\boldsymbol{w}_k^H \boldsymbol{Q}_{ki} \boldsymbol{w}_k = \text{Tr}(\boldsymbol{w}_k \boldsymbol{w}_k^H \boldsymbol{Q}_{ki})$ in (5b) by the linear terms $\text{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki})$, where the rank-one matrices $\boldsymbol{w}_k \boldsymbol{w}_k^H$ are replaced by the PSD matrices $\boldsymbol{W}_k$ of arbitrary $\text{rank}(\boldsymbol{W}_k) \leq N_t$. The approximated problem is thus given by

$$\max_{\substack{\boldsymbol{W}_i \in \mathbb{H}^{N_t}, R_i \geq 0, \\ i=1,\ldots,K}} U(R_1,\ldots,R_K) \tag{6a}$$

$$\text{s.t. } \rho_i \exp\left(\frac{(2^{R_i}-1)\sigma_i^2}{\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii})} \right)$$

$$\times \prod_{k \neq i} \left(1 + \frac{(2^{R_i}-1)\text{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki})}{\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii})} \right) \leq 1, \tag{6b}$$

$$\text{Tr}(\boldsymbol{W}_i) \leq P_i, \tag{6c}$$

$$\boldsymbol{W}_i \succeq \boldsymbol{0}, \quad i = 1,\ldots,K. \tag{6d}$$

We should mention that SDR has been widely used in various beamforming design problems (see [29] for a review), where, in most cases, a convex semidefinite program (SDP) approximation formulation can be directly obtained via SDR and thus can be efficiently solved. Problem (6), however, is still not convex yet due to the constraints in (6b). Therefore, further approximations are needed for problem (6).

In contrast to SDR that essentially results in a larger problem feasible set, the second approximation is *restrictive*, in the sense that the obtained solution must also be feasible to problem (6).

To illustrate this restrictive approximation, let us consider the following change of variables:

$$e^{x_{ki}} \triangleq \mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}), \quad e^{y_i} \triangleq 2^{R_i} - 1,$$

$$z_i \triangleq \frac{2^{R_i} - 1}{\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii})} = e^{y_i - x_{ii}}, \quad (7)$$

for $i, k = 1, \ldots, K$, where x_{ki}, y_i, z_i are slack variables. By substituting (7) into (6), one can reformulate problem (6) as the following problem

$$\max_{\substack{\{\boldsymbol{W}_i\} \in \mathcal{S}, R_i \geq 0, \\ x_{ki}, y_i, z_i \in \mathbb{R}, \\ k, i = 1, \ldots, K}} U(R_1, \ldots, R_K), \quad (8a)$$

$$\text{s.t. } \rho_i e^{\sigma_i^2 z_i} \prod_{k \neq i} \left(1 + e^{-x_{ii} + x_{ki} + y_i}\right) \leq 1, \quad (8b)$$

$$\mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) \leq e^{x_{ki}}, \quad (8c)$$

$$\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \geq e^{x_{ii}}, \quad (8d)$$

$$R_i \leq \log_2(1 + e^{y_i}), \quad (8e)$$

$$e^{y_i - x_{ii}} \leq z_i, \quad \forall k \in \mathcal{K}_i^c \ i = 1, \ldots, K, \quad (8f)$$

where $\mathcal{K}_i^c \triangleq \{1, \ldots, K\} \setminus \{i\}$, and the set $\mathcal{S}$ is defined in (9) below. Notice that we have replaced the equalities in (7) with inequalities as in (8c) to (8f). It can be verified by the monotonicity of the objective function that all the inequalities in (8b) to (8f) would hold with equalities at the optimal points. We also note that, for example, if the optimal solution satisfies $\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) = 0$ in (8d), then the optimal x_{ii} has to be minus infinity which is not attainable. Similar issues also occur in $\mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki})$ and x_{ki}. In view of this, in (8) we have enforced $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K$ to lie in the subset

$$\mathcal{S} \triangleq \left\{ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \succeq \boldsymbol{0} \,\middle|\, \begin{array}{l} \mathrm{Tr}(\boldsymbol{W}_i) \leq P_i, \ \mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) \geq \delta \\ \forall i, k = 1, \ldots, K \end{array} \right\}, \quad (9)$$

where $\delta > 0$. As long as δ is set to a small number, the rate loss due to (9) would be negligible.

It is interesting to see that constraint (8b) is now convex; constraints (8d) and (8f) are also convex. Constraints (8c) and (8e) are not convex; nevertheless, they are relatively easy to handle compared with the original (6b). Let $(\bar{\boldsymbol{w}}_1 \bar{\boldsymbol{w}}_1^H, \ldots, \bar{\boldsymbol{w}}_K \bar{\boldsymbol{w}}_K^H, \bar{R}_1, \ldots, \bar{R}_K)$ be a feasible point of problem (8). Define

$$\bar{x}_{ki} \triangleq \ln\left(\bar{\boldsymbol{w}}_k^H \boldsymbol{Q}_{ki} \bar{\boldsymbol{w}}_k\right), \ \bar{y}_i \triangleq \ln(2^{\bar{R}_i} - 1), \bar{z}_i \triangleq e^{\bar{y}_i - \bar{x}_{ii}}, \quad (10)$$

for $i, k = 1, \ldots, K$. Then, $\{\bar{x}_{ki}\}, \{\bar{y}_i\}, \{\bar{z}_i\}$ together with $\{\bar{R}_i\}$ and $\bar{\boldsymbol{W}}_i \triangleq \bar{\boldsymbol{w}}_i \bar{\boldsymbol{w}}_i^H, i = 1, \ldots, K$, are feasible to problem (8). Here we conservatively approximate (8c) and (8e) at the point $(\{\bar{x}_{ki}\}_{k, i \neq k}, \{\bar{y}_i\})$. Since both of $e^{x_{ki}}$ and $\log_2(1 + e^{y_i})$ are convex, their first-order lower bounds at $\bar{x}_{ki}$ and $\bar{y}_i$ are respectively given by

$$e^{\bar{x}_{ki}}(x_{ki} - \bar{x}_{ki} + 1) \text{ and } \log_2(1 + e^{\bar{y}_i}) + \frac{e^{\bar{y}_i}(y_i - \bar{y}_i)}{\ln 2 \cdot (1 + e^{\bar{y}_i})}. \quad (11)$$

Consequently, restrictive approximations for (8c) and (8e) are given by

$$\mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) \leq e^{\bar{x}_{ki}}(x_{ki} - \bar{x}_{ki} + 1), \ k \in \mathcal{K}_i^c, \quad (12a)$$

$$R_i \leq \log_2(1 + e^{\bar{y}_i}) + \frac{e^{\bar{y}_i}(y_i - \bar{y}_i)}{\ln 2 \cdot (1 + e^{\bar{y}_i})}. \quad (12b)$$

By replacing (8c) and (8e) with (12a) and (12b), we obtain the following approximation of problem (5):

$$\max_{\substack{\{\boldsymbol{W}_i\} \in \mathcal{S}, R_i \geq 0, \\ x_{ki}, y_i, z_i \in \mathbb{R}, \\ k, i = 1, \ldots, K}} U(R_1, \ldots, R_K), \quad (13a)$$

$$\text{s.t. } \rho_i e^{\sigma_i^2 z_i} \prod_{k \neq i}(1 + e^{-x_{ii} + x_{ki} + y_i}) \leq 1, \quad (13b)$$

$$\mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) \leq e^{\bar{x}_{ki}}(x_{ki} - \bar{x}_{ki} + 1), \quad (13c)$$

$$\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \geq e^{x_{ii}}, \quad (13d)$$

$$R_i \leq \log_2(1 + e^{\bar{y}_i}) + \frac{e^{\bar{y}_i}(y_i - \bar{y}_i)}{\ln 2 \cdot (1 + e^{\bar{y}_i})}, \quad (13e)$$

$$e^{y_i - x_{ii}} \leq z_i, \ \forall k \in \mathcal{K}_i^c, \ i = 1, \ldots, K. \quad (13f)$$

Problem (13) is a convex optimization problem; it can be efficiently solved by standard convex solvers such as CVX [30].

Let $(\hat{\boldsymbol{W}}_1, \ldots, \hat{\boldsymbol{W}}_K)$ and $(\hat{R}_1, \ldots, \hat{R}_K)$ denote the optimal beamforming matrices and achievable rates yielded by the approximation problem (13). Since the lower bounds in (11) may not be exactly tight, it may hold, for $(\hat{\boldsymbol{W}}_1, \ldots, \hat{\boldsymbol{W}}_K)$ and $(\hat{R}_1, \ldots, \hat{R}_K)$ and for some $i = 1, \ldots, K$, that

$$\rho_i \exp\left(\frac{(2^{\hat{R}_i} - 1)\sigma_i^2}{\mathrm{Tr}(\hat{\boldsymbol{W}}_i \boldsymbol{Q}_{ii})}\right) \prod_{k \neq i}\left(1 + \frac{(2^{\hat{R}_i} - 1)\mathrm{Tr}(\hat{\boldsymbol{W}}_k \boldsymbol{Q}_{ki})}{\mathrm{Tr}(\hat{\boldsymbol{W}}_i \boldsymbol{Q}_{ii})}\right) < 1, \quad (14)$$

i.e., the rate outage probability is strictly less than ϵ_i and thus the outage constraint is over satisfied. Alternatively, one can obtain a tight rate tuple $(\tilde{R}_1, \ldots, \tilde{R}_K)$, where $\tilde{R}_i \geq \hat{R}_i$ for all $i = 1, \ldots, K$, by solving the equations

$$\rho_i \exp\left(\frac{(2^{R_i} - 1)\sigma_i^2}{\mathrm{Tr}(\hat{\boldsymbol{W}}_i \boldsymbol{Q}_{ii})}\right) \prod_{k \neq i}\left(1 + \frac{(2^{R_i} - 1)\mathrm{Tr}(\hat{\boldsymbol{W}}_k \boldsymbol{Q}_{ki})}{\mathrm{Tr}(\hat{\boldsymbol{W}}_i \boldsymbol{Q}_{ii})}\right) = 1, \quad (15)$$

for $i = 1, \ldots, K$. Note that each equation in (15) can be efficiently solved by simple line search. The obtained $(\hat{\boldsymbol{W}}_1, \ldots, \hat{\boldsymbol{W}}_K)$ and $(\tilde{R}_1, \ldots, \tilde{R}_K)$ then serve as an approximate solution for problem (6).

In summary, the reformulation above consists of two approximation steps: a) the rank relaxation of $\boldsymbol{w}_k \boldsymbol{w}_k^H$ to $\boldsymbol{W}_k \succeq \boldsymbol{0}$ by SDR, and b) constraint restrictions of (8c) and (8e) by (13c) and (13e). Note that if problem (13) yields a rank-one optimal $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$, a rank-one beamforming solution can be readily obtained by rank-one decomposition of $\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H$ for all $i = 1, \ldots, K$. It is then straightforward to verify by the restrictiveness of (13c) and (13e) that this rank-one beamforming solution $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K)$ is also feasible to the original problem (5) [i.e., problem (2)], thereby satisfying the desired rate outage requirement. In view of this, it is important to investigate the conditions under which problem (13) can yield rank-one optimal $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$. The following proposition provides one such condition:

Proposition 1: *Assume that (13) is feasible. Then there exists an optimal $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$ satisfying $\mathrm{rank}(\boldsymbol{W}_i) \leq 1$, $i = 1, \ldots, K$, provided that the number of users is no larger than three, i.e., $K \leq 3$.*

Proof: Let $(\{\hat{\boldsymbol{W}}_i\}, \{\hat{R}_i\}, \{\hat{x}_{ik}\}, \{\hat{y}_i\}, \{\hat{z}_i\})$ denote an optimal solution of problem (13). Consider

$$\max_{\boldsymbol{W}_i \succeq \boldsymbol{0}} \mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \quad (16a)$$

$$\text{s.t. } \delta \leq \mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) \leq e^{\bar{x}_{ik}}(\hat{x}_{ik} - \bar{x}_{ik} + 1), \ k \in \mathcal{K}_i^c \quad (16b)$$

$$\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \geq \delta, \ \text{Tr}(\boldsymbol{W}_i) \leq P_i, \tag{16c}$$

for all $i = 1, \ldots, K$. By (9) and (13c), $\hat{\boldsymbol{W}}_i$ is also feasible to the above problem (16). Moreover, by (13b), (13d), (13e), (13f) and by the monotonicity of $U(R_1, \ldots, R_K)$, one can show, by contradiction, that $\hat{\boldsymbol{W}}_i$ is actually optimal to problem (16), for all $i = 1, \ldots, K$. Let $\boldsymbol{W}_i'$ be an optimal solution to (16), for $i = 1, \ldots, K$. Then, one can also verify that $(\boldsymbol{W}_1', \ldots, \boldsymbol{W}_K')$ is optimal to problem (13). We hence focus on problem (16). Firstly, since problem (16) is assumed to be feasible and the objective is to maximize $\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii})$, the constraint $\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \geq \delta$ in (16c) is actually irrelevant and can be dropped without affecting the optimal solution. Secondly, it is easy to observe that, for each $k \in \mathcal{K}_i^c$, it is either $\delta < \text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) = e^{\bar{x}_{ik}}(\hat{x}_{ik} - \bar{x}_{ik} + 1)$ or $\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) = e^{\bar{x}_{ik}}(\hat{x}_{ik} - \bar{x}_{ik} + 1) = \delta$ at the optimum; that is, for either case, it is equivalent to having one equality constraint in (16b) at the optimum for each $k \in \mathcal{K}_i^c$. As a result, problem (16) equivalently has only K constraints. According to [31, Theorem 3.2], there always exists an optimal solution $\boldsymbol{W}_i$ of problem (16) satisfying

$$\text{rank}(\boldsymbol{W}_i) \leq \sqrt{K} \text{ for } i = 1, \ldots, K. \tag{17}$$

Therefore, if $K \leq 3$, there always exists a rank-one optimal $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$ for problem (13). ∎

We should mention that $K \leq 3$ is only a sufficient condition but not a necessary condition. For $K > 3$, there may exist other conditions under which a rank-one optimal solution exists for problem (13). If the optimal $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$ is not of rank one, then one may resort to the rank-one approximation procedures such as Gaussian randomization [26], [29] to obtain an approximate solution to (2). Note that, in that case, the utility achieved by the randomized solution would be no larger than $U(\tilde{R}_1, \ldots, \tilde{R}_K)$. Surprisingly, in our computer simulations, we found that problem (13) is always solved with rank-one optimal $\{\boldsymbol{W}_i\}$. Some insightful analyses, which explain why problem (16) is often solved with rank-one optimal $\boldsymbol{W}_i$ for randomly generated problem instances, can be found in [17].

B. Successive Convex Approximation (SCA)

Formulation (13) is obtained by approximating problem (8) at the given feasible point $(\{\bar{\boldsymbol{w}}_i \bar{\boldsymbol{w}}_i^H\}, \{\bar{R}_i\})$, as described in (10). This approximation can be further improved by successively approximating problem (8) based on the optimal solution $(\{\boldsymbol{W}_i\}, \{R_i\})$ obtained by solving (13) in the previous approximation. Specifically, let $(\hat{\boldsymbol{W}}_1[n-1], \ldots, \hat{\boldsymbol{W}}_K[n-1])$ be the optimal beamforming matrices obtained in the $(n-1)$th iteration, and, similar to (15), let $(\tilde{R}_i[n-1], \ldots, \tilde{R}_K[n-1])$ be the corresponding achievable rate tuple obtained by solving the following K equations

$$\rho_i \exp\left(\frac{(2^{R_i}-1)\sigma_i^2}{\text{Tr}\left(\hat{\boldsymbol{W}}_i[n-1]\boldsymbol{Q}_{ii}\right)}\right)$$
$$\times \prod_{k \neq i}\left(1 + \frac{(2^{R_i}-1)\text{Tr}\left(\hat{\boldsymbol{W}}_k[n-1]\boldsymbol{Q}_{ki}\right)}{\text{Tr}\left(\hat{\boldsymbol{W}}_i[n-1]\boldsymbol{Q}_{ii}\right)}\right) = 1,$$
$$i = 1, \ldots, K. \tag{18}$$

Moreover, let

$$\bar{x}_{ki}[n-1] = \ln\left(\text{Tr}\left(\hat{\boldsymbol{W}}_k[n-1]\boldsymbol{Q}_{ki}\right)\right), \tag{19a}$$

$$\bar{y}_i[n-1] = \ln\left(2^{\tilde{R}_i[n-1]} - 1\right), \ i, k = 1, \ldots, K. \tag{19b}$$

By replacing $\bar{x}_{ki}$ and $\bar{y}_i$ in (13) with $\bar{x}_{ki}[n-1]$ and $\bar{y}_i[n-1]$ in (19) for $k \in \mathcal{K}_i^c, i = 1, \ldots, K$, we solve, in the nth iteration, the following convex optimization problem

$$\left(\left\{\hat{\boldsymbol{W}}_i[n]\right\}, \left\{\hat{R}_i[n]\right\}, \{\hat{x}_{ik}[n]\}, \{\hat{y}_i[n]\}, \{\hat{z}_i[n]\}\right)$$
$$= \arg \max_{\substack{\{\boldsymbol{W}_i\} \in \mathcal{S}, R_i \geq 0 \\ x_{ik}, y_i, z_i \in \mathbb{R} \\ i, k = 1, \ldots, K}} U(R_1, \ldots, R_K) \tag{20a}$$

$$\text{s.t. } \rho_i e^{\sigma_i^2 z_i} \prod_{k \neq i}(1 + e^{-x_{ii}+x_{ki}+y_i}) \leq 1, \tag{20b}$$

$$\text{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) \leq e^{\bar{x}_{ki}[n-1]}(x_{ki} - \bar{x}_{ki}[n-1] + 1), \tag{20c}$$

$$\text{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ii}) \geq e^{x_{ii}}, \tag{20d}$$

$$R_i \leq \log_2\left(1 + e^{\bar{y}_i[n-1]}\right) + \frac{e^{\bar{y}_i[n-1]}(y_i - \bar{y}_i[n-1])}{\ln 2 \cdot \left(1 + e^{\bar{y}_i[n-1]}\right)}, \tag{20e}$$

$$e^{y_i - x_{ii}} \leq z_i, \ \forall k \in \mathcal{K}_i^c, \ i = 1, \ldots, K. \tag{20f}$$

Note that the rate $\tilde{R}_i[n-1]$ obtained by (18) is no less than $\hat{R}_i[n-1]$ for all $i = 1, \ldots, K$, and thus the former is used to compute $\{\bar{y}_i[n-1]\}$ as the point for successive approximation. In fact, successive approximation ensures monotonic improvement of the utility $U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])$. Let us define

$$\bar{z}_i[n-1] \triangleq e^{\bar{y}_i[n-1] - \bar{x}_{ii}[n-1]}, \ i = 1, \ldots, K. \tag{21}$$

Then, by (18), (19) and (21), one can show that $(\{\hat{\boldsymbol{W}}_i[n-1]\}, \{\tilde{R}_i[n-1]\}, \{\bar{x}_{ik}[n-1]\}, \{\bar{y}_i[n-1]\}, \{\bar{z}_i[n-1]\})$ is a feasible point of (20). As a result, we have

$$U\left(\tilde{R}_1[n], \ldots, \tilde{R}_K[n]\right) \geq U\left(\hat{R}_1[n], \ldots, \hat{R}_K[n]\right)$$
$$\geq U\left(\tilde{R}_1[n-1], \ldots, \tilde{R}_K[n-1]\right), \quad \forall n \geq 1. \tag{22}$$

The proposed *successive convex approximation* (SCA) algorithm is summarized in Algorithm 1.

Algorithm 1: SCA algorithm for solving problem (2)

1: **Given** $(\bar{\boldsymbol{w}}_1 \bar{\boldsymbol{w}}_1^H, \ldots, \bar{\boldsymbol{w}}_K \bar{\boldsymbol{w}}_K^H) \in \mathcal{S}$ and $(\bar{R}_1, \ldots, \bar{R}_K)$ that are feasible to (6).
2: Set $\hat{\boldsymbol{W}}_i[0] = \bar{\boldsymbol{w}}_i \bar{\boldsymbol{w}}_i^H$ and $\tilde{R}_i[0] = \bar{R}_i$ for all $i = 1, \ldots, K$, and set $n = 0$.
3: **repeat**
4: $\quad n := n + 1$.
5: $\quad$ Obtain $\{\bar{x}_{ki}[n-1]\}$ and $\{\bar{y}_i[n-1]\}$ by (19), and solve problem (20) to obtain the optimal solution $\hat{\boldsymbol{u}}[n] \triangleq$ $(\{\hat{\boldsymbol{W}}_i[n]\}, \{\hat{R}_i[n]\}, \{\hat{x}_{ik}[n]\}, \{\hat{y}_i[n]\}, \{\hat{z}_i[n]\})$.
6: $\quad$ Compute $(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])$ by solving (18).
7: **until** the stopping criterion is met.
8: Obtain $\boldsymbol{w}_i^\star$ by decomposition of $\hat{\boldsymbol{W}}_i[n] = \boldsymbol{w}_i^\star(\boldsymbol{w}_i^\star)^H$ for all i, if $\hat{\boldsymbol{W}}_i[n]$ are all of rank one; otherwise perform Gaussian randomization [29] to obtain a rank-one approximate solution of (2).

C. Convergence Analysis

Convergence properties of Algorithm 1 are given in the following theorem:

Theorem 1: *Suppose that the utility $U(R_1, \ldots, R_K)$ is differentiable and strictly increasing with respect to R_i, for $i = 1, \ldots, K$. Then, the sequence, $\{U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])\}_{n=1}^{\infty}$ generated by Algorithm 1, converges, and any limit point of the sequence $\{(\hat{\boldsymbol{W}}_1[n], \ldots, \hat{\boldsymbol{W}}_K[n]), (\tilde{R}_1[n], \ldots, \tilde{R}_K[n])\}_{n=1}^{\infty}$ is a stationary point of problem (8) as well as a stationary point of problem (6) with extra constraints $\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) \geq \delta$ for $i, k = 1, \ldots, K$ (see (9)).*

Proof: As discussed earlier, the utility $U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])$ is nondecreasing with n. Moreover, due to the individual power constraints, the sequence $\{U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])\}_{n=1}^{\infty}$ is bounded, which implies the convergence of $U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])$.

Let $\hat{\boldsymbol{u}}[n] \triangleq (\{\hat{\boldsymbol{W}}_i[n]\}, \{\hat{R}_i[n]\}, \{\hat{x}_{ik}[n]\}, \{\hat{y}_i[n]\}, \{\hat{z}_i[n]\})$, denote the optimal solution of (20). To prove that any limit point of $\hat{\boldsymbol{u}}[n]$ is a stationary point of (8), two key observations are needed. Firstly, we note that the proposed SCA algorithm is in fact an inner approximation algorithm in the nonconvex optimization literature [32]. In particular, the nonconvex constraints in (8c) and (8e), i.e.,

$$\Psi_{ki}(\boldsymbol{W}_k, x_{ki}) \triangleq \mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) - e^{x_{ki}} \leq 0, \ k \in \mathcal{K}_i^c,$$

$$\Phi_i(R_i, y_i) \triangleq R_i - \log_2(1 + e^{y_i}) \leq 0, \ i = 1, \ldots, K,$$

are respectively replaced by

$$
\begin{aligned}
&\bar{\Psi}_{ki}\left(\boldsymbol{W}_k, x_{ki} | \bar{x}_{ki}[n-1]\right) \\
&\triangleq \mathrm{Tr}(\boldsymbol{W}_k \boldsymbol{Q}_{ki}) - e^{\bar{x}_{ki}[n-1]}\left(x_{ki} - \bar{x}_{ki}[n-1] + 1\right) \\
&\leq 0, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (23)\\
&\bar{\Phi}_i\left(R_i, y_i | \bar{y}_i[n-1]\right) \\
&\triangleq R_i - \log_2\left(1 + e^{\bar{y}_i[n-1]}\right) + \frac{e^{\bar{y}_i[n-1]}\left(y_i - \bar{y}_i[n-1]\right)}{\ln 2 \cdot \left(1 + e^{\bar{y}_i[n-1]}\right)} \\
&\leq 0, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (24)
\end{aligned}
$$

for all $k \in \mathcal{K}_i^c$, $i = 1, \ldots, K$. One can verify that $\bar{\Psi}_{ki}(\boldsymbol{W}_k, x_{ki} | \bar{x}_{ki}[n-1])$ and $\bar{\Phi}_i(R_i, y_i | \bar{y}_i[n-1])$ satisfy

$$
\begin{aligned}
&\Psi_{ki}\left(\hat{\boldsymbol{W}}_k[n-1], \bar{x}_{ki}[n-1]\right) \\
&= \bar{\Psi}_{ki}\left(\hat{\boldsymbol{W}}_k[n-1], \bar{x}_{ki}[n-1] | \bar{x}_{ki}[n-1]\right) = 0 \quad (25)
\end{aligned}
$$

$$\frac{\partial \Psi_{ki}(\boldsymbol{W}_k, x_{ki})}{\partial \boldsymbol{W}_k} = \frac{\partial \bar{\Psi}_{ki}\left(\boldsymbol{W}_k, x_{ki} | \bar{x}_{ki}[n-1]\right)}{\partial \boldsymbol{W}_k} \quad (26)$$

$$
\begin{aligned}
&\left.\frac{\partial \Psi_{ki}(\boldsymbol{W}_k, x_{ki})}{\partial x_{ki}}\right|_{x_{ki} = \bar{x}_{ki}[n-1]} \\
&= \left.\frac{\partial \bar{\Psi}_{ki}\left(\boldsymbol{W}_k, x_{ki} | \bar{x}_{ki}[n-1]\right)}{\partial x_{ki}}\right|_{x_{ki} = \bar{x}_{ki}[n-1]} \quad (27)
\end{aligned}
$$

$$
\begin{aligned}
&\Phi_i\left(\hat{R}_i[n-1], \bar{y}_i[n-1]\right) \\
&= \bar{\Phi}_i\left(\hat{R}_i[n-1], \bar{y}_i[n-1] | \bar{y}_i[n-1]\right) = 0 \quad (28)
\end{aligned}
$$

$$\frac{\partial \Phi_i(R_i, y_i)}{\partial R_i} = \frac{\partial \bar{\Phi}_i\left(R_i, y_i | \bar{y}_i[n-1]\right)}{\partial R_i} \quad (29)$$

$$\left.\frac{\partial \Phi_i(R_i, y_i)}{\partial y_i}\right|_{\substack{y_i = \\ \bar{y}_i[n-1]}} = \left.\frac{\partial \bar{\Phi}_i\left(R_i, y_i | \bar{y}_i[n-1]\right)}{\partial y_i}\right|_{\substack{y_i = \\ \bar{y}_i[n-1]}}, \quad (30)$$

for all $k \in \mathcal{K}_i^c$ and $i = 1, \ldots, K$.

Secondly, the restrictive approximations made in (20c) and (20e) are asymptotically tight as $n \to \infty$ as stated in the following claim:

Claim 1: *It holds true that*

$$\lim_{n \to \infty} |\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1]| = 0, \quad (31)$$

$$\lim_{n \to \infty} |\hat{y}_i[n] - \bar{y}_i[n-1]| = 0, \quad \lim_{n \to \infty} \left(\tilde{R}_i[n] - \hat{R}_i[n]\right) = 0, \quad (32)$$

for all $k \in \mathcal{K}_i^c$, $i = 1, \ldots, K$.

Claim 1 is proved in Appendix A. Moreover, by the monotonicity of $U(R_1, \ldots, R_K)$ and due to (9), it is not difficult to verify that:

Claim 2: *The sequence $\{\hat{\boldsymbol{u}}[n]\}_{n=0}^{\infty}$ generated by Algorithm 1 is bounded.*

Now let us consider the KKT conditions of (20). Denote $\mathcal{L}(\hat{\boldsymbol{u}}[n], \boldsymbol{\lambda}[n] | \{\{\bar{x}_{ki}[n-1]\}_{k \neq i}\}_i, \{\bar{y}_i[n-1]\})$ as the Lagrangian of (20). For ease of explanation, let

$$\Theta_i(\{x_{ki}\}_k, y_i, z_i) \triangleq \rho_i e^{\sigma_i^2 z_i} \prod_{k \neq i}(1 + e^{-x_{ii} + x_{ki} + y_i}) - 1$$

denote the constraint function in (20b), and consider the following Lagrangian-stationarity condition:

$$
\begin{aligned}
&\frac{\partial \mathcal{L}\left(\hat{\boldsymbol{u}}[n], \boldsymbol{\lambda}[n] | \left\{\{\bar{x}_{ki}[n-1]\}_{k \neq i}\right\}_i, \{\bar{y}_i[n-1]\}\right)}{\partial x_{ki}} \\
&= \lambda_i^{\mathrm{b}}[n] \frac{\partial \Theta_i\left(\{\hat{x}_{ji}[n]\}_j, \hat{y}_i[n], \hat{z}_i[n]\right)}{\partial x_{ki}} \\
&\quad + \lambda_{ki}^{\mathrm{c}}[n] \frac{\partial \bar{\Psi}_{ki}\left(\hat{\boldsymbol{W}}_k[n], \hat{x}_{ki}[n] | \bar{x}_{ki}[n-1]\right)}{\partial x_{ki}} = 0, \quad \forall k \neq i, \\
&\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (33)
\end{aligned}
$$

where $\boldsymbol{\lambda}[n] \triangleq (\{\lambda_i^{\mathrm{b}}[n]\}, \{\{\lambda_{ik}^{\mathrm{c}}[n]\}_{k \neq i}\}_i, \{\lambda_i^{\mathrm{d}}[n]\}, \{\lambda_i^{\mathrm{e}}[n]\}, \{\lambda_i^{\mathrm{f}}[n]\}, \{\lambda_i^P[n]\}, \{\lambda_{ik}^{\delta}[n]\})$ are dual variables associated with constraints (20b)–(20f), the transmit power constraint and $\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) \geq \delta$. Since problem (20) satisfies the Slater's condition, the dual variables are bounded [33]. Moreover, $\hat{\boldsymbol{u}}[n]$ is bounded as well by Claim 2. Therefore, there exists a subsequence $\{n_1, \ldots, n_\ell, \ldots\} \subseteq \{1, \ldots, n, \ldots\}$ and a primal-dual limit point, denoted by $\hat{\boldsymbol{u}}^\star \triangleq (\{\hat{\boldsymbol{W}}_i^\star\}, \{\hat{R}_i^\star\}, \{\hat{x}_{ik}^\star\}, \{\hat{y}_i^\star\}, \{\hat{z}_i^\star\})$ and $\boldsymbol{\lambda}^\star \triangleq (\{\lambda_i^{\mathrm{b}\star}\}, \{\{\lambda_{ik}^{\mathrm{c}\star}\}_{k \neq i}\}_i, \{\lambda_i^{\mathrm{d}\star}\}, \{\lambda_i^{\mathrm{e}\star}\}, \{\lambda_i^{\mathrm{f}\star}\}, \{\lambda_i^{P\star}\}, \{\lambda_{ik}^{\delta\star}\})$ such that

$$\lim_{\ell \to \infty} \hat{\boldsymbol{u}}[n_\ell] = \hat{\boldsymbol{u}}^\star, \quad \lim_{\ell \to \infty} \boldsymbol{\lambda}[n_\ell] = \boldsymbol{\lambda}^\star, \quad (34)$$

where $(\hat{\boldsymbol{u}}^\star, \boldsymbol{\lambda}^\star)$ is primal-dual feasible to (20). Consider (33) over the subsequence $\{n_1, \ldots, n_\ell, \ldots\}$. By taking $\ell \to \infty$ in (33), and by (27), (31) and (34), we obtain

$$\lambda_i^{\mathrm{b}\star} \frac{\partial \Theta_i\left(\{\hat{x}_{ji}^\star\}_j, \hat{y}_i^\star, \hat{z}_i^\star\right)}{\partial x_{ki}} + \lambda_{ki}^{\mathrm{c}\star} \frac{\partial \Psi_{ki}\left(\hat{\boldsymbol{W}}_k^\star, \hat{x}_{ki}^\star\right)}{\partial x_{ki}} = 0,$$

which is the Lagrangian-stationarity condition of (8) corresponding to x_{ki}. By applying similar arguments above to all the other KKT conditions of (20) and by Claims 1 and 2, we

end up with the conclusion that $\hat{\boldsymbol{u}}^{\star}$ satisfies the KKT conditions of problem (8) and thus is a stationary point.

According to (32), $\hat{R}_i^{\star}$ is also the limit point of $\{\tilde{R}_i[n_\ell]\}_{\ell=1}^{\infty}$ for all $i = 1, \ldots, K$. Thus, what remains is to show that any stationary point of (8) is also a stationary point of (6) if the constraint set (9) is added in (6). This can be proved by carefully examining the equivalence of the KKT conditions of the two problems. Due to the space limitation, the detailed derivations are omitted here. ∎

As the SCA algorithm only guarantees to provide a stationary point, the approximation accuracy depends on the initial point $(\{\hat{\boldsymbol{W}}_i[0]\}, \{\tilde{R}_i[0]\})$. A possible choice is to initialize Algorithm 1 via some heuristic transmission strategies. For example, one can obtain an initial point $(\{\bar{\boldsymbol{w}}_i\}, \{\bar{R}_i\})$ of problem (5) through the simple maximum-ratio transmission (MRT) strategy. In this strategy, the beamforming vectors $\{\bar{\boldsymbol{w}}_i\}$ are simply set to $\bar{\boldsymbol{w}}_i = \sqrt{P_i}\boldsymbol{q}_i$ where $\boldsymbol{q}_i \in \mathbb{C}^{N_t}$ is the principal eigenvector of $\boldsymbol{Q}_{ii}$ for $i = 1, \ldots, K$, with $\|\boldsymbol{q}_i\| = 1$. The corresponding rate $\tilde{R}_i$ can be obtained by solving (15) with $\{\hat{\boldsymbol{W}}_i\} = \{\bar{\boldsymbol{w}}_i\bar{\boldsymbol{w}}_i^H\}$. Analogously, one can also obtain an initial point by the zero-forcing (ZF) transmission strategy, provided that the column space of $\boldsymbol{Q}_{ii}$ is not subsumed by the column space of $\sum_{k\neq i}^{K}\boldsymbol{Q}_{ik}$, for all $i = 1, \ldots, K$.

IV. DISTRIBUTED IMPLEMENTATION

For Algorithm 1, we have implicitly assumed that there exists a control center in the network, collecting all the CDI of users and computing the beamforming solution in a centralized manner. In this section, we propose a distributed version for Algorithm 1, where each transmitter i only needs to optimize its own beamformer, using only its local CDI, i.e., $\{\boldsymbol{Q}_{ik}\}_k$, and some information obtained from the other transmitters. Since each of the subproblems involved has a much smaller problem size than the original problem (8), even for a centralized implementation, the proposed distributed optimization algorithm can be used to reduce the computation overhead of the control center.

The idea of the proposed distributed algorithm is to solve problem (8) from one transmitter to another, in a round-robin fashion (i.e., the Gauss-Seidel fashion). Suppose that transmitter 1 optimizes its beamformer first, followed by transmitter 2 and so on, and let n denote the index of the current round. Then, in the nth round, transmitter i solves the following problem

$$\left(\hat{\boldsymbol{w}}_i[n], \hat{R}_1[n,i], \ldots, \hat{R}_K[n,i]\right)$$
$$= \arg \max_{\substack{\|\boldsymbol{w}_i\|^2 \leq P_i \\ R_1,\ldots,R_K \geq 0}} U(R_1, \ldots, R_K) \qquad (35\text{a})$$

$$\text{s.t. } \rho_i \exp\left(\frac{(2^{R_i}-1)\sigma_i^2}{\boldsymbol{w}_i^H\boldsymbol{Q}_{ii}\boldsymbol{w}_i}\right)$$
$$\times \prod_{k\neq i}\left(1+\frac{(2^{R_i}-1)e^{\bar{x}_{ki}[n-u_{ki}]}}{\boldsymbol{w}_i^H\boldsymbol{Q}_{ii}\boldsymbol{w}_i}\right) \leq 1, \qquad (35\text{b})$$

$$\rho_j \exp\left(\frac{(2^{R_j}-1)\sigma_j^2}{e^{\bar{x}_{jj}[n-u_{ji}]}}\right)\left(1+\frac{(2^{R_j}-1)\boldsymbol{w}_i^H\boldsymbol{Q}_{ij}\boldsymbol{w}_i}{e^{\bar{x}_{jj}[n-u_{ji}]}}\right)$$
$$\times \prod_{\substack{k\neq j \\ k\neq i}}\left(1+\frac{(2^{R_j}-1)e^{\bar{x}_{kj}[n-u_{ki}]}}{e^{\bar{x}_{jj}[n-u_{ji}]}}\right) \leq 1, \ j\in\mathcal{K}_i^c, \quad (35\text{c})$$

where $\bar{x}_{kj}[n - u_{ki}] = \ln(\hat{\boldsymbol{w}}_k^H[n - u_{ki}]\boldsymbol{Q}_{kj}\hat{\boldsymbol{w}}_k[n - u_{ki}])$, and u_{ki} is equal to one if $k > i$ and zero otherwise.

Note that for (35), only $\hat{\boldsymbol{w}}_i[n]$ is optimized while $\{\bar{x}_{kj}[n - u_{kj}]\}_{k\neq i,j}$ are given. Once the beamforming solution of (35) is obtained, $\{\bar{x}_{ik}[n]\}_k$ are updated according to the optimal $\hat{\boldsymbol{w}}_i[n]$ and then passed to all the other transmitters for their subsequent beamforming optimization.[1] There are two interesting points to note here. Firstly, as can be seen from (35b) and (35c), transmitter i not only optimizes its rate R_i, but also takes into account the rate outage constraints for all the other users. The constraints in (35c) indicate that transmitter i needs to regulate its own transmission in order not to violate the outage requirement of the other users. Secondly, to solve (35), transmitter i only needs the local CDI, i.e., $\{\boldsymbol{Q}_{ik}\}_k$.

Similar difficulties arise here as in problem (5) since problem (35) is not convex. We hence apply the same approximation techniques in Section III-A to approximate (35). In particular, we first apply SDR, followed by the reformulation as described by (7), and the first-order approximations in (12). The resulting convex optimization problem is given by

$$\left(\hat{\boldsymbol{W}}_i[n], \left\{\hat{R}_k[n,i]\right\}, \{\hat{x}_{ik}[n]\}_k, \{\hat{y}_k[n,i]\}, \{\hat{z}_k[n,i]\}\right)$$
$$= \arg \max_{\substack{\boldsymbol{W}_i\in\mathcal{S}_i, R_k, x_{ik}, y_k, z_k \\ k=1,\ldots,K}} U(R_1, \ldots, R_K) \qquad (36\text{a})$$

$$\text{s.t. } \rho_i e^{\sigma_i^2 z_i} \prod_{k\neq i}\left(1+e^{-x_{ii}+\bar{x}_{ki}[n-u_{ki}]+y_i}\right) \leq 1, \quad (36\text{b})$$

$$\rho_j e^{\sigma_j^2 z_j}\left(1+e^{-\bar{x}_{jj}[n-u_{ji}]+x_{ij}+y_j}\right)$$
$$\times \prod_{k\neq j,k\neq i}\left(1+e^{-\bar{x}_{jj}[n-u_{ji}]+\bar{x}_{kj}[n-u_{ki}]+y_j}\right) \leq 1, \ j\in\mathcal{K}_i^c$$
$$(36\text{c})$$

$$\text{Tr}(\boldsymbol{W}_i\boldsymbol{Q}_{ii}) \geq e^{x_{ii}}, \qquad (36\text{d})$$

$$\text{Tr}(\boldsymbol{W}_i\boldsymbol{Q}_{ik}) \leq e^{\bar{x}_{ik}[n-1]}\left(x_{ik}-\bar{x}_{ik}[n-1]+1\right), \ k\in\mathcal{K}_i^c$$
$$(36\text{e})$$

$$R_j \leq \log_2\left(1+e^{\bar{y}_j[n,i-1]}\right)$$
$$\times \frac{e^{\bar{y}_j[n,i-1]}\left(y_j-\bar{y}_j[n,i-1]\right)}{\ln 2\cdot\left(1+e^{\bar{y}_j[n,i-1]}\right)}, \ j=1,\ldots,K, \qquad (36\text{f})$$

$$e^{y_i-x_{ii}} \leq z_i, \ e^{y_j-\bar{x}_{jj}[n-u_{ji}]} \leq z_j, \ j\in\mathcal{K}_i^c, \qquad (36\text{g})$$

where $\mathcal{S}_i \triangleq \{\boldsymbol{W}_i \succeq \boldsymbol{0}|\text{Tr}(\boldsymbol{W}_i) \leq P_i, \text{Tr}(\boldsymbol{W}_i\boldsymbol{Q}_{ik}) \geq \delta, k = 1, \ldots, K\}$,

$$\bar{x}_{kj}[n - u_{ki}] = \ln\left(\text{Tr}\left(\hat{\boldsymbol{W}}_k[n - u_{ki}]\boldsymbol{Q}_{kj}\right)\right), \qquad (37)$$

$$\bar{y}_j[n, i - 1] = \ln\left(2^{\tilde{R}_j[n,i-1]} - 1\right), \qquad (38)$$

for $j, k = 1, \ldots, K$. Similar to (18), $\tilde{R}_j[n, i] \geq \hat{R}_j[n, i]$ is obtained by solving the following equations

$$\rho_j \exp\left(\frac{(2^{R_j}-1)\sigma_j^2}{e^{\bar{x}_{jj}[n-u_{ji}]}}\right)$$
$$\times \prod_{k\neq j}\left(1+(2^{R_j}-1)e^{\bar{x}_{kj}[n-u_{ki}]-\bar{x}_{jj}[n-u_{ji}]}\right) = 1, \quad (39)$$

---

[1] In this paper, we assume that the communication between transmitters for message exchange is error-free.

for $j = 1, \ldots, K$. It is worth mentioning that problem (36) is only solved once and successive approximation is not performed as in Algorithm 1. As long as problem (36) is solved by transmitter i, the algorithm directly goes to the next optimization problem performed by transmitter $i + 1$. Successive approximation is now performed implicitly from one transmitter to another in a round-robin fashion. We summarize the proposed distributed SCA algorithm in Algorithm 2.

Algorithm 2: Distributed SCA algorithm for solving problem (2)

1: **Given** an initial beamforming matrix $\hat{W}_i[0]$ at transmitter i, for $i = 1, \ldots, K$.
2: For all $i = 1, \ldots, K$, transmitter i computes $\{\bar{x}_{ik}[0]\}_k$ by (37), and pass them to the other transmitters.
3: Set $n = 0$
4: **repeat**
5:     $n = n + 1$
6:     **for** $i = 1, \ldots, K$ **do**
7:        User i solves (39) to obtain $\{\tilde{R}_j[n, i-1]\}_j$ and compute $\{\bar{y}_j[n, i-1]\}_j$, followed by solving (36) to obtain the solution $(\hat{W}_i[n], \{\hat{R}_k[n,i]\}, \{\hat{x}_{ik}[n]\}_k, \{\hat{y}_k[n,i]\}, \{\hat{z}_k[n,i]\})$.
8:        User i computes $\{\bar{x}_{ik}[n]\}_k$ by (37) and passes them to all the other transmitters.
9:     **end for**
10: **until** the predefined stopping criterion is met.
11: For $i = 1, \ldots, K$, each transmitter i decomposes $\hat{W}_i[n] = \hat{w}_i \hat{w}_i^H$, if $\hat{W}_i[n]$ is of rank one; otherwise perform Gaussian randomization to obtain a rank-one approximate solution.

Analogous to Algorithm 1, we can show that Algorithm 2 generates a stationary point of problem (8) as stated in the following theorem.

Theorem 2: *Suppose that $U(R_1, \ldots, R_K)$ is differentiable and is strictly increasing with respect to R_i, for $i = 1, \ldots, K$. Then, the sequence $\{U(\tilde{R}_1[n, i], \ldots, \tilde{R}_K[n, i])\}_{n=1}^\infty$ generated by Algorithm 2 converges to a common value for all $i = 1, \ldots, K$. Moreover, for all i, any limit point of the sequence $\{(\hat{W}_1[n], \ldots, \hat{W}_k[n], \tilde{R}_1[n, i], \ldots, \tilde{R}_1[n, i])\}_{n=1}^\infty$ is a stationary point of problem (8) as well as a stationary point of problem (6) (with the extra constraints in (9)).*

Different from the proof for Theorem 1, the proof for Theorem 2 is more involved, since the beamforming vectors of transmitters are not simultaneously optimized as in Algorithm 1 but are individually optimized in a round-robin manner. The detailed proof of Theorem 2 is presented in Appendix B.

Remark 1: An important issue concerning distributed optimization algorithms is the communication overhead introduced by message exchange between transmitters. To address this, we compare the communication overhead of the proposed Algorithm 2 with the following two schemes. Scheme 1: All the transmitters directly exchange their CDI so that the design problem (2) can be handled independently by each transmitter. Scheme 2: A control center gathers the CDI from all transmitters, optimizes the beamforming vectors, and distributes the beamforming solutions to the transmitters. We consider a cellular system where all the transmitters (i.e., BSs) are connected by dedicated backhaul links (e.g., optical fibers) and the BSs

exchange messages in a point-to-point fashion. Since, in Algorithm 2, transmitter i needs to inform $\{\bar{x}_{ik}[n]\}_k$ (K real values) to all the other $K - 1$ transmitters in each round, the communication overhead due to transmitter i is quantified by the amount of $K(K-1)$ real values. Hence, the total communication overhead of Algorithm 2 is $N \times K \times K(K-1) = K^2(K-1)N$ real values, where N is the number of rounds run by Algorithm 2. For scheme 1, each transmitter needs to send K covariance matrices (which contain $K N_t^2$ real values) to all the other $K - 1$ transmitters. Therefore, the associated total communication overhead is given by $K \times (K-1) \times K N_t^2 = K^2(K-1)N_t^2$ real values. Therefore, for scheme 1, if $N < N_t^2$, then the proposed Algorithm 2 has a smaller amount of communication overhead. For scheme 2, there are K^2 covariance matrices sent from the transmitters to the control center, and the optimal solution $\{w_i^\star, R_i^\star\}$ passed from the control center to transmitter i for $i = 1, \ldots, K$, respectively. Hence, the communication overhead is $K^2 N_t^2 + K(2N_t + 1)$ real values. Therefore, for scheme 2, the proposed Algorithm 2 has a smaller amount of communication overhead if $N < \frac{N_t^2}{(K-1)} + \frac{(2N_t+1)}{(K^2-K)}$. As we show in the simulation section, Algorithm 2 in general can converge in less than 15 rounds for a system with $K \le 6$ and $N_t = 8$.

Remark 2: It should be noticed that, while in general Algorithm 2 is more efficient in terms of computation and communication overhead, it may result in larger transmission delays and thus reduced throughput (due to the round-robin optimization in Algorithm 2) compared to the centralized schemes. Nonetheless, since the proposed beamforming design is based on the users' statistical channel information, which usually changes much more slowly compared to the instantaneous CSI, beamforming optimization needs not be performed frequently. As a result, the induced throughput loss should not be a serious concern.

V. SIMULATION RESULTS

In the section, we demonstrate the performance of the proposed Algorithm 1 and Algorithm 2 for solving the outage constrained coordinated beamforming problem in (2). In the simulations, we consider $\beta = 0$, $\beta = 1$, and $\beta = 2$ for the objective function $U_\beta(R_1, \ldots, R_K)$, corresponding to maximization of the weighted sum rate, the weighted geometric mean rate, and the weighted harmonic mean rate, respectively. All receivers are assumed to have the same noise power, i.e., $\sigma_1^2 = \cdots = \sigma_K^2 \triangleq \sigma^2$, and all power constraints are set to one, i.e., $P_1 = \cdots = P_K = 1$. The parameter δ in (9) is set to 10^{-5}. The channel covariance matrices Q_{ki} are randomly generated. We normalize the maximum eigenvalue of Q_{ii}, i.e., $\lambda_{\max}(Q_{ii})$, to one for all i, and normalize $\lambda_{\max}(Q_{ki})$ to a value $\eta \in (0, 1]$ for all $k \in \mathcal{K}_i^c$, $i = 1, \ldots, K$. The parameter η, thereby, represents the relative cross-link interference level. If not mentioned specifically, all Q_{ki} are of full rank, and the outage probability requirements are set to the same value, i.e., $\epsilon_1 = \cdots = \epsilon_K = 0.1$, indicating a 10% outage probability. The stopping criterion of Algorithm 1 is

$$\frac{\left| U\left(\tilde{R}_1[n], \ldots, \tilde{R}_K[n]\right) - U\left(\tilde{R}_1[n-1], \ldots, \tilde{R}_K[n-1]\right) \right|}{U\left(\tilde{R}_1[n-1], \ldots, \tilde{R}_K[n-1]\right)} < 0.01.$$
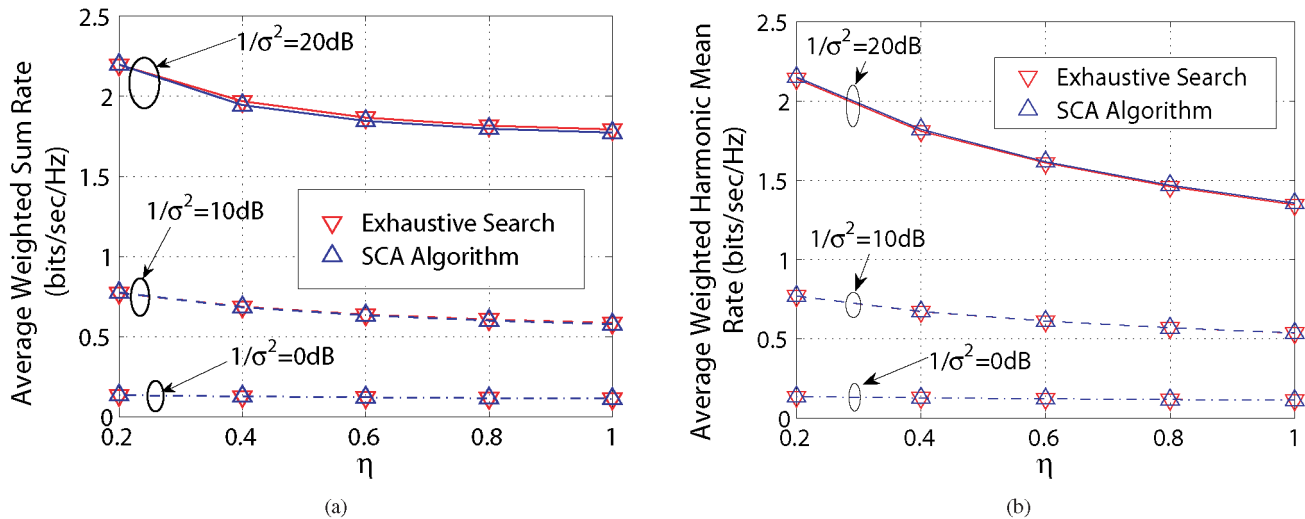
Fig. 1. Simulation results of the proposed SCA algorithm (Algorithm 1), for $K = 2$, $N_t = 4$, and $(\alpha_1, \alpha_2) = (\frac{1}{2}, \frac{1}{2})$; (a) weighted sum rate versus η, (b) weighted harmonic mean rate versus η. Each of the results is obtained by averaging over 500 realizations of $\{\boldsymbol{Q}_{ki}\}$.
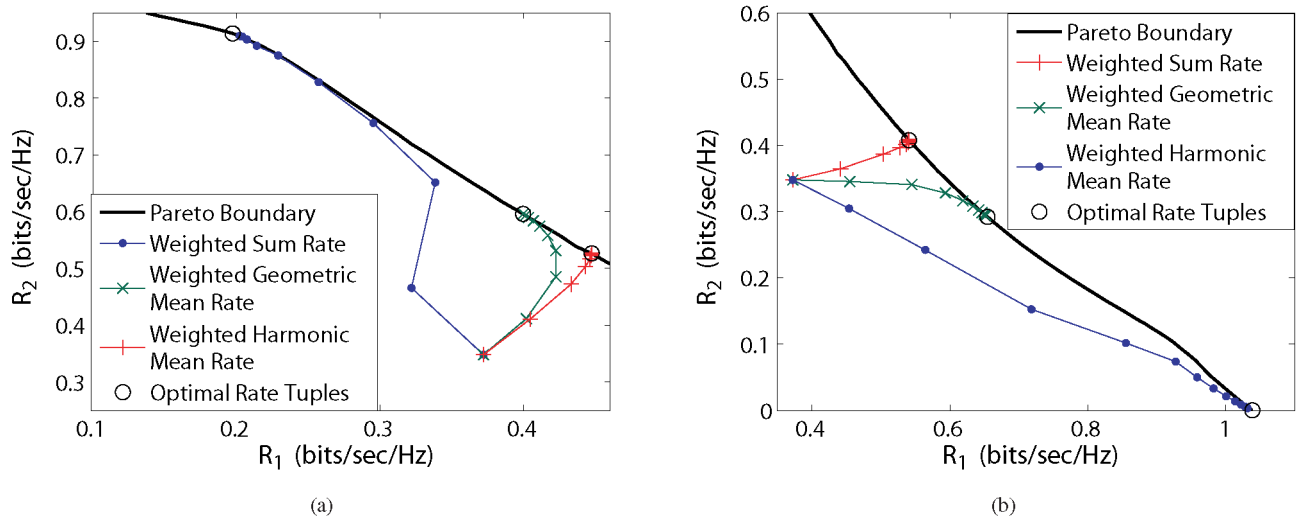

Fig. 2. Converge trajectories of the proposed SCA algorithm. $K = 2$, $N_t = 4$, $\eta = 0.4$; (a) $(\alpha_1, \alpha_2) = (\frac{1}{2}, \frac{1}{2})$, (b) $(\alpha_1, \alpha_2) = (\frac{2}{3}, \frac{1}{3})$. The results are obtained using a typical set of randomly generated $\{\boldsymbol{Q}_{ki}\}$.

That is, Algorithm 1 stops if the improvement in system utility is less than 1% of the system utility achieved in the previous iteration. The simple MRT solution is used to initialize both Algorithm 1 and Algorithm 2. The convex solver CVX [30] is used to solve the convex problems (20) and (36).

Example 1: We first examine the approximation performance of the proposed SCA algorithm, by comparing it with the exhaustive search method in [22]. In view of the tremendous complexity overheads of this exhaustive search method, we consider a simple case where only two transmitter-receiver pairs are present, i.e. $K = 2$, and set $N_t = 4$. Fig. 1(a) shows the simulation results for the comparison of the achievable weighted sum rate between the proposed SCA algorithm and the exhaustive search method against the cross-link interference level η, where the weights are given by $(\alpha_1, \alpha_2) = (\frac{1}{2}, \frac{1}{2})$. Each simulation curve is obtained by averaging over 500 realizations of randomly generated $\{\boldsymbol{Q}_{ki}\}$. From this figure, we can observe that, for $\frac{1}{\sigma^2} = 0$ dB and $\frac{1}{\sigma^2} = 10$ dB, the proposed SCA algorithm can attain almost the same average sum rate performance

as the exhaustive search method, indicating that the proposed SCA algorithm yields near-optimal solutions for the outage constrained beamforming design problem (2). For $\frac{1}{\sigma^2} = 20$ dB, it can be observed that there is a small gap between the rate achieved by the proposed SCA algorithm and that by the exhaustive search method. Nonetheless, this gap is relatively small and is within 2% of the sum rate achieved by the exhaustive search method. Fig. 1(b) displays the simulation results under the same setting as in Fig. 1(a) except that the objective function is now the average harmonic mean rate. As the mean rate performance of SCA algorithm is almost the same as that of the exhaustive search method, its solution is nearly optimal for problem (2).

To examine how the proposed SCA algorithm converges, we illustrate in Fig. 2(a) the trajectories of the optimal rate tuple of problem (20) in each iteration of Algorithm 1, where the weighted sum rate, the geometric mean rate, and the harmonic mean rate are all considered. The user priority weights are set to $(\alpha_1, \alpha_2) = (\frac{1}{2}, \frac{1}{2})$, and the Pareto boundary is obtained by
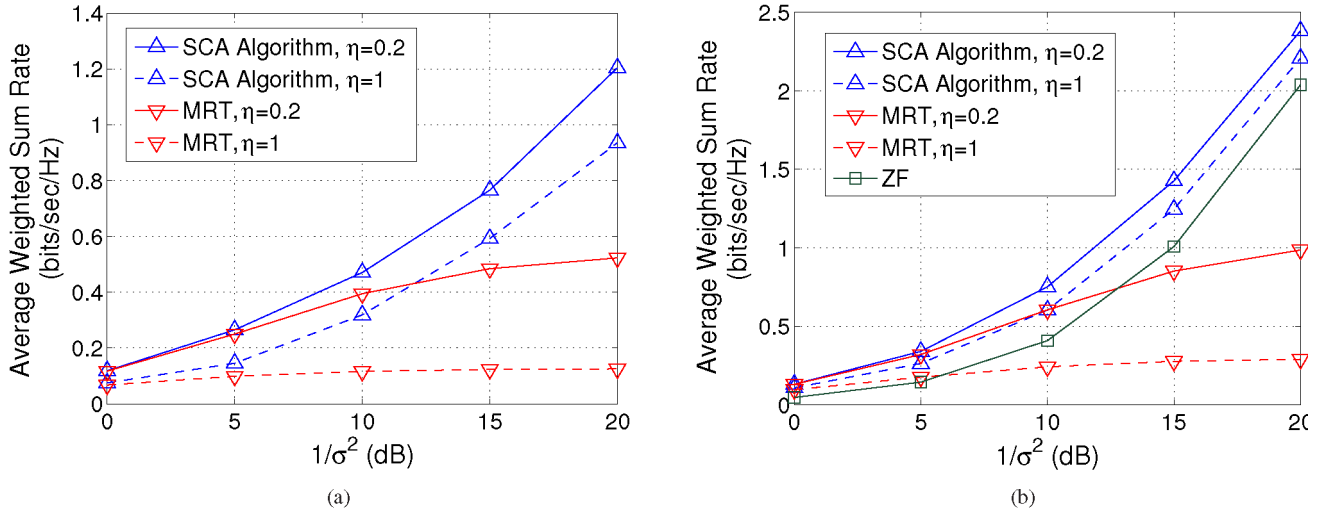
Fig. 3. Simulation results of average achievable sum rate versus $\frac{1}{\sigma^2}$; (a) $K = N_t = 4$, and full rank $\{\boldsymbol{Q}_{ki}\}$, (b) $K = 4$, $N_t = 8$ and $\text{rank}(\boldsymbol{Q}_{ki}) = 2$ for all k, i. The priority weights are set to $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. The results are obtained by averaging over 500 realizations of $\{\boldsymbol{Q}_{ki}\}$.
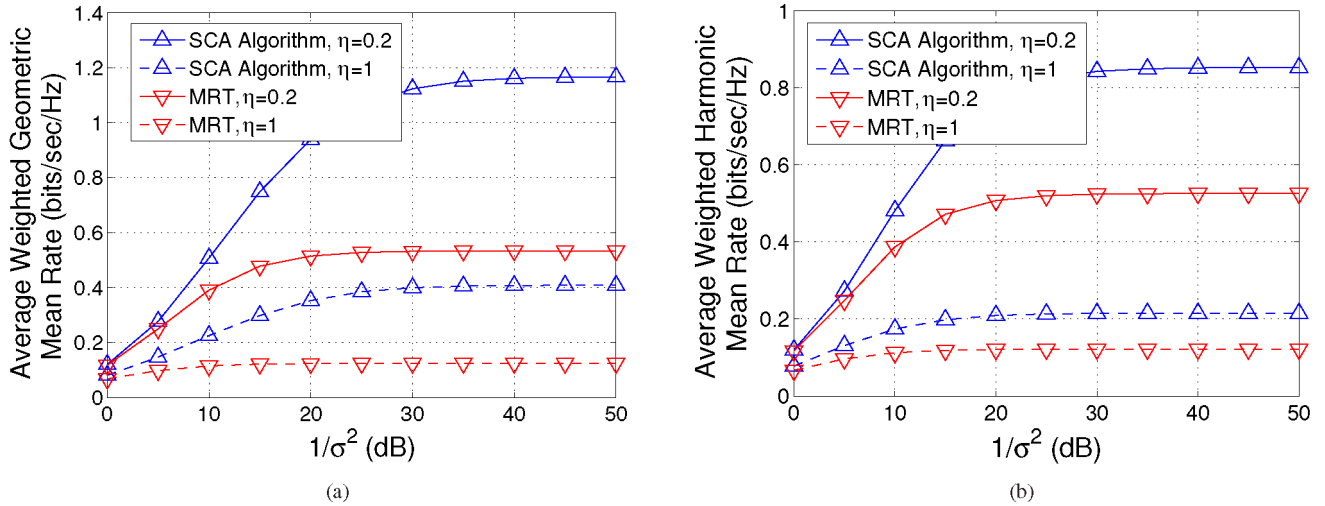

Fig. 4. Simulation results of the proposed SCA algorithm (Algorithm 1), for $K = N_t = 4$ and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2})$; (a) weighted geometric mean rate versus $\frac{1}{\sigma^2}$, (b) weighted harmonic mean rate versus $\frac{1}{\sigma^2}$. Each of the results is obtained by averaging over 500 realizations of $\{\boldsymbol{Q}_{ki}\}$.

the exhaustive search method in [22]. One can see from this figure that, for all rate utility functions, the proposed SCA algorithm first approaches the Pareto boundary and then converges to the corresponding optimal rate tuple along the boundary. In Fig. 2(b), we display similar results with an asymmetric user priority, i.e., $(\alpha_1, \alpha_2) = (\frac{2}{3}, \frac{1}{3})$. It can be observed that the SCA algorithm still converges to the optimal rate tuples in a similar fashion.

Example 2: To further demonstrate the effectiveness of the proposed SCA algorithm, we evaluate the performance of the SCA algorithm for the case of $K = N_t = 4$ in this example. (Since under this setting, the exhaustive search method in [22] is too complex to implement, and, to the best of our knowledge, there is no existing method for comparison, we can only compare the proposed SCA algorithm with the heuristic MRT and ZF schemes.) Fig. 3(a) shows the simulation results of the average achievable sum rate versus $\frac{1}{\sigma^2}$. From this figure, one can observe that the proposed SCA algorithm yields better sum rate performance than the MRT scheme, especially when $\frac{1}{\sigma^2} > 5$ dB. For $\frac{1}{\sigma^2} \leq 5$ dB, the two methods exhibit comparable per-

formance. In Fig. 3(b), we have shown the simulation results for $K = 4$, $N_t = 8$ and $\text{rank}(\boldsymbol{Q}_{ki}) = 2$ for all k, i. Under this setting, the ZF scheme is feasible and its average sum rate performance is also shown in Fig. 3(b). It can be observed from this figure that the ZF scheme outperforms the MRT scheme for high $\frac{1}{\sigma^2}$ or when the cross-link interference is strong ($\eta = 1$). Nevertheless, as can be seen from Fig. 3(b), the proposed SCA algorithm still outperforms both the MRT and the ZF schemes.

Fig. 4 demonstrates the simulation results for the weighted geometric mean rate and the weighted harmonic mean rate, for $K = N_t = 4$ and for an asymmetric weighting $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2})$. Performance comparison results similar to those in Fig. 3 can also be observed in this figure. In addition, it is interesting to note from Fig. 4 that, in contrast to the sum rate performance as shown in Fig. 3, the weighted geometric mean rates and weighted harmonic mean rates achieved by the proposed SCA algorithm in Fig. 4(a) and Fig. 4(b) saturate for high $\frac{1}{\sigma^2}$. These phenomena might result from the fact that user fairness plays a more prominent role in the geometric mean rate and the harmonic mean rate; and
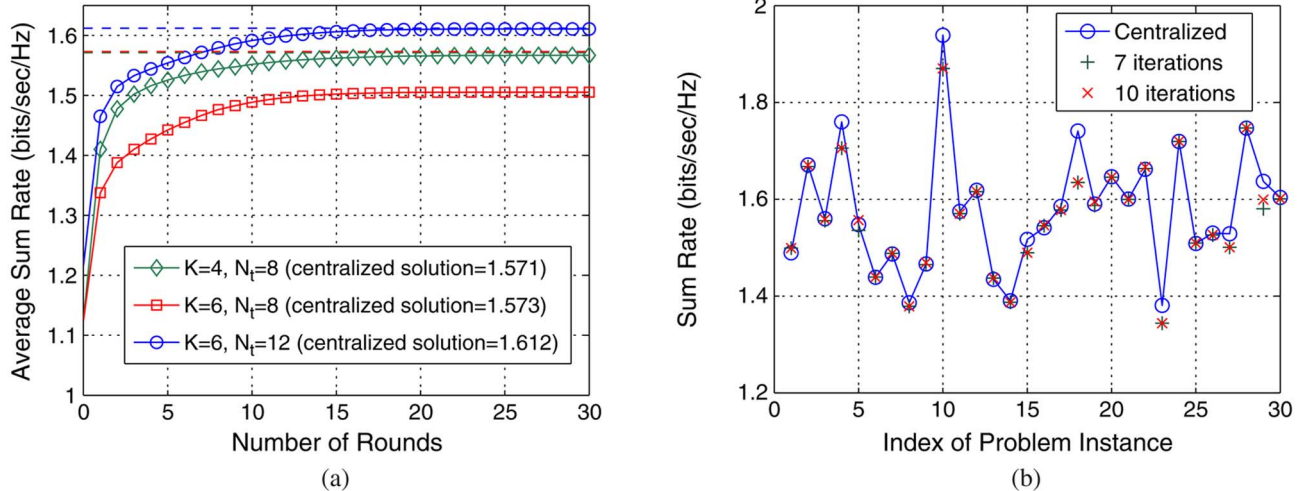
Fig. 5. Performance of Algorithm 2, for $\frac{1}{\sigma^2} = 10$ dB and $\eta = 0.4$; (a) convergence curves versus round number for $N_t = 8$, $K = 4, 6$, and for $N_t = 12$, $K = 6$, averaged over 500 sets of randomly generated $\{Q_{ki}\}$, (b) comparison with Algorithm 1 for $N_t = 8$, $K = 4$ over 30 sets of randomly generated $\{Q_{ki}\}$.

thereby in the interference dominated region (i.e., when $\frac{1}{\sigma^2}$ or η is large), the geometric mean rate and the harmonic mean rate cannot increase as fast as the weighted sum rate.

Example 3: In this example, we examine the performance of the proposed distributed SCA algorithm (Algorithm 2). Fig. 5(a) shows the convergence behaviors (the evolution of sum rate at each round) of the distributed SCA algorithm for $N_t = 8$, $K = 4, 6$, and for $N_t = 12$, $K = 6$, where $\frac{1}{\sigma^2} = 10$ dB, $\eta = 0.4$. Each curve in Fig. 5(a) is obtained by averaging over 500 sets of randomly generated $\{Q_{ki}\}$. It can be observed from Fig. 5(a) that the sum rate performance of the distributed SCA algorithm is almost the same as its centralized counterpart for $N_t = 8$, $K = 4$; whereas there is a gap between the sum rates achieved by the centralized and distributed SCA algorithms for $N_t = 8$, $K = 6$. One explanation for this gap is that, when the system is nearly fully loaded (i.e., when K is close to N_t), the distributed SCA algorithm, which updates only the variables associated with one transmitter at a time, is more likely to get stuck at a stationary point that is not as good as that achieved by the centralized SCA algorithm which optimizes all the variables in each iteration. As also shown in Fig. 5(a), when we increase N_t to 12, the decentralized algorithm again converges to the centralized solution. Fig. 5(b) shows that, for $N_t = 8$, $K = 4$ the distributed SCA algorithm yields performance similar to that achieved by its centralized counterpart for almost all of the 30 tested problem instances within 10 round-robin iterations.

VI. CONCLUSIONS

In this paper, we have presented two efficient approximation algorithms for solving the rate outage constrained coordinated beamforming design problem in (2). In view of the fact that the original design problem involves complicated nonconvex constraints, we first presented an efficient SCA algorithm (Algorithm 1) based on SDR and first-order approximation techniques. We have shown that the proposed SCA algorithm, which involves solving convex problem (20) iteratively, can yield a stationary point of the outage constrained beamforming design problem, provided that problem (20) can yield a rank-one beamforming solution. We further presented a distributed SCA algorithm (Algorithm 2) that can yield approximate beamforming solutions of problem (2) in a distributed, round-robin

fashion, using only local CDI and a small amount of messages exchanged among the transmitters. The distributed SCA algorithm was also shown to provide a stationary point of (2) provided that problem (36) can yield a rank-one beamforming solution. Finally, our simulation results demonstrated that the proposed SCA algorithm yields near-optimal performance for $K = 2$, and significantly outperforms the heuristic MRT and ZF schemes. Furthermore, the distributed SCA algorithm was also shown to exhibit performance comparable to its centralized counterpart within 10 rounds of round-robin iterations for most of the problem instances.

APPENDIX A
PROOF OF CLAIM 1

Since constraint (20e) holds with equality at the optimal point, we have

$$\hat{R}_i[n] = \log_2\left(1 + e^{\bar{y}_i[n-1]}\right) + \frac{e^{\bar{y}_i[n-1]}\left(\hat{y}_i[n] - \bar{y}_i[n-1]\right)}{\ln 2 \cdot \left(1 + e^{\bar{y}_i[n-1]}\right)}$$
$$\leq \log_2\left(1 + e^{\hat{y}_i[n]}\right); \tag{A.1}$$

similarly, from (20c), we have

$$e^{\bar{x}_{ik}[n]} = \mathrm{Tr}\left(\hat{W}_i[n]Q_{ik}\right)$$
$$= e^{\bar{x}_{ik}[n-1]}\left(\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1] + 1\right)$$
$$\leq e^{\hat{x}_{ik}[n]}, \tag{A.2}$$

for all $k \in \mathcal{K}_i^c$, $i = 1, \ldots, K$. We also note from (20d) and (19) that $\hat{x}_{ii}[n] = \bar{x}_{ii}[n]$ for all i, n. On the other hand, by (18), the definition of $\bar{x}_{ik}[n]$, $\bar{y}_i[n]$ in (19), and the fact that (20b), (20f) hold with equality at the optimum, we can obtain

$$1 = \rho_i \exp\left(\sigma_i^2 e^{\bar{y}_i[n] - \bar{x}_{ii}[n]}\right)$$
$$\times \prod_{k \neq i}\left(1 + e^{-\bar{x}_{ii}[n] + \bar{x}_{ki}[n] + \bar{y}_i[n]}\right)$$
$$= \rho_i \exp\left(\sigma_i^2 e^{\hat{y}_i[n] - \hat{x}_{ii}[n]}\right)$$
$$\times \prod_{k \neq i}\left(1 + e^{-\hat{x}_{ii}[n] + \hat{x}_{ki}[n] + \hat{y}_i[n]}\right). \tag{A.3}$$

Combining the above observations, i.e., (A.2), (A.3) and $\hat{x}_{ii}[n] = \bar{x}_{ii}[n]$, and by the monotonicity of the exponential function, we obtain that $\bar{y}_i[n] \geq \hat{y}_i[n]$, which implies

$$\hat{R}_i[n] \leq \log_2\left(1 + e^{\hat{y}_i[n]}\right) \leq \log_2\left(1 + e^{\bar{y}_i[n]}\right) = \tilde{R}_i[n] \; \forall i, n. \tag{A.4}$$

Suppose that $e^{\hat{x}_{ki}[n]} - e^{\bar{x}_{ki}[n]}$ does not converge to zero for some i and $k \in \mathcal{K}_i^c$. Then there exists an $\epsilon > 0$ such that, for all $N \geq 1$, $e^{\hat{x}_{ki}[n]} > e^{\bar{x}_{ki}[n]} + \epsilon$ for some $n \geq N$. From (A.1) to (A.4), we must have $e^{\bar{y}_i[n]} > e^{\hat{y}_i[n]} + \epsilon'$ and thus $\tilde{R}_i[n] > \hat{R}_i[n] + \epsilon''$, where $\epsilon', \epsilon'' > 0$, which, together with (22), implies that the utility $U(\tilde{R}_1[n], \ldots, \tilde{R}_K[n])$ diverges as n goes to infinity however. Therefore, we must have

$$\lim_{n\to\infty}\left(e^{\hat{x}_{ik}[n]} - e^{\bar{x}_{ik}[n]}\right) = 0 \; \forall i, k, \tag{A.5}$$

$$\lim_{n\to\infty}\left(\tilde{R}_i[n] - \hat{R}_i[n]\right) = 0 \; \forall i. \tag{A.6}$$

Now we use (A.5) to prove (31). It follows from (A.2) and (A.5) that

$$\lim_{n\to\infty}\left(e^{\hat{x}_{ik}[n]} - e^{\bar{x}_{ik}[n-1]}\left(\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1] + 1\right)\right) = 0 \tag{A.7}$$

for all i and $k \in \mathcal{K}_i^c$. Consider the 2nd-order Taylor series expansion [34] of $e^{\hat{x}_{ik}[n]}$ at $\bar{x}_{ik}[n-1]$, i.e.,

$$e^{\hat{x}_{ik}[n]} = e^{\bar{x}_{ik}[n-1]}\left(\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1] + 1\right) \\ + e^{\theta[n]\hat{x}_{ik}[n] + (1-\theta[n])\bar{x}_{ik}[n-1]}\left(\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1]\right)^2,$$

where $0 \leq \theta[n] \leq 1$ for all $n \geq 1$. Substituting it into (A.7) gives rise to

$$\lim_{n\to\infty} e^{\theta[n]\hat{x}_{ik}[n] + (1-\theta[n])\bar{x}_{ik}[n-1]}\left(\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1]\right)^2 = 0.$$

Since both $\bar{x}_{ik}[n]$ and $\hat{x}_{ik}[n]$ are bounded by Claim 2, we conclude that (31) is true.

To show (32), we note from (A.1), (A.4) and (A.6) that

$$\lim_{n\to\infty}\Bigg(\ln\left(1 + \exp\left(\hat{y}_i[n]\right)\right) - \ln\left(1 + \exp\left(\bar{y}_i[n-1]\right)\right) \\ - \frac{\exp\left(\bar{y}_i[n-1]\right)}{1 + \exp\left(\bar{y}_i[n-1]\right)}\left(\hat{y}_i[n] - \bar{y}_i[n-1]\right)\Bigg) = 0. \tag{A.8}$$

Analogously, by considering the 2nd-order Taylor series expansion of $\ln(1 + e^{\hat{y}_i[n]})$ at $\bar{y}_i[n-1]$, i.e.,

$$\ln\left(1 + e^{\hat{y}_i[n]}\right) \\ = \ln\left(1 + e^{\bar{y}_i[n-1]}\right) \\ + \frac{\exp\left(\bar{y}_i[n-1]\right)}{1 + \exp\left(\bar{y}_i[n-1]\right)}\left(\hat{y}_i[n] - \bar{y}_i[n-1]\right) \\ + \frac{\exp\left(\theta[n]\hat{y}_i[n] + (1-\theta[n])\bar{y}_i[n-1]\right)}{\left(1 + \exp\left(\theta[n]\hat{y}_i[n] + (1-\theta[n])\bar{y}_i[n-1]\right)\right)^2} \\ \times \left(\hat{y}_i[n] - \bar{y}_i[n-1]\right)^2,$$

where $0 \leq \theta[n] \leq 1$ for all $n \geq 1$, and substituting it into (A.8), we obtain

$$\lim_{n\to\infty} \frac{\exp(\theta[n]\hat{y}_i[n] + (1-\theta[n])\bar{y}_i[n-1])(\hat{y}_i[n] - \bar{y}_i[n-1])^2}{(1 + \exp(\theta[n]\hat{y}_i[n] + (1-\theta[n])\bar{y}_i[n-1]))^2} = 0.$$

Again, since $\bar{y}_i[n]$ and $\hat{y}_i[n]$ are bounded by Claim 2, we obtain (32). ∎

APPENDIX B
PROOF OF THEOREM 2

Define $\bar{z}_k[n, i-1] = e^{\bar{y}_k[n,i-1] - \bar{x}_{kk}[n-u_{k(i-1)}]}$ for all $k = 1, \ldots, K$. Then it can be shown that

$$\bar{\boldsymbol{u}}[n-1, i] \triangleq \Big(\hat{\boldsymbol{W}}_i[n-1], \{\tilde{R}_k[n, i-1]\}_k, \{\bar{x}_{ik}[n-1]\}_k, \\ \{\bar{y}_k[n, i-1]\}_k, \{\bar{z}_k[n, i-1]\}_k \Big),$$

is a feasible point of (36). Hence,

$$U(\hat{R}_1[n, i], \ldots, \hat{R}_K[n, i]) \geq U(\tilde{R}_1[n, i-1], \ldots, \tilde{R}_K[n, i-1])$$

for all $i = 1, \ldots, K$. In addition, analogous to (15), we have $\tilde{R}_j[n, i] \geq \hat{R}_j[n, i]$ for all i, j, n, and thus

$$U(\tilde{R}_1[n, i], \ldots, \tilde{R}_K[n, i]) \geq U(\tilde{R}_1[n, i-1], \ldots, \tilde{R}_K[n, i-1]),$$

for all $i = 1, \ldots, K$, which implies that the sequence $\{U(\tilde{R}_1[1, 1], \ldots, \tilde{R}_K[1, 1]), \ldots, U(\tilde{R}_1[1, K], \ldots, \tilde{R}_K[1, K]), U(\tilde{R}_1[2, 1], \ldots, \tilde{R}_K[2, 1]), \ldots\}$ is nondecreasing. Since it is also bounded, $U(\tilde{R}_1[n, i], \ldots, \tilde{R}_K[n, i])$, $i = 1, \ldots, K$, converge as $n \to \infty$.

Now let us look at the KKT conditions of problem (36). Recall the definitions of $\bar{\Psi}_{ki}(\cdot)$ and $\bar{\Phi}_j(\cdot)$ in (23) and (24) and their inner approximation properties in (25) to (30). Let

$$\Theta_i^{[i]}\left(x_{ii}, y_i, z_i, \{\bar{x}_{ki}[n-u_{ki}]\}_{k\neq i}\right) \\ \triangleq \rho_i e^{\sigma_i^2 z_i} \prod_{k\neq i}\left(1 + e^{-x_{ii} + \bar{x}_{ki}[n-u_{ki}] + y_i}\right) - 1, \tag{A.9}$$

$$\Theta_j^{[i]}\left(x_{ij}, y_j, z_j, \{\bar{x}_{kj}[n-u_{ki}]\}_{k\neq i}\right) \\ \triangleq \rho_j e^{\sigma_j^2 z_j}\left(1 + e^{-\bar{x}_{jj}[n-u_{ji}] + x_{ij} + y_j}\right) \\ \times \prod_{\substack{k\neq j \\ k\neq i}}\left(1 + e^{-\bar{x}_{jj}[n-u_{ji}] + \bar{x}_{kj}[n-u_{ki}] + y_j}\right) - 1, \quad j \in \mathcal{K}_i^c. \tag{A.10}$$

Moreover, let

$$\hat{\boldsymbol{u}}[n, i] \triangleq \Big(\hat{\boldsymbol{W}}_i[n], \{\hat{R}_k[n, i]\}, \{\hat{x}_{ik}[n]\}_k, \{\hat{y}_k[n, i]\}, \{\hat{z}_k[n, i]\} \Big)$$

be the optimal solution of (36), and let

$$\boldsymbol{\lambda}[n, i] \triangleq \Big(\lambda_i^{\mathrm{b}}[n, i], \{\lambda_k^{\mathrm{b}}[n, i]\}_{k\neq i}, \lambda^{\mathrm{d}}[n, i], \\ \{\lambda_k^{\mathrm{e}}[n, i]\}_{k\neq i}, \{\lambda_k^{\mathrm{f}}[n, i]\}_k, \lambda_i^{\mathrm{g}}[n, i], \\ \{\lambda_k^{\mathrm{g}}[n, i]\}_{k\neq i}, \lambda^P[n, i], \{\lambda_k^\delta[n, i]\}_k \Big) \succeq \boldsymbol{0},$$

where $\lambda_i^{\mathrm{b}}[n, i]$, $\{\lambda_k^{\mathrm{b}}[n, i]\}_{k\neq i}$, $\lambda^{\mathrm{d}}[n, i]$, $\{\lambda_k^{\mathrm{e}}[n, i]\}_{k\neq i}$, $\{\lambda_k^{\mathrm{f}}[n, i]\}_k$, $\lambda_i^{\mathrm{g}}[n, i]$, $\{\lambda_k^{\mathrm{g}}[n, i]\}_{k\neq i}$ denote the dual variables associated with constraints in (36b) to (36g), and $\lambda^P[n, i]$, $\lambda_k^\delta[n, i]$, denote the dual variables associated with constraint $\mathrm{Tr}(\boldsymbol{W}_i) \leq P_i$ and $\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{Q}_{ik}) \geq \delta$, respectively. Let $\mathcal{L}^{[i]}(\hat{\boldsymbol{u}}[n, i], \boldsymbol{\lambda}[n, i])$ be the Lagrangian function. The KKT conditions of (36) are given by (A.11) and (A.12) shown on the the next page. Note that we have omitted the complementary slackness conditions for constraints (36b)–(36g) since they are trivially satisfied at $\hat{\boldsymbol{u}}[n, i]$.

To show the desired results, we also need the following two claims:

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial \boldsymbol{W}_i}$$

$$= \lambda^P[n,i]\mathbf{I}_{N_t} - \left(\lambda^{\mathrm{d}}[n,i] + \lambda_i^{\delta}[n,i]\right)\boldsymbol{Q}_{ii} + \sum_{k \neq i}\left(\lambda_k^{\mathrm{e}}[n,i]\frac{\partial \bar{\Psi}_{ik}\left(\hat{\boldsymbol{W}}_i[n], \hat{x}_{ik}[n]|\bar{x}_{ik}[n-1]\right)}{\partial \boldsymbol{W}_i} - \lambda_k^{\delta}[n,i]\boldsymbol{Q}_{ik}\right) \succeq \mathbf{0}, \quad \text{(A.11a)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial R_j}$$

$$= -\frac{\partial U\left(\hat{R}_1[n,i],\ldots,\hat{R}_K[n,i]\right)}{\partial R_j} + \lambda_j^{\mathrm{f}}[n,i]\frac{\partial \bar{\Phi}_j\left(\hat{R}_j[n,i], \hat{y}_j[n,i]|\bar{y}_j[n,i-1]\right)}{\partial R_j} \geq 0 \;\forall j, \quad \text{(A.11b)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial x_{ii}}$$

$$= \lambda_i^{\mathrm{b}}[n,i]\frac{\partial \Theta_i^{[i]}\left(\hat{x}_{ii}[n], \hat{y}_i[n,i], \hat{z}_i[n,i], \{\bar{x}_{ki}[n-u_{ki}]\}_{k \neq i}\right)}{\partial x_{ii}} + \lambda^{\mathrm{d}}[n,i]e^{\hat{x}_{ii}[n]} - \lambda_i^{\mathrm{g}}[n,i]e^{\hat{y}_i[n,i]-\hat{x}_{ii}[n]} = 0, \quad \text{(A.11c)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial x_{ij}}$$

$$= \lambda_j^{\mathrm{b}}[n,i]\frac{\partial \Theta_j^{[i]}\left(\hat{x}_{ij}[n], \hat{y}_j[n,i], \hat{z}_j[n,i], \{\bar{x}_{kj}[n-u_{ki}]\}_{k \neq i}\right)}{\partial x_{ij}} + \lambda_j^{\mathrm{e}}[n,i]\frac{\partial \bar{\Psi}_{ij}\left(\hat{\boldsymbol{W}}_i[n], \hat{x}_{ij}[n]|\bar{x}_{ij}[n-1]\right)}{\partial x_{ij}} = 0 \;\forall j \in \mathcal{K}_i^c,$$

$$\text{(A.11d)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial y_i}$$

$$= \lambda_i^{\mathrm{b}}[n,i]\frac{\partial \Theta_i^{[i]}\left(\hat{x}_{ii}[n], \hat{y}_i[n,i], \hat{z}_i[n,i], \{\bar{x}_{ki}[n-u_{ki}]\}_{k \neq i}\right)}{\partial y_i} + \lambda_i^{\mathrm{f}}[n,i]\frac{\partial \bar{\Phi}_i\left(\hat{R}_i[n,i], \hat{y}_i[n,i]|\bar{y}_i[n,i-1]\right)}{\partial y_i}$$

$$+ \lambda_i^{\mathrm{g}}[n,i]e^{\hat{y}_i[n,i]-\hat{x}_{ii}[n]} = 0, \quad \text{(A.11e)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial y_j}$$

$$= \lambda_j^{\mathrm{b}}[n,i]\frac{\partial \Theta_j^{[i]}\left(\hat{x}_{ij}[n], \hat{y}_j[n,i], \hat{z}_j[n,i], \{\bar{x}_{kj}[n-u_{ki}]\}_{k \neq i}\right)}{\partial y_j} + \lambda_j^{\mathrm{f}}[n,i]\frac{\partial \bar{\Phi}_j\left(\hat{R}_j[n,i], \hat{y}_j[n,i]|\bar{y}_j[n,i-1]\right)}{\partial y_j}$$

$$+ \lambda_j^{\mathrm{g}}[n,i]e^{\hat{y}_j[n,i]-\bar{x}_{jj}[n-u_{ji}]} = 0 \;\forall j \in \mathcal{K}_i^c, \quad \text{(A.11f)}$$

$$\frac{\partial \mathcal{L}^{(i)}\left(\boldsymbol{u}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial z_i}$$

$$= \lambda_i^{\mathrm{b}}[n,i]\frac{\partial \Theta_i^{[i]}\left(\hat{x}_{ii}[n], \hat{y}_i[n,i], \hat{z}_i[n,i], \{\bar{x}_{ki}[n-u_{ki}]\}_{k \neq i}\right)}{\partial z_i} - \lambda_i^{\mathrm{g}}[n,i] = 0, \quad \text{(A11.g)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial z_j}$$

$$= \lambda_j^{\mathrm{b}}[n,i]\frac{\partial \Theta_j^{[i]}\left(\hat{x}_{ij}[n], \hat{y}_j[n,i], \hat{z}_j[n,i], \{\bar{x}_{kj}[n-u_{ki}]\}_{k \neq i}\right)}{\partial z_j} - \lambda_j^{\mathrm{g}}[n,i] = 0, \; j \in \mathcal{K}_i^c, \quad \text{(A.11h)}$$

$$\lambda^P[n,i] \cdot \left(\mathrm{Tr}\left(\hat{\boldsymbol{W}}_i[n]\right) - P_i\right) = 0, \quad \frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial R_j}\hat{R}_j[n,i] = 0 \;\forall j, \quad \text{(A.12a)}$$

$$\frac{\partial \mathcal{L}^{[i]}\left(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i]\right)}{\partial \boldsymbol{W}_i} \cdot \hat{\boldsymbol{W}}_i[n] = \mathbf{0}, \quad \lambda_j^{\delta}[n,i] \cdot \left(\delta - \mathrm{Tr}\left(\hat{\boldsymbol{W}}_i[n]\boldsymbol{Q}_{ik}\right)\right) = 0 \;\forall j. \quad \text{(A.12b)}$$

Claim 3: *It holds true that*

$$\lim_{n \to \infty} |\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1]| = 0 \; \forall i,k, \tag{A.13a}$$

$$\lim_{n \to \infty} |\hat{x}_{ik}[n] - \bar{x}_{ik}[n]| = 0 \; \forall i,k, \tag{A.13b}$$

$$\begin{cases} \lim_{n \to \infty} |\hat{y}_k[n,1] - \bar{y}_k[n-1,K]| = 0 \\ \lim_{n \to \infty} |\hat{y}_k[n,i] - \bar{y}_k[n,i-1]| = 0 \end{cases} \forall i,k, \tag{A.13c}$$

$$\lim_{n \to \infty} |\hat{y}_k[n,i] - \bar{y}_k[n,i]| = 0 \; \forall i,k, \tag{A.13d}$$

$$\begin{cases} \lim_{n \to \infty} \left| \tilde{R}_k[n,1] - \tilde{R}_k[n-1,K] \right| = 0 \\ \lim_{n \to \infty} \left| \tilde{R}_k[n,i] - \tilde{R}_k[n,i-1] \right| = 0 \end{cases} \forall i,k, \tag{A.13e}$$

$$\lim_{n \to \infty} \left| \hat{R}_k[n,i] - \tilde{R}_k[n,i] \right| = 0 \; \forall i,k. \tag{A.13f}$$

Claim 4: *For each i, $\hat{\boldsymbol{u}}[n,i]$ generated by Algorithm 2 is bounded for all n.*

The proof of Claim 3 is presented in Appendix C. Similar to Claim 1, (A.13a) to (A.13d) imply that the restrictive approximations in (36e) and (36f) are asymptotically tight as $n \to \infty$. Since problem (36) satisfies the Slater's condition, the dual variable vector $\boldsymbol{\lambda}[n,i]$ is bounded [33]. Moreover, $\hat{\boldsymbol{u}}[n,i]$ is also bounded by Claim 4. Now let us consider the primal-dual solution pair $(\hat{\boldsymbol{u}}[n,i], \boldsymbol{\lambda}[n,i])$ for all $i = 1, \ldots, K$. Since they are all bounded, there exists a subsequence $\{n_1, \ldots, n_\ell, \ldots\} \subseteq \{1, \ldots, n, \ldots\}$ and limit points $\hat{\boldsymbol{u}}^\star[i] \triangleq (\hat{\boldsymbol{W}}_i^\star, \{\hat{R}_k^\star[i]\}, \{\hat{x}_{ik}^\star\}_k, \{\hat{y}_k^\star[i]\}, \{\hat{z}_k^\star[i]\})$ and $\boldsymbol{\lambda}^\star[i] \triangleq (\lambda_i^{\mathrm{b}\star}[i], \{\lambda_k^{\mathrm{b}\star}[i]\}_{k \neq i}, \lambda^{\mathrm{d}\star}[i], \{\lambda_k^{\mathrm{e}\star}[i]\}_{k \neq i}, \{\lambda_k^{\mathrm{f}\star}[i]\}_k, \lambda_i^{\mathrm{g}\star}[i], \{\lambda_k^{\mathrm{g}\star}[i]\}_{k \neq i}, \lambda^{P\star}[i], \{\lambda_k^{\delta\star}[i]\}_k) \succeq \boldsymbol{0}$ for all i, such that

$$\lim_{\ell \to \infty} \hat{\boldsymbol{u}}[n_\ell, i] = \hat{\boldsymbol{u}}^\star[i], \quad \lim_{\ell \to \infty} \boldsymbol{\lambda}[n_\ell, i] = \boldsymbol{\lambda}^\star[i] \tag{A.14}$$

for $i = 1, \ldots, K$. By (A.13e) and (A.13f), we see that both $\hat{R}_k[n_\ell, i]$ and $\tilde{R}_k[n_\ell, i]$ converge to the same limit point, and they are the same for all i, i.e.,

$$\hat{R}_k^\star[1] = \hat{R}_k^\star[2] = \cdots = \hat{R}_k^\star[K] \triangleq \tilde{R}_k^\star, \; k = 1, \ldots, K. \tag{A.15}$$

Analogously, by (A.13a) to (A.13d), we have that

$$\hat{y}_k^\star[1] = \hat{y}_k^\star[2] = \cdots = \hat{y}_k^\star[K] \triangleq \hat{y}_k^\star, \; k = 1, \ldots, K, \tag{A.16}$$

$$\hat{z}_k^\star[1] = \hat{z}_k^\star[2] = \cdots = \hat{z}_k^\star[K] \triangleq \hat{z}_k^\star, \; k = 1, \ldots, K. \tag{A.17}$$

Then, it follows from the inner approximation properties in (25) to (30), (A.13a), (A.13c), and (A.14) to (A.17) that the KKT conditions in (A.11) and (A.12) converge along the subsequence $\{n_1, \ldots, n_\ell, \ldots\}$ to

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial \boldsymbol{W}_i}$$
$$= \lambda^{P\star}[i] \mathbf{I}_{N_t} - \left(\lambda^{\mathrm{d}\star}[i] + \lambda_i^{\delta\star}[i] \right) \boldsymbol{Q}_{ii}$$
$$+ \sum_{k \neq i} \left(\lambda_k^{\mathrm{e}\star}[i] \frac{\partial \Psi_{ik} \left(\hat{\boldsymbol{W}}_i^\star, \hat{x}_{ik}^\star \right)}{\partial \boldsymbol{W}_i} - \lambda_k^{\delta\star}[i] \boldsymbol{Q}_{ik} \right) \succeq \boldsymbol{0}, \tag{A.18a}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial R_j}$$
$$= -\frac{\partial U \left(\tilde{R}_1^\star, \ldots, \tilde{R}_K^\star \right)}{\partial R_j} + \lambda_j^{\mathrm{f}\star}[i] \frac{\partial \Phi_j \left(\tilde{R}_j^\star, \hat{y}_j^\star \right)}{\partial R_j} \geq 0 \; \forall j, \tag{A.18b}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial x_{ii}}$$
$$= \lambda_i^{\mathrm{b}\star}[i] \frac{\partial \Theta_i^{[i]} \left(\hat{x}_{ii}^\star, \hat{y}_i^\star, \hat{z}_i^\star, \{\hat{x}_{ki}^\star\}_{k \neq i} \right)}{\partial x_{ii}}$$
$$+ \lambda^{\mathrm{d}\star}[i] e^{\hat{x}_{ii}^\star} - \lambda_i^{\mathrm{g}\star}[i] e^{\hat{y}_i^\star - \hat{x}_{ii}^\star} = 0, \tag{A.18c}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial x_{ij}}$$
$$= \lambda_j^{\mathrm{b}\star}[i] \frac{\partial \Theta_j^{[i]} \left(\hat{x}_{ij}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \{\hat{x}_{kj}^\star\}_{k \neq i} \right)}{\partial x_{ij}}$$
$$+ \lambda_j^{\mathrm{e}\star}[i] \frac{\partial \Psi_{ij} \left(\hat{\boldsymbol{W}}_i^\star, \hat{x}_{ij}^\star \right)}{\partial x_{ij}} = 0 \; \forall j \in \mathcal{K}_i^c, \tag{A.18d}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial y_j}$$
$$= \lambda_j^{\mathrm{b}\star}[i] \frac{\partial \Theta_j^{[i]} \left(\hat{x}_{ij}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \{\hat{x}_{kj}^\star\}_{k \neq i} \right)}{\partial y_j}$$
$$+ \lambda_j^{\mathrm{f}\star}[i] \frac{\partial \Phi_j \left(\tilde{R}_j^\star, \hat{y}_j^\star \right)}{\partial y_j} + \lambda_j^{\mathrm{g}\star}[i] e^{\hat{y}_j^\star - \hat{x}_{jj}^\star} = 0 \; \forall j, \tag{A.18e}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial z_j}$$
$$= \lambda_j^{\mathrm{b}\star}[i] \frac{\partial \Theta_j^{[i]} \left(\hat{x}_{ij}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \{\hat{x}_{kj}^\star\}_{k \neq i} \right)}{\partial z_j} - \lambda_j^{\mathrm{g}\star}[i] = 0 \; \forall j, \tag{A.18f}$$

and

$$\lambda^{\mathrm{i}\star}[i] \cdot \left(\mathrm{Tr} \left(\hat{\boldsymbol{W}}_i^\star \right) - P_i \right) = 0, \tag{A.19a}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial R_j} \tilde{R}_j^\star = 0 \; \forall j, \tag{A.19b}$$

$$\frac{\partial \mathcal{L}^{[i]} \left(\hat{\boldsymbol{u}}^\star[i], \boldsymbol{\lambda}^\star[i] \right)}{\partial \boldsymbol{W}_i} \cdot \hat{\boldsymbol{W}}_i^\star = \boldsymbol{0}, \tag{A.19c}$$

$$\lambda_j^{\delta\star}[i] \cdot \left(\delta - \mathrm{Tr} \left(\hat{\boldsymbol{W}}_i^\star \boldsymbol{Q}_{ik} \right) \right) = 0 \; \forall j. \tag{A.19d}$$

It can be observed from (39) that, for $\rho_i < 1$, $\tilde{R}_j[n,i]$ is strictly greater than zero for all i, j, n; therefore, $\tilde{R}_j^\star > 0$ for all j, which indicates that $\frac{\partial \mathcal{L}^{[i]}(\hat{\boldsymbol{u}}^\star, \hat{\boldsymbol{\lambda}}^\star[i])}{\partial R_j} = 0$ for all i, j by (A.19b). Substituting this into (A.18b) for all $i = 1, \ldots, K$, gives rise to

$$\lambda_j^{\mathrm{f}\star}[1] = \cdots = \lambda_j^{\mathrm{f}\star}[K]$$
$$= \frac{\partial U \left(\tilde{R}_1^\star, \ldots, \tilde{R}_K^\star \right)}{\partial R_j} \left(\frac{\partial \Phi_j \left(\tilde{R}_j^\star, \hat{y}_j^\star \right)}{\partial R_j} \right)^{-1}$$
$$\triangleq \lambda_j^{\mathrm{f}\star}, \; j = 1, \ldots, K. \tag{A.20}$$

In addition, one can verify that

$$
\frac{\partial \Theta_j^{[1]}\left(\hat{x}_{1j}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \left\{\hat{x}_{kj}^\star\right\}_{k \neq 1}\right)}{\partial y_j} = \cdots
$$

$$
= \frac{\partial \Theta_j^{[K]}\left(\hat{x}_{Kj}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \left\{\hat{x}_{kj}^\star\right\}_{k \neq K}\right)}{\partial y_j},
$$

$$
\frac{\partial \Theta_j^{[1]}\left(\hat{x}_{1j}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \left\{\hat{x}_{kj}^\star\right\}_{k \neq 1}\right)}{\partial z_j} = \cdots
$$

$$
= \frac{\partial \Theta_j^{[K]}\left(\hat{x}_{Kj}^\star, \hat{y}_j^\star, \hat{z}_j^\star, \left\{\hat{x}_{kj}^\star\right\}_{k \neq K}\right)}{\partial z_j},
$$

which, together with (A.18e) (A.18f) and (A.20), lead to

$$
\lambda_j^{\mathrm{b}\star}[1] = \cdots = \lambda_j^{\mathrm{b}\star}[K] \triangleq \lambda_j^{\mathrm{b}\star}, \qquad (\text{A.21a})
$$

$$
\lambda_j^{\mathrm{g}\star}[1] = \cdots = \lambda_j^{\mathrm{g}\star}[K] \triangleq \lambda_j^{\mathrm{g}\star} \ \forall j. \qquad (\text{A.21b})
$$

Finally, by (A.18), (A.19), (A.20) and (A.21), we conclude that $(\{\hat{\boldsymbol{W}}_i^\star\}, \{\tilde{R}_k^\star\}, \{\hat{x}_{ik}^\star\}_k, \{\hat{y}_k^\star\}, \{\hat{z}_k^\star\})$ and $(\{\lambda_i^\star\}, \{\lambda^{\mathrm{d}\star}[i]\}, \{\{\lambda_k^{\mathrm{e}\star}[i]\}_{k \neq i}\}_i, \{\lambda_k^{\mathrm{f}\star}\}, \{\lambda_j^{\mathrm{g}\star}\}, \{\lambda^{P\star}[i]\}, \{\lambda^{\delta\star}[i]\})$ satisfy the KKT conditions of problem (8). The proof is completed. ∎

APPENDIX C
PROOF OF CLAIM 3

The ideas of the proof are similar to that of Claim 1. Because constraints (36e) and (36f) hold with equality at the optimum, we have

$$
e^{\bar{x}_{ik}[n]} \leq e^{\hat{x}_{ik}[n]}, \ \hat{R}_k[n,i] \leq \frac{1}{\ln 2}\ln\left(1 + e^{\hat{y}_k[n,i]}\right) \quad (\text{A.22})
$$

for all $k \in \mathcal{K}_i^c, i = 1, \ldots, K$. Also by (36c), (36g) and (39), we have

$$
1 = \rho_j \exp\left(\sigma_j^2 e^{\hat{y}_j[n,i] - \bar{x}_{jj}[n - u_{ji}]}\right)
$$

$$
\times \left(1 + e^{-\bar{x}_{jj}[n - u_{ji}] + \hat{x}_{ij}[n] + \hat{y}_j[n,i]}\right)
$$

$$
\times \prod_{\substack{k \neq j \\ k \neq i}} \left(1 + e^{-\bar{x}_{jj}[n - u_{ji}] + \bar{x}_{kj}[n - u_{ki}] + \hat{y}_j[n,i]}\right)
$$

$$
= \rho_j \exp\left(\sigma_j^2 e^{\bar{y}_i[n,i] - \bar{x}_{ii}[n - u_{ji}]}\right)
$$

$$
\times \left(1 + e^{-\bar{x}_{jj}[n - u_{ji}] + \hat{x}_{ij}[n] + \bar{y}_j[n,i]}\right)
$$

$$
\times \prod_{\substack{k \neq j \\ k \neq i}} \left(1 + e^{-\bar{x}_{jj}[n - u_{ji}] + \bar{x}_{kj}[n - u_{ki}] + \bar{y}_j[n,i]}\right),
$$

for all $j \in \mathcal{K}_i^c$. Using the above equation and (A.22) and the monotonicity of exponential function, we obtain $\hat{y}_j[n,i] \leq \bar{y}_j[n,i]$. Thus,

$$
\hat{R}_j[n,i] \leq \frac{1}{\ln 2}\ln\left(1 + e^{\hat{y}_j[n,i]}\right)
$$

$$
\leq \frac{1}{\ln 2}\ln\left(1 + e^{\bar{y}_j[n,i]}\right) = \tilde{R}_j[n,i] \ \forall j \in \mathcal{K}_i^c, \quad (\text{A.23})
$$

Similarly, by (36b), (36g) and (39), we have

$$
\hat{R}_i[n,i] \leq \frac{1}{\ln 2}\ln\left(1 + e^{\hat{y}_i[n,i]}\right)
$$

$$
= \frac{1}{\ln 2}\ln\left(1 + e^{\bar{y}_i[n,i]}\right) = \tilde{R}_i[n,i]. \quad (\text{A.24})
$$

Using the same arguments as in obtaining (A.3) to (A.6) in Appendix A, we can show that (A.13f), (A.13b), (A.13d), (A.13c) and

$$
\lim_{n \to \infty} |\hat{x}_{ik}[n] - \bar{x}_{ik}[n-1]| = 0 \ \forall i, \ k \in \mathcal{K}_i^c, \quad (\text{A.25})
$$

which is (A.13a) for $k \neq i$, are true. What remains is to prove (A.13e) and $\lim_{n \to \infty} |\hat{x}_{ii}[n] - \bar{x}_{ii}[n-1]| = 0 \ \forall i$.

It follows from (A.13c), (A.13d) and the triangle inequality that

$$
\lim_{n \to \infty} |\bar{y}_k[n,1] - \bar{y}_k[n-1,K]| = 0,
$$

$$
\lim_{n \to \infty} |\bar{y}_k[n,i] - \bar{y}_k[n,i-1]| = 0 \ \forall i, k,
$$

which, by the definition in (38), is equivalent to (A.13e). By considering (39) for transmitter $i - 1$, and the fact that (36b) holds with equality at the optimal point for transmitter i, we can obtain

$$
1 = \rho_i \exp\left(\sigma_i^2 e^{\bar{y}_i[n,i-1] - \bar{x}_{ii}[n-1]}\right)
$$

$$
\times \prod_{k \neq i}\left(1 + e^{-\bar{x}_{ii}[n-1] + \bar{x}_{ki}[n - u_{ki}] + \bar{y}_i[n,i-1]}\right)
$$

$$
= \rho_i \exp\left(\sigma_i^2 e^{\hat{y}_i[n,i] - \hat{x}_{ii}[n]}\right)
$$

$$
\times \prod_{k \neq i}\left(1 + e^{-\hat{x}_{ii}[n] + \bar{x}_{ki}[n - u_{ki}] + \hat{y}_i[n,i]}\right) \ \forall i.
$$

Since both $\{\hat{y}_i[n,i]\}_{n=1}^\infty$ and $\{\bar{y}_i[n, i-1]\}_{n=1}^\infty$ are bounded, and by (A.13c), we obtain from the above equation that

$$
\lim_{n \to \infty} |\hat{x}_{ii}[n] - \bar{x}_{ii}[n-1]| = 0 \ \forall i.
$$

Thus the proof of Claim 3 is complete. ∎

REFERENCES

[1] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi, "A convex approximation approach to weighted sum rate maximization of multiuser MISO interference channel under outage constraints," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 22–27, 2011, pp. 3368–3371.

[2] 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, pp. 155–165, Jan. 2008.

[3] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1380–1408, Dec. 2010.

[4] E. Björnson, N. J. Jaldén, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Trans. Signal Process.*, vol. 59, pp. 6086–6101, Dec. 2011.

[5] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1748–1759, May 2010.

[6] L. Ventruino, N. Prasad, and X.-D. Wang, "Coordinated linear beamforming in downlink multi-cell wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1451–1461, Apr. 2010.

IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 61, NO. 5, MARCH 1, 2013

[7] L. Ventruino, N. Prasad, and X.-D. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2835–2848, Jul. 2009.

[8] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, pp. 60–70, Jan. 1978.

[9] X. Shang and B. Chen, "Achievable rate region for downlink beamforming in the presence of interference," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 4–7, 2007, pp. 1684–1688.

[10] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime," *IEEE Trans. Inf. Theory*, vol. 57, pp. 2565–2581, May 2011.

[11] E. G. Larsson, E. A. Jorswieck, J. Lindblom, and R. Mochaourab, "Game theory and the flat-fading Gaussian interference channel," *IEEE Signal Process. Mag.*, vol. 26, pp. 18–27, Sep. 2009.

[12] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, pp. 5292–5296, Jul. 2008.

[13] X. Shang, B. Chen, and H. V. Poor, "Multiuser MISO interference channels with single-user detection: Optimality of beamforming and the achievable rate region," *IEEE Trans. Inf. Theory*, vol. 57, pp. 4255–4273, Jul. 2011.

[14] R. Mochaourab and E. A. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Trans. Signal Process.*, vol. 59, pp. 1128–1141, Mar. 2011.

[15] Y.-F. Liu and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 59, pp. 1142–1157, Mar. 2011.

[16] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual SINR framework," in *Proc. Int. ITG Workshop Smart Antennas*, Berlin, Germany, Feb. 16–18, 2009, pp. 1–7.

[17] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, pp. 5450–5458, Oct. 2010.

[18] E. G. Larsson and E. A. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1059–1069, Sep. 2008.

[19] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed resource allocation schemes: Pricing algorithms for power control and beamformer design in interference networks," *IEEE Signal Process. Mag.*, vol. 26, pp. 53–63, Sep. 2009.

[20] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, pp. 4298–4310, Aug. 2010.

[21] E. Karipidis, A. Gründinger, J. Lindblom, and E. G. Larsson, "Pareto-optimal beamforming for the MISO interference channel with partial CSI," in *Proc. 3rd IEEE Int. Workshop Comput. Adv. Multi-Sens. Adapt. Process.*, Aruba, Dutch Antilles, Dec. 13–16, 2009, pp. 5–8.

[22] J. Lindblom, E. Karipidis, and E. G. Larsson, "Outage rate regions for the MISO IFC," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 1–4, 2009, pp. 1120–1124.

[23] J. Lindblom, E. Karipidis, and E. G. Larsson, "Outage rate regions for the MISO interference channel: Definitions and interpretations," [Online]. Available: http://arxiv.org/abs/1106.5615v1

[24] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 46–55, Jan. 2002.

[25] S. Ghosh, B. D. Rao, and J. R. Zeidler, "Outage-efficient strategies for multiuser MIMO networks with channel distribution information," *IEEE Trans. Signal Process.*, vol. 58, pp. 6312–6324, Dec. 2010.

[26] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, pp. 20–34, May 2010.

[27] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, pp. 556–567, Oct. 2000.

[28] T. Bonald and L. Massoulié, "Impact of fairness on internet performance," in *Proc. ACM SIGMETRICS*, Cambridge, MA, Jun. 16–20, 2001, pp. 82–91.

[29] 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*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2010, ch. 4.

[30] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, Version 1.21 [Online]. Available: http://cvxr.com/cvx Apr. 2011

[31] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, pp. 664–678, Feb. 2010.

[32] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operat. Res.*, vol. 26, pp. 681–683, 1978.

[33] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Cambridge, MA: Athena Scientific, 2003.

[34] R. A. Gamboa and B. Middleton, "Taylor's formula with remainder," presented at the 3rd Int. Workshop ACL2 Theorem Prover and Its Appl., Grenoble, France, April 8–9, 2002.

Wei-Chiang Li received the B.S. degree in electrical engineering from the National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2009.

Currently, he is pursuing the Ph.D. degree in communications engineering at NTHU. His research interests are in signal processing problems in wireless communications, and convex optimization methods and its applications.
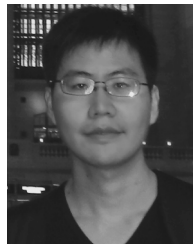
Tsung-Hui Chang (S'07–M'08) received the B.S. degree in electrical engineering and the Ph.D. degree in communications engineering from the National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2003 and 2008, respectively.

Since September 2012, he has been with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology (NTUST), Taipei, as an Assistant Professor. Before joining NTUST, he held research positions with NTHU from 2008 to 2011, and the University of California at Davis, CA, USA, from 2011 to 2012. He was also a visiting scholar of the University of Minnesota, Twin Cities, MN, USA, and the Chinese University of Hong Kong. His research interests are widely in signal processing problems in wireless communications and smart grid, and convex optimization methods and its applications.

Che Lin (S'02–M'08) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, in 1999. He received the M.S. degree in electrical and computer engineering in 2003, the M.S. degree in math in 2008, and the Ph.D. degree in electrical and computer engineering in 2008, all from the University of Illinois at Urbana-Champaign, IL, USA.

Since 2008, he has been at National Tsing Hua University (NTHU), Hsinchu, Taipei, where he is currently an Assistant Professor. He holds a U.S. patent, which has been included in the 3GPP LTE standard. His research interests include feedback systems, distributed algorithms in networks, practical MIMO code design in interference networks, optimization theory, information theory, systems biology, and synthetic biology.

Dr. Lin received the Excellent Teaching Award for the College of Electrical Engineering and Computer Science, NTHU, in 2012. He received a two-year Vodafone Graduate Fellowship in 2006 and the E. A. Reid Fellowship Award in 2008.

Chong-Yung Chi (S'83–M'83–SM'89) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1983.

From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, CA, USA. 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 co-sponsored workshops, symposiums, and conferences on signal processing and wireless communications, including Co-Organizer and General Co-Chairman of the 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), and Co-Chair of the 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 IEEE SIGNAL PROCESSING LETTERS (June 2006–May 2010), and a member of Editorial Board of EURASIP *Signal Processing Journal* (June 2005–May 2008), and an editor (July 2003–December 2005) as well as a Guest Editor (2006) of 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 the IEEE Signal Processing Committee on Signal Processing for Communications and Networking, a member of the IEEE Signal Processing Committee on Sensor Array and Multichannel Signal Processing, and an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.